United States Environmental Protection Agency | Office of Research and Development, Cincinnati, OH 45268 | EPA/625/R-00/010, August 2000

Environmental Problem Solving with Geographic Information Systems

1994 and 1999 Conference Proceedings

About this CD:

This CD, EPA Document Number EPA/625/R-00/010, contains conference proceedings documents from the 1994 and 1999 *Environmental Problem Solving with Geographic Information Systems* conferences. The 1994 papers are also contained in a Seminar Publication, EPA Document Number EPA/625/R-95/004, available from EPA.

Accessing CD contents:

All of the files contained on this CD are in Adobe Acrobat 4.0 format. A copy of Adobe Acrobat Reader 4.05 for Windows has been included on this CD for your convenience if you do not have this software. The software is also available free of charge on the Adobe homepage at: http://www.adobe.com/. If an older version of the Acrobat Reader is used, error messages will appear when files are opened.

To install Adobe Acrobat from the file on this CD, double-click the file entitled rs405eng.exe and follow the installation instructions that will appear on the screen. You will be asked to read and accept the Electronic End-User License Agreement as part of this installation.

Contents of CD:

- **1999 Agenda.pdf –** This file contains the agenda of the 1999 conference with links to the available presentation papers. It is organized by date and time of presentation at the conference.
- **1994 Contents.pdf –** This file contains a list of presentations from the 1994 conference with links to the available presentation papers. It is organized by topic area.
- **1999 Attendees.pdf –** This file contains the list of attendees from the 1999 conference, with links to papers by attending authors.

- **Papers –** This folder contains PDF files of all of the available papers from the 1994 and 1999 conferences. Every paper in this folder may be accessed through the links in the 1999 Agenda and 1994 Contents documents.
- **Readme.pdf -** This file contains information about the CD contents, access, and navigation. Start here when using this CD for the first time.
- **rs405eng.exe –** This file contains the Adobe Acrobat 4.05 installation files.

Navigating between files:

The 1994 Contents, 1999 Agenda, and 1999 Attendees List may all be accessed from this file. All the papers may be accessed by using the links from the 1994 Contents, 1999 Agenda, and 1999 Attendees List. When a link is followed to a paper, the originating file is closed. To return to the contents or agenda file, click on the title of the open paper. Every PDF document on this CD may also be accessed using your Windows Explorer or File Manager.

For more information:

Contact the United States Environmental Protection Agency at: www.epa.gov

Environmental Problem Solving with Geographic Information Systems

September 21-23, 1994
Cincinnati, Ohio

Other GIS Applications

Environmental Problem Solving With Geographic Information Systems

Tuesday, September 21, 1999

Pre-registration and Cash Bar Reception (5:00 PM - 8:00 PM)

Day 1 - Wednesday, September 22, 1999

Grand Ballroom A-B

PLENARY SESSION

7:30 - 9:00	Registration and Name Badge Pickup
9:00 - 9:15	Welcome & Overview *Sue Schock, USEPA, ORD, NRMRL* *Daniel J. Murray, USEPA, ORD, NRMRL*
9:15 - 10:00	New Directions in Environmental Problem Solving *Michael F. Goodchild, Ph.D., Chair* *National Center for Geographic Information and Analysis, and* *Department of Geography,University of California - Santa Barbara*
10:00 - 10:20	BREAK
10:20 - 10:50	GIS Workgroup: An Overview GIS - QA *George M.Brilis, J.D., USEPA, ORD, NERL*
10:50 - 11:30	Environmental Visioning with Geographic Information Systems *Sudhir R. Kshirsagar, Ph.D.,Global Quality Corporation, and* *Paul Koch, Pacific Environmental Services Inc.*
11:30 - 1:00	LUNCH

Day 1 - Wednesday, September 22, 1999 (continued)

Room	Grand Ballroom A		Grand Ballroom B	
SESSION *Moderator*	SESSION A DIFFUSE SOURCE *Lyn Kirschner* *Conservation Technology Information Center*		SESSION B ASSESSMENT / REMEDIATION *Jill Neal, USEPA, ORD, NRMRL*	
TIME	PRESENTATION	SPEAKER(S)	PRESENTATION	SPEAKER(S)
1:00 - 1:25	GIS Watershed Delineation Tools	James Goodrich, Ph.D., USEPA, ORD, NRMRL	GIS and GPS in Environmental Remediation Oversight at Federal Facilities in Ohio	Kelly Kaletsky and Bill Lohner, Ohio EPA
1:25 - 1:50	Nonpoint Pollutant Loading Application for ArcView GIS	Laurens van der Tak, P.E., CH2MHill	The Impact of Spatial Aggregation on Environmental Modeling: A GIS Approach	Lin Liu, Ph.D., University of Cincinnati
1:50 - 2:15	Application of DEM and Land Cover Data in Estimating Atmospheric Deposition to the Northeast and Mid-Atlantic Regions: Model Development and Applications	James A. Lynch, Ph.D., Pennsylvania State University	Characterizing the Hydrogeology of Acid Mine Discharges from the Kempten Mine Complex, West Virginia and Maryland	Benjamin R. Hayes, Bucknell University
2:15 - 2:40	Assessing the Impact of Landuse/Landcover on Stream Chemistry in Maryland	Gabriel Senay, Ph.D., PAI/SAIC	The GIS Connection to Residential Yard Soil Remediation	Jennifer Deis, Black & Veatch
2:40 - 3:00	BREAK			
3:00 - 3:25	Assessing the Long-Term Impact of Land Use Change On Runoff and Non-Point Source Pollution Using a GIS-NPS Model	Budhendra Bhaduri, Oak Ridge National Laboratory	GIS in the Confirmation Process	Raymond E. Bailey, Ph.D., MK-Ferguson
3:25 - 3:50	A Web-Based GIS Model for Assessing the Long-Term Hydrologic Impacts of Land Use Change (L-THIA GIS WWW): Motivation and Development	Jon Harbor, Purdue University (Bernie Engel, Ph.D.)	Using a Geographic Information Systems Application to Implement Risk Based Decisions in Corrective Action	Lesley Hay Wilson, P.E., The University of Texas - Austin
3:50 - 4:15	Using a Geographic Information System for Cost-Effective Reductions in Nonpoint Source Pollution: The Case of Conservation Buffers	Mark S. Landry, Virginia Tech University	Determining the Accuracy of Geographic Coordinates for NPDES Permittees in the State of Ohio	Bhagya Subramanian, USEPA, NERL
4:15 - 4:40	Putting Geospatial Information Into the Hands of the "Real" Natural Resource Managers: Lessons from the NEMO Project in Educating Local Land Use Decision Makers	Joel Stocker, Cooperative Extension, University of Connecticut	No presentation	

4:30 - 6:30 Reception (cash bar)

Day 2 - Thursday, September 23, 1999

8:00 - 9:00 Registration and Name Badge Pickup

Room	Grand Ballroom A		Grand Ballroom B	
SESSION *Moderator*	SESSION C APPLICATIONS *Doug Grosse, USEPA, ORD, NRMRL*		SESSION D URBAN / BROWNFIELDS / COMMUNITY *Jim Kreissl, USEPA, CERI, TTB*	
TIME	PRESENTATION	SPEAKER(S)	PRESENTATION	SPEAKER(S)
8:30 - 8:55	Evaluating Soil Erosion Parameter Estimates from Different Data Sources	Gabriel Senay, Ph.D., PAI/SAIC	Merging Transportation and Enviromental Planning Using GIS	Elizabeth Lanzer, Washington Dept. of Transportation
8:55 - 9:20	A Planning Strategy for Siting Animal Confinement Facilities: The Integrated Use of GIS and Digital Image Simulation Technologies	Thora Cartlidge, AICP, ASLA, University of Minnesota	Use of GIS Tools for Conducting Community On-Site Septic Management Planning	David Healy, Stone Environmental, Inc.
9:20 - 9:45	Lake Superior Decision Support Systems: GIS Databases and Decision Support Systems for Land Use Planning	George E. Host, Ph.D., Natural Resources Research Institute	Management and Reuse of Contaminated Soil -- The SoilTrak Method	Edward Rogers, Jr., BEM Systems, Inc.
9:45 - 10:10	Update of GIS Land Use Attributes from Land Surface Texture Information Using SIR-C Images	Francisco J. Artigas, Ph.D., Rutgers University	Using GIS to Rank Environmentally Sensitive Land in Orange County, Florida	Michael J. Gilbrook, HDR Engineering, Inc.
10:10 - 10:30	BREAK			
10:30 - 10:55	Onsite Wastewater Management Program in Hamilton County, Ohio--An Integrated Approach to Improving Water Quality and Preventing Disease	Timothy I. Ingram, Hamilton County General Health District, Ohio	Use of GIS for the Investigation and Classification of Land Being Redeveloped Under the Ohio Voluntary Action Program	Andrew Rawnsley, Ravensfield Geographic Resources, Ltd.
10:55 - 11:20	Modeling Combined Sewer Overflow (CSO) Impact: The Use of a Regional GIS in Facilities Planning	Michael D. Witwer, Metcalf & Eddy, Inc.	Assessing and Managing the Impacts of Urban Sprawl on Environmentally Critical Areas: A Case Study of Portage County, Ohio	Jay Lee, Ph.D., Kent State University
11:20 - 11:45	Building a Shared and Integrated GIS to Support Environmental Regulatory Activities in South Carolina	Guang Zhao, Ph.D., South Carolina Dept. of Health & Env. Control	Building a Brownfield Sitebank With Internet Map Server Technology	Alan Rao, Ph.D., Vanasse Hangen Brustlin, Inc
11:45 - 1:30	LUNCH			

Day 2 - Thursday, September 23, 1999 (Continued)

Room	Grand Ballroom A		Grand Ballroom B	
SESSION *Moderator*	<u>SESSION E</u> ECOLOGY / RESTORATION *Scott Minamyer, USEPA, CERI, TTB*		<u>SESSION F</u> RISK / ENVIRONMENTAL JUSTICE / EXPOSURE *Tom Brennan, USEPA, OPPT*	
TIME	PRESENTATION	SPEAKER(S)	PRESENTATION	SPEAKER(S)
1:30 - 1:55	Targeting the Knowledge Assembly Process of the Flora of North America (FNA): Biological Resource Problem Solving Using GIS	Leila M. Shultz,Ph.D., Harvard University and Utah State University	Using GIS to Analyze the Spatial Distribution of Environmental, Human Health, and Socio-Economic Characteristics in Cincinnati	Xinhao Wang, Ph.D. and Chris Auffrey, University of Cincinnati
1:55 - 2:20	GIS Standards for Environmental Restoration and Compliance	Bobby G. Carpenter, P.E., Tri-Service CADD/GIS Tech. Center	Public Participation GIS Applications for Environmental Justice Research and Community Sustainability	David Padgett, Ph.D., Tennessee State University
2:20 - 2:45	Reporting on the Development of an Environmental GIS Application - Wetlands Restoration in the Central Valley of California	David Hansen, US Bureau of Reclamation	Quantifying Risk in Watershed Assessment Using GIS & Stochastic Field-Scale Modeling	Conrad Heatwole, Ph.D., P.E., Virginia Tech University
2:45 - 3:10	Habitat Filters, GIS, and Riverine Fish Assemblages: Sifting Through the Relationships Between Fishes and Their Habitat	Douglas A. Nieman, Normandeau Associates	Methodological Issues in GIS-Based Environmental Justice Research	Jeremy Mennis, Pennsylvania State University
3:10 - 3:30	BREAK			

Day 2 - Thursday, September 23, 1999 (Continued)

Room	Grand Ballroom A		Grand Ballroom B	
SESSION *Moderator*	SESSION E ECOLOGY / RESTORATION *Scott Minamyer, USEPA, CERI, TTB*		SESSION F RISK / ENVIRONMENTAL JUSTICE / EXPOSURE *Tom Brennan, USEPA, OPPT*	
TIME	PRESENTATION	SPEAKER(S)	PRESENTATION	SPEAKER(S)
3:30 - 3:55	Using a GIS Model to Predict the Extent of Common Reed Encroachment into Two Tidal Wetland Areas in Northeastern New Jersey	Karla Hyde and Robin Dingle, Northern Ecological Associates, Inc.	Using GIS to Evaluate the Effects of Flood Risk on Residential Property Values	Alena Bartošová and David E. Clark, Ph.D., Marquette University
3:55 - 4:20	The Application of GIS in the Development of Regional Restoration Goals for Wetland Resources in the Greater Los Angeles Drainage Area	Charles Rairdan, US Army Corps of Engineers	Environmental Justice in Kentucky: Examining the Relationships Between Low-Income and Minority Communities and the Location of Landfills, and TSD Facilities	Larisa J. Keith, Northern Kentucky Area Planning Commission
4:20 - 4:45	Fractal Dimension as an Indicator of Human Disturbance in Galveston Bay, Texas	Amy Liu, PAI/SAIC	Application of GIS to Address Environmental Justice: Needs and Issues	Babafemi A. Adesanya, Environmental Equity Information Institute

4:30 - 6:30 Reception (cash bar)

Day 3 - Friday, September 24, 1999

8:00 - 9:00   Registration and Name Badge Pickup

Room	Grand Ballroom A		Grand Ballroom B	
SESSION *Moderator*	<u>SESSION G</u> WATERSHEDS *Mike Troyer, Ph.D., USEPA, ORD, NRMRL*		<u>SESSION H</u> MODELS / SYSTEMS *Randall Ross, Ph.D., USEPA, ORD, NRMRL*	
TIME	PRESENTATION	SPEAKER(S)	PRESENTATION	SPEAKER(S)
8:30 - 8:55	The National Hydrography Dataset – Status and Applications	Thomas G. Dewald, USEPA, and Keven S. Roth, USGS	Strategic Planning for GIS	Parrish Swearingen, Robins AFB
8:55 - 9:20	Sustainable Developments: Definition, Location, and Understanding	Michael E. Troyer, Ph.D., USEPA, ORD, NRMRL	No More 3-Ring Binders!	Margaret B. Martin, P.E., US Army Corps of Engineers
9:20 - 9:45	Development of a National Watershed Boundaries Dataset	Alan Rea, USGS	Pollution Exposure Index Model Measures Airborne Pollutants in National Forests	Michael V. Miller, CH2MHill
9:45 - 10:10	A Watershed-Based Approach to Source Water Assessment and Protection Utilizing GIS-Based Inventories: A Case Study in South Carolina	James M. Rine, Ph.D., Earth Sciences Research Institute	A GIS-Based Approach to Predicting Wetland Drainage & Wildlife Habitat Loss in the Prairie Pothole Region of South-Central Canada	David Howerter, Institute for Wetland and Waterfowl Research
10:10 - 10:30	BREAK			
10:30 - 10:55	Using an ARC/INFO GIS to Analyze Forest Patches for Watershed-Based Conservation and to Present Data on a Web Site	Lonnie Darr, Montgomery County, MD, Dept. of Env. Protection	Application of a Geographic Information System for Containment System Leak Detection	Randall R. Ross, Ph.D., USEPA, ORD, NRMRL
10:55 - 11:20	Application of a Water Balance Model and GIS for Sustainable Watershed Management	Thomas H. Cahill, P.E., and Susan Pagano, Cahill Associates	A High-Resolution Hydrometeorological Data System for Environmental Modeling and Monitoring	David R. Legates, Ph.D., University of Delaware
11:20	CONFERENCE CONCLUDES			

Eric Adams
OFFICE: 614-644-2752
FAX: 614-644-2909

Ohio EPA
P.O. Box 1049
Columbus OH 43216-1049

Babafemi A. Adesanya
OFFICE: 7578658950
FAX:
e2i2@eeii.org

Environmental Equity Information Institute
P.O. Box 189
Hampton VA 23669
USA

Larry Alber
OFFICE: 518-457-3143
FAX: 518-457-0738
lalber@gw.dec.state.ny.us

NYS Dept. of Environmental Conservation
50 Wolf Road
Albany NY 12233

Lora Alberto
OFFICE: 513-861-7666
FAX: 513-559-3155
mcrp@one.net

Mill Creek Restoration Project
42 Calhoun Street
Cincinnati OH 45219

Melanie Allamby
OFFICE: 216-881-6600x446
FAX: 216-881-2738
allambym@neorsd.org

N.E. Ohio Regional Sewer District
4415 Euclid Avenue
Cleveland OH 44103

Dr. James K. Andreasen (Jim)
OFFICE: 202-564-3293
FAX: 202-565-0076
andreasen.james@epa.gov

USEPA ORD Natl. Center for Env. Assessment
401 M Street, SW, Mail Code (8623D)
Washington DC 20460

Roger Anzzolin
OFFICE: 202-260-7282
FAX: 202-401-3041
anzzolin.roger@epa.gov

U.S. EPA OGWDW 4606
401 M. St. SW
Washington DC 20460

Gary Arnold
OFFICE: 541-686-7838x247
FAX: 541-686-7551
arnold.gary@deq.state.or.us

OREGON DEQ
1102 Lincoln
Eugene OR 97401

Francisco J. Artigas
OFFICE: 9733531069
FAX:
artigas@cimic.rutgers.edu

Center for Info. Mgmt., Integration & Connectivity **PAPER**
Rutgers University
180 University Avenue
Newark NJ 07102 USA

Chris Auffrey
OFFICE: 513-556-0579
FAX: 513-556-1274
chris.auffrey@uc.edu

University of Cincinnati **PAPER**
P.O. Box 210016
Cincinnati OH 45221-0016

Margaret Aycock
OFFICE: 409-880-8897
FAX: 409-880-1837
aycockma@hal.lamar.edu

Gulf Coast Hazardous Substance Research Center
P.O. Box 10671
Beaumont TX 77710-0671

Russ Baier | **Texas Natural Resource Conservation Comm.**
OFFICE: 512-239-1746 | 12100 Park 35 Circle
FAX: 512-239-5687 | Austin TX 78753
rbaier@tnrcc.state.tx.us

Jeanne M. Bailey | **American Water Works Association**
OFFICE: 202-628-8303 | 1401 New York Ave., NW #640
FAX: 202-628-2846 | Washington DC 20005
jbailey@awwa.org

Raymond E. Bailey | **MK-Ferguson** | **PAPER**
OFFICE: 6364418086 | 7295 Highway 94 South
FAX: | St. Charles MO 63304
ray_bailey@wssrap-host.wssrap.com | USA

Thomas E. Bailey | **Cinergy Corp.**
OFFICE: 513-287-2596 | 139 E. Fourth St., Room 552a
FAX: 513-287-3499 | Cincinnati OH 45201
tbailey@cinergy.com

Kim Baker | **Ohio Dept. of Natural Resources**
OFFICE: 614-265-6411 | 1952 Belcher Drive, Bldg. C-2
FAX: 614-267-2981 | Columbus OH 43224-1386
kim.baker@dnr.state.oh.us

M.C. Baldwin | **Navajo Nation Environmental Protection Agency**
OFFICE: 520-871-7690 | Water Quality Program - PO Box 339
FAX: 520-871-7599 | Window Rock AZ 86515
mcb4gis@juno.com

Todd Baldwin | **Island Press**
OFFICE: 202-232-7933 | 1718 Connecticut Avenue, NW, Suite 300
FAX: 202-234-1328 | Washington DC 20009
tbaldwin@islandpress.org

Brian Balsley | **BHE Environmental, Inc.**
OFFICE: 513-326-1500 | 11733 Chesterdale Rd.
FAX: 513-326-1550 | Cincinnati OH 45246
bbalsley@bheenv.com

Robert Bamford | **Nevada Division of Environmental Protection**
OFFICE: 775-687-4670 | 333 West Nye Lane
FAX: 775-687-6396 | Carson City NV 89706
rbamford@ndep.carson-city.nv.us

Quinn Barker | **Department of Natural Resources and Environmental Management**
OFFICE: 765-285-2327 | Ball State University
FAX: 765-285-2606 | Muncie IN 47306
jeflin1@gw.bsu.edu

Chris Barnett | **CARES - University of Missouri**
OFFICE: 573-882-9291 | 130 Mumford Road
FAX: 573-884-2199 | Columbia MO 65211-6200
barnett@cares.missouri.edu

Ralph P. Barr
OFFICE: 301-258-8651
FAX: 301-258-8679
barrR@ttnus.com

Tetra Tech NUS, Inc.
910 Clopper Road, Suite 400
Gaithersburg MD 20878

Alena Bartosova
OFFICE: 4142885128
FAX:
5x99Bartosov@marquette.edu

Dept. of Civil & Env. Engineering **PAPER**
Marquette University
P.O. Box 1881
Milwaukee WI 53201-1881 USA

Bruce Battles
OFFICE: 405-744-8974
FAX: 405-744-7008
bruce@seic.lse.okstate.edu

Oklahoma State University
201 CITD
Stillwater OK 74078

Bruce Bauch
OFFICE: 502-493-1945
FAX:
bbauch@usgs.gov

U.S. Geological Survey
9818 Bluegrass Parkway
Louisville KY 40299

Stephanie Beak
OFFICE: 614-644-4852
FAX: 614-728-1245
stephanie.beak@epa.state.oh.us

Ohio EPA
122 South Front Street
Columbus OH 43215

Michael L. Bechdol
OFFICE: 214-665-7133
FAX: 214-665-2191
bechdol.michael@epamail.epa.gov

U.S. EPA - Region 6WQ-SG
1445 Ross Avenue
Dallas TX 75202

Glynn Beck
OFFICE: 270-827-3414
FAX: 270-827-1117
ebeck@kgs.mm.uky.edu

Kentucky Geological Survey
P.O. Box 653
Henderson KY 42419

Tara L. Beckman
OFFICE: 4129218358
FAX:
beckmant@ttnus.com

TetraTech NUS, Inc.
Foster Plaza 7, 661 Andersen Dr
Pittsburgh PA 15220
USA

Bob Bednar
OFFICE: 405-702-8197
FAX: 405-702-8101
bobby.bednar@deqmail.state.ok.us

Oklahoma Dept. of Environmental Quality
707 N. Robinson Ave.
Oklahoma City OK 73102

Brian Begley
OFFICE: 502-564-6716
FAX: 502-564-2705
brian.begley@mail.state.ky.us

Kentucky Division of Waste Management
14 Reilly Road
Frankfort KY 40601

Dick Behr
OFFICE: 207-287-6828
FAX: 207-287-7826
richard.s.behr@state.me.us

Maine Department of Environmental Protection
17 State House Station
Augusta ME 04333

Milovan Beljin — **M.S. Beljin and Assoc.** — **PAPER**
OFFICE: 513-729-1602
FAX:
9416 Shadyoak Ct.
Cincinnati OH 45231

Kevin Benck — **Water Resources Institute-South Dakota St. Univ.**
OFFICE: 605-688-4776
FAX:
kbenck@brookings.net
SDSU Box 2120 AE 211
Brookings SD 57007

Malcolm Bender — **U.S. EPA / SEE Program**
OFFICE: 214-665-8378
FAX: 214-665-6660
bender.malcolm@epa.gov
1445 Ross Avenue
Dallas TX 75202

Clif Benoit — **US Forest Service**
OFFICE: 801-625-5594
FAX: 801-625-5483
cbenoit/r4@fs.fed.us
Federal Building
140 E. 3275 N.
N. Ogden UT 84414

Jerry Bernard — **USDA-Natural Resource Conservation Service**
OFFICE: 202-720-5356
FAX: 202-720-0428
jerry.bernard@usda.gov
P.O. Box 2890, Room 6123
Washington DC 20013

Budhendra Bhaduri — **Oak Ridge National Laboratory** — **PAPER**
OFFICE: 423-241-9272
FAX: 423-241-6261
bhaduribl@ornl.gov
P.O. Box 2008, MS 6237
Oak Ridge TN 37831-6237

Taher Bishr
OFFICE: 0020-12-2166800
FAX: 00203-4835337
20 Salah Salem Str. D. Toun
Alex - EGYPT

Ben Blaney — **USEPA /NRMRL**
OFFICE: 513-569-7852
FAX: 513-569-7680
blaney.ben@epa.gov
26 West Martin Luther King Dr.
Cincinnati OH 45268

Brian Bohl — **U.S. EPA**
OFFICE: 513-271-4182
FAX: 513-569-7160
bohl.brian@epamail.epa.gov
26W. Martin Luther King Drive
Cincinnati OH 45268

Steve Bolssen — **KY Dept. of Env. Protection, Division of Water**
OFFICE: 502-564-3410
FAX: 502-564-4245
steven.bolssen@mail.state.ky.us
14 Reilly Road
Frankfort KY 40601-1189

Bill Boria — **Chautauqua County Health Department**
OFFICE: 716-753-4481
FAX: 716-753-4344
billb@health.co.chatauqua.ny.us
7 North Erie Street
Mayville NY 14757

Joshua Bowen
OFFICE: 800-355-9429
FAX: 770-399-0535
jbowen@ftr.com

United Capitol Insurance Company
400 Perimeter Center Terrace, Suite 345
Atlanta GA 30346

Virgil Brack
OFFICE: 513-236-1163
FAX: 513-236-1178
vbrack@bheenv.com

BHE Environmental, Inc.
11733 Chesterdale Road
Cincinnati OH

John Bradley
OFFICE: 517-335-3146
FAX: 517-373-9657
bradlejn@state.mi.us

Michigan DEQ
300 South Washington Square
Lansing MI 48933

Tom Brankamp
OFFICE: 606-431-8579
FAX: 606-431-8581

Woolpert LLP
525 West Fifth Street, Suite 213
Covington KY 41011

Don Brannen
OFFICE: 513-681-8247
FAX: 513-681-1594

Cincinnati Recreation Commission
1655 Chase Avenue
Cincinnati OH 45206

Tom Brennan
OFFICE: (202) 260-3920
FAX:
brennan.thomas@epa.gov

U.S. EPA, OPPT
(MC7406)
Washington DC 20460

Timothy Bricker
OFFICE: 765-214-0088
FAX:
tjbricker@bsuvc.edu

Ball State Univeristy
3556 North Tillotson #205
Muncie IN 47304

James Bridges
OFFICE: 513-489-6611
FAX: 513-489-6619
jbridges@cin.pes.com

Pacific Environmental Services, Inc.
7209 E. Kemper Rd.
Cincinnati OH 45247

Jan W. Briede
OFFICE: 513-247-8000
FAX: 513-247-8010
jbriede@scientech.com

Scientech, NES, Inc.
11400 Grooms Road
Cincinnati OH 45242

George M. Brilis
OFFICE: 702-798-3128
FAX:
brilis.george@epamail.epa.gov

USEPA
P.O. Box 93478
Las Vegas NV 89193

Linda Briscoe
OFFICE: 513-641-3081
FAX: 513-641-0508
LBri938500@aol.com

Ohio/Cincinnati Women's Health Project
4860 Winneste Ave.
Cincinnati OH 45232

Thomas M. Brody
OFFICE: 312-353-8340
FAX: 312-353-4755
brody.tom@epa.gov

Region 5 Office of Information services
77 W. Jackson Blvd.
Chicago IL 60604

Joyce A. Broka
OFFICE: 517-686-8025x8371
FAX: 517-684-9799
brokaj@state.mi.us

Michigan Department of Environmental Quality
503 N. Euclid Ave.
Bay City MI 48706

Hugh J. Brown
OFFICE: 765-285-5788
FAX: 765-285-2606
hbrown@gw.bsu.edu

Ball State University
2000 University
Muncie IN 47306

Mike Bruening
OFFICE: 513-3261500
FAX:
mbruening@bheevin.com

BHE Environmental
Cincinnati OH

Jason Buck
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental
Ball State University
Muncie IN 47306

John Burckle
OFFICE: 513-569-7496
FAX: 513-569-7471
burckle.john@epamail.epa.gov

U.S. EPA / NRMRL
26 West Martin Luther King Drive
Cincinnati OH 45220

David Burden
OFFICE: 580-436-8606
FAX: 580-436-8614
burden.david@epa.gov

U.S. EPA/NRMRL/SPRD
P.O. Box 1198
ADA OK 74820

Tracy Burke
OFFICE: 717-236-3006
FAX: 717-233-0994
tab151@yahoo.com

GTS Technologies
851 South 19th Street
Harrisburg PA 17104

David Butler
OFFICE: 502-564-6716x339
FAX: 502-564-2705
david.butler@mail.state.ky.us

Kentucky Division of Waste Management
14 Reilly Road
Frankfurt KY 40601

Neill Cade
OFFICE: 513-357-7211
FAX: 513-357-7262
neill.cade@chdburn.rcc.org

Cincinnati Health Dept.
3101 Burnet Ave. Room 324
Cincinnati OH 45229

Thomas H. Cahill, P.E.
OFFICE: 610-696-4150
FAX: 610-696-8608
tcahill@thcahill.com

Cahill Associates
104 South High Street
West Chester PA 19382
USA

Tom Cain
OFFICE: 304-636-1800x289
FAX: 304-636-1875
tcain/r9_monong@fs.fed.us

Monongahela National Forest
200 Sycamore St.
Elkins WV 26241

Leslie M. Calderon
OFFICE: 817-608-2341
FAX: 817-695-9191
lcalder@dfwinfo.com

North Central Texas Council of Governments
616 Six Flags Drive, Suite 200, Centerpoint Two
Suite 200, Centerpoint Two
Arlington TX 76005-5888

Guy Cameron
OFFICE: 513-556-9740
FAX:
g.cameron@uc.edu

University of Cincinnati **PAPER**
Department of Biological Science
Cincinnati OH 45221

John Capillo
OFFICE: 606-986-0868
FAX: 606-986-2695
kefcapil@acs.eku.edu

KY Environmental Foundation
PO Box 467
Berea KY 40403

Bobby G. Carpenter
OFFICE: 601-634-4572
FAX: 601-634-4584
carpenb@wes.army.mil

Tri-Service CADD/GIS Technology Center
USACE Waterways Experiment Station, 3909 Halls Ferry Road
Vicksburg MS 39180-6199
USA

Thora Cartlidge
OFFICE: 612-624-9273
FAX: 612-624-1704
crd@tc.umn.edu

University of MN-Center for Rural Design **PAPER**
217-1518 Cleveland Avenue
St. Paul MN 55108

Erman Caudill
OFFICE: 606-257-4093
FAX:
elcaud00@pop.uky.edu

University of Kentucky
342 Waller Avenue #3C
Lexington KY 40504

Jean Caudill
OFFICE: 614-644-7181
FAX: 614-466-4556
jcaudill@gw.ohd.state.oh.us

Ohio Dept. of Health
P.O. Box 118
Columbus OH 43266-0118

Stephane Chalifoux
OFFICE: 514-287-8606
FAX: 514-287-8643
s.chalifoux@tecsult.com

TECSULT
85 St. Catherine Street West
Montreal, Quebec
CANADA H2X 3P4

Yu-mei Chang
OFFICE: 513-558-2744
FAX:
changym@email.uc.edu

University of Cincinnati
202 Ruth Lyons Way, Suite #261
Cincinnati OH 45267-0840

James Chapman
OFFICE: 765-213-1269
FAX:
jdchapman31@hotmail.com

Delaware County Indiana
100 W. Main Room 206
Muncie IN 47305

Jonathan W. Chapman
OFFICE: 410-260-8514
FAX: 410-260-8595
jchapman@dnr.state.md.us

MD Dept. of Natural Resources-Forest Service
Tawes State Office Building E1 580 Taylor Avenue
Annapolis MD 21401

David Christenson
OFFICE: 303-312-3345
FAX: 303-312-6065
christenson.dave@epa.gov

USEPA Region VIII
999 18th Street, Suite 500
Denver CO 80202

David Christian
OFFICE: 613-787-3879
FAX: 613-787-3884

ATCO Frontec Corporation
100-170 Laurier Avenue West
Ottawa, Ontario K1P 5V5

Kendra Cipollini
OFFICE: 312-886-1432
FAX: 312-886-9697
cipollini.kendra@epa.gov

USEPA Region 5, Critical Ecosystems Team
77 West Jackson Avenue T13J
Chicago IL 60604

John Ckmanec
OFFICE: 513-569-7481
FAX: 513-569-7585
ckmanec.john@epa.gov

U.S. EPA (MS G-75)
26 Martin Luther King Drive
Cincinnati OH 45268

Tara Clapp
OFFICE: 502-852-8152
FAX: 502-852-4558
tlclapp@rcf.usc.edu

University of Louisville
426 W. Bloom Street, 202
Louisville KY 40208

David E. Clark
OFFICE: 414-288-3339
FAX: 414-288-5757
clarkde@marquette.edu

Marquette University, Dept. of Economics **PAPER**
P.O. Box 1881
Milwaukee WI 53201-1881

Tim Clarke
OFFICE: 502-573-2886
FAX: 502-573-2355
tim.clarke@mail.state.ky.us

Kentucky State Nature Preserves Commision
801 Schenkel Lane
Frankfort KY 40601-1403

Richard Cochran
OFFICE: 615-532-0997
FAX: 615-532-0046
RCochran2@mail.state.tn.us

Tennessee Department of Environmental and Conservation
401 Church Street, 7th Floor L & C Annex, Div of Water Pollution Control
Nashville TN 37243

James C. Coleman II
OFFICE: 513-583-1249
FAX: 513-583-1250
jcoleman@environcorp.com

ENVIRON International Corp.
6443 Lewis Road
Loveland OH 45140

Jim Coon
OFFICE: 937-285-6038
FAX: 937-285-6404
jim.coon@epa.state.oh.us

Ohio EPA **PAPER**
401 E. Fifth St.
Dayton Oh 45402

Amy Covert
OFFICE: 502-573-2886
FAX: 502-573-2355
amy.covert@mail.state.ky.us

Kentucky State Nature Preserves Commission
801 Schenkel Lane
Frankfort KY 40601-1403

Christopher Cox
OFFICE: 309-782-2887
FAX: 309-782-5038
wildmanf@ria.army.mil

Rock Island Arsenal
510R1-AO, Bldg 220
Rock Island IL 61299

Michael J. Cramer
OFFICE: 513-556-9740
FAX: 513-556-5299
michaeljcramer@hotmail.com

University of Cincinnati
P.O. Box 210006
Cincinnati OH 45221-0006

Irene M. Crawford
OFFICE: 513-569-7167
FAX:
crawford.irene@epa.gov

SoBran, Inc.
26 West Martin Luther King Drive
Cincinnati OH 45268

Pat Curley
OFFICE: 770-673-3640
FAX: 770-396-9495
pcurley@brwncald.com

Brown and Caldwell
41 Perimeter Center East, Suite 400
Atlanta GA 30346

Darrin L. Curtis
OFFICE: 501-973-0760
FAX:
dlc2@engr.uark.edu

USAF
2335 East Yvonne Drive
Fayetteville AR 72703

Bernie Daniel
OFFICE: 513-569-7401
FAX: 513-569-7609
daniel.bernie@epa.gov

USEPA -NERL
26 West Martin Luther King Drive
Cincinnati OH 45242

Lonnie Darr
OFFICE: 2407777703
FAX:
darrl@co.mo.md.us

Watershed Management Div., Montgomery Cty Dept. of Env. Protection
Suite 120, 155 Rockville Pike
Rockville MD 20850
USA

Bruce De Young
OFFICE: 616-336-3234
FAX: 616-336-2436
e10771@iserv.net

Kent County Health Department
700 Fuller, N.E.
Grand Rapids MI 49503

Kevin L. DeFosset
OFFICE: 606-624-4471
FAX:
studefok@acs.eku.edu

Ewers Water Consultants / Eastern KY University
326-10 Lancaster Avenue
Richmond KY 40475

Jennifer Deis, P.G.
OFFICE: 913-458-6585
FAX: 913-458-2934
deisj@bv.com

Black & Veatch Corporation **PAPER**
11401 Lamar Avenue
Overland Park KS 66211

Ed Delisio
OFFICE: 312-886-1303
FAX: 312-353-5374
delisio.edward@epa.gov

US EPA Region 5
77 West Jackson Blvd., B-19J
Chicago IL 60604

Phil Dennis
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental
Ball State University
Muncie IN 47306

Tommy G. Dewald
OFFICE: 2022602488
FAX:
dewald.tommy@epamail.epa.gov

Office of Water
401 "M" Street, SW (4503F)
Washington DC 20460
USA

Martin Diaz-Zorita
OFFICE: 606-257-3655
FAX: 606-257-2185
mdzori2@pop.uky.edu

University of Kentucky - Agronomy Department
N122-Agric.Sci.Center North
Lexington KY 40546-0091

Robin Dingle
OFFICE: 2078799496
FAX:
Rdingle@neamaine.com

Northern Ecological Associates, Inc.
386 Fore St.
Portland OR 04101
USA

Harold B. Dirschl
OFFICE: 613-787-9611
FAX: 613-787-3884
hkdir@sympatico.ca

ATCO Frontec Corporation
100-170 Laurier Avenue West
Ottawa, Ontario K1P 5V5

David Dixon
OFFICE: 513-247-8000
FAX: 513-247-8010
ddixon@scientech.com

Scientech, NES, Inc.
11400 Grooms Road
Cincinnati OH 45242

Mohsen Dkhili
OFFICE: 573-751-1300
FAX: 573-526-5797
nrdkhim@mail.dnr.state.mo.us

DNR / WPCP
P.O. Box 176
Jefferson City MO 65102

Doug Dobransky
OFFICE: 614-644-2752
FAX: 614-644-2909

Ohio EPA
P.O. Box 1049
Columbus OH 43216-1049

Kevin Doniere
OFFICE: 513-281-2211
FAX: 513-281-2243
humannature@fuse.net

Human Nature
990 St. Paul Place
Cincinnati OH 45206

Damon Dougherty
OFFICE: 361-883-6016
FAX: 361-883-7417
damon@moorhousecc.com

Moorhouse Associates, Inc.
5826 Bear Lane
Corpus Christi TX 78405

Joan Douglas
OFFICE: 352-378-6517
FAX: 352-338-8247
FACA@aol

Florida Rural Community Assistance Project
6212 NW 43 Street, Suite A
Gainsville FL 32653

Mark E. Duewell
OFFICE: 573-526-5214
FAX: 573-751-6417
duewem@mail.health.state.mo.us

Missouri Department of Health
930 Wildwood, P.O. Box 570
Jefferson City MO 65102-0570

Jeffrey Duke
OFFICE: 216-881-6600x456
FAX: 216-881-2738
dukej@neodrsd.org

N. E. Ohio Regional Sewer District
4415 Euclid Avenue
Cleveland OH 44103

John Dunham R.S.
OFFICE: 513-564-1788
FAX: 513-564-1776
John.Dunham@igwmail.rcc.org

Cincinnati Health Dept.
3845 W.P. Dooley By-Pass
Cincinnati OH 42223

Scott Dyer
OFFICE: 513-627-1163
FAX: 513-627-1208
dyer.sd@pg.com

Procter & Gamble Co.
P.O. Box 538707
Cincinnati OH 45253-8707

Don Ebert
OFFICE: 702-798-2158
FAX: 702-798-2158
ebert.donald@epa.gov

U.S. EPA
944 E. Harmon Avenue
Las Vegas NV 89119

T. J. Edwards
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental Management
Ball State University
Muncie IN 47306

James Eflin
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental Management
Ball State University
Muncie IN 47306

Keith Egan
OFFICE: 513-576-0009
FAX: 513-574-9756
egank@srwenvironmental.com

SRW Environmental Services
55 West TechneCenter, Suite C
Milford OH 45150

Lisa E. Enderle
OFFICE: 703-645-6950
FAX: 703-698-6101
enderlel@saic.com

SAIC
2222 Gallows Road, Suite 300
Dunn Loring VA 22027

Bernie Engel
OFFICE: 765-494-1198
FAX: 765-496-1115
engelb@ecn.purdue.edu

Purdue University **PAPER**
1146 ABE
W. Lafayette IN 47907-1146

Sendar Ertep
OFFICE: 404-562-9683
FAX: 404-562-9598
ertep.sendar@epamail.epa.gov
USEPA Region 4
61 Forsyth St., SW
Atlanta GA 30303

Alan Evereson
OFFICE: 513-569-7046
FAX:
U.S. EPA / ORD / NRMRL
26 W. Martin Luther King Drive
Cincinnati OH 45268

Dr. Ralph O. Ewers
OFFICE: 606-623-8464
FAX: 606-623-6464
glyewers@acs.eku.edu
Dept. of Earth Science Eastern KY University
160 Redwood
Richmond KY 40475

Susan Fagan
OFFICE: 202-260-9477
FAX: 202-260-2941
fagan.susan@epa.gov
U.S. EPA, Office of Water
401-M Street, SW - Mailcode 4501F
Washington DC 20460

Todd Falter
OFFICE: 402-471-6571
FAX: 402-471-6436
tfalter@hhs.state.ne.us
NE Health and Human Services
P.O. Box 95007
Lincoln NE 68509

Ian Farrar
OFFICE: 304-545-4388
FAX:
ifarrar@columbiaenerorcnup.com
Collumbia Gas
1700 MacConkle Ave.
Charleston WV

Terry Felkerson
OFFICE: 573-308-3725
FAX: 573-308-3652
jfelkerson@usgs.gov
U.S. Geological Survey
1400 Independence Road
Rolla MO 65401

Don Ficklen
OFFICE: 210-671-4844
FAX: 210-671-2241
holmes.ficklen@lockland.af.mil
U.S. Air Force
Lockland Air Force Base

William S. Fischer
OFFICE: 513-564-1787
FAX: 513-564-1776
wfish@mailcity.com
Cincinnati Health Dept.
3845 William P. Doley By-Pass
Cincinnati OH 45223

Jeff Flege
OFFICE: 614-265-6686
FAX:
jeff.flege@dnr.state.oh.us
Ohio Department of Natural Resources
Fountain Square, Building C-2
Columbus OH 43224-1386

Terry Flum
OFFICE: 513-569-7715
FAX: 513-569-7609
USEPA-NERL MS 642
26 W. Martin Luther King Drive
Cincinnati OH 45268

Charles Frank
OFFICE: 610-935-4703
FAX: 610-935-5583
cfrank@envstd.com
Environmental Standards Inc.
1140 Valley Forge Road
Valley Forge PA 19482-0810

William T. Frederick
OFFICE: 716-942-2563
FAX: 716-942-2247
frederw@wv.doe.gov
Dames & Moore
10282 Rock Springs Road
West Valley NY 14171-9799

Dr. Robert C. Frey
OFFICE: 614-466-1069
FAX: 614-644-7440
rfrey@gw.odh.state.oh.us
Ohio Department of Health
246 North High Street
Columbus OH 43266-0588

Lawrence Friedl
OFFICE: 202-564-6933
FAX: 202-565-2431
friedl.lawrence@epa.gov
US EPA
401 M St., SW (8101R)
Washington DC 20460

Joseph Frizado
OFFICE: 419-372-7202
FAX: 419-372-7205
frizado@bgnet.bgsu.edu
Bowling Green State University
Bowling Green State University
Dept. of Geology
Bowling Green OH 43403

Kevin Frysinger
OFFICE: 610-935-5577
FAX: 610-935-5583
Environmental Standards
1140 Valley Forge Road
Valley Forge PA 19482

Florence Fulk
OFFICE:
FAX:
fulk.florence@Epa.gov
EPA/NERL
26 W. Martin Luther King Drive
Cincinnati OH 45268

Richard Futrell
OFFICE: 606-622-1581
FAX:
antfutre@acs.eku.edu.
Eastern Kentucky University Sociology Dept.
223 Keith
Richmond KY 40475

Robert Galbraith
OFFICE: 513-576-0009
FAX: 513-576-9756
galbraib@srwenvironmental.com
SRW Environmental Services
55 West TechneCenter, Suite C
Milford OH 45150

Dr. Achal Garg
OFFICE: 513-357-7209
FAX: 513-357-7262
achal.garg@chdburn.rcc.org
Cincinnati Health Department
3101 Burnet Ave.
Cincinnati OH 45229

Donald Gasper
OFFICE: 304-472-3704
FAX:
WV Department of Natural Resources
4 Ritchie Street
Buchannon WV 26201

Glee H. Gasper
OFFICE: 304-942-3704
FAX:
Upshur County Litter Control Bd.
4 Ritchie St.
Buckhannon WV 26201

David R. German
OFFICE: 605--688-5611
FAX: 605-688-4917
wrisdsu@mg.sdstate.edu
South Dakota Water Resources Institute
SDSU Box 2120
Brookings SD 57007

Patricia Germany
OFFICE: 502-574-3645
FAX: 502-574-1389
pgermany@louky.org
City of Louisville / Health & Environment
1514 Hale Avenue
Louisville KY 40210

Tim Gessner
OFFICE: 513-326-1500
FAX:
BHE Envir.
11733 Chesterdale Rd.
Cincinnati OH 45204

Nick Giardino
OFFICE: 210-536-6128
FAX: 210-536-1130
nicholas.giardino@brooks.af.mil
IERA/RSRE
2513 Kennedy Circle
San Antonio TX 78235-5123

Michael J. Gilbrook
OFFICE: 407-872-7801
FAX: 407-872-0603
mgilbroo@hdrinc.com
HDR Engineering, Inc. **PAPER**
201 S. Orange Avenue, Suite 925
Orlando FL 32801-3413

Rebecca Glos
OFFICE: 703-318-4797
FAX: 703-736-0826
glosr@saic.com
SAIC
11251 Roger Bacon Drive
Reston VA 20190

Haynes Goddard
OFFICE: 513-569-7685
FAX: 513-569-7111
goodard.haynes@epamail.epa.gov
USEPA
26 West Martin Luther King Drive
Cincinnati OH 45268

Maria A. Gomez-Balandra
OFFICE: 52-73-19-4000x407 or x410
FAX: 52-73-20-8638
magomez@tlaloc.imta.mx
Water Technology Mexican Institute
Paseo Cuauhnahuac 8532
Morelos, Mexico 62550

James A. Goodrich
OFFICE: 513-569-7605
FAX: 513-569-7185
goodrich.james@epa.gov
USEPA, NRMRL/WSWRD/WQMB **PAPER**
26 W. Martin Luther King Drive
Cincinnati OH 45268

Walter M. Grayman
OFFICE: 513-281-6138
FAX: 513-281-6139
grayman@fuse.net
W.M. Grayman Consulting Engineer
730 Avon Fields Lane
Cincinnati OH 45229

Don Green
OFFICE: 615-837-5492
FAX: 615-837-5025
dgreen@mail.state.tn.us

TN Nonpoint Source Program
Ellinton Ag. Center, Box 40627, Holeman Bldg.
Nashville TN 37204

Chris Griffith
OFFICE: 513-732-8036
FAX:
ccghd@fuse.net

Clermont Co. General Health District
2275 Bauer Road Suite 300
Batavia OH 45103

Corey Gullion
OFFICE: 765-288-4057
FAX: 765-288-4057
cgullion@home.com

Ball State Univ. Natural Resource Env. Mgr.
1616 W. Gilbert G-100
Muncie IN 47303

Beth Hailstock
OFFICE: 513-357-7206
FAX: 513-357-7262
beth.hailstock@chdburn.rcc.org

Cincinnati Health Dept.
3101 Burnet Ave
Cincinnati OH 45229

Tonia Hampton
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental Management
Ball State University
Muncie IN 47306

David T. Hansen
OFFICE: 916-978-5268
FAX: 916-978-5290
dhansen@mp.usbr.gov

U.S. Bureau of Reclamation--MPGIS **PAPER**
2800 Cottage Way
Sacramento CA 95825-1898

Todd Hanson
OFFICE: 218-335-7415
FAX: 218-335-7430
lldrm@paulbunyan.net

Leech Lake Reservation Water Resources
6530 Hwy 2 NW
Cass Lake MN 56633

Jon Harbor
OFFICE: 7654949610
FAX:
jharbor@purdue.edu

Dept. of Earth & Atmospheric Sciences **PAPER**
Purdue University
West Lafayette IN 47906-1397
USA

Deborah D. Harris
OFFICE: 513-791-8330
FAX: 513-791-7335
logictree@aol.com

Natl. Technology Assoc.-Cincinnati
P.O. Box 42356
Cincinnati OH 45242

Jasper L. Harris
OFFICE: 919-530-6394
FAX: 919-530-7966
jasperharris@yahoo.com

North Carolina Central University
1801 Fayetteville Street
Durham NC 27707

Roderick L. Harris, R.S.
OFFICE: 615-353-8363
FAX:
rlh1914@yahoo.com

Meharry Medical College
1408 Mountain Valley Bend
Nashville TN 37209

Joel Harrison
OFFICE: 304-637-0245
FAX: 304-637-0250
jharrison@dnr.state.wv.us
WV Division of Natural Resources
Ward Road
Elkins WV 26241

Robin Harrover
OFFICE: 425-649-7232
FAX: 425-649-7098
rhar461@ecy.wa.gov
Washington State Department of Ecology
3190 160th Avenue, SE
Bellevue WA 98008-5452

Paul Harten
OFFICE: 513-569-7045
FAX: 513-569-7471
harten.paul@epamail.epa.gov
U.S. EPA
26 W. Martin Luther King Drive
Cincinnati OH 45268

Megan Hartgrove-Del Gaudio
OFFICE: 410-974-7276
FAX: 410-974-7200
mhart@menv.com
Maryland Environmental Service
2011 Commerce Park Drive
Annapolis MD 21401

Patrick L. Havens
OFFICE: 317-337-3465
FAX: 317-337-3235
phavens@dowagro.com
Dow AgroSciences LLC
9330 Zionsville Road 306/A2
Indianapolis IN 46468

Fred Hayden
OFFICE: 510-526-7140
FAX:
Haydens@Earthlink.net
Hydrosource
639 Madison Street
Albany CA 94706

Benjamin R. Hayes
OFFICE: 570-372-4215
FAX: 570-372-2726
bhayes@susqu.edu
Dept. of Geo. and Envir. Sc. Susquehanna U. **PAPER**
514 University Ave.
Selinsgrove PA 17870-1001
USA

Linda Haynie
OFFICE: 512-239-6821
FAX: 512-239-5687
LHAYNIE@TNRCC.STATE.TX.US
TX Natural Resource Conservation Commission
P.O. Box 13087
Austin TX 78711

David J. Healy
OFFICE: 802-229-1879
FAX: 802-229-5417
dhealy@stone-env.com
Stone Environmental, Inc. **PAPER**
58 East State Street
Montpelier VT 05602

Richard H. Heath
OFFICE: 207-287-7637
FAX: 207-287-7826
richard.h.heath@state.me.us
Maine Department of Environmental Protection
State House Station #17
Augusta ME 04333

Conrad Heatwole
OFFICE: 5402314858
FAX:
heatwole@vt.edu
Biological Systems Engineering
Virginia Tech
Blacksburg VA 24061-0303 USA

Pat Heider
OFFICE: 614-644-2752
FAX: 614-644-2909

Ohio EPA
P.O. Box 1049
Columbus OH 43216-1049

Beverly Henderson
OFFICE: 614-644-8065
FAX: 614-644-7740
bhenders@gw.odh.state.oh.us

Ohio Dept. of Health Assessment Section
246 N. High Street, 8th Floor
Columbus OH 43215

Nereida Hernández
OFFICE: 787-764-8824
FAX: 787-766-0150
jcaemer@prtc.net

Puerto Rico Environmental Quality Board
P.O. Box 11488
San Juan PR 00910

Bob Hilbert
OFFICE: 716-942-2417
FAX:
hilbert@wv.doe.gov

West Valley Nuclear Services
10282 Rock Springs Road
West Valley NY 19171

Sara Hines
OFFICE: 502-573-2886
FAX: 502-573-2355
sara.hines@mail.state.ky.us

Kentucky State Nature Preserves Commission
801 Schenkel Lane
Frankfort KY 40601-1403

Terri Hoagland
OFFICE: 513-569-7783
FAX: 513-569-7111
hoagland.theresa@epamail.epa.gov

U.S. EPA
26 W. Martin Luther King Drive MS-466
Cincinnati OH 45268

Molly Hodgson
OFFICE: 216-910-1941
FAX: 216-910-2010
molly-hodgson@aquaalliance.com

Metcalf & Eddy, Inc. **PAPER**
Suite 1215, 1300 E. Ninth Stree
Cleveland OH 44114
USA

Matthew Hopton
OFFICE:
FAX:
hopton@toast.net

Dept. of Biological Sc., University of Cincinnati
Cincinnati OH 45221-0006

George E. Host, Ph.D.
OFFICE: 218-720-4264
FAX: 218-720-4328
ghost@sage.nrri.umn.edu

Natural Resources Research Institute; U of MN **PAPER**
5013 Miller Truck Highway
Duluth MN 55811

Kevin House
OFFICE: 502-564-3080
FAX: 502-564-9195
kevin.house@mail.state.ky.us

Kentucky Division of Conservation
663 Teton Trail
Frankfort KY 40601

David Howerter
OFFICE: 204-467-3292
FAX: 204-467-9426
d_howerter@dveb.com

Institute for Wetland and Waterfowl Research **PAPER**
P.O. Box 1160
Oak Hammock Marsh
Stonewall MB ROC2ZO

Hsuan-Tsung (Sean) Hsieh
OFFICE: 405-744-8974
FAX: 405-744-7008
hsieh@seic.lse.okstate.edu
Oklahoma State University--SEIC
Oklahoma State University
201 CITD
Stillwater OK 74074

Karla Hyde
OFFICE: 2078799496
FAX:
Rdingle@neamaine.com
Northern Ecological Associates, Inc.
386 Fore St. Ste. 40
Portland OR 04101
USA

Peter J. Idstein
OFFICE: 606-624-8722
FAX: 606-622-2876
idstein@gateway.net
Ewers Water Consultants
971 Villa Drive, Apartment #27
Richmond KY 40475

Timothy I. Ingram
OFFICE: 513-326-4503
FAX: 513-772-6405
Tim.Ingram@health.hamilton-co.org
Health Commissioner, Hamilton County General Health District
Chester Towers, Suite 600, 11499 Chester Road **PAPER**
Cincinnati OH 45246
USA

Brian Jacobson
OFFICE: 724-349-5733
FAX:
BAJ133@PSU.EDU
Penn State University
905 F W. Aaron Drive
State College PA 16803

Karen Jarocki
OFFICE: 505-768-7706
FAX: 505-768-7601
jarocki@mrcabq.com
Mission Research Company
5001 Indian School Road NE
Albuquerque NM 87110

Tony Jasek
OFFICE: 210-536-5448
FAX: 210-536-1130
tony.jasek@brooks.af.mil
US Air Force
2903 Oak Falls
San Antonio TX 78231

Bruce Jeffries
OFFICE: 517-335-0183
FAX: 517-335-6565
jeffries@mshda,cis.state.mi.us
MSHDA
401 S. Washington Square
Lansing MI 48933

Becky Jenkins
OFFICE: 614-265-6631
FAX: 614-263-8144
becky.jenkins@dnr.state.oh.us
Ohio Division of Wildlife
1840 Belcher Dr., Bldg. G-2
Columbus OH 43224

Ralph Johanson
OFFICE: 502-584-4244
FAX: 502-584-4246
rjohanson@grwinc.com
GRW Engineers
433 South 5th Street
Louisville KY 40202

Brian C. Johnson
OFFICE: 518-474-5488
FAX: 518-473-2534
brian.johnson@oag.state.ny.us
NY State Attorney Gen. - Env. Protection Bureau
The Capitol
Albany NY 12224

Robert L. Johnson
OFFICE: 614-466-1060
FAX: 614-564-2410
bjohnson@gw.odh.state.oh.us

Ohio Department of Health
246 North High Street
Columbus OH 43215

J. Kent Johnson, PhD
OFFICE: 319-626-6053
FAX: 319-626-6053
JKENT@inav.net

Iowa Institute of Hydraulic Research
University of Iowa
Iowa City IO 52242

Rick Jones
OFFICE: 317-247-3105
FAX: 317-247-3414
jonesrw@in-arng.ngb.army.mil

Military Dept. of Indiana
2002 S. Holt Road
Indianapolis IN 46241

Terri Justice
OFFICE: 719-567-4035
FAX: 719-567-4036
terri.justice@schriever.af.mil

US Air Force 50 Ceslcecr
300 O'Malley Ave., Suite 19
Schriever AFB CO 80912-5091

Jeff Kaczka
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental Management;
Ball State University
Muncie IN 47306

K.E. (Kim) Kalen
OFFICE: 613-992-0757
FAX: 613-996-3925

Dept. of National Defense-North WarningSystem
219 Laurier Avenue West
Ottawa, Ontario K1A OK2

Kelly Kaletsky
OFFICE: 5132856075
FAX:
kelly.kaletsky@epa.state.oh.us

Office of Federal Facilities Oversight - Ohio EPA **PAPER**
401 East Fifth Street
Dayton OH 45402-2911
USA

Sumana Keener
OFFICE: 513-556-2542
FAX: 513-556-2522
skeener@uceng.uc.edu

University of Cincinnati
P.O. Box 210071
Cincinnati OH 45227

Bud Keesee
OFFICE: 410-278-6755
FAX: 410-278-6779
bkeesee@dshe.apg.army.mil

Div. of Safety, Health, and Environment
STEAP-SH-ER - Bldg 5650
Aberdeen Proving Group Garrison
MD 21005

Larisa J. Keith
OFFICE:
FAX:
larisa_j_k@yahoo.com

Northern Kentucky Area Planning Commission in Kenton County
146 Stafford Ridge Rd. **PAPER**
Sanders KY 41083
USA

Jay H. Kim
OFFICE: 304-285-6140
FAX: 304-285-6111
jhk0@cdc.gov

CDC / NIOSH
1095 Willowdale Road
Morgantown WV 26505

Carl Kinkade
OFFICE: 402-441-6246
FAX: 402-441-8323
ckinkade@ci.lincoln.ne.us

Lincoln-Lancaster County Health Dept.
3140 N Street
Lincoln NE 68510

Lyn Kirschner
OFFICE: (765) 494-9555
FAX:
kirschner@ctic.purdue.edu

Conservation Technology Information Center
1220 Potter Dr., Ste. 170
West Lafayette IN 47906

Karen Klima
OFFICE: 202-260-7087
FAX: 202-260-7024
klima.karen@epa.gov

USEPA / Office of Water
401 M Street, SW Mail Code 4503F
Washington DC 20460

Randy Knight
OFFICE: 319-398-5893
FAX: 319-398-5894
rknight@kirkwood.cc.ia.us

Kirkwood Community College
6301 Kirkwood Blvd. SW
Cedar Rapids IA 52406

Jill Korach
OFFICE: 513-921-5124
FAX: 513-921-5136
imago@one.net

IMAGO, Inc.
553 Enright Avenue
Cincinnati OH 45205

Angel J. Kosfiszer
OFFICE: 214-665-2187
FAX: 214-665-6490
kosfiszer.angel@epamail.epa.gov

US EPA Region 6
1445 Ross Avenue
Dallas TX 75202-2733

Jeff Kreider
OFFICE: 608-266-0856
FAX: 608-267-2800
kreidj@dnr.state.wi.us

Wisconsin Dept. of Natural Resources
101 S. Webster Street
Madison WI 53707

Jim Kreissl
OFFICE: (513) 569-7611
FAX:
kreissl.james@epa.gov

U.S. EPA, CERI, TTB
26 W. Martin Luther King Dr.
Cincinnati OH 45268

Dennis Kreitzburg
OFFICE: 216-739-0555
FAX: 216-739-0560
DJK&HaleyAldrich.com

Haley & Aldrich, Inc.
5755 Granger Road, Suite 100
Independence OH 44131

Eric J. Kroger
OFFICE: 513-648-4473
FAX: 513-648-4473
eric.kroger@fernald.gov

Fluor Daniel Fernald, Inc.
P.O. Box 538704
Cincinnati OH 45253

Rosanne Kruzich
OFFICE: 502-895-4559
FAX: 502-895-4559
rkruzich@bellsouth.net

RKX Consulting, Inc.
750 Zorn Avenue #56
Louisville KY 40206

Dr. Sudhir R. Kshirsagar OFFICE: 513-474-9780 FAX: 513-474-9781 *president@gqc.com*	**Global Quality Corporation** 7828 Beechmont Avenue, Suite 202 Cincinnati OH 45255	
Tony Lafferty OFFICE: 314-949-6620 FAX: 314-949-6735 *tlafferty@esri.com*	**ESRI St. Louis** 207 South Main St. St. Charles MO 63301	
Mark S. Landry OFFICE: 540-951-4899 FAX: 540-951-4896 *m-landry@vt.edu*	**VA Polytechnic Institute and State University** Dept. of Agric. & Applied Economics Blacksburg VA 24061-0401	**PAPER**
Robert Langstroth OFFICE: 608-264-8801 FAX: 608-264-8795 *rlangstroth@commerce.state.wi.us*	**Wisconsin Dept. of Commerce** 201 W. Washington Ave. Madison WI 53701	
Elizabeth L. Lanzer OFFICE: 360-705-7476 FAX: 360-705-6833 *lanzere@wsdot.wa.gov*	**WA Department of Transportation** P.O. Box 47331 Olympia WA 98504-7331	**PAPER**
Jill Leale OFFICE: 616-696-1606 FAX: *lealej@river.it.gvsu.edu*	**Grand Valley State University** 14990 Simmons Ave. Cedar Springs MI 49319	
Lawrence A. Lee OFFICE: 919-560-6344 FAX: 919-530-7973 *larrylee@wpo.nccu.edu*	**North Carolina Central University** 1801 Fayetteville Street Durham NC 27707	
Sang-Suk Lee OFFICE: 614-292-0585 FAX: 614-292-7688 *lee.1099@osu.edu*	**The Ohio State University** 223 Mendenhall Lab., 125 South Oval Mall Columbus OH 43210	
Jay Lee, Ph.D. OFFICE: 3306723222 FAX: *jlee@kent.edu*	**Dept. of Geography** Kent State University Kent OH 44242-0001 USA	**PAPER**
David R. Legates, Ph.D. OFFICE: 302-831-2294 FAX: 302-831-6654 *legates@udel.edu*	**University of Delaware, Center for Climatic Research** 216 Pearson Hall Newark DE 19716-2541	**PAPER**
Geoff Leking OFFICE: 419-373-3092 FAX: 419-373-3125 *geoff.leking@epa.state.oh.us*	**OHIO EPA** 347 N. Dunbridge Rd. Bowling Green OH 42402	

April Lewis
OFFICE: 937-255-5654x3512
FAX: 937-255-4645
april.lewis@afit.af.mil

Air Force Institute of Technology
AFITICEL 2950 P Street Bldg. 643
Wright Patterson AFB OH 45433-7765

Norma Lewis
OFFICE: 513-569-7665
FAX: 513-569-7471
lewis.norma.@epamail.epa.gov

U.S. EPA
26 W. Martin Luther King Drive
Cincinnati OH 45268

Xaan Li
OFFICE: 614-292-0585
FAX: 614-292-7688
li@geology.ohio-state,edu

Dept. of Geological Studies Ohio State U.
Columbus OH 63210

Jun Liang
OFFICE: 513-569-7619
FAX:
liang_jun@hotmail.com

EPA NERL GIS Lab
435 Riddle Rd. Apt.9
Cincinnati OH 45220

Susan Licher
OFFICE:
FAX:
sslicher@hotmail.com

Ball State Graduate Student

Jason Linn
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental Management
Ball State University
Muncie IN 47306

Beth Linton
OFFICE: 770-677-0330
FAX: 770-399-0535

United Capitol Insurance Company
400 Perimeter Center Terrace, Suite 400
Atlanta GA 30346

Amy Liu
OFFICE: 5135697171
FAX:
liu.amy@epamail.epa.gov

PAI/SAIC, US EPA, NRMRL **PAPER**
MS 690, 26 W. Martin Luther King Dr
Cincinnati OH 45268
USA

Lin Liu, Ph.D.
OFFICE: 513-556-3429
FAX: 513-556-3370
lin.liu@uc.edu

University of Cincinnati
Department of Geography
714 Swift Hall
Cincinnati OH 45221-0131

Bill Lohner
OFFICE: 937-285-6051
FAX: 937-285-6404
bill.lohner@epa.state.oh.us

Ohio EPA **PAPER**
401 E. 5th Street
Dayton OH 45402

Cornell Long
OFFICE: 210-536-6121
FAX: 210-536-1130
cornell.long@brooks.af.mil

USAF Institute for ESOH Risk Analysis
2513 Kennedy Circle
Brooks AFB TX 78235-5123

Colleen Lovett
OFFICE: 937-656-3637
FAX: 937-656-3663
colleen.lovett@wpafb.af.mil

U.S. Air Force
4225 Logistics Ave., Suite 23
Wright - Patterson AFB OH 45433-5762

Damon Lowe
OFFICE: 317-598-8742
FAX:
dlowe@worldnet.att.net

Ball State University
6836 Challenge Lane
Indianapolis IN 46250

James A. Lynch, Ph.D.
OFFICE: 814-865-8830
FAX: 814-863-7193
jal@psu.edu

Penn State University
311 Forest Resource Lab
University Park PA 16802

Stan Lynn
OFFICE: 513-564-8349
FAX: 513-241-0354
lynns@ttemi.com

Tetra Tech EM Inc.
625 Eden Park Drive, Suite 100
Cincinnati OH 45202

Debbie E. Maes
OFFICE: 505-846-8568
FAX: 505-853-1793
debbie.maes@ao.dtra.mil

Defense Threat Reduction Agency
1680 Texas Street, SE
Kirtland AFB NM 87117-5669

Robert Magai
OFFICE: 573-522-3779
FAX: 573-525-5797
nrmagar@mail.dnr.state.mo.us

Missouri DNR
205 Jefferson Street
Jefferson City MO 65102

Laura Mahoney
OFFICE: 615-255-2370x406
FAX: 615-256-8332
lmahoney@brwncald.com

Brown & Caldwell
227 French Landing Drive
Nashville TN 37228

Sarada Majumder
OFFICE:
FAX:
majumder.sarada@epa.gov

SBI Env./USEPA
26 W. Martin Luther King Drive
Cincinnati OH 45219

Miguel A. Maldonado
OFFICE: 787-764-8824
FAX: 787-766-0150
jcaemer@prtc.net

PR EQB
P.O. Box 11488
San Juan PR 00910

Richard M. Males
OFFICE: 513-871-8566
FAX:
males@iac.net

RMM Technical Services, Inc.
3319 Eastside Avenue
Cincinnati OH 45208

Christopher Mantia
OFFICE: 937-776-7724
FAX: 937-586-3059
cmantia@corbus.com

Corbus, LLC
33 W. First Street, Suite 300
Dayton OH

Renante Marante
OFFICE: 312-742-0123
FAX: 312-744-6451
rmarante@ci.chi.il.us
Dept. of Environment Chicago Brownfields
30 N. La Salle, Room 2500
Chicago IL 6060

Joanne Markert
OFFICE: 3607057444
FAX:
markertj@wsdot.wa.gov
Washington Dept. of Transportation
310 Maple Park Ave., SE
Olympia WA 98504-7331
USA

Lawrence Martin
OFFICE: 202-564-6497
FAX: 202-565-2926
martin.lawrence@epa.gov
EPA / ORD
401 M Street. NW (8103R)
Washington DC 20460

Robert C. Martin, Jr.
OFFICE: 404-894-8446
FAX: 404-894-2184
bob.martin@gtri.gatech.edu
Georgia Tech Research Institute - Electro-Optics
Environment and Materials Laboratory
Atlanta GA 30332-0837

Margaret B. Martin, P.E.
OFFICE: 410-962-3500
FAX: 410-962-2318
margaret.b.martin@nab02.usace.army
USACE / HTRW Design Center
P.O. Box 1715
ATTN: CENAB-EN_
Baltimore MD 21203

Dana Martin-Hayden
OFFICE: 419-373-3067
FAX: 419-352-8468
dana.martin-hayden@epa.state.oh.us
OHIO - EPA
347 North Dunbridge Road
Bowling Green OH 43402

Carmen Maso
OFFICE: 312-886-1070
FAX: 312-353-6519
maso.carmen@epa.gov
US EPA
77 West Jackson
Chicago IL 60604

Jason Masoner
OFFICE: 580-436-8508
FAX:
jmasoner@usgs.gov
USGS
P.O. 1198
Ada OK 74820

Gregg Matthews
OFFICE: 613-541-6000x6099
FAX: 613-541-6596
matthews-g@rmc.ca
Environmental Sciences Group, RMC
P.O. Box 17000 Stn. Forces, Bldg. 62
Kingston - Canada ON K7K7B4

Robert Matzner
OFFICE: 703-305-5975
FAX: 703-305-6309
matzner.robert@epa.gov
US EPA/OPPTS/OPP/EFED
401 M St., SW
Mail Code 7507c
Washigton DC 20460

Dr. Eric F. Maurer
OFFICE: 606-257-7652
FAX: 606-257-1717
efmaur@pop.uky.edu
Univ. of Kentucky, Biological Sciences
101 Morgan Building
Lexington KY 40506

Chris Mazza
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu
Department of Natural Resources and Environmental
Ball State University
Muncie IN 47306

R.B. (Rusty) McAllister
OFFICE: 501-682-0022
FAX: 501-682-0010
mcallister@adeq.state.ar.us
Arkansas Dept. of Environmental Quality
8001 National Drive
Little Rock AR 72209

Andy McCammack
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu
Department of Natural Resources and Environmental
Ball State University
Muncie IN 47306

Larry McCanoless
OFFICE: 513-648-3772
FAX: 513-648-4029
Flour Daniel Fernald
P.O. Box 538704
Cincinnati OH 45253

Jeff McCormack
OFFICE: 810-239-1154
FAX: 810-239-1180
jmccormack@ssoe.com
SSOE, Inc.
111 E. Court St.
Flint MI 48502

Seth McCoy
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu
Department of Natural Resources and Environmental
Ball State University
Muncie IN 47306

George McKee
OFFICE: 513-684-3073
FAX: 513-684-3039
george.m.mckee@1rdor.usace.army.mil
Army Corps of Engineers
5717 Nickview Drive
Cincinnati OH 45247

Iver McLeod
OFFICE: 207-287-8010
FAX: 207-287-7826
iver.j.mcleod@state.me.us
Maine Department of Environmental Protection
State House Station #17
Augusta ME 04333

Christine McMahon
OFFICE: 6186921478
FAX:
mcmahon@icon-stl.net
209 N. Fillmore Street
Edwardville IL 62025
USA

Robert B. McMaster
OFFICE: 6126259883
FAX:
mcmaster@atlas.socsci.umn.edu
Dept. of Geography
414 Social Sciences Bldg.
Minneapolis MN 55455
USA

Stephanie McSpirit
OFFICE: 606-622-3070
FAX:
antmcspi@acs.eku
Eastern Ky. University
223 Keith
Richmond KY 40475

Jeremy Mennis
OFFICE:
FAX:
jmennis@gis.psu.edu
Dept. of Geography, Pennsylvania State Univ. **PAPER**
302 Walker Building
University Park PA 16802
USA

Jim Metzger
OFFICE: 812-665-2207
FAX: 812-665-5041
jmetzger@osmre.gov
Indiana DNR/Div. of Reclamation
RR 2 Box 129
Jascksonville IN 47438

Jim Meyer
OFFICE: 573-882-9291
FAX: 573-884-2199
jcmeyer@showme.missouri.edu
CARES - University of Missouri
130 Mumford Hall
Columbia MO 65211-6200

Michael V. Miller
OFFICE: 5417524271
FAX:
mmiller@ch2m.com
CH2M Hill **PAPER**
2300 NW Walnut Blvd.
Corvallis OR 97330
USA

Bruce Milligan
OFFICE: 317-290-3200x347
FAX: 317-290-3225
bruce.milligan@in.usda.gov
Natural Resources Conservation Service
6013 Lakeside Blvd.
Indianapolis IN 46278

Scott Minamyer
OFFICE: (513) 569-7175
FAX: (513) 569-7585
minamyer.scott@epa.gov
U.S. EPA, CERI, TTB
26 W. Martin Luther King Dr.
Cincinnati OH 45268

Ken Mix
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu
Department of Natural Resources and Environmental Management
Ball State University
Muncie IN 47306

Sara Moola
OFFICE:
FAX:
smoola@esri.com
ESRI
2205 A West Balboa Blvd.
Newport Beach CA 92663

Mike Moore
OFFICE: 717-783-7258
FAX: 717-783-7267
mmoore@dcnr.state.pa.us
PA Geological Survey
1500 North 3rd St.
Harrisburg PA 17102-1910

Shawn Moore
OFFICE: 937-255-5654
FAX: 937-255-4645
shawn.moore@afit.af.mil
AFIT / CEM
2950 P Street Bldg. 643
Wright Patterson AFB OH 45433-7765

Maggie Moorhouse
OFFICE: 361-883-6016
FAX: 361-883-7417
maggie@moorhousecc.com
Moorhouse Associates, Inc.
5826 Bear Lane
Corpus Christi TX 78405

Bruce Motsch
OFFICE: 614-265-6772
FAX:
bruce.motsch@dnr.state.oh.us
Ohio Department of Natural Resources
1952 Belcher Drive C-2
Columbus OH 43224

Erin T. Mutch
OFFICE: 209-532-0361
FAX: 209-532-0773
emutch@condorearth.com
Condor Earth Technologies, Inc.
21663 Brian Lane
Sonora CA 95370

Wolf Naegeli
OFFICE: 423-584-4806
FAX:
wnn@utk.edu
Energy, Environment & Resource Center, U of TN
4425 Balraj Ln
Knoxville TN 37921-2938

Jill Neal
OFFICE: (513) 569-7277
FAX:
niel.jill@epamail.epa.gov
U.S. EPA, ORD, NRMRL **PAPER**
26 W. Martin Luther King Dr.
Cincinnati OH 45268

Dean Nelson
OFFICE: 423-483-3191
FAX: 423-483-8617
dnelson@usit.net
Peer Consultants, P.C.
78 Mitchell Road
Oak Ridge TN 37830

Tamas Nemeth
OFFICE:
FAX:
lacus@rissac.hu
GIS Lab, Research Inst. For Soil Science & Agri. Chemistry
Hungarian Academy of Sciences
H-1022 Bud Herman Otto ut 15
Hungary

Bruce Nielsen
OFFICE: 317-290-3200
FAX: 317-290-3225
bruce.nielsen@in.usda.gov
USDA-Natural Resources Conservation Service
6013 Lakeside Blvd.
Indianapolis IN 46278

Douglas A. Nieman
OFFICE:
FAX:
Dnieman@normandeau.com
Normandeau Associates **PAPER**
3450 Schuylkill Rd.
Spring City PA 19475
USA

Eugene O'Brien
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu
Department of Natural Resources and Environmental Management
Ball State University
Muncie IN 47306

Toby Obenauer
OFFICE: 305-242-7892
FAX: 305-242-7836
toby_obenauer@nps.gov
National Park Service - Everglades
40001 SR 9336
Homestead FL 33034-6733

Ramon A. Olivero
OFFICE: 919-572-2764
FAX: 919-572-2765
rolivero@lmepo.com
Lockheed Martin Environmental Services **PAPER**
100 Capitola Drive, Suite 111
Durham NC 27713

Deborah Olszowka
OFFICE: 513-231-7719
FAX: 513-231-7761
dmo@orsanco.org
ORSANCA
5735 Kellogg Avenue
Cincinnati OH 45228

Lindell Ormsbee
OFFICE: 606-257-1302
FAX:
lormsbee@engr.uky.edu
University of Kentucky
Dept. of Civil Engineering
Lexington KY 40506-0281

Steven D. Ortiz
OFFICE: 512-239-2008
FAX: 512-239-4007
sortiz@tnrcc.state.tx.us
TX Natural Resource Conservation Commission
12100 Park 35 Circle, MC - 108
Austin TX 78753

Kurt Overmyer
OFFICE: 616-336-3086
FAX: 616-336-2436
overmyer@iserv.net
Kent County Health Department
700 Fuller, N.E.
Grand Rapids MI 49503

David A. Padgett
OFFICE: 615-963-5471
FAX: 734-939-5813
padgettdavid@netscape.net
Tennessee State University **PAPER**
3500 John A. Merritt Blvd.
Nashville TN 37219

David Padgett, Ph.D.
OFFICE: 4407758747
FAX:
david.padgett@oberlin.edu
Oberlin College
USA

Mercedes Padilla
OFFICE: 787-767-8181x2243
FAX: 787-766-0150
jcaemer@prtc.net
PR Environemtal Quality Board
P.O. Box 192785
San Juan PR 00919

Susan Pagano
OFFICE: 610-696-4150
FAX: 610-696-8608
spagano@thcahill.com
Cahill Associates
104 South High Street
West Chester PA 19382

Harry Parrott
OFFICE: 414-297-3342
FAX: 414-297-3808
hparrott/r9@fs.fed.us
USDA-Forest Service
310 W. Wisconsin Ave.
Milwaukee WI 53203

Laszlo Pasztor, Ph.D.
OFFICE: 361-356-3694
FAX: 361-212-1891
lacus@rissac.hu
GIS Lab, Research Inst. for Soil Science & Agri. Chemistry
Hungarian Academy of Sciences
Herman Otto ut 15
H-1022 Bud Hungary

Clair (Pat) Patterson
OFFICE: 410-260-8512
FAX: 410-260-8595
ppatterson@dnr.state.md.us
Maryland DNR-Forest Service
Tawes State Office Building E1 580 Taylor Avenue
Annapolis MD 21401

Michael J. Pavilonis OFFICE: 412-442-4066 FAX: 412-442-4194 *pavilonis.michael@dep.state.pa.us*	**Pennsylvania DEP** 400 Waterfront Drive Pittsburgh PA 15222-4745
Alan Peacock OFFICE: 317-337-3414 FAX: 317-337-3235 *apeacock@dowagro.com*	**Dow AgroSciences** 9330 Zionsville Road Indianapolis IN 46268
Norman E. Peters OFFICE: 770-903-9145 FAX: 770-903-9199 *npeters@usgs.gov*	**U.S. Geological Survey** 3039 Amwiler Rd., Suite 130 Atlanta GA 30360-2824
Carl Petersen OFFICE: 502-564-8390 FAX: 502-564-2088 *cpete00@pop.uky.edu*	**Unversity of Kentucky** 1854 Bellefonte Drive Lexington KY 40503
Herb Petitjean OFFICE: 502-564-6717 FAX: 502-564-5096 *herb.petitjean@mail.state.ky.us*	**Kentucky Dept. for Environmental Protection** 14 Reilly Road Frankfort KY 40601
Rebecca Petty OFFICE: 614-466-4801 FAX: 614-466-4556 *rpetty@gw.odh.state.oh.us*	**Ohio Department of Health** 246 N. High Street, 5th Floor Columbus OH 43266-0588
Kenneth Pew OFFICE: 216-881-6600x814 FAX: 216-881-2738 *pewk@neorsd.org*	**N.E. Ohio Regional Sewer District** 4415 Euclid Avenue Cleveland OH 44103
Robert R. Pierce OFFICE: 770-903-9113 FAX: 770-903-9199 *rrpierce@usgs.gov*	**USGS/WRD** 3039 Amwiler Road, Suite 130 Atlanta GA 30360-2824
Michael Plastino OFFICE: 202-260-0048 FAX: 202-260-7926 *plastino.michael@epa.gov*	**EPA Office of Water** 401 M Street, SW (MC 4102) Washington DC 20460
Mark Porembka OFFICE: 412-442-4327 FAX: 412-442-3428 *porembka.mark@dep.state.pa.us*	**PA Dept. of Environmental Protection** 400 Waterfront Drive Pittsburgh PA 15222
Jacob Prater OFFICE: 765-759-6898 FAX:	**Ball State University** 6613 W. Jackson St. Muncie IN 47304

Garry Price
OFFICE: 717-772-5630
FAX: 717-787-5259
price.garry@dep.state.pa.us

PA Dep.
400 Market Street
Harrisburg PA 17105

Eric R. Pruitt
OFFICE: 812-323-8871
FAX:
erpruitt@indiana.edu

Indiana University
714 E Cottage Grove West
Bloomington IN 47408

Tom Quinn
OFFICE: 513-489-6611
FAX: 513-489-6619
tomq@pes.cin.com

P.E.S.
7209 E Kemper Rd.
Cincinnati OH 45241

Nancy J. Radle
OFFICE: 651-779-5102
FAX: 651-779-5109
nancy.radle@dot.state.mn.us

Minnesota Department of Transportation
3485 Hadley Avenue, N -- Mail Stop 620
Oakdale MN 55128

Charles Rairdan
OFFICE: 916-557-7833
FAX: 916-557-7848
crairdan@spk.usace.army.mil

USACE - Sacramento District **PAPER**
1325 "J" Street
Sacramento CA 95814

Alexandrine Randriamahefa
OFFICE: 256-726-7436
FAX: 256-726-7055
drrand@oakwood.edu

Oakwood College
7000 Adventist Blvd.
Huntsville AL 35896

Ravi Rao
OFFICE: 404-562-8349
FAX:
rao.ravi@epamail.epa.gov

USEPA Region 4, OPM, PAB
61 Forsyth Street
Atlanta GA 30303

Alan Rao, Ph.D.
OFFICE: 6179241770
FAX:
arao@vhb.com

Vanasse Hangen Brustlin, Inc.
101 Walnut Street
P.O. Box 915
Watertown MA 02471
USA

Andrew Rawnsley
OFFICE:
FAX:
ronz@ravensfield.com

Ravensfield Geographic Resources, Ltd. **PAPER**
P.O. Box 410
Granville OH 43023
USA

Alan Rea
OFFICE: 208-387-1323
FAX: 208-387-1372
ahrea@usgs.gov

US Geological Survey
230 Collins Road
Boise ID 83702-4520

Jeffrey Reese
OFFICE: 765-213-1269
FAX: 765-747-7744
masine@iquest.net

Delaware County, Indiana GIS
100 W. Main St., Room 204
Muncie IN 47305

Jon Reid
OFFICE: 513-558-1723
FAX:
jon.reid@uc.edu.
University of Cincinnati
3223 Eden Ave. ML 056
Cincinnati OH 45267-0056

Chad K. Rhodes
OFFICE: 313-237-3093
FAX: 313-224-1547
rhodesc@EnvAfrs.ci.detroit.mi.us
City of Detroit Dept. of Environmental Affairs
660 Woodward Avenue, Suite 1590
Detroit MI 48226

Carl Rich
OFFICE: 303-202-4316
FAX: 303-202-4354
clrich@usgs.gov
U.S. Geological Survey
P.O. Box 25046, MS-516, Denver Federal Center
Denver CO 80225

Ian Richardson
OFFICE: 519-884-0510
FAX: 519-884-0525
irichardson@rover.com
CRA
651 Colby Dr.
Waterloo
Ontario

John Richardson
OFFICE: 402-562-8290
FAX: 404-562-8269
richardson.john@epa.gov
US EPA Region 4
61 Forsyth Street
Atlanta GA 30303

James M. Rine, Ph.D.
OFFICE: 803-777-7792
FAX: 803-777-6437
jrine@esri.esri.sc.edu
Earth Sciences & Resources Institute **PAPER**
University of South Carolina
901 Sumter Road, Room 401
Columbia SC 29208

Jim Rocco
OFFICE: 330-562-9391
FAX: 330-562-9391
roccojl@worldnet.att.net
Sage Risk Solutions LLC
360 Heritage Road
Aurora OH 44202

Edward Rogers, Jr.
OFFICE: 9085982600
FAX:
erogers@bemsys.com
BEM Systems, Inc. **PAPER**
100 Passaic Ave.
Chatham NJ 07928
USA

Cynthia Root
OFFICE: 319-398-5678
FAX: 319-398-1250
croot@kirkwood.cc.ia.us
Kirkwood Community College
6301 Kirkwood Blvd. SW
Cedar Rapids IA 52406

Lloyd Ross
OFFICE: 216-739-0555
FAX: 216-739-0560
lsr@haleyaldrich.com
Haley & Aldrich, Inc.
5755 Granger Road, Suite 100
Independence OH 44131

Randall Ross, Ph.D.
OFFICE: (580) 436-8611
FAX:
ross.randall@epamail.epa.gov
R.S. Kerr Environ. Research Center U.S. EPA/ORD/NRMRL **PAPER**
919 Kerr Research Dr., P.O. Box 119
Ada OK 74820

Holly Roten
OFFICE: 419-353-1531
FAX:
rotenh@bgnet.bgsu.edu

Bowling Green State University
455 South Enterprise, Apartment H.
Bowling Green OH 43402

Chris Rudnick
OFFICE: 502-564-6716
FAX: 502-564-5096
chris.rudnick@mail.state.ky.us

Kentucky Environmental Protection-State Superfund
14 Reilly Road
Frankfort KY 40601

Nel Ruffin
OFFICE: 606-257-1299
FAX:
nruffin@engr.uky.edu

Kentucky Water Resources Research Institute
233 Mining & Minerals Resource Building
University of Kentucky
Lexington KY 40506

Steve Saines
OFFICE: 614-644-2752
FAX: 614-644-2909

Ohio EPA
P.O. Box 1049
Columbus OH 43216

Scott Samson
OFFICE: 606-257-3767
FAX: 606-257-4354
ssamson@pop.uky.edu

University of Kentucky
Rural Sociology Program, Garrigus 500
Lexington KY 40546

Vicki Sandiford
OFFICE: 919-541-2629
FAX: 919-541-7690
sandiford.vicki@epa.gov

US EPA
MD-13
RTP NC 27711

George Sarapa
OFFICE: 412-262-5400
FAX: 412-262-3036
sarapg@lrkimball.com

L. Robert Kimball & Associates, Inc.
415 Moon Clinton Road
Moon Township PA 15108

Gary Schaal
OFFICE: 614-265-6769
FAX: 614-267-2981
gary.schaal@dnr.state.oh.us

Ohio Department of Natural Resources
1952 Belcher Drive C-4
Columbus OH 43224

Devin Scheak
OFFICE: 513-921-5124
FAX: 513-921-5136
IMAGO@ONE.net

IMAGO, Inc.
553 Enright Avenue
Cincinnati OH 45205

Tom Schneider
OFFICE: 937-295-6466
FAX: 937-285-6404
tom.schneider@epa.state.oh.us

Ohio EPA
401 E. 5th Street
Dayton OH 45402

Theresa Schnorr
OFFICE: 513-558-5984
FAX:
ths1@cdc.gov

N.I.O.S.H
4676 Columbia Parkway
Cincinnati OH 45220

Susan Schock
OFFICE: 513-569-7551
FAX: 513-487-2513
schock.sue@epamail.epa.gov

U.S. EPA / ORD / NRMRL
26 W. Martin Luther King Drive
Cincinnati OH 45268

Chris Schroeder
OFFICE: 402-441-6272
FAX: 402-441-8323
cschroed@ci.lincoln.ne.us

Lincoln-Lancaster County Health Dept.
3140 "N" Street
Lincoln NE 68510

Joe Schubauer-Berigan
OFFICE: 513-569-7734
FAX:
Schubauer-Berigan.joseph@epa.gov

U.S. EPA National Center for Envir. Assessment
26 W. Martin Luther King Drive
Cincinnati OH 45220

Susan Schultz
OFFICE: 513-861-7666
FAX: 513-559-3155
schultz23@fuse.net

Mill Creek Restoration Project
42 Calhoun Street
Cincinnati OH 45219

Robert Scott
OFFICE: 404-675-1753
FAX: 404-675-6246
bob_scott@mail.dnr.state.ga.us

Georgia EPD
4220 International Parkway, Suite 101
Atlanta GA 30354

Irena Scott, PhD
OFFICE: 614-644-8020
FAX: 614-644-7740
iscott@gw.odh.state.us

Ohio Department of Health
246 North High Street
Columbus OH 43266-0118

Tom Seibert
OFFICE: 502-564-6716
FAX: 502-564-7484
tom.seibert@mail.state.ky.us

Kentucky Division for Waste Management
14 Reilly Road
Frankfort KY 40601

Gabriel Senay, Ph.D.
OFFICE: 5135697096
FAX:
SENAY.GABRIEL@epa.gov

PAI/SAIC, US EPA **PAPER**
26 W. Martin Luther King Dr.
Cincinnati OH 45268
USA

Barbara Seymour
OFFICE: 301-622-9696x108
FAX: 301-622-9693
bseymour@plexsc.com

Plexus Scientific
12501 Prosperity Drive, # 401
Silver Spring MD 20904

Jo Ann Shaw
OFFICE: 573-751-9370
FAX:
nrshawj@mail.dnr.state.mo.us

MO Dept. of Natural Resources
205 Jefferson St.
Jefferson City MO 65102

Jerry Shi
OFFICE: 614-292-6193
FAX: 614-292-7688
shi.19@osu.edu

The Ohio State University
125 South Oval Mall
Columbus OH 43210

Andy Shirmeyer
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental
Ball State University
Muncie IN 47306

Massoud Shoa
OFFICE: 502-564-6717x219
FAX: 502-564-2705
massoud.shoa@mail.state.ky.us

Kentucky Division of Waste Management
14 Reilly Road
Frankfort KY 40601

Leila Shultz
OFFICE: 617-327-4294
FAX: 617-327-4294
shultz@oeb.harvard.edu

Harvard University
22 Divinity Avenue
Cambridge MA 02138

Amardeep Singh
OFFICE: 916-653-2726
FAX: 916-653-6366
asingh@fnd.csus.edu

CSUS
905 23rd Street #3
Sacramento CA 95816

Nathan Sloan
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental
Ball State University
Muncie IN 47306

Ryan Sloan
OFFICE: 765-289-7816
FAX:
resloan@hotmail.com

Dept. of Natural Resources , Ball State Univ.
323 1/2
Muncie IN 47304

Mike Sloop
OFFICE: 770-673-3624
FAX: 770-396-9495
msloop@brwncald.com

Brown and Caldwell
41 Perimeter Center East
Suite 400
Atlanta GA 30346

Kathleen Smaluk-Nix
OFFICE: 502-574-1358
FAX: 502-574-1389
ksmaluk@louky.org

Office of Health and Environment
600 W. Main St. , 4th Floor
Louisville KY 40202

Art Smith
OFFICE: 502-574-2511
FAX: 502-574-1389
asmith@louky.org

City of Louisville Office of Health and Environment
600 W. Maine Street, 4th Floor
Louisville KY 40202

Christopher Smith
OFFICE: 608-221-6330
FAX: 608-221-6353

Wisconsin DNR
1350 Femrite Drive
Monona WI 53716

Gerry Snyder
OFFICE: 303-312-6623
FAX: 303-312-6063
snyder.gerry@epamail.epa.gov

US EPA
999 18th Street, Suite 500, 8IG
Denver CO 80202-2466

Marilyn Sommer
OFFICE: 502-564-3256
FAX:
maras425@ntr.net

Kentucky Resource Cabinet
1933 Deerwood
Louisville KY 40210

Puneet Srivastava
OFFICE: 501-682-0018
FAX: 501-682-0010

ADEQ
800 National Drive
P.O. Box 8913
Little Rock AR 72205

Greg Starkebaum
OFFICE: 303-763-7188
FAX: 303-763-4896
gstarkebaum@techlawinc.com

TechLaw, Inc.
300 Union Blvd. #600
Lakewood CO 80228

Steven K. Starrett, Ph.D
OFFICE: 785-532-1583
FAX: 785-532-7717
steveks@ksu.edu

Kansas State University
160 Seaton Hall
Manhattan KS 66506-2905

Dr. Gerry L.
OFFICE: 301-415-5265
FAX: 301-415-5348
gls3@nrc.gov

U.S. Nuclear Regulatory Commission
TWFN-MS T7C6-11545 Rockville Pike
Rockville MD 20852

Joel Stocker
OFFICE: 860-345-4511
FAX: 860-345-3357
jstocker@canr.uconn.edu

UConn CES, NEMO Project
1066 Saybrook Rd., Box 70
Haddam CT 06438-0070

Brad Stone
OFFICE: 502-564-6716
FAX: 502-564-4049
brad.stone@mail.state.ky.us

KY Div. of Waste Management
14 Reilly Road
Frankfort KY 40601

William Story
OFFICE: 775-687-4670
FAX: 775-687-6396
bstory@ndep.carson-city.nv.us

Nevada Division of Environmental Protection
333 W. Nye Lane
Carson City NV 89706

Steven V. Strausbauch
OFFICE: 210-536-6134
FAX: 210-536-1130
steven.strausbauch@brooks.af.mil

IERRA/RSRE
2513 Kennedy CIR
Brooks AFB TX 78235

Ryan Stults
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental
Ball State University
Muncie IN 47306

Prathiba Subramaniam
OFFICE: 513-474-9780
FAX:
prathiba@gqc.com

Global Quality Corporation
3304 Jefferson Avenue #104
Cincinnati OH 45220

Bhagya Subramanian
OFFICE: 5135697349
FAX:
subramanian.bhagya@epa.gov
US EPA / NERL **PAPER**
26 W. Martin Luther King Dr.
Cincinnati OH 45268
USA

Parrish Swearingen
OFFICE: 912-926-1197x114
FAX: 912-926-9642
parrish.swearingen@robins.af.mil
WR-ALC/EMX
455 Byron Street
Robins AFB GA 31098

Jane C. Thapa
OFFICE: 518-402-7713
FAX: 518-402-7599
jct02@health.state.ny.us
NY DOH
547 River Street, Room 400
Troy NY 12180-2216

Geoff Thompson
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu
Department of Natural Resources and Environmental Management
Ball State University
Muncie IN 47306

Gordon Thompson
OFFICE: 513-782-4556
FAX:
gwthompson@theitgroup.com
The IT Group
312 Directors Drive
Knoxville TN 37923

Richard E. Thornburg
OFFICE: 513-564-1785
FAX: 513-564-1776
Rick.Thornburg@cinhlthe.rcc.org
Cincinnati Health Dept.
3845 Wm. P. Dooley By-Pass
Cincinnati OH 45223

Larry Tinney
OFFICE: 702-897-3270
FAX: 702-897-3285
ltinney@lmepo.com
Lockheed Martin Environmental Services
980 Kelly Johnson Dr.
Las Vegas NV 89119

Alejandro F. Tongco
OFFICE: 405-744-8974
FAX: 405-744-7008
al@seic.1se.okstate.edu
Oklahoma Spatial & Environmental Information Clearinghouse, OSU
201 CITD
Stillwater OK 74076

Oury Traore
OFFICE: 606-986-2373
FAX: 606-986-2619
otraore@maced.org.
MACED
433 Chestnut St.
Berea KY 40403

Michael E. Troyer, Ph.D.
OFFICE: (513) 569-7399
FAX:
troyer.michael@epamail.epa.gov
U.S. EPA, ORD, NRMRL **PAPER**
26 W. Martin Luther King Dr.
Cincinnati OH 45268

Amy Tucker
OFFICE: 765-281-1026
FAX:
amyetucker@prodigy.net
Ball State University
100 S. Calvert Ave.
Muncie IN 47303

W. Lance Turley — **Hull and Associates, Inc.** — **PAPER**
OFFICE: 513-459-9677
FAX: 513-459-9869
lturley@hullinc.com
4700 Duke Drive, Suite 172
Mason OH 45040

Rachel Tyler — **Florida Rural Community Assistance Project**
OFFICE: 352-378-6517
FAX:
faca@aol.com
6212 NW 43rd St. Suite A
Gainesville FL 32697

Tom Ungar — **Montgomery Watson**
OFFICE: 216-623-2738
FAX: 216-621-4972
thomas.ungar@mw.com
1300 East 9th Street, Suite 2000
Cleveland OH 44114

Andre van Delft — **TNO Building and Construction Research**
OFFICE: 31152695286
FAX: 31152695335
a.vandelft@bouw.tno.nl
P.O. Box 49
Delft - Netherlands 2600 AA

Laurens Van der Tak — **CH2M Hill** — **PAPER**
OFFICE: 703-471-6405
FAX: 703-471-1508
lvandert@ch2m.com
13921 Park Center Rd., Suite 60
Herndon VA 20171
USA

Robert J. van Waasbergen — **Applied Environmental Data Services**
OFFICE: 408-737-7697
FAX: 408-737-7978
robertvw@aeds.com
201 W. California Ave. #206
Sunnyvale CA 94086

Dirk Vandervoort — **Concurrent Technologies Corporation**
OFFICE: 360-475-6915
FAX: 360-475-6901
vandervoort@ctc.com
510 Washington Avenue, Suite 120
Bremerton WA 93777-1844

James M. Vanek — **PA Dept. of Environmental Protection**
OFFICE: 412-442-4031
FAX: 412-442-4328
vanek.james@dep.state.pa.us
400 Waterfront Drive
Pittsburgh PA 15222

Paul Vermaaten — **City of Jackson**
OFFICE: 517-788-4075
FAX: 517-788-4641
161 W. Michigan Ave.
Jackson MI 49201

Janet Vick — **Gwinnett County**
OFFICE: 770-822-7473
FAX: 770-822-7479
75 Langley Drive
Lawrenceville GA 30045

Todd Vikan — **KPMG LLP**
OFFICE: 937-259-9850
FAX: 937-259-9869
tvikan@kpmg.com
3139 Research Blvd. Suite 200
Dayton OH 45420

Jack Wachter
OFFICE: 513-352-6992
FAX: 513-352-4970
jack.wachter@cinems.rcc.org
City of Cincinnati
Office of Environmental Mgmt.
805 Central Ave., Suite 610
Cincinnati OH 45202-1947

Scott Wade
OFFICE: 734-332-1200
FAX: 734-332-1212
swade@limno.com
LIMNO-Tech, Inc.
501 Avis Drive
Ann Arbor MI 48108

Tim Wade
OFFICE: 702-798-2117
FAX: 702-798-2208
wade.timothy@epa.gov
U.S. EPA
944 E. Harmon Avenue
Las Vegas NV 89119

Jerry Wager
OFFICE: 614-265-6619
FAX: 614-262-2064
jerry.wager@dnr.state.oh.us
Ohio Department of Natural Resources
1939 Fountain Square Ct. E-2
Columbus OH 43224

James Wang
OFFICE: 404-562-8280
FAX:
wang.james@epamail.epa.gov
USEPA Region 4
61 Forsyth Street
Atlanta GA 30303

Lizhu Wang
OFFICE: 608-221-6335
FAX: 608-221-6353
wangl@mailo.dnr.state.wi.us
Wisconsin DNR
1350 Femrite Drive
Monona WI 53716

Xinhao Wang, Ph.D.
OFFICE: 513-556-0497
FAX: 513-556-1274
xinhao.wang@uc.edu
School of Planning, University of Cincinnati **PAPER**
6210 DAAP Building
Cincinnati OH 45221-0016
USA

Robin Wankum
OFFICE: 913-458-6538
FAX: 913-458-6633
wankumrd@bv.com
Black and Veatch Special Projects **PAPER**
6601 College Blvd.
Overland Park KS 66211

Stephen Want
OFFICE: 301-951-4681
FAX: 301-652-1273
stephen.want@lexis-nexis.com
ENVIRONMENTAL ABSTRACTS
Congressional Information Service, Inc.
4520 East-West Highway
Bethesda MD 20814-3389

Mark Warrell
OFFICE: 502-564-6716
FAX: 502-564-2705
mark.warrell@mail.state.ky.us
Kentucky Division of Waste Management
14 Reilly Road
Frankfort KY 40601

Steve Webb
OFFICE: 405-702-8195
FAX: 405-702-8101
steve.webb@deqmail.state.ok.us
Oklahoma Dept. of Environmental Quality
707 N. Robinson Ave.
Oklahoma City OK 73102

Kevin Wehrly
OFFICE: 513-569-7301
FAX: 513-569-7609
wehrly.kevin@epa.gov
USEPA-NERL MS-642
26 W. Martin Luther King Drive
Cincinnati OH 45268

Tom Weir
OFFICE: 215-685-9436
FAX: 215-685-7593
thomas.weir@phila.gov
Air Management Services
321 University Avenue / Spelman Building
Philadelphia PA

Joseph P. Wellner
OFFICE: 513-251-2730
FAX: 513-251-0200
wellnerj@ttnus.com
Tetra Tech NUS, Inc.
1930 Radcliff Drive
Cincinnati OH 45204

Russell Weniger
OFFICE: 210-671-4844
FAX: 210-671-2241
russell.weniger@lackland.af.mil
US Air Force
2240 Walker Avenue
Lackland AFB TX 78236-5637

Bill Wheaton
OFFICE: 919-541-6158
FAX: 919-541-5929
Research Triangle Industries
3040 Cornwallis Road - P.O. Box 12194
Research Triangle Park NC 27709-2194

Brad White
OFFICE: 937-384-4215
FAX: 937-384-4201
whiteb@mail.rfweston.com
Roy F. Weston, Inc.
2566 Kohnle Drive
Miamisburg OH 45342

Ron White
OFFICE: 513-648-5920
FAX: 513-648-4029
Ronald_white@fernald.gov
Fluor Daniel Fernald
P.O. Box 538704
Cincinnati OH 45253-8704

Charlotte White-Hull
OFFICE: 513-627-1197
FAX: 513-627-1208
whitehull.ce@pg.com
Procter & Gamble Company
11810 East Miami River Road, Room 1A04T BTF
Ross OH 45061

D. Whitmire
OFFICE: 765-289-1099
FAX:
Ball State University
6613 W. Jackson St.
Muncie IN 47304

Derik Whitmire
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu
Department of Natural Resources and Environmental
Ball State University
Muncie IN 47306

Duane Wilding
OFFICE: 410-974-7276
FAX: 410-974-7200
dwild@menv.com
Maryland Environmental Service
2011 Commerce Park Drive
Annapolis MD 21401

Frances Wildman
OFFICE: 309-782-7907
FAX: 309-782-5038
wildmanf@ria.army.mil

Rock Island Arsenal
5I0RI-SEV, Bldg 210
Rock Island IL 61299

Dave Williams
OFFICE: 270-827-3414
FAX: 270-827-1117
williams@kgs.mm.uky.edu

Kentucky Geological Survey
P.O. Box 653
Henderson KY 42419

Matt Williams
OFFICE: 765-285-2327
FAX: 765-285-2606
jeflin1@gw.bsu.edu

Department of Natural Resources and Environmental Management
Ball State University
Muncie IN 47306

Steve Williams
OFFICE: 614-644-2752
FAX: 614-644-2909

Ohio EPA
P.O. Box 1049
Columbus OH 43216-1049

Lesley Hay Wilson
OFFICE: 512-367-2952
FAX: 512-367-2953
hay_wilson@mail.utexas.edu

University of Texas at Austin **PAPER**
10100 Burnet Road, PRC Building 119
Austin TX 78758

James Wolfe
OFFICE: 304-696-6042
FAX: 304-696-5454
jawolfe@marshall.edu

Marshall University
400 Hal Greer Blvd.
Huntington WV 25755-2585

Tim T. Wright
OFFICE: 513-732-7499
FAX: 513-732-7969
CCGHD@FUSE.NET

Clermont County GHD
2275 Bauer Road Suite 300
Batavia OH 45103

Patricia Yaden
OFFICE: 502-564-6716
FAX: 502-564-4049
patty.yaden@mail.state.ky.us

NREPC
14 Reilly Road
Frankfort KY 40601

Thomas K. Yeager
OFFICE: 717-772-4018
FAX: 717-772-3249
yeager.thomas@dep.state.pa.us

Pennsylvania DEP-Water Supply Management
P.O. Box 8467, 11th Floor, RCSOB
Harrisburg PA 17105-8467

Greg Young
OFFICE: 513-576-0009
FAX: 513-576-9756
youngg@srwenvironmental.com

SRW Environmental Services
55 West TechneCenter, Suite C
Milford OH 45150

Pao-Chiang Yuan
OFFICE: 601-968-2466
FAX: 601-968-2344
pcyuan@yahoo.com

Jackson State University
P.O. Box 18489
Jackson MS 39217

Donald Zelazny
Platinum International Inc.
5350 Shawnee Road, Suite 200
Alexandria VA 22312
OFFICE: 703-916-7987
FAX: 703-916-7984
dzelazny@plaii.com

David Zellmer
Virginia Tech
428 Winston Avenue
Blacksburg VA 24060
OFFICE: 540-552-8185
FAX: 540-552-8189
dzellmer@vt.edu

Guang Zhao, Ph.D.
Information Technology Section, SC Dept. of Health & Env. Control
2600 Bull Street
Columbia SC 29201
USA
PAPER
OFFICE: 8037344833
FAX:
zhaog@columb30.dhec.state.sc.us

Peter Max Zimmerman
People's Counsel for Baltimore County
Courthouse 47, 400 Washington Ave.
Towson MD 21204
OFFICE: 410-887-2188
FAX: 410-887-3182

Environmental Problem Solving with Geographic Information Systems

September 21 - 23, 1999

This presentation was not made available.

Please contact the presenter with questions and comments.

Back to
1999 Agenda

The Watershed Assessment Project: Tools for Regional Problem Area Identification

Christine Adamus
St. Johns River Water Management District, Palatka, Florida

The St. Johns River Water Management District of Florida recently completed a major water resources planning effort. As part of this planning effort, the St. Johns River Water Management District created a geographic information systems (GIS) project called the Watershed Assessment, which included a nonpoint source pollution load model. This paper introduces the planning project and the Watershed Assessment, and describes how the results of the model are being used to guide water management activities in northeast Florida.

Background

The St. Johns River Water Management District (District), one of five water management districts in Florida, covers 12,600 square miles (see Figure 1). The St. Johns River starts at the southern end of the District and flows north; it enters the Atlantic Ocean east of the city of Jacksonville. The cities of Orlando, Daytona Beach, and Jacksonville are partially or entirely within the District boundaries. Ad valorem taxes provide primary funding for the District.

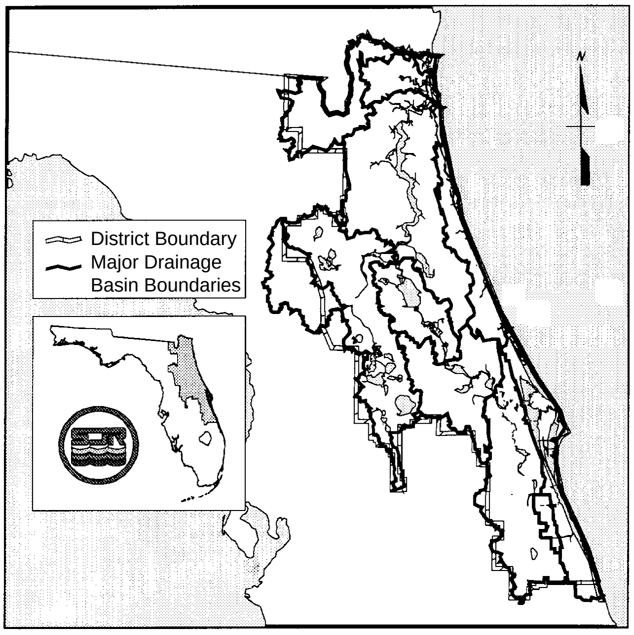


Figure 1. St. Johns River Water Management District, Florida.

The District boundaries are somewhat irregularly shaped because Florida water management districts are organized on hydrologic, not political, boundaries, which greatly improves the District's ability to manage the resources. On the north, the District shares the St. Mary's River with the state of Georgia and on the south, shares the Indian River Lagoon with another water management district. Most of the water bodies the District manages, however, have drainage basins that are entirely contained within the District's boundaries.

Water management districts in Florida have amassed extensive GIS libraries, which they share with local and statewide agencies. These libraries include basic data layers such as detailed land use, soils, and drainage basins. Districts also coordinate data collection and management to ensure data compatibility.

District Water Management Plan

All activities and programs of the water management districts are related to one or more of the following responsibilities: water supply, flood protection, water quality management, and natural systems management.

Each water management district recently completed a district water management plan (Plan). The main purpose of these Plans is to provide long-range guidance for the resolution of water management issues. The Florida Department of Environmental Protection will use these five Plans as the basis for a state water management plan. Each water management district used the same format, which comprised the following components:

- Resource assessment: What are the problems and issues related to each of the four responsibilities listed above?

- Options evaluation: What options are available for addressing the problems?
- Water management policies: What existing District policies influence the decisions that must be made?
- Implementation strategy: What is the best plan for addressing the problems?

The Watershed Assessment Project

The District created the Watershed Assessment project as part of is resource assessment. This GIS project examines the entire District to identify problems related to flood protection, ecosystems protection, and surface water quality.

The flood protection component is the only part of the Watershed Assessment that is not complete. It will involve simple overlays of floodplain boundaries with existing and future land use. Floodplain boundaries are defined as Federal Emergency Management Agency (FEMA) flood insurance rate map 100-year flood hazard areas. In many areas, these designations are not very accurate, yet we decided to proceed with their use because they are the best available information for many parts of the District. In areas where little hydrologic information is available and where the District has not conducted any related studies, the FEMA data are a helpful starting point. This echoes a theme of the Watershed Assessment project: the assessment is primarily intended to fill in gaps where we have not performed previous resource assessments, not to supplant existing information.

The ecosystems protection component of the Watershed Assessment is based heavily on a project identifying priority habitat in Florida, conducted by the Florida Game and Fresh Water Fish Commission (1). It is similar to gap analyses that the U.S. Fish and Wildlife Service currently is conducting in many parts of the country. For the Watershed Assessment, we modified the data somewhat and examined ways to protect the habitat in cooperation with local agencies.

The surface water quality component of the Watershed Assessment has two main parts. The first uses water quality data from stations that have been spatially referenced so that we can map them and combine the information with other information, such as the second part of the water quality component. This second part is a nonpoint source pollution load model, which is discussed in more detail below.

The Pollution Load Screening Model

The nonpoint source pollution load model is the Pollution Load Screening Model (PLSM), a commonly used screening tool in Florida. It is an empirical model that estimates annual loads to surface waters from stormwater runoff. Our goal in designing this model was to identify pollution load “problem areas” for examination in the Plan.

In these types of models, annual pollutant loads are a function of runoff volume and mean pollutant concentrations commonly found in runoff. Runoff volume varies with soil and land use, while pollutant concentrations vary with land use. For the PLSM, pollutant concentrations were derived from studies conducted solely in Florida. A report describing the model in detail is available (2).

Usually, this kind of model combines GIS with a spreadsheet: the GIS supplies important spatial information that is input into a spreadsheet where the actual calculations are made. The PLSM is different, however, because we programmed it entirely within GIS. The District’s GIS software is ARC/INFO, and the model employs an ARC/INFO module called GRID, which uses cell-based processing and has analytical capabilities (3). All the model calculations are done in the GIS software, resulting in a more flexible model with useful display capabilities.

Model input consists of grids, or data layers, with a relatively small cell size (less than 1/2 acre). We chose this cell size based on the minimum mapping unit of the most detailed input data layer (land use) and the need to retain the major road features. The model has four input grids: land use, soils, rainfall, and watershed boundaries. For any given cell, the model first calculates potential annual runoff based on the land use, soil, and rainfall in that cell. It then calculates annual loads by applying land-use-dependent pollutant concentrations to the runoff.

For this model:

- Land use is from 1:24,000-scale aerial photography flown in 1988 and 1989. The model incorporates 13 land use categories.
- Soils are the Soil Conservation Service (SCS) SSURGO database, which corresponds to the county soil surveys. The PLSM uses the hydrologic group designation of each soil type.
- Rainfall was taken from a network of long-term rainfall stations located throughout the District.
- Watersheds were delineated by the United States Geological Survey (USGS) on 1:24,000-scale, 7.5-minute maps and digitized.

Model output consists of a runoff grid and six pollutant load grids. We calculated loads for total phosphorus, total nitrogen, suspended solids, biochemical oxygen demand, lead, and zinc. We chose these pollutants because reliable data were available and because they characterize a broad range of nonpoint pollution-generating land uses, from urban to agricultural. The model calculates runoff and loads for any point in space, allowing the user to see the spatial distribution of loads. An example of a total phosphorus load grid for one subbasin in the Jacksonville, Florida vicinity is shown in Figure 2.

The grids themselves provide a detailed view of model output. Model results can also be summarized by watershed, using the watershed boundary grid, and the information can be examined from a basinwide perspective.

We have applied PLSM results in other useful ways at the District. For example, District staff felt that previous sediment sampling sites were not appropriately located, so the District water quality network manager used model results to locate new sampling sites, focusing on problem areas as well as areas where we expect to see little or no nonpoint impact.

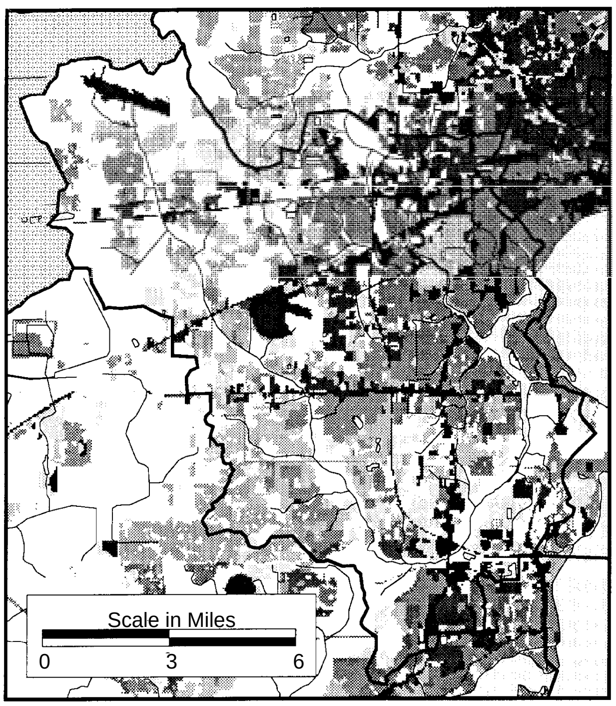


Figure 2. Distribution of total phosphorus loads, Ortega River subbasin (darker areas represent higher loads).

Application of Model Results in the Plan

Because the goal of the model was to identify potential stormwater runoff problem areas, we needed to simplify, or categorize, the model results for use in the Plan. We calculated the per acre watershed load for each pollutant and defined “potential stormwater runoff problem areas” as those individual watersheds with the highest loads for all pollutants. Problem areas for one major basin in the District, the lower St. Johns River basin, are depicted in Figure 3.

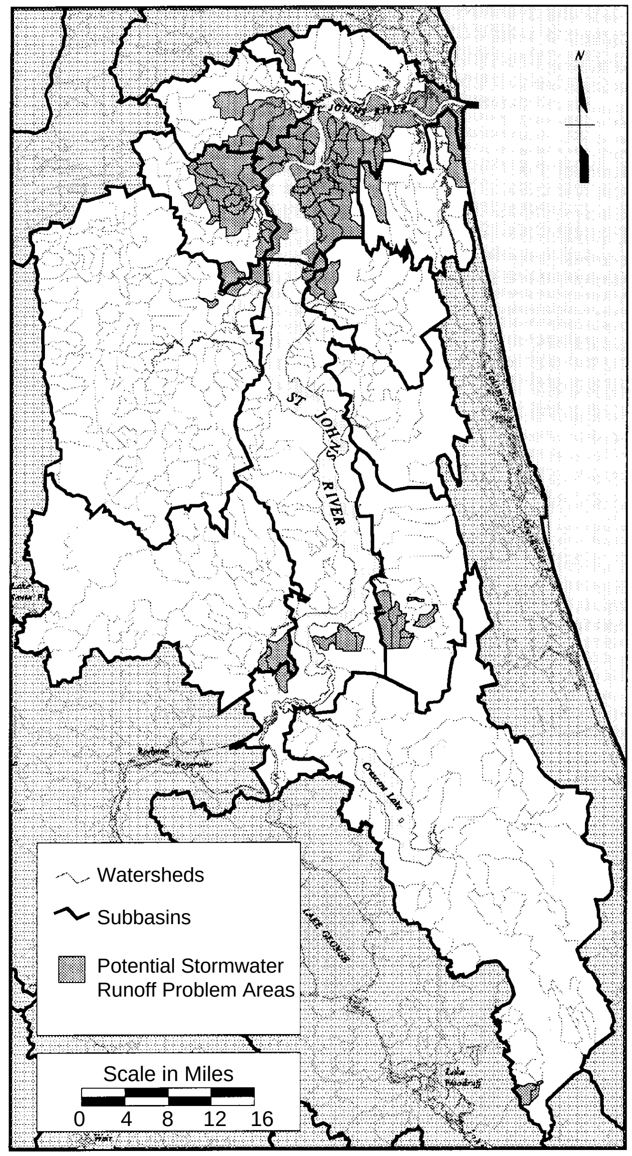


Figure 3. Potential stormwater runoff problem areas, lower St. Johns River basin.

We also ran the model with future land use data obtained from county comprehensive plans. Because the county maps are guides to future development, and not predictions of actual development, we exercised caution when using the results. Problem areas were defined as those watersheds with projected loads greater than or equal to existing problem areas. Also, District planners combined model results with information about individual counties’ regulations and policies to evaluate where problems are most likely to occur.

Prior to compiling the Plan, the District conducted workshops in each county in the District, in which problem areas identified by the PLSM were discussed with local agency staff, officials, and the public. We provided large, hard copy maps depicting stormwater runoff problem areas combined with results of a separate water quality analysis on county-based maps. These maps proved to be powerful tools for initiating discussions and gathering feedback.

In the Plan, stormwater runoff problem areas were reported for each of the 10 major drainage basins in the District. The information was also repackaged in a county-based format to create a quick reference for local agencies. District planners recommended strategies for addressing problems; these strategies vary as appropriate for each county. Examples include the need to assess compliance with existing stormwater permits, encourage stormwater reuse during the stormwater and consumptive use permitting processes, coordinate with municipalities that are implementing stormwater management plans, encourage and assist significantly affected municipalities to create stormwater utilities, and improve monitoring in problem areas that do not have sufficient water quality data.

In conclusion, the Watershed Assessment GIS project has proved to be useful not only to the St. Johns River Water Management District, but also to local governments. Large projects such as this could not be completed in a reasonable time without the use of GIS. Also, for ARC/INFO users who have been restricted to vector processing, the cell-based processing available in GRID is a powerful modeling tool.

References

1. Cox, J., R. Kautz, M. MacLaughlin, and T. Gilbert. 1994. Closing the gaps in Florida's wildlife habitat conservation system. Florida Game and Fresh Water Fish Commission, Office of Environmental Services, Tallahassee, FL.
2. Adamus, C.L., and M.J. Bergman. 1993. Development of a nonpoint source pollution load screening model. Technical Memorandum No. 1. Department of Surface Water Programs, St. Johns River Water Management District, Palatka, FL.
3. ESRI. 1992. Cell-based modeling with GRID. Redlands, CA: Environmental Systems Research Institute, Inc.

Update of GIS land use attributes from land surface texture information using SIR-C images

Francisco J. Artigas[1]

Abstract

The Meadowlands in northern New Jersey were used as dumping grounds for decades and today dense canopies of common reed (Phragmites australis) cover most of the District's open spaces. Our objective is to evaluate the utility of multi-frequency SAR in updating GIS land use information by means of prospecting for anomalously high backscatter in open spaces that could indicate the presence of building or metallic waste concentrations near the surface. We combined a land use vector coverage with a co-registered SIR-C C-HV image acquired October 10, 1994 to isolate officially designated "open space" parcels. Groups of four or more pixels with anomalously high backscatter values were converted to a vector coverage and draped over a very fine spatial resolution color infrared digital orthophoto of the District. We discuss the implications for the use of operational imaging radar for monitoring land use and management of open spaces within a dynamic and complex urban environment.

Introduction

The Hackensack Meadowlands Development Commission (HMDC) overseas the orderly development of the Meadowlands District (District) which is a 82 square kilometer degraded urban estuary four kilometers west of New York City in northern New Jersey. The unregulated use of District lands as disposal sites for solid and industrial waste for more than 150 years has turned these meadows and wetlands into one of the most environmentally assaulted areas in the U.S. (Grossman 1992). The HMDC (a New Jersey State Agency) oversees the preservation and development of more than 4,375 parcels of land including 1,012 hectares of landfill. This study presents the results of an effort to use Geographical Information Systems (GIS) in combination with Synthetic Aperture Radar (SAR) images to update the District's current land use database. This was accomplished by detecting and documenting areas in open spaces that showed unusual surface texture roughness as a consequence of land use change or disturbance by disposal of solid waste.

[1] Rutgers University, Center for Information Management, Integration and Connectivity. 180 University Ave., Newark NJ 07102. artigas@cimic.rutgers.edu

Land use research based on SAR images is emerging as a promising new operational technology for the applied earth sciences. Images of this area from the Shuttle Imaging Radar (SIR) with a ground resolution of 12.5 m were first made available in 1994. In 1995, Canada launched Radarsat-1 (ground resolution 8-100 meters) which some consider the starting point for the true application and commercialization of radar images (Glackin 1997). By the year 2002, NASA's LightSAR will be collecting radar images of the earth from a space platform on a continuous basis. In just a few years it is expected that there will be an abundance of high-resolution radar imagery which have some notable advantages over traditional spectral sensors.

The main advantage of radar over spectral remote sensing sensors (e.g. Landsat, Spot, SPIN-2) is that it can capture images of the earth's surface under any weather conditions, day and night. Radar can also penetrate vegetation canopies and at certain wavelengths penetrate the first centimeters of the soil (Xia et al 1997). The intensity of backscatter from radar microwave pulses is highly correlated to surface texture. A useful rule-of-thumb in analyzing radar images is that the brighter the backscatter on the image, the rougher the surface being imaged (Freeman 1996).

Since SIR-C images became available there has been great interest by scientists to document radar backscatter from diverse earth surfaces (Alpers and Holt 1995, Beaudoin et al 1994, Cordey et al 1996, Freeman and van de Broek 1995). The areas of interest have predominantly been large (10 to 5000 square kilometers), and focused on relatively homogeneous surfaces (e.g. deserts, boreal forest, crops, ocean etc.). There have been a less number of studies reported for urban areas (Taket et. al, 1991; Xia et al. 1997) and no reported studies that look at spots of less than 0.1 hectares (0.25 acres) within a complex urban matrix.

Our research used an image from an October 1994 Space Shuttle flight (SIR-C) that captured the New York City and North New Jersey Metropolitan areas. As far as radar scattering is concerned, such an urban scene is a target of considerable complexity. It provides a variety of surfaces with sharp edges that go from scales of several hundred feet (buildings) to just a few centimeters (building surfaces and fields with rubble).

Our specific objective was to evaluate the use of radar images in locating structures and debris fields in the District and lay out an effective methodology for updating GIS land use attributes based on landscape surface texture information. We discuss the implications of this technology for monitoring land use and management of open spaces within a dynamic and complex urban environment.

Methods

The HMDC maintains a GIS system to manage zoning, land use and block and lot information of District properties (parcels) for the purpose of land use management and planning. A geo-referenced parcel information coverage (Figure 1A.) was used to identify open space areas. The areas selected included: landfill, parkland, riparian, vacant and open water. A binary vector polygon coverage was created where open area parcel id's acquired the value of one while the rest of District parcels acquired a value of zero. The binary coverage was converted to a raster image (pixel size 12.5 meters on each side) where pixels representing open space areas (black) have a value of one and all other pixels have a value of zero (Figure 1B.).

A set of images for the District were documented and manipulated with raster software. The SIR-C sensor acquired images at three microwave wavelengths: L-band ~ 24 cm; C-band ~ 6 cm and X-band ~ 3 cm resolution (Freeman et al 1995). However, only L and C bands were available for the study. Pixel size is 12.5 meters on each side. Incidence angle was 64 degrees with illumination from the southwest. SAR transmits pulses of microwaves in either horizontal (H) or vertical (V) polarization and receives in either H or V. The polarization available for each band (L and C) were HH (horizontally transmitted and horizontally received) and HV. A third image for each band labeled total power (TP) is the combination of all polarizations in one image. Images used in this study were C-band HH, C-band HV, C-band TP, L-band HH, L-band HV and L-band TP (figure 3). The radar backscatter value for each pixel is measured in decibels (dB). In our case values ranged from -40dB, very smooth surfaces in light gray indicating open water, to +5dB very rough surfaces in dark gray indicating building structures (Figure 3).

Determining the accuracy of each radar pixel on the ground is important since we used the x, y coordinates of radar pixels to navigate to spots on the ground and verify anomalous returns. The original radar images were geo-referenced to the New Jersey State Plane Coordinate System by using 11 control points from a geo-referenced GIS vector coverage of District parcels (Figure 1C). The average RMS error of the re-sampling procedure was 1.3 m. However, the

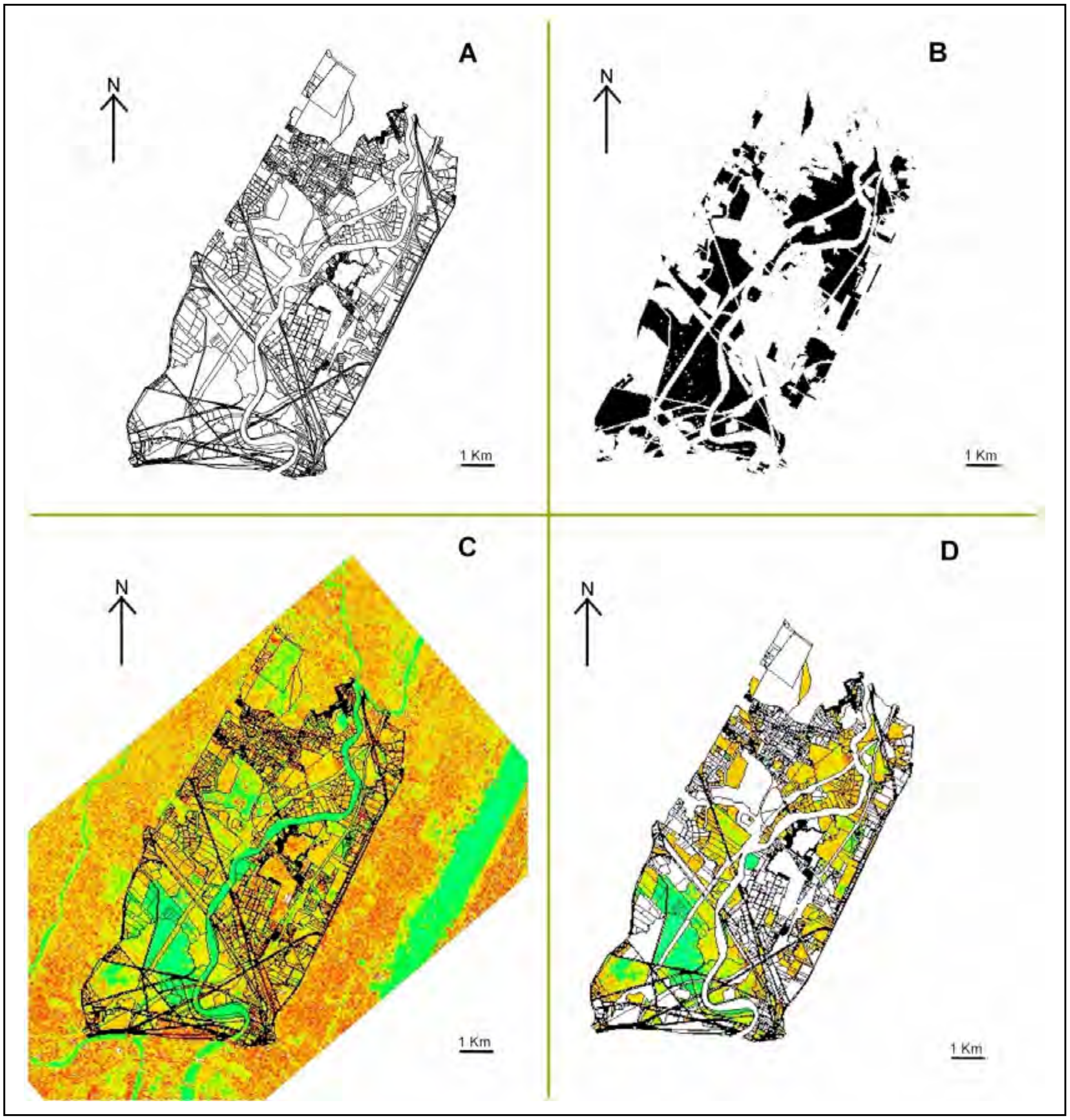


Figure 1. A.- GIS parcel vector coverage of the District. B.- Binary raster image of the District, black areas correspond to open spaces. C.- Geo-referenced radar image with an overlay of District parcels. D.- Image containing radar pixel values for open space areas.

parcel vector coverage also had digitizing and tolerance value errors. The original parcel map was digitized from a 1'': 200' scale map of the district. The resolution of the digitizing table was 0.0127 m (0.005 inches). The digitizing error on the ground was in the order of 30 cm. The tolerance level for the coverage (minimum distance before vector lines snap) was originally set to 4.2 meters. Therefore, the estimated error of parcel limits is 4.5 m. When the RMS error is included, then the total error on the ground of radar pixels for this study is at least 6 meters or half a pixel.

Once the images were geo-registered, the open area binary layer (Figure 1B) multiplied them each. As a result, all pixels not corresponding to open spaces were made zero. All other pixels maintained their dB values (Figure 1D). To avoid working with negative numbers, pixel values (dB) were re-scaled to fall within the range of 0-255. Pixel frequency distributions for all six radar images were graphed (Figure 2.). Backscatter distributions of images were inspected for speckle effects and contrast.

In order to select an image for field verification, two main criteria were used: 1- The image should clearly discriminate between water and different vegetation types, and 2- The image should have a minimum of speckle or background noise. The speckle or background noise is usually associated with strong reflecting surfaces from structures in the vicinity of open spaces.

Backscatter values from known power line towers in the middle of open spaces were used to select a threshold pixel value for an anomalous return. An anomalous return in an open space area with vegetation would be brighter than normal and most likely due to a corner reflector (i.e. two surfaces in 90-degree angle). Known power line tower locations in the middle of open spaces provided good examples of corner reflectors within a vegetation patch. High energy double-bounce scatter mechanisms prevail from rectangular metallic surfaces compared to the less energetic volumetric scatter mechanism that prevails from a canopy of vegetation. A similar reflectance mechanism (double bounce) would be expected from rubble (e.g. concrete slabs or metal artifacts) such as old cars and drums exposed or hidden under vegetation.

All pixels exceeding the threshold value were extracted and clusters of more than four pixels plotted. Clusters of at least four were selected to avoid selecting single pixels with unusual brightness values located exactly on parcel boundaries that in reality do not correspond to true open spaces. These clusters were converted to vector polygons and draped over a 1m resolution geo-referenced digital orthophoto (USGS 1995). The draped orthophoto images were visually inspected for obvious structures on the ground that may have produced the bright backscatter return (Figure 6 and 7). Parcels classified as open space that contained structures were re-classified in the GIS database. Spots were selected for ground verification when visual inspection of the draped orthophoto images failed to reveal structures on the ground that could explain the anomalous return. Moreover, ease of ground access and clusters size were important considerations in selecting specific spots for field verification. Radar image coordinates were used to navigate to the sites. Known reference points in the field were selected and azimuth and distance from these points to the center of the anomalous spots were used to navigate in the field. Sites were finally located in the field by using a 30-meter measuring tape and a compass. Historical aerial photography from 1966 was used to confirm and explain some of the disturbances detected by radar.

Results

Although several theoretical models have been developed to describe how ground objects reflect radar energy (Taket 1991, Evans et al 1988), most knowledge comes from practical observations. There is no strong set of rules indicating what bands or polarizations work best on different surfaces and under specific sensor conditions. Given the time and resources available, it was important to find one best image out of the six available to field verify.

An inspection of the pixel brightness frequency distribution values for all images (Figure 2) shows that the distributions tend to be bimodal except for HH polarization. Cross-polarization (HV) images function better to capture the difference between water and vegetation surface by clearly separating water and vegetation into two distinct peaks. Of all images, C-band HV has the best contrast (a greater valley between peaks) as well as a sharper break between the end of the vegetation values (tall peak) and the tail of the extreme bright values.

The amount of speckle was another criteria for image selection. Speckle effects, or "noise", are common in radar images capturing complex surfaces. Speckle can be caused by an object with complex dielectric properties that behaves as a very strong reflector at a particular alignment between itself and the spacecraft. Structures such as metal bridges, elevated highways, and power line towers behave as strong reflectors in urban areas. These structures influence the value of neighboring pixels making them appear much brighter than the surfaces they actually represent. In our case speckle effect is clear in C-band HH, C-band TP and L-Band HH (Figure 2B, 2C and 2E.). In all these cases there is an unusual high frequency of bright returns at the high end of the distribution. Unfortunately, speckle effects affect open areas near strong reflector objects. In our case this is very clear for C-band HH (Figure 2B). Speckle effect is also clear in L-band TP. However, in this situation background noise manifests itself by cross and

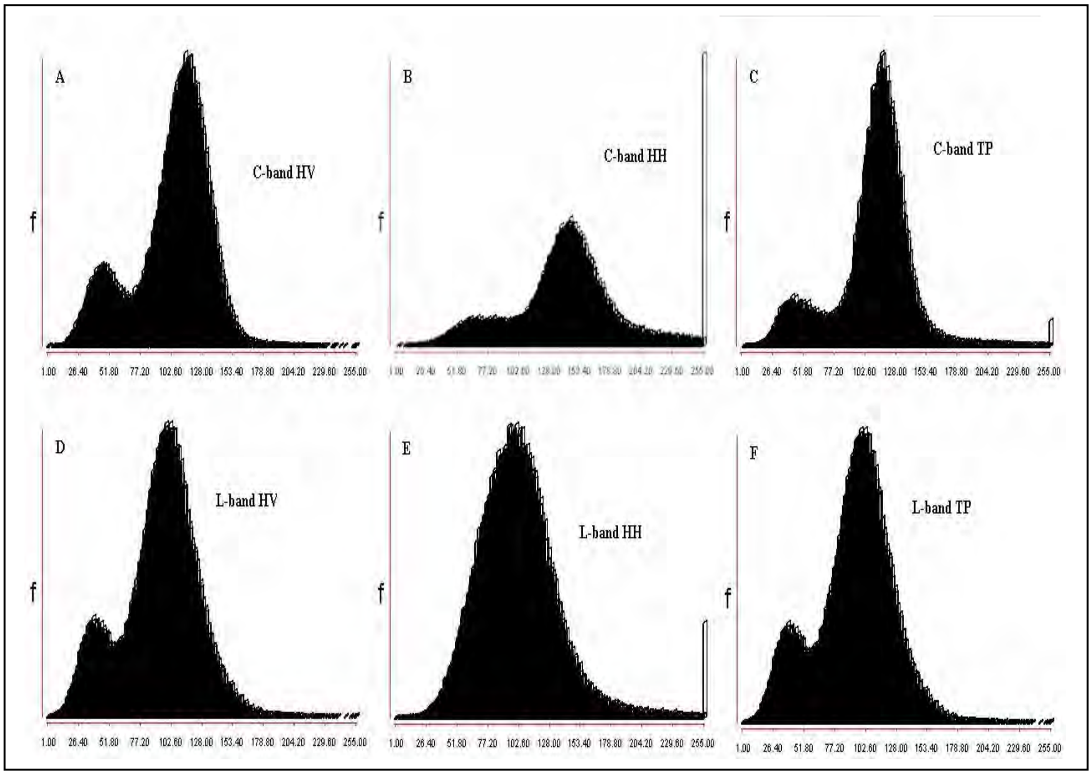


Figure 2. - Pixel brightness frequency distributions for C and L bands HH and HV polarizations. HV and TP polarizations for both bands, clearly separate open water and vegetation surface texture (bi-modal distribution). C-band HV shows the greatest contrast between surface textures (valley between peaks) and a sharp break between the vegetation peak and the upper tail of the brightness distribution.

star like patterns that have their origin at points with rough features and high dielectric constants (Figure 3).

The final criteria used in selecting an image were based on radar wavelength. Shorter wavelength (C) have better resolution than longer wavelength (L), thus improving the overall spatial resolution and the ability to delimit detail and boundaries (Xia and Henderson 1997). Based on these observations and since our objective is to determine surface texture differences within small areas, of the two best images available (C-band HV and L-Band HV) we choose the smaller wavelength image C-band HV.

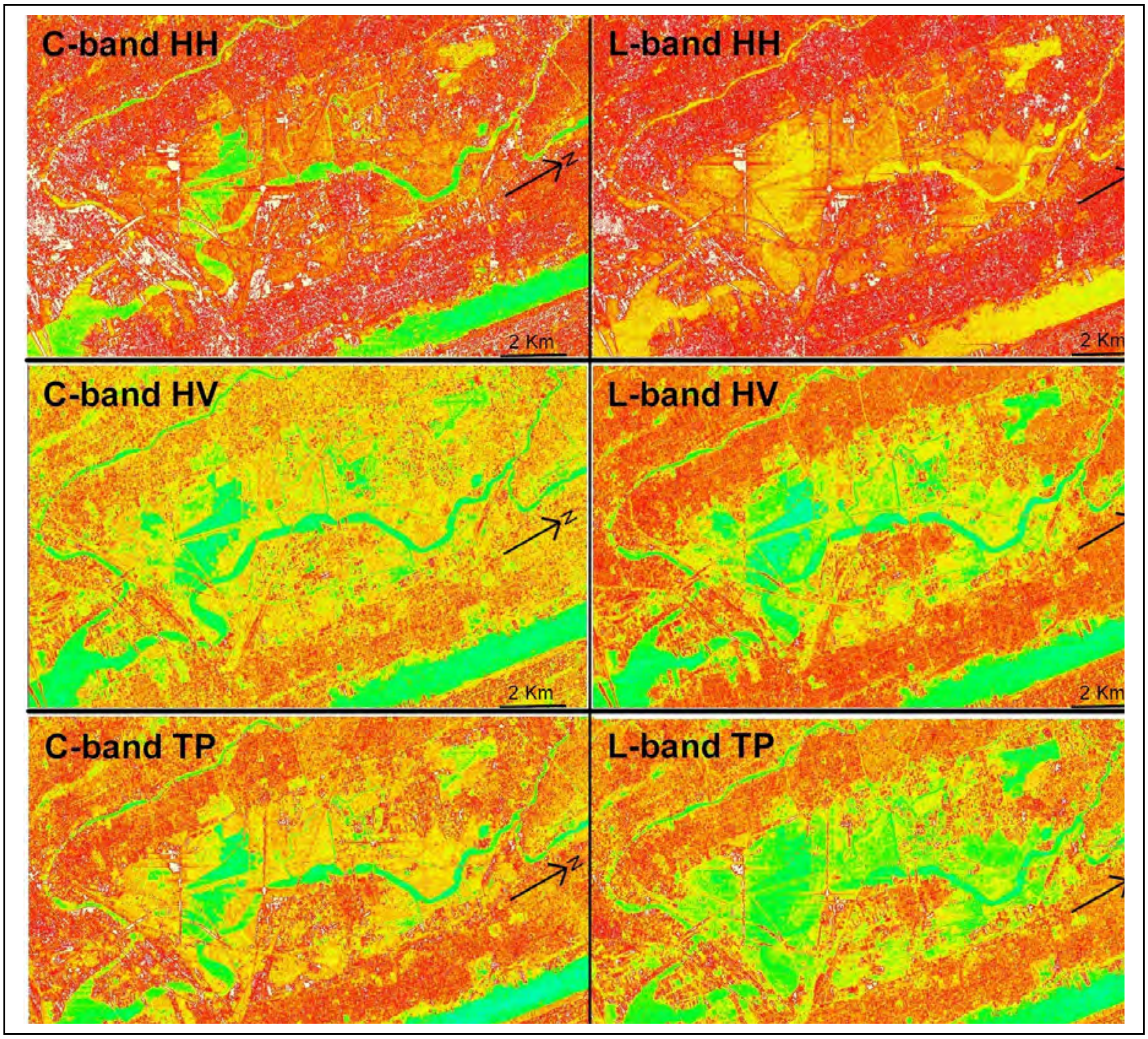


Figure 3. - Images of the Hackensack Medowlands (C and L bands, HH and HV polarizations. Flat smooth surfaces in gray represent mainly open water bodies. The lower right corner of the images shows the Hudson river and a part of Manhattan. West of the Hudson is the Hackensack and Passaic rivers draining into Newark Bay. Other white-gray surfaces represent wetlands. Dark gray and black areas represent developed urban areas.

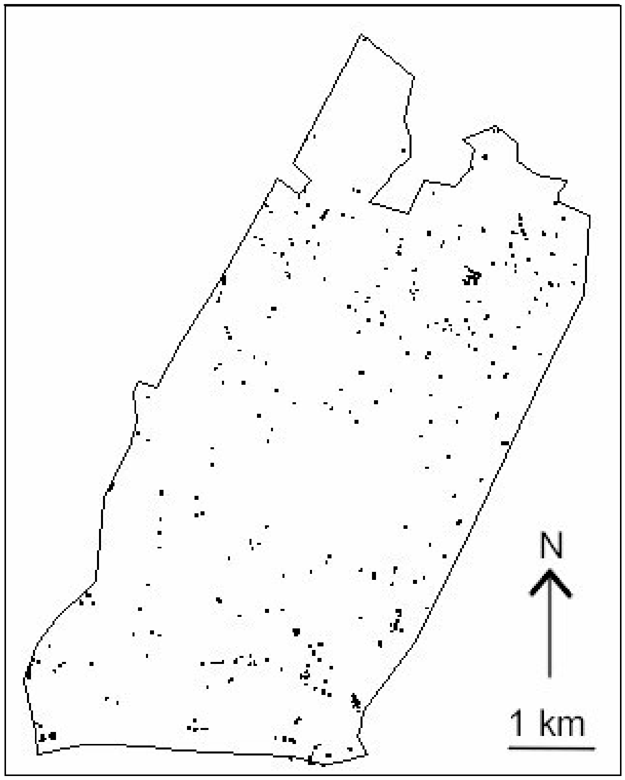


Figure 4. -Pixel clusters from open spaces that were classified as anomalous returns (brightness value greater than 176 out of 255).

"vegetation pixels" from brighter backscatter returns. For this study, a value of 176 was selected as the threshold value for an anomalous return. Figure 4 shows all pixels from open spaces that exceeded the threshold value and were classified as anomalous returns for the District.

There were a total of 128 clusters of anomalous returns. These clusters varied from four to 53 pixels per cluster. The greatest number of clusters (61) were made out of only four pixels (Figure 5). The most common clusters (82%) were made out of 10 pixels or less. The most extensive cluster was made out of 53 pixels representing an area of 0.8 hectares.

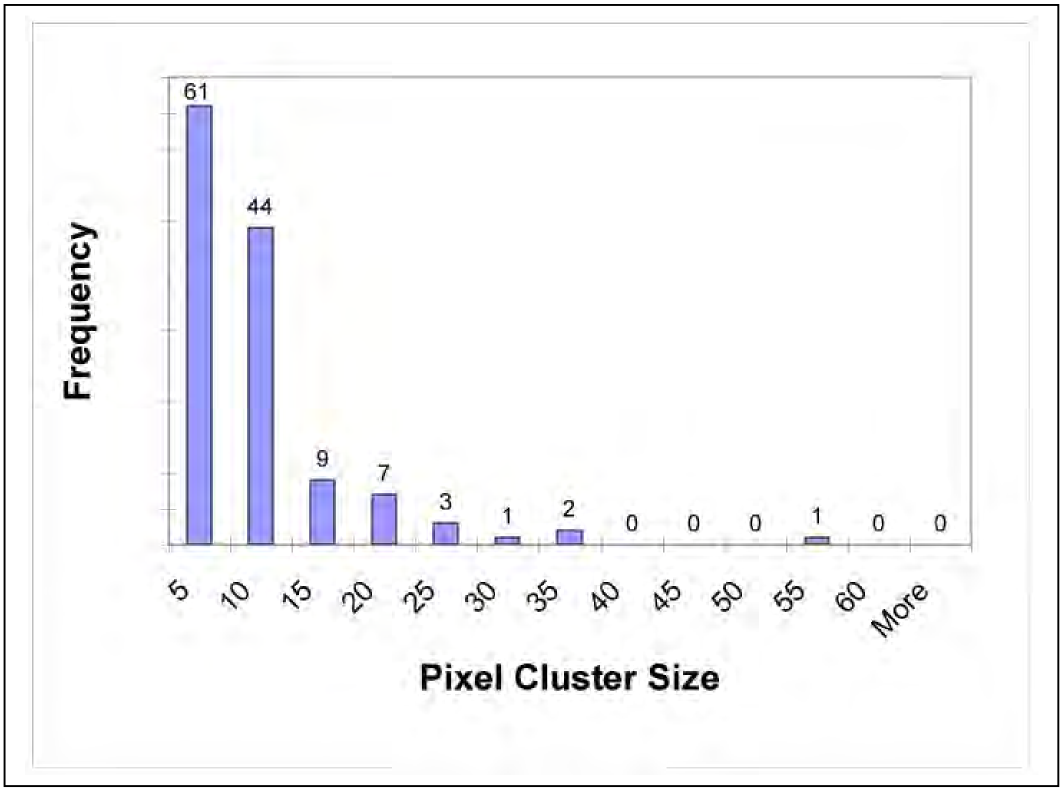


Figure 5. - Frequency distributions for clusters of pixels that exceeded the brightness threshold value of 176. More than 80% of the clusters were 10 pixels or less.

Figure 6 shows how a several clusters of radar anomalous returns overlay at least three distinct parcels of land at the southwest corner of the District. These parcels were classified as vacant in the GIS. Parcels range from 0.5 to 1.5 hectares. By zooming in the ortho-photo one can actually see that these parcels contain structures (trailers and construction equipment) and are not vacant.

Similarly, Figure 7, shows another anomalous return from a border of a vacant lot (arrow) that turned out to be a 0.1 hectares (0.25 acre), seven meter high bulldozed mound of earth mixed with metal rubble, tires, old batteries, cloth, etc. It is impossible to tell by zooming in the ortho-photo that this structure actually exists since it blends with the colors of the surrounding area. A list of similar sites and their coordinates were detected and documented.

Figure 6. - Anomalous radar clusters are shown as white polygons draped over a 1-meter resolution orthophoto. White lines represent parcel boundaries. Red circles mark parcels classified as vacant by the GIS yet they contain structures (metallic rigs and machinery) which offer corner reflector surfaces detected by radar.

The most interesting anomalous returns originated from what appeared to be a field with a dense homogeneous canopy of <u>Phragmatis australis</u>. (Figure 7 circled area). After navigating to the anomalous spots in the field, the only consistent difference in surface texture was a much lower and sparse canopy of the common reed. These "open" spots seemed to be correlated to differences in soil compaction. The effect is a sharp boundary between the tall reed canopy and the sparsely vegetated areas inside the selected spots. In some of these spots the soil was littered with waste such as leather scraps, plastic wallpaper, roofing material, tiles and even junked cars. Soil compaction, sharp canopy boundaries and scattered rubble provided the surface features that favored double bounce reflections as opposed to volumetric scatter from the immediate surrounding areas.

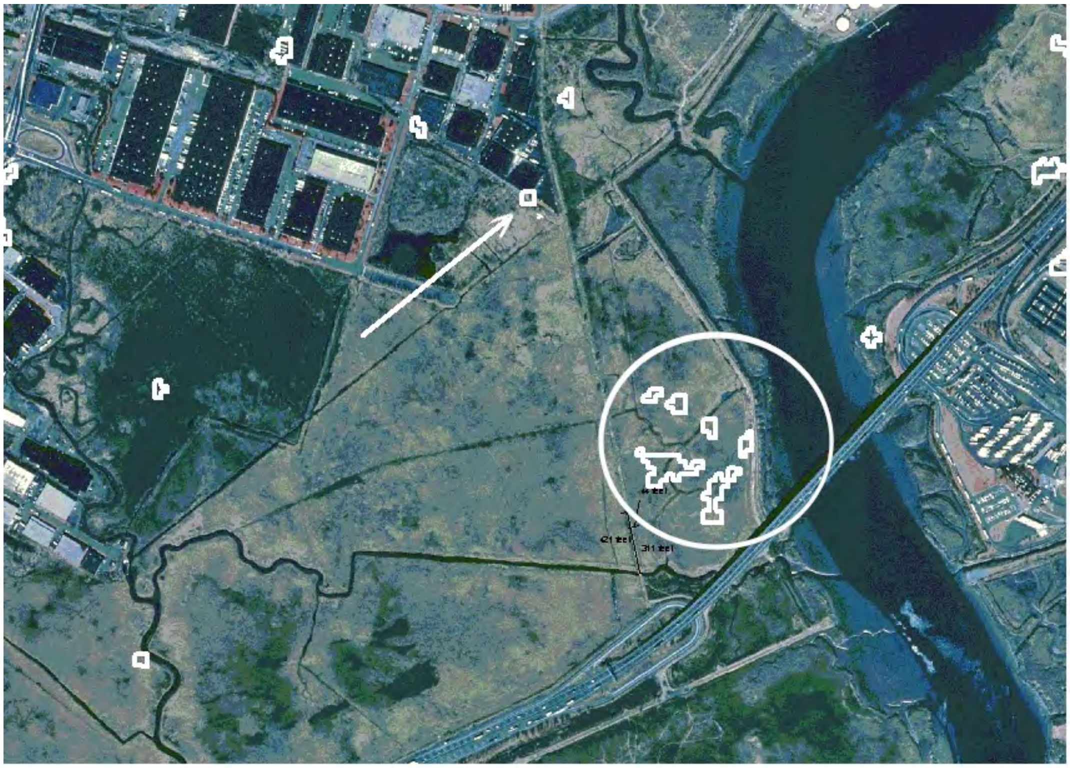


Figure 7. - Radar anomalous returns (white polygons) for the South Hackensack part of the District. The white arrow shows a 40X40-meter area where there is a bulldozed mound of rubble. This mound, hidden under vegetation would be difficult to separate from the colors of the surrounding vegetation by using traditional spectral sensors. The anomalous returns included in the white circle indicate an area of what seems to be a homogenous canopy of the common reed. However, field inspection of these spots revealed differences in surface texture due to past disturbances

Recently declassified intelligence satellite photography from 1966 (3-m resolution) of the same area (Figure 8) revealed a construction staging area exactly where the anomalous returns from radar originate. What is today a "dense" stand of <u>Phragmites australis</u>, was in the late 1960's a staging area for the construction of the New Jersey Turnpike. Roads and trails can be identified in this area from the 1966 satellite photograph. This explains the differences in soil compaction and the presence of debris in what seems today to be the middle of a common reed field.

Figure 7. - Declassified spy satellite photograph from 1970 showing the same area of Figure 6. The brighter area inside the white circle shows a surface disturbance created by heavy machinery and rubble disposal that created differences in surface texture hidden under vegetation and detected by radar 24 years later

Discussion and Conclusions

The use of SAR images to detect disturbed or incorrectly classified land uses based on surface texture proved to be a reliable and effective approach. Detection and navigation to the previously unrecognized disturbed areas was done with less than 10 meters of error on the ground. This in itself demonstrates the practicality and effectiveness of the method. One critical part of the study was selecting the appropriate band and polarization to identify sites on the ground since resource constraints prohibited the evaluation from both bands and all polarizations in the field. Our criteria for image selection was in agreement with most research to date which seems to prefer cross-polarized (HV) images for urban and land use cover mapping over co-polarization (Bryan 1974, Henderson 1985). In our case the C-band HV distribution better separates between "vegetation pixels" and a class of brighter pixel values than L-band HV (Figure 2A and 2D). Contrary to findings reported in the literature (Xia 1997), surface features appear smoother at shorter wavelengths (C-band HV) than at larger wavelengths (L-band HV). Selecting C-band over L- band may have resulted in a loss of information associated with surface texture within vegetation patches. However, overall, we found that shorter wavelength best separated known targets from the surrounding vegetation. Selecting a threshold value, in other words, determining which pixels were brighter than normal for a given surface was a critical decision of the study. There are many models for scatter mechanisms (Evans 1988, Jacob 1993,) and they all tend to agree that the double bounce mechanism carries the greatest amount of energy back to the receiver creating brighter pixels. In this case the microwave bounces off one surface, hits another surface and returns back to the receiver (corner reflector effect). We considered this mechanism to prevail in the case of a target such as an electrical power tower, or slabs of concrete and metal rubble hidden under vegetation or ground that has been compacted in open spaces surrounded by hard stems. Our data suggests that volumetric scatter would be the dominant scatter mechanism from a canopy of Phragmatis australis. Incident radiation would bounce off stems, branches and leaves scattering in all directions and less radiation would return to the receiver creating less bright returns from these surfaces.

Incidence angle and "look" angle for each image vary according to the position of the sensor in space in relation to its ground target. In this case, our image had a fixed 64 degrees incidence angle illuminating from the southwest. Different fly-over passes of the sensor will create images with different incidence angles and therefore different backscatter patterns for the same area. Images will clearly have to be selected according to the most favorable incidence angles and illumination direction. Topography will also influence the backscatter pattern, as shadows from mountains will hide surfaces that can not be imaged by radar. These factors emerged as important limitations to surface texture detection using radar. In our specific case, the relatively flat topography of the District would favor the use of radar.

Radar was able to identify specific parcels that were incorrectly classified in the GIS database. Some parcels contained reflector surfaces from structures such as trailers, metallic rigs and machinery which were easily recognizable from the ortho-photos.

In other cases (Figure 7 arrow), radar was able to detect rubble hidden under vegetation. The 40X40-meter mound could not have been separated from the surrounding vegetation by using any other type of spectral sensor i.e. aerial photo or satellite image. In this case, the orientation height (7 m) and the flat surfaces of metallic rubble associated with the mound made it possible for radar to detect and separate this particular structure from the surrounding vegetation.

Finally, radar was able to detect past disturbances within certain areas of normal looking reed fields. In these sites, soil compaction by heavy machinery and dumping of construction rubble and waste from almost 40 years ago altered the natural hydrology and influenced community plant development. Again, with radar it was possible to separate surface feature roughness and detect these disturbances.

One mission of the HMDC is to balance development with preservation. This process involves making informed decisions regarding what areas are to be developed and what areas are to be set aside for preservation. To make these decisions, agency executives and technicians need to be well aware of the location, quality and characteristics of each site. Currently, the land use management team updates land use of District parcels by field recognizance every four to five years. Under the current method, there may be enough time in between updates for an undetected illegal land use practice to create considerable damage.

Our methodology proved that based on surface texture detection it is possible to monitor and update parcel information from areas as small as 0.1 hectare. Once periodic radar images become available, standard backscatter returns for District parcels would be documented and become a template for the current surface texture pattern. New images (with same incidence angle and illumination direction) would be checked against the template. Changes in surface texture due to constructions of structures, disturbance of soil surface, change in vegetation cover or waste dumping would create a different backscatter pattern from the template that would translate into the identification of parcels where changes took place. The challenge ahead is to integrate these systems and methodologies into an expert system with the ability to detect change in surface texture, update the GIS database automatically, and continuously repeat this process as new images become available. This approach should be able to keep up with the dynamic and complex land use changes in the District and help managers make informed decisions based on timely land use information.

Acknowledgment

This work was preformed with funding from a grant from the Hackensack Meadowlands Development Commission and sponsored by the Rutgers University NASA Regional Application Center and the Center for Information Management Integration and Connectivity (CIMIC). I would like to thank Dr. Geoffrey Henebry for providing the images and for his comments on radar backscatter mechanism. Finally, I would like to thank Matthew Ceberio and Alon Frumer for their help with field verification.

References

Alpers, W., and B. Holt, 1995. Imaging of ocean features by SIR-C/X-SAR: An overview, Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'95), July 10-14, Florence Italy, pp. 1588-1590.

Beaudoin, A., T. Le Toan, S. Gozc. E. Nezry, A. Lopes, E. Mougin, C. C. Hsu, J. A. Kong and R. T. Shin. 1994. Retrieval of forest biomass from SAR data. Int. J. Rem. Sens. vol. 15, pp 2777-2794.

Bryan M.L. 1974. Extraction of urban land cover data from multiplexed synthetic aperture radar imagery. in Proc. Ninth Int. Symp. Rem. Sens. of Environment. Environmental Research Institute of Michigan, Ann Arbor. MI, pp. 271-288.

Cordey, R., J. R. Baker, S. Quegan, G. M. Foody, N.J. Veck, and A. Wielogorska 1996. A study of the potential of multi-component SAR imagery for agricultural and forest studies, in Science results from the Spaceborne Imaging Radar C/X-Band Synthetic Aperture Radar (SIR-C/X-SAR): Progress Report, Jet Propulsion Laboratory Publication No. 96-7, pp. 112-14.

Evans L. Diane, Tom G. Farr, Jakob J. Van Zyl and Howard A. Zebeker 1988. Radar Polarimetry: Analysis tools and applications. IEEE Transactions on Geoscience and Remote Sensing, Vol.26. No. 6.

Freeman, A., M. Alves, B. Chapman, J. Cruz, Y. Kim, S. Schaffer, J. Sun, E. Turner, and K. Sarabandi 1995. SIR-C data quality and calibration results, IEEE Transactions on Geoscience and Remote Sensing, Vol. 33, no. 4, pp. 848-857.

Freeman, A., and B. van den Broek, 1995. Mapping vegetation types using SIR-C data, Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS'95), July 10-14, Florence Italy, pp. 921-923.

Freeman A., 1996. What is imaging radar?. Spaceborne imaging Radar-C SIR-CED, Jet Propulsion Laboratory. Pasadena California.

Glackin L. David. 1997. Earth observation in transition: An international overview, Acta Astronomica Vol. 41, Nos. 4-10. pp. 413-420.

Grossman and Associates, Inc. 1992. Stage 1-A Archeological and Historical Sensitivity evaluation of the Hackensack Meadowlands. Grossman and Associates Inc. Technical Report, Lyndhurst. NJ.

Henderson F. M. 1984. Confusion errors among urban land-cover types on SAR imagery. International Journal Remote Sensing, vol.6, no 10. pp. 1607-1622.

Jakob J. van Zyle, 1993. The effect of topography on radar scattering from vegetated areas. IEEE Trans. on Geoscience and Remote Sensing, Vol. 31, No 1.

Richards J.A. 1986. Remote Sensing Image Analysis. Berling, Springer.

Taket N.D., S.M. Howarth, and R.E. Burge 1991. A model for the imaging of urban areas by synthetic aperture radar. IEEE Transactions on Geoscience and Remote Sensing, Vol. 29 No. 3.

USGS 1995. Digital Orthophoto Quadrangles . US Geological Survey. Reston, VA.

Xia Zong-Guo and F.M Henderson 1997. Understanding the relationship between radar response patterns and the Bio and Geophysical parameters of urban areas. IEEE Trans. on Geoscience and Remote Sensing, Vol. 35, No. 1. pp 93-102.

GIS in the Confirmation Process

Dr. Raymond E. Bailey and Mr. Madhukar Mohan

The ultimate goal of remediation of the Department of Energy (DOE) Weldon Spring mixed waste site is to release the site for unrestricted use to the extent possible. This dictates that an accurate assessment of post cleanup activities is performed to confirm that contaminated soils have been successfully treated or removed. The assessment begins by developing an accurate 3-dimensional picture of the spatial distribution of contamination prior to the start of cleanup activities. A geographic information system (GIS) was selected over traditional manual methods to map the initial spatial distribution of contamination (pre-cleanup levels) and to manage the confirmation database. The Weldon Spring site, consisting of approximately 217 acres, was divided into eleven remedial units (RU). Each RU was divided into approximately 2,000 m2 areas known as confirmation units (CU). Upon completion of remedial activities within each RU, Environmental Safety and Health technicians conducted a walkover survey with a 2-inch by 2-inch sodium iodide (NaI) scintillation detector to establish removal of surface contamination to levels at or below background levels of gamma-emitting radioactivity. Following the surface scan, soil samples were collected and analyzed for radiological and chemical contaminants of concern listed in the Record of Decision. Results of laboratory radiological and chemical analyses were used to populate the attribute portion of the GIS database. Geographic locations in the database were obtained from surveys of the sample locations. This established an accurate location for each sample, provided confirmation that the as-built excavation had achieved the designed depths, identified areas that had achieved cleanup levels, and, if above target cleanup levels, identified locations of contamination requiring additional remediation.

The Department of Energy (DOE) Weldon Spring Site Remedial Action Project (WSSRAP) is a mixed waste site located in St. Charles County, Missouri, approximately 48 km (30 mi.) west of St. Louis (Figure 1).

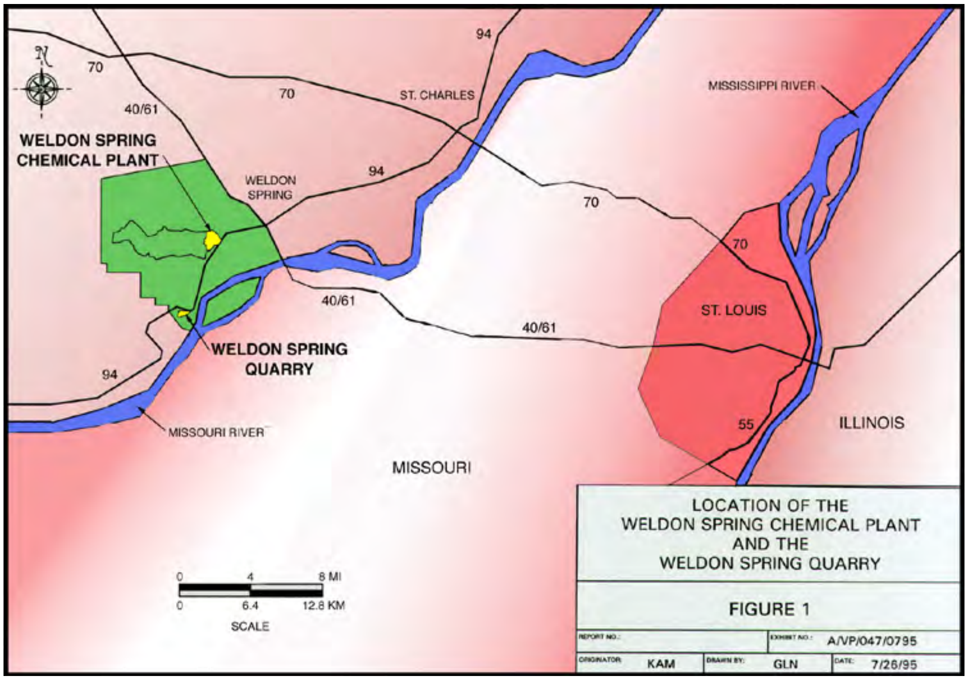


Figure 1

The WSSRAP consists of a 217-acre chemical plant area initially used by the U.S. Department of the Army during the 1940s to produce the explosives trinitrotoluene (TNT) and dinitrotoluene. After World War II, the structures were razed, decontaminated, and the site was re-graded. The U. S. Atomic Energy Commission (predecessor to the DOE) built a chemical plant upon the former Army site to process uranium and thorium ore concentrates. Production operations proceeded throughout the 1950s and 1960s, resulting in the disposal of radioactively and chemically contaminated waste on site.

Contaminated areas at the Weldon Spring Site included material from 40 building foundations, four raffinate pits, two ponds, and two former dump areas (Figure 2).

The contaminants of concern requiring treatment are radioactive contaminants (primarily radionuclides of the natural uranium and thorium-232 decay series) and chemical contaminants (including naturally occurring metals and inorganic anions, as well as organic compounds such as polychlorinated biphenyls and nitro-aromatic compounds). The remediation alternative selected consists of removing material from contaminated areas, treatment as appropriate by chemical stabilization and/or solidification, and disposal in an engineered disposal facility

Figure 2

constructed on site. The ultimate goal of site remediation activities is the release of the site for unrestricted use to the extent possible. To achieve this goal, requires an accurate assessment of post-cleanup activities (the confirmation process).

The site geographic information system, utilizing ArcInfo software operating on a Sun Unix workstation, was selected to create an accurate 3-dimensional picture of the spatial distribution of contamination prior to the initiation of remediation activities and to manage the confirmation database. The geographic database was populated with ArcInfo coverages of the Weldon Spring Site topography in 1954, aerial surveys of WSSRAP in 1993 and 1998, coordinates of anomalies identified during geophysical surveys, and the location of characterization drilling and sampling points. A complete picture of the spatial distribution of the pre-cleanup contamination was provided by linking the analytical laboratory results of the characterization samples with their geographic locations.

The confirmation database was created by dividing the 217-acres site into eleven remedial units (RU). Each RU was further divided into approximately 2,000 m^2 (0.5 acres) areas known as confirmation units (CU) (Figure 3).

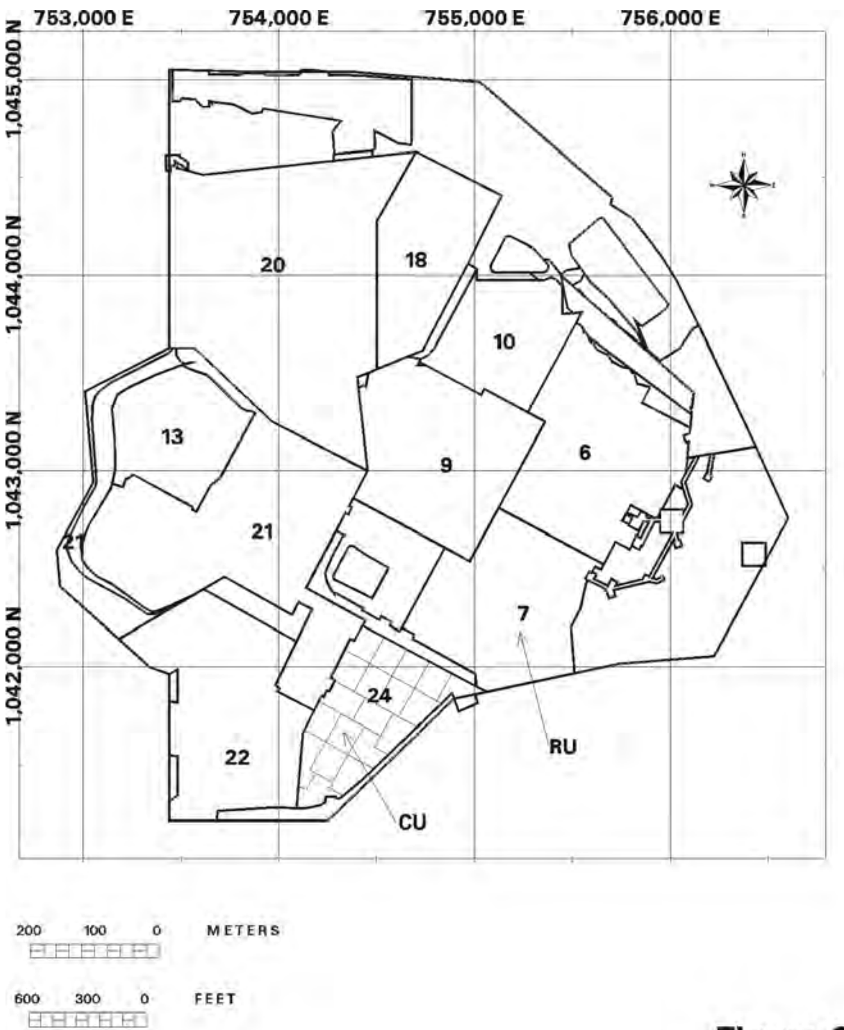


Figure 3

The CU is the area for which a decision is made as to whether cleanup standards have been attained. The size of the CU was selected to provide an area of approximately the same size as that used in the risk assessment for a future residential lot. This size also provided manageable areas capable of supporting the construction schedule when an excavated area needed to remain open pending the confirmation that cleanup standards had been attained.

Upon completion of remedial activities within each RU, Environmental Safety and Health technicians conducted a walkover survey with a 2-inch by 2-inch sodium iodide (NaI) scintillation detector to establish removal of surface contamination to levels at or below background levels of gamma-emitting radioactivity. Areas showing elevated readings greater than 1.5 times background were designated as “hot spots”, and additional material was removed. After obtaining a surface scan at or below background readings, a 10 meter by 10 meter grid was surveyed and soil samples were collected and analyzed for radiological and chemical contaminants of concern listed in the Record of Decision. Results of laboratory radiological and chemical analyses were used to populate the attribute portion of the GIS confirmation database. Geographic locations in the database were obtained from surveys of the sample locations and compared with design excavation limits. This established an accurate location for each sample, provided confirmation that the as-built excavation had achieved the designed depths, identified areas that had achieved cleanup levels, and, if above target cleanup levels, identified locations of contamination requiring additional remediation.

Vulnerability Assessment of Missouri Drinking Water to Chemical Contamination

Christopher J. Barnett, Steven J. Vance, and Christopher L. Fulcher
Center for Agricultural, Resource, and Environmental Systems, University of Missouri, Columbus, Missouri

Introduction

In 1991, the Missouri Department of Natural Resources (MDNR) implemented the Vulnerability Assessment of Missouri Drinking Water to Chemical Contamination project. MDNR's Public Drinking Water Program (PDWP) contracted with the Center for Agricultural, Resource, and Environmental Systems (CARES) to conduct this assessment. They designed the project to determine which, if any, public water supplies are threatened by chemicals being tested under the Safe Drinking Water Act.

Under Phase II of the Safe Drinking Water Act, the United States Environmental Protection Agency (EPA) required that all public drinking water systems be routinely monitored for 79 contaminants beginning January 1, 1993. If a selected chemical parameter is not detected in an area that would affect a water supply (where "detected" is defined as used, stored, manufactured, disposed of, or transported regardless of amount), then the water supply need not be tested for that chemical. Instead, that system would be granted a use waiver, meaning that the state would not test for that chemical. EPA grants use waivers for 43 of the 79 contaminants. Use waivers can result in considerable cost savings.

Because use waivers are granted based on the spatial relationship between drinking water sources and contaminant sources, accurate positional data needed to be collected for those items. A geographic information system (GIS) was used to store and analyze this information in a spatial context.

Water Sources

Water sources, as defined for this study, are the points where water is drawn from a river, lake, or aquifer for use in a public water supply. Our efforts focused primarily on the development of the water source layers for the GIS. These layers, containing wellheads, impoundment intakes, and river intakes, were created in house or obtained from state and federal agencies. MDNR regional office personnel inspected these water source layers in the spring of 1993. Since these personnel routinely inspect Missouri public drinking water supplies, their knowledge of these locations is exceptional. The updated water source information was mapped on 1:24,000-scale USGS topographic quadrangles at the regional offices, then entered into the GIS. MDNR's PDWP provided available attribute information, which was associated with these layers. The layers offer the most accurate and current information available. Only the community (e.g., cities, subdivisions, mobile home parks) and the nontransient, noncommunity (e.g., schools, large businesses) water supply systems were considered for water source mapping. This study did not consider private wells.

The information is stored in the GIS in the form of geographic data sets or layers. The wellhead layer contains 2,327 public wells and their attributes (e.g., well depth, casing type). The majority of the wellheads are located in the Ozarks and Southeast Lowlands. Naturally poor ground-water quality prohibits a heavy reliance on ground water for drinking water in other areas of the state. The surface water impoundment layer contains 105 points representing the intake locations for systems that rely on lake water. Additionally, the drainage basin and lake area are mapped for these systems. The majority of the systems that rely on lake water are located in northern and western Missouri. The final layer represents the systems that use river water. The majority of the 50 intakes are located on the Mississippi and Missouri Rivers and on the major streams in the Grand and Osage River basins.

Contaminant Sources

Contaminant sources, as defined for this study, are the points or areas where existing databases indicate the presence of a chemical contaminant. Incorporation of contaminant data into the GIS proved to be the most difficult task. These data usually contained very precise information about what contaminants were found at a site and who was responsible, but the quality of the locational information was often poor.

Ninety-three state and federal databases were reviewed for contaminant information before performing the final use waiver analysis. The contaminant information was broken into two separate types, contaminant sites and pesticide dealerships. The contaminant sites were locations at which certain chemicals were known to exist. The pesticide dealerships were dealerships licensed to distribute restricted use pesticides. Information about contaminant sites was extracted from the databases and entered into Microsoft Excel, a spreadsheet program. The small amount of data with coordinate (latitude/longitude) or map information was readily converted to the GIS. The majority of the contaminant records, however, contained only address information, often appearing as a rural route address or post office box number.

While the water source locations were being verified, personnel at the MDNR regional offices reviewed the contaminant site records. The regional office personnel were familiar with their respective territories and could assist CARES personnel in locating the contaminant sites. The Missouri Department of Agriculture pesticide use investigators provided additional information about the locations of contaminant sites. All contaminant source information was also mapped on the 1:24,000-scale USGS topographic quadrangles and transferred to the GIS.

Of more than 2,800 contaminant sites found in these databases, 88 percent were geographically located and used in the study. At this time, the contaminant site layer contains 2,493 points representing the information collected on the 43 chemical contaminants required by MDNR. Each point contains a seven-digit chemical code indicating the chemical it represents and serving as a link to the chemical contaminant files. The contaminant sites tend to be concentrated more in urban areas than rural areas. Even though this layer is being continually updated, the basic distribution of contaminant sites remains the same.

A second contaminant source layer represents Missouri's licensed pesticide dealers. This information is included to indicate potential contamination even though specific chemicals at dealership locations are not known. At this time, we have been able to locate 1,344 dealerships out of 1,650. Two types of dealerships are included in the layer, active dealers and inactive dealers. Of the active dealerships in 1991, 91 percent were found and entered into the GIS. Of the inactive dealerships, 79 percent were located.

Spatial Analysis

The final parameters for the use waiver analysis were developed from EPA and MDNR guidelines and account for the capabilities of the GIS. These parameters were designed to present a conservative list of the systems that needed to be tested for the possible presence of studied chemicals. Parameters for the wellhead analysis are as follows:

- A ¼-, ½-, and 1-mile radius around each wellhead was searched for contaminant sites and pesticide dealerships (see Figure 1). Any contaminant sources found within those radii were reported to PDWP. (PDWP requested that the results of the three radius analyses be reported, but the ½-mile radius was used to determine the issue of the use waiver.)
- Any wellheads found within a contaminant area were denied a use waiver for that contaminant.
- Each highway and railroad within 500 feet of a wellhead was recorded. This indicates the threat posed by the transport of chemicals near wellheads.
- Additionally, the percentage of the county planted in corn, soybeans, wheat, sorghum, tobacco, cotton, and rice was listed for each well to indicate the threat posed by agricultural chemical use within that county.

The parameters for the systems relying on lake water are as follows:

- Any contaminant sources found within a surface water impoundment drainage basin caused the associated intake(s) to fail use waiver analysis for those contaminants.

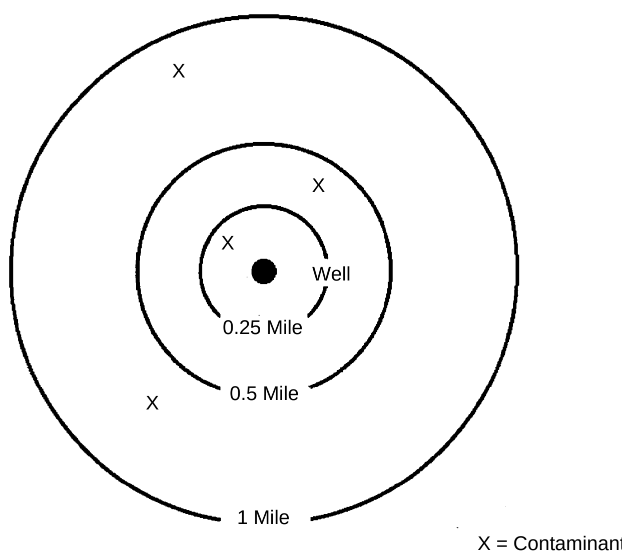


Figure 1. Use waiver search radius distances.

- Any area of contamination overlapping a drainage basin caused the associated intake to fail use waiver analysis for that contaminant.
- Transportation corridors passing through a drainage basin were noted to indicate the threat posed by transport of chemicals within the basin.
- The percentage of the county planted in the seven crops mentioned above was listed to indicate agricultural chemical use within the drainage basin.

Many of the rivers that supply water to systems in Missouri have their headwaters outside the state. To fully evaluate the potential for contamination within those drainage basins, we would have to collect data for large areas outside of the state. For example, the Mississippi and Missouri River drainage basins cover large portions of the United States. Because collecting data for those areas would be impractical, we have recommended to MDNR that use waivers not be granted to river supplies.

The following provides details on how the analysis was performed. The GIS searches around each wellhead for each radius and notes which contaminant sites affect which wellheads. If a contaminant falls within that radius, we recommend that the wellhead be monitored. In this example, the well is affected by one contaminant within the ¼-mile radius, two within the ½-mile radius, and four within the 1-mile radius.

Results

The results of the use waiver analysis indicate which systems may be affected by the use of a chemical near a water source. Several results show the substantial savings realized from our analysis. For example, the analysis showed that only five wells serving four public drinking water systems were potentially affected by dioxin and should be monitored. By not testing the remaining systems for dioxin, the state can realize a considerable cost savings, as the test for dioxin is the most expensive test to perform.

The final wellhead system analysis shows that the ½-mile buffer analysis affected a total of 447 wellheads in 241 systems. That is, a chemical site or pesticide dealership was found within ½ mile of 447 public wellheads. A result form was generated for each of the 1,340 systems in the state listing each well or intake and the potential threat posed by nearby contaminant sources.

The cost of testing all wellhead systems for all 43 contaminants without issuing use waivers is more than $15 million (see Table 1). According to our analysis, CARES estimates that only $1.8 million need be spent to monitor vulnerable wells. Therefore, the state can save more than $13.5 million in monitoring costs.

Table 1. Estimated Cost Savings for Public Drinking Water Systems

Method	Estimated Total Cost	Estimated Mean Cost per System	Estimated Total Cost Savings
No use waiver	$15,533,100	$12,200	$0
With use waiver	$1,813,900	$1,400	$13,719,200

Summary and Recommendations

To date, the investment the state made in the vulnerability assessment project has provided many benefits. The state saved several million dollars in testing costs and developed several spatial and nonspatial databases that will have many uses. In addition, the project established a basic framework for future assessments, which EPA requires on a regular basis.

The basic data required for use waiver analysis are the locations of water sources and the locations of potential contamination sources. CARES determined that the available data did not contain the information necessary to map these locations or that the data were of questionable quality. Many layers required update and correction. Considerable effort was necessary to improve existing locational information for both water source layers and chemical contaminant files. Local knowledge of an area was heavily relied upon to determine accurate locations, particularly contaminant sites. The vast majority of these sites contained only the address as the geographic reference. An address is not a coordinate system; it does not indicate a fixed location on a map. Because the location of any chemical detection site is of vital importance, state and federal agencies that collect these data need to record more complete geographic information. Ideally, a global positioning system could be employed to generate coordinates. Realistically, the recording of legal descriptions or directions from an easily located point would substantially improve the quality of the current databases.

In many cases, data resided in digital format; however, due to regulations or lack of agency cooperation, they could only be distributed in paper format. Reentering data from paper format into digital format required considerable time and expense. Interagency cooperation should be emphasized to reduce unnecessary data entry.

Using GIS to Evaluate the Effects of Flood Risk On Residential Property Values

Alena Bartošová, David E. Clark, Vladimir Novotny, Kyra S. Taylor
Marquette University, Milwaukee, WI 53201-1881

1. Introduction

Annually, flooding causes more property damage in the United States than any other type of natural disaster. One of the consequences of continued urbanization is the tendency for floodplains to expand, increasing flood risks in the areas around urban streams and rivers. Hedonic modeling techniques can be used to estimate the relationship between residential housing prices and flood risks. One weakness of hedonic modeling has been incomplete controls for locational characteristics influencing a given property. In addition, relatively primitive assumptions have been employed in modeling flood risk exposures.

We use GIS tools to provide more accurate measures of flood risks, and a more thorough accounting of the locational features in the neighborhood. This has important policy implications. Once a complete hedonic model is developed, the reduction in property value attributed to an increase in flood risks can, under certain circumstances, be interpreted as the household's willingness to pay for the reduction of flood risk. Willingness to pay estimates can in turn be used to guide policymakers as they assess community-wide benefits from flood control projects.

2. Hedonic Theory and Literature

The hedonic price model used in this study has its roots in the works of Lancaster (1966) and Rosen (1974). It is based on the premise that individuals can choose consumption levels of local public goods such as environmental quality through their residential location choice. The model views the price of individual houses as dependent on a bundle of housing characteristics. These characteristics include those related to the structure (e.g., lot size, number of bathrooms, etc.); the neighborhood (e.g., average commute time, median household income, etc.); the environment (e.g., variables related to flood risk); and fiscal factors (e.g., property tax rates).

There are several underlying assumptions in this model. The model assumes that the study area is a single market for housing services. It also assumes that all buyers and sellers have perfect information on the alternatives that exist and that the housing market is in equilibrium. This last assumption means that all households have made their utility maximizing choice in terms of residential location given the prices of alternatives, all of which just clear the market. The relationship outlined here can be linear only when repackaging of the house is possible, and in general, this is not the case. When an individual makes a residential location decision, they are accepting the entire bundle of housing characteristics. It is not possible to trade a house with two full baths upstairs for the exact same house with one full bath upstairs and one downstairs. Thus, the function is nonlinear.

Given the previous assumptions, the market clearing price of the house is treated as parametric and can be represented as p(Z), where $Z = z_{1,} z_{2, \ldots ,} z_n$ is a vector of n structural, neighborhood, and environmental characteristics. The housing market implicitly reveals the hedonic function, p(Z), which relates prices and characteristics. This price function p(Z) is a reduced form equation representing both supply and demand influences in the housing market. The implicit price of attribute n is given by the partial derivative of p(z) with respect to attribute n, or $p_n(z) = \partial p/\partial z_n$. That is to say, the partial derivative with respect to any of the aforementioned characteristics in the function can be interpreted as a marginal implicit price of that characteristic. This marginal implicit price is the additional amount that must be paid by any household to move to a bundle of housing services with a higher level of that characteristic. For example, the coefficient on the number of rooms in a home may be interpreted as the price that must be paid by the household to move from a house with eight total rooms to the same house with nine total rooms, all else constant. Since the function for housing is nonlinear, the marginal implicit price depends on the quantity of the characteristic being purchased.

Several hedonic studies specifically address the issue of flooding including the effect of floodplain regulations on residential property values (Schaefer 1990), the impact of subsidized and non-subsidized flood insurance on property values (Shilling et al., 1987), and the influence of flood risk on property values (Barnard 1978; Park and Miller 1982; Thompson & Stoevener 1983; Donnelly 1989; Speyrer and Ragas 1991; Shabman and Stephenson 1996). For the most part, the results from these studies indicate that location in a floodplain, or proxies for flood risk, negatively impacts residential property values. One study examined a major flood event (Babcock and Mitchell 1980); however, this was done by a comparison of prices before and after the event, and thus was vulnerable to bias due to omitted factors in the analysis. None of these studies measure flood risks directly, nor do they investigate the impact of a specific flooding event in a hedonic framework.

3. **Definition of Flood Risks**

A flood is defined for the purpose of this paper as a stream discharge greater than the capacity flow of the channel. This is obviously a very simplistic definition. For example, Williams (1978) presented 11 definitions of the channel bankfull flow, from which the flow that reaches the valley active floodplain is the one accepted by most river morphologists. A flood of certain magnitude occurs or is exceeded with a certain frequency. The most common flow used for delineation of floodplain is the flow with the recurrence interval T_r = 100 years, i.e. the risk of flooding is $r = 1 / T_r = 1/100 = 0.01$.

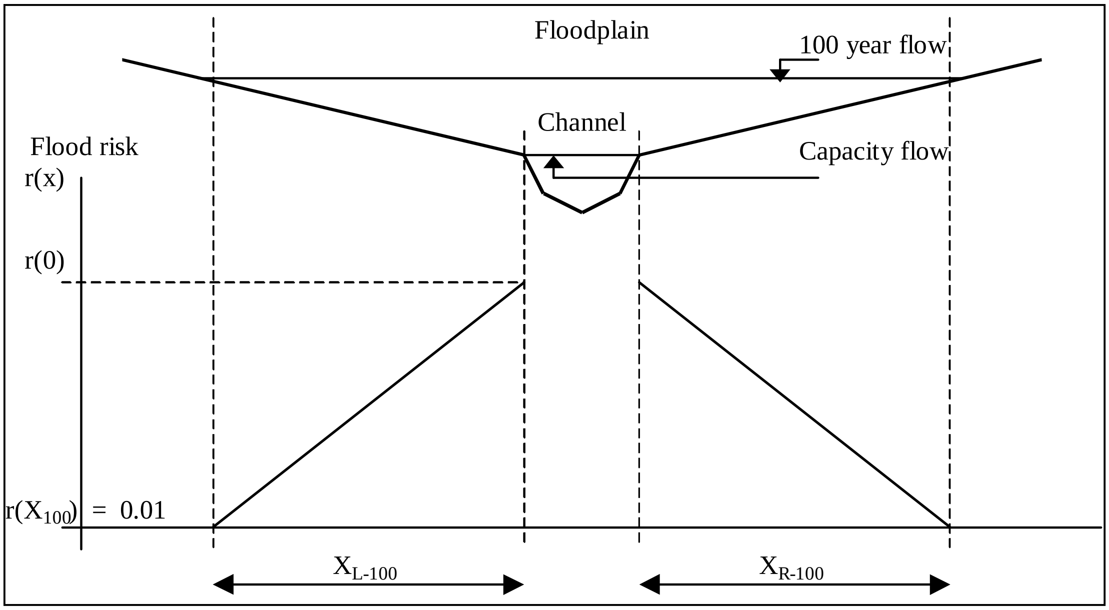
The delineation of the floodplain for a flow of given frequency is a tedious task. Such tasks usually involve the development of a complex hydrologic/hydraulic model. Once calibrated, the model can be used to simulate a wide range of flows and the flow-elevation relationship can be obtained. Hydraulic models can be combined with GIS systems to delineate a floodplain for any recurrence interval (e.g., McLin, 1993, Correia et al., 1998). However, this requires a considerable amount of data and substantial effort. Thus, a simplifying alternative has been proposed in this study.

The extent of 100-year floodplain, often used for engineering and flood insurance purposes, is delineated by Federal Emergency Management Agency (FEMA). The flood risk varies within the floodplain and decreases with increased distance from the channel. The properties located within the 100 years floodplain are under different risks of flooding and hence there is a need to express a flood risk relation in the urban floodplain.

A schematic representation of the following concept is shown in Figure 1. The channel can contain a flow with a certain recurrence interval. This flow is called a capacity flow, or bankfull flow. As one moves away from the river's edge, the probability of flooding decreases, and at some point at a distance x from the river the recurrence interval of flooding becomes 100 years, i.e., the risk of flooding is $r(x) = 0.01$. This is the extent of the 100-year floodplain that is useful for many engineering and flood insurance purposes.

Channels of natural streams are in equilibrium with the flow. Leopold, Wolman, and Miller (1995) document that channels of rivers in eastern and Midwestern US have a channel capacity that can contain a flow that has an approximate recurrence interval of about 1 ½ years. For example, if the smallest flow that leaves the channel is about a 2-year flow before urbanization, then the risk of flooding at the edge of the river is r(0) = 1 / 2 = 0.5.

Figure 1: Concept of flood risk



The scale of the risk function *r*(*x*) should be logarithmic, i.e., a zero risk of flooding is expected to occur at an infinitely large distance *x* from the river edge. The logarithmic form of the risk function is selected for convenience and simply expresses the fact that floods on rare occasions may extend further than the 100-year floodplain limits. The logarithmic risk function can be expressed as

$$r(x) = C\,10^{-Kx} \qquad \text{Eq. 1}$$

The function parameters in Eq. 1 can be easily estimated from the knowledge of the risk of exceeding the bankfull capacity flow and from the extent of the 100-year floodplain: *C* corresponds to the risk of exceeding the bankfull flow, or, *C* = *r*(0). The risk function can be integrated across the floodplain cross-section, as shown in the following equation, in which subscripts *L* and *R* correspond to the left and right bank floodplains:

$$R = \int_0^\infty r_L(x)dx + \int_0^\infty r_R(x)dx = r(0)\int_0^\infty [10^{-K_L x_L} + 10^{-K_R x_R}]dx \qquad \text{Eq. 2}$$

The magnitude of the floodplain shape coefficient, *K*, can be obtained from the extent of the 100-year floodplain at the point of interest on the river, denoted as X_{100}, and from the risk of exceeding the bankfull discharge, *r*(0):

$$\log\left[\frac{r(X_{100})}{C}\right] = \log\left[\frac{0.01}{r(0)}\right] = -K\,X_{100} \qquad \text{Eq. 3}$$

and

$$K = \frac{\log[r(0)] + 2}{X_{100}} \qquad \text{Eq. 4}$$

Finally, substituting for *K* in Eq. 2 from Eq. 4 yields the following expression for the floodplain risk parameter:

$$R = \frac{r(0)}{2.3\left(2 + \log r(0)\right)}\left[X_{L-100} + X_{R-100}\right] \qquad \text{Eq. 5}$$

The dimension of the floodplain risk parameter *R* is length/time, and a possible unit is meter/day. However, the unit does not have a physical meaning, as *R* is only a measure of the flood risk over a floodplain. *R* increases with an increase in the size of the floodplain and with an increase in the risk of overbank flow. This floodplain risk parameter changes along the stream. The integration of the flood risk over the watershed represents an overall risk of flooding of the watershed, the flood risk factor that can be used in comparing watershed management alternatives.

This characterization of flood risks will be used to assign unique values of flood risk to each property within the floodplain. The flood risk measure, *FRM*, calculated in a GIS environment, is a negative logarithm of the flood risk *r*(x). The anti-logarithm of the flood risk measure is basically a recurrence interval, i.e., *FRM* = 2 for T_r = 100 years.

4. Empirical Model

a. Study Area

The study area for this analysis is located approximately 11.5 miles (18.5 km) along the middle to lower sections of the Menomonee River through the cities of Wauwatosa and Milwaukee, Wisconsin. The Menomonee River is a 71.85 (15.5 km) mile river system and discharges into the Milwaukee River about 0.9 mile upstream of where the Milwaukee River enters Lake Michigan. This region was selected to encompass two significant areas, the city of Wauwatosa and the Valley Park neighborhood in Milwaukee. Wauwatosa makes up a great portion of the study area and lies within the Menomonee River watershed boundaries. Located west of Milwaukee in northern Milwaukee County, Wauwatosa is just over 13 square miles (34 km^2) with a population of 49,300. Furthermore, it is a high density residential area, with more than 22.8 persons per net residential acre (55 persons/ha). Valley Park, the other area of concern, is the smallest and most isolated neighborhood in Milwaukee. The study area is shown in Figure 2.

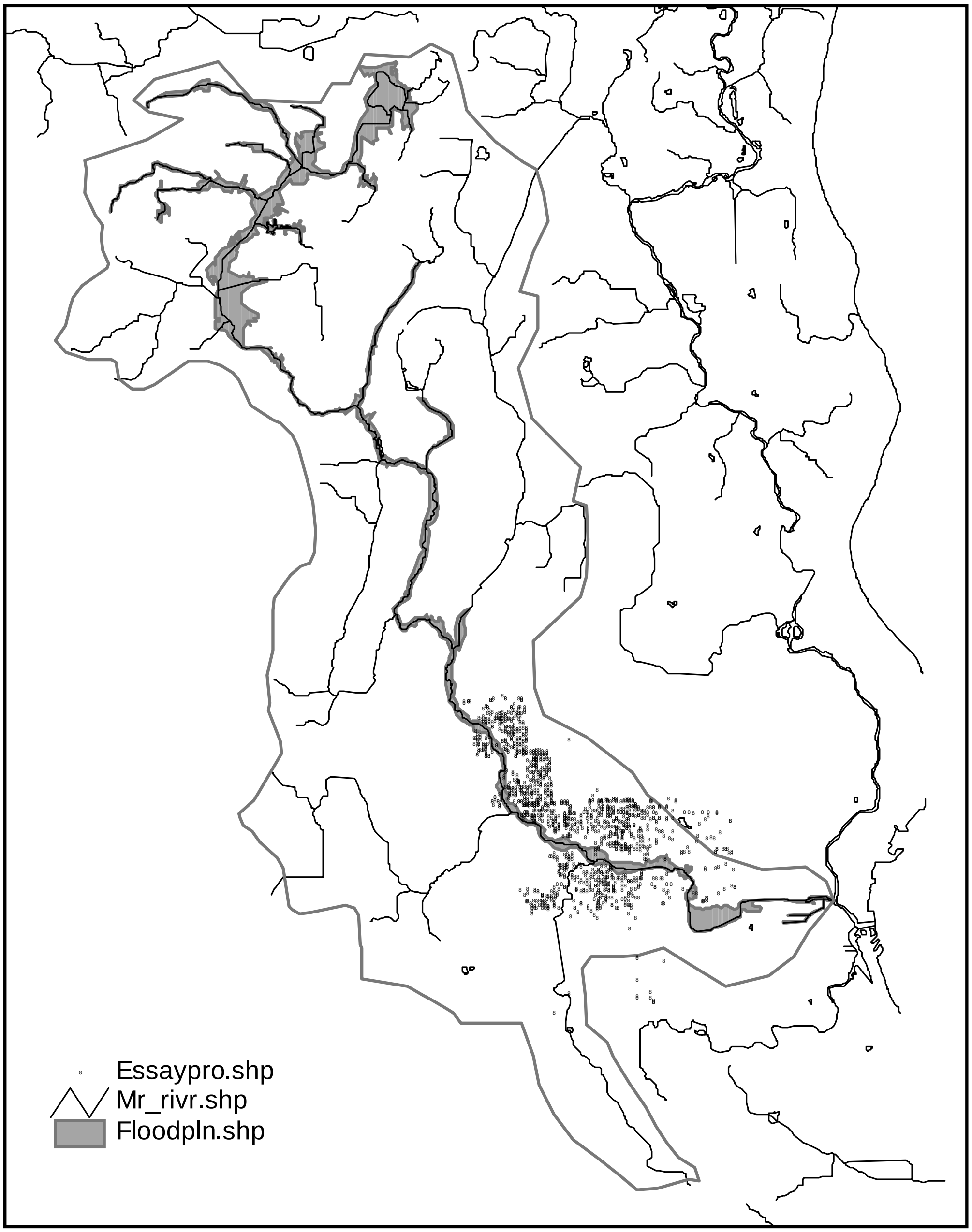


Figure 2: Menomonee River watershed. Location of properties in 100-year floodplain.

These two areas are significant for this study as a result of their susceptibility to flooding. Specifically, the study examines the short and intermediate run impacts of a 100-year flood that occurred in June of 1997. The flood was the worst rain for the Milwaukee Metropolitan area since August 6, 1986. After the first night of the rainfall, totals ranged as high as 9.78 inches (25 cm), indicating a flood recurrence interval exceeding 100 years. Roads were shut down and many residents lost power. Damage for Milwaukee County alone was estimated to be $37 million, including $24 million to residential property. About 70 homes in the County incurred major damage including collapsed basements and roofs forcing residents to evacuate their homes. Approximately 2100 homes sustained lesser damage. As a result of the flood, Wauwatosa submitted a Hazard Mitigation Grant Program application for the acquisition of a number of structures located in the floodway on the Menomonee River. They used Community Development Block Grant funds to acquire flood prone structures as a means of creating open space in the riverfront floodway. Of the 20,289 structures in Wauwatosa, about 738 are located in the special flood hazard area, 669 of which are residential. Due to its susceptibility to flood disaster, Wauwatosa was invited by FEMA in June of 1998 to participate in a nationwide effort to become a “Project Impact” community. This program would develop efforts to minimize the risk of damage from natural disasters. Valley Park also suffered from the flood in terms of water levels. However, there is a great sense of community in the neighborhood, which became evident in the recovery period following the disaster. Both Wauwatosa and the city of Milwaukee, in which “Valley Park” resides, are participants in the National Flood Insurance Program (NFIP); Wauwatosa entering in 1978 and Milwaukee in 1982. The NFIP implements floodplain management regulations which ensure that development in flood-prone areas is protected from flood damages. However flood insurance is mandatory only for those properties residing within the 100-year floodplain. This increase in cost associated with location in the floodplain may reduce property value for those houses.

b. GIS Analysis

ArcView, a Geographical Information System (GIS), was used in several aspects of this study. First, it was used to spatially define flood risks. Second, properties were geocoded to the street address, and finally location specific data were matched to each property. We describe each of these activities below.

The properties were geocoded to the precise street address using the ArcView GIS package. A key to the geocoding process is the accuracy of addresses, the geographic files, and matching of the addresses to the geographic files. The addresses and geographic files received from outside sources (MLS and Wisconsin Department of Transportation) are believed to be accurate given the sources' own incentive for accuracy of the files. ArcView assigns a score to each match made for the properties. Of the 1475 observations, 1402 of them (or approximately 95%) were given a score of 75 or above on a 100 point scale. The majority of these received a score between 98-100.[1] The resulting sample size is 1431, as 44 were unable to be geocoded and eliminated from the sample. Once geocoding of properties was completed boundary files for geographic areas were digitized if they were not already available as ArcView shape files. For example, the 100-year floodplain was geocoded from FEMA maps and maps provided by the Southeastern Wisconsin Regional Planning Commission (SEWRPC). Other spatial boundary data (e.g., school district boundaries, historic preservation district data) were also manually digitized.

Once the geocoding was completed, properties were matched to locational attributes of the neighborhood using one of three techniques. When a neighborhood characteristic was defined by a point in space (e.g., proximity to air quality monitors), straight line distance calculations between the property and the attribute were used. If the attribute was defined by a polygon (e.g., school districts, census block groups), then individual properties were mapped to the underlying polygon, and attributes of the polygon were attached to the property. Finally, buffers were defined for various types of line data (e.g., roads, railroads) and properties falling within the buffer zone were identified.

Turning to the calculation of property specific flood risks, two basic approaches were considered. The first is a vector-based approach that employed a custom developed ArcView Avenue scripts program. This approach permits estimation of risks only at specific points rather than for complete areas. The second more general approach works in a grid (raster) environment, and makes use of the Spatial Analyst Extension for ArcView. It permits flood risk to be calculated for the entire watershed, and specified points can be assigned the corresponding value from the underlying polygon. The second approach was selected because of its future applicability in watershed management applications.

---

[1] A possible reason for a score at the lower end of the spectrum would be misspellings. For example, if an address appears as “Menomone Pkwy” and the correct spelling would be “Menomonee Pkwy,” the addresses may still be matched and assigned a lower score as a result. For this reason, the matches receiving a score of less than 80 were interactively re-matched by the author to ensure accuracy and minimize error.

When we refer to the floodplain in this paper, it should be understood as the 100-year floodplain. The width of the floodplain is the key parameter in calculation of the flood risk, when r(0) is kept constant. The floodplain width for any specified point, both inside and outside the floodplain, is the distance of the flood fringe from the river bank for the river cross-section on which this point is located. The calculation of the floodplain width corresponding to the selected locations had to be done separately for inside and outside of the floodplain. The floodplain width is calculated as

$$X_{100} = X_W + X_F \qquad \text{Eq. 6}$$

or

$$X_{100} = X_W - X_F \qquad \text{Eq. 7}$$

where X_W is the distance from the river channel and X_F is the distance from the floodplain (see Figure 3)

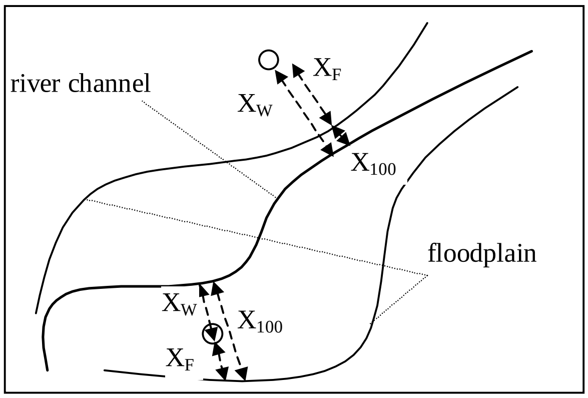


Figure 3: Calculation of floodplain width for locations inside and outside the floodplain

The floodplain was digitized as a polygon and used as such in calculations for the areas outside the floodplain. For the areas inside the floodplain, it had to be converted into a polyline and divided into several reaches. The calculation of the floodplain width for points inside the floodplain was calculated separately also for left and right banks, although the calculation followed the same procedure. The data essential for risk calculations include digitized maps of the river channel and 100-year floodplain, as well as the watershed boundaries. The risk associated with the capacity flow has been estimated separately using the information from USGS on capacity flow and the annual maximum series for the gage station in Wauwatosa. This station is located in the same area as the majority of the properties. The recurrence interval associated with the capacity flow is approximately 1 year, i.e., r(0) = 1.

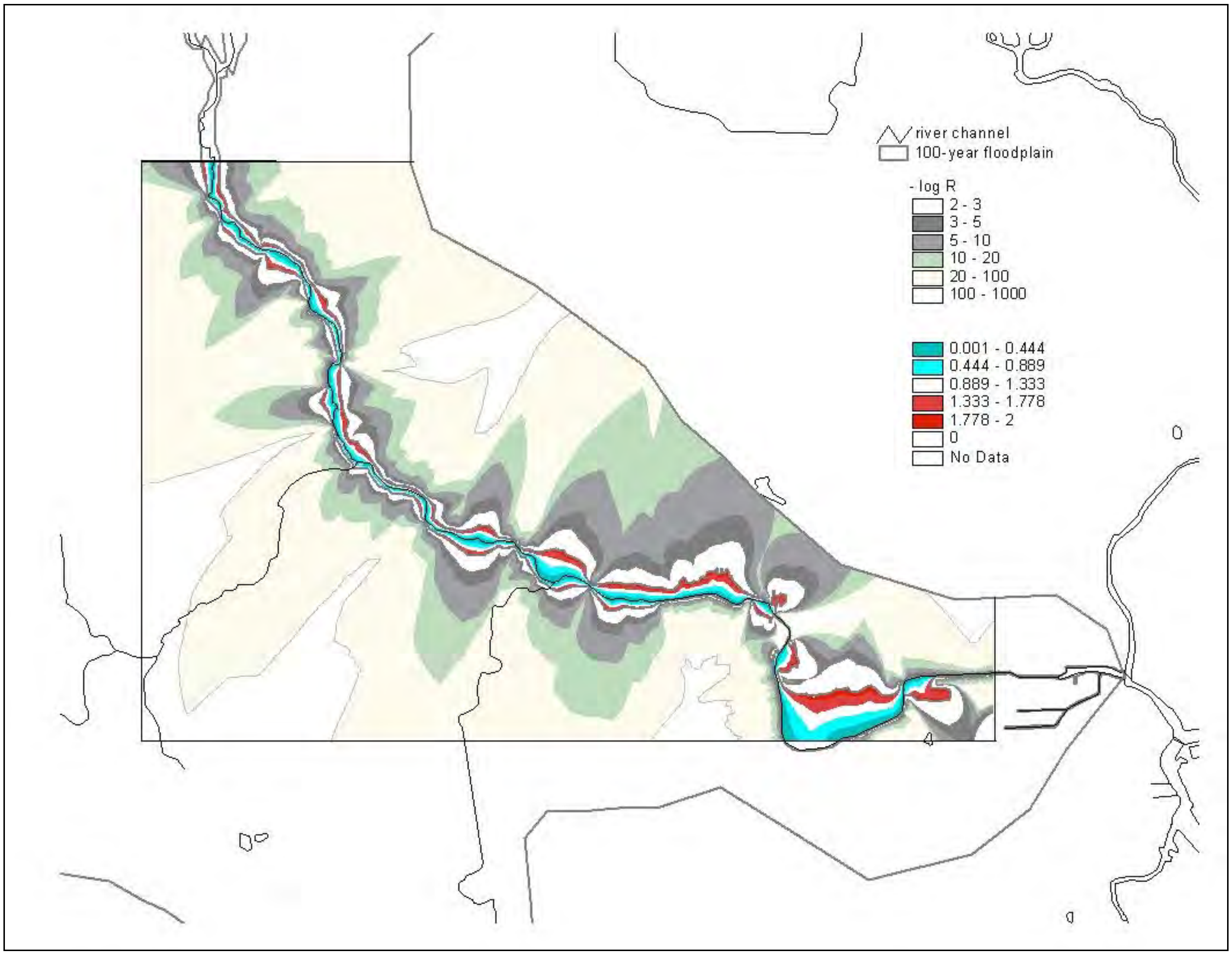


Figure 4: Flood risk measure

Figure 4 shows the flood risk measure, i.e., the negative logarithm of the flood risk, in the area where the properties are located. Individual properties were assigned a value corresponding to the underlying cell. The higher is this value, the lower is the likelihood of flooding for the specific property. An increase in this variable of one implies that flood risks decrease by an order of magnitude. For example, as you move from flood risk measure of 2 to 3 you move from a risk of 0.01 (i.e., once per 100 years) to 0.001 (i.e., once per 1000 years).

c. Description of the Data

Detailed house attribute data as well as the sales prices of the houses were obtained from the Multiple Listing Service (MLS) for the Milwaukee Metropolitan Area. Information was collected for each transaction, listed through the MLS, for the time period January of 1995- July of 1998. This time frame provides an adequate period for property value fluctuation to occur as a result of the flooding event in June of 1997, if this is the case. A total of 1,965 properties were listed through the MLS in the study area for the time period examined. From this total, properties were eliminated as a result of missing data for: the lot size (290), age of the house (198) and taxes (2). Furthermore, the MLS database only includes properties sold through realtors, and thus leaves out of the sample properties sold directly by the owner. This may reduce the possibility of including "non-market" transactions in the sample, assuming that properties sold to relatives or close friends may be transacted by this means. Finally, as noted above, 44 properties were lost as a result of geocoding difficulties, yielding a total sample of 1431 properties.

The variables in the model are organized into six categories: *Structural, Neighborhood*, *Fiscal, Disequilibrium, Time Related* and *Flood.* Many influences are controlled within the neighborhood category in order to avoid misspecification biases and to account for spatial influences. For simplicity, the fiscal variable (tax rate) and the disequilibrium control (days on the market) are included in the *Neighborhood* category for the specification. Following Cropper (et al.) a semi-log specification is chosen, and the model is specified by Eq. 8.

LnRPRICE = *f* (*Structural,Neighborhood,Time Related,Fiscal,Disequilibrium,Flood*) Eq. 8

The variable definitions and data sources are reported in Table 1, and descriptive statistics are in Table 2. The dependent variable is the log of real sale price of housing and is deflated by the housing component of the CPI (1982-84) for the month in which the property sold.

i. *Structural* Variables

The structural characteristics include the number of bedrooms, bathrooms, other rooms, presence of an attached garage, as well as square footage of the lot and the property. It is expected that an increase in any one of the previous characteristics will increase the sale price, assuming that these attributes increase the housing services a property provides. Measures of area are included in linear and quadratic form to account for non-linearity in these variables. Finally, the age of the house is included expecting a negative relationship between the age of the house and the sale price. This is based on an assumption that older homes may have dated technology lacking several beneficial features that would increase the housing service provided by the property.

ii. ***Locational*** **Variables**

Each property was matched to numerous locational variables, including those in the *Neighborhood* category. To account for various demographic characteristics, census data was attached accordingly to the appropriate property. The census block group data captures the racial and ethnic mix of the neighborhood. The sign for these variables cannot be predicted without knowledge of a home purchaser's cultural preferences. The characteristics also include measures of income and poverty, home occupancy, age of the neighborhood. Also, the model controls for the travel time to work and the population density of the neighborhood. The latter variable is included to control for aspects of the neighborhood correlated with density which are not measured (e.g., crime, cultural amenities).

The property tax is included to account for fiscal effects, expecting that increases in taxes would decrease the sale price. Also capturing fiscal impacts is the teacher student ratio for the high school district in which the property resides. A dummy variable is included to account for residence within Wauwatosa or Milwaukee, which may capture a submarket influence and perceptions associated with living in Wauwatosa (versus Milwaukee). The number of days a property was on the market is used in the model as a disequilibrium control variable.

Past studies have found historical preservation districts to positively impact property values (Clark and Herrin 1997; Coffin 1989). The coefficients may be positive in the case that creation of the district provides people with additional information about the housing stock and revitalizes the neighborhood, yet also may be negative if the structural restrictions reduce housing demand. There are a total of six preservation districts in this study area, three in Milwaukee and three in Wauwatosa. Dummy variables are included for each of the districts.

As indicated in the theoretical review of the hedonic price model, one of the influences on the property sale price is environmental quality. Several variables controlling for environmental quality factors are included within the neighborhood category including measures of air quality, and proximity to Toxic Release Inventory sites. Accounting for the impact of local annoyance factors is the proximity of a residence to both highways and rail lines, as well as being located on a major road. One would expect these factors to negatively affect property sale price in most cases. A variable is also included to capture scenic benefits of residing along the river, a positive environmental attribute. This is measured by a dummy variable for those properties residing on the Menomonee River Parkway. While some of the properties along the Menomonee River Parkway may also be susceptible to flooding, only 7 of the 13 properties along the Parkway are also in the 100-year floodplain. Thus, the effect of this variable should pick up the scenic benefits of the river, while holding constant the risk associated with flooding (accounted for by variables in the *Flood* category).

iii. *Time Related* Variables

The model also includes dummy variables in the *Time Related* category for both the year and season in which the property was sold. Business cycles may affect property values, and the year variables are incorporated to capture the possibility of that influence. Furthermore, the year variables may capture an interest rate effect. Similarly, the season dummies control for trends that may be associated with time. There are no expected signs for the variables relating to time.

iv. *Flood* Variables

Finally, variables representing the focus of this study are included in the *Flood* category and also capture environmental quality. Other studies (Speyer and Ragas 1991, Schaefer 1990, Donnelly 1989, Park and Miller 1982, Thompson and Stoevener 1983) have used dummy variables accounting for a property's location inside or outside of the 100-year floodplain. All, with the exception of Schaefer, have found a significant negative relationship between location in the floodplain and the sale price of a property. This study differs from the previous studies in that a continuous measure of risk is derived. This permits floodplains of any periodicity to be defined. We investigate floodplains in 100-year increments from 100-500 year floodplains. Over the 3-year period, 15 properties sold in the 100-year floodplain, and 32 sold within the 500-year floodplain. In addition, we examine the rate at which property values change within each increment.

A second objective is to analyze the short run and intermediate run effects of a specific flood event that occurred in June of 1997. To do so, two different measures are used. First, to measure the short run impact, the floodplain dummy is interacted with a dummy variable for whether the property was sold after the flood event. Of the 1431 properties in the sample, 512 of them were sold after the flood event and 4 of these were within the 100-year floodplain whereas 12 were within the 500-year floodplain. Second, to measure intermediate run effects, the floodplain dummy is interacted both with the dummy for whether the property was sold after the flooding event and the number of days between the flooding event and the sale of the house. If present, one would expect short run effects to be stronger than intermediate impacts, assuming that the consequences of the flood event will taper off in the minds of homeowners and buyers as time passes.

5. **EMPIRICAL FINDINGS**

The coefficients on control variables in the *structural*, *neighborhood, fiscal, disequilibrium* and *time related* categories differ minimally among the tables. To conserve space, these variables are reported only once, with subsequent regressions reporting only the flood category variables. Heteroskedasticity, a non-constant variance in the model's error term, is expected in this sample of data since variance in selling price is likely to differ between the low-end and high-end of the market. To test for the presence of heteroskedasticity, White's test is used and the null hypothesis of no heteroskedasticity is rejected at the 95% level of confidence for each regression (Gujarati, 1995). White's correction is employed to generate consistent estimates of the standard errors. All models estimated explained approximately 91% of the variation in the real housing price.

i. ***Structural* Variables**

All structural variables are significant at the 99% level of confidence, except the dummy accounting for whether the garage is attached. The number of garage spaces is significant, with each additional space increasing the value of the home by 4.8%. The number of bedrooms, other rooms, half baths, and full baths all positively impact property sale price. One additional half bath, full bath, bedroom, and other room, increases the property value by 11.2%, 6.2%, 5.0%, and 5.8% respectively. The large magnitude of the coefficient on the half bath variable suggests that it may be serving as a proxy for other structural features of the house. Both square footage variables, interior and lot, increase property value at a decreasing rate reflected by positive linear terms and negative quadratic terms. The partial derivative of sale price with respect to the interior square footage (∂Real Price/∂Building area) is equal to [$\$_{AREA} + 2 * \$_{AREASQ}$*Building area]. Evaluated at the mean for interior square footage (705.7 sq.ft. or 0.65 m^2), property value increases by 6.8% for an increment of 100 square feet (or 0.72%/m^2). Similarly, an increment of 1000 square feet for the lot size increases sale price by 1.7% (or 0.18%/m^2 evaluated at the mean). Finally, other things equal, age has a negative effect on property value (i.e., 1.6% for each additional 10 years). Inclusion of a quadratic term for age made both the linear and quadratic terms insignificant.

ii. ***Locational*** **Variables**

Evaluating the demographic variables taken from the block group data, many coefficients appear to be significant at the 99% confidence level. Exceptions include population density and the percent of occupied housing units, and percent owner occupied units. Population density has a negative relationship with property value suggesting that on the net, urban scale related disamenities have a stronger influence than that of amenities, yet the variable is insignificant. The racial variables reveal that higher concentrations of Asian (as compared to nonwhite other race) populations in a neighborhood positively affect property values. Specifically, a 1% increase in the Asian population increases property value by 3%. The impact of Hispanic populations, on the other hand, decrease real home sale prices by 2.5%. Percent White is positive and significant, raising prices 1.3% per 1% increase, whereas percent Black is not significant. Note, that most of the neighborhoods in the study areas have relatively few minority households. As expected, higher poverty rates in a neighborhood decrease home sale price, yet the effect is not great. Median household income, also reflecting socioeconomic dimensions of the neighborhood, positively impacts property values. Measured by the median year of houses built in the neighborhood, older neighborhoods have significantly higher priced housing in the study area. This is somewhat contrary to the sign on the age variable, yet it may suggest that people prefer historic surroundings in a neighborhood along with the benefits of a technologically advanced home. Finally, in line with the existing theory, each additional 10 minutes of commute time decreases the home sale price by 9%.

The tax rate, incorporating fiscal effects into the model, negatively impacts property value. Specifically, a 1% increase in the property tax rate (e.g. 4.3% to 5.3%) decreases the property sale price by 2.0%. The teacher student ratio included to proxy the quality of education does have a positive effect, yet is insignificant. Also insignificant is the number of days a house was on the market. The dummy variable accounting for city jurisdiction is significant indicating higher sales prices (by a magnitude of 19%) in Wauwatosa than in Milwaukee. However, Valley Park is only one small area in Milwaukee and the dummy accounting for location in Valley Park was insignificant.

The effect of historic preservation districts was positive in all cases confirming that historic preservation districts provide home buyers with additional information regarding the housing stock and serve the purpose of revitalizing the neighborhood. The influence of five of the six districts was significant. The most dramatic of all influences was that of The McKinley Boulevard Historic District in Milwaukee, increasing property value by 49%. The Concordia Historic District, also in Milwaukee, has a similar effect with 41% increase in property value as a result of residing within the district. The one historic preservation district that did not have a significant impact was The Wauwatosa Avenue Historic District. These districts were also interacted with age, yet the resulting variables were insignificant and doing so overwhelmed the significance of the individual dummies. Therefore, they were not included in the final regression.

Several other variables in the neighborhood category were indicative of the surrounding environmental quality. The quality of the air measured by the sulfur dioxide reading negatively impacts property sale price as we would expect, and this effect is significant at the 99% level of confidence. Furthermore, location within one mile of a Toxic Release Inventory site has the effect of reducing home sale prices by 2.8%, all else constant. Two of the variables representing local annoyance factors significantly reduce the sale price of a home. Specifically, residence on a major road and residence within a quarter of a mile of rail lines reduce home sale prices by 5.7% and 6.0% respectively. On the other hand, residence within a quarter of a mile of Interstate 94 increased sales prices for homes by 8.5%. It is possible that this variable is controlling for non-work related travel accessibility in addition to an annoyance factor. Finally, residence along the scenic Menomonee River Parkway has the significant effect of increasing property value by 7.1%, all else constant.

iii. ***Time Related*** **Variables**

The seasonal dummy variables are insignificant indicating that the season in which a house is sold has no impact on the sales price. The year dummy variables indicate that real housing prices have fallen over the time period 1995- July of 1998. The effect in 1996 is insignificant; however, housing prices significantly decreased for both 1997 and 1998.

iv. Flood Variables

There are two objectives in terms of flood risk for this study. The first objective is to determine the effect that flood hazard in general has on property value. In the first regression reported in

Table 3, we proxy flood risk using the negative log (base 10) of the expected flood frequency , i.e., flood risk measure (see
Figure 4). The log of the value is included due to the rapid rate at which flood risks fall as distance from the river increase, and elevation rise. The findings indicate a clear relationship between reduced flooding risk, and increased property values. However, the value of the coefficient is extremely low. This finding is not surprising, given that the vast majority of properties are well beyond even the 1000-year floodplain. Hence a reduction of risk from say 10E-23 to 10E-24 is of negligible value to those residents.

To investigate the variation of flood risks within floodplains, we explore several different specifications. First, we examine the 100-year floodplain. Although flood risk is continuously defined, lenders only require that properties in the 100-year floodplain purchase flood insurance. In Table 4, we report the findings on a regression that includes a dummy variable for whether the property lies within the 100-year floodplain. In addition, we interact that variable with the recurrence interval, i.e., anti-log of the flood risk measure. The recurrence interval takes on values between 6.3 (i.e., a flood is expected with a probability of 1/6.3) for the property closest to the river, and 100 for a property at the edge of the 100-year floodplain. Both the dummy variable and the risk interaction term are statistically significant. The findings suggest that properties at the edge of the river would sell for approximately 7.8% than those outside the floodplain. However, as flood risk diminishes by 10 years (e.g., from a one-year flood frequency to an 11-year frequency) property values would increase by 2.3%. This implies that the detrimental effect of the flood risk is eliminated after the expected flood risk falls to once every 33.3 years.

In Table 5, we add a second interaction term to consider the effect of a flooding event. The variable *Days since* is the number of days since the flood in June of 1997. Hence, it measures the effect of the flooding event on the impact of the 100-year floodplain. The inclusion of this variable renders the floodplain dummy variable insignificant, although it remains negative. This is due to multicollinearity between the two variables. Treating the coefficient on the dummy variable as point estimate, it suggests that properties (at the edge of the river) selling in the floodplain prior to the flood sold for 5.1% less than comparable properties outside the floodplain prior to the flood. Those selling a year after the flood would sell for 18.9% less than properties outside the floodplain. The pattern did not appear to be nonlinear, although note that it was not possible to capture longer-term effects due to the fact that the sample did not extend further into the future. Thus, it appears that at least over the short term, the flooding event did reduce property values beyond what they were prior to the flood.

In the final model presented in Table 6, we explore whether wider floodplains generate detrimental effects on properties within those areas. Thus, we define floodplains between 100 and 200 hundred years, 200 and 300 years, and so on. Given that the detrimental effects of flood risk appear to dissipate within the 100-year floodplain, it is not surprising that none of the other floodplain categories are negative and significant. Indeed, the region between the 300 and 400-year floodplain sells at a premium over those outside the floodplains. We also explored whether the flooding event negatively influenced any of the property values within the 200 year and beyond areas, and found no evidence of detrimental impacts.

6. Conclusions

This study employed GIS tools to more accurately characterize flood risks in an urban watershed. An interpolation scheme to evaluate the level of flood risk in the watershed has been developed and applied to the Menomonee River watershed. Together with a wide range of other locational attributes, flood variables were matched to geocoded properties to investigate impacts on housing prices. Our findings support the hypothesis that increases in flood risk decrease values for residential properties within the 100-year floodplain. Unlike other studies which conclude that there are uniform impacts within the floodplain, we find declining effects with reduced risk. Furthermore, there is evidence suggesting that flooding events heighten sensitivity to such risks and raise the property price premium associated with a given level of flood risk. Negative impacts beyond the 100-year floodplain are not found.

The use of GIS tools to complement statistical analyses of urban spatial problems will continue to grow as PC-based GIS software becomes more powerful, and geographic data sources more abundant. In addition, GIS tools can serve as a conduit for interdisciplinary work as geographic modeling in the physical sciences and engineering is integrated with spatial modeling by social scientists.

Acknowledgements

This research is sponsored by the EPA/NSF Watershed Management Program in a form of a grant to Marquette University. Views and findings included in the presentation are those of the authors and not of the funding agency.

Table 1: Variable Definitions and Data Sources
Dependent Variable and Variables in the *Structural* Category

Variable Name	**Definition [mean, standard deviation]**	**Source**	**Predicted Sign**
Real Price	Real sale price of the property (1982-84 dollars)	MLS	LnRPRICE is the dependent variable
Age house	Age of the house in years	MLS	-
Full bath	Number of full baths in house	MLS	+
Half bath	Number of half baths in house	MLS	+
Bedrooms	Number of bedrooms in house	MLS	+
Other rooms	Total rooms minus number of bedrooms	MLS	+
Building area	Area of the master bedroom+bedroom2+livingroom+kitchen in square feet Note: Due to data limitations, all of the square footage is not captured	MLS	+
Garage spaces	Number of garage spaces	MLS	+
Garage attached	1 = garage attached, 0 = otherwise	MLS	+
Lot size	Lot area in square feet	MLS	+

Variables in the *Neighborhood, Fiscal, and Disequilibrium Control* Categories

Variable Name	**Definition [mean, standard deviation]**	**Source**	**Predicted Sign**
Sulphur Dioxide	Distance weighted value of the nearest air monitor, computed as sulfur dioxide/distance of monitor to property	LandView III	-
Major road	1 = property resides on a primary road, 0 = otherwise	ArcView	-
Menomonee Parkway	1= property resides on the Menomonee River Parkway, 0 = otherwise	ArcView	+
¼ mile I94	1= property within a quarter of a mile of Interstate 94, 0 = otherwise	ArcView	-
Commute time	Average household travel time to work for the block group in minutes	1990 Census of Population and Housing	-
¼ railroad	1= property within a quarter of a mile of railroad tracks, 0 = otherwise	ArcView	-
Toxic Release Inv.	1= property within a quarter of a mile of a manufacturing facility on the Toxic Release Inventory, 0 = otherwise	BASINS	-

Variable Name	Definition [mean, standard deviation]	Source	Predicted Sign
Historic Preservation Districts	HPDTOSA 1= resides within The Wauwatosa Avenue Historic District, 0= otherwise HPDCHURCH 1= resides within The Church Street Historic District, 0= otherwise HPDWASH-HIGH 1= resides within The Washington Highlands Historic District, 0= otherwise HPDCONCORD 1= resides within The Concordia Historic District, 0=otherwise HPDMCKINLEY 1=resides within The McKinley Boulevard Historic District, 0=otherwise HPDHIMOUNT 1= resides within The Washington-Hi Mount Boulevards Historic District, 0=otherwise	Maps received from: Wauwatosa City Planning (first three) Milwaukee City Planning (last three)	
TS ratio	Teacher student ratio for the school district in which the property resides	Respective High Schools	+
Pop density	Population density in the block group, measured as people per square mile	1990 Census of Population and Housing	?
Median year built	Median year of houses built in the block group	1990 Census of Population and Housing	?
Median HH income	Median household income of the block group	1990 Census of Population and Housing	+
%Asian	Percent of the block group population that is Asian	1990 Census of Population and Housing	?
%Black	Percent of the block group population that is Black	1990 Census of Population and Housing	?

Variable Name	Definition [mean, standard deviation]	Source	Predicted Sign
%Hispanic	Percent of the block group population that is Hispanic	1990 Census of Population and Housing	?
%Other	Percent of block group population which falls into the "other" category	1990 Census of Population and Housing	+
%Occupied units	Percent of the block group housing units that are occupied	1990 Census of Population and Housing	+
%Owner occupied	Percent of block group housing units that are owner occupied	1990 Census of Population and Housing	+
%Poverty	Percent of block group population that is below the poverty line	1990 Census of Population and Housing	-
Tax rate	Tax payment / [sale price/1000]	MLS	-
Wauwatosa	1 = property resides in Wauwatosa, 0 = Milwaukee	MLS	+
Valley Park	1 = property resides in Valley Park, 0 = otherwise	ArcView	?
Days on market	Number of days the house was on the market	MLS	-

Time Related Variables

Variable Name	Definition [mean, standard deviation]	Source	Predicted Sign
Seasonal Dummy Variables	SPRING=1 (March-May),0=otherwise SUMMER=1 (June-Aug), 0=otherwise FALL=1 (Sept-Nov), 0=otherwise WINTER=1 (Dec-Feb), 0=otherwise	MLS	? Winter is omitted variable
Year	1= dwelling sold in ith year, 0=otherwise i = 1995, 1996, 1997, 1998	MLS	? 1995 is omitted variable

Variables in the *Flood* Category

Variable Name	Definition [mean, standard deviation]	Source	Predicted Sign
$Floodplain_{100}$	1= resides in the 100-year, 0=otherwise	ArcView	**-**
$Floodplain_{200}$	1= resides in space beyond 100 year flood and within 200 year flood, 0=otherwise		
$Floodplain_{300}$	1= resides in space beyond 200-year and within 300 year flood, 0=otherwise		
$Floodplain_{400}$	1= resides in space beyond 300-year and within 400 year flood, 0=otherwise		
$Floodplain_{500}$	1=resides in space beyond 400-year and within 500 year flood, 0=otherwise		
Flood Risk Measure	Minus log of flood risk	Arcview	**+**
Recurrence Interval	The expected number of years between flooding events	ArcView	**+**
After	1= after the June 1997 flood, 0 = otherwise	ArcView	**?**
Days since	The number of days since the June 1997 flood.	ArcView	**?**

Table 2: Descriptive Statistics

Dependent Variable and *Structural Characteristics:*

Variable	Mean	SD	Maximum	Minimum
RPRICE	79048.1	34708.90	360962.6	7348.029
Agehouse	59.970	16.678	138	1
Full bath	1.278	0.487	4	1
Half bath	0.423	0.497	2	0
Bedrooms	3.211	0.741	7	2
Other rooms	3.488	0.990	8	0
Building area	705.214	155.137	1917	400
Garage space	1.793	0.639	4	0
Garage attached	0.193	0.395	1	0
Lot size	7081.323	3768.827	58344	1381

Variables in *Neighborhood, Fiscal, and Disequilibrium Control Categories*

Variable	Mean	SD	Maximum	Minimum
Sulpher Dioxide	153080	53632.03	504252	91485.71
Major road	0.062	0.241	1	0
Menomonee Parkway	0.009	0.094	1	0
¼ mile I94	0.042	0.200	1	0
Commute time	16.991	2.239	32.633	12.435
¼ mile railroad	0.093	0.291	1	0
Toxic Release Inv.	0.468	0.499	1	0
HPD Tosa	0.006	0.078	1	0
HPD Church	0.006	0.078	1	0
HPD Wash. Highlands	0.003	0.058	1	0
HPD Concord	0.003	052	1	0
HPD McKinley	0.004	0.064	1	0
HPD Himount	0.008	0.087	1	0
TS ratio	0.118	0.082	0.21	0.03
Pop. Density	7247.6333	3530.725	27743.90	752.500
Median Year Built	1945.530	7.017	1975	1939
Median HH income	40259.25	11716.96	66,649	7557
%ASIAN	1.137	1.754	18	0
%BLACK	2.970	11.040	90	0
%HISPANIC	1.369	1.537	13	0
%OTHER	0.460	0.906	9	0
%OCCUPIED	0.977	0.024	1	0.765
%OWNEROCC	72.808	18.028	99	5
%POVERTY	5.021	9.20	81	0
Taxrate	0.028	2.181	0.077	0.009

Valley Park	0.009	0.094	1	0
Wauwatosa	0.633	0.482	1	0
Days on Market	54.023	67.673	1095	0

Time Related Variables

Variable	Mean	SD	Maximum	Minimum
Spring	0.282	0.450	1	0
Summer	0.336	0.472	1	0
Fall	0.234	0.424	1	0
Winter	0.401	0.490	1	0
Year95	0.157	0.364	1	0
Year96	0.302	0.459	1	0
Year97	0.321	0.467	1	0
Year98	0.220	0.414	1	0

Flood Related Variables

Variable	Mean	SD	Maximum	Minimum
Flood Risk Measure	24.562	26.104	179.42	0.8
Recurrence Interval_{100}	36.9	29.258	100	6.8
Recurrence Interval_{500}	174.102	167.652	489.778	6.8
Floodplain_{100}	0.0105	0.102	1	0
After	0.358	0.479	1	0
Days since	69.317	113.661	397	0

Table 3 – Hedonic Regression with Log Flood Risk

Variable	***Coefficient***	***t-score***	***Variable***	***Coefficient***	***t-score***
Intercept	10.81558	3.3085			
Structural Characteristics			*Time Dummy Variables*		
Agehouse	-0.001594	-3.149	Year 1996	-0.014904	-1.295
Bedrooms	0.049593	7.0307	Year 1997	-0.075591	-6.212
Full bath	0.061932	6.0275	Year 1998	-0.079498	-5.296
Half bath	0.112181	12.078	Spring quarter	-0.00728	-0.595
Other rooms	0.057908	11.015	Summer quarter	-0.009696	-0.845
Garage space	0.047633	6.6189	Fall quarter	-0.001184	-0.093
Garage attached	0.013503	1.1273	*Historic Preservation Districts and locational variables*		
Building area	0.001224	10.133	HPD Church	0.063261	2.982
Building area * Building area	-3.85E-07	-5.542	HPD Concordia	0.412596	3.312
Lotsize	2.10E-05	6.7995	HPD High Mount	0.141946	2.039
Lotsize*Lotsize	-2.49E-10	-4.832	HPD McKinley	0.486035	5.299
Neighborhood and Fiscal Characteristics			HPD Wauwatosa	0.069102	1.198
Sulpher Dioxide	-1.16E-06	-3.134	HPD Wash. Highlands	0.213099	8.95
Major road	-0.057245	-3.99	Wauwatosa	0.198344	10.31
¼ mile I94	0.084733	3.1272	Valley Park	-0.023755	-0.264
¼ mile railroad	-0.059753	-3.279	Menomonee Pkwy	0.071265	1.795
Commute time	-0.008686	-4.69	*Flood Risk Variables*		
Toxic Release Inv.	-0.027812	-2.633	Flood Risk Measure	0.000253	2.003
Teacher Student ratio	0.028262	0.3231	*Disequilibrium Control*		
Population Density	-2.91E-06	-1.355	Days on market	-8.17E-06	-0.115
Median HH Income	3.07E-06	3.9097			
%Asian	0.030403	4.2097			
%Black	0.006825	1.1918			
%Hispanic	-0.02546	-3.941			
%White	0.013137	2.3295			
%Owner occupied	-0.000667	-1.26			
% Occupied units	-0.001439	-0.003			
% Poverty	-0.004957	-3.852			
Tax rate	-0.020374	-19.32			
Median year built	-0.003079	-1.894			
R-squared	0.917731		Adjusted R-squared	0.914996	
Mean dep. variable	6.574281		S.E. of regression	0.137611	
F-statistic	335.6265		Log likelihood	831.532	

Table 4: Model II—Flood Risk within the floodplain

LnRPRICE = *f* (*Structure, Neighborhood, Time Sold, Flood*),

Variable	Coefficient	t-statistic
$Floodplain_{100}$	-0.078337	-1.931
$Floodplain_{100}$*Recurrence Interval	0.002332	3.4425

Table 5: Model III—Flood Risk and a Flooding Event

LnRPRICE = *f* (*Structure, Neighborhood, Time Sold, Flood*),

Variable	Coefficient	t-statistic
$Floodplain_{100}$	-0.050991	-1.041
$Floodplain_{100}$*Recurrence Interval	0.002091	2.6966
$Floodplain_{100}$*Days Since Flood	-0.000378	-2.233

Table 6: Model III—Flood Risk in Expanded Flood Zones

LnRPRICE = *f* (*Structure, Neighborhood, Time Sold, Flood*),

Variable	Coefficient	t-statistic
$Floodplain_{100}$	-0.05261	-1.064
$Floodplain_{100}$*Recurrence Interval	0.002184	2.5027
$Floodplain_{100}$*Days Since Flood	-0.000366	-2.177
$Floodplain_{200}$	-0.020201	-0.323
$Floodplain_{300}$	-0.046497	-1.366
$Floodplain_{400}$	0.143638	4.87
$Floodplain_{500}$	-0.007187	-0.118

REFERENCES

Babcock, M., and Mitchell, B. 1980. "Impact of Flood Hazard on Residential Property Values in Galt (Cambridge), Ontario." *Water Resources Bulletin*, 16: 532-537.

Barnard, J.R. 1978. "Externalities from Urban Growth: The Case of Increased Storm Runoff and Flooding." *Land Economics,* 54: 298-315.

Clark, D.E., and Herrin, W.E. 1997. "Historical Preservation Districts and Home Sales Prices: Evidence from the Sacramento Housing Market." *The Review of Regional Studies*, 27: 29-48.

Coffin, D.A. 1989. "The Impact of Historical Districts on Residential Property Values." *Eastern Economic Journal*, 15: 221-28.

Correia, F.N., Rego, F.C., Da Graca Saraiva, M., and Ramos, I. 1998. "Coupling GIS with hydrologic and hydraulic flood modeling." Water Resources Management, 12(3):229-249.

Cropper, M.L., L.B. Deck, and K.E. McConnell. 1988. "On the Choice of Functional Form for Hedonic Price Functions." *Review of Economics and Statistics,* 70: 668-675.

Damianos, D., and Shabman. L.A. 1976. "Land Prices in Flood Hazard Areas: Applying Methods of Land Value Analysis." Offices of Water Resources Research. Project A-054-VA.

Donnelly, W.A. 1989. "Hedonic Price Analysis of the Effect of a Floodplain Location on Property Values." *Water Resources Bulletin*, 25: 581-586.

Gujarati, Damodar N. 1995. *Basic Econometrics* McGraw-Hill, Inc., New York.

Kiel, K.A., and McClain, K.T. 1995. "House Prices during Siting Decision Stages: The Case of an Incinerator from Rumor through Operation," *Journal of Environmental Economics and Management,* 28: 241-55.

Lancaster, K.J. 1966. "A New Approach to Consumer Theory," *Journal of Political Economy*, 74:132-157.

McLin, S.G. 1993. "Combined GIS-HEC procedure for flood hazard evaluation." Report No. LAUR933008, Los Alamos National Lab., NM, USA

Park, W.M., and Miller, W.L. 1982. "Flood Risk Perceptions and Overdevelopment in the Floodplain." *Water Resources Bulletin,* 18: 89-94.

Rosen, S. 1974. "Hedonic prices and implicit markets: product differentiation in pure competition." *Journal of Political Economy*, 82: 132-57.

Shabman, L. and Stephenson, K. 1996. "Searching for the Correct Benefit Estimate: Empirical Evidence for an Alternative Perspective." *Land Economics*, 72: 433-39.

Shaefer, K.A. 1990. "The Effect of Floodplain Designation/Regulations on Residential Property Values: A Case Study in North York, Ontario." *Canadian Water Resources Journal,* 15: 319-333.

Shilling, J.D., Sirmans, C.F., Benjamin, J.D. 1987. "Flood Insurance, Wealth Distribution, and Urban Property Values." *Journal of Urban Economics,* 26: 43-53.

Speyrer, J.F., and Ragas, W.R. 1991. "Housing Prices and Flood Risk: An Examination Using Spline Regression." *Journal of Real Estate Finance and Economics,* 4: 395-407.

Thompson, M.E., and Stoevener, H.H. 1983. "Estimating Residential Flood Control Benefits Using Implicit Price Equations." *Water Resources Bulletin,* 19: 889-95.

Tobin, G.A., and Newton, T.G. 1986. "A Theoretical Framework of Flood Induced Changes in Urban Land Values." *Water Resources Bulletin,* 22: 67-71.

US EPA (1998): http://www.epa.gov/ …BASINS

Williams, G.P. (1978): Bank-full Discharge of Rivers. *Water Resources Research* 14(6):1141-1154

Zimmerman, R. 1979. "The Effect of Flood Plain Location on Property Values: Three Towns in Northeastern New Jersey." *Water Resources Bulletin*, 15: 1653-6

A GIS Demonstration for Greenbelt Land Use Analysis

Joanna J. Becker
Environmental Planning Services, Santa Rosa, California

Purpose

The goal of this project was to demonstrate what analyses could be undertaken with a GIS program without substantial GIS training or time input. The demonstration attempted to show how planning staff and decision-makers could easily and usefully employ GIS. It was not intended as a complete study of all possible variables. Only available data were used. Diverse techniques were presented while keeping the content as simple and relevant as possible. The project was designed as a demonstration using regional scale data and was combined with another parcel-specific demonstration that showed urban GIS applications.

The demonstration showed the following modeling techniques:

- Buffer zones
- Combination of variables (overlays)
- Weighting of values
- Absolute value variables
- Reclassification of final values

Site Location

The San Luis Obispo watershed comprises an area of approximately 84 square miles. The watershed drains into the Pacific Ocean at Avila, California. The major creek in the watershed, San Luis Obispo Creek, is a perennial creek, but many of its tributaries have only seasonal flow. Agriculture and grazing are the major land uses in the watershed, although a significant number of areas have been developed. Growth of these areas is moderate to limited but has a pronounced effect on the watershed. The watershed also supports a large amount of riparian and other natural vegetation. Figure 1 demonstrates the distribution of land cover/land use within the watershed.

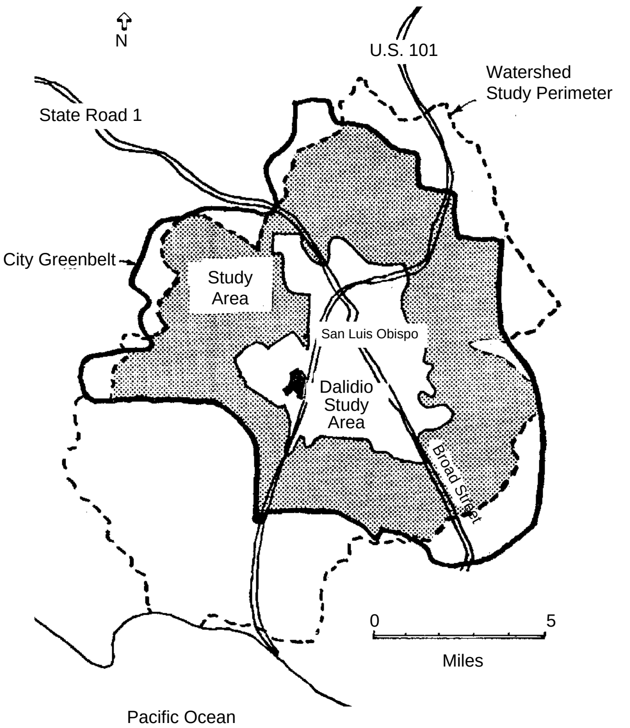


Figure 1. A generalized map of the San Luis Obispo area showing the location of the two ARC/INFO GIS study areas: (1) the parcel analysis in the Dalideo area and (2) the regional analysis in the city greenbelt stippled area surrounding the city. All boundaries are approximate and are for schematic purposes only (1, 2).

Background

The City Council approved the open space element of the San Luis Obispo General Plan in January 1994 and identified a greenbelt area that extended from the Urban Reserve Line approximately to the boundaries of the San Luis Obispo Creek watershed. The intent of this greenbelt area is to provide a buffer to the city and to preserve the agricultural and natural resources of the area.

The data for the watershed were already available through the work of the Landscape Architecture Department of California Polytechnic State University at San Luis Obispo for a study of San Luis Obispo Creek (1), and the city's greenbelt area lay approximately within those boundaries (see Figure 1). Variables were selected that could be extracted from the available data.

The data for the creek study were initially entered into workstation ARC/INFO in a polygon format. They were then transferred to MacGIS, a PC raster program, for simplicity of use. The final product was then transferred back to ARC/INFO as a grid format and viewed in a PC version of ARC/VIEW using DOS files.

In interpreting the overlay of values, the assumption was made that the occurrence of high values for the most variables would result in the most suitable land for that land use. This was presented as a range of values derived from the total values divided into three approximately equal groups of high, medium, and low. In addition to providing a composite analysis, however, any one of the data sets can be queried separately such that, for example, slopes greater than 20 percent could be identified or two layers such as storie index and distance from roads could be compared.

Criteria

Note that the ratings of high, medium, and low are based on available data, and the rating of low implies no suitable use. In addition, these ratings do not imply that categories rated low could not be used for a particular land use but rather that other land uses might be more appropriate. For example, open space use was rated low for flatter slopes but only because this category would likely be more suitable for agriculture.

The demonstration used an existing cell size of the data on the MacGIS program of 75 meters per side, which is assigned during the initial conversion process. Therefore, buffer zones are in aggregates of 75 meters. This size cell does not allow for minute analysis but reduces the size of the files, which may become extensive in raster format.

In presenting the final analysis, land contained within the urban reserve limit line has been excluded.

Procedure

Initially, eight variables from the available data were deemed suitable for this analysis:

- Slope
- Storie index (indicating soil fertility)
- Distance from major roads
- Distance from creeks
- Erosion hazard[1]
- Oak woodlands
- Land use compatibility
- Grasslands[2]

After selecting the six variables, the categories were recoded to conform to a rating of high, medium, or low.

Each land use was then evaluated separately for every variable except for combining the variables of slope and distance from creeks for rangeland analysis. In this case, composite values were assigned to the two variables, then recoded to produce a high, medium, or low rating.

After obtaining maps for each of the variables according to land use, the maps were compiled to indicate the density of overlays for each land use category. In assessing the suitability of land for the three land uses, the values of all the variables, except land use, were aggregated and a rating system developed. In addition, a double weight was assigned to the storie index in evaluating agriculture (because this is a primary index for considering prime agricultural land). If less than 75 meters, the distance from creeks also received added weight in consideration of open space preservation (because this is likely to ensure the least erosion and pollution to the waterways). The weighting then altered the scores as follows:

Land Use	Attributes	Number of Values
Agriculture	5	6
Rangeland	4	4
Open space	5	6

After the values had been assigned for each ranking, further recoding established three categories of high, medium, and low for each land use.

The land use buffer was added to this recoded aggregate map and resulted in an additional three values due to the interaction of the buffer with each category. These additional values were recoded according to each land use to produce a final map with three values.

The Urban Reserve Area was overlaid on the final map to exclude urban areas.

Assumptions

In determining what properties would be most suitable for each land use, the following assumptions were made.

[1] After reviewing the material, erosion hazard was eliminated because it was similar to the Soil Conservation Service (SCS) storie index data while the identification of native grasslands fell only within the area currently designated as open space land, so it was not included.

[2] See above note.

Open Space

This land is desirable to preserve as open space because of the existence of scenic or significant natural resources. It could also be land that is inappropriate for other uses due to the presence of such factors as steep slopes or poor soils. A distinction is made in the final map between land that is designated open space for recreational uses, such as parks, and land preserved for habitat or species protection. Open space adjacent to an urban area would be rated high if public accessibility was desirable but low if its purpose was resource protection or preservation. Separate maps based on two types of proposed uses present the contrast in analysis.

The analysis of the variables thus was rated as follows:

A. Land with steep slopes and therefore less suited for other purposes.

B. Land that has oak woodland vegetation resources.

C. Riparian land.

D. Low storie index indicating a lack of suitability for agriculture or rangeland.

E. At least approximately one-eighth of a mile from a major road to avoid negative impact on wildlife.

F. Either approximately one-quarter of a mile from urban areas if designated to protect resources or adjacent to urban areas if designated to serve as parks and recreation.

Agriculture

This land use includes all forms of agricultural activity; obviously, its suitability for specific crops and practices would vary. The determination of suitability would need to be made on a site-specific basis.

For the purpose of general agricultural suitability, the highest land suitability for agriculture was a rating of the variables as follows:

A. Land that does not have oak woodland.

B. Land that is not close to perennial creeks (to avoid fertilizer/pesticide runoff contamination).

C. The flattest slopes.

D. The highest storie index.

E. Proximity to a major road (considered an advantage for trucking and farm equipment access).

Rangeland

Some types of livestock can graze under most conditions, but for purposes of this analysis, land more suited for either open space or agricultural designation was rated above that of rangeland. The major limitation to suitability of land for rangeland was a combination of steep slopes and proximity to creeks. A rating of medium for the other variables was considered the most desirable for rangeland purposes.

Details of the Variables

Storie Index

In determining the most suitable uses according to soil fertility, the SCS storie index rating was used, with a modification of the categories to three to accommodate the ratings of high, medium, and low that were used throughout the analysis. Therefore, the first two SCS categories of excellent and good were combined into Category 1. Categories fair and poor were combined into Category 2, and categories very poor, nonagricultural, urban, and mines were combined to compose Category 3.

Subsequently, recoding was undertaken to prioritize these categories according to land use:

Agriculture	High for Category 1
Rangeland	High for Category 2
Open space	High for Category 3

Roads

The five principal arterials of the watershed were used in this analysis:

- Highway 1
- Highway 227
- Los Osos Valley Road
- U.S. 101
- Avila Valley/San Luis

A buffer on each side of the road was created, which was then recoded into the three categories of more than 75 meters, 75 meters to 150 meters, and greater than 150 meters. Each land use was then evaluated according to these criteria, with agriculture deemed the most suitable closest to the roads and open space the least suitable.

Slope

The slope categories in the San Luis Obispo watershed data set were divided into a number of classes, which were recoded into the most appropriate grouping of three classes for the analysis of the three land uses. The existing categories were not altered for the purpose of the demonstration, so they do not necessarily represent the most ideal slopes for the particular uses. A separate category of less than 10 percent slopes was provided for agriculture because most agricultural practices require flat land.

Slopes between 10 and 21.5 percent would be limited to such activities as orchards or vineyards.

Agriculture:	< 10 percent	High (H)
	10 to 21.5 percent	Medium (M)
	> 21.5 percent	Low (L)
Open space:	< 10 percent	L
	10 to 46 percent	M
	> 46 percent	H

No slope analysis was undertaken for rangeland because this category was combined with that for streams (see Streams section).

Urban Adjacency

Urban adjacency was treated as an overlay of the aggregate map of the other variables because it is an absolute value. That is, this variable has no ranking. Land is either within or outside the buffer. The suitability of each land use adjacent to urban areas was determined, then the aggregate map was adjusted according to a comparison of the recoded aggregate values with the designated ranking of land use suitability.

The first step was to recode the existing data on land uses (interpreted from 1989 aerial photography obtained from the United States Department of Agriculture [USDA], Agricultural Stabilization and Conservation Service, Atascadero, California) into urban/commercial areas and nonurban/noncommercial areas. A buffer zone of approximately one-quarter of a mile was then applied around the urban/commercial areas and an analysis undertaken of suitability for the three land uses to be located within this buffer.

In making the analysis, the following assumptions were made:

Land Use	Ranking	Reason
Agriculture	L	Conflict with dust, noise, pesticides, and urban use
Rangeland	M	Fire hazard of open grassland near buildings
Open space	H	Most suitable if used as parks/recreation
	L	Least suitable if designated for habitat/wildlife protection

Therefore, the analysis provided for two planning alternatives for open space, with the scenarios presented as separate maps overlaid on the aggregate map for the other rangeland variables.

In interpreting this map, the combined values were rated according to the above criteria, with any values lower than the desired ranking receiving a low value no matter what the aggregate value had been. This produced the following results:

- **Agriculture:** A rating of low for any land within the buffer zone.
- **Rangeland**: A rating of medium for any land within the buffer zone except that rated low for the aggregate map.
- **Open Space:** A rating equivalent to the rating of the aggregate map if the land was to be used for parks/recreation, or a rating of low for any land within the buffer zone if the land was to be used for habitat protection.

Streams

The original data for streams (from 1:24,000-scale USGS maps), which included both intermittent and perennial creeks, were used for rangeland analysis. A buffer of 150 meters was then applied to these streams, and these data were combined with the slope analysis. This combination was important in evaluating the erosion hazard and resulting stream pollution caused by nitrogen waste and hoof disturbance. The complete stream complex for the watershed area was therefore evaluated using the following matrix:

Stream Buffer	Slope		
	< 21.5 percent	21.5 to 46 percent	> 46 percent
< 75 meters	1/L	2/M	3/L
75 to 150 meters	4/M	5/H	6/M
> 300 meters	7/H	8/H	9/M

The stream data were then modified for agricultural and open space land use analysis to indicate only the perennial creeks as defined by the California Department of Fish and Game. Buffers were created for this as follows:

Stream Buffer	Open Space	Agriculture
< 150 meters	H	L
150 to 300 meters	M	M
> 300 meters	L	H

Oak Woodlands

The recoding of oak woodland data for agriculture was different than for the other two land uses because the presence of oak woodlands is not conducive to agriculture:

No oak woodlands	H
< 10 percent	M
> 10 percent	L

The suitability of oak woodlands for open space and rangeland was ranked as follows:

	Open Space	Rangeland
< 33 percent	L	M
33 to 75 percent	M	H
> 75 percent - high	H	L

References

1. Hallock, B.G., L.S. Bowker, W.D. Bremer, and D.N. Long. 1994. Nutrient objectives and best management practices for San Luis Obispo Creek. San Luis Obispo, CA: Coastal Resources Institute, California Polytechnic State University.
2. Community Development Department of San Luis Obispo. 1994. Open space element for the city of San Luis Obispo general plan. San Luis Obispo, CA.

Small Is Beautiful: GIS and Small Native American Reservations—Approach, Problems, Pitfalls, and Advantages

Jeff Besougloff
Upper Sioux and Lower Sioux Indian Communities, Redwood Falls, Minnesota

Background

The Lower Sioux Indian Reservation

The Lower Sioux Indian Reservation covers 1,743 acres in southwestern Minnesota bordering the Minnesota River. The land base consists of several hundred acres of prime, flat agricultural land, a large wetlands slough complex, prairie pothole wetlands, bottom land wetlands, small lakes, and approximately 250 acres of timber and brush. The elevation ranges from Minnesota River level to the adjacent bluffs several hundred feet higher.

This rural reservation contains a moderate amount of infrastructure, including paved and dirt roads, 12-acre sewage lagoon serving a moderately sized casino, community water system composed of a tower and small treatment plant for the 90 mostly single-family dwelling homes, convenience store/gas station/gift shop, community center, small two-room schoolhouse, pottery works with gift shop, warehouse, and church. The casino-fueled economic boost to the community recently resulted in improvements to infrastructure and plans for additional projects.

Office of the Environment

The tribal government was formed under the Indian Reorganization Act of 1934. The governing body is an elected five-person tribal community council that administers several government departments and is responsible for all government activities.

Under a U.S. Environmental Protection Agency (EPA) Region 5 multimedia grant, the Upper Sioux and Lower Sioux Office of the Environment (OE) was formed in late 1992. This unique joint venture between two tribes and EPA envisioned moving the tribal governments into compliance with major federal environmental legislation. At the present time, only the Lower Sioux are developing a tribal geographic information system (GIS). Therefore, this article is solely applicable to this community, although adoption of an Upper Sioux Community GIS would likely follow a similar lifeline.

Environmental Regulation in Indian Country

Reservations are subject to a bewildering array of environmental regulations. Numerous meetings, publications, projects, and court decisions are devoted to determining what law does or does not apply on any particular reservation. In very general terms, the following can be stated: state environmental regulations do not apply, federal regulations do apply, and tribal regulations may apply. From a tribal environmental employee's point of view, numerous environmental regulations (whether federal or tribal) do exist that apply to reservation activities and land, and they require compliance.

The Problem and the Solution

The OE's responsibility is to bring the reservation into compliance with the 14 major pieces of environmental legislation administered through EPA and directly applicable to tribes. The OE finds itself responsible for any and all other applicable environmental regulations and all other less-regulated environmental media. The OE currently has a staff of one.

In addition to the responsibility of moving the tribes into compliance with federal environmental regulations, the OE also develops environmental infrastructure, institutes environmental programs, and performs grant writing. Lower Sioux programs currently include Clean Water Act (CWA) and Safe Drinking Water Act (SDWA) compliance, solid waste planning including development and institution of a household recycling program, wetlands regulations compliance, wetlands mapping and restoration, National Environmental Policy Act (NEPA) compliance and site assessments, basic hydrological data gathering and mapping, radon testing and mitigation, environmental education as necessary, SARA

Title III compliance and planning, and a variety of related tasks.

Contracts or grants are currently being administered under several Bureau of Indian Affairs (BIA) programs, U.S. Geological Survey (USGS) and U.S. Army Corps of Engineers (COE) matching funds programs, two EPA programs, one Federal Emergency Management Agency (FEMA) program, one Administration of Native Americans (ANA) program, one Great Lakes Intertribal Council (GLITC) program, and a cooperative project with the National Tribal Environmental Council (NTEC).

Needless to say, responsibilities of the OE are limited by staff hours rather than need. As distressing as the reservation's unaddressed environmental needs are, equally distressing (prior to GIS development) was the helter skelter manner in which the OE digested the data and information flowing into the office. Because of its broad responsibilities and the administrative problems being encountered, the OE began to investigate developing a tribal GIS.

The Lower Sioux GIS

System Selection

The Lower Sioux GIS system is a networked PC station through the Bureau of Indian Affairs Geographic Data Service Center's (BIA GDSC's) two Sun MP690 SparcServers in Golden, Colorado. The GDSC is the hub of BIA's GIS and remote sensing program, known as the Indian Integrated Resources Information Program (IIRIP). The purpose of the IIRIP is twofold: first, make GIS and remote sensing technology available to tribes and BIA personnel; second, transfer these technologies to tribal organizations.

Database development and management functions, technical support, development of simplified user interfaces, remote sensing interpretation, and implementation of equipment directives are performed by approximately 30 GDSC employees for approximately 230 GDSC users. User technical support is also available through BIA field offices, each of which has a designated GIS coordinator. Simplified user interfaces for specialized programs have been developed including the Lightening Display System and the Land Title Mapping System. Quality control is provided for non-BIA-produced data that will be inputted.

The GDSC has standardized on the ARC/INFO family of software produced by Environmental Systems Research Institute (ESRI). GDSC has developed a number of hardware/software configuration options depending on tribal needs and financial resources and based upon GDSC experience. The OE happily relies on this experience to avoid the familiar horror stories related to equipment and software incompatibility.

Based upon GDSC configuration advice, the initial GIS setup will be on the OE's existing Compaq PC using Tektronix Terminal Emulation software (EM4105) and a Multitech modem (MT932BA). The system can use the OE's Hewlett Packard (HP) Deskjet 500, although a significant upgrade, possibly to an HP Paint or HP Excel Paint, is soon expected. Initial startup hardware and software costs are minimal in this configuration. Costs for the above equipment and introductory training are less than $5,000.

GIS Users at Lower Sioux

Initial setup and data loading will be in the OE, and the OE employee will receive introductory training on the system. Because the OE is formed through a cooperative agreement between two tribes, the Upper Sioux and the Lower Sioux, the OE is centrally located between the reservations. The system will probably be relocated to the Lower Sioux Community Center within 1 year. A tribal government employee will receive advanced GIS training and be available for all tribal government departments and businesses.

Funding

In addition to tribal contributions, funding has come through several sources and joint agreements with the tribe and BIA, EPA, and ANA.

Training

The GDSC supplies no-cost training to tribes. The *Geographic Data Service Center 1995 Training Catalog* (no federal document number available) offers eight formal courses repeatedly throughout the year, a 5-week intern program, and a cooperative student program. Courses are held at the GDSC or by request at BIA field offices and tribal locations.

The GDSC also produces the monthly *The Service Center Review* (ISSN 1073-6190), a helpful compilation of current issues, available resources, system bugs, and other items of interest to GDSC users.

Data Collection and Input

Data collection can be divided into three categories: aerial photography, portable global positioning system (GPS) data, and ARC/INFO export files created under contracted studies.

Aerial Photography

Surface features and topography will be obtained using aerial photography reduced to GIS format, then downloaded to the GDSC. Coverages will consist of 62 categories of features on a scale of 1 inch = 100 feet with 2-foot contour intervals.

Global Positioning Systems

Use of a portable Trimble, Inc., GPS Pathfinder Pro XL submeter GPS mapping system purchased with assistance from an ANA grant will allow updating of surface features and addition of nonsurface features as necessary. It will also facilitate input of attribute data.

The GPS will also be used during field work by USGS on the Lower Sioux hydrological mapping project to obtain data that otherwise would not be put into ARC/INFO export file for any reason (i.e., it might not be directly related to the project at hand or outside the agreed upon data to be converted to ARC/INFO export file form but nevertheless is of importance to the OE). The alternative is that this type of information never makes it into the GIS and is lost.

ARC/INFO Export Files

Fortunately, most federal agencies that supply funding to tribes for environmental work are well versed in GIS applications and the need for GIS-ready data. The OE now requires all information and mapping to be delivered as an ARC/INFO export file with data registered to a real world coordinate system. Downloading of this data to the GDSC mainframes allows for direct input of the data. The OE has contemplated, but not acted on, conversion of existing data for the GIS. This is an expensive and time consuming process that must be weighed in comparison with recollecting the data. Ironically, the lack of reservation data therefore becomes a benefit because time consuming and expensive data conversion is unnecessary.

The Intertribal GIS Council

Information gathering, networking, and addressing uniquely tribal problems were some of the accomplishments at the first annual meeting of the Intertribal GIS Council (IGC) held in June 1994. Vendors as well as BIA regional office and GDSC representatives answered questions and presented panels. This annual conference is likely to become a major benefit to the tribe as it continually develops the GIS.

The Future

As the tribal government becomes more familiar with the GIS, its uses, and advantages, recognized governmental needs will likely drive the development of further coverages. The OE also expects to access existing governmental data of importance to the tribe in an effort to expand the GIS database and is actively seeking sources of such information.

Philosophical Caveat

Albert Einstein stated that, “The significant problems we face cannot be solved at the same level of thinking we were at when we created them.” Some assume that GIS is the next level of reasoning in the environmental profession because we can accomplish tasks more quickly, more efficiently, with more variables accounted for, and beyond what we could have hoped to accomplish prior to GIS.

Essentially, what we have gained is speed and the capacity to include additional data, which is not what Einstein was referring to when he spoke of the next level. Wisdom, in the sense of a higher level of understanding, is the necessary ingredient to the solution of current environmental problems; in other words, movement beyond the paradigm that created the problem. GIS may be the tool that pushes the environmental professional to the next level of wisdom by presenting the data and information in a manner that allows the user to stand back and see more clearly on a higher plane. But that level can be found only within the environmental professional himself or herself and not within GIS.

Reach File 3 Hydrologic Network and the Development of GIS Water Quality Tools

Stephen Bevington
Water Quality Section, Division of Environmental Management, North Carolina Department of Environment, Health, and Natural Resources, Raleigh, North Carolina

Introduction

The application of geographic information system (GIS) tools to water quality management is limited by the lack of geographically referenced data describing the surface water environment. Ongoing efforts at the local, state, and federal level are producing a multitude of GIS data coverages describing land use/cover and relevant water quality data files. As these data coverages become available, water quality managers will need to develop new analysis techniques to take advantage of the vast amount of geographically referenced data. A key step in the development of analytical tools for water quality management will be the development and maintenance of a coverage describing the structure and hydrology of surface waters.

Reach File 3 (RF3) is one potential source of surface water maps and topology for the development of a GIS-based water quality analysis tool. This paper describes a pilot project designed to examine the suitability of RF3 as a network system for the collection, integration, and analysis of water quality data.

To be considered an appropriate water quality analysis tool, RF3 should provide the following functions:

- Present a working environment that allows users to explore geographic relationships between surface water features, landmark features, and data coverages.
- Allow users to select specific stream segments, including all points upstream and downstream of a given point.
- Provide tools to assist users in partitioning water quality databases into hydrologically meaningful subsets.

Reach File 3

RF3 is a hydrographic database of the surface waters of the United States. The database contains 3 million river reaches mapped at 1:100,000 scale. The source for RF3 arcs were digital line graphs (DLGs).

Attribute data for RF3 arcs include the major-minor DLG pairs, stream name, water-body type, stream order, and a unique identifying reach number. The unique reach numbers are structured in such a way as to provide a logical hydrologic framework. Reach numbers can be used to sort the database for all reaches in any specified watershed or locate all upstream or downstream reaches.

The U.S. Environmental Protection Agency (EPA) originally designed RF3 as a tabular data set. It evolved into a GIS data coverage, and EPA and the U.S. Geological Survey (USGS) will likely maintain it as a surface water mapping standard. At present, RF3 as a GIS data layer is not widely used for water quality applications.

RF3 Pilot Study: Upper Yadkin River Basin

The Upper Yadkin River basin (USGS h03040101) was selected to test RF3 water quality applications (see Figures 1 and 2). The Upper Yadkin was chosen because of the availability of water quality and stream flow data layers in that area. Also, the Upper Yadkin RF3 file contained arcs depicting lakes and double-line rivers as well as simple stream networks. These two-dimensional water features present interesting complications to network routing and path-finding.

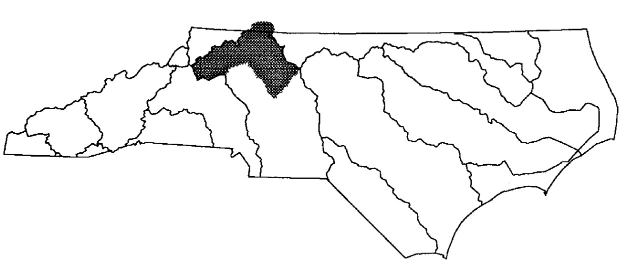


Figure 1. The Upper Yadkin River watershed, North Carolina and Virginia.

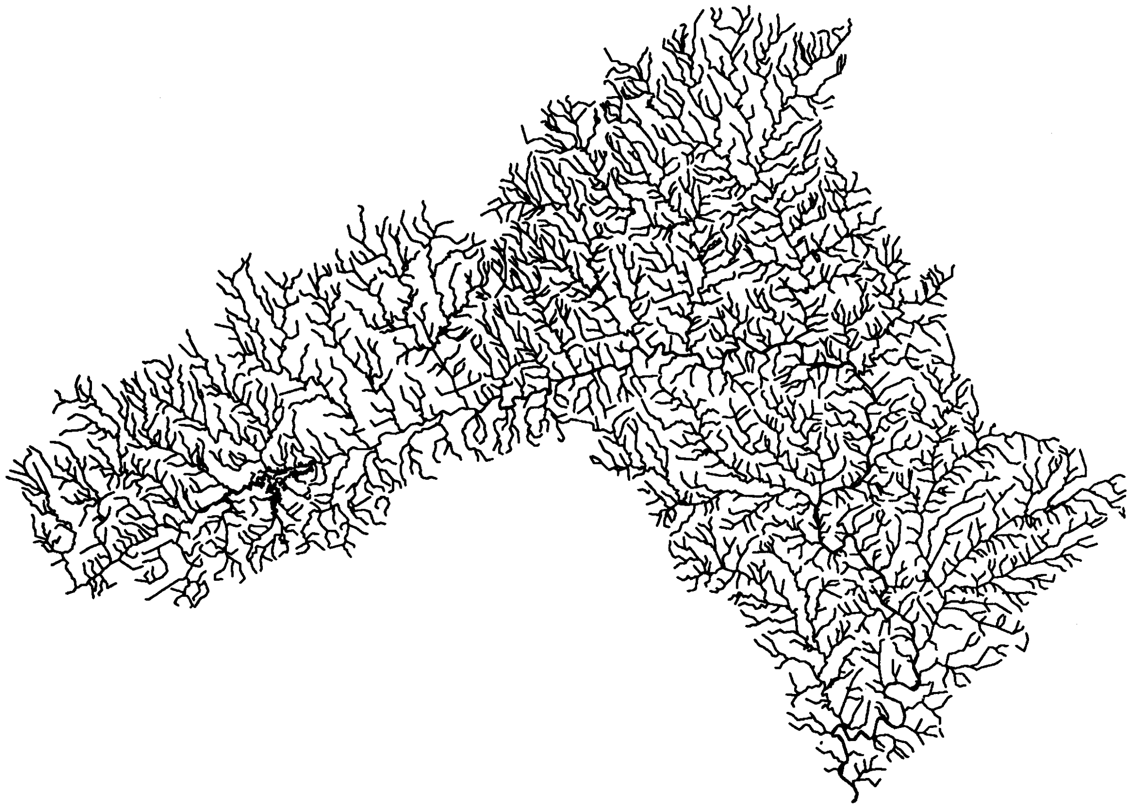


Figure 2 RF3 hydrography for the Upper Yadkin River basin.

Two forms of point source data were used in the study: National Pollutant Discharge Elimination System (NPDES) wastewater discharge points and USGS gages. The NPDES coverage includes data on the permit limits such as daily flow, dissolved oxygen, biochemical oxygen demand (BOD), and ammonia. The USGS gage coverage includes data on several flow statistics for each USGS gage in the basin. Both data layers contain information about the location of the site and stream with which it is associated.

Coverages of counties and cities were also made available for geographic orientation.

Preparing the Network

The original RF3 file received from the USGS had several topological issues that needed to be addressed before RF3 could function as a stream network. First, not all arcs were connected to each other (see Figure 3). The ARC/INFO command TRACE was used to select all connected arcs. This revealed three major blocks of connected arcs and many isolated arcs. The three major blocks were easily connected in ARCEDIT by extending the main tributary links between the blocks. Processing of the isolated arcs was not pursued for this study. Complete processing of arcs for this RF3 basin would not be difficult or time consuming, with the possible exception of the many arcs surrounding the lake. A functional network encompassing a high percentage of the arcs was not difficult to achieve, however.

The second network issue concerned the direction of the arcs. RF3 has all arcs oriented toward the top of the watershed, with the exception of one side of double-line streams. Arcs that make up double-line streams are oriented up one side of the double-line section and down the other (see Figure 3). Clearly, this complicates routing. To allow for accurate downstream routing, arcs on the downward-facing side of the stream were flipped using ARCEDIT. With all arcs in the network facing upstream, most hydrologic routes can be traced. Given the network system alone, upstream routing from double-line streams does not function properly, ignoring all tributaries on one side of the double-line stream.

Double-Line Stream Routing

Many possible solutions exist for the problems caused by double-line streams. Some involve improving the network (e.g., by adding center-line arcs down the middle of double-line streams). This would involve not only adding arcs but establishing conductivity with all tributaries. This option will involve significant topological changes to RF3. To maintain compatibility with other

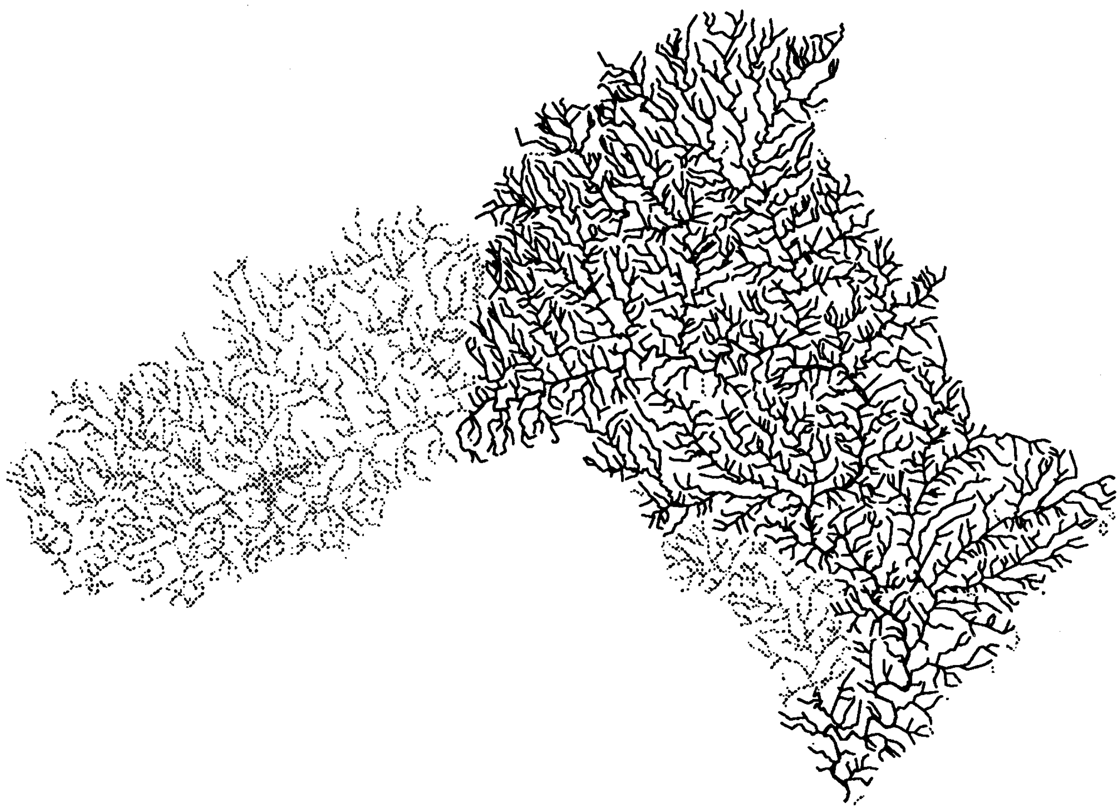


Figure 3 Original conductivity of RF3 hydrograph

RF3 and DLG sources, this option should be considered only as part of a major RF3 upgrade.

At the other end of the technological spectrum, one could simply instruct users to watch for double-line streams and select arcs from both sides of the river. Users may have trouble with this option, however, if they are not working at an appropriate scale to easily differentiate between double- and single-line streams.

A third option is to program an arc macro language (AML) to check for double-line streams and run upstream traces from both sides of the stream. The difficulty in this method is to find the appropriate starting place on both banks. The algorithm developed to do this goes as follows:

- Select stream segment and trace upstream. (Results in incomplete trace.)
- Find the minimum segment and mile of selected double-line streams.
- Unselect all double-line streams below minimum segment and mile.
- Add to selection all non-double-line streams.
- Trace from original point both upstream and downstream. (Results in completed upstream trace.)

Results and Conclusions

AMLs and menus were written that can perform upstream and downstream traces on the RF3 stream network and select data points within 500 feet of the stream. Lists of attributes can be returned to the screen. This system is easy to use and can be used to quickly identify general watersheds and water quality data points. An AML can be used to trace upstream from a double-line stream given only one point on the stream (see Figure 4). The success of these methods suggests that two-dimensional surface water features can be successfully integrated into RF3 water quality analyses.

This system could be further developed to support polygon analysis using the ARC command BUFFER. Other developments could include the procedures to write selected attributes to files and increased flexibility for the screen environment and outputs.

This pilot project demonstrates only a few of the potential applications of RF3 to water quality. Success in this pilot project suggests that RF3 is a potentially valuable water quality analysis tool. It may also be a valuable tool for demonstrating the results of water quality analyses to managers or the public.

Because RF3 will require some processing before network algorithms can be run, it is important to plan for the integration of RF3 into other GIS tools and data coverages.

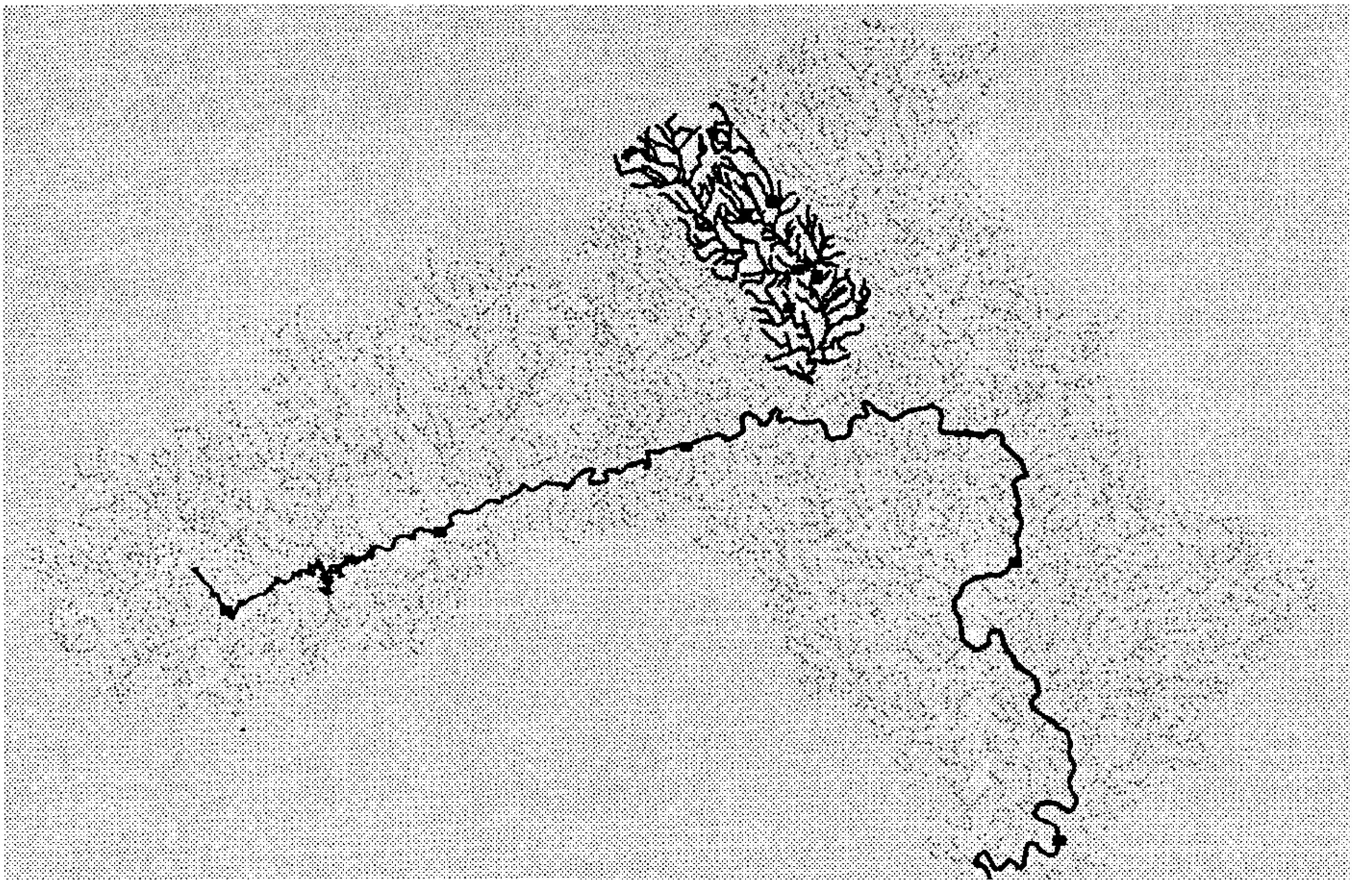


Figure 4   Upstream and downstream traces of RF3 hydrograph

Ongoing efforts to update RF3 may address some of these problems. If RF3 is to be developed into a productive water quality management tool, it is important to proceed in a way that is compatible with ongoing efforts to update RF3 and the development of new data sources.

Assessing the Long-Term Hydrologic Impact of Land Use Change Using a GIS-NPS Model and the World Wide Web*

BUDHENDRA BHADURI[1]
JON HARBOR[2]
BERNIE ENGEL[3]
KYOUNG.J. LIM[3]
DON JONES[3]

---

* This paper covers the contents of two presentations in the Diffuse Source session: "Assessing Long-Term Impact of Land Use Change on Runoff and Non-Point Source Pollution Using a GIS-NPS Model" and "A Web-based GIS Model for Assessing the Long-Term Impact of Land Use Change (L-THIA GIS WWW): Motivation and Development".

[1] Oak Ridge National Laboratory, PO Box 2008, MS 6237, Oak Ridge, TN 37831-6237; Phone: (423) 241 9272; Email: bhaduribl@ornl.gov
[2] Department of Earth & Atmospheric Sciences, Purdue University, West Lafayette, IN 47906-1397; Phone: (765) 494 9610; Email: jharbor@purdue.edu
[3] Department of Agricultural & Biological Engineering, Purdue University, West Lafayette, IN 47907-1146; Phone: (765) 494 1198; Email: engelb@ecn.purdue.edu. kjlim@ecn.purdue.edu, jonesd@ecn.purdue.edu

ABSTRACT:

Assessment of the long-term hydrologic impacts of land use change is important for optimizing management practices to control runoff and non-point source (NPS) pollution associated with watershed development. Land use change, dominated by an increase in urban/impervious areas, can have a significant impact on water resources. Non-point source (NPS) pollution is the leading cause of degraded water quality in the US and urban areas are an important source of NPS pollution. Despite widespread concern over the environmental impacts of land use changes such as urban sprawl, most planners, government agencies and consultants lack access to simple impact-assessment tools that can be used with readily available data. Before investing in sophisticated analyses and customized data collection, it is desirable to be able to run initial screening analyses using data that are already available. In response to this need, we developed a long-term hydrologic impact assessment technique (L-THIA) using the popular Curve Number (CN) method that makes use of basic land use, soils and long-term rainfall data. Initially developed as a spreadsheet application, the technique allows a user to compare the hydrologic impacts of past, present and any future land use change. Consequently, a NPS pollution module was incorporated to develop the L-THIA/NPS model.

Long-term daily rainfall records are used in combination with soils and land use information to calculate average annual runoff and NPS pollution at a watershed scale. Because of the geospatial nature of land use and soils data, and the increasingly widespread use of GIS by planners, government agencies and consultants, the model is linked to a Geographic Information System (GIS) that allows convenient generation and management of model input and output data, and provides advanced visualization of the model results. Manipulation of the land use layer, or provision of multiple land use layers (for different scenarios), allows for rapid and simple comparison of impacts. To increase access to L-THIA, we have begun development of a WWW-accessible version of the method. Using databases housed on our computers, the user can select any location in the US and perform L-THIA/NPS analyses.

In this paper we present applications of the WWW-based L-THIA/NPS and L-THIA/NPS GIS model on the Little Eagle Creek (LEC) watershed near Indianapolis, Indiana. Three historical land use scenarios for 1973, 1984, and 1991 were analyzed to track land use change in the watershed and to assess the impacts of land use change on annual average runoff and NPS pollution from the watershed and its five sub-basins. Comparison of the two methods highlights the effectiveness of the L-THIA approach in assessing the long-term hydrologic impact of urban

sprawl. The L-THIA/NPS GIS model is a powerful tool for identifying environmentally sensitive areas in terms of NPS pollution potential and for evaluating alternative land use scenarios to enhance NPS pollution management. Access to the model via the WWW enhances the usability and effectiveness of the technique significantly. Recommendations can be made to community decision makers, based on this analysis, concerning how development can be controlled within the watershed to minimize the long-term impacts of increased stormwater runoff and NPS pollution for better management of water resources.

INTRODUCTION:

For decision makers, such as land use planners and watershed managers, it is important to assess the effects of land use changes on watershed hydrology. At present numerous hydrologic models are available that focus on event-specific assessment and management of hydrologic impacts of land use change. Traditionally the focus of these urban surface water management models has been on the control of peak discharges from individual, high magnitude storm events that cause flooding. Models such as those developed by the US Army Corps of engineers (HEC-1, 1974), the US Department of Agriculture (TR-20, 1983; TR-55, 1986), and the US Environmental Protection Agency (Huber and Dickinson, 1988) are routinely used in assessing how proposed land use changes will affect runoff quantity. Although hydrologic impact assessment based on individual, high magnitude storm events is an appropriate approach for designing runoff control facilities for reducing local flooding and improving water quality, it is of limited use for attempts to understand the long-term hydrologic impacts of land use change. However, it has been increasingly realized that there is a long-term hydrologic impact associated with land use change, and that this is dominated by runoff generated from frequently occurring, smaller storm events rather than extreme, high magnitude storms (Harbor, 1994; McClinktock et al., 1995).

Realizing the importance of NPS pollution, over the last 25 years, models including SWMM (Huber and Dickinson, 1988), AGNPS (Young et al., 1989), and WEPP (Nearing et al., 1989, Flanagan and Nearing, 1995) have been created with capabilities to assess the impacts of NPS pollution on runoff quality in addition to standard assessments of peak discharges. Because NPS pollution from agricultural areas was originally identified as the major cause of water quality degradation, most NPS pollution models focus on typical agricultural pollutants such a sediment, nutrients (nitrogen and phosphorus), and organic compounds (pesticides and herbicides). However, heavy metal pollution from urban areas has recently been identified as a leading cause of NPS pollution problems (Novotny and Olem, 1994) but estimation of heavy metal pollution from urban areas with existing hydrologic/NPS pollution models is quite limited.

In assessing the long-term hydrologic impacts of land use change planners, developers, and community decision makers usually avoid using the existing hydrologic models because these models are too complex, data intensive, time consuming, expensive, and requires considerable user-expertise (Harbor, 1994). To overcome the limitations of traditional hydrologic models, the Long-Term Hydrologic Impact Assessment (L-THIA) model was developed as a user-friendly tool for long-term runoff estimation (Harbor, 1994). L-THIA is built around the Natural Resources Conservation Service's Curve Number (CN) technique that is the core component of many sophisticated hydrologic models (Williams et al., 1984; Young et al., 1989). Curve numbers or CN values represent surface characteristics of a soil-land use complex. In L-THIA a long-term (typically 30 years) daily precipitation record is used along with soil and land use information to compute daily runoff for estimating annual average runoff. The model was initially developed as a simple spreadsheet application (Harbor, 1994; Bhaduri et al., 1997). Subsequently a C program was developed for the model to facilitate input data handling and model application. The L-THIA model was further expanded to L-THIA/NPS by adding a NPS pollution assessment module. To enhance spatial data management, spatial analyses, and advanced visualization of model results Geographic Information Systems (GIS) have been utilized. L-THIA GIS (Grove, 1997) and L-THIA/NPS GIS have been developed as customized applications of commercial GIS software. Recently, a WWW-based version of the L-THIA/NPS model has been developed. In the WWW-based implementation of the L-THIA/NPS model, the user provides land use and hydrologic soil group information and L-THIA/NPS is run using long-term daily precipitation data queried from an ORACLE database. By determining and comparing the average annual runoff depths and NPS pollutant loads for land use scenarios from different time periods, it is possible to assess the absolute and relative changes in runoff and NPS pollution due to land use change.

METHODOLOGY

Structure of L-THIA and L-THIA/NPS Model

The L-THIA model was originally developed as a preliminary hydrologic impact assessment tool that focused on predicting the percent increase in annual average runoff from a watershed due to some land use change represented by a change in the CN value for the watershed. The model utilizes a lumped parameter design to minimize model complexity and to reduce the expense and time involved in data collection. For a watershed with multiple land use categories and/or sub-watersheds, the model can be applied as a lumped (composite CN) as well as a distributed (distributed CN) approach (Grove, 1997, Grove et al., 1998).

Runoff Calculation:

Daily runoff is calculated using the USDA NRCS Curve Number (CN) method for a daily precipitation data set spanning many years (typically 30 years). The CN method is an empirical set of relationships between rainfall, land use characteristics, and runoff depth. CN values, ranging from 25 to 98, represent land-surface conditions and are a function of land use, hydrologic soil group (or soil permeability), and antecedent moisture condition (USDA SCS, 1986). The basic equations used in the CN method for standard or average conditions are:

$$\boldsymbol{R} = \frac{(\boldsymbol{P} - \mathbf{0.2}\boldsymbol{S})^2}{(\boldsymbol{P} + \mathbf{0.8}\boldsymbol{S})} \qquad R = 0 \text{ for } P \le 0.2\,S \qquad (1)$$

$$\boldsymbol{S} = \left(\frac{\mathbf{1000}}{\boldsymbol{CN}}\right) - \mathbf{10} \qquad (2)$$

where:

R = runoff depth (inches)

P = precipitation depth (inches)

S = potential maximum retention (inches)

CN = Curve Number

Antecedent Moisture Condition (AMC) and CN Variation

The effect of antecedent rainfall and associated soil moisture conditions has long been recognized as a primary source of variability in runoff determination. To account for this, the Natural Resources Conservation Services (NRCS) introduced the concept of an antecedent moisture condition (AMC), also referred to as antecedent runoff condition (ARC).

Three AMCs are defined as a step function of 5-day antecedent rainfall, and an AMC remains constant for the specific range of antecedent rainfall values. Definitions of growing and dormant seasons are not easily available and to keep calculations simple and consistent, growing and dormant seasons were assumed to begin on April 15 and on October 15 of any year, respectively. CN values for AMC 1 and 3 are determined by the following relationship as described in NEH-4 (USDA SCS, 1985).

$$CN_1 = \frac{4.2CN_2}{10 - 0.058CN_2} \quad \text{and} \quad CN_3 = \frac{23CN_2}{10 + 0.13CN_2}$$

[where, CN_1, CN_2, and CN_3 represent CN values for AMC 1, 2, and 3 respectively.]

Runoff analyses were performed with the CN values for AMC 1, 2, and 3 where AMC is a step function of 5-day antecedent rainfall (Table 1).

	5-day Antecedent Rainfall (mm)	
AMC	Dormant Season	Growing Season
1	< 13	< 36
2	13 -38	36 - 53
3	> 28	> 53

Table 1: Criteria for determination of Antecedent Moisture Conditions (SCS, 1972).

NPS Pollution Calculation:

L-THIA/NPS GIS: Pollutant Build-up and Washoff

The most common urban NPS pollutant estimation technique in current deterministic water quality models including STORM and SWMM is the pollutant “buildup-washoff” function (Huber, 1986; Barbé et al., 1996). “Buildup” refers to all dry-weather processes that lead to accumulation of solids and associated pollutants on a watershed surface which are “washed off” during subsequent storm events. In developing a NPS pollution sub-model for L-THIA, it was assumed that pollutants accumulate on a land surface as a linear function of time.

$$M_i = (\text{number of days}) \times (L_i)$$

[L_i = accumulation rate for pollutant i (mass/area/day);
M_i = Ultimate pollutant accumulation (mass/area);]

For this study, daily accumulation rates of solid particles (dust and dirt) for urban land uses (low and high density residential, industrial, and commercial) were adopted from the SWMM manual. Daily dust and dirt accumulation values are reported as a function of curb (road) density, and thus road densities for the urban land uses were required to produce dust and dirt accumulation values as mass/area. Although road density values for different urban areas have been reported in the literature, for this study values of road density for Tulsa, Oklahoma (Heany et al., 1977) were chosen as representative of the Indianapolis, Indiana area where the model was applied. Daily build up values of pollutants on non-urban land uses (agricultural, grass/pasture, and forest) could not be found in literature and were not included in the daily simulations of NPS pollution analyses. However, annual average loading rates for non-urban land uses were taken from literature and used to calculate NPS pollution in the GIS analysis. The NPS pollutant loading values are reported in Table 2.

Pollutant	Annual average loading rate (kg/ha/year)					
	Agricultural	Grass/Pasture	Forest	Low-density Residential	High-density Residential	Commercial
Total N	26.00	6.20	6.20	0.86	1.92	2.30
Total P	1.05	0.50	0.50	0.09	0.20	0.33
Lead	0.10	0.10	0.10	2.90	6.93	12.80
Copper	0.02	0.02	0.02	0.17	0.25	0.52
Zinc	0.08	0.08	0.08	0.58	0.98	3.78

Table 2. Annual average pollutant loading values used in L-THIA/NPS GIS simulations.

For the washoff function, a non-linear washoff equation was used. The washoff relationship is an exponential function of the runoff depth. This approach has been successfully used in numerous studies (Haith and Shoemaker, 1987; Dikshit and Loucks, 1996) and was utilized in the NPS simulation for the L-THIA model because daily runoff depths are calculated in the runoff sub-model which then can be used in the washoff function. The washoff function is expressed as:

$$w_{k,t} = 1 - \exp(-1.81\ Q_{kt})$$

where:

$w_{k,t}$ = fraction of the pollutant mass removed from the land use k on day t;

Q_{kt} = runoff from land use k on day t (cm);

WWW-based L-THIA/NPS: Event Mean Concentration (EMC)

In the Web-based version of L-THIA/NPS, Event Mean Concentration (EMC) data were introduced to predict NPS pollutants for non-urban areas as well as urban areas (Baird and

Jennings, 1996). The EMC data used were compiled by the Texas Natural Resource Conservation Commission (Baird and Jennings, 1996). Numerous literature and existing water quality data were reviewed by Baird and Jennings (1996) with respect to eight categories of land use and several parameters. Land use categories defined were (1) industrial; (2) transportation; (3) commercial; (4) residential; (5) agricultural cropland (dry land and irrigated); (6) range land; (7) undeveloped/open; and (8) marinas. The total pollutant load for a NPS pollutant divided by runoff volume during a runoff event yielded the Event Mean Concentration for that pollutant. EMCs should be reliable for determining average concentrations and calculating constituent loads (Table 3).

NPS Pollutant	Land use classification						
	Residen -tial	Comm -ercial	Indus - trial	Transt- ion	Mixed	Agricu- ltural	Range
Total Nitrogen (mg/L)	1.82	1.34	1.26	1.86	1.57	4.4	0.7
Total Phosphorus (mg/L)	0.57	0.32	0.28	0.22	0.35	1.3	0.01
Total Lead (g/L)	9	13	15	11	12	1.5	5
Total Copper (g/L)	15	14.5	15	11	13.9	1.5	10
Total Zinc (g/L)	80	180	245	60	141	16	6

Table 3. Event Mean Concentration for each land use classification

(Baird and Jennings, 1996)

L-THIA/NPS GIS Setup

The L-THIA model has been linked with Arc/INFO® GIS software as a GIS application (Grove, 1997). The ArcView® GIS software was chosen for L-THIA/NPS GIS application because ArcView® is the dominant desktop GIS, and it has a friendlier graphical user interface than Arc/INFO®. The GIS application is implemented through a linked-model approach that utilizes both the graphical and spatial data handling capabilities of a GIS as well as the speed and flexibility offered by a standard executable program. The required input data is initially selected in the GIS before the L-THIA/NPS executable is called by the GIS. The executable calculates the annual average runoff depths for all land uses and annual average dust and dirt amounts (kg/km^2) for all CN values for urban land uses (low density residential, high density residential, industrial, and commercial). These calculations are based on daily rainfall data spanning many years. The output file created by the L-THIA/NPS executable is then read back into the GIS and used to produce final results.

WWW-Based L-THIA/NPS Setup

A user-friendly L-THIA WWW interface was developed using Java/Java Script, HTML, and CGI scripts (http://pasture.ecn.purdue.edu/~sprawl/lthia2). This interface provides easy access to the model and potentially improves understanding of the results through graphical representation. Figure 1 shows the L-THIA/NPS WWW interface.

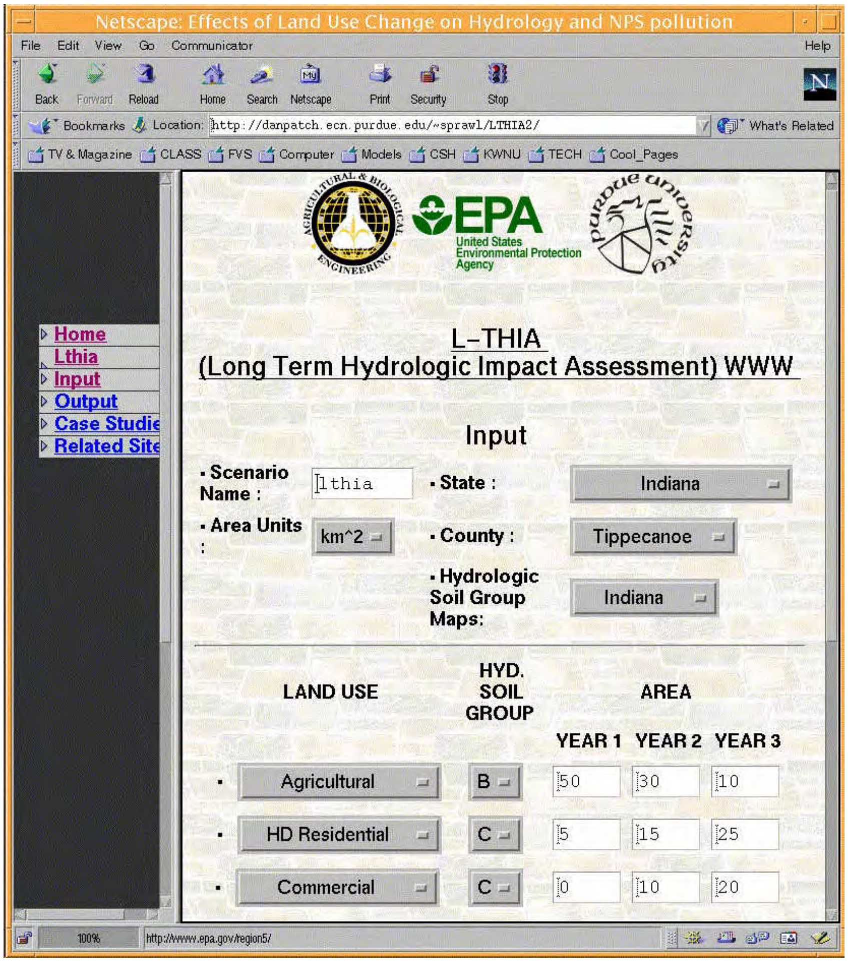


Figure 1. LTHIA/NPS WWW Interface.

Depending on the location the user selects, weather data for the nearest weather station are queried from the database and reformatted for the L-THIA run. The user selects one of the eight land use classifications, hydrologic soil group information and provides the area for this combination for each time step the L-THIA/NPS WWW system is to be run. Tables, bar charts, and pie charts for runoff and NPS pollution are generated on the fly for display in the user's WWW browser. The tabular output provides all information the user provided in the input interface, the Curve Number, runoff depth, and runoff volume for each time step. Bar graphs provide runoff depth, runoff volume, total volume, average runoff depth, and NPS pollution information. Pie charts provide land use and runoff volume for each time of interest. LTHIA/NPS WWW has several advantages over the traditional model and decision support system approach: 1) It is accessible through the Internet using only a WWW browser, 2) Database and GIS data are maintained at a single location, 3) All model users access the same version of the model, and 4) All data are verified by the model maintainer so errors due to input data can be minimized.

STUDY AREA

The L-THIA/NPS model was applied to the Little Eagle Creek (LEC) watershed, a rapidly urbanizing watershed in the northwest section of Indianapolis, Indiana and its suburbs. The LEC watershed is 70.5 km^2 in size and consists of five smaller sub-basins (Figure 2). This watershed has experienced extensive urbanization over the past three decades. Land uses ranging from non-urban natural grass and forested areas and agricultural areas to typical urban residential, commercial, and industrial categories exist in the LEC watershed.

As part of a long-term hydrologic impact assessment study (Grove, 1997), digital land use data were generated from LANDSAT satellite imagery (80m resolution Landsat Multi-Spectral Scanner imagery) for 1973, 1984, and 1991 and these three images represented temporal changes in land use in the watershed. In this study, these land use coverages along with the Soil Survey Geographic (SSURGO) soils data (1:20,000) were used to analyze the long-term impact of land use change on runoff and non-point source pollution. Only hydrologic soil groups B and C are present in the watershed. The watershed and sub-watershed boundaries were delineated from a Digital Elevation Model (DEM) using the Arc/INFO GRID module (Grove, 1997). Curve Numbers ranged from approximately 60 to 97 for all sub-basins in the watershed. Six land use categories were delineated using ERDAS Imagine software and were used in L-THIA/NPS GIS simulations. These areas of these land use categories and hydrologic soil B and C groups were used in the WWW L-THIA/NPS simulations.

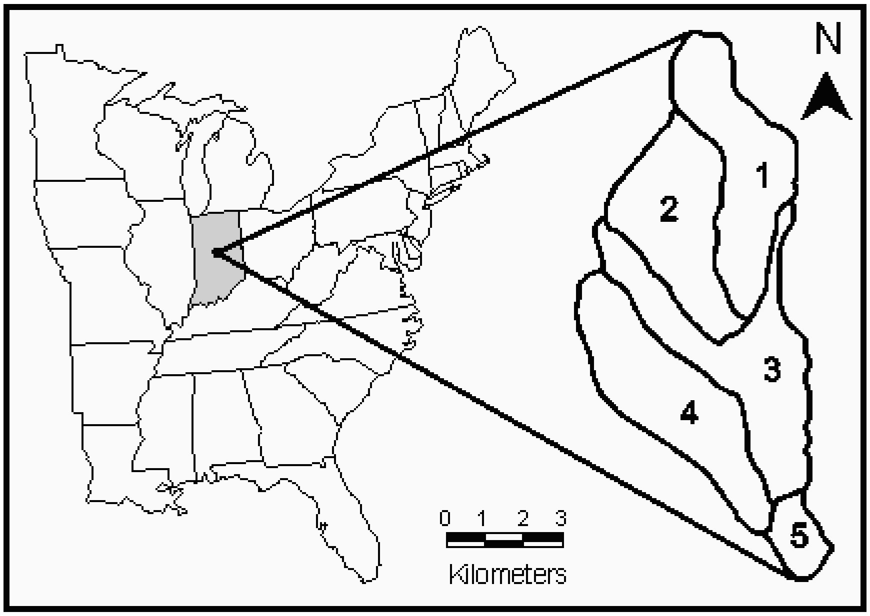


Figure 2. Location of the Little Eagle Creek watershed.

RESULTS AND DISCUSSION

Land Use Change

There is a significant increase in urban land uses between 1973 and 1991 with the majority of the changes taking place between 1973 and 1984 (Table 4). Grouping agricultural, forest, and grass/pasture as non-urban and low density residential, high density residential/industrial, and commercial as urban land uses, 49.3%, 63.5%, and 68.1% area of LEC watershed was urban in 1973, 1984, and 1991 respectively. Thus, there was a 14.2% increase in urban land uses between 1973 and 1984 and a 4.6% increase in urban areas between 1984 and 1991. The increase in urban land uses is not uniformly reflected in all the urban land use categories.

Land Use	**Area (km^2)**			**% Change in Individual Category**		
	1973	**1984**	**1991**	**1973-1984**	**1984-1991**	**1973-1991**
Agricultural	10.82	10.21	9.23	-5.66%	-9.58%	-14.70%
Commercial	5.82	10.56	11.31	81.40%	7.12%	94.32%
Forest	13.74	5.72	5.14	-58.37%	-10.19%	-62.61%
Grass/Pasture	10.90	9.64	7.76	-11.62%	-19.47%	-28.83%
HD Residential / Industrial	8.12	19.25	21.44	137.21%	11.34%	164.10%
LD Residential	20.83	14.91	15.30	-28.42%	2.63%	-26.53%
Water	0.27	0.21	0.32	-21.84%	52.73%	19.38%

Table 4. Land use distributions in Little Eagle Creek watershed for 1973, 1984, and 1991.

For individual land use categories, high density residential and commercial areas show tremendous increase in the watershed while low density residential areas show a 28.4% decrease between 1973 and 1984 and a 2.6% increase between 1984 and 1991. The initial decrease in low density residential areas is possible conversion of low density residential to high density residential areas. The increase in urban areas is followed by an equivalent decrease in non-urban areas. However, all the non-urban land uses decrease at the same rate. Forested areas show the greatest loss with a 62.6% decrease, followed by grass/pasture with a 28.8% decrease between 1973 and 1991. Agricultural areas show minimum change (14.7% decrease) during the same time period.

Impact of Urbanization on Annual Average Runoff and NPS Pollution

L-THIA/NPS analyses were performed to assess the impact of land use change on average annual runoff and NPS pollution for the LEC watershed. There are significant changes in average annual runoff volumes and NPS pollution loads from the LEC watershed as a result of land use change. The results from L-THIA/NPS GIS and L-THIA/NPS web-versions are presented in Table 5 and Table 6 respectively. However, changes in runoff volume or NPS pollution do not have a simple or linear relationship with land use change.

Change in runoff and NPS pollution using L-THIA/NPS GIS simulation:						
Pollutant	Year			% Change		
	1973	1984	1991	1973 to 1984	1984 to 1991	1973 to 1991
Runoff (m^3)	2547736	4255457	4581833	67.03%	7.67%	79.84%
Nitrogen (kg)	47848.90	43183.45	39722.04	-9.75%	-8.02%	-16.98%
Phosphorus (kg)	2887.40	2682.04	2526.83	-7.11%	-5.79%	-12.49%
Lead (kg)	18608.45	30374.48	32855.31	63.23%	8.17%	76.56%
Copper (kg)	888.21	1296.23	1387.92	45.94%	7.07%	56.26%
Zinc (kg)	4313.06	6756.24	7238.03	56.65%	7.13%	67.82%

Table 5. Average annual runoff volume and NPS pollution from LEC watershed using L-THIA/NPS GIS that uses daily pollutant build-up and washoff functions for pollution calculation.

Change in Runoff and NPS pollution using WWW L-THIA/NPS simulation:						
Pollutant	Year			% Change		
	1973	1984	1991	1973 to 1984	1984 to 1991	1973 to 1991
Runoff (m^3)	5012456	7926728	8379366	58.14%	5.71%	67.17%
Nitrogen (kg)	8599	13051	13684	51.77%	4.85%	59.13%
Phosphorus (kg)	2307	3530	3713	53.01%	5.18%	60.94%
Lead (kg)	46	81	87	76.09%	7.41%	89.13%
Copper (kg)	50	89	95	78.00%	6.74%	90.00%
Zinc (kg)	581	1032	1104	77.62%	6.98%	90.02%

Table 6. Average annual runoff volume and NPS pollution from LEC watershed using web-based L-THIA/NPS that uses Event Mean Concentrations (EMC) for pollution calculation.

The annual average runoff volumes predicted by L-THIA/NPS GIS are approximately half of those predicted by the web-version of the model. This is primarily because two different sets of daily precipitation data were used for the two simulations and the one used for the web version had several large storm events that produced significantly more runoff. However, considering the relative change in runoff volume, very similar results were obtained from both simulations. The amounts of urban or impervious areas dominantly control the volume of runoff produced from a watershed. For example, in LEC watershed, 87% of the total runoff volume (81% with web-based L-THIA/NPS) was produced from urban areas that occupied only 49% of the total watershed area in 1973. In 1984 and 1991, urban areas occupied less than 70% of the total watershed area but contributed over 93% of the annual average runoff volume (over 90% with web-based L-THIA/NPS). Percent increase in average annual runoff volume is greater between 1973 and 1984 than between 1984 and 1991 because a greater percentage of non-urban land use is changed to urban (more impervious) land use during the former time interval (Figure 3).

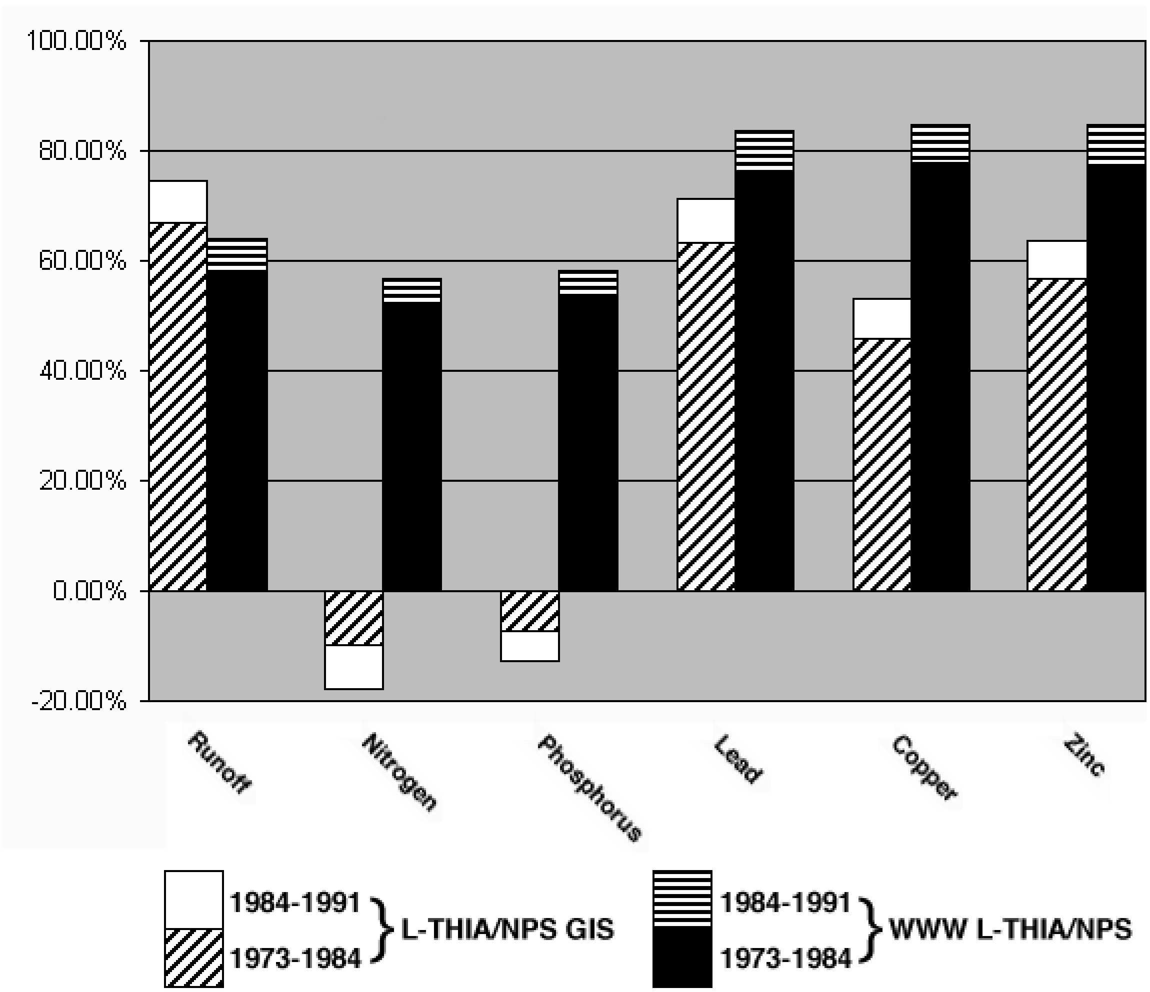


Figure 3. Changes in annual average runoff and NPS pollution from the Little Eagle Creek watershed.

For the NPS pollutants, the relative change in annual average NPS pollution from LEC watershed is not only controlled by the nature of land use change but also by the nature of the pollutants. The time period between 1973 and 1984 experienced a much greater amount of urbanization than the time period between 1984 and 1991. This pattern of land use change is also reflected in changes in average annual runoff volume and metal pollution from the LEC watershed. However, total pollutant loads predicted by web-based L-THIA/NPS are roughly higher by an order of magnitude than those predicted by the L-THIA/NPS GIS model. This difference can be attributed primarily to the different methods of pollution calculations by the two simulations and also the different concentration values of the pollutants used. Using EMC values in the web-version of the model, two different days with the same amount of runoff will produce the same pollutant loads. In the GIS version, that uses pollutant build-up and washoff functions, two similar runoff events can produce significantly different pollutant loads depending upon masses of pollutants that accumulated before those two runoff events. Moreover, Bhaduri (1998) showed that more than 90% of the days in the study area are dry (AMC 1), and thus before any runoff event there will be a significant amount of pollutant accumulated on the watershed.

One significant difference between the two approaches of pollution calculations can be observed in the predicted changes in nutrient pollution. L-THIA/NPS GIS predicts decreasing nutrient pollution with increasing urbanization in the watershed. Nitrogen and phosphorus are dominantly produced from agricultural areas. Moreover, the other non-urban land uses (forest and grass/pasture) show significant decrease between 1973 and 1984. Thus, this small change is nutrient loading between 1973 and 1984 is most plausibly related to the small reduction in agricultural area in the watershed. On the contrary, the web-version of L-THIA/NPS predicts changes in nitrogen and phosphorus loads that conform to the increasing trend in urbanization. Because nitrogen and phosphorus are typically identified as non-urban pollutants, it might be assumed that conversion of non-urban land uses to urban areas would significantly reduce nutrient pollution from a watershed. Our analyses on LEC watershed indicate that, between 1973 and 1991, a conversion of 19% areas from non-urban to urban land uses results in annual average nitrogen and phosphorus loads being increased by about 60%. This is primarily because there is only a small reduction of agricultural area and a large increase in urban areas in the watershed. Although urban areas produce nutrients at a much lower rate than non-urban areas, but increases in urban land uses produce runoff at a significantly higher rate and thus, the web-version predicts very high nutrient loads.

Heavy metals, such as lead, copper, and zinc, are considered “urban” pollutants because urban land uses contribute a major portion of the metal pollution from a watershed. For the LEC watershed using L-THIA/NPS GIS, we found that only 49% of the area had urban land uses but they contributed 98% of total lead load, 92% of total copper load, and 93% of the total zinc load from the watershed in 1973. However, for individual metal pollutants, this 18% increase in urban areas (or an equivalent decrease in non-urban areas) between 1973 and 1991 results in 76.5%, 56.2%, 67.8% increase in lead, copper, and zinc loads from the watershed respectively. Predictions of changes in metal pollution from web-based L-THIA/NPS simulation were similar to those from the GIS version (Figure 3).

CONCLUSIONS

Assessment of the long-term hydrologic impacts of land use change is important for optimizing management practices to control runoff and non-point source (NPS) pollution from urban sprawl. The L-THIA/NPS model uses the popular curve number technique and empirical relationships between land uses and pollutant accumulation and wash off processes to estimate the relative impacts of land use change on annual average runoff and NPS pollution. L-THIA/NPS uses readily available data to overcome the difficulties of long-term modeling by existing hydrologic models because of their complexity and extensive input data requirements. Moreover, most traditional hydrologic/NPS pollution models do not emphasize the changes in loads of typical urban pollutants such as heavy metals, which can be addressed by L-THIA/NPS. The model is linked to a GIS to enhance input data generation, data management, and advanced visualization of model results. The GIS version computes NPS pollution using daily pollutant build-up and washoff functions. A World Wide Web based version of the model has also been developed that provides easy access to the model through the Internet. This web-based version of the model uses Event Mean Concentrations (EMC) of pollutants for predicting NPS pollution.

L-THIA/NPS was applied to the Little Eagle Creek (LEC) watershed, an urbanizing watershed near Indianapolis, Indiana, to provide estimates of changes in annual average runoff volumes and NPS pollution loads for three time periods: 1973, 1984, and 1991. Increases in urban land uses were much higher between 1973 and 1984 than between 1984 and 1991. Non-urban land uses, particularly agricultural areas, are the dominant sources of nutrient (nitrogen and phosphorus) pollution but the majority of the metal pollution is contributed from urban areas. Overall, increasing urbanization resulted in increases in annual average runoff volume and metal loads. The L-THIA/NPS GIS simulations predicted decreases in nitrogen and phosphorus loads from the LEC watershed. However, the web-based version of the model indicated increases in nutrient pollution with increasing urbanization. This difference can be explained by the two different methods of pollution calculations by the GIS and web-based versions of the model. This difference in pollution calculation is also reflected in the absolute values of pollution predicted by the two versions. However, considering relative change in runoff and NPS pollution from urbanization, both simulations indicate a very similar trend and direction of changes in NPS pollutants for the Little Eagle Creek watershed.

L-THIA/NPS GIS is a simple and user-friendly model that makes it attractive for applications to other watersheds. Although L-THIA/NPS GIS is designed to run with easily available data, such data is often not readily available for most of the watersheds. Thus, compilation of model input data sets through field experiments for a variety of watersheds characterized by different geography, climate, and land uses will greatly enhance model applications and performance in a wider range of watersheds. These field-measured data sets should be used to calibrate the L-THIA/NPS model and validate the model predictions. Future work should also explore the sensitivity of the L-THIA/NPS model to the spatial and temporal scales of input data. Work in progress is aimed at allowing a user to access a modified web-based L-THIA/NPS model that will take advantage of GIS functionality in the analysis. In this modified web-version, the users will not only be able to access the model through a web-browser, but will also be able to select or define a watershed using system-supplied maps, and then run L-THIA/NPS analyses run using land use, soils and climate databases stored on our server. The user will then be able to manipulate the GIS land use data in the browser environment or on a remote computer, and run multiple L-THIA analyses to compare hydrologic impacts from different land use scenarios.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Darrell Leap and Dr. Darrell Norton of Purdue University for their valuable suggestions, comments, and review of the manuscript. Daniel Pack of Oak Ridge National Laboratory provided valuable help with some of the graphics. Funding for this work was provided in part by the Purdue Research Foundation and the United States Environmental Protection Agency.

REFERENCES:

Baird, C., and M. Jennings. 1996. Characterization of Nonpoint Sources and Loadings to the Corpus Christi Bay National Estuary Program Study Area, Texas Natural Resource Conservation Commission.

Barbé, D. E., J. F. Cruise, and X. Mo. 1996. Modeling the buildup and washoff of pollutants on urban watersheds. Water Resources Bulletin, Vol. 32, No. 3, p. 511-519.

Bhaduri, B. 1998. A geographic information system based model of the long-term impact of land use change on non-point source pollution at a watershed scale. Unpublished Ph.D. dissertation, Purdue University, West Lafayette, Indiana 47907-1397.

Bhaduri, B., M. Grove, C. Lowry, and J. Harbor. 1997. Assessing the long-term hydrologic impact of land use change. Journal of the American Water Works Association. Vol. 89, No. 11, pp. 94-106.

Flanagan, D. C. and Nearing, M. A., 1995. USDA-Water Erosion Prediction Project: Hillslope profile and watershed model documentation. NSERL Report No. 10. West Lafayette, Indiana: National Soil Erosion Research laboratory, USDA-ARS.

Grove, M. 1997. Development and application of a GIS-based model for assessing the long-term hydrologic impacts of land-use change. Unpublished MS Thesis, Purdue University, West Lafayette, IN 47907-1397.

Grove M. and J. Harbor, and B. Engel. 1998. Composite versus distributed curve numbers: effects on estimates of storm runoff depths. Journal of the American Water Resources Association. Vol. 34, No. 5, pp. 1015-1023.

Haith, D. A. and L. L. Shoemaker. 1987. Generalized watershed loading functions for stream flow nutrients. Water Resources Bulletin, Vol. 23, No. 3, pp. 471-478.

Harbor, J. 1994. A practical method for estimating the impact of land use change on surface runoff, groundwater recharge and wetland hydrology. Journal of American Planning Association*,* Vol. 60, pp.91-104.

Heany, J. P., W. C. Huber, M. A. Medina, Jr., M. P. Murphy, S. J. Nix, and S. M. Haasan. 1977. Nationwide evaluation of combined sewer overflows and urban stormwater discharges – Vol. II: Cost assessment and impacts. EPA-600/2-77-064b (NTIS PB-266005), Environmental Protection Agency, Cincinnati, Ohio.

Huber, W. 1986. Deterministic modeling of urban runoff quality. In Torno, H. C., Marsalek, J, and Desbordes, M. (Eds.) *Urban Runoff Quality*. Springer-Verlag. 893 p.

Huber, W. C., and R. E. Dickinson. 1988. Storm Water Management Model, Version 4: User's Manual. US Environmental Protection Agency, Report No. EPA-600/3-88-001a. 569 pp.

McClintock, K., J. Harbor, and T. Wilson. 1995. Assessing the hydrologic impact of land use change in wetland watersheds, a case study from northern Ohio, USA. In: McGregor, D. and Thompson, (eds.) *Geomorphology and Land Management in a Changing Environment*, pp.107-119.

Nearing, M. A., Foster, G. R., Lane, L. J., and Finkner, S. C., (1989). A process-based soil erosion model for USDA-water erosion prediction project technology. Transactions of the ASAE, Vol. 32, No. 6, pp. 1587-1593.

Novotny, V. and H. Olem. 1994. *Water Quality*. Van Nostrand Reinold. New York. 1054 pp.

U.S. Army Corps of Engineers, 1985. Flood hydrograph package HEC-1. The Hydrologic Engineering Center, Davis, CA.

U.S. Department of Agriculture, Soil Conservation Service, 1983. Computer programs for project formulation - hydrology. Technical Release 20.

U.S. Department of Agriculture, Soil Conservation Service, 1986. Urban hydrology for small watersheds. Technical Release 55.

USEPA, 1983. Results of the National Urban Runoff Program. Vol. 1. USEPA, Washington, D.C.

Williams, J. R., Jones, C. A., and Dyke, P. T., 1984. The EPIC model and its application. Proc. ICRISAT-IBSNAT-SYSS Symposium on Minimum Data Sets for Agrotechnology Transfer, March 1983, Hydrabad, India, p. 111-121.

Young, R. A., C. A. Onstad, D. D. Bosch, and W. P. Anderson. 1989. AGNPS: A non-point source pollution model for evaluating agricultural watersheds. Journal of Soil and Water Conservation Society, Vol. 44, No. 2, p. 164-172.

Ecological Land Units, GIS, and Remote Sensing: Gap Analysis in the Central Appalachians

Ree Brannon, Charles B. Yuill, and Sue A. Perry
West Virginia University, Morgantown, West Virginia

Abstract

The gap analysis team in West Virginia is assessing the state's natural communities as part of a nationwide, comprehensive planning effort. Underrepresented or unrepresented habitats represent gaps in the present network of conservation lands and conservation activities. After identifying these gaps, we can assess whether our current management direction will maintain natural diversity and will prevent additional species from being classified as threatened or endangered.

The relationship between vegetation and ecological variables serves as the basis for classifying ecological land units. To characterize the ecological land units, many layers of physical data can be integrated in a geographic information system (GIS). Satellite imagery and videography map existing conditions over the state. The existing vegetation is classified to reflect physiognomic and floristic elements to correlate with vertebrate and butterfly habitat requirements. This correlation of vegetation and wildlife habitat creates mappable habitat types. Analysis of these habitat types with land-ownership data indicates where the species-rich areas occur in the landscape and whether the most species-rich areas are protected.

Introduction

To respond to the urgency of habitat loss and its effect on species diversity, scientists must implement a methodology for rapid assessment and documentation of natural communities at a scale pertinent for regional management activities.

Geographic information systems (GIS) and remote sensing support the development of ecological land classifications over large regions. GIS-based mapping of ecological land classes allows users to combine and display environmental variables for spatial modeling and refinement of ecological land units (1).

The Gap Analysis Project is a comprehensive planning effort for land conservation in the United States. The objective of the Gap Analysis Project is to identify species, species-rich areas, and vegetation types underrepresented or unrepresented in existing biodiversity management areas. Unprotected communities are the gaps in the conservation strategy. The Gap Analysis Project is not merely identifying communities with the largest number of species; its ultimate goal is to identify clusters of habitats that link the greatest variety of unique species.

Local areas with considerable diversity of habitat or topography usually have richer faunas and floras (2). Nature reserves, which incorporate a variety of habitats, may be the best guarantee for long-term protection of biodiversity. By protecting species-rich regions, we can reduce the enormous financial and scientific resources needed to recover species on the brink of extinction.

The West Virginia Gap Analysis Project began in 1991. The objective is to map existing vegetation and to use that as the foundation to model potential distribution of vertebrate and butterfly species. High cost precludes intensive field inventory and monitoring of wildlife. Therefore, habitat modeling is critical to predict wildlife species composition and potential ranges over the various landscape types of West Virginia. Lastly, the vegetation map will provide a record of the existing habitat to use in monitoring changes due to human activities and natural disturbances.

Pilot Study Area

Distribution of wildlife and plant communities will be modeled for the entire state. Initially, we will focus on a smaller pilot study area. This region includes approximately 50,000 hectares in the central Appalachian Mountains and spans several physiographic provinces and vegetative communities. Generally, soils in the pilot study area are of two kinds: acidic soils that develop a clay horizon from extensive leaching over time and younger soils that are found on steep slopes and where environmental conditions, such as cold climate, limit soil development (3, 4).

The vegetation types include spruce-fir, oak-pine, high elevation bogs, northern hardwoods, Appalachian oak, mixed-mesophytic, open heath barrens, and grass balds. The mixed-mesophytic, the most diverse in West Virginia, lies primarily west of this area, but localized stands do occur in the lower elevations. Cover types within the Appalachian oak and mixed-mesophytic types are not discrete and will be difficult to delineate.

The pilot study area includes a variety of land uses, such as residential, commercial, industrial, mining, and agriculture. Portions of the Monongahela National Forest in this area are the Fernow Experimental Forest, and the Otter Creek and Laurel Forks Wilderness Areas.

Methods

The following discussion describes the methods formulated and data compiled for the West Virginia Gap Analysis Project.

Describe Ecological Land Units With Existing Vegetation

Davis and Dozier (1) note that a landscape can be partitioned by ecological variables, which contributes to an ecological land classification. This process is applied frequently to analysis and mapping of natural resources. Davis and Dozier classified vegetation in California based on the documented associations of vegetation with terrain variables. They based this approach on the assumptions that the arrangement of natural landscape features is spatially ordered by an ecological interdependence among terrain variables and that actual vegetation is a reliable indicator of these ecological conditions. Similar documentation exists for the distribution of vegetation types in West Virginia, and the gap analysis team is proceeding along a similar course.

West Virginia lies in two major provinces, the Eastern Broadleaf Forest and the Central Appalachian Broadleaf Forest-Coniferous Forest-Meadow Provinces (5). Within these provinces are several broad vegetation types. The gradient diagram in Figure 1 (6) illustrates the range of these types. The vertical axis represents elevation in feet. Three vegetation types emerge distinctly along the elevation gradient. The horizontal axis is not quite as explicit. This gradient spans moist, protected slopes to dry, exposed ridgetops, and the vegetation types are much less distinct. Drier oak and pine types occur almost exclusively on exposed ridgetops. The vegetation types along the horizontal axis are the mixed mesophytic forest association of the Allegheny and Cumberland Mountains and can have 20 to 25 overstory and understory species per hectare in North America (7).

The distribution of the vegetation along gradients such as elevation and soil moisture lends itself to a GIS analysis. Physical data such as elevation and soil moisture regime can be incorporated into a GIS. These layers of information can be manipulated graphically or mathematically to model the spatial distribution of vegetation types or to provide useful ancillary data for classification of satellite imagery (see Figure 2). Much of this data is digital and can be used to substantially reduce the time required to develop a database. To standardize output, members of the national gap team have specified ARC/INFO as the software to generate final products.

Classify Satellite Imagery To Create a Map of Current Distribution

Remote sensing provides an effective means to classify forests, and the Gap Analysis Project has successfully used it in the western United States (1, 8, 9).

The Gap Analysis Project is using Landsat Thematic Mapper imagery in all states to standardize the baseline information. The hypothesis is that the spectral data from the imagery is related to the distribution of the ecological land units and land use across the landscape. The data include all spectral bands, except thermal, for the entire state. The West Virginia project is using two seasons of data, spring and fall. Temporal changes, which record phenologic variation in the deciduous species, enhance classification accuracy. The spectral resolution is 30-meter pixels. This is equivalent to approximately 1/6 hectare (1/2 acre). Our final product will be a series of 1:100,000 maps. The minimum mapping unit is 40 acres.

The mountainous terrain in the Appalachian Mountains offers disadvantages and advantages for using remotely sensed data. Irregular topography can cause inconsistencies in the spectral data that diminish the classification accuracy. Similar cover types may have different spectral signatures; for instance, if one stand is in sunlight and the other is shaded. Also, phenology can vary due to microclimatic influences. Conversely, topographic features influence the distribution of vegetation types, and ancillary data, such as digital elevation models (DEMs), enhance classification results of the imagery. The West Virginia gap analysis team selected the following strategy for image classification.

1. Stratify the imagery using ecological units based on a hierarchical scheme. Bauer et al. (10) found that an initial stratification of physiographic regions was necessary to reduce the effect of broadscale environmental factors caused by changes over latitude. Therefore, stratification enhanced the

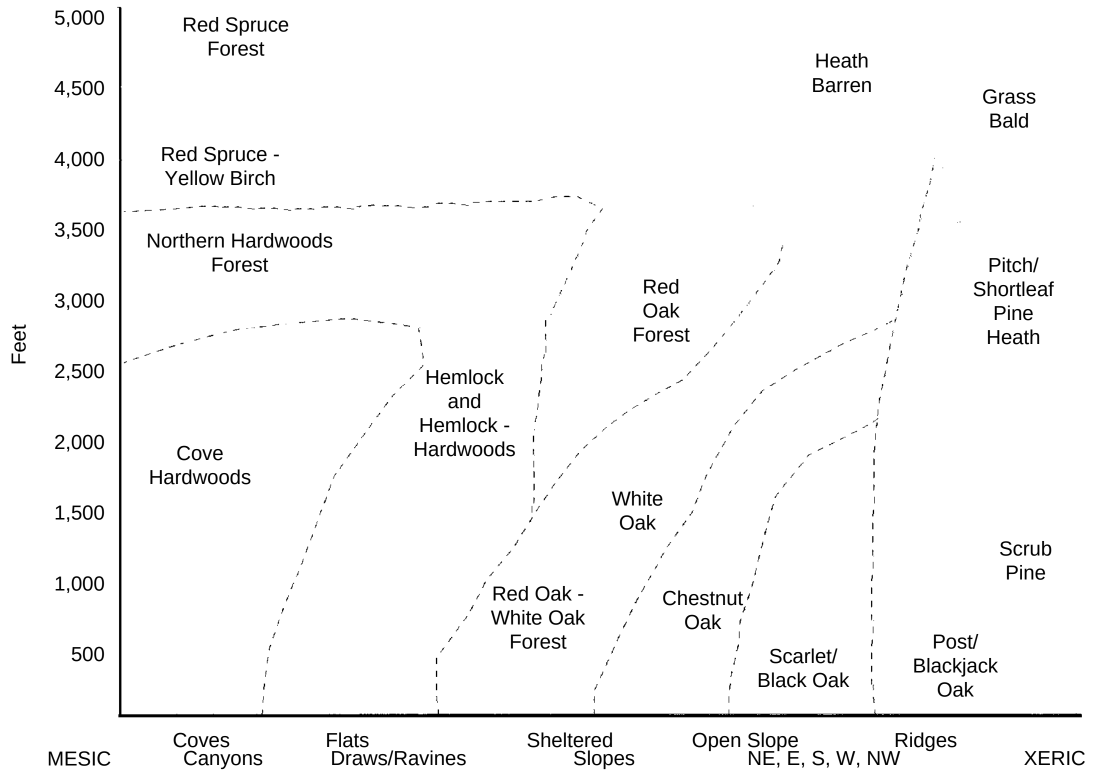


Figure 1. Environmental gradients for vegetation.

efficiency of the training data. An interagency committee, including ecologists from the Monongahela National Forest, West Virginia Division of Forestry, and geologists from the State Geologic Survey, generated a draft map of physiographic regions. They delineated sections based on geomorphology and climate. Sections were divided into subsections: those most typical of the section or those that are transitional, or irregular, to the section. Figure 3 is a draft map of these sections and subsections in West Virginia.

2. Classify stratified imagery using the ancillary data. High resolution imagery has not been used widely in the eastern United States, where forests are not homogeneous stands of relatively few tree species as they are in the West. Researchers who classified eastern forests from satellite imagery attempted to find the most distinctive spectral band combinations for discriminating cover types (10-13). One recent technique (14) uses a nonparametric approach that combines all spectral and informational categories to classify imagery. We are testing a variety of methods such as nonparametric processes, traditional clustering techniques, and use of derived vegetation indexes to find the most successful method.

3. Assess accuracy with random plots from videography. Videography will be acquired in the spring of 1995. Aerial transects, which extend the length of the state, will be flown with approximately 30-kilometer spacing. By regulating flight altitudes, the resolution per frame can be captured at 1 kilometer per frame. About 7,000 frames will be collected, which make up a 3 percent sample of the state. About 2 to 3 percent of the videography frames will be field verified. With this strategy, we will test the effectiveness of using videography, instead of intensive field plots, to verify classification of satellite imagery. Areas of special interest, which the systematic transects may not capture, will require separate flights. The bulk of the videography will provide training data for supervised image classification. The remaining frames will be used to assess the accuracy of the classification.

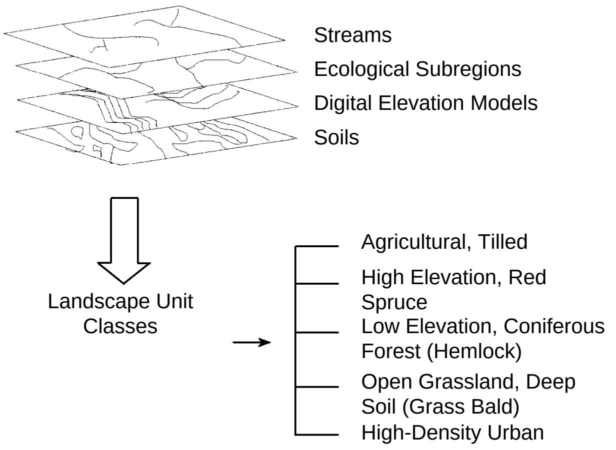


Figure 2. GIS and the development of ecological land units.

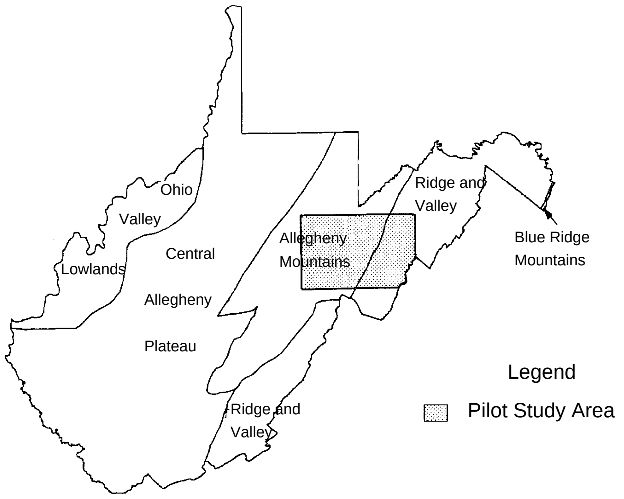


Figure 3. Physiographic regions of West Virginia and pilot study area.

In summary, 100 percent of the state will be classified using the Thematic Mapper imagery. Aerial videography, covering approximately 3 percent of the state, will help to verify the image classification, and 2 to 3 percent of the videography will be verified from transects on the ground.

4. Determine sources of data for image classification. Due to the increasing interest in GIS, digital data are more readily available from a variety of sources, such as the federal government, state agencies, and private companies. Acquisition of available data sets can substantially reduce the time and cost of database development. Users must bear in mind, however, that databases are developed with differing objectives and techniques, so one must consider scale and standards of production when deciding which data sets are appropriate for project design. The West Virginia gap team determined that the following GIS coverages are important for image classification.

The U.S. Geological Survey's (USGS's) graphic information retrieval and analysis system (GIRAS) earmarked land use/land cover data. The classification was done several years ago, and although these data are not current, they provide excellent information concerning urban and agricultural land use. Land-cover categories represent Level II classifications from Anderson's (15) system. The maps are produced at a 1:250,000 scale, so they require few GIS operations to piece together a regional coverage of land use.

The Southern Forest Experiment Station mapped U.S. forestland using advanced very high resolution radiometer (AVHRR) satellite imagery (16). This imagery is relatively current (1991 to 1992), but the resolution is coarse at 1 kilometer per pixel (100 hectares or 247 acres). The classes are based on Forest Inventory and Analysis plots established by the U.S. Department of Agriculture (USDA) Forest Service and Küchler's (17) potential natural vegetation types. We are using the maps to depict broad changes in forest type over a region, such as the state of West Virginia. This coverage does not show land use.

The eastern region of the Nature Conservancy has completed a draft of the classification of the terrestrial community alliances (18). The classification hierarchy is that prescribed by the national gap team, and as such, reflects physiognomic and floristic characteristics necessary for correlating vegetation structure and floristic composition with vertebrate habitat requirements. The descriptions include the range of alliances and characteristic species of the overstory, understory, and herbaceous layer. This provides information on associated species not detectable by image classification.

The National Wetlands Inventory data are available digitally. Maps have been digitized at a scale of 1:24,000, and the classification scheme is from Cowardin et al. (19). Coverages come with attribute data for each polygon, arc, or point as needed. A labor-intensive effort is required to join the maps in a GIS for an area the size of West Virginia, but the information will be invaluable for masking water and forested wetlands on the satellite imagery. The U.S. Fish and Wildlife Service includes detailed instructions for converting the data to coverages.

Field data, much of it already digital, has been acquired from many sources. Commercial timber companies provided data for timber stand composition and age groups. The USDA Forest Service ecologist conducted transects throughout the Monongahela National Forest to characterize ecological land units. Forest inventory and analysis plots are also available. We acquired these data to verify videography classification.

The USGS has digital data of terrain elevations. The West Virginia gap analysis team acquired 3 arc-second DEMs as an additional band in the satellite imagery. We will use these data to generate coverages of slope, aspect, and elevation classes to further stratify the physiographic regions of the state. This will increase the accuracy of the classification.

The Soil Conservation Service created a statewide database called STATSGO, produced at a scale of 1:250,000. For West Virginia, the map of the soil mapping units consists of approximately 450 polygons. Each mapping unit is an aggregation of soil components that occupy a certain percentage of the mapping unit area. The database is extensive and includes information on soil attributes such as soil taxonomy, soil chemistry, soil structure, and interpretations for erodability and wildlife habitat. For an attribute such as soil temperature, each component has an individual value so that each mapping unit may have several different values for soil temperature. Attribute information is difficult to query in ARC/INFO, where there is a one-to-many relationship between polygons and database entries (for instance, each mapping unit, or polygon, has several soil components). We found that exporting the attribute information from ARC/INFO to another software package such as Excel was easier. The values can be aggregated by attribute and then imported into ARC/INFO to produce individual coverages such as soil texture, soil depth, or soil group. STATSGO data can provide useful information for the physical variables that influence vegetation, such as soil moisture and nutrient availability.

To review, the project researchers will first identify the physical parameters that govern the distribution of ecological land units and the existing vegetation in the state. Then, the team will gather applicable ancillary data of physical data in GIS to support the image classification. Once the imagery has been classified, the wildlife models can be incorporated.

Integrate Terrestrial Vertebrate Models

Once classification is completed for the state, the gap analysis team will integrate terrestrial vertebrate models. Concurrent with the image classification, the team will develop a species profile for each vertebrate species known to occur in the state. These profiles, when completed, will be condensed into rule-based models for associated species that can be linked to the ecological land units (see Figure 4). This step will create habitat types. After integration, the team will generate maps that display species richness of vertebrates for each habitat type (see Figure 5). These maps link spatial data to the species database. This enables users to identify areas in the landscape that combine habitats with the greatest number of unique species. When a coverage of landownership is overlaid on this map, land managers or conservation groups can take a proactive stance to seek protection of critical habitats. Additional analyses that users can perform are displays of the potential distribution of vertebrate groups, such as upland salamanders. Another analysis would be to report the species that occur in the fewest habitats and that would be most vulnerable to landscape changes. Clearly, GIS provides a powerful environment for quick and efficient retrieval of spatial data for management decisions.

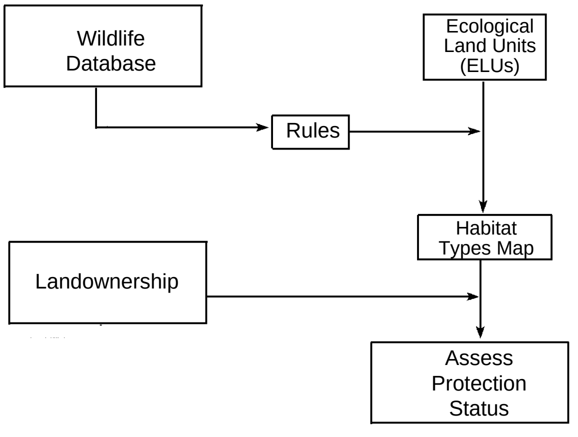


Figure 4. Informational flow chart for wildlife data.

Summary

To summarize, the West Virginia gap analysis team is assessing the natural communities in the state as a part of the national comprehensive planning effort. We need to conduct the assessment rapidly, compiling existing information and integrating these data with GIS and remotely sensed data. Ecological land units are classified according to the relationship between vegetation and ecological variables. Satellite imagery is used to map existing conditions over the state. The existing vegetation is classified to reflect physiognomic and floristic elements to correlate with vertebrate and butterfly habitat requirements.

The gap analysis team is using many widely available data sets such as DEMs, land use/land cover, wetlands inventory, and soils data. While these can reduce the time and cost of developing an ecological database, they do present implications for project design and accuracy. When the user combines maps of different scales, accuracy is constrained by the map with the smallest scale. Additionally, data sets may be constructed with objectives for an intended use that is not compatible with project needs. The classification of AVHRR data reflects forest types but not land use, so another source may be required for these data.

The Gap Analysis Project is not a substitute for intensive biological studies at a fine scale. It is merely a quick assessment at a broad scale to provide information on existing conditions. While accumulating data and modeling potential wildlife distributions, we will identify where inventory data may be lacking. Additional work must be done to verify wildlife models and the classification of vegetation, but this preliminary analysis will be a valuable framework that will direct future studies of biological diversity. Finally, this effort will provide a data set that can be used to monitor changes to land cover and land use.

Acknowledgments

This study is being funded by the Cooperative Research Units Center, National Biological Survey, United States Department of the Interior. This manuscript is published as Scientific Article No. 2499 of the West Virginia Agriculture and Forestry Experiment Station.

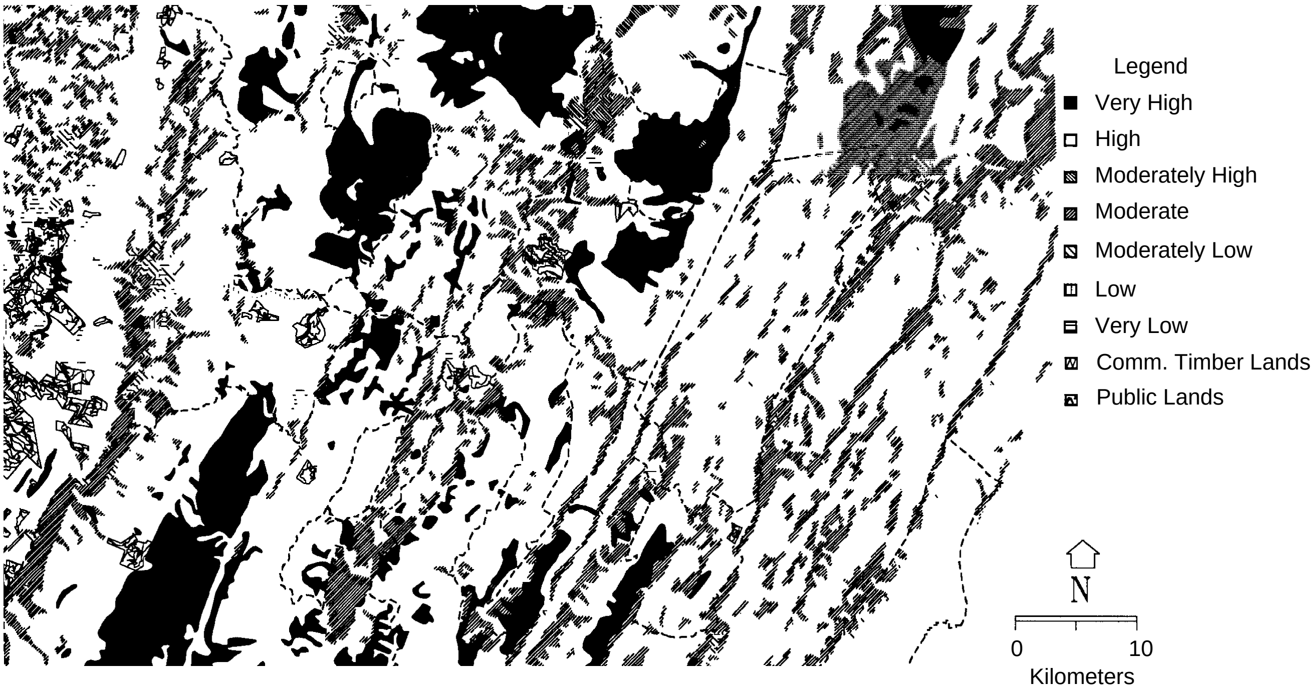


Figure 5. Relative species richness.

References

1. Davis, F.W., and J. Dozier. 1990. Information analysis of a spatial database for ecological land classification. Photogramm. Eng. and Remote Sensing 56:605-613.

2. Scott, J.M., B. Csuti, K. Smith, J.E. Estes, and S. Caicco. 1991. Gap analysis of species richness and vegetation cover: an integrated biodiversity conservation strategy. In: Kohm, K., ed. Balancing on the brink of extinction: the Endangered Species Act and lessons for the future. Washington, DC: Island Press. pp. 282-297.

3. Stephenson, S. 1993. Upland forests of West Virginia. Parsons, WV: McClain Printing Co.

4. Buol, S.W., F.D. Hole, and R.J. McCracken. 1980. Soil genesis and classification, 2nd ed. Ames, IA: Iowa State University Press.

5. Bailey, R.G. 1994. Ecoregions and subecoregions of the United States (map). U.S. Department of Agriculture, Forest Service. Revised February 23.

6. Fortney, R. 1994. Vegetation pattern of central Appalachian Mountains. (Unpublished).

7. Barbour, M.G., J.H. Burk, and W.D. Pitts. 1987. Terrestrial plant ecology, 2nd ed. Menlo Park, CA: Benjamin/Cummings Publishing Company.

8. Davis, F.W., D.M. Stoms, J.E. Estes, J. Scepah, and M. Scott. 1990. An information systems approach to the preservation of biological diversity. Int. J. Geogr. Info. Syst. 4:55-78.

9. Scott, J.M., F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, H. Anderson, S. Caicco, F. D'Erchia, T.C. Edwards, Jr., J. Ulliman, and R.G. Wright. 1993. Gap analysis: A geographic approach to protection of biological diversity. Wildl. Monogr. 123:1-41.

10. Bauer, M.E., T.E. Burk, A.R. Ek, P.R. Coppin, S.D. Lime, T.A. Walsh, D.K. Walters, W. Befort, and D.F. Heinzen. 1994. Satellite inventory of Minnesota forest resources. Photogramm. Eng. and Remote Sensing 60:287-298.

11. Benson, A.S., and S.D. DeGloria. 1985. Interpretation of Landsat-4 thematic mapper and multispectral scanner data for forest surveys. Photogramm. Eng. and Remote Sensing 51:1,281-1,289.

12. Hopkins, P.F., A.L. Maclean, and T.M. Lillesand. 1988. Assessment of thematic mapper imagery for forestry application under lake states conditions. Photogramm. Eng. and Remote Sensing 54:61-68.

13. Schriever, J.R., and R.G. Congalton. 1993. Mapping forest covertypes in New Hampshire using multitemporal Landsat Thematic Mapper data. Papers presented at ASCM/ASPRS Annual Convention and Exposition, New Orleans, LA (February). pp. 333-342.

14. Cetin, H., T.A. Warner, and D.W. Levandowski. 1993. Data classification, visualization, and enhancement using n-dimensional probability density functions (nPDF): AVIRIS, TIMS, TM, and geophysical applications. Photogramm. Eng. and Remote Sensing 59:1,755-1,764.

15. Anderson, J.R., E.E. Hardy, J.T. Roach, and R.E. Witmer. 1976. A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 965. Washington, DC: U.S. Geological Survey.

16. Zhu, Z., and D.L. Evans. 1992. Mapping midsouth forest distributions: AVHRR satellite data and GIS help meet RPA mandate. J. Forestry 27-30.

17. Küchler, A.W. 1964. Potential natural vegetation of the conterminous United States (map and manual). American Geographical Society Special Publication 36. New York, NY: American Geographical Society.

18. Sneddon, L., M. Anderson, and K. Metzler. 1994. A classification and description of terrestrial community alliances in the Nature Conservancy's eastern region: First approximation. Washington, DC: U.S. Department of the Interior Fish and Wildlife Service.

19. Cowardin, L.M., V. Carter, F.C. Golte, and T.E. LaRoe. 1979. Classification of wetlands and deepwater habitats of the United States. Report No. FWS/OBS-79/31. Washington, DC: U.S. Department of the Interior Fish and Wildlife Service.

GIS Watershed Applications in the Analysis of Nonpoint Source Pollution

Thomas H. Cahill, Wesley R. Horner, and Joel S. McGuire
Cahill Associates, West Chester, Pennsylvania

Introduction

Geographic information systems (GIS) have been used to evaluate the impact of nonpoint source (NPS) pollution in a variety of watersheds and drainage systems over the past 20 years (1-6). During that period, our understanding of the sources and hydrologic transport mechanisms of NPS pollutants, both in particulate and soluble forms, has greatly increased (7-9). Our ability to create and manipulate land resource data, however, has advanced at a far more dramatic rate. Whereas 20 years ago, both computer system capabilities and peripheral hardware limited the process of encoding, storing, and displaying spatial data, today we can encode land resource data, analyze it, and produce stunning visual displays at a relatively low cost.

The question is: what has this experience told us regarding the yet unresolved problem of water quality degradation from NPS pollution in our streams, lakes, and coastal waters (10, 11)?

The purpose of this paper is to report on several recent studies of this nature that created a GIS as a tool to analyze NPS pollution. This paper will not cover all aspects of these studies; detailed reports on each project are available from the authors or respective clients. The objects of these studies were:

- A medium-sized lake draining a fairly small watershed
- A riverine system with multiple use impoundments
- A 100-mile stretch of Atlantic coastal estuary

These water bodies all have one common ingredient: NPS pollution significantly affects them. While the primary focus of these studies was to understand the dynamics of surface water quality, and specifically the NPS component, the further objective was to document the causal link between identified water resource problems and the watershed-wide management actions needed for their remediation. Thus, GIS serves not only as a mechanism for analysis of NPS pollution sources but also as the tool by which to evaluate alternative methods that would reduce or prevent this pollution.

Study Concepts

These three studies illustrate different approaches to both aspects of this problem. In the 93-square-mile Upper Perkiomen Creek watershed (UPW) study, the objective was to develop a management program that would reduce nutrient load in a system of reservoirs at the base of the watershed. An essential element in the analysis underlying GIS design (ARC/CAD) was to be able to differentiate and evaluate pollution sources in the watershed, while providing the technical basis for an innovative and far-reaching management program on all levels of government; that is, GIS was used not only to analyze the problem but to help formulate the solution.

In the more focused Neshaminy Creek study, Cahill Associates (CA) designed a detailed pixel/raster format for GIS to support detailed hydrologic modeling (12) and NPS loading analysis. This study, carried out under Pennsylvania's Act 167 stormwater management program, was under a legal requirement to translate technical findings into subdivision regulations that all 30 watershed municipalities would adopt. This mandate required much more geographically specific rigor in the GIS approach and in the management recommendations the law stipulated.

These two projects (see Figure 1), when taken together, illustrate the critical relationship between understanding the appropriate level of detail in GIS system design, GIS development with modeling and other analytical requirements, and ultimately, the proposed management actions for watershed-wide implementation.

In the New Jersey Atlantic Coastal Drainage (ACD) study, the objective was to document more completely the magnitude and sources of NPS pollutants, especially nutrients, entering New Jersey estuarine coastal waters. The GIS design placed special attention on the role of urban or developed land uses situated along the coastal fringe, particularly the maintained or landscaped portions of developed sites. Most previous studies have largely ignored this factor. Instead, they have focused water quality analysis typically on NPS loadings as a

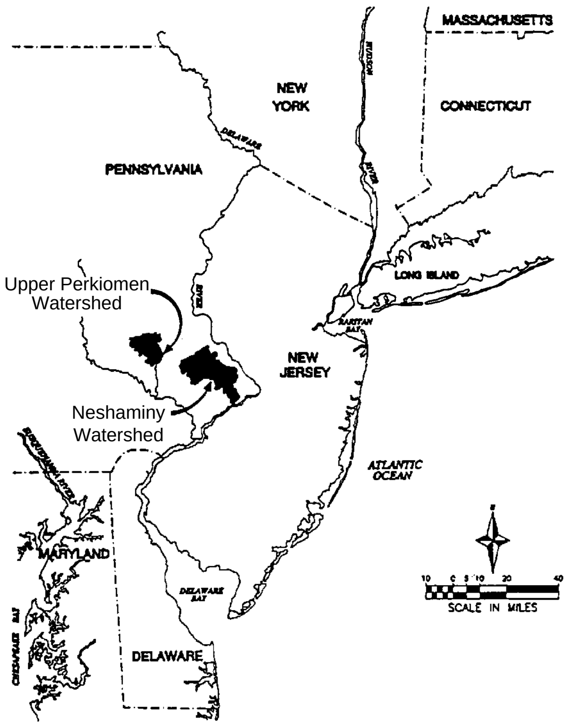


Figure 1. Regional location of Upper Perkiomen and Neshaminy basins.

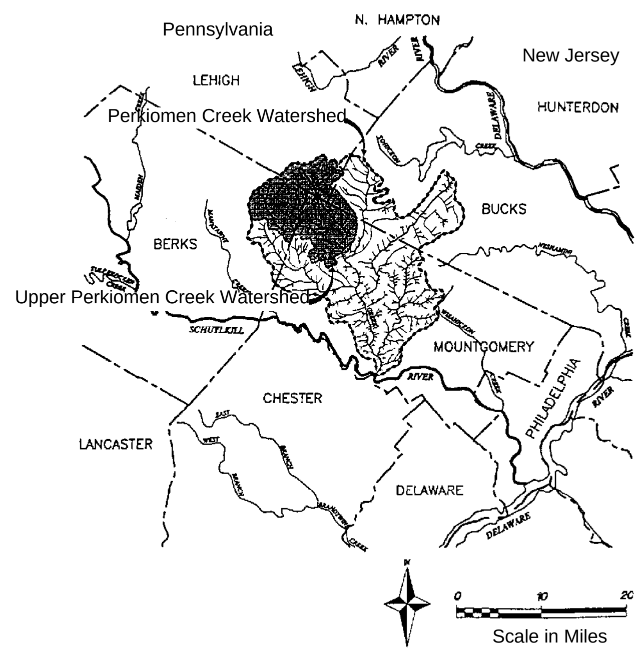


Figure 2. The Perkiomen Creek watershed in the Delaware River basin.

function of impervious area coverage, with the assumption that loadings increase as imperviousness increases.

On the contrary, the CA thesis states that certain pollutant loadings, such as nutrients, maximize in areas with relatively moderate densities (1/2- to 1-acre lots) and percentage impervious cover but with large maintained lawnscapes. Because sandy soils allow soluble NPS pollutants to pass as interflow to points of surface discharge with surprising ease, they exacerbate the problem of nutrient applications in typical coastal drainage areas. GIS application in this case enabled estimation of the nutrient loading to coastal waters. Existing fertilized lawn areas were calculated to be a significant source of nutrient pollution, with loadings from new land development posited as an even more serious problem for New Jersey's coastal waters. GIS was then applied to evaluate the suitability of various best management practices (BMPs), based on the physical and chemical properties of the soil mantle and the existing and anticipated land use.

The Upper Perkiomen Creek Watershed Study

Background

The UPW in southeastern Pennsylvania is a tributary of the Schuylkill River in the Delaware River basin (see Figure 2). Serious eutrophication problems occurring in the system of reservoirs lying at the base of this relatively rural watershed prompted the study. The study effort evolved from concerns on the part of the Delaware Riverkeeper, a private nonprofit environmental organization dedicated to promoting the environmental well-being of the Delaware River watershed. The Upper Perkiomen Creek has experienced various water quality problems, especially the eutrophication of Green Lane Reservoir, a large raw water supply storage reservoir (see Figure 3). Green Lane's highly eutrophic condition has been a constant since shortly after initial construction over 35 years ago, but the relative importance of NPS inputs has dramatically increased. Whereas 10 years ago point source input was the major source of phosphorus, elimination of some point sources and advanced waste treatment for others has greatly reduced that component of pollutant loading, while NPS sources have remained constant or increased. Current analysis indicates that NPS pollution constitutes over 80 percent of the annual load of phosphorus (see Figure 4) into the Green Lane Reservoir and is well in excess of the desired loading to restore water quality (see Figure 5).

Nonpoint Source Analysis

Calculating the NPS load was an essential ingredient in the study and relied on developing accurate measurement of NPS transport during stormwater runoff periods. Certain pollutants, specifically those associated with sediment and particulate transport such as phosphorus, have produced a "chemograph" that parallels but does not exactly follow the traditional form of the hydrograph (see Figure 6). The pollutant mass transport associated with this runoff flux frequently constitutes the major fraction of NPS discharge in a given watershed (8, 13). In

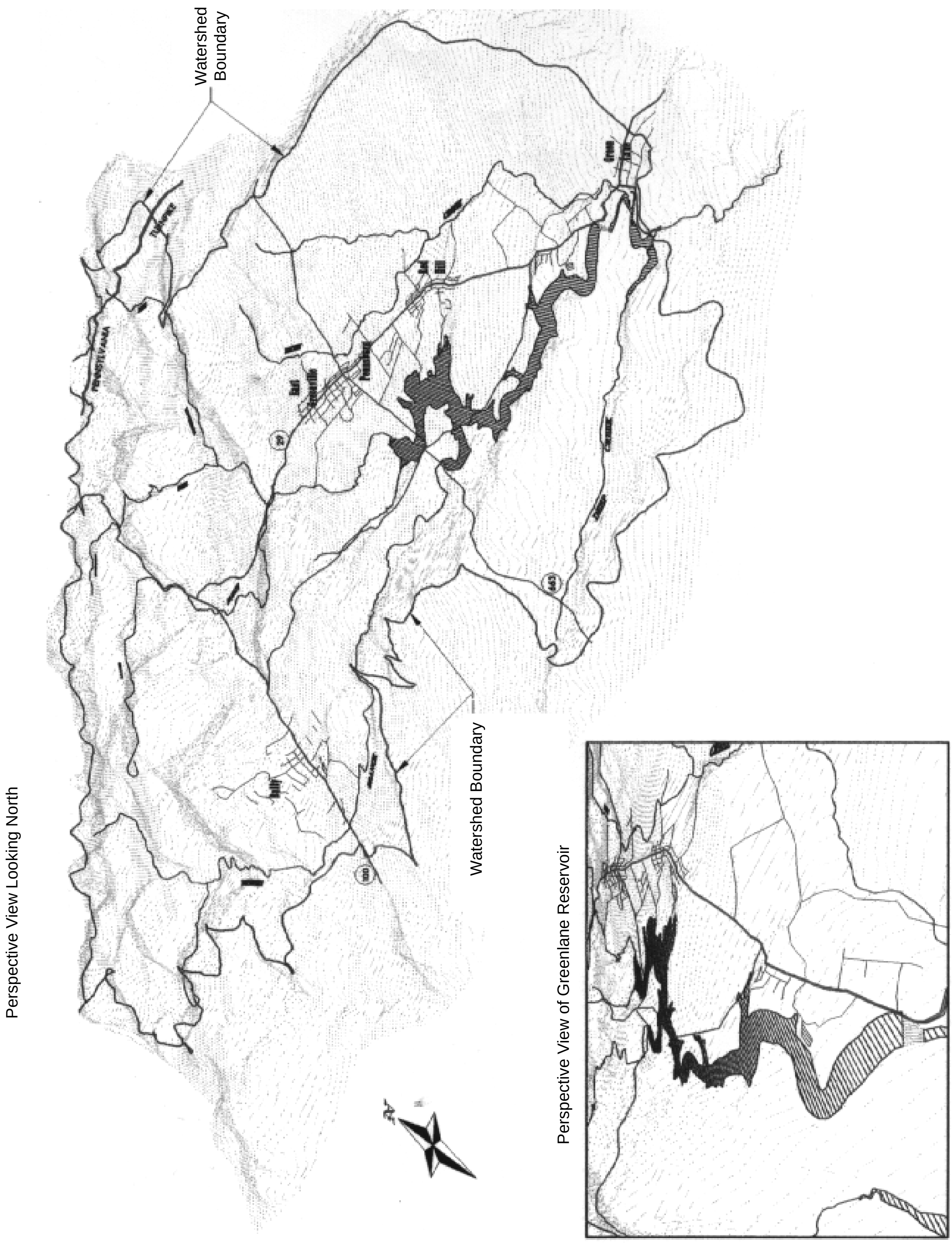


Figure 3. Green Lane Reservoir in the Upper Perkiomen watershed, 814 acres, 4.3 BG.

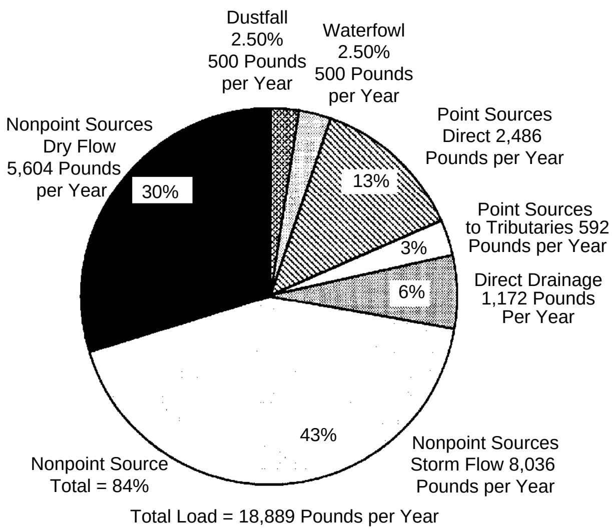


Figure 4. Sources of total phosphorus mass transport into the Green Lane Reservoir from the Upper Perkiomen watershed (71 square miles) in an average flow year—in pounds per year.

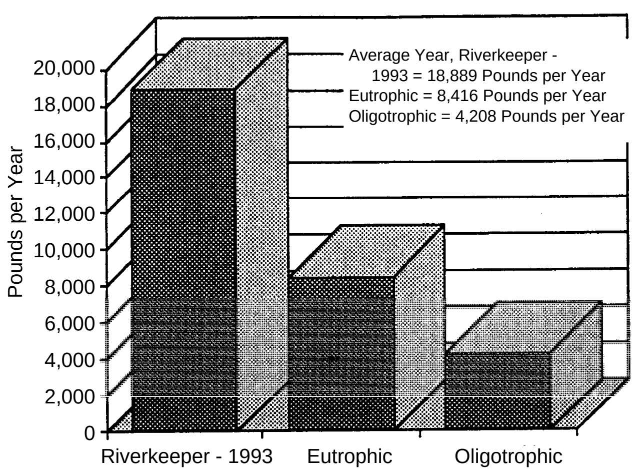


Figure 5. Reduction in annual phosphorus load required to achieve improved trophic level.

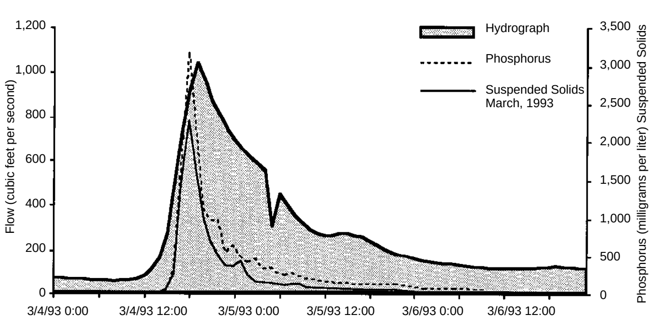


Figure 6. Storm hydrograph in the Upper Perkiomen watershed illustrating the dramatic increase in total phosphorus and suspended sediment during runoff.

the UPW study, operating continuous sampling stations at two key gage locations above the reservoir allowed the measurement of stormwater chemistry of this type and produced estimates of wet weather transport of phosphorus and sediment. Surprisingly, the NPS transport during dry weather, calculated by subtracting the point sources, was also significant and is attributed to livestock discharges and septage drainage.

But the wet weather proportion of NPS pollution still dominates lake water quality. Many have said that water quality in a given watershed is a function of land use, but that statement is as unsatisfying as saying that runoff is a function of rainfall. Experience has taught us that neither process is quite that simplistic, nor does either follow a direct linear relationship of cause and effect. The causal mechanisms that generate a certain mass load of pollutant in a drainage basin certainly result from how much mass of that pollutant is applied to the landscape within the drainage, which in turn is scoured from the landscape during periods of surface saturation, transported in, and diluted by runoff. The end result is a concentration of pollutant in the stormwater that might be several orders of magnitude greater than during dry weather flow, the hydrologic period traditionally used to measure and define water quality.

Developing NPS analysis or algorithms for stormwater quality modeling requires replicating the specific hydrograph and its associated chemograph, as well as defining the mechanisms by which pollutants are scoured from the land surface, transported in runoff, and pass through the river system. Total phosphorus (TP), for example, is transported with the colloidal soil particles (see Figure 7), so sediment transport and deposition constitute a key mechanism.

Adding to these complications is the question of whether to model single or multiple events. Is the chemodynamic process one in which the transport takes place over a series of storm events, so that each storm moves the pollutant mass a given distance in the drainage and then allows it to settle in the channel only to resuspend it with the next peak of flow? Or does the total mass transport occur in one single dynamic, from corn field or suburban lawn to lake, estuary, or other sink, that is hours or days downstream in the drainage? The issue of how stormwater transport of pollutants takes place is of paramount importance in current planning and regulatory implementation (11) because many of our current BMPs are relatively ineffective in removing NPS pollutants. This understanding is critical even as we attempt to intervene in the pollutant generation process by changing the way we cultivate the land, fertilize our landscapes, or for that matter, how we alter the land surface during growth.

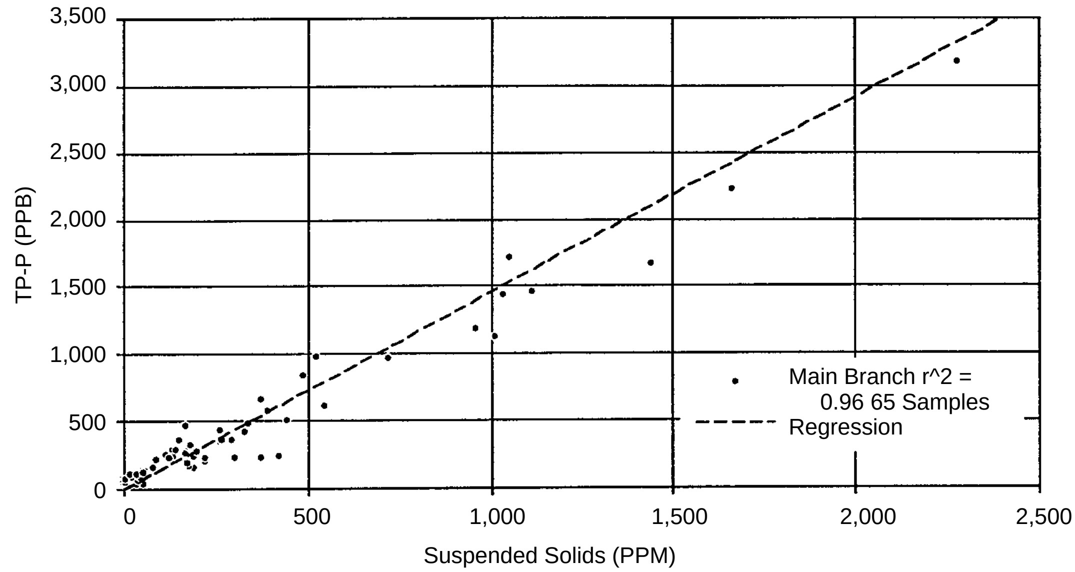


Figure 7. The relationship between total phosphorus and suspended sediment concentrations during runoff is strong but varies with different watersheds.

GIS Evaluation

The GIS data files on land use/land cover that were created for the UPW show that the bulk of the area is still quite undeveloped and rural (see Figure 8), with the steeply sloped and igneous rock areas in the headwaters in forest cover (38 percent) and the valleys in mixed agriculture (44 percent). The urban/suburban land composes the remaining 18 percent and largely consists of several older, historic boroughs linked together in a lineal pattern with widely scattered, low-density residential areas. Much of the existing housing is turn-of-the-century at quite high densities, mixed with a variety of commercial and other uses. This pattern contrasts sharply with typical large-lot suburban subdivisions. In fact, these watershed boroughs resemble the “village” concepts that innovative planning theorists advocate in a variety of important ways.

The watershed (see Figure 9) is blessed, or cursed, depending upon one’s perspective, with a multiplicity of local governments including four different counties and 18 different municipalities. This arrangement poses special challenges for management program implementation. Population projections indicate that additional development will occur at moderate rates throughout the watershed, reflecting recent trends.

Farming, both crop cultivation and dairying, is a major existing land use in the watershed, although agriculture is not especially robust and appears to be declining. This lack of agricultural vibrancy becomes a major factor in determining how to impose additional management measures on agricultural pollution sources. GIS tabulation of agricultural land totals some 19,000 acres above the reservoir, which can be compared with the estimated TP and suspended solids (SS) mass transport reaching the lake. Considering only the agricultural land to be the source of this NPS input (not quite true) suggests an average annual yield of 180#/acre/year-SS and 0.22#/acre/year-TP.

This sediment/phosphorus yield is more than sufficient to maintain a eutrophic condition in the reservoir system. The problem with this yield, however, is that it is two orders of magnitude less than commonly accepted methodologies of soil erosion, such as the universal soil loss equation (14), would suggest might come from such a watershed. Analysis of the cultivation practices taking place on farmland in the watershed estimates soil erosion to be approximately 5 to 10 tons per acre or more per year, far more than is observed passing out of the basin into the reservoir. The phosphorus applications on both cultivated and maintained residential landscapes also appear much greater than the mass transport actually measured in the flowing streams, which represent perhaps 7 percent or less of the annual land application.

The implication for NPS analysis is that the standard shopping list of either agricultural or urban BMPs might only reduce the mass transport by a relatively small fraction, even if successfully applied throughout the drainage. As Figure 7 illustrates, most of the phosphorus transport occurs on the colloidal fraction of sediment particles, which tend to remain in suspension as stormwaters pass through conventional detention structures, terraces, or grassed swales.

To consider more radical measures, GIS was used to determine possible alternatives, such as creating a stream buffer system (see Figure 10) with various setback distances from the perennial stream network, and to evaluate how great an impact this might have on agricultural land use and urban development. Land

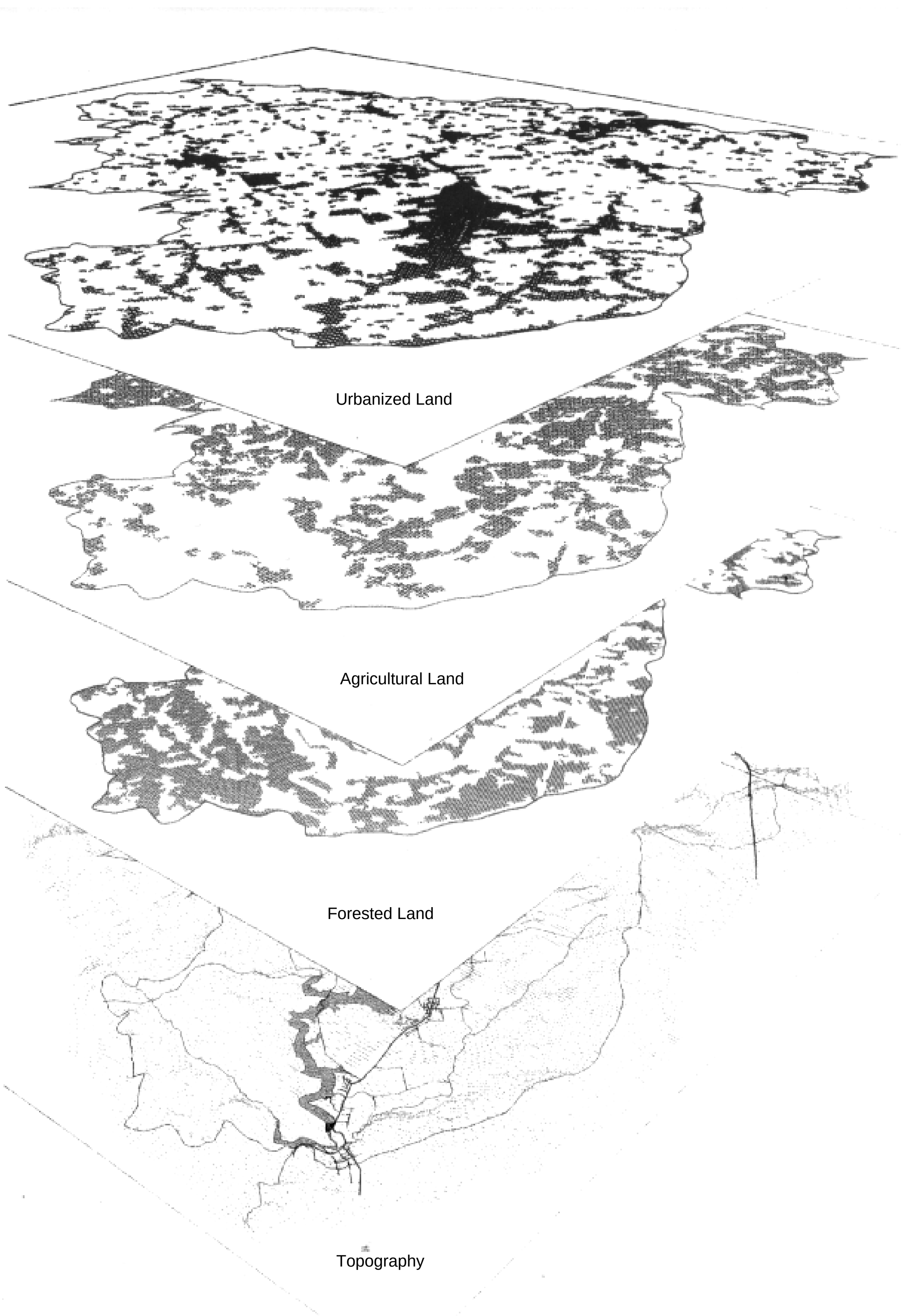


Figure 8. GIS data files showing land use/land cover characteristics for the Upper Perkomien watershed.

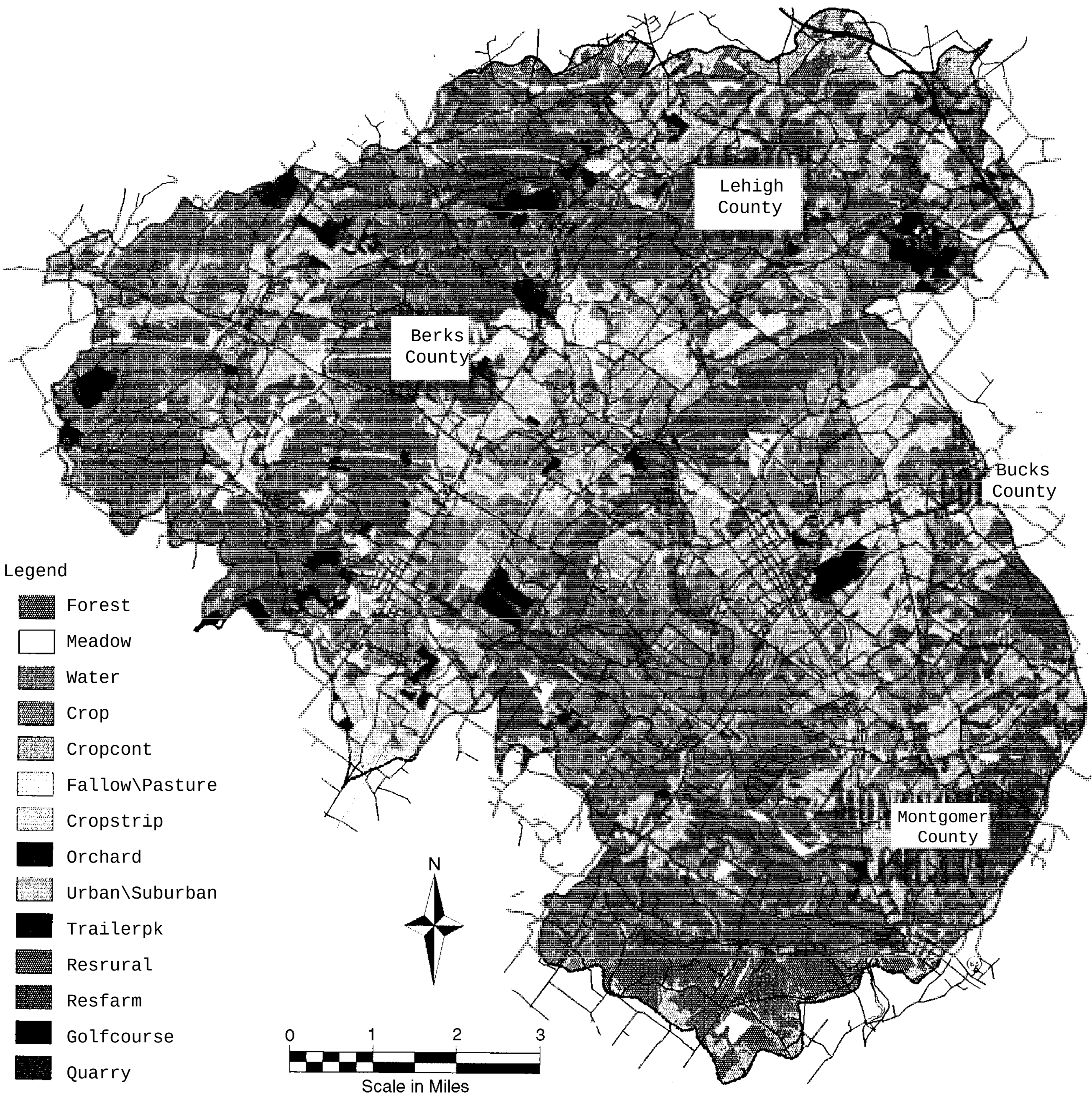


Figure 9. Existing land use/land cover GIS file for the Upper Perkiomen watershed. The 95-square-mile basin includes portions of four counties and 18 municipalities.

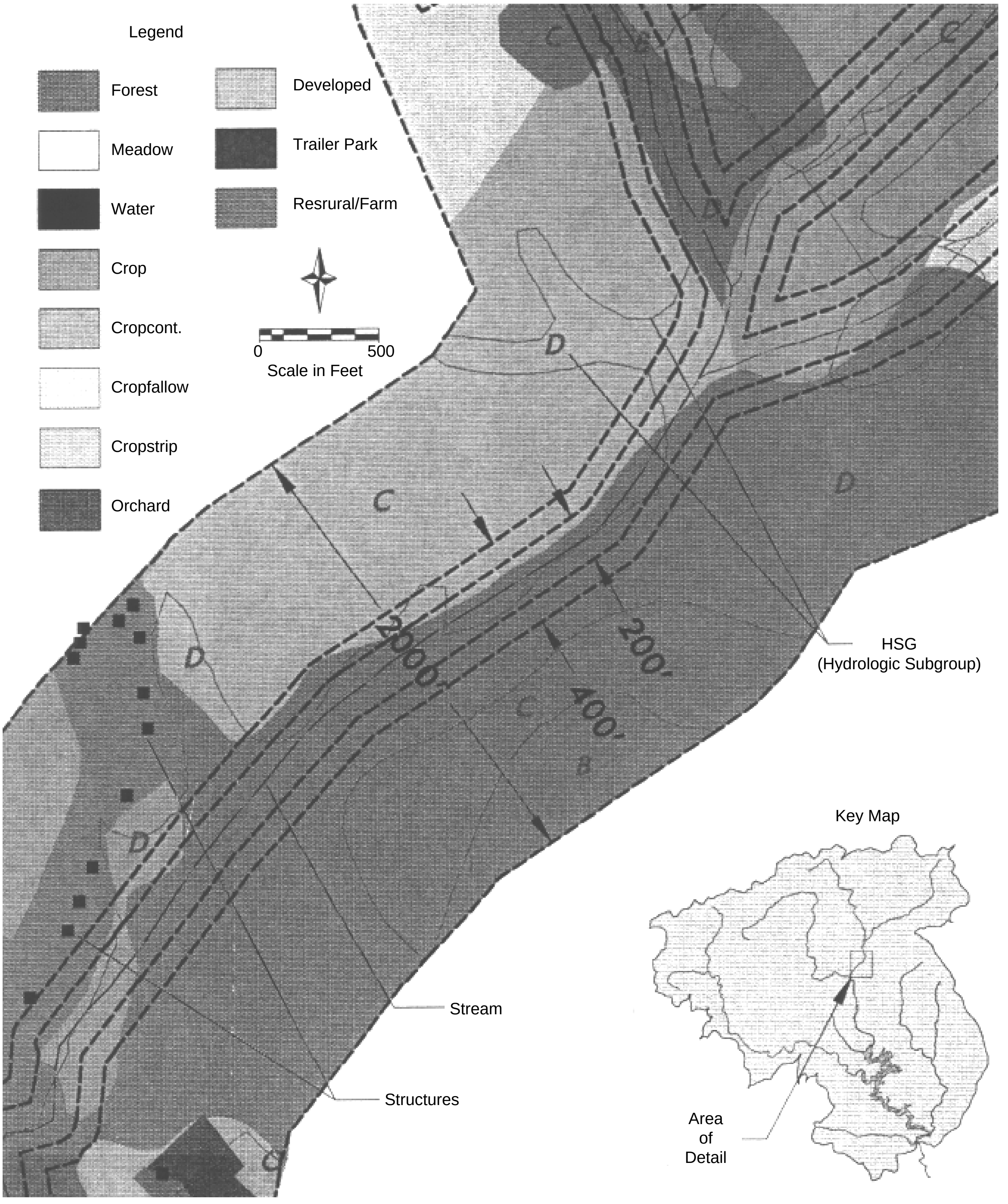


Figure 10. GIS analysis of stream corridors allows evaluation of riparian buffer systems, potential agricultural land loss, potential septic system discharges, and related NPS reduction with selected best management practices.

use at varying distances (100 feet, 200 feet, and 1,000 feet) from streams was tabulated, including all land area in the "active" agriculture categories. This GIS documentation allowed estimation of the significant NPS reduction in loadings that a riparian corridor management program could achieve.

In the same way, GIS analysis helped estimate pollutant loadings from malfunctioning onsite septic systems. Counts of structures in nonpublicly sewered areas within varying distances from the stream system were developed using GIS data files. The nearly 300 potential systems within a 200-foot radius of those streams draining into the Green Lane Reservoir identified in this manner, with pollutant generation factors applied, became the basis of a dry weather pollutant estimation. Although this approach was dependent on a variety of assumptions, alternative approaches of evaluating the problem, such as field visits to actual onsite systems throughout the watershed, would not have been feasible.

For urban and suburban development, the management focus was to estimate NPS loadings from future land development. GIS was used to demonstrate NPS pollutant load implications of future growth envisioned in the watershed's keystone municipality, Upper Hanover Township. Here, an increase of 15,000 residents would convert 1,772 acres into residential, commercial, and industrial uses. Nonpoint pollutant loadings generated by this new land development constituted significant increases in phosphorus, suspended solids, metals, oil/grease, and other pollutants, and would reverse any improvements in Green Lane Reservoir water quality that recent wastewater treatment plant upgrades achieved.

From a water quality perspective, future alternative land use configurations that concentrate development and minimize ultimate disturbance of the land surface yielded would substantially reduce NPS pollutant loadings into the reservoirs. This entire process of testing land use implications of different management approaches for their water quality impacts indicated that pollutant loads could be minimized far more cost effectively through management actions, both structural and nonstructural, which varied from the areawide to the site-specific.

Neshaminy Creek Watershed Stormwater Management Study

Background

The Neshaminy Creek watershed, including 237 square miles of mixed urban and rural land uses, lies primarily in Bucks County, Pennsylvania, and flows directly into the Delaware River (see Figure 1). The 1978 Pennsylvania 167 Stormwater Management Act, which required that counties prepare stormwater management plans for all 353 designated watersheds in the state, mandated the Neshaminy study. This act further stipulated that municipalities then needed to implement the watershed plans through adopting the necessary municipal ordinances and regulations. In fact, the Neshaminy study had three water resource management objectives:

- Prevent worsened flooding downstream caused by increased volumes of runoff from land development.
- Increase ground-water recharge.
- Reduce NPS pollutant loadings from new development.

In the initial study design, water quality and NPS issues were secondary to flooding concerns. When Pennsylvania's stormwater management program was conceived, the state focused on preventing watershed-wide flooding. Clearly, detention basins have become the primary mode of managing peak rates of stormwater discharge site-by-site in most communities. Because detention basins only control peak rates of runoff and allow significantly increased total volumes of water discharged from sites, however, the increased stormwater volumes can theoretically combine and create worsened flooding downstream. Consequently, most Act 167 planning has focused on elaborate hydrologic modeling designed to assess the seriousness of potential cumulative flooding in watersheds under study.

In the case of the Neshaminy, however, the record suggested that although localized flooding could be an issue, an existing network of eight multipurpose flood control structures constructed during the 1960s served to prevent significant flooding. Water quality certainly was a serious stormwater concern, however, especially in the areas flowing into the reservoirs where recreational use had become intense. Several of the existing impoundments were multipurpose, their permanent pools providing critical recreational functions for a burgeoning Bucks County population. At the same time, the proliferation of development in the watershed, with its increased point and nonpoint sources, had degraded streams and seriously affected the reservoirs. While the total stream system in the watershed was of concern, the future of the reservoirs came to be particularly important in developing the total stormwater management program for the Neshaminy watershed.

The Neshaminy lies at the heart of Bucks County, Pennsylvania's primary population and employment growth county (see Figure 11). Although the Neshaminy watershed has already experienced heavy development, especially in the lower or southern portions, farmsteads and large areas of undeveloped land still exist, especially in headwater areas. Agriculture has been a major land use in the past, but farms rapidly are converting to urban uses as the wave of urbanization moves outward from Philadelphia and from the Princeton/Trenton metropolitan areas. Growth projections indicate continu-

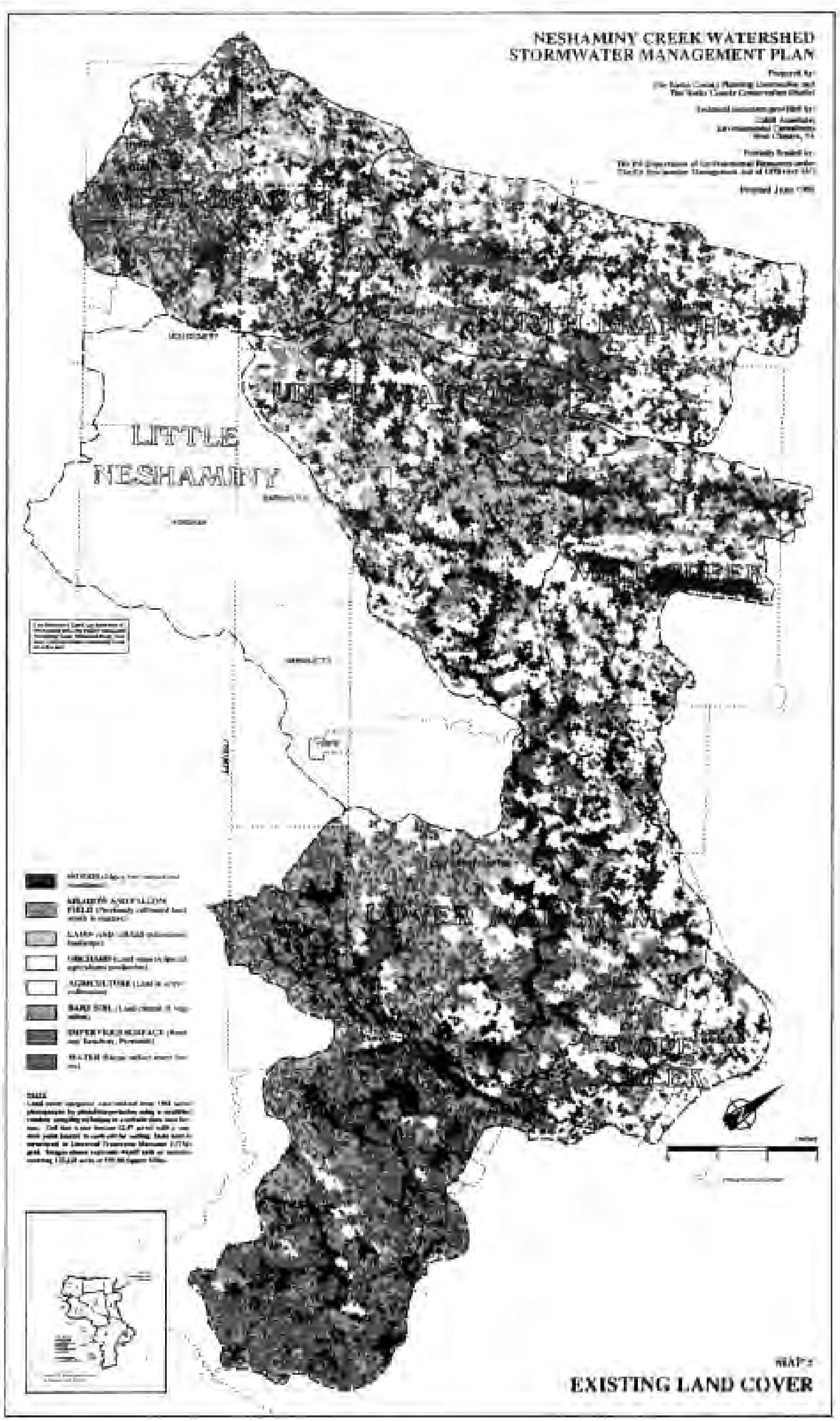


Figure 11. Land use/land cover in the Neshaminy basin of Bucks County, Pennsylvania. The watershed covers 237 square miles in southeast Pennsylvania.

ation of this rapid growth and a continuing change in existing land use/land cover, together with projected development within the required 10-year planning horizon.

Physiographically, the watershed spans both the Piedmont and Atlantic coastal plain provinces, with rolling topography and relatively steep slopes underlain by Triassic formation rock, including the Lockatong, Brunswick, and Stockton formations. This bedrock ranges from being a poor aquifer (Lockatong) to an excellent aquifer (Stockton) where the many rock fractures allow for considerable ground-water yields. Soils are quite variable, ranging from good loam (hydrologic soil group B) to clays and other types with poor drainage characteristics (e.g., high water table, shallow depth to bedrock). A large proportion of the soils in the watershed are categorized as hydrologic soil group C, which is marginal for many stormwater management infiltration techniques (see Figure 12) and produces a relatively large proportion of direct runoff. With an annual rainfall of 45 inches, base flow accounts for about 12 inches and direct runoff accounts for 10 inches.

The system of eight stormwater control structures, which were built over the past three decades under the federal PL 566 program, have altered the hydrology of the watershed (15). In addition, in heavily developed portions of the watershed, impervious surfaces combined with numerous detention basins prevent the bulk of the precipitation from being recharged, and the volume of total runoff proportionally increases. An elaborate system of municipal and nonmunicipal wastewater treatment plants also adds to this alteration of the hydrologic cycle. These plants discharge wastewater effluent that, in some cases, constitutes the bulk of the stream flow during dry periods. While the impact of NPS was evident throughout the drainage, it was of special interest in the impoundment network, especially those impoundments that were conceived as multipurpose in function and constitute major recreational resources in the watershed.

GIS Design

Act 167 requirements and the needs of the hydrologic and other modeling used in planning both heavily influenced the GIS developed for the Neshaminy. Spatial data files, including existing land use, future land use, and soil series aggregated by hydrologic soil groups, were created by digitizing at a 1-hectare (2.5-acre) cell resolution. The encoding process that helped design the GIS used a stratified random point sampling technique that similar studies had developed and applied (1, 3). The encoding process used a metric grid of 5-kilometer sections, subdivided into 2,500 1-hectare cells (100 meters on a side), aligned with the Universal Transverse Mercator (UTM) Grid System. This grid appears in blue on U.S. Geological Survey (USGS) topographic maps. These maps served as the framework of reference for all data compilation. Within each 100-meter cell, a randomly located point was chosen (see Figure 13) at which the specific factor was encoded as representative for the cell, using a digitizer tablet. This approach allowed extraction of the data from the respective source documents with some rectification necessary for many types of source maps and photographs.

The combination of soil series and cover in each cell helped to calculate the curve number and unit runoff per cell. The 45,000-cell data file was then used to calculate total runoff for a range of events in each of 100 subbasins that averaged 1.95 square miles each. The resultant hydrographs, used in combination with a separate linear data file in GIS describing the hydrographic network of stream geometry, routed and calibrated the hydrologic model (TR-20). NPS mass transport loadings were estimated on an annual basis by cell, again using the land use/land cover data file, and total loads summed by groups of subbasins above critical locations. This issue was particularly important with respect to the drainage areas above the impoundments, where NPS pollutants were of greatest concern.

The soil properties data file was especially useful in evaluating certain management objectives, such as the opportunity for recharging ground-water aquifers. The spatial variation in relative effectiveness of infiltration BMPs was considered for both quantity and quality mitigation because the best methods for NPS reduction usually include recharge where possible. The soil series corresponding with new growth areas were classified regarding their suitability for these BMPs, which are most efficient on well-drained or moderately well-drained soil. Thus, the alternative impacts of future growth could be considered in terms of potential generation (or management) of NPS loads. A BMP selection methodology (see Figure 14), which was developed for the 30 municipalities within the watershed, focused on new land development applications and considered both water quantity and quality management objectives. BMP selection is a function of several factors, including:

- The need for further peak rate reduction.
- The recharge sensitivity of the project site (defined as a function of headwaters stream location, areawide reliance on ground water for water supply, or presence of effluent limited streams).
- The need for priority NPS pollution controls (location within reservoir drainage).

Development of two “performance” levels of BMP selection techniques gave municipalities some degree of flexibility in developing their new stormwater management programs. This system required only the minimally acceptable techniques but recommended the more fully effective ones, hoping that municipalities would strive to incorporate

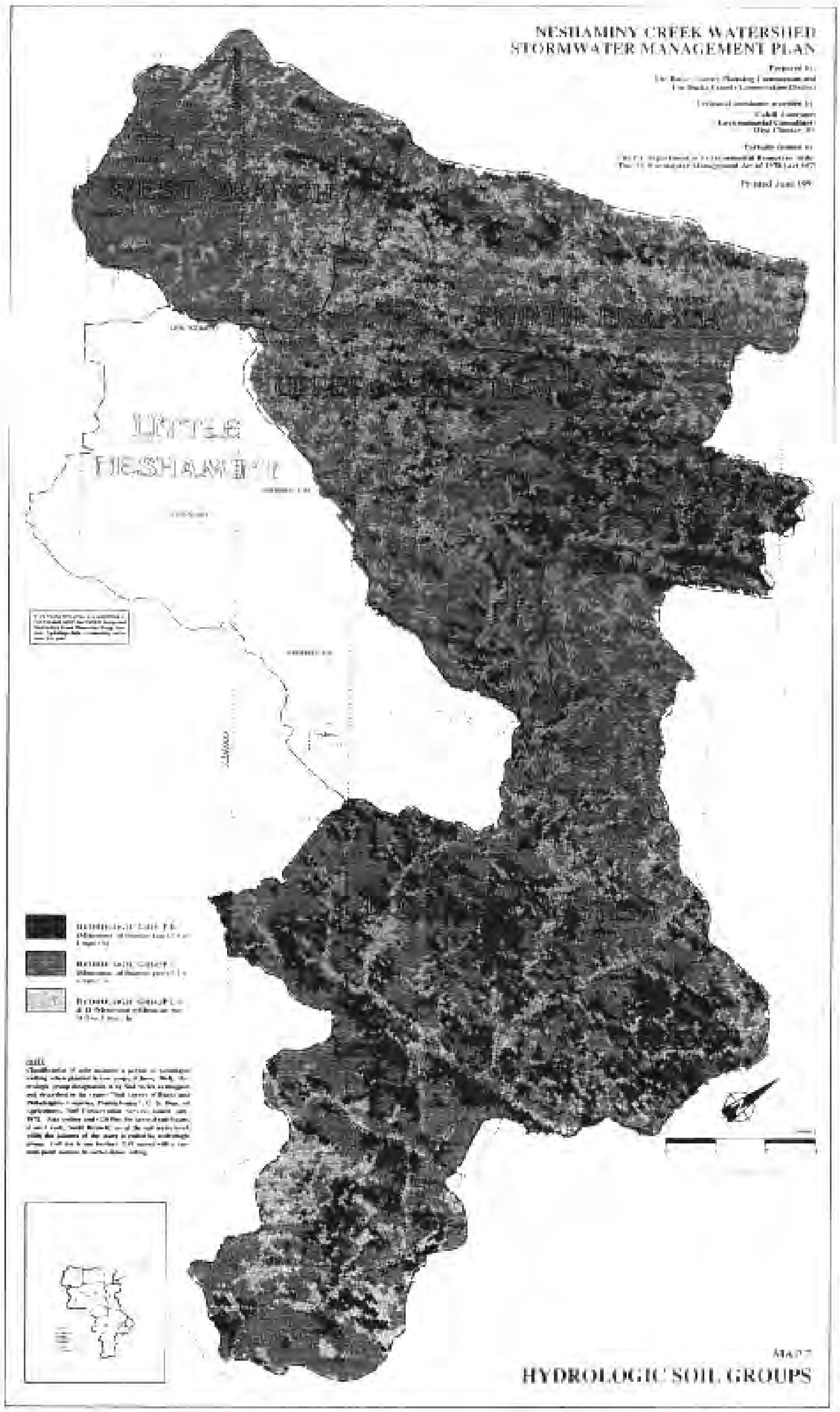


Figure 12. GIS file of hydrologic soil groups in the Neshaminy basin. The 31 soil series are digitized in 45,000 pixels of 1-hectare size.

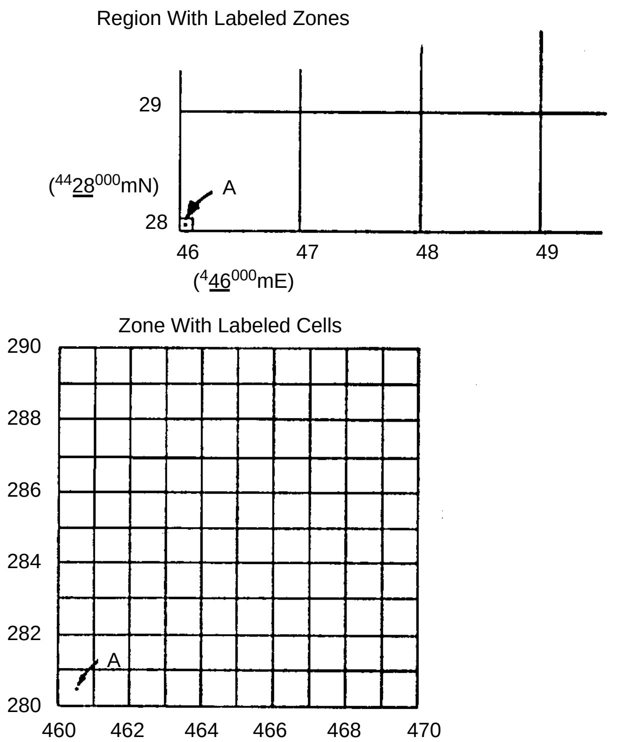


Figure 13. Raster/pixel design of GIS for Neshaminy modeling study. Each pixel is 1 hectare (2.47 acres).

recommended management measures wherever possible. The BMP selection methodology also was sensitive to type of land use or proposed development, assigning typical single-family residential subdivisions different BMPs than, for example, multifamily and other nonresidential proposals (including commercial and industrial proposals). The selection process also determined size of site to be a factor, differentiating between sites of 5 acres or more because of the varying degrees of cost and effectiveness of different BMP approaches. The methodology, if properly and fully implemented, should achieve the necessary stormwater-related objectives—both quantity and quality—that the analysis had deemed necessary (16).

GIS was especially important in its ability to test how reasonable the BMP selection methodology was. Such tests included the ability to evaluate, for each municipality, the following factors:

- The nature and extent of the projected development.
- The size of development/size of site assumptions.
- Other vital BMP feasibility factors such as soils and their appropriateness for different BMP techniques.

GIS also enabled analysis of the water quantity and quality impacts of projected growth on a baseline basis, assuming continuation of existing stormwater management practices. Water quality loadings to individual reservoirs and to the stream system could be readily demonstrated. Because overenrichment of the reservoirs was so crucial, researchers could estimate phosphorus and nitrogen loadings from projected development assuming existing stormwater practices, even on a municipality by municipality basis.

New Jersey Atlantic Coastal Drainage Study

Background

The third study considered a much larger coastal watershed in New Jersey (see Figure 15). The New Jersey

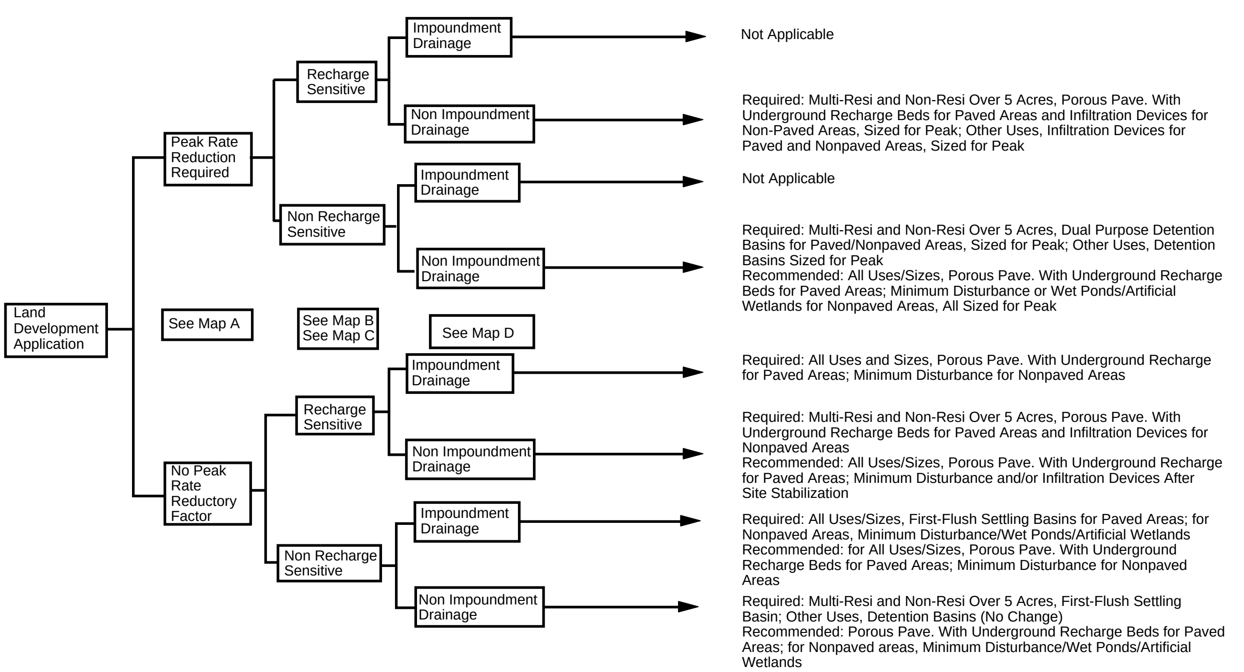


Figure 14. BMP selection methodology used with the GIS database in the Neshaminy basin modeling study.

Atlantic Coastal Drainage (ACD) includes an area of 2,086 square miles, with barrier islands (50 square miles), wetlands/bays/estuaries (285 square miles), and a unique scrubby pitch pine-cedar forest, known as the Pine Barrens, largely covering the 1,750 square miles of mainland interior (see Figure 16). This flat coastal plain comprises a series of unconsolidated sedimentary deposits of sand, marl, and clay, which increase in thickness toward the coastline. Over the past 16,000 years, as the ocean level has risen, the water's edge has progressed inland to its present position. Ocean currents and upland erosion and deposition have created a long, narrow series of barrier islands that absorb the energy of ocean storms and buffer the estuary habitats from the scour of waves and currents. Between the mainland and barrier islands are embayments and estuaries of different sizes and configurations. Inland erosion and marine sediments have gradually filled many of these areas, creating extensive wetlands (17).

In this ACD region, new land development and population growth have caused significant degradation of water quality from an increase in both point source and NPS pollution. Although the array of pollutants is ominously broad, increased nitrogen and phosphorus loadings have resulted in enrichment of back bays, estuaries, and nearshore waters, contributing to algal blooms, declining finfish and shellfish populations, diminished recreational

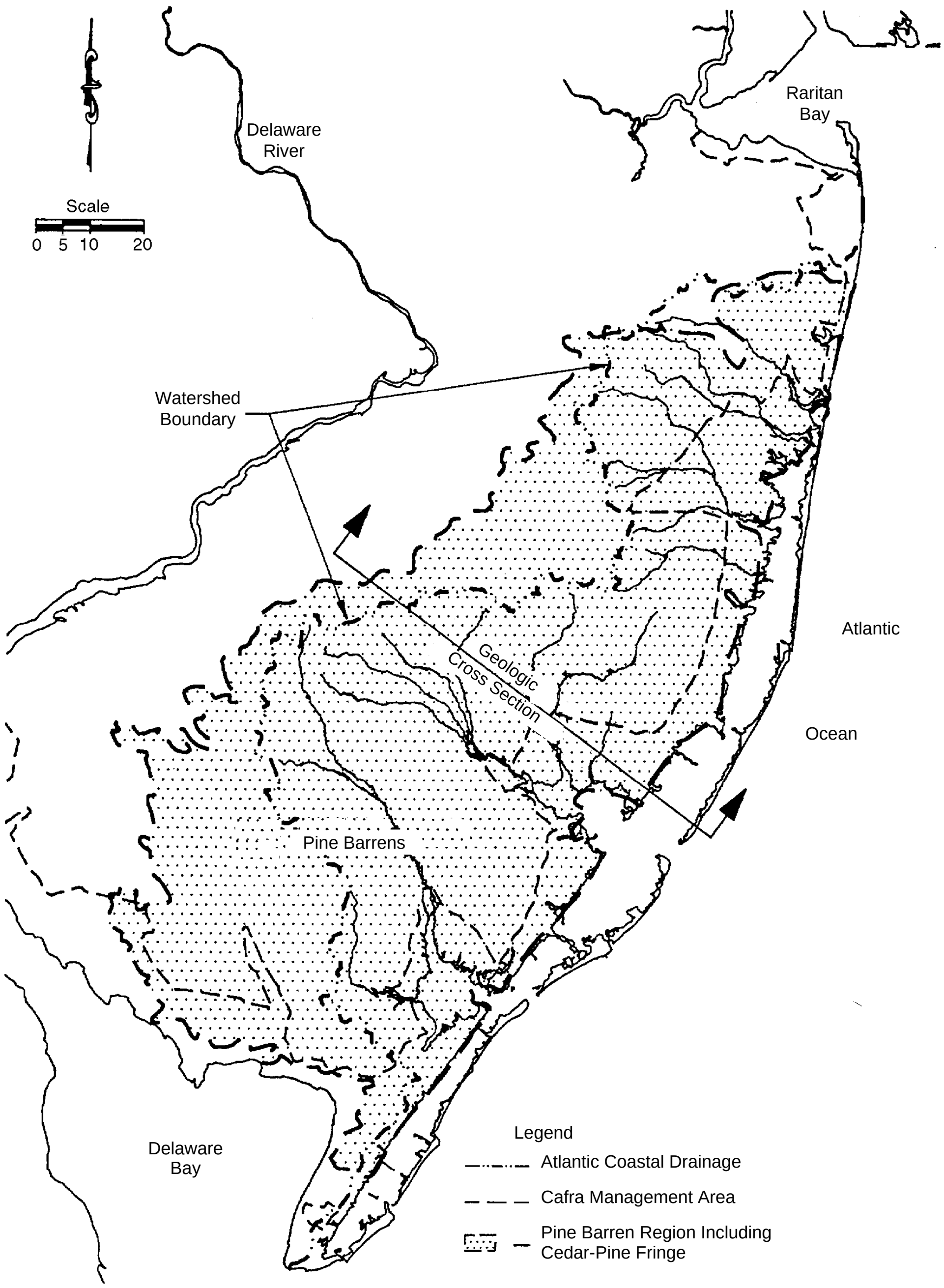


Figure 15. The ACD of New Jersey includes approximately 2,000 square miles of land area from the Manasquan River to Cape May.

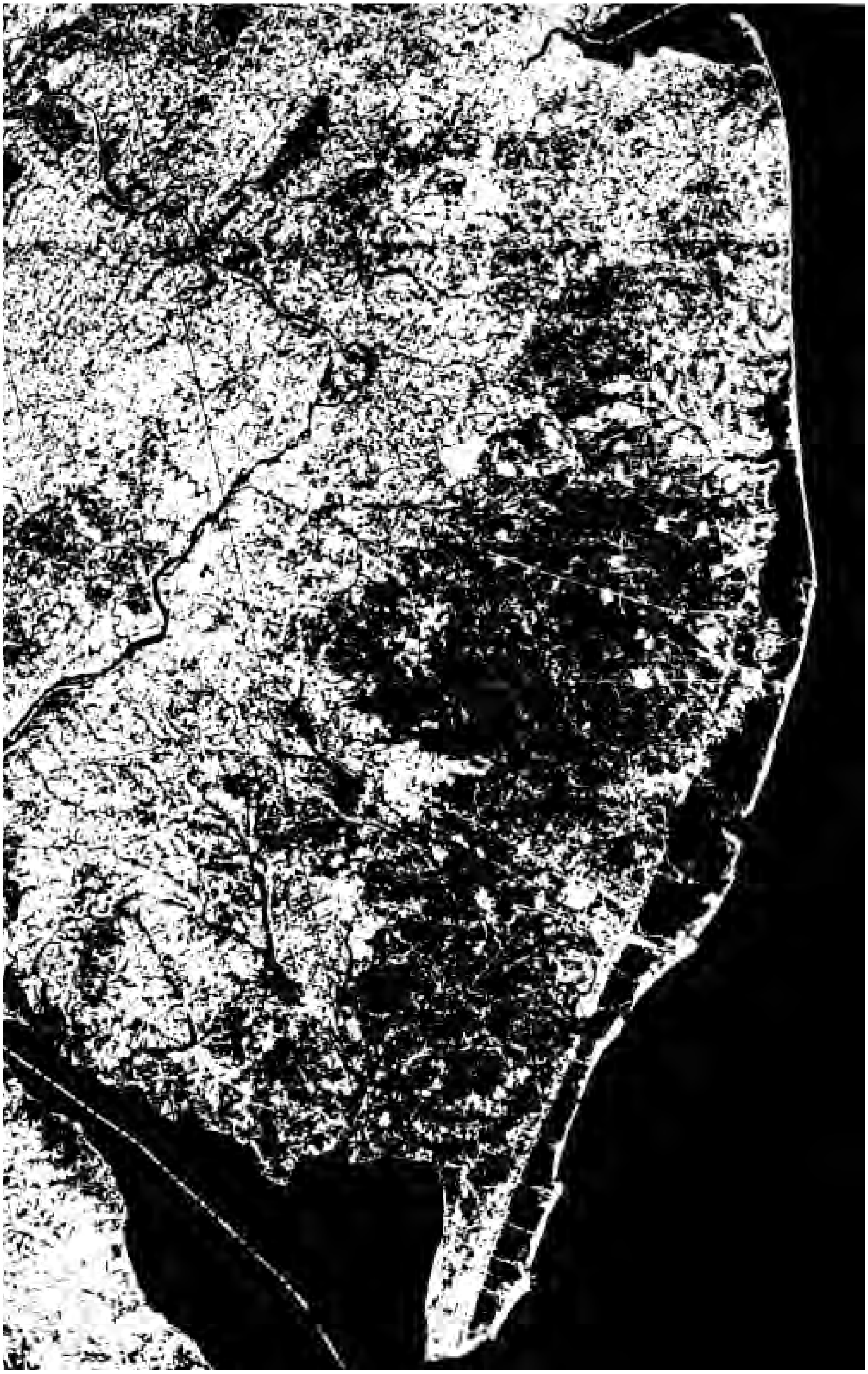


Figure 16. Aerial photograph of New Jersey illustrating the Barrier Islands and estuary system situated along the Atlantic coast.

opportunities, and a variety of other problems (18). A major source of these nutrients is point source sewage treatment plants (STPs), but the effluent outfalls of almost all these STPs discharge into nearshore ocean waters beyond the barrier islands. Thus, NPS pollutants almost totally dominate the water quality in the estuaries and back bays (19, 20).

These NPS pollutants, which rain scours from the land surface and flushes into coastal waters with each rainfall, comprise a largely unmeasured and unmanaged flux of contaminants. Prior research on coastal water quality has given considerable attention to NPS pollution generated from paved or impervious surfaces, particularly roadways and parking lots where hydrocarbons, metals, suspended solids, biologic oxygen demand (BOD), and other pollutants have been measured.

Although these NPS pollutants are certainly of concern in New Jersey's coastal waters, the enrichment issue has led to a focus on NPS pollution produced when creating large areas of pervious and heavily maintained landscape, such as lawns and other landscaped areas, in the sandy soil context of the coastal area. Typically, significant quantities of fertilizer and other chemicals, which are applied on these new pervious surfaces, are naturally low in nutrients. Although a modest portion of the applied fertilizer runs off directly into surface waters, larger quantities of soluble pollutants, such as nitrates and herbicides, quickly percolate down through the sandy soil, then move rapidly as interflow to the estuary system.

In this coastal drainage of unconsolidated sediments, the hydrologic cycle differs from inland watersheds. Of the 45-inch average annual rainfall, only a small fraction (2.5 inches per year) becomes direct runoff, with the balance rapidly infiltrating into the sand strata (21). Most of the infiltration that reaches the ground water (20 inches per year) discharges to surface streams (17 inches per year) within a few hours following rainfall, producing a lagging and attenuated hydrograph. This rapid infiltration, combined with the sand texture of the soil, has a major bearing on the water quality implications of new land development. Thus, urbanization of coastal regions has dramatically altered hydrologic response, with every square foot of new impervious surface converting what had been approximately 41.5 inches of infiltration into direct runoff to bays and estuaries, with a turbid soup of NPS pollutants.

Even in areas that have maintained infiltration, the coastal soils do not remove NPS pollutants as efficiently as other areas of New Jersey that overlie consolidated formations with heavier clay soils. These soils provide a much more thorough removal of NPS pollutants through physical, chemical, and biologic processes, as rainfall percolates through the soil mantle.

With development of coastal areas, increased impervious areas and changing flow pathways (inlets and storm sewers) convey nonpoint pollutants introduced by development (from both pervious and impervious surfaces) directly to the coastal waters. In addition, freshwater recharge to the underlying aquifer decreases with the increase in impervious surfaces, with resulting increases in saltwater intrusion into the sand aquifers and contamination of ground-water supply wells along the coast. Further compounding the loss of the stormwater for ground-water recharge are increased ground-water withdrawals necessary for new water supply. In sum, urban growth within the ACD, with its 1.13 million permanent residents (and still growing) and an additional 1.5 million summer tourists, has dramatically altered the natural drainage system (and landscape) in a way that significantly increases the discharge of NPS pollutants (22).

GIS Approach

New Jersey's Department of Environmental Protection already had developed a computerized GIS system (ARC/INFO) for environmental analysis and resource planning, so this study aimed to use existing GIS work and to refine this GIS system. Although data files for municipal boundaries, watershed areas, and a variety of other factors already existed, land use/land cover data had not been developed and constituted a major work task. The subsequent land use/land cover file included the entire 2,000 square miles of the ACD, but this focused on the urbanized area (212 square miles) that occupied about 11 percent of the coastal fringe. The end product was a polygon file that described about 2,500 polygons of urban/suburban land, each averaging about 0.1 square miles (see Figure 17).

Using aerial photographs combined with USGS base maps and extensive field reconnaissance, each polygon was classified by:

- Land use type.
- Percentage of impervious cover and maintained areas.
- Degree of maintenance (fertilization) being provided to these maintained areas.

Although classifying land use type and extent of impervious cover/maintained areas was a relatively straightforward evaluation process (rated within one of 11 categories by percentage, 0 to 5 percent, and so forth), the third variable, degree of maintenance, required special treatment and data development procedures. Degree of maintenance was translated into high, medium, and low categories, with high maintenance exemplified by golf courses or other intensively maintained areas. Medium maintenance assumed chemical application rates comparable with those recommended by Rutgers University state agronomists. Finally, low maintenance was typified by a wooded or otherwise naturally vegetated

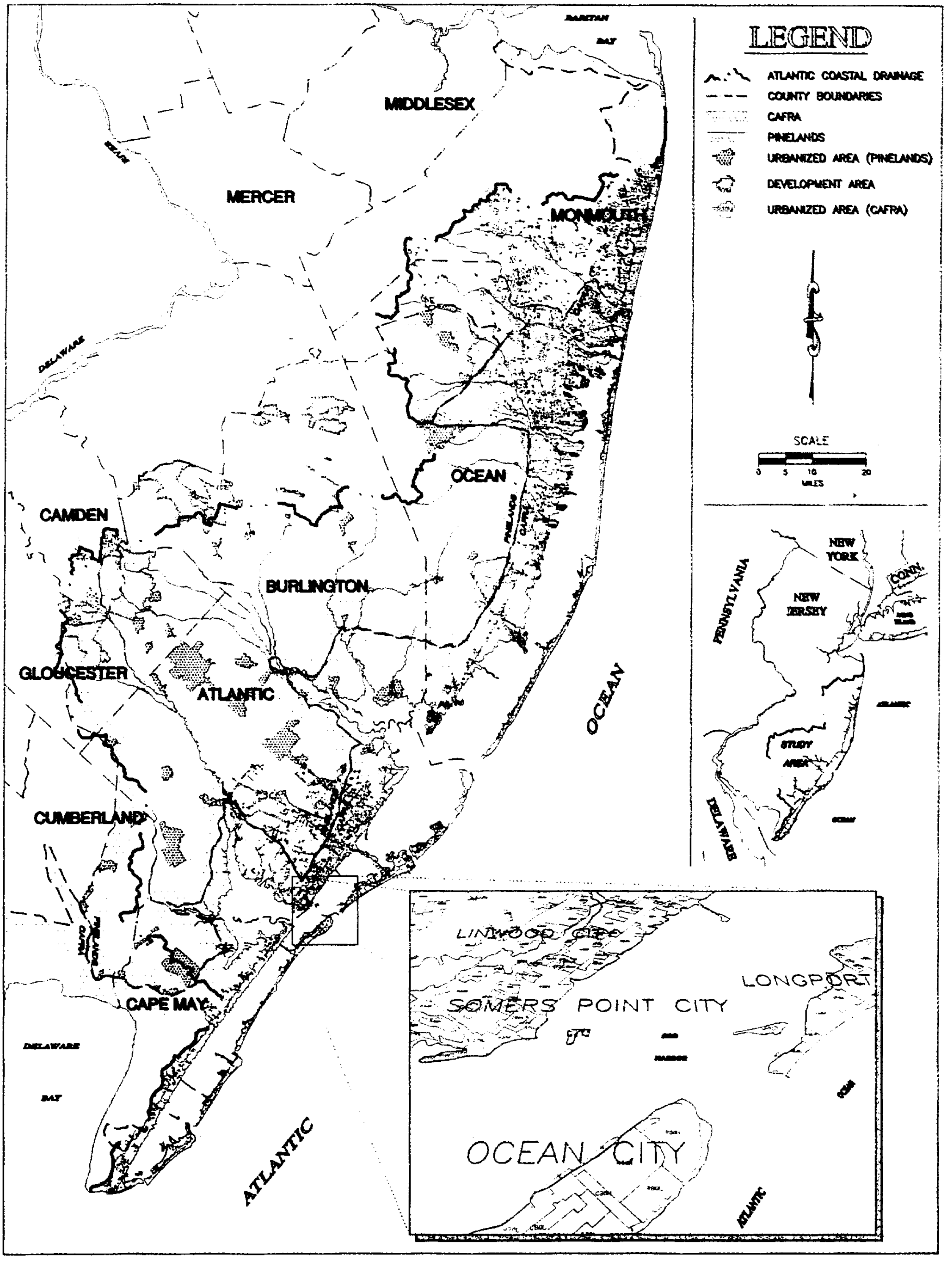


Figure 17. Urban land use polygons digitized for the New Jersey coastal drainage. The 2,500 polygons shown cover approximately 212 square miles (11 percent) of the ACD area of 2,000 square miles.

lot and assumed little or no regular chemical application. Research staff executed considerable field reconnaissance to objectify this judgment-based rating technique (see Figure 18).

Nonpoint Source Analysis

Because the drainage is almost entirely estuarine, the hydrologic aspects in this study were almost irrelevant except as a tool to describe the pollutant transport process. Such coastal drainage systems do not allow the measurement of hydrographs and chemographs (see Figure 6), except on inland riverine segments or selected infrastructure points of discharge (storm sewer outfalls). Thus, the NPS analysis focused on the pollutant production and transport process, especially the nutrients applied to the maintained landscapes, which are a major part of coastal urbanization.

This study required a great deal of effort to produce an index table relating urban cover characteristics (percentage impervious, amount of chemical application) to NPS production potential. For each of the 2,500 urban land polygons GIS described, estimates of the NPS loading for a number of pollutants were generated. Potential loadings were then aggregated by subwatershed. Total NPS loadings could then be compared with point source loadings for the entire coastal drainage. Interestingly, the NPS loading dominated water quality in the estuarine drainage while the point sources, discharged by ocean outfalls to nearshore waters beyond the barrier islands, were the major source of nutrient pollution in this portion of the coastal environment (see Figure 19).

Given the estimates of NPS pollution, the major question involves how to control or reduce these loads. The suitability of selected BMPs for the reduction/prevention of pollutant generation was then evaluated and spatially identified within the drainage (see Figure 20). This figure evaluated the use of constructed wetland systems as a structural measure. That is, GIS allowed state regulators to identify not only what works best in terms of water quality protection measures, but also where these methods could work successfully. This analysis was driven by a detailed evaluation of the combinations of natural conditions GIS identified within the study area. For example, certain BMPs can be applied on soils that have a certain set of characteristics (permeability, depth to seasonal high water table) and that are presently in a given land cover and planned for urbanization.

GIS also aided in evaluating alternative BMP techniques, including reduction in nutrient applications and land management BMPs such as elimination of artificial landscapes, again using a combination of natural features and land use patterns (see Figure 21). The result of this analysis considered the relative proximity of urban land uses to the coastal waters as significantly increasing the potential for NPS transport. State regulatory programs establish minimum setback criteria for development in sensitive areas, and these criteria may be modified to consider pollutant production potential based on GIS delineation of pollutant production.

New Jersey has been striving to develop NPS management programs for coastal areas to reduce existing sources of pollution as well as prevent the creation of

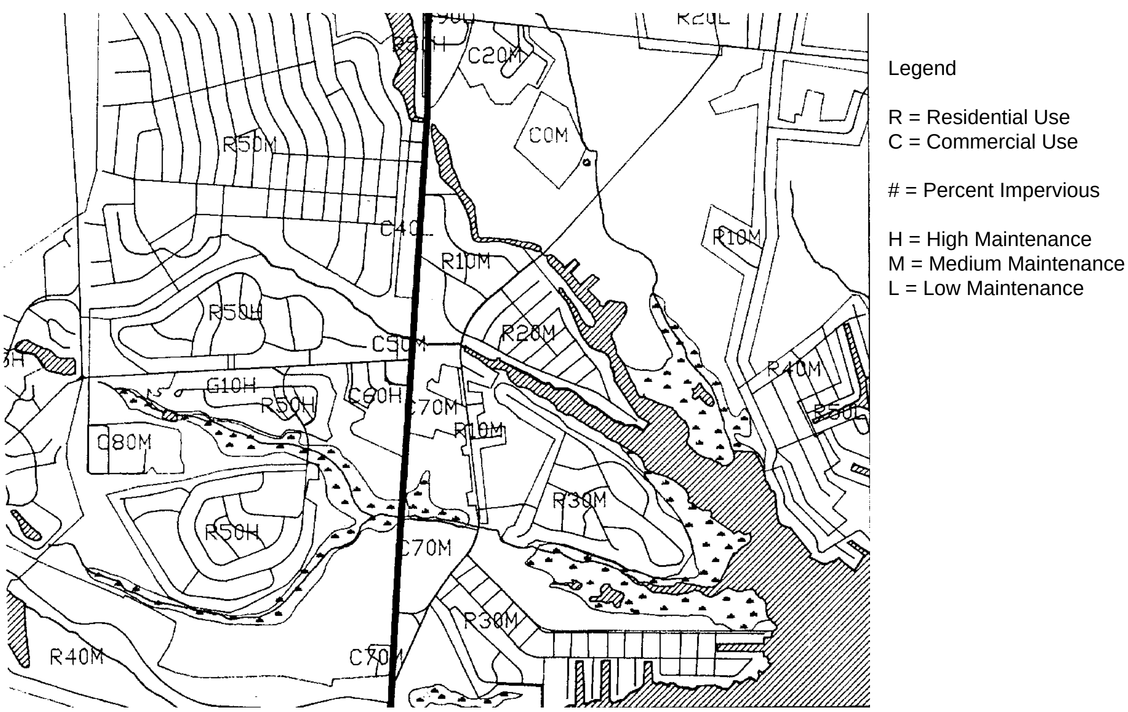


Figure 18. Classification of urban polygons by land use, percentage impervious cover, and degree of land fertilization.

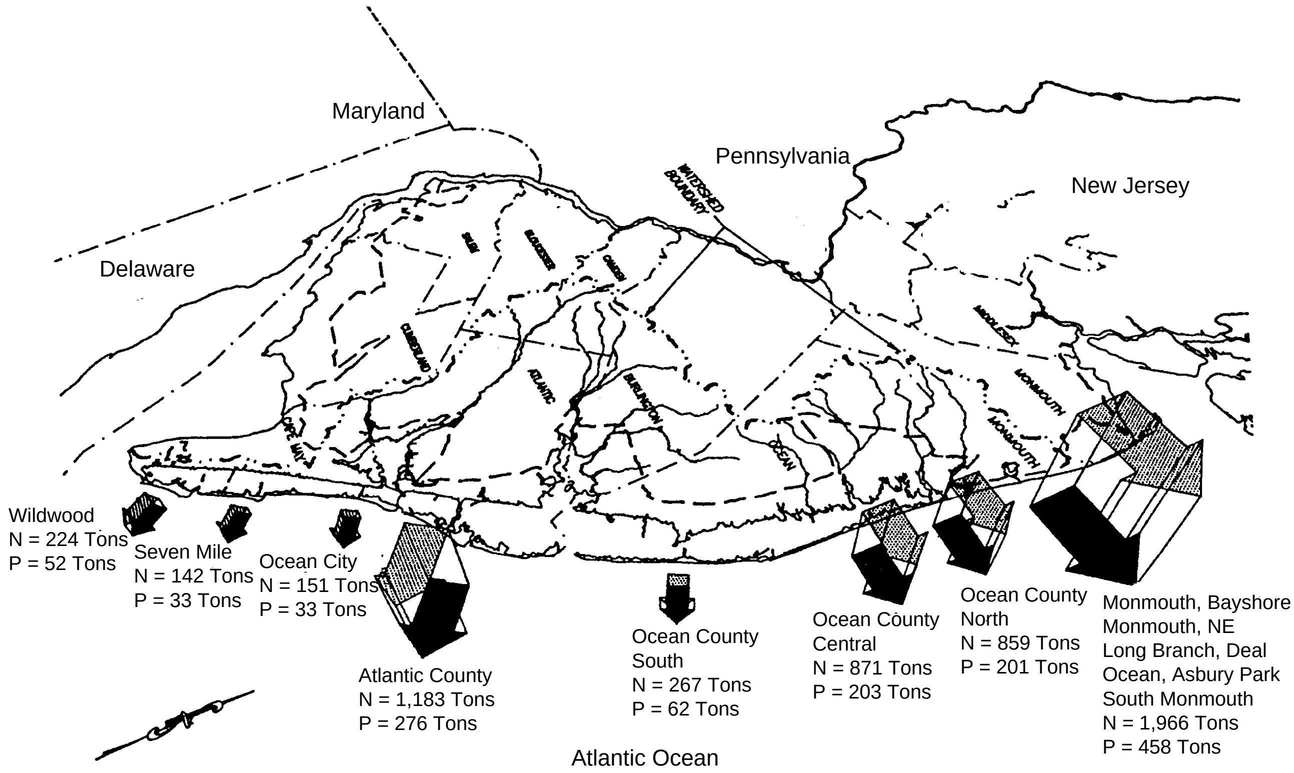


Figure 19. Point and NPS discharges to the ACD. Data shown are in tons of TP and NO_3-N per year.

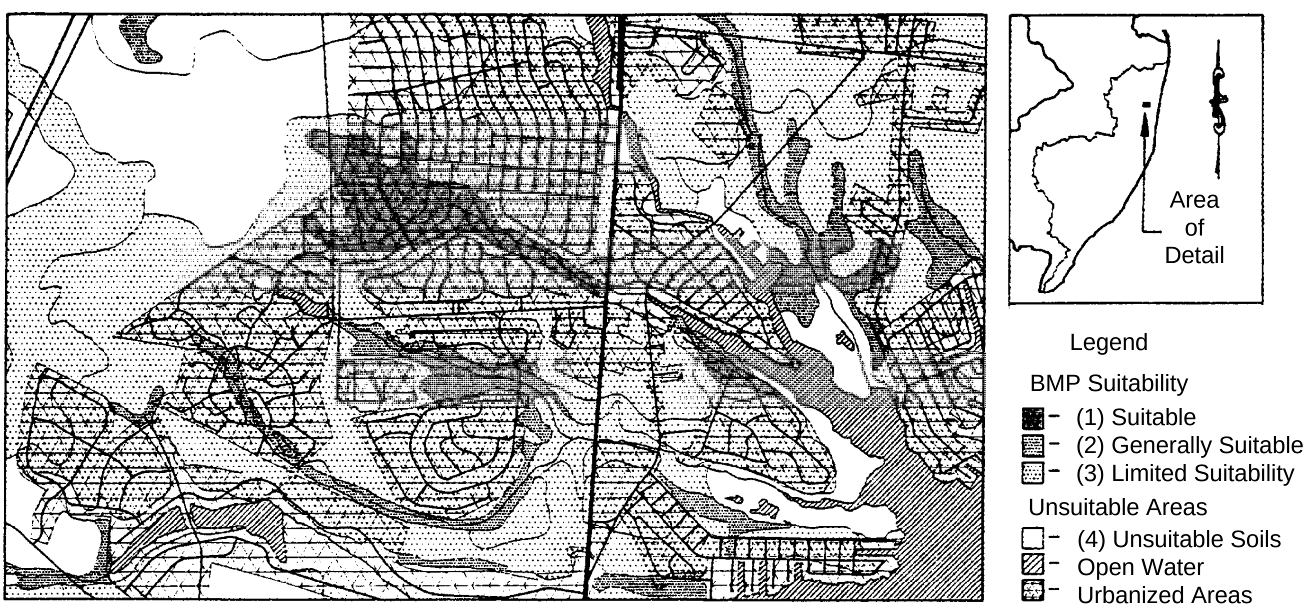


Figure 20. BMP analysis using GIS. Files consider soil suitability, current vegetative cover, and BMP criteria on vacant and developed land parcels.

pollution. As most regulatory agencies have discovered, NPS management programs can be difficult to implement, especially when confronting issues of land use management. To substantiate the need for new management programs amidst these controversies, the ability to document causal linkages (i.e., to generate data and statistics that make the case for NPS pollutant generators and resultant water quality degradation) is very important. The need for documentation of various types is especially great given the less than perfect data record of water quality in coastal and other waters. All of these factors come together to make the value of a GIS system for water quality management very real.

Conclusion

This GIS-driven analysis indicates that NPS pollutants, especially the nutrients phosphorus and nitrogen, generated from fertilized fields or maintained landscapes surrounding new residential, commercial, and other types of development in drainage systems, contribute significantly to water quality degradation. In effect, the particulate-associated phosphorus and the soluble nitrates serve as surrogates for the full spectrum of NPS pollutants that each rainfall washes from the land. A comprehensive water quality management program must include structural measures to remove pollutants this runoff conveys, as well as management of the contributing landscape to reduce (and perhaps eliminate) the application of these chemicals within the drainage.

In planning new development, management actions should occur on a variety of levels or tiers. On an areawide basis, growth should proceed (with guidance and management) in a manner that would reduce total pollutant discharges; therefore, the total amount of maintained area being created should be as concentrated as possible. On the more site-specific level, measures and construction techniques that reduce the quantity of pollutants generated are essential. Required development guidelines must include, but not be limited to:

- Prevention of excessive site disturbance and ongoing site maintenance (described as a policy of minimum disturbance and minimum maintenance).
- Use of special materials for reduction of stormwater runoff (porous pavement and ground-water recharge).
- Use of stormwater treatment systems (water quality detention basins, artificial wetlands).

In sum, the regulatory framework must contain both "how to build" guidelines, as well as "where not to build" guidelines. GIS can be a powerful tool in both of these processes.

While inland lakes serve as nutrient traps for these NPS pollutants, perhaps the greatest potential impact is the gradual process of excessively enriching our coastal waters. As population continues to migrate to coastal areas, the importance of protecting this fragile ecosystem

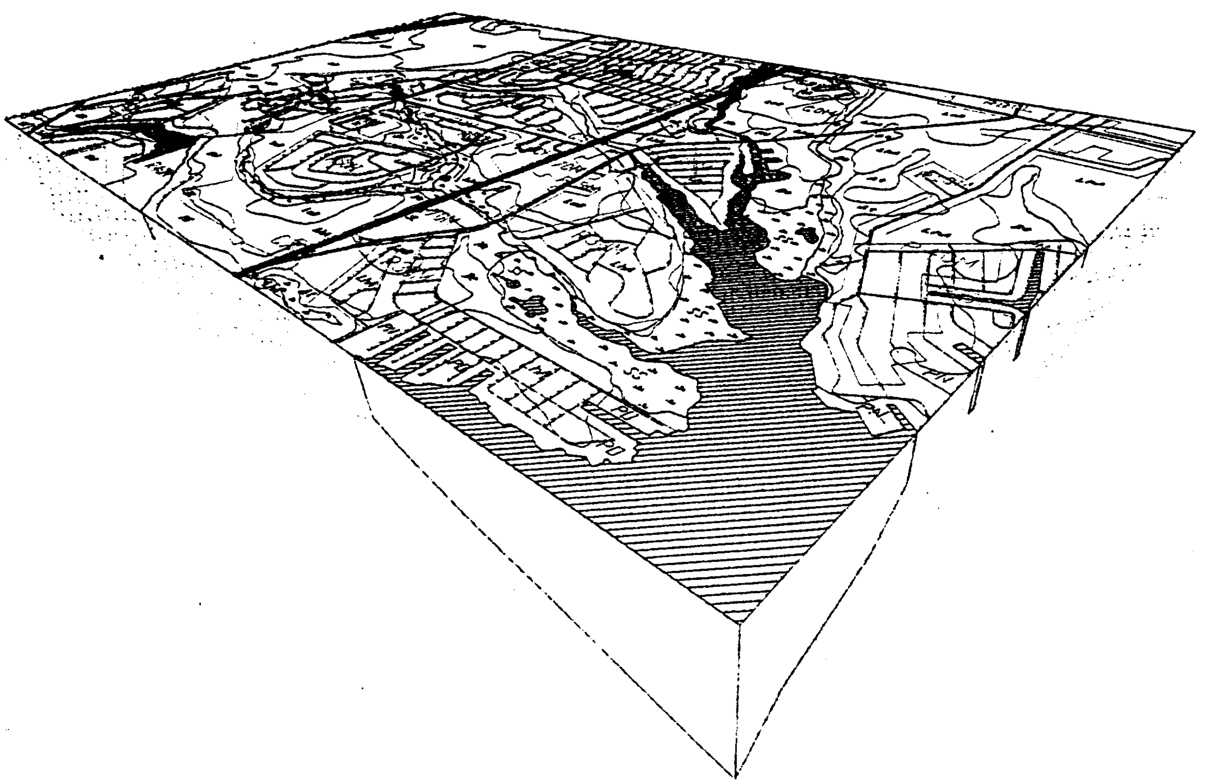


Figure 21. For certain regulatory criteria, the proximity of land uses to the water's edge was a consideration in BMP selection.

increases. The pollution that new land development generates, including the discharge of point source wastes, should not be allowed to enter coastal waters; it should not be allowed to destroy the natural balance that exists between land and water. The concept of stormwater management takes on an entirely different meaning when viewed as one of the basic mechanisms of this NPS pollution transport. For centuries, engineering of the shoreline has intensively focused on protecting human developments from the ravages of ocean storms. Now, however, the converse seems to be emerging: ocean waters need protection from the impacts of human development.

References

1. Bliss, N., T.H. Cahill, E.B. MacDougall, and C.A. Staub. 1975. Land resource measurement for water quality analysis. Chadds Ford, PA: Tri-County Conservancy of the Brandywine.
2. Cahill, T.H., R.W. Pierson, Jr., and B.R. Cohen. 1978. The evaluation of best management practices for the reduction of diffuse pollutants in an agricultural watershed. In: Lohr, R.C., ed. Best management practices for agriculture and silviculture. Ann Arbor, MI: Ann Arbor Science.
3. Cahill, T.H., R.W. Pierson, Jr., and B.R. Cohen. 1979. Nonpoint source model calibration in the Honey Creek watershed. #R-805421-01. Athens, GA: U.S. EPA Environmental Research Laboratory.
4. Cahill, T.H. 1980. The analysis of relationships between land use and water quality in the Lake Erie basin. Burlington, Ontario: International Association of Great Lakes Research.
5. Cahill, T.H., J. McGuire, and C. Smith. 1993. Hydrologic and water quality modeling with geographic information systems. Proceedings of the Symposium on Geographic Information Systems and Water Resources, AWRA, Mobile, AL (March).
6. U.S. EPA. 1991. Remote sensing and GIS applications to nonpoint source planning. Proceedings of the U.S. EPA Workshop for Region 5 and Northeast Illinois Planning Commission, Chicago, IL (October 1990).
7. Baker, D.B., and J.W. Kramer. 1973. Phosphorus sources and transport in an agricultural basin of Lake Erie. Proceedings of the 16th Conference, Great Lakes Research, Ann Arbor, MI (September).
8. Cahill, T.H., and T.R. Hammer. 1976. Phosphate transport in river basins. Proceedings of the International Joint Committee on Fluvial Transport Workshop, Kitchener, Ontario (October).
9. Cahill Associates. 1989. Stormwater management in the New Jersey coastal zone. Trenton, NJ: State of New Jersey, Department of Environmental Protection, Division of Coastal Resources.
10. Delaware Riverkeeper. 1993. Upper Perkiomen Creek watershed water quality management plan. Lambertville, NJ: Delaware Riverkeeper/Watershed Association of the Delaware River.
11. U.S. EPA. 1993. Guidance specifying management measures for sources of nonpoint pollution in coastal waters. EPA/840/B-92/002. Washington, DC.
12. Soil Conservation Commission. 1982. TR-20, project formulation—Hydrology. Technical Release No. 20. PB83-223768. Landham, MD: Soil Conservation Service.
13. Cahill, T.H., M.C. Adams, and W.R. Horner. 1990. The use of porous paving for groundwater recharge in stormwater management systems. Presented at the 1988 Floodplain/Stormwater Management Symposium, State College, PA (October).
14. Soil Conservation Commission. 1974. Universal soil loss equation. Technical notes, Conservation Agronomy No. 32. Portland, OR: West Technical Service Center.
15. Cahill, T.H., M. Adams, S. Remalie, and C. Smith. 1988. The hydrology of flood flow in the Neshaminy Creek basin, Pennsylvania. Jamison, PA: The Neshaminy Water Resources Authority (May).
16. BCPC. 1992. Neshaminy Creek watershed stormwater management plan, Vol. 1: Policy document, and Vol. II: Plan implementation. Doylestown, PA: Bucks County Planning Commission (January).
17. Clark, J. 1977. Coastal ecosystems: Ecological considerations for management of the coastal zone. Washington, DC: The Conservation Foundation.
18. New Jersey Department of Environmental Protection. 1988. The state of the ocean: A report by the blue ribbon panel in ocean incidents. Trenton, NJ.
19. Cahill, T.H., M. Adams, C.L. Smith, and J.S. McGuire. 1991. GIS analysis of nonpoint source pollution in the New Jersey coastal zone. With Whitney, S., and S. Halsey, New Jersey Department of Environmental Protection, Division of Coastal Resources. Presented at the National Conference on Integrated Water Information Management, Atlantic City, NJ (August).
20. Cahill, T.H., M. Adams, C.L. Smith, and J.S. McGuire. 1991. Living on the edge: Environmental quality in the coastal zone. With Whitney, S., and S. Halsey, New Jersey Department of Environmental Protection, Division of Coastal Resources. Presented at the International Conference on Integrated Stormwater Management, Singapore (July).
21. Martin, M. 1989. Ground-water flow in the New Jersey coastal plain. Open File Report 87-528. West Trenton, NJ: U.S. Geological Survey.
22. National Oceanic and Atmospheric Administration. 1989. Selected characteristics in coastal states, 1980-2000. Rockville, MD: U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Strategic Assessment Branch, Ocean Assessment Division (October).

GIS Standards for Environmental Restoration and Compliance

Bobby G. Carpenter, P.E.
The CADD/GIS Technology Center
Waterways Experiment Station
Army Engineer Research and Development Center
Vicksburg, Mississippi

The CADD/GIS Technology Center (formerly Tri-Service CADD/GIS Technology Center) was established at the Information Technology Laboratory (ITL), U.S. Army Engineer Waterways Experiment Station (WES), Vicksburg, Mississippi in October 1992. The CADD/GIS Center's primary mission is to serve as a multi-service vehicle to set computer-aided design and drafting (CADD) and geographic information system (GIS) standards; coordinate CADD/GIS facilities systems within the Department of Defense (DoD); promote CADD/GIS system integration; support centralized CADD/GIS hardware and software acquisition; and provide assistance for the installation, training, operation, and maintenance of CADD and GIS systems.

The term *geospatial data* refers to data that can be referenced to a specific geographic location on the Earth. The specific geographic location can be depicted by a graphic *feature* on a map or drawing (e.g., a building, monitoring well, road, location where an environmental sample is collected, etc.).

Environmental restoration activities involve the investigations and cleanup efforts associated with the identification and removal of chemical or radioactive contaminants present in the soil, groundwater, surface water, or sediment. The most common environmental restoration programs within the DoD include the Installation Restoration Program (IRP) (involves all Air Force, Army, and Navy installations) and the Defense Environmental Restoration Program - Formerly Used Defense Sites (DERP-FUDS) (administered by USACE).

Environmental compliance activities involve conducting everyday business practices in a manner which complies with U.S. Environmental Protection Agency (EPA) and state/local environmental regulatory agency laws and regulations. Environmental compliance activities include the management or removal of "toxic substances" (asbestos containing materials, lead paint, PCBs), proper handling of hazardous materiels, proper handling and disposal of hazardous wastes, monitoring and management of regulated storage tanks, monitoring and management of permitted surface water discharges and permitted air emissions.

Why Do We Need Data Standards? The collection, storage, management, and analysis of geospatial data are critical components of environmental restoration and compliance activities. Geospatial data can be stored in a number of ways (i.e., paper, microfilm, and/or electronically) which may not be readily accessible and usable, or easily shared with, or reported to others. CADD and GIS technology can provide cost-effective and efficient tools to apply and manage such data. However, careful planning and the use of consistent data storage and CADD/GIS system implementation standards are necessary to achieve the full potential offered by CADD and GIS technology.

Tri-Service Spatial Data Standards (TSSDS) and Tri-Service Facility Management Standards (TSFMS) Development. One of the CADD/GIS Center's (http://tsc.wes.army.mil) major initiatives has been development of the TSSDS and TSFMS. This project involves the development of graphic and non-graphic standards for GIS and facility management implementations at Air Force, Army, and Navy installations, and Army Corps of Engineers Civil Works activities. The TSSDS/TSFMS are the only "nonproprietary" standard designed for use with the predominant commercially available "off-the-shelf" GIS (e.g., ESRI ARC/INFO & ArcView; Intergraph MGE & GeoMedia; AutoDesk AutoCAD, Map and World; and Bentley MicroStation & GeoGraphics) and relational database software (e.g., Oracle & Microsoft Access). This design, in conjunction with it's universal coverage, have propelled the TSSDS into the standard for GIS implementations throughout the DoD, and well as the De Facto standard for GIS implementations in other Federal, State, and Local Government organizations; public utilities; and private industry throughout the United States and the World.

The TSSDS/TSFMS are distributed via CD-ROM and the Internet (http://tsc.wes.army.mil). A "User Friendly" interactive Microsoft Windows based software application installs the TSSDS on desktop computers and networks, provides viewing and printing capability, and generates SQL code for construction of the GIS database. The CADD/GIS Center annually updates and expands the TSSDS and TSFMS data coverages. Release 1.80 of the TSSDS/TSFMS was completed in February 1999.

Contributors and Coordination. The TSSDS and TSFMS have been developed based on input from various technical experts; review and analysis of existing working DoD and state GIS; review and analysis of various existing database management systems used throughout DoD and the federal government; and content contributions from federal, state, local, and private sector sources. The CADD/GIS Center has organized Field Working Groups (FWGs) whose membership is composed of subject matter, CAD, and GIS technical experts to assist in the development of the Tri-Service Standards and other CADD/GIS projects. The CADD/GIS Center's Environmental FWG has been very active is assisting in the development of the TSSDS and TSFMS.

The CADD/GIS Center is coordinating development of the TSSDS/TSFMS with other DoD and Federal standards initiatives such as the Defense Environmental Security Corporate Information Management (DESCIM) program, the Federal Geographic Data Committee (FGDC), the Defense Information Standards Agency (DISA), the National Imagery and Mapping Agency (NIMA), and the Environmental Protection Agency (EPA).

Some of the specific DoD and Federal initiatives contributing to the environmental restoration and compliance content of the TSSDS/TSFMS include: (1) Air Force “Environmental Restoration Program Information Management System”, (ERPIMS) (formerly called IRPIMS); (2) USACE Alaska District “Environmental Data Management System” (EDMS); (3) Army Environmental Center (AEC) “Installation Restoration Data Management Information System” (IRDMIS); (4) Southwest Division Naval Facilities Engineering Command, “Navy Environmental Data Transfer Standard” (NEDTS); (5) USACE “Formerly Used Defense Site (FUDS) Database”; (6) Air Force Aeronautical Systems Center (ASC) and USACE District, Louisville “Draft System Specification for the Technical Data Management System”; (7) DESCIM Cleanup, Explosive Safety, and other data modeling work groups; (8) Edwards AFB, Patuxent River Naval Air Station, and other DoD GIS; (9) CADD/GIS Center's Environmental FWG, (10) Environmental Protection Agency; (11) FGDC Facilities Standards Working Group, and (12) EPA's Environmental Data Registry.

TSSDS/TSFMS Data Model Structure. Both graphic (i.e., symbols, text fonts, line styles/types, and level/layer schemas) and nongraphic (e.g., database attribute tables and domains) geospatial data requirements are addressed in the TSSDS/TSFMS. The TSSDS/TSFMS data model consists of five basic levels of hierarchy: Entity Sets, Entity Classes, Entity Types (includes Entities) (TSSDS only), Attribute Tables, and Domain Tables.

Entity Sets (or Themes) are broad groupings of features and related data. The TSSDS/TSFMS structure currently includes the following twenty-five themes: (1) Auditory, (2) Boundary, (3) Buildings, (4) Cadastre, (5) Climate, (6) Common, (7) Communications, (8) Cultural, (9) Demographics, (10) Environmental Hazards, (11) Ecology, (12) Fauna, (13) Flora, (14) Geodesy, (15) Geology, (16) Hydrography, (17) Improvements, (18) Landform, (19) Land Status, (20) Military Operations, (21) Olfactory, (22) Soil, (23) Transportation, (24) Utilities, (25) and Visual.

Entity Classes are logical groupings of features and data within an Entity Set for data management purposes.

The TSSDS Entity Classes contain logical groupings of "real-world", geographically referenced (geospatial) features (entity types & entities) with related (graphic) database attribute tables. Each Entity Class consists of a separate map or drawing file (i.e., category or design file in MGE; workspace in ARC/INFO; design file in MicroStation; drawing file in AutoCAD). The current TSSDS Entity Classes in the Environmental Hazards Entity Set include: (1) characterization, (2) surface water pollution, (3) munitions remediation, (4) emergency preparedness (spills, etc.), (5) general, (6) groundwater pollution, (7) hazardous materiels/hazardous waste, (8) munitions materiel/munitions waste, (9) pollution remediation, (10) regulated tanks, (11) sediment pollution, (12) sites, (13) building environmental concerns, (14) solid waste, (15) air pollution, and (16) soil pollution.

The TSFMS Entity Classes contain logical groupings of (non-graphic) database attribute tables which contain temporal or event data for specific "business" activities or functions. The TSFMS Classes in the Environmental Hazards Entity Set include: (1) hazardous materiel management, (2) munitions waste management, (3) asbestos containing materiel management, (4) surface water discharge management, (5) hazardous waste management, (6) regulated tank management, (7) PCB (polychlorinated biphenyl) management, (8) lead paint management, (9) indoor air management, (10) field measurements management, (11) remediation management, (12) environmental management, (13) munitions materiel management, and (14) air quality management.

Entity Types are a grouping or collection of like, or similar, features (entities) which appear graphically on a map or drawing. Each entity type has an associated attribute table. Entities can be represented as one of the following three categories:

- Boundary (Polygon) - A line string (or group of arcs) which forms the perimeter of an area. An example would be the boundary of a lake.

- Point - A single point representing the geographical location of a feature; e.g., a well. Points are normally represented on a map by a symbol. The TSSDS provide symbols in the native formats of AutoCAD, MicroStation, and ARC/INFO.

- String/Chain - A line or group of arcs.

An *Attribute Table* is a relational database table containing non-graphic, or attribute, information about an entity. Attribute Tables which are linked directly to a graphic entity and contain data directly related to that entity can be classified as "graphic" (i.e., TSSDS) attribute tables. Attribute Tables not directly linked to an entity but which contain data required for a "business process" or function, along with data and relationships linked through specific data field ids which may be queried for geospatial and relational analysis, can be classified as "nongraphic" (i.e., TSFMS) attribute tables.

Domain Tables contain lists of codes (i.e., permissible or valid values) for populating specific fields in the Attribute Tables; e.g., units of measure, material types, etc.

Join relationships are mechanisms by which relational databases link multiple records of a common attribute or item and provide access to the records through the use of queries. Join relationships are established in the TSSDS/TSFMS through the use of "Primary Key" attribute fields in a "parent" attribute table and "Foreign Key" attribute fields in related "child" attribute tables.

Integration of Approved FGDC Geospatial Data Standards into the TSSDS. Executive Order 12906, "Coordinating Data Acquisition and Access: The National Spatial Data Infrastructure" (NSDI), which was signed by the President on 11 April 1994, requires that all Federal agencies use the FGDC Metadata Standard to document new geospatial data and make them electronically accessible through the use of a National Geospatial Data Clearinghouse. Executive Order 12906 also assigned authority for the development of national geospatial data standards to the FGDC. The FGDC standards development program ensures that standards are created under an open consensus, with participation by non-federal and federal communities.

The FGDC geospatial data standards provides a “Logical Data Model” consisting of descriptive feature names (entity), attribute names, and domain names. However, this data model must be fully developed into a “Physical Data Model” before it can be implemented in a GIS. That is, all symbology (e.g., symbols, colors, fonts, line types); level/layer schemas; coverages; file table, attribute, and domain names which are compatible with commercially available GIS and relational database management systems must be developed. The TSSDS provides the “Physical Data Model” for implementation of the approved FGDC geospatial data standards in a GIS. The TSSDS has been designed to comply with the Spatial Data Transfer Standard (SDTS) data model, and has been updated to permit compliance with the recently revised FGDC Metadata Standard. Provisions of the FGDC Bathymetric Geospatial Standard (International Hydrographic Standard (IHO S-57)) were incorporated into the TSSDS Release 1.6. The FGDC Vegetation, Wetlands, and Soils standards have been incorporated into the CADD/GIS Center's TSSDS/TSFMS Release 1.8. In addition, two of the standards currently under development by the FGDC Facilities Working Group (Environmental Hazards Geospatial Standard and Utilities Geospatial Standard) originated from the TSSDS.

Planning Strategies for Siting Animal Confinement Facilities: The Integrated Use of Geographic Information Systems and Landscape Simulation Technologies

T. L. Cartlidge[1], B. Chamberlain[2], D. G. Pitt[1], M. Olson[3],
B. Halverson[2] and T. Harikrishnan[1]

[1]Department of Landscape Architecture, University of Minnesota, Minneapolis, MN
[2]Hoisington-Koegler Group, Minneapolis, MN
[3]Department of Landscape Architecture, The Pennsylvania State University, University Park, PA

Abstract

Through presentation of a case study, a series of planning strategies for siting animal confinement facilities in the rural landscape are compared and contrasted. The strategies use geographic information systems (GIS) technology to develop an environmental protection framework for a 31,000-hectare watershed in west central Minnesota. The framework is based on desires to maintain landscape integrity as reflected in enhanced biological diversity, conserved soil and improved water quality as well as to maintain neighborhood cohesion among farm and non-farm neighbors in the rural landscape. Design of the production landscape is compared and contrasted under three alternative scenarios: a) the use of Euclidean zoning; b) the use of overlay zoning; and c) the use of technical assistance to small-scale operators to enhance adoption of whole farm planning. Low elevation aerial oblique renderings are presented as a means of communicating to stakeholder groups the spatial organization and visual appearance of agriculture in the rural landscape following implementation of the strategies.

Introduction

The location and operation of animal confinement facilities in rural landscapes is an issue of large environmental concern in states having economic bases that include livestock production. In the recent gubernatorial elections in Minnesota, for example, all candidates were expected to voice their position on a legislatively proposed moratorium on animal confinement facility siting. The State's Environmental Quality Board is preparing a Generic Environmental Impact Statement (GEIS) on animal confinement facilities in Minnesota. Among other concerns, the

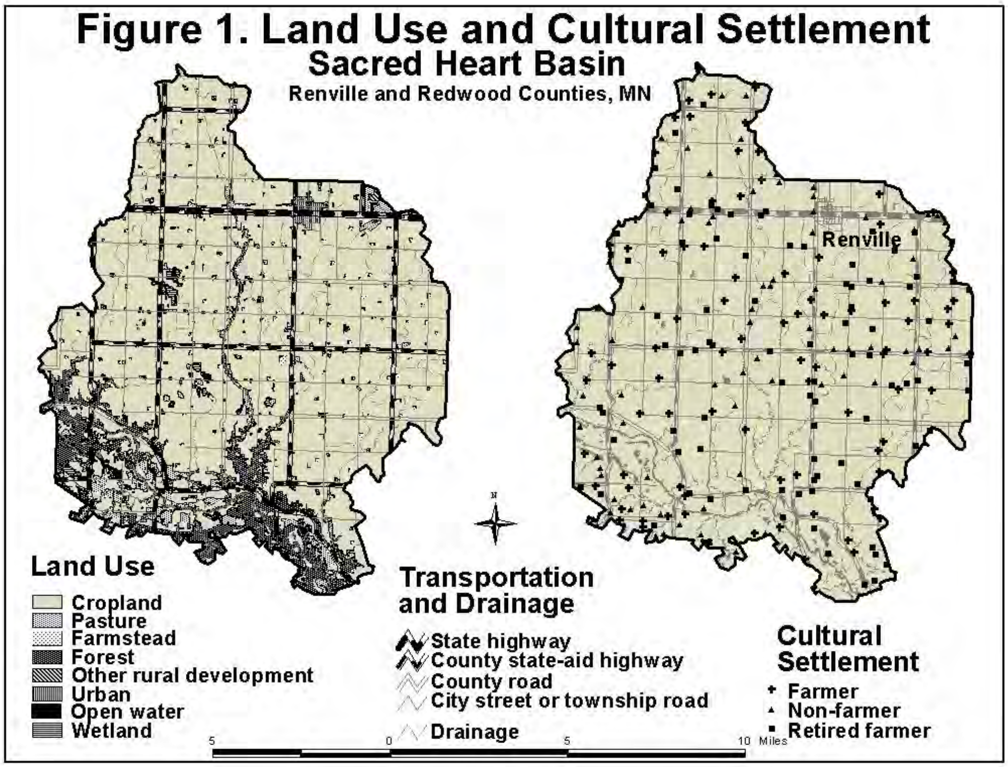
GEIS is examining the effects of animal confinement facilities on water quality, air quality, the structure of local economies, animal health, human health and the changing characteristics of rural communities. In addition, the GEIS will also examine the effectiveness and capability of local land use planning strategies in considering animal confinement facility siting.

The diversity of issues being considered in the GEIS is symptomatic of the complexity of the issues surrounding animal agriculture in the contemporary rural landscape. No longer can the focus rest solely on the impacts of production facilities on physical and biological characteristics of the environment. Evaluation of animal agricultural issues in the twenty-first century also requires assessment of social, economic and political concerns. The changing structure of American agriculture has affected the manner in which farmers conduct their business as well as the definition of who is a farmer. Scale of operation has increased significantly. Corporate structures of farm enterprises often break the traditional land-based agrarian ties of operators as land becomes little more than another factor of production. Rural community structure deteriorates as seemingly strangers operate land once cared for by trusted neighbors. Small township governments sometimes find themselves overpowered by the corporate structures with which they must often interact in issues related to animal agriculture.

The physical, biological, social economic and political issues associated with animal agriculture must be examined at the scale of the individual production facility, the immediate landscape context of the facility as well as the region within which the facility is located. Examining the spatial dimensions of these issues at the landscape and regional levels lends itself directly to the use of geographic information systems (GIS) technology. Planners attempting to resolve the complexities of animal agriculture issues can use GIS technology to integrate the diverse issues across space and spatial scales, and they can use the technology to develop sophisticated map representations of their findings and recommendations. The map representations can serve as a basis for creating realistic image simulations to present recommendations to various stakeholder groups.

The University of Minnesota Extension Service sponsored a Rural Landscape Project to demonstrate how local units of government could use geographic information systems (GIS) technology and image simulation technologies to enhance their abilities to plan for animal confinement facilities. The demonstration project was located in the 31,000-hectare Sacred

Heart drainage basin, a tributary of the Minnesota River in west central Minnesota. Figure 1 illustrates existing land use and cultural settlement patterns in the basin.



Objectives of the demonstration project were:

- To evaluate at the local level – in rural communities- the land management alternatives that can be used to sustain animal confinement agriculture in the rural landscape. This objective sought the identification of strategies to conserve soil, maintain water quality, enhance biological diversity, contribute to regional economic health, maintain the viability of individual farm enterprises and enhance the well being of people living and working in the landscape. The objective sought means of building and sustaining healthy ecological systems, healthy economies and healthy communities in the Sacred Heart basin.

- To broaden input and further discussion of critical issues related to animal agriculture and the rural landscape between producers and their neighbors, policy-makers and communities, and state and local governments.

- To help foster consensus on principles for sustaining the rural landscape of west central Minnesota.

The case study is proceeding in three phases. The first phase involved a series of in-depth interviews and focus group discussions with selected animal agriculture operators in west central Minnesota. Along with an examination of agricultural statistics for the region and a review of relevant literature, the interviews and discussions enabled the demonstration team to better understand the issues associated with animal agriculture operation in the region. The second phase of the case study, and the subject of this paper, developed a series of planning strategies to enable continuation of sustainable animal agriculture in one of the region's watersheds. In developing the strategies, plan view mapped representations of the designs were created. To better communicate the design strategies to different sets of stakeholders during a subsequent round of workshops, the strategies were also represented as hand drawn, low-elevation aerial oblique renderings. The renderings were prepared to offer stakeholders a sense of how the design strategies would affect spatial organization of agriculture in the rural landscape as well as engender a sense of the landscape's visual appearance following implementation of each strategy. The third phase, to be conducted in the fall of 1999, will engage a series of stakeholder groups in the region in conversation about the planning strategies. The workshops will involve use of the GIS mapped information as well as the aerial perspective renderings of the various landscape design alternatives.

Designing an Environmental Protection Framework

The first step in creating a planning strategy for siting animal confinement facilities involved development of an environmental protection framework. The purpose of the framework was to identify those components of the landscape in the Sacred Heart basin that warrant special protection in meeting the objectives of conserving soil, protecting water quality, enhancing biological diversity and promoting community values. The framework was defined in terms of two dimensions. Landscape integrity, a measure of the ability of landscape structure to maintain ecological function, defined a balance between environmental quality and agricultural production values. Neighborhood cohesion defined relationships between farm and non-farm residents of the rural landscape.

Defining Landscape Integrity

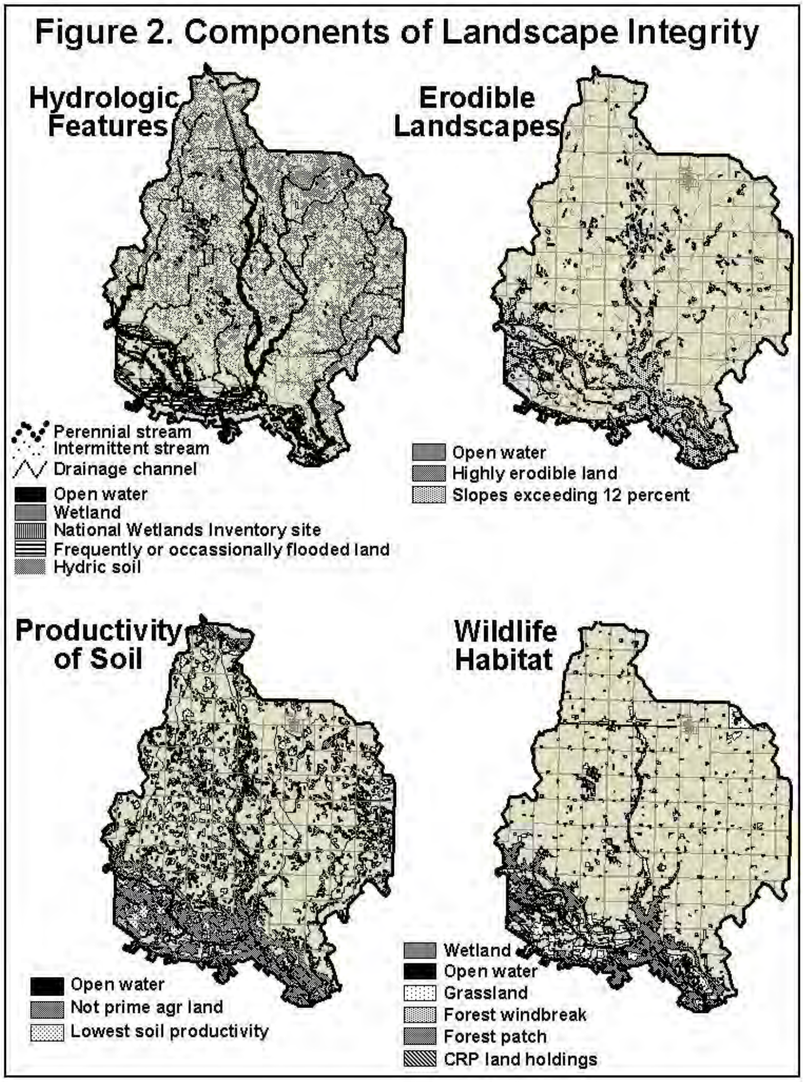
Landscape dimensions critical to defining integrity included:

a) slopes exceeding 12% slope;
b) highly erodible soils;
c) existing surface hydrology;
d) soils that are occasionally or regularly flooded;
e) hydric soils as defined by USDA-NRCS criteria.
f) existing forest, wetland, open water and grassland communities;
g) National Wetland Inventory sites;
h) land enrolled in the Conservation Reserve Program;
i) that one-fourth of the watershed’s total area whose soils contain the lowest potential productivity for corn and soybean;
j) locations of existing animal confinement facilities; and
k) existing field pattens.

Data related to these dimensions were gathered from the Soil Survey Manual of Renville County, Minnesota, digital aerial ortho-photographs of the watershed, digital raster graphic images of the 7-1/2 minute Topographic Quadrangles for the watershed and various existing data sources maintained by agencies of state government (e.g. National Wetland Inventory sites, existing animal confinement facilities). The data were compiled into a GIS data base using Arc-Info™ technology. Figure 2 illustrates the data used in defining landscape integrity.

Determinants of Soil Conservation and Water Quality. Soils that were steeply sloping, highly erodible, occasionally or regularly flooded or contained hydric conditions, especially where they were located in close proximity to surface hydrologic features, were defined as being critical to maintaining surface and ground water quality.

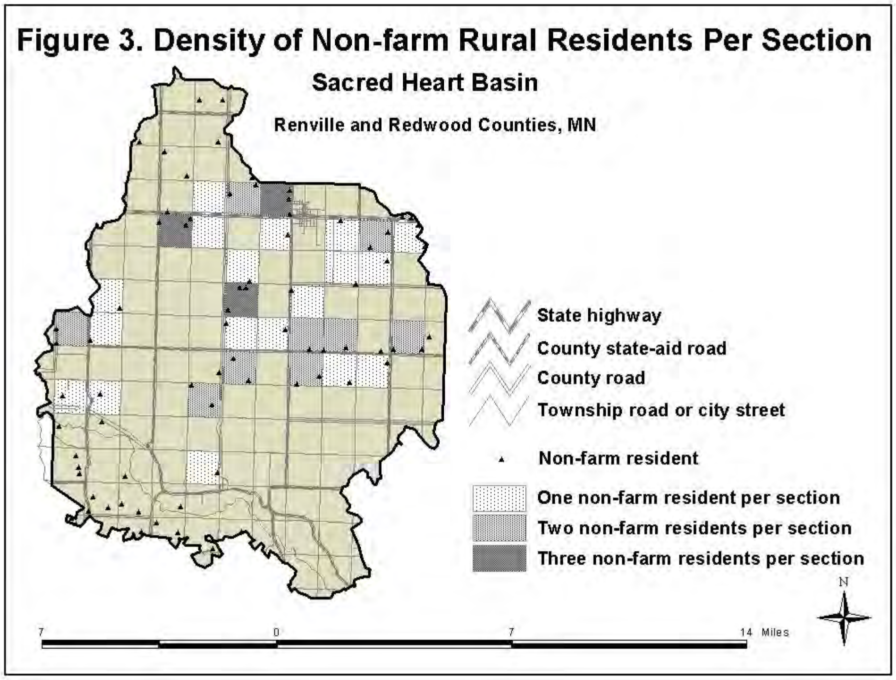
Determinants of Biological Diversity. The original pre-settlement vegetation of the watershed contained upland conditions of xeric, mesic and hydric prairie and bottomland conditions predominated by cottonwood, elm, silver maple and other floodplain species. oak forest and oak openings existed on the south-facing bluffs along the Minnesota River. One hundred and twenty years of agriculture in the watershed has essentially eliminated the prairie vegetation from the landscape, and over 80% of the wetlands have been drained. The cultural landscape of farming introduced upland forest vegetation to the late nineteenth century landscape in the form of



shelterbelt patches and strips of hedgerow and fencerow vegetation. As economic scale of farm operations increased in the last thirty years, many of the hedgerows and fencerows were removed. Similarly, many shelterbelt patches associated with abandoned farmsteads were removed. The result of these cultural influences was the creation of a mono-typical landscape matrix characterized by increasingly larger fields of either soybean or corn. This pattern was broken only where farmers enrolled less productive land in the Conservation Reserve Program (CRP) during the 1980's. Identification of remnant wetland systems, forest patches and strips, grassland communities and CRP land holdings became a priority step in defining patterns of biological diversity in the watershed. Forest patches were differentiated from strips using a 200-meter forest patch width criterion.

Defining Neighborhood Cohesion

Assessing neighborhood cohesion involved mapping the locations of farmers, retired farmers and rural non-farm residents in the landscape. The density of non-farm residents in each land section was mapped. This strategy assumed that non-farm rural residents have greater aversion to the externalities of animal confinement operation (e.g. odor, impacts of transportation, etc.). Figure 3 illustrates the density of non-farm residents in each land section.



Design of Environmental Protection Framework

The dimensions of landscape integrity and neighborhood cohesion provided the raw material for creation of the environmental protection framework (see Figure 4). Within the context of these dimensions, design of the framework pursued five principles:

- **Define and protect riparian corridors within the watershed.** Accomplishment of this design principle ranged from protection of the extensive floodplain systems that exist along the bottomlands of the Minnesota River and its tributaries to delineation of ten meter vegetative buffer strips along the constructed upland drainage systems. Definition of the riparian corridors included bluff and steeply sloped landscapes adjacent to the floodplain and drainage systems. It also included integration of any of the dimensions of either landscape integrity or neighborhood cohesion that were adjacent to the riparian system. For example, corridor definition incorporated areas that were adjacent to the riparian corridor and contained soils with low productivity values.

- **Provide a network of upland connections among riparian corridors.** This objective established an upland network of connectivity among the riparian corridors. Components of the network were defined initially on the basis of remnant wetland, woodland and grassland communities as well as existing CRP holdings. The location of such factors as the presence of low productivity or hydric soil conditions, steep slopes, highly erodible soils, or National Wetland Inventory sites connected remnant communities to themselves and to the riparian corridors.

- **Provide buffering around non-farm residents living in the rural landscape.** Areas containing high concentrations of rural non-farm residents became components of the environmental protection matrix.

- **Use existing patterns of cultural settlement to establish the boundaries of the framework.** Within the context of maintaining the integrity of the spatial pattern of the environmental protection framework, existing field patterns provided a basis of defining framework edges.

- **Define uses appropriate to landscape characteristics.** Within the environmental protection framework, land uses would be permitted based on their appropriateness to the resource characteristics of the landscape. For example, low input perennial or agro-forestry crops were considered appropriate in vegetative buffers of upland drainage riparian systems.

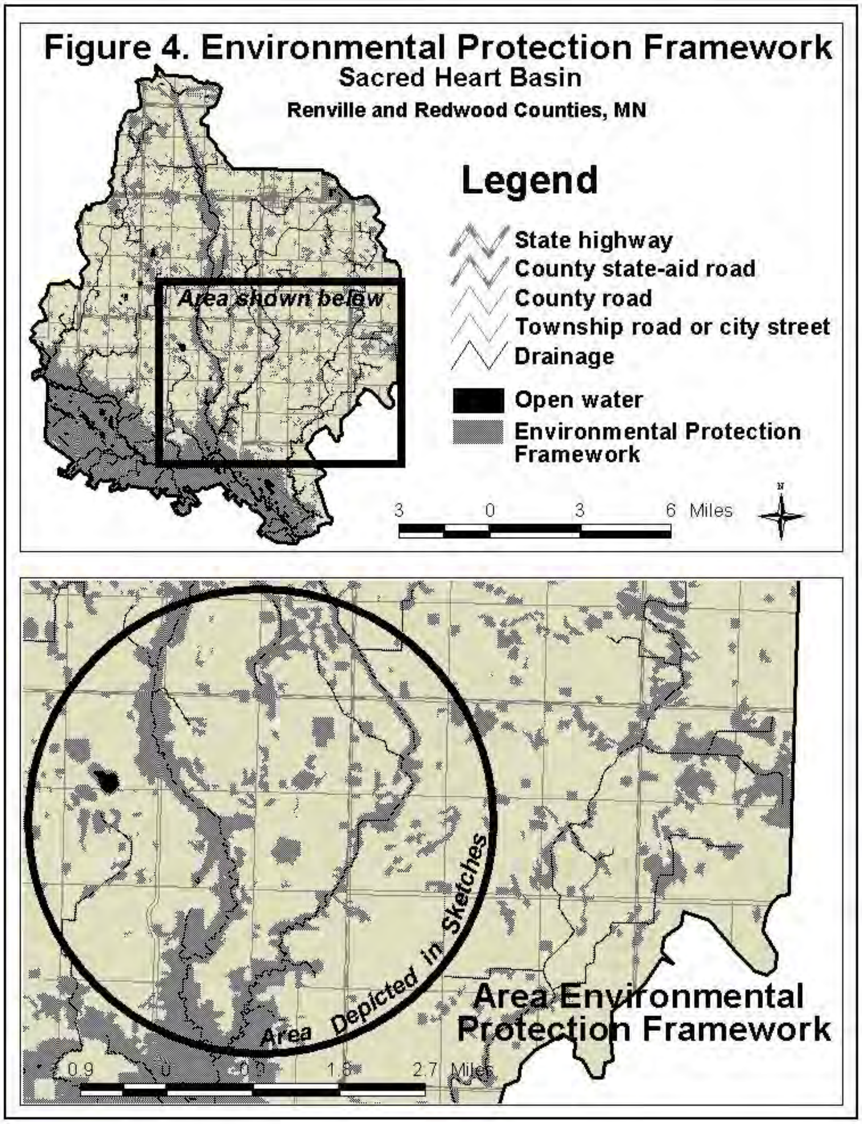


Figure 4. Environmental Protection Framework

Defining Farm Production Units in the Landscape

The environmental protection framework established an armature within which commodity production values in the landscape can be realized. Farm production units involved in pursuit of these values were classified along three continua: crop diversity; farm scale; and operation distribution. Crop diversity varied from highly specialized production of one or two cash crop commodities (e.g. corn and soybean) to farm units that combine dairy and pork production with forage and pick-your-own truck crop production. Farm scale referred to the capital intensity of the farm enterprise, and it generally correlated with geographic size of the operation. Finally, operation distribution referred to the geographic dispersion of the enterprise across the landscape. Distribution varied from farm units operating in one concentrated locus to units operating across several different locations. As illustrated in Figure 5, these three criteria were used to describe farm units operating in the Sacred Heart watershed.

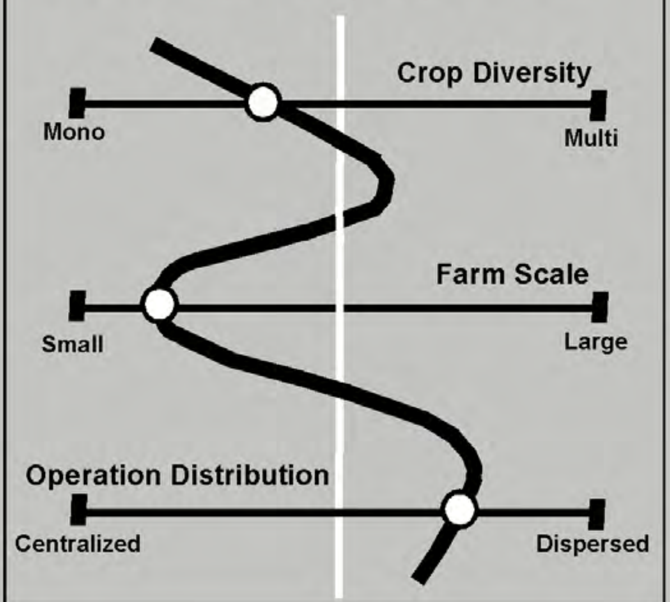


Figure 5. Farm Production Units Can Be Characterized Along Three Continua.

Designing the Production Landscape

Landscape not included in the environmental protection framework was defined as production landscape. Three principle functions occurred within the production landscape: animal production; crop production; and residential uses associated with both farm operators as well as non-farm inhabitants. The case study investigated three scenarios by which these uses might be managed and spatially organized within the production landscape to establish a sustainable rural environment. Each scenario described a prototypical set of conditions that might be found in areas wherein the policies inherent in the scenario were adopted. A low-elevation, oblique aerial perspective rendering accompanies the presentation of each scenario to depict its spatial structure and appearance in the landscape.

Scenario One: Large Scale Zoning Approach

The first design scenario (see Figure 6) involved establishment of four exclusive use zoning districts:

a) an environmental protection zone;
b) an animal agriculture enterprise district;
c) an intensive cropping district; and
d) a rural residential district.

Characteristics of the environmental protection zone were previously defined, and the characteristics of the remaining districts are described below.

Animal Agriculture Enterprise District

An animal agriculture enterprise district consisted of large-scale confinement facilities clustered within the same geographic proximity. Animal agriculture was the permitted use within the district. Location of the considered such criteria as: a) proximity to components of the environmental protection zone; b) maintaining a 1.5 mile buffer between the enterprise district and components of the rural residential district; and c) existing transportation infrastructure. A set of performance criteria relating to air and water quality, resource recycling and the enhancement of biological diversity regulated operation of the animal agriculture uses within the zone. The enterprise districts were large enough to permit establishment of centralized treatment facilities for manure management and treatment.

A manure management cooperative, structured similarly to a rural electric cooperative, managed the waste products of several enterprise districts. The scale of the cooperative's operation permitted it to become involved bye-product. The cooperative established and maintained integrated relationships between generators of manure sludge within the enterprise district and users of manure sludge in the adjacent cropping district. The cooperative also managed integrated relationships between forage producers in the cropping district and forage consumers in the enterprise district. These cooperative relationships allowed the creation of closed production systems in which size of the enterprise district was determined by the forage production and sludge handling capabilities of producers in the adjacent cropping district.

Cropping District

Within a cropping district, large scale crop production, intended in part to supply forage for and process manure sludge from adjacent animal enterprise districts, was the preferred use. The location of the cropping districts was based on three criteria: a) soil productivity, b) constraints and opportunities defined by the environmental protection framework; and c) provision of a 1.5 mile buffer between the enterprise districts and the rural residential districts. The transportation infrastructure within a cropping district was maintained primarily to foster delivery of products to market and interaction between individual producers in the cropping district and animal producers in the enterprise district. Management of this infrastructure involved the conversion of redundant township roads into field roads or the transformation of these rights-of-way into components of the environmental protection framework. Performance standards and incentives operating within the cropping district encouraged adoption of best management practices, establishment of a routine fallow program for fields and creation of the environmental protection framework. Since lands within the cropping district would be down-zoned to crop production uses, a transfer of development rights program (TDR) or a purchase of development rights program (PDR) was established to permit landowners to participate in the development windfalls that accrued to landowners in the rural residential district. Either of these programs allowed development rights transfers into the rural residential district were such non-farm uses were permitted and encouraged.

Rural Residential District

The fourth zoning district was a rural residential district. Within this district, rural residential uses, hobby farms and small-scale farm operations were primary uses. Pre-existing residential uses in the cropping or enterprise districts remained in their current locations, although they would be eligible for relocation into a rural residential district. All future rural residential development would occur within a rural residential district. Future development was encouraged to follow cluster strategies in implementation, and density bonuses became available to developers pursuing cluster strategies. Location of the rural residential districts were based on: a) constraints and opportunities defined by the environmental protection framework; b) provision of a 1.5 mile buffer between the enterprise districts and the rural residential districts; and c) the suitability of soils for development and domestic waste management (either as septic tank drainfield effluent or alternative technologies). Within the rural residential districts, the existing pattern of transportation infrastructure was maintained.

Policy Environment of Large Scale Zoning Approach

The policy environment of the large scale zoning approach involved an intensive regulatory framework. Having defined the environmental protection framework, a significant amount of centralized planning established the boundaries of the other three districts. An administrative framework managed the performance standards and incentive systems operating in the districts, and the framework established the operating procedures of the development rights

Figure 6. Large Scale Zoning Approach

transfer program. A forage production and manure sludge-handling cooperative managed integrated relationships between forage production and sludge handling.

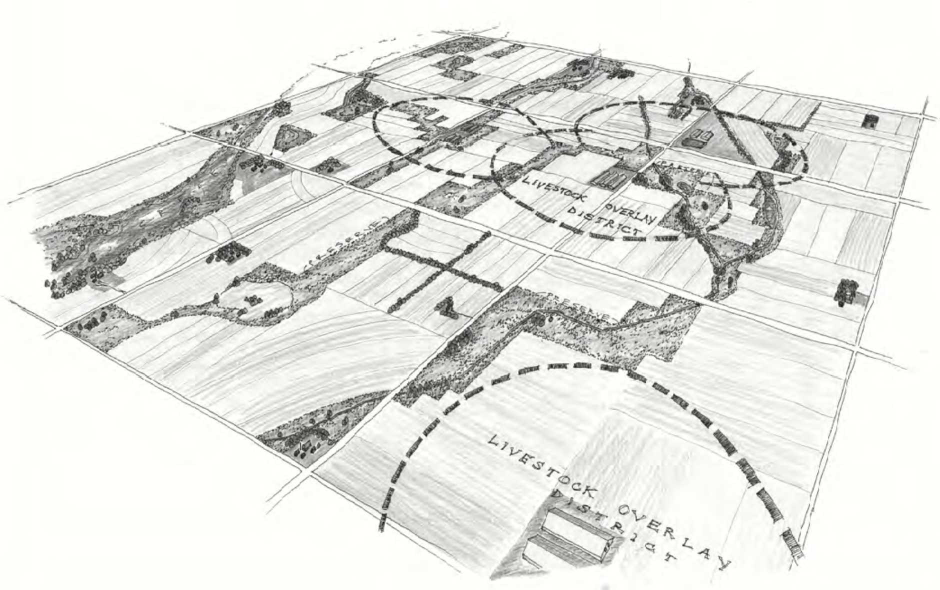
Scenario Two: Mixed Scale Overlay District Approach

Rather than relying on the rigidity of an Euclidean approach to land use separation through exclusive use districts, scenario two adopted the flexibility of overlay district zoning. In this scenario, an application for development or expansion of an animal confinement facility triggered a request to implement an overlay district. The district consisted of all lands within a 1.5 mile radius from the confinement facility (see Figure 7). The application was reviewed in a site selection process that examined the overlay district’s location relative to: a) constraints and opportunities defined by the environmental protection framework; b) existing transportation infrastructure; and c) proximity to rural non-farm residential uses.

Applicants were required to negotiate and secure operating agreements with all landowners within the confines of the overlay district. The requirement to negotiate operating agreements with all landowners within a 1.5 mile radius of the led to the agglomeration of such facilities within the rural landscape. However, rather than legislating these concentrations as was true in the zoning approach, the agglomerations evolved through market activity. Performance criteria relating to air and water quality, resource recycling and the enhancement of biological diversity governed operation of the livestock facility within the overlay district. To the extent practical, operators were encouraged to pursue bye-product recovery (e.g. capture of methane production for small-scale energy generation). The existing pattern of transportation infrastructure within the overlay district was maintained.

The overlay district concept required a less intensive regulatory framework than did the large scale zoning approach. Administration of the site selection and overlay district creation process was required along with administration of performance standards. Otherwise, administrative functions occurred in the private sector through landowner to landowner transactions or through adjudicatory processes initiated to resolve compensatory claims issues.

Figure 7. Mixed Scale Overlay District Approach



Scenario Three: Small Scale Diversified Farming Approach

The third scenario (see Figure 8) assumed a livestock commodity price and economic structure that favored a return to smaller scale more diversified farm operations. In this scenario, individual farms pursued both crop and livestock production, and alternative and low-input agricultural systems became more widely accepted. Individual operators became both forage producer and forage consumer; both manure producer and manure consumer. Farming became a closed system wherein operating scale was based on potential productivity and manure assimilative capacity of soils. Farmers pursued whole farm planning with technical assistance from USDA-NRCS personnel. Implementation of this planning activity, coupled with incentives to encourage adoption of best management practices, enabled establishment and maintenance of the environmental protection framework. Performance criteria relating to air and water quality, resource recycling and the enhancement of biological diversity governed farming operations. The policy environment of scenario three required regulatory activities associated with administration of performance standards. The technical assistance involved in whole farm planning coupled with educational efforts and incentives associated with adoption of best management practice adoption also required administrative activity.

Figure 8. Small Scale Diversified Farming Approach

Conclusions

None of the three scenarios, by itself, represents the single planning approach to accommodating animal confinement facilities in the rural landscape. Collectively, the scenarios represent a catalog of strategies that can be applied in different physiographic and cultural contexts. In many instances, the most appropriate solution may involve a combination of strategies. Regardless of how the production landscape is designed and managed, it is important to frame the designs around the establishment and maintenance of the environmental protection framework. GIS technology is well suited for design of the framework. Execution of the alternative production landscape designs is also well suited to the application of GIS technology. Maps produced from GIS technology lend themselves readily to the production of illustrative graphics, which can be used to further explain the spatial structure and appearance of a sustainable rural landscape for animal production.

A GIS-Based Approach to Characterizing Chemical Compounds in Soil and Modeling of Remedial System Design

Leslie L. Chau, Charles R. Comstock, and R. Frank Keyser
ICF Kaiser Engineers, Inc., Oakland, California

Introduction: The Problem

The cost-effectiveness of implementing a computerized geographic information system (GIS) for environmental subsurface characterization should be based on long-term remedial objectives. A GIS project was developed to characterize soil contamination and to provide design parameters for a soil vapor extraction remedial system, as part of a $120-million remediation and “land sale” project in California. The primary purposes of the GIS were to efficiently combine and evaluate (model) disparate data sets, provide “new” and more useful information to aid in short-term engineering decisions, and support the development of long-term cleanup goals.

The project had a major change in scope early on, and the schedule was expedited to allow for the development of “land sale” options and for actual site redevelopment at the earliest opportunity. Characterization of chemically affected soil would have been compromised given the above circumstances without an ambitious undertaking of concurrently developing and implementing a GIS with three-dimensional (3-D) geostatistical and predictive modeling capabilities.

The GIS Approach

Computer solutions included the use of a cross-platform (DOS and UNIX) GIS to quickly and systematically incorporate spatial and chemical data sets and to provide a distributed data processing and analysis environment (see Figure 1). Networked, DOS-based relational databases were used to compile and disseminate data for the numerous investigatory and engineering tasks. UNIX-based computer aided design (CAD) and modeling applications received data from databases, performed quantitative analyses, and provided 3-D computer graphics. Given the aggressive project schedule, exclusive use of one platform would not be realistic due mainly to the limited data modeling capacity and 3-D graphics in DOS systems. On the other hand, the high startup and operating costs of several UNIX workstations would render their exclusive use much less cost-effective.

The hardware and software configurations were integrated in a client/server Intergraph InterPro 6400 with 48 megabytes of memory. It is largely a 3-D CAD system with add-on modules of geologic mapping and 3-D visual models capable of consolidating both environmental and engineering parameters for analysis (see Figure 1). Textual environmental and geologic data were extracted by SQL queries from relational databases and were transferred to mapping and modeling modules via PC/TCP cross-platform linkage.

The GIS assisted in making short- and long-term decisions regarding health-risk-based regulatory strategy and engineering feasibility. Use of spatial statistical and predictive models was part of a GIS-based decision-making loop (see Figure 2). The process supported concurrent activities in:

- Data collection: field program.
- Numerical models of remedial system configurations.
- Development of cleanup goals from health risk assessments.
- Remedial design with CAD capability.

Site Background

In early 1993, the remedial investigation of the operable unit for soil at a former aircraft manufacturing facility in southern California was thought to be ready for remedial alternatives feasibility study. ICF Kaiser Engineers, Inc., was awarded the contract to perform feasibility studies on applicable soil cleanup technologies and to subsequently design and manage the installation and early operation of the selected technologies. After $700,000 was spent evaluating data collected by previous consultants, it was decided that an additional $5 million worth

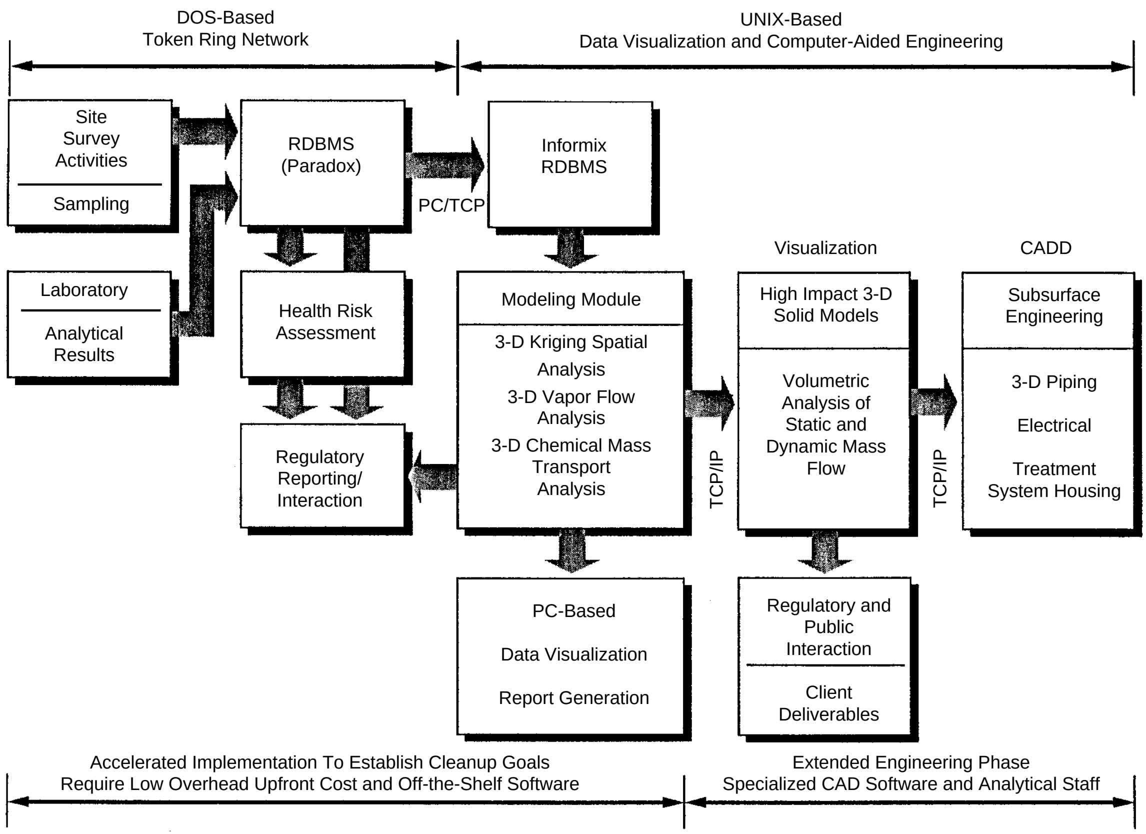


Figure 1. Multiplatform GIS project.

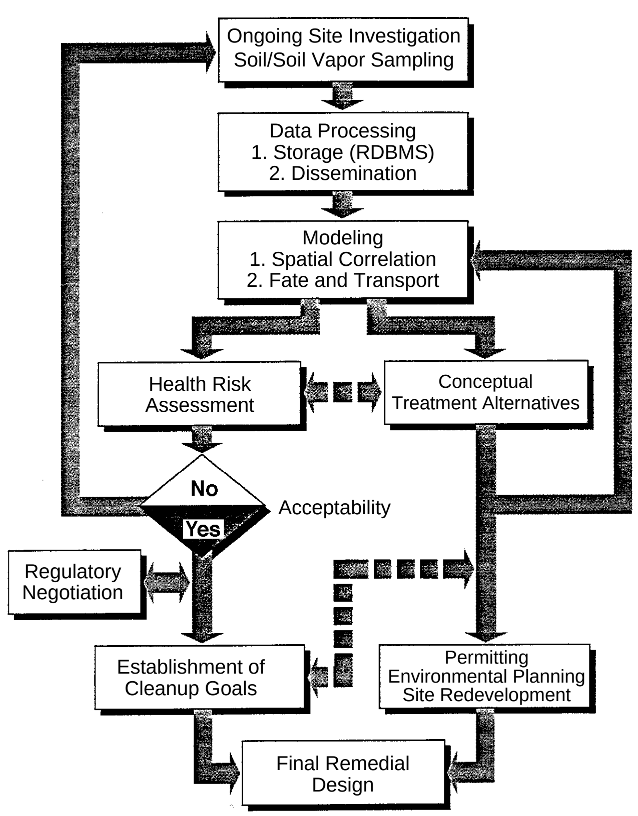


Figure 2. GIS-assisted decision tree.

of field activities were required to more definitively estimate the volume of chemically affected soil and the nature and extent of contamination at the facility. Because of the data gaps, the selection and design of alternatives could not be addressed with a high degree of certainty. Hence, computer assisted data processing was crucial to speed up the feasibility study, accelerate downstream work, and reduce the overall project schedule to the minimum.

The site is environmentally complex, covering an area of approximately 120 acres. As a result of nearly half a century of aircraft production and development, soil beneath the facility is affected by fuel and heavy oil hydrocarbons (TPH) commingled with volatile compounds, mainly perchloroethylene (PCE) and trichloroethylene (TCE) (see Figure 3). Ground water at 170 feet below ground surface is affected by TCE and PCE, but it is not part of the drinking water aquifer. The facility has been demolished, and shallow contaminated soil has been excavated and back-filled to an interim grade.

Methodology

Health-Risk-Based Cleanup Goals

Central to determining the volume and kinds of data to be collected was the question of whether chemicals in soil represented potentially unacceptable risks to human

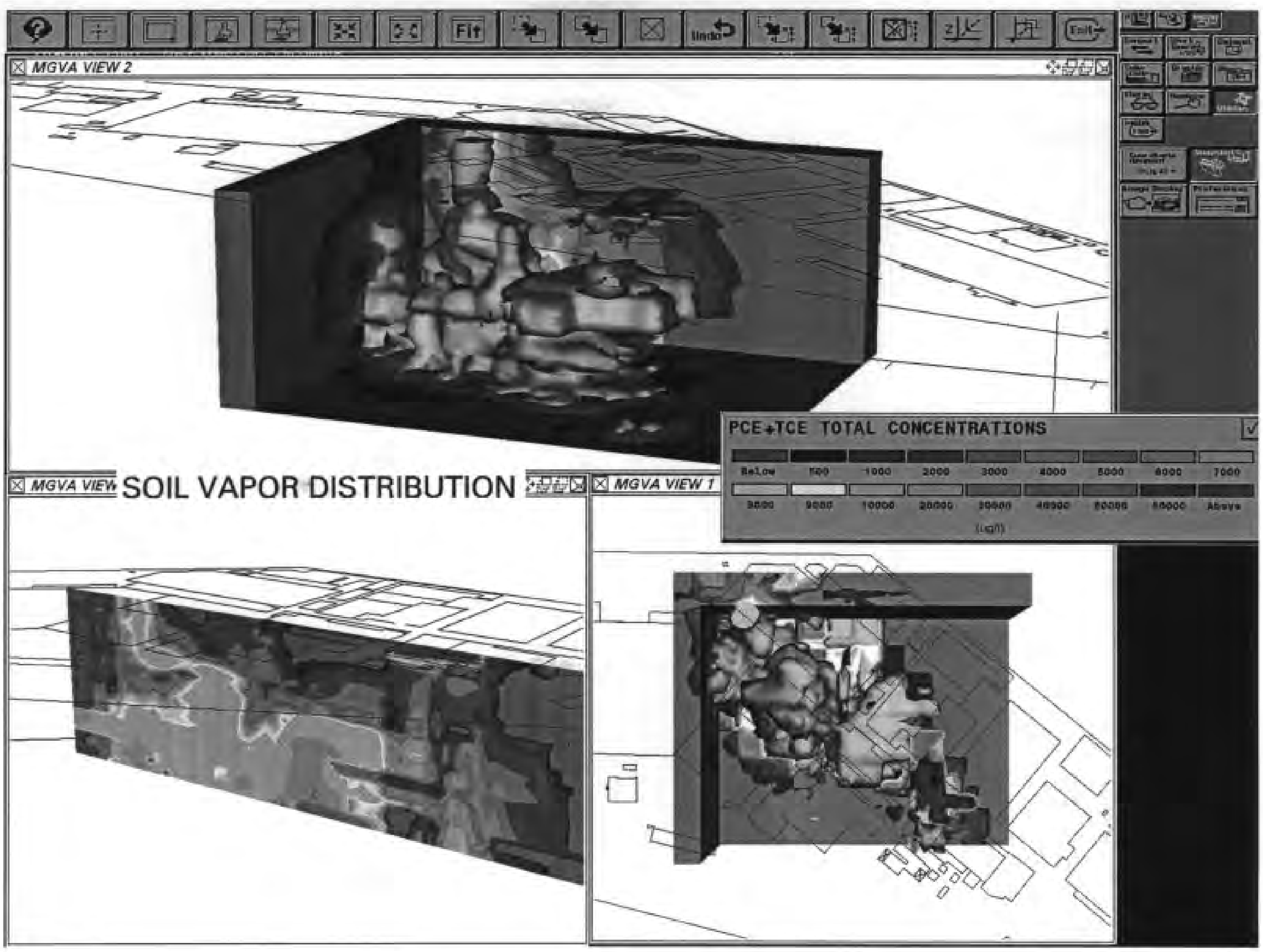


Figure 3. **Aircraft manufacturing facility in California. Outline of demolished buildings located at the 120-acre site are shown as surface features for reference. A geostatistical model of a 3-D kriged VOC soil vapor cloud in the subsurface was simulated with Intergraph's MGVA. Views displayed are: 3-D isometric, vertical section of chemical isoplaths, and a nearly plan view. Digital simulation also illustrates VOCs affecting ground water in a dispersive nature at a depth of nearly 170 feet bgs (shown at bottom of isometric view).**

health and to the environment, with the former being of particular concern to construction workers onsite during redevelopment.

Because site redevelopment was scheduled to begin in the near term, data collection and GIS analysis concentrated on shallow depths (top 20 feet), with decreasing sample density at greater depths. A health-risk-based cleanup goal (HBCG) approach to collecting more data was to establish cleanup goals for near-term remediation of the shallow soils as well as for long-term remedial measures of contaminated soils at greater depths. Further, various regulatory agencies had to approve the estimated cleanup goals in a short time. Ongoing site demolition and excavation schedules encouraged the aggressive regulatory negotiations. The shallow cleanup goals for volatile organic compounds (VOCs) and TPH determined the volume of contaminated soil to be removed. At greater depths, data gaps were minimized to more definitively characterize the nature of TPH and VOC contaminations and to facilitate the implementation of long-term remedial objectives (i.e., in situ soil vapor extraction).

In situ soil vapor and soil sampling composed the field program, which provided data to map the subsurface distribution of volatile organic compounds, including TCE and PCE. Only in situ soil sampling was used for characterizing TPH. The ratio of soil vapor to soil samples was 4:1. No previous soil vapor information was available. ICF Kaiser has been refining the technique of comparing results from paired soil vapor and soil samples in past and similar projects. Hydraulic probes were used instead of drilling to acquire soil vapor samples at shallow depths. This minimized waste and cost in the field program significantly.

Risk Assessment and Spatial Analysis

Human health risk analyses were conducted for the entire site, and risk factors were contoured and overlaid on maps of past usage and known soil contamination areas. Before the risk modeling could proceed, chemical and lithological data gathered in the past 7 years and those acquired by ICF Kaiser populated the environmental relational databases. Approximately 522 soil vapor probes were located in 100-square-foot spacings with additional probes in areas requiring better plume definitions. The database contains approximately 15,000 xyz-records of soil and soil gas laboratory analytical results. This information in text and graphics form, combined with site infrastructures and building outlines with attributes of “past usage,” were stored as map layers, making up the GIS nucleus. Accuracy of site maps was verified with aerial photographs when available. Data types combined for computerized evaluation included known locations of contaminated soil, contaminated ground water, soil types, and site features. Composite risk maps of the above data were analyzed for data gaps at discrete depth intervals. This analysis was performed while the field program was in progress and hence gave guidance to optimize the locations of additional data points and to minimize the number of samples taken.

The MGLA/MGLM mapping module and the MSM terrain modeling module tracked the earth excavation and removal of contaminated soil. Excavation was largely part of site demolition. It also expedited the removal of TPH contaminated soils, however, because no other short-term means of remediation are available for these substances. Tracking of removed soils was essential because concurrent field activities were occurring in site demolition, data gathering, and risk modeling.

The GIS coordinated all three. Geologists and surveyors provided terrain data from daily excavation activities, which were transcribed into database formats. Maps illustrated the locations of excavated soil and removed chemicals in soil at various depths. Although TCE and PCE were of foremost concern as health risks, all compounds and some metals identified in soil were screened for unacceptable risk. Terrain modeling (mapping) as part of health risk assessment may seem unusual, but results of estimated cleanup levels and accurate locations of left-in-place contamination, mostly soils at greater depths, were critical to the cost-effectiveness and proper design of long-term remedial systems.

Characterization of Subsurface VOCs

In situ soil vapor extraction (SVE) of total volatile compounds in dense nonaqueous, liquid, gaseous, and adsorbed solid forms in the subsurface produced favorable results that have been well documented in recent years. ICF Kaiser proposed a very large-scale SVE system (see Figure 4), perhaps the largest yet, as long-term remedial technology for this former aircraft manufacturing site. The primary design problem was speculating on air flow capacity and operating time of the complex system components. The SVE system comprises three fundamental elements:

- Front-end, in situ subsurface vents (totaling 193 corings).
- Applied vacuum and air transport manifolds linking the subsurface vents to the treatment compound (distance of one-quarter mile with over 100 manifolds).
- A multivessel activated carbon treatment system.

To size the pipes, carbon vessels, and vacuum required to achieve a certain rate of VOC removal, the total mass and nature of sorption had to be known. Due to the schedule-driven nature of this project, the SVE design accounted for the time needed to accomplish the cleanup goals.

To estimate the extent and total mass of VOCs in the subsurface, soil vapor data were input to a 3-D kriging algorithm (1) to produce a concentration continuum model (see Figure 3). This solid model of predicted total VOC concentrations took the form of a uniformly spaced 3-D grid-block that completely encased the site. Cell sizes ranged from 10 to 20 cubic feet, depending on the model run, number of data clusters, density of data points in areas of clustered data, and the standard deviation of variances for estimated values in all cells. The Fortran program estimated a concentration value for each cell based on the nearest field sample(s).

The validity of such “block kriging” models can be judged by the size of the variances, smoothness, and agreement with nearby field data. Because volume is a known quantity in kriging, the total mass can be calculated by incorporating soil bulk density or porosity, both of which were less than abundant for this investigation. Renderings of kriged results in 2-D plan view contour maps, cross-sectional maps, and 3-D “vapor cloud” (see Figure 3) were included in client reports and used in regulatory presentations and public forums.

Remedial Design Layout

Final Extension of a Fully Integrated GIS

With the total mass and extent of VOCs derived from 3-D kriged results, the applied vacuum at individual vent heads and the cumulative pressure (negative) necessary to extract and transport VOC vapors from the subsurface to the treatment system can be estimated. We performed 3-D air flow analysis by use of finite difference fluid flow models and chemical transport models. The Fortran codes used to approximate compressible flow and chemical transport were AIR3D (2) and VT3D (3), respectively. Air flow simulations focused on maximizing vacuums at the shallow depths down to 20 feet to expedite remediation of contaminated soils that were

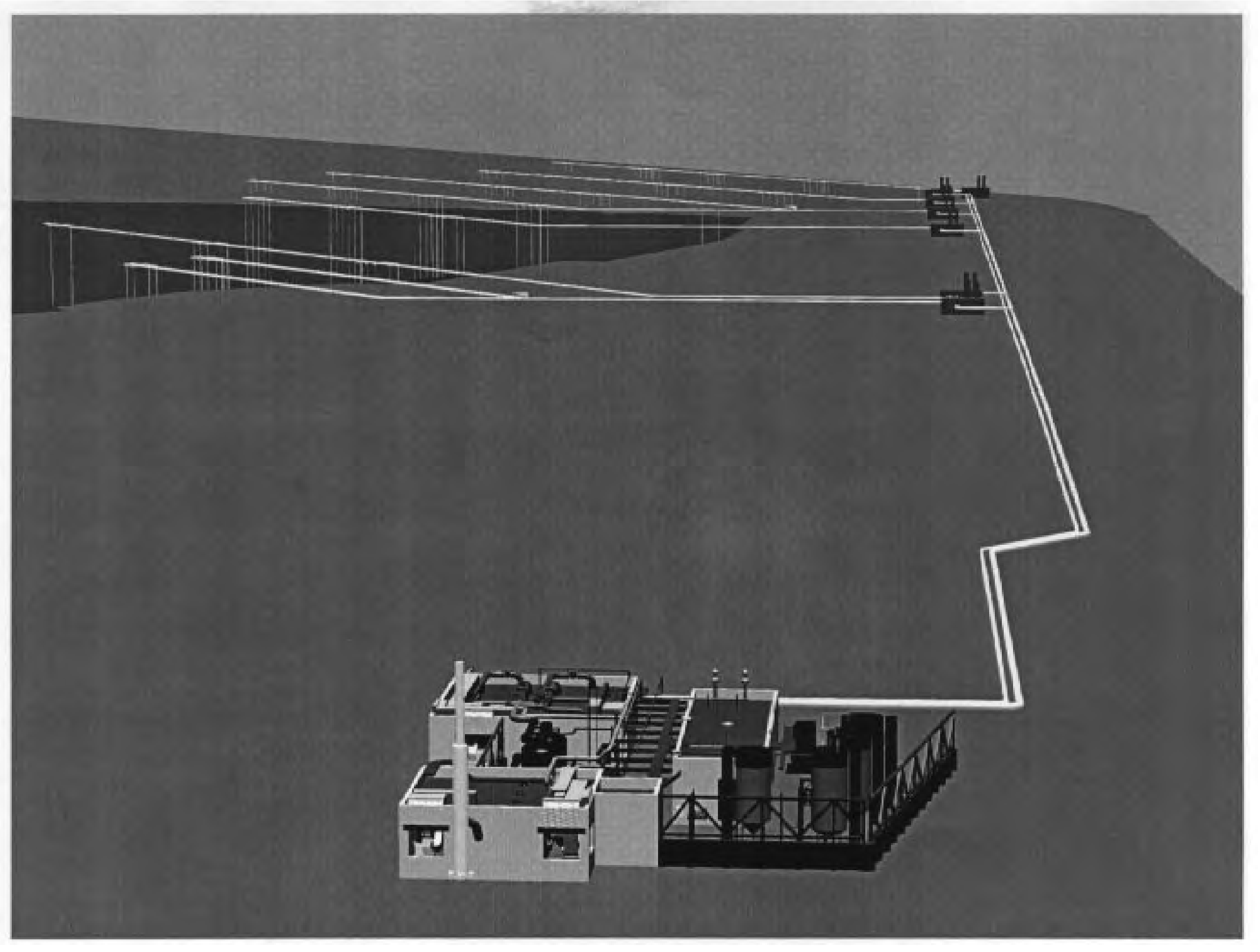


Figure 4. A rendering by the Intergraph 3-D Plant Design System of an in situ soil vapor extraction and treatment system. The cutaway section located near the upper left portion of the figure exposes some of the 193 subsurface extraction vents bottoming at 120 feet (bgs). These vents are located in a cluster for long-term extraction of the VOC vapor cloud presented in Figure 3. Vents are connected to a system of parallel airflow manifolds (right side of figure), which runs one-quarter of a mile to the treatment compound (foreground of figure).

not removed during site demolition and excavation. The lower depths were also included in each simulation. Transient mass transport models incorporate flow fields, given by flow models, and predicted cleanup times based on established HBCG cleanup goals. As VOC concentrations in an operating SVE system fall below cleanup levels in the top 20 feet, thus minimizing human risk, available vacuums thereafter will be diverted to vents at lower depths to be part of long-term extraction scenarios. Models suggested that cleanup for the top 20 feet can be accomplished within 1 year.

Numerical models prescribed vacuum levels at each vent head, which is the aboveground segment of a subsurface SVE vent. The 193 vents are connected to a system of parallel manifolds (see Figure 4) that transport vapor to the treatment system. With the vacuums known at vent heads, the size of manifolds and capacity of vacuum blowers can be determined and integrated into the overall system design. With 3-D Plant Design module as part of the Intergraph CAD/GIS, manifold layouts and treatment compound can be modeled in 3-D and easily checked for pipe routing interferences. The final layout of the SVE system was overlaid onto contour maps of total VOC concentrations to check on accuracy and completeness of vent locations and manifold layouts.

Conclusion

Maximized Visual and Analytical Responses

One goal of this project was to expedite regulatory negotiations and gain early acceptance of cleanup goals. The computerized data processing and visualization contributed generously to the rapid understanding of modeling results by expert regulators and the lay public. Likewise, the GIS facilitated the response to regulatory comments. Positive comments first came from the client's in-house review of model results and the high impact 3-D color rendering of kriged VOC distributions in the subsurface (see Figure 3).

Analytically, benefits were derived from the efficiency of electronic data access and the ability to "predict" the presence of contaminant in areas with sparse field data. The process of kriging involves the linear interpolation and extrapolation of existing data. The resultant contaminant distribution is a "conservative" model that provided the best fit with field data and validated conceptualized subsurface conditions. Further, models provided conservative estimates of mass and extent of PCE and TCE contaminations. Kriging also provided information on the uncertainty of the predicted chemical distribution, which is extremely useful for regulatory discussion and system design. The efficiency of computer models allowed investigators to perform numerous model runs with varied boundary parameters, such as cell size and search radii, in the kriging process.

Accurate mapping of excavated soil and the removal of most TPH source areas provided the incentive to critically assess the feasibility of a no-action remedial scenario for these substances at greater depths. With removal of many TPH source areas, 1-D finite difference models (4) were used to assess the mobility of TPH in NAPL and adsorbed residual phase. Specifically, models assessed the likelihood of largely residual-phase TPH affecting ground water and migrating upward to affect indoor air volumes via gaseous diffusion. Results were extremely favorable; models predicted negligible likelihood of TPH affecting ground water or indoor air volumes. Combined with GIS graphic evidence of specific areas of excavated soil and the absence of TPH sources, regulatory agencies accepted the model results, and the no-action remedial alternative for TPH was approved.

References

1. Deutsch, C.V., and A.G. Journel. 1992. Geostatistical software library and user's guide. New York, NY: Oxford University Press.
2. U.S. Department of the Interior Geological Survey. 1993. AIR3D: An adaptation of the ground-water flow code MODFLOW to simulate three-dimensional air flow in the unsaturated zone. Books and open file reports. Denver, CO.
3. Zheng, C. 1994. VT3D: Numerical model for VOC removal from unsaturated soil (draft). Bethesda, MD: S.S. Papadopulos and Associates, Inc.
4. Rosenbloom, J., P. Mock, P. Lawson, J. Brown, and H.J. Turin. 1993. Application of VLEACH to vadose zone transport of VOCs at an Arizona Superfund site. Groundwater monitoring and remediation (summer). pp. 159-169.

Using GIS/GPS in the Design and Operation of Minnesota's Ground Water Monitoring and Assessment Program

Tom Clark, Yuan-Ming Hsu, Jennifer Schlotthauer, and Don Jakes
Minnesota Pollution Control Agency, St. Paul, Minnesota

Georgianna Myers
Water Management Consultants, Denver, Colorado

Abstract

Minnesota's Ground Water Monitoring and Assessment Program (GWMAP) is administered by the Minnesota Pollution Control Agency (MPCA) to evaluate baseline ground-water quality conditions regionally and statewide. The program uses a systematic sampling design to maintain uniform geographic distribution of randomly selected monitoring stations (wells) for ground-water sampling and data analysis. In 1993, geographic information system (GIS) and global positioning system (GPS) technologies were integrated into GWMAP, automating the selection of wells and the field determination of well locations.

GWMAP consists of three components: the statewide baseline network, regional monitoring cooperatives, and a trends analysis component. In the statewide baseline network, Minnesota is divided into over 700 121-square-mile grid cells, each with a centralized, 9-square-mile sampling region. Within each target area, single-aquifer, cased and grouted wells are sampled for about 125 metals, organic compounds, and major cations and anions. We are currently finishing the second year of a 5-year program to establish the statewide grid. When complete, the statewide baseline component will consist of about 1,600 wells representing Minnesota's 14 major aquifers.

In 1993, approximately 4,000 well construction records were selected for geologic and hydrologic review, using a GIS overlay, from a database of 200,000 water well records maintained in the state's County Well Index (CWI). Using GPS, 364 wells were sampled and field located. The semiautomatic well selection process uses existing electronic coverage of public land survey (PLS) data maintained in CWI in conjunction with the digitized systematic sampling grid. GIS has greatly reduced the time needed for selecting sampling stations. With the combination of GIS and GPS, program costs have decreased, allowing more resources to be applied toward sampling, while efficiency and quality of data have improved.

Introduction

Quantitative assessment of ground-water quality conditions requires a highly organized data collection program that includes statistical evaluation of monitoring results (1, 2). States have difficulty providing the staff and financial resources necessary to generate statewide quantitative ground-water information. With the use of geographic information system (GIS) and global positioning system (GPS) technologies, however, states have the potential to improve the quality of environmental monitoring programs and to reduce the amount of staff time needed to collect and evaluate data, thus decreasing costs. The degree to which states realize these potential benefits depends largely on how effectively the technology can be incorporated into the design of the monitoring program. This paper describes how GIS and GPS technologies are being integrated into the design and operation of Minnesota's Ground Water Monitoring and Assessment Program (GWMAP) to improve overall effectiveness.

The Minnesota Pollution Control Agency (MPCA) has sampled and analyzed ambient ground-water quality in the state's 14 principal aquifers since 1978. In 1990, the MPCA began a redesign of its ground-water monitoring program to better assess water quality conditions statewide (3). Three program components resulted from the redesign: a statewide baseline network for complete geographic coverage, a trends analysis component for intensive studies of how ground-water quality in specific areas changes with time, and a regional monitoring cooperative link to governmental units such as counties to meet specific local ground-water assessment needs. This paper describes the design and operation of the statewide baseline network.

The design of the statewide network is geographically and statistically based to automate well selection and data interpretation. In 1993, the MPCA began integrating GIS and GPS technologies into this part of the program. The implementation of GIS and GPS surpassed our expectations by reducing staff time required to select wells and evaluate analytical results (see Table 1). In addition, through the elimination of previously uncontrollable variables, the use of GIS and GPS has increased the accuracy of GWMAP data.

Monitoring Program Description

Since 1992, GWMAP has selected 150 to 250 existing water supplies yearly for ground-water sampling and analysis of about 125 parameters, including major cations and anions, metals, and volatile organic compounds. Well selection is a fundamental element of GWMAP that, if efficiently performed, supports the program objectives by upholding the quality of the monitoring data and minimizing the operating costs.

A key to the interpretation of monitoring data is the technique used to select wells for sampling (2, 4, 5). Minnesota has over 200,000 active water wells with approximately 10,000 new installations annually. For each well selected for GWMAP monitoring, a hydrologist must individually review many well construction records. An automated prescreening mechanism to facilitate well selection can result in considerable time (and therefore cost) savings. GWMAP chose GIS as the best tool for this task. GIS enables the program to combine a systematic sampling technique with hydrogeologic criteria to ensure an efficient and consistent selection process. As Table 1 shows, GIS allowed us to more than triple our geographic coverage and wells initially selected, while dramatically reducing the records that must be individually reviewed. We realized a time savings of 2 months compared with the time required before GIS implementation.

In general, systematic sampling techniques use a randomly generated uniform grid to determine sampling locations in space and/or time (5). Systematic sampling was initially implemented in GWMAP in 1991 using a manually generated spatial grid defined by the public land survey (PLS) (3). Although the PLS is not 100 percent geographically uniform, it was selected for the grid to expedite well selection from existing digital databases in which wells are organized by PLS location.

Systematic Sample Site Selection Using GIS

Systematic sample site selection is a three-step process. First, a database search of Minnesota's County Well Index (CWI) (6), containing nearly 200,000 driller's records, is conducted to include all available water wells in the region of interest. Second, the candidate pool is reduced to those wells located within regularly spaced grid cells. Third, further wells are eliminated from the candidate pool by applying geologic and well construction criteria mandated in the GWMAP design (7).

Generating the Sampling Grid

The statewide sampling grid was generated from a randomly selected origin (8). This grid consists of approximately 700 square cells, 11 miles on a side (see Figure 1). The centroid of each cell is consecutively numbered and was extracted to produce the origin of the sampling zone.

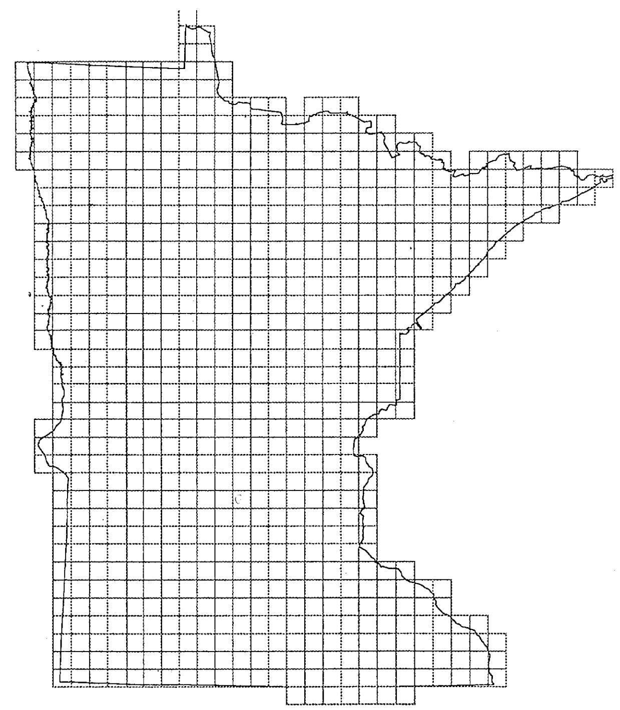


Figure 1. Statewide baseline network sampling grid.

Table 1. Well Selection in 1992 and 1993

Year	Area Covered	PLS Sections Selected	Well Logs Selected	Well Logs Reviewed	Wells Sampled	Time Spent
1992	9 counties	500	3,000	3,000	158	6 months
1993	26 counties	1,659	11,000	834	206	4 months

Establishing the Sampling Zone

Each sampling zone consists of a 3- by 3-mile box from which potential sampling sites are selected. It is generated by computing the coordinates of the four corners of the box using the grid cell's centroid as the origin. To link the sampling zone and grid cell, both are identified with the same numerical code.

These sampling "target" zones, a series of regularly spaced, 9-square-mile boxes, are then made into a GIS coverage and overlaid on top of the PLS coverage to extract those sections that are associated with each of the sampling zones. Ideally, each sampling zone should cover exactly nine PLS sections (3). Due to irregularities in the PLS system, however, portions of 16 to 20 sections usually fall within the sampling zone of each cell (see Figure 2).

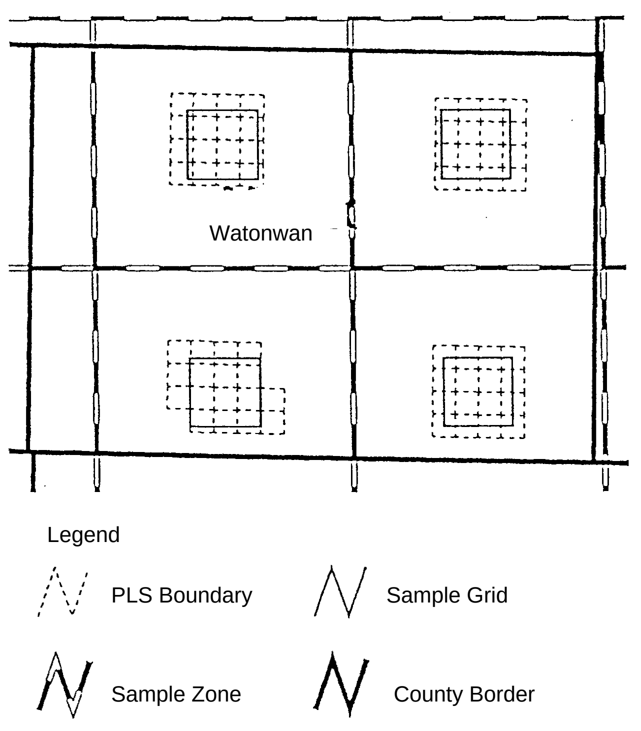


Figure 2. PLS and the sampling grid, Watonwan County.

Selection of PLS Sections

The PLS coverage was derived from the Minnesota Land Management Information Center (LMIC) "GISMO" file. It was originally created in 1979 by digitizing every section corner in Minnesota from the U.S. Geological Survey (USGS) 7.5-minute quadrangle map series.

The PLS section information is necessary in the well selection process because the original well construction logs, maintained by the Minnesota Geological Survey (MGS), are organized by PLS. Although most of the well selection process can be automated, manual file searches for well records are still necessary and require the PLS information.

Well Selection

After identifying the PLS sections within the sampling grid, the statewide well database is imported as a point coverage and overlaid with the selected PLS section coverage. Thus, all wells that fall within the 16 to 20 sections are selected as potential candidates. The accuracy of the well locations in CWI varies; most of the point locations are approximated to four quarters (2.5 acres). The CWI does not contain all well construction information, however, requiring that copies of driller's logs be made for GWMAP files.

The final well selection is done after applying the 9-square-mile sampling zone over the potential pool of candidates. For wells that fall within the zone, the well construction records are pulled from MGS files, copied, and submitted for hydrologist review. Depending on the target cell location, the number of candidate wells requiring review may range from a few to more than 100. For newly installed water wells whose records have not yet been digitized by LMIC, the PLS locations of the wells are manually plotted onto a map to confirm whether they fall into a sampling grid cell. Typically, from 5 percent to as many as 20 percent of selected wells that meet the location criteria are sampled. This accounts for the hydrogeologic and well construction criteria and the cooperation of well owners participating in the program. Currently, interest in ground-water protection programs runs high in rural Minnesota, with an acceptance rate of up to 80 percent.

The implementation of GIS in well selection helped GWMAP excel in two major areas. First, the development of the statewide GIS grid eliminated previously uncontrolled variables by removing the PLS spatial inconsistencies from the systematic grid. Second, the GIS reduced the manual workload with the automation of two important steps in the well selection process: the generation of PLS section information to facilitate the database search, and the identification of wells that meet the geographic location criteria. The success of GWMAP relies largely on the ability to use existing GIS coverages. In using coverages created by other entities, this program identified the need for a uniform standard for data conversion and transfer.

Application of Global Positioning Systems in Ground-Water Sampling

In 1991, the U.S. Environmental Protection Agency (EPA) established a policy that all new data collected after 1992 should meet an accuracy goal of 25 meters or better (9). The purpose of EPA's Locational Data Policy (LDP) is to establish principles for collecting and documenting consistently formatted locational data to facilitate cross-programmatic, multimedia analyses. Accurate geographic information is important to the spatial analysis of well sampling results. Any uncertainty in sample location can compromise hydrogeologic analysis (10). GPS is an easy, cost-effective solution.

Global Positioning System Field Application

Beginning in October 1992, GWMAP employed GPS in the field to assist in locating sample sites. Applying GPS in the field has proven to be quite easy. The program uses a multichannel C/A code receiver with internal data logging capability. Typically, the receiver is placed directly on top of the wellhead and logs 100 to 150 GPS readings into the receiver's internal memory in approximately 5 minutes.

The GPS is also used for navigation in the field to locate sampling sites. Because sampling sites are predetermined, their locations can be extracted from a topographic map. The approximate coordinates can then be loaded into a GPS receiver. In most cases, the receiver successfully led the field team within visual range of the sampling site.

Because of the inherent selective availability (SA) of the GPS, raw field data must go through a differential correction process to achieve the goal of 25-meter accuracy (9, 11).

Data Management and Processing

Once the GPS receiver is brought back from the field, data are downloaded to a personal computer (I486 processor at a speed of 50 MHz) and differentially corrected (11). The average or mean of the 100 or more readings collected onsite is calculated and reported as the site location.

The MPCA does not operate a GPS base station for the purpose of differential correction. The base station data are obtained through a computer network (Internet) from the Minnesota Department of Health (MDH) base station located in Minneapolis.

To facilitate future data integration and document data accuracy for secondary application, GWMAP proposed quality assurance codes for GPS data collected by the MPCA. The value of the accuracy proposed is a nominal value rather than an absolute number (see Table 2). Each of the seven processing methods is assigned a separate code.

In the field experience of GWMAP, a nominal accuracy of 2 to 5 meters has been consistently achieved after the postdifferential correction and averaging have been applied to the data. This technology is suitable for any program that is designed to conduct either large-area or intensive monitoring activities. It helps to cut costs by increasing efficiency and accuracy of the data. The data collected by GWMAP can be used not only in a regional study but could be used directly in a site-specific investigation as well.

Table 2. Proposed Nominal Accuracy Reference Table

Type of GPS Receiver Used	Processing Method Used To Correct Data	Nominal Accuracy (meters)
Navigational quality C/A code receiver	Postdifferential corrected	2-5
	Real-time differential corrected (RTCM)	2-5
	Autonomous mode (no correction)	15-100
Navigational quality with carrier aid receiver	Postdifferential corrected	< 1
	Real-time differential corrected (RTCM)	< 1
	Autonomous mode (no correction)	15-100
Survey quality receiver (dual or single frequency)	Postdifferential corrected	< 0.1

GWMAP also found that GPS can be used most efficiently by separating the two roles of field operator and data manager. The field operators receive only the brief instructions necessary to operate a GPS receiver before going into the field. The data manager handles the data processing details. The field operators can then concentrate their efforts on obtaining ground-water samples and conducting the hydrogeologic investigation.

Conclusions

GIS and GPS technologies made it possible for the MPCA to implement the statewide GWMAP project by optimizing the available funding and staff time. GIS minimized staff time spent on identifying sampling areas, manipulating the sampling grid, and selecting monitoring sites. In addition, GIS enabled GWMAP to integrate a variety of databases and maps of different scales.

Using GPS to locate sampling sites enabled GWMAP to efficiently obtain accurate geographic locational data with relative ease. This eliminated the degree of uncertainty that previously might have compromised the statistical evaluation of the hydrogeologic data.

GWMAP's success in integrating existing digital data to automate the well selection process clearly demonstrated the importance of the ability to share information with others and the great need for a broadly applied standard for data conversion and transfer.

Acknowledgments

The authors wish to thank Renee Johnson of the Minnesota Department of Natural Resources for her work to convert the PLS data layer to GIS coverage. Susanne Maeder of LMIC supplied the statewide CWI coverage, and Susan Schreifels of MPCA conducted research on the LDP and made valuable suggestions on implementing GPS.

References

1. Nelson, J.D., and R.C. Ward. 1981. Statistical considerations and sampling techniques for ground-water quality monitoring. Ground Water 19(6):617-625.
2. Ward, R.C. 1989. Water quality monitoring—A systems approach to design. Presented at the International Symposium on the Design of Water Quality Information Systems, Colorado State University, Ft. Collins, CO.
3. Myers, G., S. Magdalene, D. Jakes, and E. Porcher. 1992. The redesign of the ambient ground water monitoring program. St. Paul, MN: Minnesota Pollution Control Agency.
4. Olea, R.A. 1984. Systematic sampling of spatial functions. Series on Spatial Analysis, No. 7. Kansas Geological Survey, University of Kansas, Lawrence, KS.
5. Gilbert, R.O. 1987. Statistical methods for environmental pollution monitoring. New York, NY: Van Nostrand Reinhold.
6. Wahl, T.E., and R.G. Tipping. 1991. Ground-water data management—The county well index. Report to the Legislative Commission on Minnesota Resources. Minnesota Geological Survey, University of Minnesota, St. Paul, MN.
7. Clark, T., Y. Hsu, J. Schlotthauer, and D. Jakes. 1994. Ground-water monitoring and assessment program—Annual report. St. Paul, MN: Minnesota Pollution Control Agency.
8. ESRI. 1993. ARC/INFO version 6.1, ARCPLOT command references. Redlands, CA: Environmental Systems Research Institute, Inc.
9. U.S. EPA. 1992. Global positioning systems technology and its application in environmental programs. GIS Technical Memorandum 3. EPA/600/R-92/036. Washington, DC.
10. Mitchell, J.E. 1993. A characterization of the influence of (x,y) uncertainty on predicting the form of three-dimensional surfaces. Proceedings of the AWRA Spring Symposium on Geographic Information Systems and Water Resources, Mobile, AL. pp. 559-567.
11. Trimble Navigation, Ltd. 1993. GPS Pathfinder System, general reference. Sunnyvale, CA: Trimble Navigation, Ltd.

EPA's Reach Indexing Project—Using GIS To Improve Water Quality Assessment

Jack Clifford
Office of Wetlands, Oceans, and Watersheds, U.S. Environmental Protection Agency, Washington, DC

William D. Wheaton and Ross J. Curry
Research Triangle Institute, Research Triangle Park, North Carolina

Abstract

The Waterbody System (WBS), which the U.S. Environmental Protection Agency (EPA) originally developed to support preparation of the report to Congress that Section 305(b) of the Clean Water Act requires, is a potentially significant source of information on the use support status and the causes and sources of impairment of U.S. waters. Demand is growing for geographically referenced water quality assessment data for use in interagency data integration, joint analysis of environmental problems, establishing program priorities, and planning and management of water quality on an ecosystem or watershed basis.

Because location of the waterbody assessment units is key to analyzing their spatial relationships, EPA has particularly emphasized anchoring water bodies to the River Reach File (RF3). The reach file provides a nationwide database of hydrologically linked stream reaches and unique reach identifiers, based on the 1:100,000 U.S. Geological Survey (USGS) hydrography layer.

EPA began the reach indexing project to give states an incentive to link their water bodies to RF3 and to ensure increased consistency in the approaches to reach indexing. After a successful 1992 pilot effort in South Carolina, an expanded program began this year. Working with Virginia, a route system data model was developed and proved successful in conjunction with state use of PC Reach File (PCRF), a PC program that relates water bodies to the reach file. ARC/INFO provides an extensive set of commands and tools for developing and analyzing route systems and for using dynamic segmentation.

One important advantage of the route system is that it avoids the necessity of breaking arcs; this is an important consideration in using RF3 as the base coverage in a geographic information system (GIS). Using dynamic segmentation to organize, display, and analyze water quality assessment information also simplifies use of the existing waterbody system data. Because of the variability in delineation of water bodies, however, other states used a number of different approaches. Working with these states has defined a range of issues that must be addressed in developing a consistent set of locational features for geospatial analysis.

Wider use of these data also depends upon increased consistency in waterbody assessments within and between states. Several factors complicate the goal of attaining this consistency in assessment data:

- The choice of beneficial use as the base for assessment of water quality condition.
- The historical emphasis on providing flexible tools to states.
- The lack of robust standards for assessment of water quality condition.

This paper explores possible resolutions to the problem of building a national database from data collected by independent entities.

Section 305(b) of the Clean Water Act and the Waterbody System

Background of Section 305(b)

Since 1975, Section 305(b) of the Federal Water Pollution Act, commonly known as the Clean Water Act (CWA), has required states to submit a report on the quality of their waters to the U.S. Environmental Protection Agency (EPA) administrator every 2 years. The administrator must transmit these reports, along with an analysis of them, to Congress.

State assessments are based on the extent to which the waters meet state water quality standards as measured against the state's designated beneficial uses. For each use, the state establishes a set of water quality criteria or requirements that must be met if the use is to be realized. The CWA provides the primary authority to states to set their own standards but requires that all state beneficial uses and their criteria comply with the 'fishable and swimmable' goals of the CWA.

Assessments and the Role of Guidelines

EPA issues guidelines to coordinate state assessments, standardize assessment methods and terminology, and encourage states to assess support of specific beneficial uses (e.g., aquatic life support, drinking water supply, primary contact recreation, fish consumption). For each use, EPA asks that the state categorize its assessment of use support into five classes:

- *Fully supporting:* meets designated use criteria.
- *Threatened:* may not support uses in the future unless action is taken.
- *Partially supporting:* fails to meet designated use criteria at times.
- *Not supporting:* frequently fails to meet designated use criteria.
- *Not attainable:* use support not achievable.

In the preferred assessment method, the state compares monitoring data with numeric criteria for each designated use. If monitoring data are not available, however, the state may use qualitative information to determine use support levels.

In cases of impaired use support (partially or not supporting), the state lists the sources (e.g., municipal point source, agriculture, combined sewer overflows) and causes (e.g., nutrients, pesticides, metals) of the use support problems. Not all impaired waters are characterized. Determining specific sources and causes requires data that frequently are not available.

States generally do not assess all of their waters each biennium. Most states assess a subset of their total waters every 2 years. A state's perception of its greatest water quality problems frequently determines this subset. To this extent, assessments are skewed toward waters with the most pollution and may, if viewed as representative of overall water quality, overstate pollution problems.

Assessment Data Characteristics

Each state determines use support for its own set of beneficial uses. Despite EPA's encouragement to use standardized use categories, the wide variation in state-designated beneficial uses makes comparing state uses an inherent problem. This affects the validity of aggregation and use of data across state boundaries. Comparably categorizing waters into use support categories also poses a problem; different states apply the qualitative criteria for use support levels in very different ways. Further limiting the utility of Section 305(b) data is that data are aggregated at the state level and questions about the use support status of individual streams cannot be resolved without additional information. While some states report on individual waters in their Section 305(b) reports, EPA's Waterbody System (WBS) is the primary database for assessment information on specific waters.

State monitoring and assessment activities are also highly variable. States base assessments on monitoring data or more subjective evaluation. The evaluation category particularly differs among states.

Waterbody System

The WBS is a database and a set of analytical tools for collecting, querying, and reporting on state 305(b) information. It includes information on use support and the causes and sources of impairment for water bodies, identification and locational information, and a variety of other program status information.

As pointed out earlier, although some states discuss the status of specific waters in their 305(b) reports, many do not. The WBS is generally much more specific than the 305(b) reports. It provides the basic assessment information to track the status of individual waters in time and, if georeferenced, to locate assessment information in space. By allowing the integration of water quality data with other related data, the WBS provides a framework for improving assessments.

WBS has significant potential for management planning and priority setting and can serve as the foundation for watershed- and ecosystem-based analysis, planning, and management. In this respect, it can play a vital role in setting up watershed-based permitting of point sources. The primary function of WBS is to define where our water quality problems do and do not exist. WBS is increasingly used to meet the identification requirement for waters requiring a total maximum daily load (TMDL) allocation. It can serve as the initial step in the detailed allocation analysis included in the TMDL process. In addition, WBS is an important component of EPA participation in joint studies and analyses. For instance, EPA is currently participating with the Soil Conservation Service (SCS) in a joint project to identify waters that are impaired due to agricultural nonpoint source (NPS) pollution. WBS can also anchor efforts to provide improved public access at the state and national levels to information on the status of their waters.

It is important to recognize that use of WBS is voluntary. Of the 54 states, territories, river basin commissions, and Indian tribes that submitted 305(b) reports, approximately

30 used the WBS in the 1992 cycle. While submissions for the 1994 cycle are not complete, we anticipate about the same level of participation. This represents about a 60-percent rate of participation in WBS, which may be the limit for a voluntary system. This severely limits use of WBS assessment data for regional and national level analysis. If data at the national level are required, mandatory data elements, formats, and standards may be necessary.

EPA is currently attempting to achieve consistency through agreement with other state and federal agencies. The recent work of the Interagency Task Force on Monitoring offers hope for eventual consensus on the need for nationally consistent assessment data and mutually agreed upon standards for collection, storage, and transfer. The Spatial Data Transfer Standards already govern spatial data, allowing movement of data between dissimilar platforms. The Federal Geographic Data Committee provides leadership in coalescing data integration at the federal level; it provides a model for government and private sector efforts. This level of cooperation, however, has not always been present in water assessment data management. Assuming that national and regional assessment data are needed, if consistency is elusive through cooperative efforts, regulations may be necessary. Developing a national database may not be feasible without a mutual commitment by EPA and the states to using common assessment standards.

WBS was originally developed as a dBASE program in 1987. It has undergone several revisions since then, and the current Version 3.1 is written in Foxpro 2.0. The WBS software provides standard data entry, edit, query, and report generation functions. WBS has grown substantially in the years since its inception, primarily in response to the expressed needs of WBS users and EPA program offices. The program's memory requirements and the size of the program and data files, however, are of growing concern to state WBS users and the WBS program manager. Because of the wide range of WBS user capabilities and equipment, users must be equipped to support an array of hardware from high capacity Pentium computers to rudimentary 286 machines with 640 Kb of memory and small hard disks. This range makes memory problems inevitable for some users.

While WBS contains over 208 fields, exclusive of those in lookup tables, approximately 30 fields in four files provide the core data needed to comply with 305(b) requirements. These fields contain:

- Identification information for the water body.
- The date the assessment was completed.
- The status of use support for beneficial uses.
- The causes and sources of any use impairment in the water body.

The uses WBS considers are both state-designated uses and a set of nationally consistent uses (e.g., overall use, aquatic life support, recreation) specified in the 305(b) guidelines. The other essential piece of information is the geographic location of the water body, which the remainder of this paper discusses in detail.

Significant differences exist in the analytical base as well as in assessments. EPA provided little initial guidance on defining water bodies; therefore, states vary widely in their configurations of water bodies. Water bodies are supposed to represent waters of relatively homogeneous water quality conditions, but state interpretation of this guidance has resulted in major differences in waterbody definition.

Initially, many states developed linear water bodies, and these were often very small. The large number of water bodies delineated, however, created significant difficulties in managing the assessment workload and were not ideal in the context of the growing need for watershed information. Some states, such as Ohio, developed their own river mile systems.

As discussed below, some states indexed their water bodies to earlier versions of the reach file, and therefore, the density of the streams these water bodies include is fairly sparse. Recently, many states have redefined their water bodies on the basis of small watersheds (SCS basins, either 11-digit or 14-digit hydrologic unit codes [HUCs]).

Locating water bodies geographically is a necessary prerequisite to assessing water quality on a watershed or ecosystem basis. The WBS has always included several locational fields, including county name and FIPS, river basin, and ecoregion. These fields have not been uniformly populated, however. One of the WBS files includes fields for the River Reach File (RF3) reaches included in the water body. While a few states had indexed their water bodies to older versions of the reach file (RF1 and RF2), however, no state had indexed to RF3 until 1992.

In 1992, EPA initiated a demonstration of geographic information system (GIS) technology in conjunction with the South Carolina Department of Health and Environmental Control. This project involved:

- Indexing South Carolina's water bodies to RF3.
- Developing a set of arc macro languages (AMLs) for query and analysis.
- Producing coverages of water quality monitoring stations and discharge points.
- Using GIS tools in exploring ways to improve water quality assessments.

South Carolina has defined its water bodies as SCS basins.

The results have been very encouraging. First, South Carolina took the initial coverages and decided they needed much more specificity in their use support determinations and their mapping of the causes and sources of impairment. As a result, they mapped these features down to the reach level. Next, they decided that they needed better locational information, so they used global positioning satellite receivers to identify accurate locations for discharges and monitoring stations. They then used GIS query and analysis techniques to relate their monitoring and discharge data to their water quality criteria. South Carolina is using GIS to actively identify water quality problems and improve their assessments.

In 1993, EPA worked cooperatively with several states to index their water bodies to the reach file. Virginia, the next state to be indexed, demonstrated the successful use of PC Reach File (PCRF) software (described later in this paper) for indexing water bodies to the reach file. Ohio and Kansas also are essentially complete. Each of these states required a somewhat different approach than Virginia. The need for flexibility in dealing with states on reach indexing issues is essential. Existing waterbody delineations often represent considerable investment; therefore, EPA must provide the capability to link the state's existing assessment data to the reach file in order to encourage state buy-in.

Figure 1 shows the results of Ohio's indexing of a typical cataloging unit (CU). Figure 2 reflects part of the output of the Kansas work. We can link use support, cause, and source data to each of these water bodies now. In the future, we hope to map these attributes at a higher level of resolution, down to the reach segment level. GIS has proven to be a useful assessment tool. With higher resolution, it should prove to be even more helpful in identifying water quality problems, picking up data anomalies, and assessing management actions, strategies, and policies. This entire process has taught us much and has strengthened enthusiasm for place-based management.

The Reach Indexing Project—Georeferencing the Waterbody System

Purpose and Overview

The reach indexing project is designed to locate water bodies using RF3 as an electronic base map of hydrography and to code RF3 reaches with the specific waterbody identifier (WBID). After linking water bodies to their spatial representation, they can be queried and displayed with assessment data located in WBS files.

Reach indexing includes several steps. First, the state must supply waterbody locations and WBIDs. The next step entails developing a set of procedures for indexing. Finally, the coded RF3 data must be produced.

Input data to the indexing process includes:

- A list of valid WBIDs. In most cases, the state has already input these identification numbers to the WBS.
- Information about the location of each water body. Locational information may be found in marked-up paper maps showing waterbody locations or electronic files from WBS containing waterbody indexing

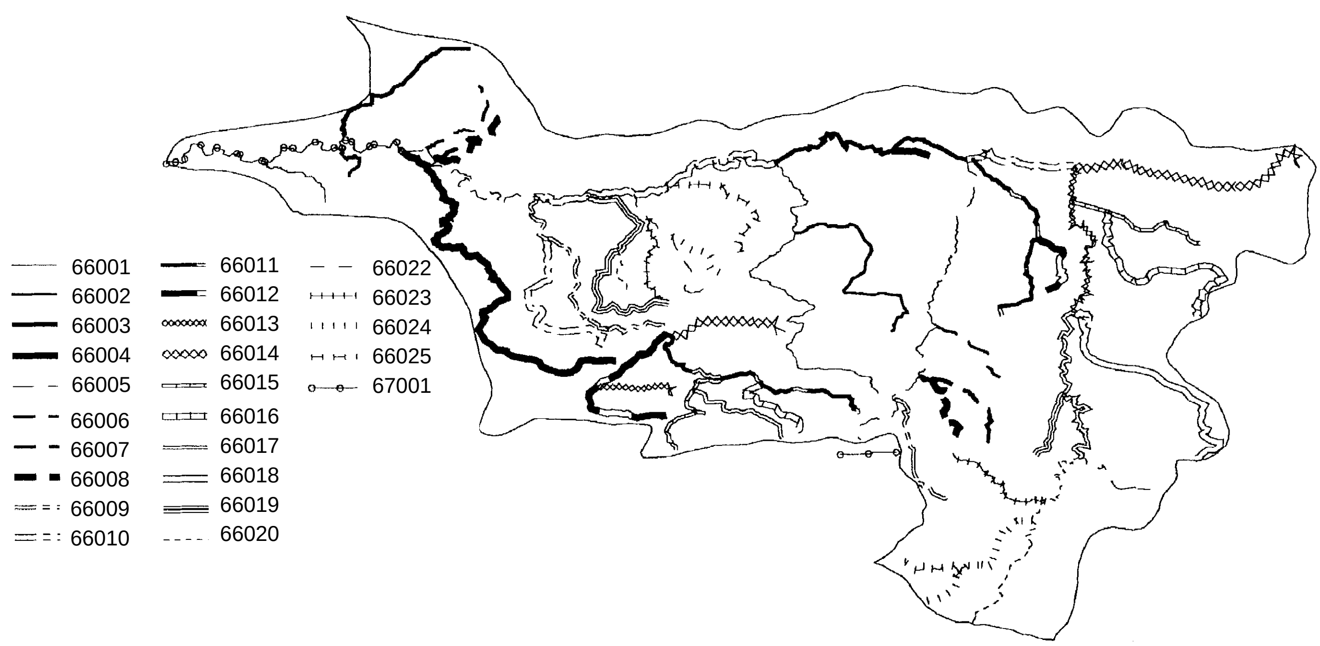


Figure 1. State of Ohio water bodies in CU 04100008.

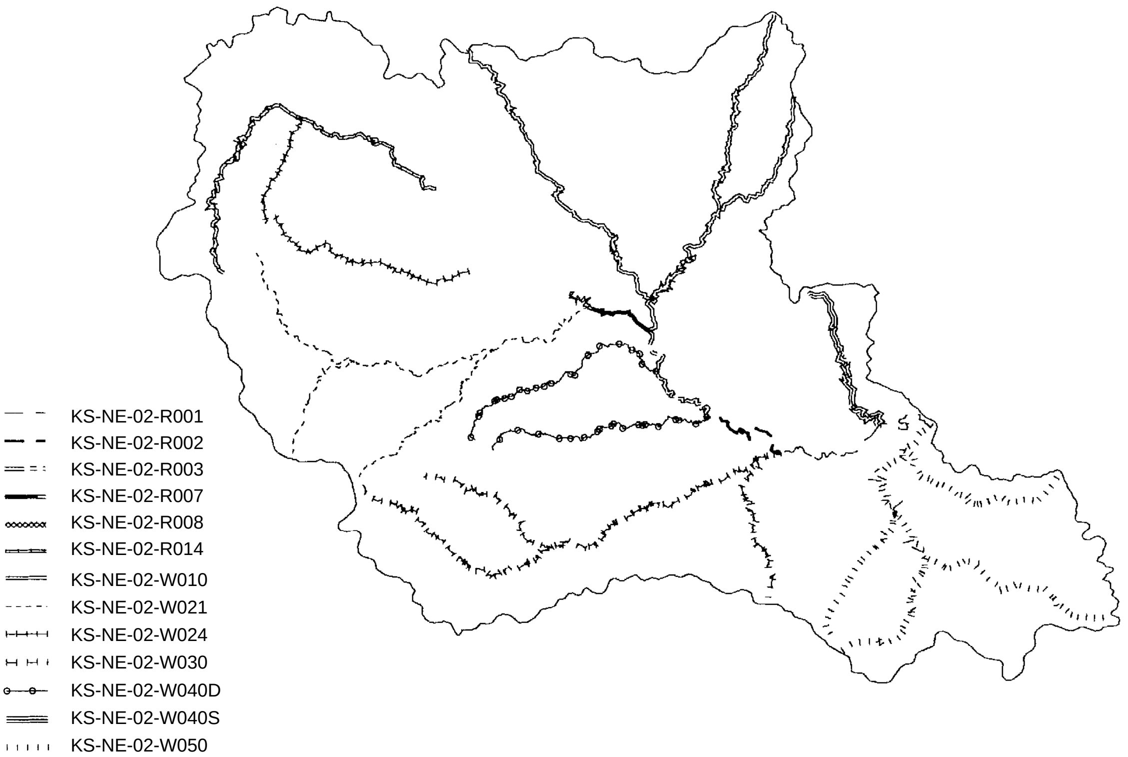


Figure 2. State of Kansas water bodies in CU 11070202.

expressions (discussed later), or it may be embedded in the WBID itself.

- A complete set of RF3 data for the state being indexed.

Depending on the type of information the state supplies, procedures used to index water bodies can be almost fully automated, semiautomated, or completely manual.

The final result of the indexing processes is a set of RF3 coverages that contain a WBID attribute. This product allows querying and displaying of assessment data, which is collected and stored by water body, in a GIS environment.

The Reach File Database

The reach file is a hydrographic database of the surface waters of the continental United States. Elements within the database represent stream segments. The elements were created for several purposes:

- To perform hydrologic routing for modeling programs.
- To identify upstream and downstream connectivity.
- To provide a method to uniquely identify any particular point associated with surface waters.

The unique reach identifier has succeeded in associating other EPA national databases, such as STORET, to surface waters. Any point within these databases can be associated with and identified by a specific location on any surface water element, such as a reservoir, lake, stream, wide river, or coastline.

There are three versions of the reach file. The first was created in 1982 and contained 68,000 reaches. The second version, released in 1988, doubled the size of Version 1. The third version (RF3) includes over 3,000,000 individual reach components.

The base geography of RF3 is derived from U.S. Geological Survey (USGS) hydrographic data (1:100,000 scale) stored in digital line graph (DLG) format. Unlike DLG data, which are partitioned by quad sheet boundaries, RF3 data are partitioned by CU. A CU is a geographic area that represents part or all of a surface drainage basin, a combination of drainage basins, or a distinct hydrologic feature. The USGS uses CUs for cataloging and indexing water-data acquisition activities.

The continental United States comprises over 2,100 CUs. CUs are fairly small; for example, 45 units fall partially or completely within the state of Virginia (see Figure 3).

RF3 is a powerful data source used in hydrologic applications for many reasons, including the following:

- RF3 has spatial network connectivity that topological upstream/downstream modeling tools use.

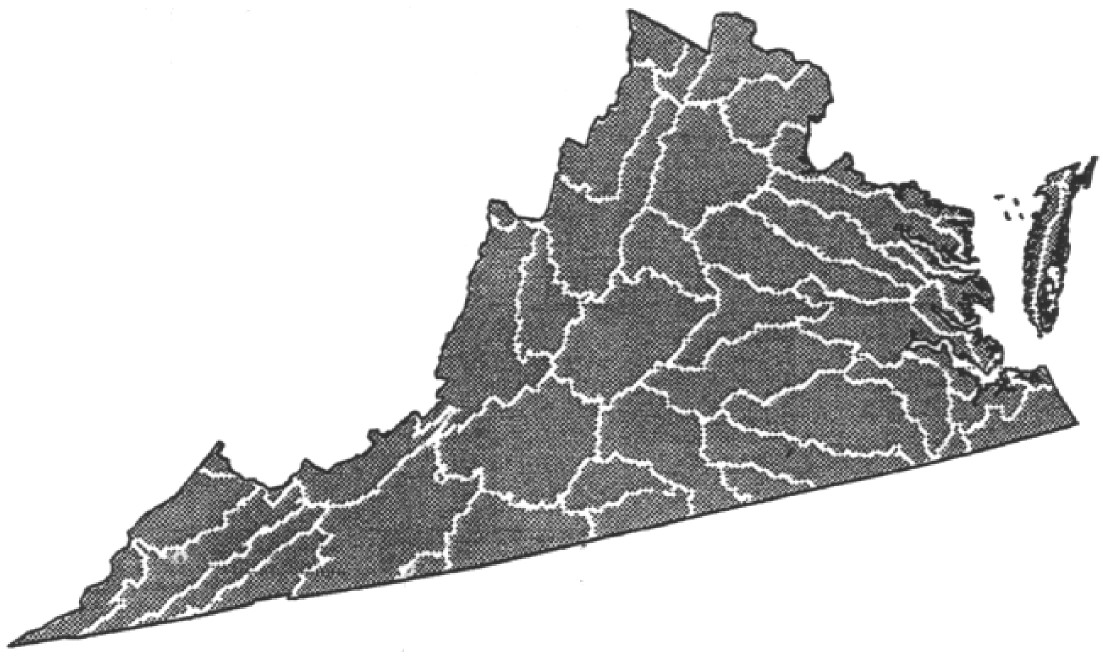


Figure 3. Cataloging units in Virginia.

- RF3 has attributes that describe connectivity, which offers the ability to accomplish upstream/downstream navigation analytically (without topological networking).
- RF3 has a simple and consistent unique numbering system for every stream reach in the United States.
- RF3 has built-in river mileage attributes that describe upstream/downstream distances along river reaches.

Use of RF3 in the Indexing Process

When importing Reach File data from EPA's mainframe computer, an arc attribute table (AAT) is automatically built for each RF3 coverage. The AAT contains the standard AAT fields, plus the items found in Table 1.

The CU item stores the USGS CU number of this piece of RF3. Every arc in the coverage has the same value for CU.

The SEG item stores the number of the stream segment to which the particular arc is assigned. SEG numbers start at 1 and increase incrementally by 1 to 'N' for each CU. A SEG could represent all the arcs of a mainstream, the arcs of a tributary, or piece of a mainstream or tributary. SEG numbers were defined in the production of RF3.

MI stores the marker index for each particular arc. The MI resembles a mile posting along a stream. In reality, the MI field does not truly measure mileage along the RF3 stream network. It does, however, represent a method of producing a unique identifier (in combination with the CU number and the SEG number) for every reach in the United States (see Figure 4).

Table 1. Fields Found in Arc Attribute Table

12070104-ID	CU	SEG	MI	UP	DOWN
1	12070104	1	0.00	-1	0
2	12070104	1	1.30	-1	0
3	12070104	1	2.10	-1	0
4	12070104	2	0.00	-1	0
5	12070104	3	0.00	-1	0
6	12070104	3	1.15	-1	0
7	12070104	4	0.00	-1	0

Together CU, SEG, and MI uniquely identify every arc in RF3 nationwide. These three items are combined in the redefined item called RF3RCHID. This provides a powerful scheme for consistently identifying locations along streams everywhere in the country.

Along with the AAT file, a second attribute file is automatically created for RF3 coverages. This file is always named COVER.DS3. The DS3 file stores a wealth of information about arcs in the coverage. Some of the important fields in the DS3 file contain:

- Upstream and downstream connectivity for navigating along reaches.
- Codes to describe the type of reach (e.g., stream, lake boundary, wide river).
- DLG major and minor attributes.

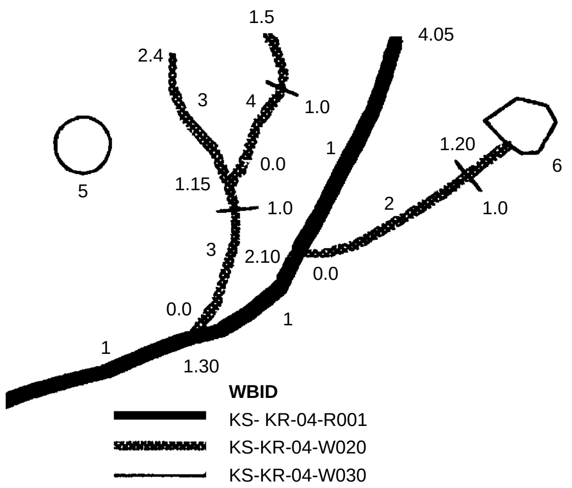


Figure 4. RF3, SEG, and MI data elements.

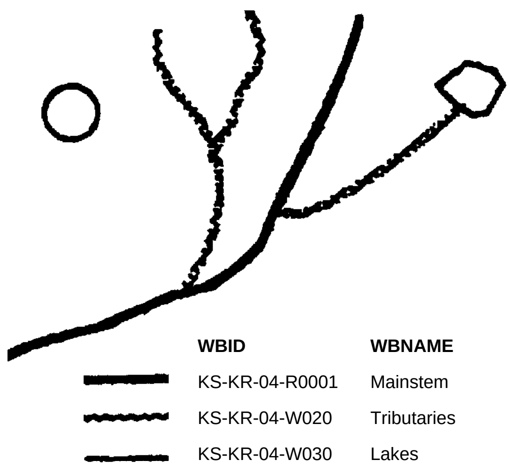


Figure 5. Potential definitions of water bodies.

Waterbody Locations

Because states define water bodies, they provide the only information on waterbody location. South Carolina was indexed to RF3 in 1992, followed by Virginia. Virginia indexed its water bodies using the PCRF program instead of indexing in a GIS environment with ARC/INFO.

PCRF is a PC-based system that indexes water bodies and locates other assessment data from WBS. PCRF stores the definitions of water bodies (including their location) in a file that is linked to other WBS database files that contain information about the assessment status and quality of the waters.

A water body is a set of one or more hydrologic features, such as streams, lakes, or shorelines, that have similar hydrologic characteristics. Water bodies are the basic units that states use to report water quality for CWA 305(b) requirements. Depending on the state's assessment goals and resources, water bodies can be defined in several ways, including (see Figure 5):

- All streams within a watershed
- All lakes and ponds within a watershed
- Sets of streams with similar water quality conditions

PCRF stores locational data for a water body with a unique WBID. WBS uses this WBID as a common field to relate the water body's definition and location to descriptive data about the water body's assessment status and quality. The two most important files used in PCRF are the SCRF1 and SCRF2 files.

The SCRF1 file simply lists the unique water bodies by state. Table 2 offers an example. The most relevant data for reach indexing in this file are the WBID, WBNAME, and WBTYPE, as defined by the state. The WBID, as stated, is a unique identifier for each water body the state has defined. The WBNAME stores a verbal description of the water body. Finally, the WBTYPE defines the type of water body; for example, R is for river, L is for lake.

The SCRF2 file contains an explicit definition of each water body. Because of the complexity involved in defining water bodies, this file may include more than one record for each water body. The SCRF2 file can be considered a waterbody definition *language* because it contains specific codes, attributes, and keys that can be converted into specific reaches on the RF3 data, if read properly (see Table 3). The WBBEGIN and WBEND fields contain explicit CU, SEG, and MI attributes to define the location of the starting point and ending point for the water body. The WBDIR field contains an attribute that describes whether to go upstream or downstream from the WBBEGIN to the WBEND. In addition, a blank WBEND field denotes that the water body should include all upstream or downstream reaches (depending on the WBDIR) of the WBBEGIN reach.

Virginia used PCRF to create an SCRF2 file that contains reach indexing expressions for all of their defined water bodies. ARC/INFO macros were then written to process this file and expand the expressions into the set of specific arcs that compose each water body. The macros will be described in more detail later.

Table 2. Example of SCRF1 File Data

WBID	WBNAME	WBTYPE	WBSIZE
KS-KR-04-R001	KANSAS RIVER	R	15.20
KS-KR-04-W020	LOWER WAKARUSA RIVER	R	61.60
KS-KR-04-W030	MUD CR	R	39.43
KS-KR-04-W040	CAPTAIL CR	R	15.63

Table 3. Example of SCRF2 File Data

WBID	WBDIR	WBBEGIN	WBEND	RFORGFLAG
KS-KR-04-R001	U	10270104001 0.00	1027010400115.20	2
KS-KR-0R-W020	D	10270104005 10.80	10270104005 0.00	2
KS-KR-04-W030	U	10270104059 12.05	10270104007 8.10	3
KS-KR-04-W040	U	10270104038 0.00		3

States that have not already generated indexing expressions in PCRF must provide locations in some other way. The most basic method is for the state to supply a set of 1:100,000 USGS quad sheets that they have marked up with locations of each water body. The maps can be used in conjunction with a digitizer to manually select the appropriate RF3 reaches and code them with the WBID.

The state of Ohio created a GIS database of its river reaches several years ago. The GIS coverage is representational in nature. The stream reaches are 'stick-figures' only. Generally, they fall along the paths of the actual streams, but they are schematic in nature and do not show the true shape of streams. The GIS layer, however, contains the attributes of Ohio's stream reach numbering system, which is used to identify water bodies as well. Ohio's river reach coverage contains information on the locations of water bodies and is being manually conflated to transfer the WBIDs to RF3. The conflation process will be covered later in this paper.

The state of Kansas had previously defined its water bodies on RF2, the precursor to RF3. Kansas defined some indexes by a set of RF3 SEG numbers in a CU and some by the RF3 reaches in a small watershed polygon within a CU. The locations were, in effect, defined within the WBID itself.

Indexing Procedures

Procedures developed for performing waterbody indexing include automated, semiautomated, and manual systems.

Automated Indexing Procedures

As stated, Virginia used PCRF to perform the indexing operation. The state delivered an SCRF2 file containing indexing expressions for all of its water bodies. AML programs were created to read the SCRF2 file and select the reaches specified by each indexing expression. The selected sets of reaches were then coded with the appropriate WBID. The macros were designed to run on one RF3 CU at a time, so the operator specified runs of up to 10 CUs at a time. The macros had to accommodate indexing expressions that included:

- Select reaches upstream of a specified location.
- Select reaches on a reach-by-reach basis.
- Select reaches within a given polygon area.
- Select shorelines of lakes or ponds given latitude and longitude coordinates.
- Select reach downstream from a specified location.

Kansas water bodies were also indexed through an automated process. Kansas supplied an ARC/INFO coverage of small watershed polygons (sub-CU polygons) containing a watershed identifier. The state's WBID contained all other information necessary to determine the RF3 CU and the set of reaches making up each water body. An example of a Kansas WBID is KS-KR-02-W030. This is explained by the following:

- KS refers to the state. All WBIDs in Kansas begin with KS.
- The second component (in this case KR) is an abbreviation of the basin in which the water body falls. KR indicates that this water body is in the Kansas-Lower Republican River basin.
- The third component contains the last two digits of the eight-digit CU number. Although basins comprise several CUs, the last two digits of each CU in a basin are unique; therefore, between the basin (e.g., KR) designation and the last two digits of the CU (e.g., 02), the complete eight-digit CU number in which the water body falls is defined.
- The next letter (in this case W) denotes whether the water body is defined by a watershed polygon (W), an RF2 SEG (R), or a lake or pond shoreline (L).
- Finally, the WBID ends with the number of the polygon (in this case 030) that contains the reaches for the water body in the watershed coverage.

The completed macros could index the entire state in a single run provided that all the WBIDs were contained in single file.

In all cases, Kansas has indexed to RF2 reaches. Only RF3 reaches originally created in RF2 production, therefore, are coded with a WBID.

Manual Indexing Procedures

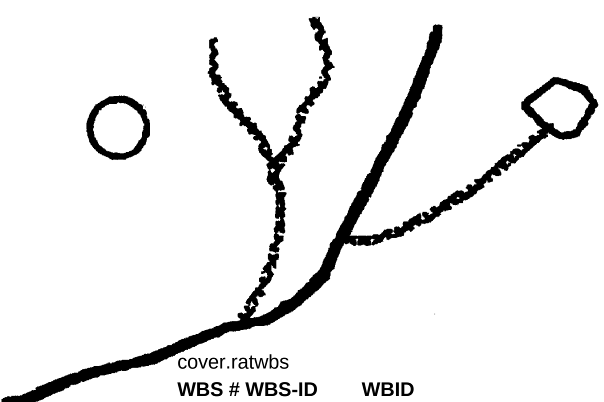
Because Ohio already has a coverage of river reach codes, WBIDs from this coverage had to be transferred to the RF3 reaches they represent. This entailed using a manual conflation process. The operator displayed a CU of RF3 along with the Ohio river reach system for the same area. In a simple process of 'pointing and clicking,' the operator first selected an Ohio river reach arc, then the RF3 arcs that seemed to coincide. As each RF3 arc was selected, it was coded with the WBID of the previously selected Ohio river reach arc.

Other states that have no means of describing water bodies in electronic files may have to mark up paper maps to show waterbody locations. These maps can then be used in a manual process of selecting RF3 reach and coding them with WBIDs either in ARC/INFO or in PCRF.

Using the Route System Data Model To Store Water Bodies

Because water bodies can be defined as noncontiguous sets of arcs and portions of arcs, a robust linear database model is necessary to model these entities. ARC/INFO's route system data model seems well suited for this application. The route system data model allows one to group any set of arcs or portions of arcs into routes. Each route is managed as a feature in itself. Attributes of water bodies are stored in a route attribute table (RAT) and relate to all the arcs defined as the water body. Figure 6 helps illustrate the route system model.

Each route comprises one or more arcs or sections of arcs. ARC/INFO manages the relationship between arcs and routes in the section table (SEC). The structure of the SEC, which is an INFO table, is defined in Table 4. Table 5 reflects how the sections that make up the above routes would appear.



cover.ratwbs

	WBS #	WBS-ID	WBID
	1	1	KS-KR-04-R0001
	2	2	KS-KR-04-W020
	3	3	KS-KR-04-W030

Figure 6. The route attribute table containing waterbody data.

Table 4. Definition of Structure of SEC INFO Table

ROUTELINK#	The route upon which the section falls
ARCLINK#	The arc upon which the section falls
F-MEAS	The measurement value at the beginning of the section
T-MEAS	The measurement value at the end of the section
F-POS	The percentage of the distance along the arc at which the section begins
T-POS	The percentage of the distance along the arc at which the section ends
SEC#	The internal identifier of the section
SEC-ID	The user identifier of the section

Table 5. How Sections Appear in SEC INFO Table

ROUTE-LINK#	ARCLINK#	F-MEAS	T-MEAS	F-POS	T-POS	SEC#	SEC-ID
1	1	0	1.30	0	100	1	1
1	2	1.30	2.10	0	100	2	2
1	3	2.10	4.05	0	100	3	3
2	4	0	1.20	0	100	4	4
3	5	0	1.15	0	100	5	5
3	6	1.15	2.4	0	100	6	6
3	7	0	2.5	0	100	7	7

Representing Water Bodies as Routes

ARC/INFO offers several ways of grouping sets of arcs into discrete routes. One can use ARCEDIT to select a set of arcs to group them into a route, or ARCSECTION or MEASUREROUTE in ARC to group arcs into routes. The method described here uses the MEASUREROUTE command. This method requires that the AAT or a related table has an attribute containing the identifier of the route to which an arc should be assigned. In the application the authors employed, they converted the SCRF2 file into an INFO table containing, for each arc in the coverage, the RF3RCHID of the arc and the WBID to which the arc should be assigned. The WBID item is used to group arcs into routes. One route exists for each unique WBID. Table 6 illustrates the table used in the MEASUREROUTE command method. This table is related to the AAT of the RF3 coverage by the RF3RCHID.

An RAT is automatically created for the coverage, which now can be related to other WBS assessment files for display and query. Figure 4 illustrates the RAT. The most important characteristic of the file is that it has only one record for each water body. This simplifies the display and query of water bodies based on water quality data.

Using EVENTS for Subwaterbody Attributes

Water bodies, as states define them, often constitute a gross aggregation of the water in an area. States often have more specific data about particular stretches of streams within a water body. A system is needed to

Table 6. Table Used in MEASUREROUTE Command Method To Group Arcs Into Routes

$RECNO	RF3RCHID	WBID
1	10270104 1 0.00	KS-KR-04-R0001
2	10270104 1 1.30	KS-KR-04-R0001
3	10270104 1 2.10	KS-KR-04-R0001
4	10270104 2 0.00	KS-KR-04-W020
5	10270104 3 0.00	KS-KR-04-W020
6	10270104 3 1.15	KS-KR-04-W020
7	10270104 4 0.00	KS-KR-04-W020
8	10270104 5 0.00	KS-KR-04-W030
9	10270104 6 0.00	KS-KR-04-W030

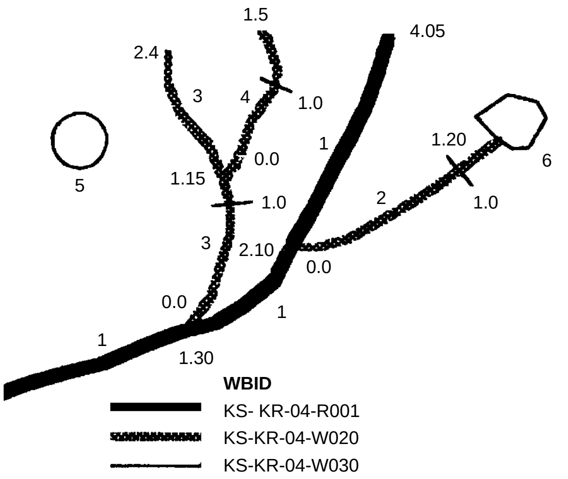


Figure 7. Measurements along SEGs.

query and display data at the subwaterbody level. ARC/INFO's dynamic segmentation tools and event tables are useful for this application. Once water bodies have been defined and reporting methods have been set up based on those water bodies, the task of redefining them is cumbersome.

Event tables can help to keep these waterbody definitions yet still offer the ability to store, manage, and track data at the subwatershed level. Event tables are simple INFO files that relate to route systems on coverages. This concept and data structure can work in conjunction with the predefined waterbody system. We have already seen how a route system called WBS is created in RF3 to group arcs into waterbody routes. This works quite well when displaying water bodies and querying their attributes. A route system based on the WBID cannot, however, act as an underlying base for subwaterbody events because the measures used to create the WBS route system are not unique for a particular route. For example, in the route depicted in Figure 7, three locations are defined as being on WBID KS-KR-04-W020 and having measure 1.0.

The mileage measurements along SEG, however, are always unique (see Figure 8). To use EVENTS, therefore, a second route system must be created based on the RF3 SEG attribute, which provides a unique code for each CU.

The ARCSECTION command, instead of the MEASUREROUTE command, is used to create the SEG route system. This is because the measurement items (MI on the AAT and UPMI on the DS3) already store the summed measures along particular SEGs. Table 7 lists the contents of the resulting RAT table.

Because the name of the route system is SEG, the SEG# and SEG-ID are the names of the internal and user IDs. The SEG item contains the actual SEG number in the RF3 coverage. Because the SEG numbers for each RF3 CU coverage start at 1 and increase incrementally by 1, the SEG item looks much like the SEG-ID and SEG#.

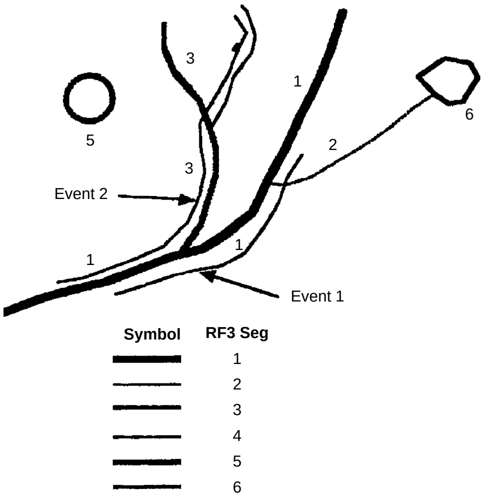


Figure 8. Events located on RF3 data.

Table 7. Route Attribute Table

SEG-ID	SEG#	SEG
1	1	1
2	2	2
3	3	3

Event tables contain a key item, the WBID or SEG, to relate them to the appropriate route system (see Figure 8). They also contain locational information on where to locate the events on the route (either WBID to indicate the water body on the WBS route or SEG to identify the route in the SEG route system). Separate event tables can then relate use support, causes, and sources as linear events. FROM and TO store the starting and ending measures for each event. Using event tables allows us to apply many useful cartographic effects (e.g., hatching, offsets, text, strip maps). Events can be queried both in INFO and graphically. Event data can help in producing overlays of two or more event tables.

An event table can display use support information (see Table 8). WBS users can update their event tables using RF3 maps supplied by EPA without having proficiency in ARC/INFO. ARC/VIEW2 is expected to support events and route systems. This will give users powerful tools for spatial query of assessment data. Developing event tables would also display and query data on the causes and sources of use impairment. These events can be offset and displayed to show the areas of interaction. More permanently, preparing line-on-line overlays can show intersections and unions.

An alternative approach is to use an EVENT-ID as a unique identifier for each event. The SEG field stores the number of the route (SEG) upon which the event occurs. FROM and TO store the beginning and end measures along the route upon which the event occurs. WBID contains the identifier of the water body upon which the event occurs (see Table 9). An event can occur within a single SEG, across two or more SEGs, within a single water body, or across two or more water bodies.

Additional attribute tables can be created to store descriptive attributes for each event. These tables would resemble the SCRF5 and SCRF6 files except that instead of using the WBID to relate to a water body, a field called 'EVENT-ID' would link the use, cause, and source data to a particular event (see Table 10).

Both approaches offer some advantages. In either case, they allow us to map our water quality assessment data and communicate it in a meaningful and useful way.

Table 8. Event Table That Reflects Use Support Information

SEG	FROM	TO	WBID	USE	USE SUPPORT
1	0.80	1.30	KS-KR-04-R0001	21	Fully
1	1.30	2.10	KS-KR-04-R0001	21	Partial
1	2.10	2.31	KS-KR-04-R0001	21	Not supported
1	0.50	1.30	KS-KR-04-R0001	40	Threatened
3	0.00	1.15	KS-KR-04-W030	21	Fully
4	0.00	2.5	KS-KR-04-W040	40	Not supported

Table 9. Using EVENT-ID as a Unique Event Identifier

EVENT-ID	SEG	FROM	TO	WBID
1	1	0.80	1.30	KS-KR-04-R0001
1	1	1.30	2.10	KS-KR-04-R0001
2	4	0.00	2.5	KS-KR-04-W040

Table 10. Using EVENT-ID To Link Use, Cause, and Source Data to an Event

EVENT-ID	ASCAUSE	ASSOURC
1	900	1200
1	-9	1100
1	0500	1100
2	1200	9000
2	0900	8100

Nonpoint Source Water Quality Impacts in an Urbanizing Watershed

Peter Coffin, Andrea Dorlester, and Julius Fabos
University of Massachusetts at Amherst, Amherst, Massachusetts

Abstract

As part of the larger Narragansett Bay Estuary Project, the University of Massachusetts Cooperative Extension Service contracted with the university's METLAND research team to develop a geographic information system (GIS) database, generate watershed-wide maps, perform analyses, and develop a modeling procedure. The objective was to educate local officials about the impacts of development on water quality and to help local boards minimize the effect of nonpoint sources of pollution.

Because the receiving waters of the Narragansett Bay are located far downstream in Rhode Island, the upstream communities in Massachusetts are reluctant to enact measures to improve water resources outside of their jurisdictions. A GIS was used to create awareness of existing downstream problems and to show the upstream communities how development will ultimately affect water resources in their own backyards.

To nurture this awareness, a "buildout" analysis was conducted for an entire upstream subwatershed, the Mumford River watershed, containing parts of four towns, and roughly 50 square miles. This buildout was coupled with a loading model using Schueler's Simple Method to illustrate the potential impacts of future development, and encourage local boards to minimize future nonpoint sources of pollution.

GIS proved its usefulness by developing customized maps for each town, by generating several "what if" scenarios showing the impacts of different zoning changes, by facilitating long-range planning for small towns without professional staff, and by encouraging a regional perspective on development issues. The entire planning process was most successful in creating a series of partnerships that will continue after the grant expires. The university shared coverages with the state GIS agency, creating new coverages not previously available, specifically soils, ownership, and zoning. Small towns learned about the potential of the new technology. Students gained from hands-on experience with real-world problems. State agencies saw their efforts understood at the local level, especially as they reorganize on a basin approach and begin to implement a total mass daily loading (TMDL) procedure to coordinate permitted discharges and withdrawals.

As greater emphasis is placed on controlling nonpoint sources of pollution, more attention needs to be focused on local boards, who control land use decisions in New England.

Introduction

Project Description

Narragansett Bay is a vital resource for southern New England. The health of its waters is critical to the regional economy, supporting fisheries, tourism, and quality of life. Increased development along the bay's shorelines and throughout its drainage basin threatens the quality of these waters, however. The U.S. Environmental Protection Agency (EPA) recognized the threats to this important water body and designated the Narragansett Bay under its National Estuary Program in 1985.

Completing a Comprehensive Conservation and Management Plan (CCMP) for Narragansett Bay took 7 years. The CCMP identified seven priority areas for source reduction or control, including the reduction of agricultural and other nonpoint sources of pollution. The nonpoint source strategy identified United States Department of Agriculture (USDA) agencies, conservation districts, and other public and private organizations as having principal roles in nonpoint source management.

Whereas the vast majority of Narragansett Bay lies within the boundaries of Rhode Island, a significant portion of its pollution load originates in Massachusetts. Recognizing that the watershed extends beyond state boundaries, the USDA provided 3 years of funding to Cooperative Extension and the Soil Conservation Service (SCS) in both Massachusetts and Rhode Island to coordinate their efforts in an innovative attempt to reduce the impact of nonpoint sources of pollution on Narragansett Bay. While water quality is a relatively new focus for Cooperative Extension, it fits well with the historic mission of extending the knowledge base of the land-grant colleges out into the community, and providing training and capacity building for local officials and community organizations.

With such a large area of concern, the management team decided to focus on a smaller subwatershed area in each state for the first 2 years. The strategy was first to develop a program for the mitigation of nonpoint source pollution on the smaller scale of a watershed of roughly 50 square miles, then take the lessons learned and apply the most appropriate efforts throughout the larger watershed. By using similar strategies in Rhode Island and Massachusetts, but choosing subwatersheds that differ in terms of location relative to the receiving water, size, staffing, and sophistication, the two states gained from each other's experience, sharing the successful techniques and avoiding each other's mistakes.

For its pilot study, Rhode Island chose Aquidneck Island, home of Newport, Portsmouth, and Middletown, with a special focus on protecting surface water supply reservoirs. Massachusetts chose an upstream watershed in the Blackstone Valley, somewhat rural in character, but rapidly undergoing a transformation to suburbia.

Watershed Description

The Blackstone River drops 451 feet in its 48-mile journey from Worcester, Massachusetts, to Pawtucket, Rhode Island. In the 19th century, this drop of roughly 10 feet per mile was ideally suited to providing power to mills during the early years of the industrial revolution. By the Civil War, every available mill site was developed, earning the Blackstone River the name "The Hardest Working River."

The Blackstone has a long history of pollution. First, the textile industry, then steel, wire, and metal finishing industries used the river for power, in their manufacturing process, and for waste disposal.

In Massachusetts, the Blackstone River is the major source of many pollutants to Narragansett Bay. Based on total precipitation event loading calculations, the Blackstone River is the principal source of solids, cadmium, copper, lead, nitrate, orthophosphate, and PCBs to the bay (1). The Blackstone River has an average flow of 577 million gallons per day or 23.2 percent of the freshwater input to the bay.

The watershed area in Massachusetts equals 335 square miles; with a population of 255,682, this results in a density of 763 people per square mile. The Blackstone Valley has 9,000 acres in agricultural use, with more land in hay (4,500 acres) than crops (3,700 acres) to support its 4,400 animals.

Based on aerial mapping flown in 1987, the Blackstone Valley has lost 5 percent of its cropland, 9 percent of its pasture, and 21 percent of its orchards since 1971. The valley remains more than 60 percent forested, but that represents a decrease of 5 percent. The forest and farmlands were lost to development as low density housing grew by 45 percent, commercial use grew by 15 percent, and transportation grew by 54 percent. Waste disposal grew 52 percent to 582 acres, and mining, which in this region represents gravel pits, grew 22 percent to 1,100 acres.

Watershed soils consist mainly of compact glacial till on rolling topography, with 3 to 15 percent slopes. The river and stream valleys are underlain by glacial-derived sand and gravel outwash, which provide drinking water to all towns in the area except Worcester and support the large gravel pits. The high clay content in the till soils of the uplands makes for a high water table, which is beneficial for growing corn but causes problems for septic systems.

Following a preliminary study of the subwatersheds, the Mumford River in the Blackstone Valley was selected as the focus watershed based on its size, location, land use, and existing water quality (see Figure 1). The Mumford River watershed has an area of 57 square miles, with a length of 13 miles, and lies within the towns of Douglas, Northbridge, Sutton, and Uxbridge. These towns share the attributes of small, rural communities undergoing rapid development, with no professional planning staff (see Figure 2). According to the 1990 Census, Douglas grew 46 percent in 10 years to 5,438; Uxbridge experienced 24 percent growth to 10,415; Sutton increased 17 percent to 6,824; and Northbridge grew 9 percent to 13,371.

Project Strategy

Because the generation of nonpoint sources of pollution is so closely tied to land use, and because local boards composed of citizen volunteers have principal control over land use in New England, the key focus of this program is to train local boards to recognize and begin managing the threat that nonpoint sources of pollution pose to water quality. Local planning boards, conservation commissions, and boards of health address land use issues and can regulate and shape existing and proposed development. By developing a program to train local officials, Cooperative Extension can focus its outreach where it will have the greatest impact in both the short and long term. Local boards have the strongest opportunity to comment on how land is to be used as it undergoes development. Therefore, this project focused on preventing future deterioration as opposed to fixing

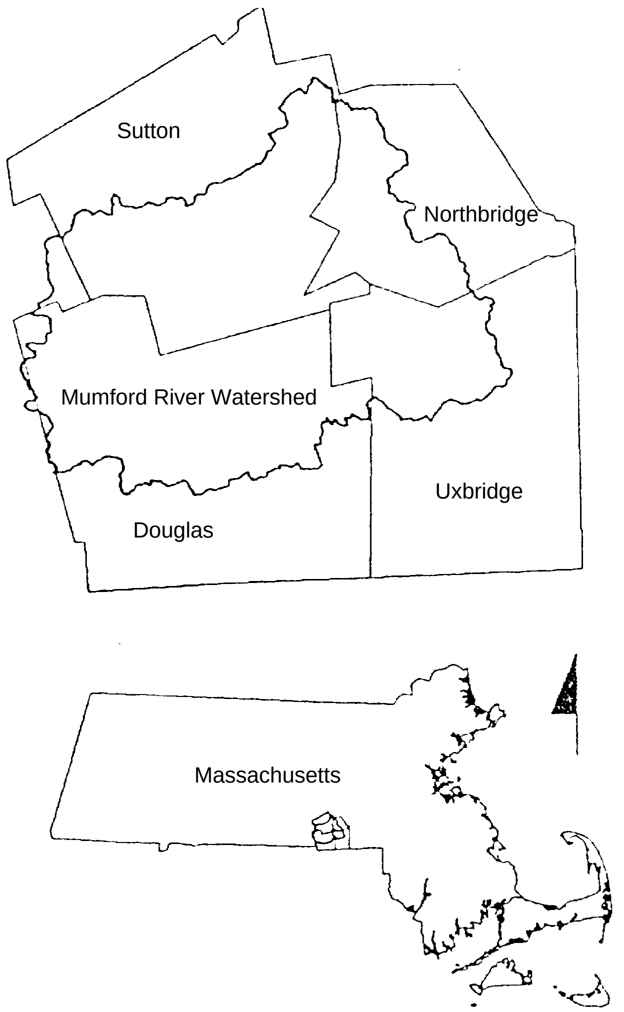


Figure 1. Map of Mumford River watershed study area.

existing problems. This is especially appropriate in a rapidly urbanizing setting.

Both Massachusetts and Rhode Island chose to utilize GIS technology because of its ability to store, analyze, transform, and display geographic, or spatial information. Its database management and analytical capabilities make it a useful tool for pollution load modeling and buildout scenario development, while its mapping capabilities make it an excellent tool for sharing information with local officials. This paper documents a case study on how GIS technology was used to apply a watershed-wide pollution loading model and to develop buildout scenarios for demonstrating to local officials the potential impacts of future development on water quality.

This project used GIS in four different applications:

- *Printing customized, large-scale maps:* This most basic application of a GIS proved the most useful for local officials. It was a revelation for some officials to see how their current zoning related to actual land use. In one town, these maps inspired a change in zoning to protect the area of a future water supply reservoir. These maps helped officials see how their towns fit into the regional picture and how their zoning and land use affected the adjoining towns, and vice-versa.
- *Performing "buildout" analysis:* A "buildout" analysis demonstrates the consequences of existing zoning. It assumes that all land that can be developed will be developed at some future date. In essence, it is a spreadsheet that divides the land available for development in each zone by the required lot size, subtracting

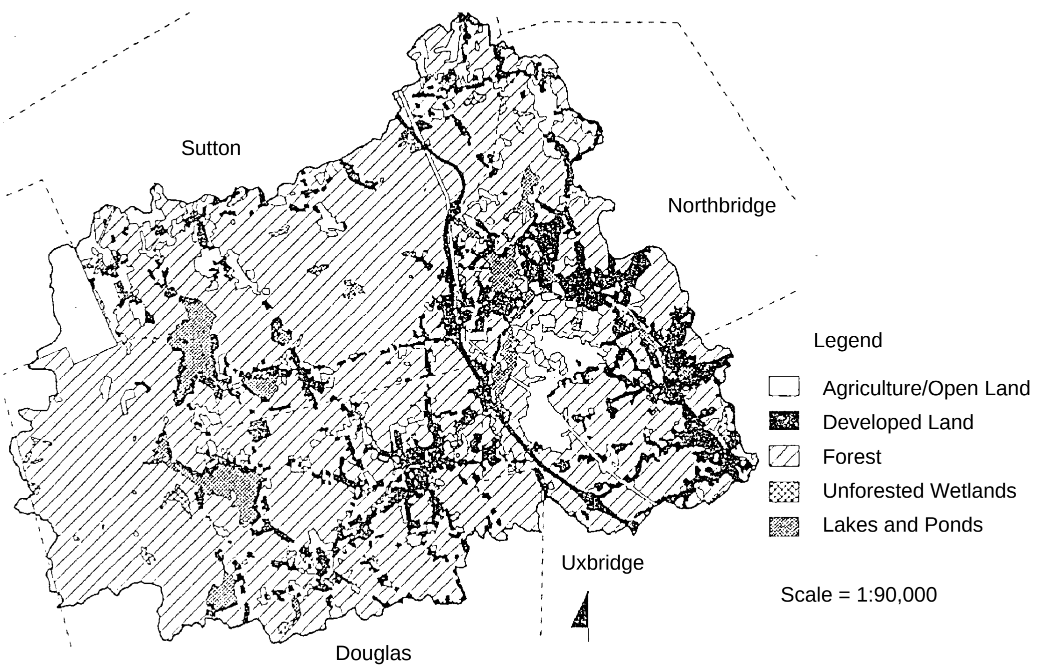


Figure 2. Land use/land cover map of Mumford River watershed.

a certain percentage for the road network and steep slopes. It is best used to evaluate different development scenarios, substituting different zoning requirements.

- *Applying a watershed-wide pollutant loading model:* GIS provided the input needed to apply the "Simple Method" for estimating existing and potential pollutant loads. Future pollution loading was estimated using a buildout with existing zoning and again assuming the implementation of cluster zoning. The Simple Method was compared in one subwatershed with the Galveston Bay Method, which accounts for the hydrologic class of the soils.
- *Promoting planning for a greenway:* Land use maps were overlaid with parcel ownership to show the existing network of preserved open space and to identify those parcels of land having significant wildlife habitat and recreational value. In one town, these maps were used to gain funding for planning a river walk.

Database Development

The most daunting aspect of using a GIS is the prospect of spending a great deal of time and money creating a useful database. Fortunately for Massachusetts, many of the basic coverages needed for regional planning are housed in a state agency, MASS GIS, and are available for a small processing fee. These coverages include most of what appears on the standard United States Geological Survey (USGS) map: roads, streams, town boundaries, as well as watershed boundaries and land use data generated from the interpretation of aerial photographs. The university entered into an agreement whereby we gained access to this data at no charge, in return for sharing the new coverages that the project would generate.

New coverages needed for the study included: zoning, soils, sewer and water lines, and land ownership, or parcels taken from the assessor's maps. The soils maps were obtained from the SCS, digitized by hand, then the scale was converted with a computer program, "rubber-sheeting," to achieve a uniform scale of 1:25,000. All other new coverages were transferred onto a USGS topographical map at a scale of 1:25,000, then digitized directly into the computer. We obtained elevation data, but the triangulation process used to convert elevation data to slopes would require so much time and memory that, for our purpose, deriving a slope map from the four classes identified on the soils map was sufficient.

While GIS computer programs are powerful enough to perform most overlay and analysis functions necessary in nonpoint source pollution load modeling, database development and accuracy issues can limit the effectiveness of such modeling. The choice of which model to use is a function of which data are available for input. Physics-based distributed models are more precise but require detailed input parameters, beyond the scope of this project. The extent of our database limited us to lumped-parameter empirical models. We chose two such models, the Simple Method and the Galveston Bay Method.

GIS Applications

The Simple Method

Schueler (2) developed the Simple Method, one of the simplest lumped-parameter empirical models. The input data necessary to compute pollutant loading with the Simple Method are land use, land area, and mean annual rainfall. Land use determines which event mean concentration (EMC) values and percentage of imperviousness to use in the computation. The amount of rainfall runoff is assumed to be a function of the imperviousness of various land uses. More densely developed areas have more impervious surfaces, such as rooftops and paving, which cause stormwater to run off the land rather than be absorbed into the soil. The Simple Method can generate rough figures for annual pollutant loading within a watershed and can effectively show relative increases in pollutant levels as land is developed.

The formula used in the Simple Method is as follows:

$$L = [(P)\,(Pj)\,(Rv)/12] \;*\; (C) \;*\; (A) \;*\; (2.72)$$

$$(\text{load}) \;=\; (\text{runoff}) \;*\; (\text{EMC}) \;*\; (\text{area})$$

where:

L = pounds of pollutant load per year
P = rainfall depth (inches) over the desired time interval (1 year)
Pj = percentage of storms that are large enough to produce runoff (90 percent)
Rv = fraction of rainfall that is converted into runoff ($Rv = 0.05 + 0.009\,(I)$, where I represents the percentage of site imperviousness)
C = flow-weighted mean concentration (EMC) of the targeted pollutant in runoff (milligrams per liter)
A = area (in acres) of the study region

The Simple Method can be applied using a hand-held calculator or a computer spreadsheet program. For this project, the calculations were performed entirely within the ARC/INFO GIS environment, where the input data were stored. Results were exported to the Excel spreadsheet program for presentation purposes.

The application of the Simple Method consists of three major steps.

Step 1: Aggregate Land Uses and Obtain Area Figures for Land Use Categories Within Each Subbasin

The land use coverage in our database has 21 categories. For the purpose of applying the Simple Method, these were aggregated into the following six major categories, based on development density: undeveloped forest and other open land, large-lot single-family residential, medium-density residential, high-density residential, commercial, and industrial. The aggregated land use categories were matched with study basins from the Nationwide Urban Runoff Program (NURP) for the purpose of assigning EMC values.

Step 2: Enter Percentage Imperviousness and Event Mean Concentrations for Each Land Use Type

The TABLES module of ARC/INFO was used to assign percentage of imperviousness and EMC values to individual land use polygons within the watershed's subbasins. The estimated percentage of imperviousness was obtained from Schueler's guide to using the Simple Method (2). EMC values for three pollutants—phosphorous, nitrogen, and lead—were taken from selected NURP study basins and were assigned to the aggregated land uses within the watershed.

Step 3: Input the Simple Method's Mathematical Loading Formula, Calculate Loading Results for Each Distinct Land Use Area, and Sum Results by Watershed Subbasin

Finally, the pollutant load was calculated for each distinct land use area within the Mumford River watershed by inputting the loading formula through the TABLES module of ARC/INFO. The mean annual rainfall figure was assumed to be that of Worcester, Massachusetts, or 47.6 inches. After calculating loading figures for phosphorous, nitrogen, and lead for each distinct land use area, these numbers were summed for each watershed subbasin, using the ARC/INFO frequency table reporting capability.

The Galveston Bay Method

As an experiment, we applied the Galveston Bay Method to one of the subbasins to compare results with the Simple Method. The slightly more sophisticated Galveston Bay model considers soil drainage characteristics in addition to land use/imperviousness to determine rainfall runoff. This method is similar to the Simple Method, in that amount of rainfall runoff and EMCs for particular land uses are multiplied by land area to determine total pollutant load (3). Runoff in this method, however, is calculated using the USDA SCS's TR 55 runoff curve model. The SCS model calculates runoff as a function of both land use and soil type. Runoff equals total rainfall minus interception by vegetation, depression storage, infiltration before runoff begins, and continued infiltration after runoff begins (4).

The formula used with the Galveston Bay Method has a structure similar to that of the Simple Method and is as follows:

$$L = \frac{P - 0.2\,[(1000/CN) - 10]^2}{P + 0.8\,[(1000/CN) - 10]} * EMC * A$$

$$\text{(load)} = \qquad \text{(runoff)} \quad * \quad \text{(EMC)} \quad * \quad \text{(area)}$$

where:

L = milligrams of pollutant load per year
P = mean annual rainfall amount
CN = runoff curve number, which is a function of soil type and land use
EMC = event mean concentration
A = area (in acres) of the study region

The application of the Galveston Bay Method consists of four major steps.

Step 1: Aggregate Land Uses and Obtain Area Figures for Land Uses Categories. Aggregate Soils According to Drainage Classes

Land use types were aggregated into the same six major categories as the Simple Method in order to match EMC values and to allow for later comparison of the two pollution loading methods. Soils were aggregated according to drainage classes for use with the USDA SCS TR 55 runoff formula. The SCS identifies four classes of soils according to their drainage capacity:

Class A = excessively to well-drained sands or gravelly sands.
Class B = well to moderately drained, moderately coarse soils.
Class C = moderately to poorly drained fine soils.
Class D = very poorly drained clays or soils with a high water table.

Step 2: Overlay Soils Data With Land Use Data and Clip This New Coverage Within the Subbasin

The ARC/INFO GIS overlay capability was used to overlay land use and soils maps for the Mumford River watershed on top of each other. This created new, distinct areas of different land use and soils combinations. Because we were only applying this model in one subbasin, the subbasin boundary was used in conjunction with the ARC/INFO "clip" command to cut out (like a cookie cutter) that portion of the watershed within the subbasin.

Step 3: Assign Runoff Curve Numbers and EMC Values to Each New Land Use/Soils Polygon Within the Subbasin

EMC values were assigned to each distinct land use/soils area in the same manner as they were assigned to land use areas using the Simple Method.

Runoff curve numbers were assigned to each distinct land use/soils area within the subbasin according to values established by the USDA SCS.

Step 4: Calculate Loading Results for Each Distinct Land Use/Soils Area and Sum Results for the Entire Subbasin

Finally, the pollutant load was calculated for each distinct land use/soils area within the subbasin by inputting the loading formula through the TABLES module of ARC/INFO. After calculating loading figures for phosphorous, nitrogen, and lead for each distinct land use/soils area, these figures were then summed for the subbasin using the ARC/INFO frequency table reporting capability. Results of this modeling were converted from milligrams per acre per year to pounds per acre per year to facilitate later comparisons.

Buildout Scenarios

For planning purposes, GIS is most useful in its ability to quickly generate alternative scenarios. When these development scenarios are coupled with a pollutant load model as described above, alternative scenarios can be evaluated according to their impact on water quality. This project generated two different scenarios for each of the four towns in the watershed: a maximum buildout with existing zoning and a maximum buildout with clustered development.

Maximum Buildout

A maximum buildout scenario was used to show the worst case for development according to current zoning regulations (see Figure 3). The result of this buildout is expressed both in the number of new residential units to be built and in the area of land to be converted from undeveloped to residential and other urban uses.

Step 1: Eliminate From Consideration All Land That Is Already Developed

Step 2: Eliminate From Consideration All Land That Is Under Water

Step 3: Eliminate From Consideration All Land That Is Protected From Development

These protected lands included cemeteries, parks, and all land permanently restricted from development.

Step 4: Reduce the Remaining Amount of Land by 20 Percent To Account for New Roadways and Extremely Steep Slopes

The remaining land was considered to have "developable" status. Wetlands were included in this category because while a house probably would not be built on a wetland, wetlands can and often do constitute portions of the required lot size of large residential lots.

Step 5: Overlay the Land Use Coverage With Zoning and Minimum Lot Size Information

This created new land use areas as a function of zoning. All forests and fields were converted to a developed status.

Step 6: Divide Net Developable Land Area Within Each Zone by Minimum Lot Size Allowed To Obtain the Number of New Units

Results from the buildout are expressed in the number of new units. Results can also be shown spatially by shading in areas on the map according to future density

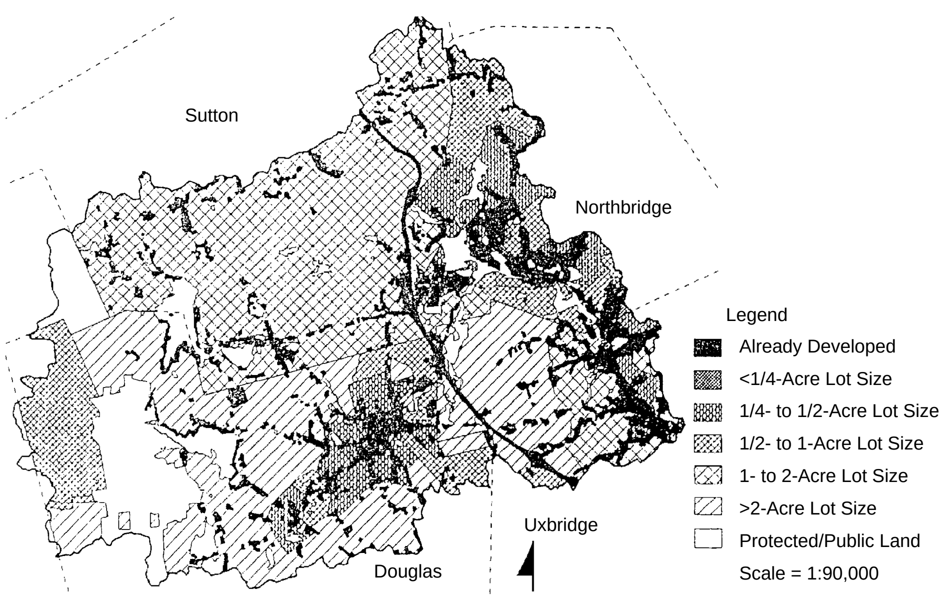


Figure 3. Maximum buildout scenario within the Mumford River watershed.

of development (darker shades for higher density, lighter shades for lower density).

Clustered Buildout

Another alternative development scenario was generated assuming the implementation of clustered development. All areas zoned for lots larger than 1 acre were changed to cluster zones, where three-fourths of the land area remains undeveloped, and the remaining one-fourth of the land area is developed at a density of 1/2-acre lot size. With clustering, an area zoned for 2-acre house lots still supports the same number of new units, but three-quarters of the land area remains open space for passive recreation, protected wildlife habitat, and as a buffer zone to filter runoff.

Step 1: Select All Land Available for Development Zoned for 1-Acre Lots or Larger

Step 2: Multiply Selected Land by 0.75 and Add to the Category of Protected Land

Step 3: Multiply Selected Land by 0.25 and Change the Minimum Lot Size to One-Half an Acre

Step 4: Divide Step 3 by 20 Percent To Allow for New Roads and Steep Slopes

Step 5: Divide Step 4 by 21,780 (One-Half an Acre) To Determine Number of New Housing Units

Results

Lumped-parameter empirical models were chosen for this project and were applied to watershed subbasins ranging in size from 1 to 20 square miles and having an average of 4 square miles. The application of the Simple Method to existing land use conditions allowed for a comparison of the Mumford River watershed's subbasins for the purpose of identifying the subbasins that contribute the highest levels of pollutants per acre per year. The development of a maximum buildout scenario identified those areas within the watershed that will sustain the greatest amount of new growth. The application of the Simple Method to this maximum buildout scenario revealed that pollutant levels in surface water runoff would increase substantially for all subbasins in the watershed. This finding supports the theory of a positive relationship between development and increased pollutant levels from surface water runoff.

The development of a customized buildout scenario for future development identified those areas that are currently zoned for large-lot residential "sprawl" and that can support higher development density under cluster zoning, while protecting a significant amount of open space that can support a variety of beneficial uses. The application of the Simple Method to the customized buildout scenario revealed that the use of cluster development can reduce future levels of water pollution, especially from nutrients (see Figures 4 and 5).

Results determined by applying the Galveston Bay Method to one subbasin were compared with those obtained using the Simple Method. The predicted pollutant loading from current conditions differed significantly

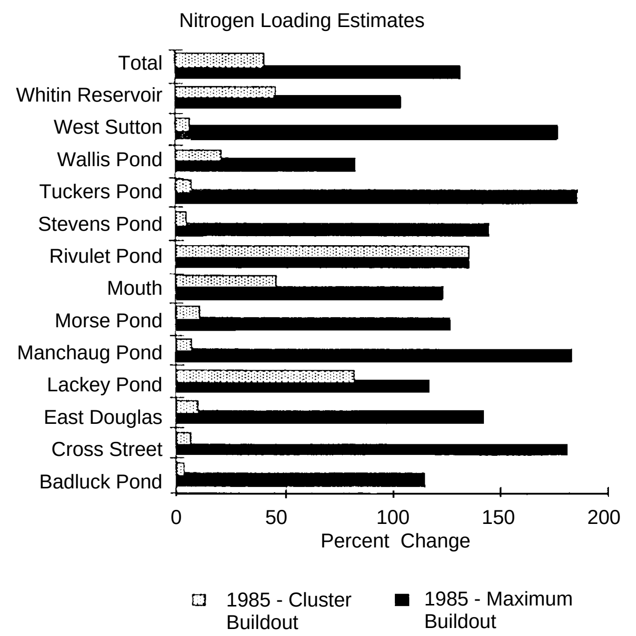


Figure 4. Chart showing difference in simple method results for nitrogen loading between maximum and customized buildout scenarios.

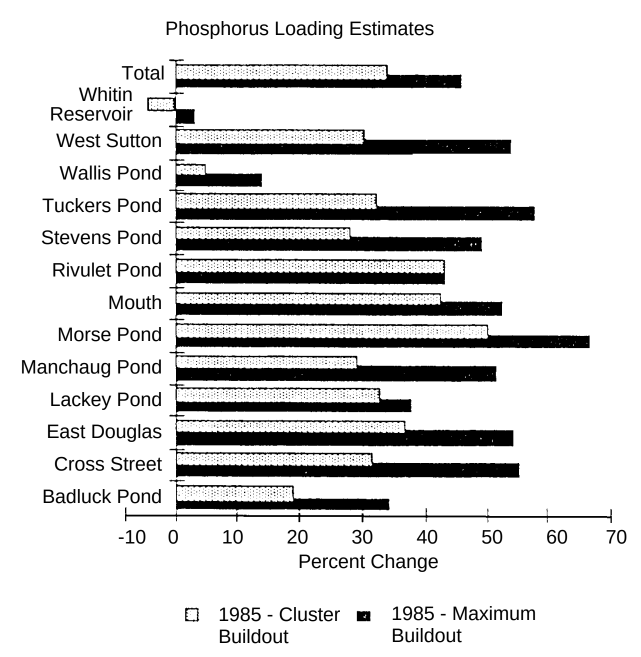


Figure 5. Chart showing difference in simple method results for phosphorus loading between maximum and customized buildout scenarios.

between the two methods; the Simple Method consistently predicted five times the amounts generated by the Galveston Bay Method.

When the two methods were applied to both the maximum and customized buildout scenarios, however, the percentage growth of predicted pollutant loadings was remarkably similar for both methods; the Simple Method consistently predicted loadings 10 to 15 percent greater than the Galveston Bay Method. This indicates that while the Galveston Bay Method may provide more accurate results in predicting actual pollutant loading, the Simple Method is adequate enough for evaluating and comparing different development scenarios (see Figures 6 and 7).

Discussion

As states begin to implement a TMDL approach to regulating water quality, they face the quandary of how to determine the extent of nonpoint source pollution in the rivers. The crudest method is to subtract from the total load those quantities generated by point sources and call all the rest nonpoint source. While this is appropriate in some settings, it is unacceptable in a watershed with a long history of pollution because a significant source of pollution is the resuspension of historical sediments stirred up by storms. The situation demands the development of a model to predict the loading from nonpoint sources. Only a computer can handle the multiple factors that interact to generate nonpoint sources of pollution.

As greater emphasis is placed on watershed planning, the abilities of a GIS to input, store, manipulate, analyze, and display geographic information become indispensable. As the scientific community improves its knowledge base for determining the critical factors influencing nonpoint source pollution, GIS technology is improving in its ability to store and handle large amounts of data.

While a detailed, physics-based distributed model would be more accurate than the lumped-parameter models used for this project, they are difficult to apply at the watershed scale. The real limiting factor is the provision of all the data coverages needed to apply complex models. Lumped-parameter models, such as the Simple Method and the Galveston Bay Method, are ineffective for accurately predicting pollutant loads, but they are suitable for comparing and evaluating alternative development scenarios.

Time, and the development community, will not wait until all the answers are known. Local officials continue to approve development with no thought to the impacts on water quality. These officials need to be informed about the implications of haphazard growth. A GIS, with its ability to generate customized maps and quickly evaluate alternative development scenarios, is a powerful tool to help local officials visualize how the decisions they make on paper today will have an impact on the land tomorrow.

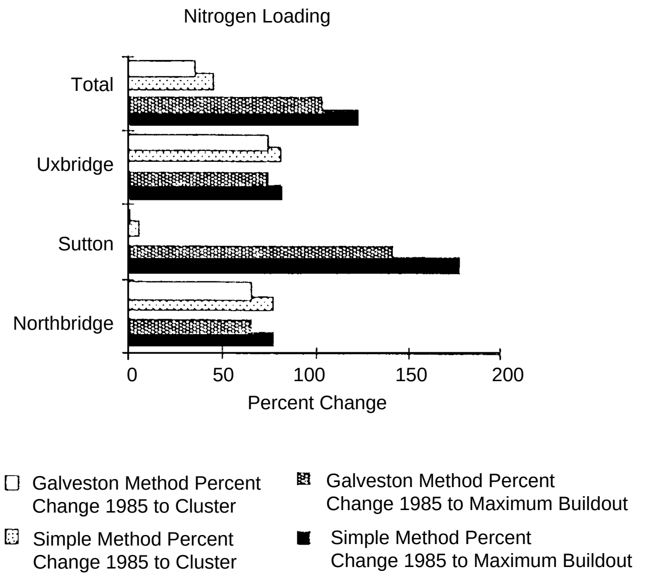


Figure 6. Chart showing difference between Simple Method results and Galveston Bay Method results for nitrogen loading.

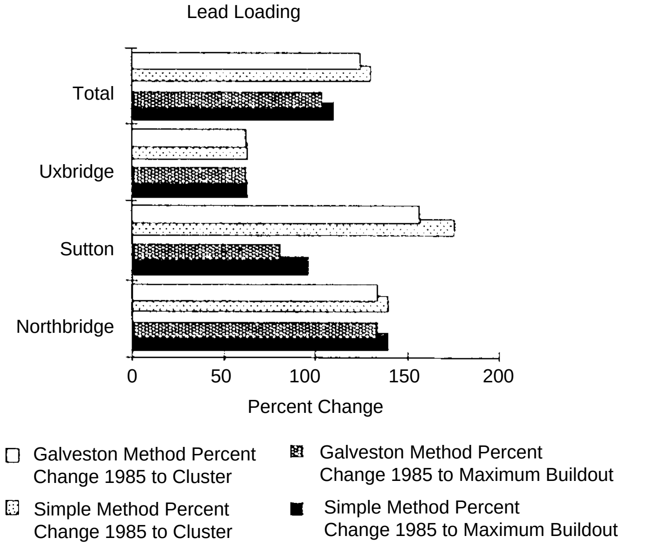


Figure 7. Chart showing difference between Simple Method results and Galveston Bay Method results for lead loading.

References

1. Metcalf & Eddy, Inc. 1991. Assessment of pollutions in Narragansett Bay. Draft report to the Narragansett Bay Project.

2. Schueler, T. 1987. Controlling urban runoff: A practical manual for planning and designing urban BMPs. Washington, DC: Metropolitan Washington Council of Governments.

3. Newell, C., et al. 1992. Characterization of nonpoint sources and loadings to Galveston Bay, Vol. I, Technical Report: Galveston Bay National Estuary Program (March).

4. U.S. Department of Agriculture (USDA) Soil Conservation Service. 1975. Urban hydrology for small watersheds. Technical Release 55. Springfield, VA: U.S. Department of Agriculture.

Additional Reading

Arnold, C.L., et al. 1993. The use of Geographic Information System images as a tool to educate local officials about the land use/ water quality connection. Proceedings of Watershed Conference, Alexandria, VA.

Chesebrough, E. 1993. Massachusetts nonpoint source management plan. Massachusetts Department of Environmental Protection, Office of Watershed Management (October).

Joubert, L., et al. 1993. Municipal training for water quality protection. Contribution #2845. College of Resource Development, University of Rhode Island/Rhode Island Agricultural Experiment Station.

Massachusetts Geographic Information System. 1993. MassGIS Datalayer descriptions and guide to user services. Boston, MA: Executive Office of Environmental Affairs.

McCann, A., et al. 1994. Training municipal decision-makers in the use of Geographic Information Systems for water resource protection. Contribution #2927. College of Resource Development, University of Rhode Island/Rhode Island Agricultural Experiment Station.

Narragansett Bay Project. 1992. Comprehensive conservation and management plan: A summary (January).

U.S. EPA. 1983. Results of the Nationwide Urban Runoff Program (NURP), Vols. I and II. Final report prepared by Water Planning Division. NTIS PB84185537. Springfield, VA: National Technical Information Service.

U.S. EPA. 1992. Compendium of watershed-scale models for TMDL development. EPA/841/R-92/002. Washington, DC (June).

The GIS Connection to Residential Yard Soil Remediation

Jennifer L. Deis and Robin D. Wankum, P.E.
Black & Veatch Special Projects Corp.
Overland Park, Kansas 66211

Abstract

By using innovative approaches to geographic information system (GIS) applications, the U.S. Environmental Protection Agency (EPA) and Black & Veatch Special Projects Corp. (BVSPC), as a contractor to EPA, were able to implement a site investigation concurrently with the site's cleanup. This effort ultimately saved the EPA approximately $1.2 million dollars. The purpose of the investigation was to locate and prioritize the residential yards that were adversely affected by mining activities. BVSPC used an environmental data management system (EDMS) to consolidate x-ray fluorescence (XRF), global positioning system (GPS), and laboratory analytical data into a unique and flexible electronic, GIS-compatible database. The application of the EDMS with a desktop GIS allowed effective completion of the investigation of the Oronogo-Duenweg Mining Belt Site. This paper will present the GIS approach used to expedite the investigation and cleanup activities at the site and will identify the benefits of this process.

Site History

The Site is part of the Tri-State Mining District, which covers hundreds of miles in southwestern Missouri, southeastern Kansas, and northeastern Oklahoma. Mining, milling, and smelting of lead and zinc ore began in the 1850s and continued in the District until the 1970s. The activities generated approximately 9 million tons of mine and mill wastes and smelter related materials. These wastes, which contain high levels of lead, cadmium, and zinc, were deposited throughout Jasper County. Additionally, air emissions during smelting operations resulted in the contamination of soil surrounding the smelters. Approximately 6,500 residences are now located within the 60 square-mile Superfund Site where lead and zinc were mined.

The Missouri Department of Health (MDOH) conducted an exposure study at the Site to evaluate health effects on residents living in the area (1). The study concluded that the most significant source of contamination resulting in elevated blood-lead levels was residential yard soils. Preliminary characterization of the surface soils within a 3/4-mile radius of the largest smelter indicated that 86 percent of those yards exceeded the 800 parts per million (ppm) target cleanup level for lead. In addition, previous investigation activities conducted by others concluded that 50 percent of the homes within 200 feet of mine and mill waste pile locations also exceeded the target cleanup level for lead (2). The site location and its designated areas, areas where mining activities and/or wastes were known to occur, are shown on Figure 1. The above findings led the EPA to develop an overall strategy to prioritize and expedite Site cleanup. EPA contracted the U.S. Army Corps of Engineers (USACE) to manage the cleanup activities at the site concurrently with the BVSPC investigation of the site.

Investigation Objectives

The ultimate goal of the investigation was to identify and locate those properties exceeding the established 800 ppm surface soil cleanup level for lead within the designated areas. These areas included the radius of the main “smelter zone," other small smelter areas, and specific mine and mill waste areas. The area to be investigated around the smelter zone was estimated by comparing the data from previous characterization activities at the Site. The lead smelter location and the historical smelter stack height were compared with the prevailing southeast wind direction to determine the impacted zone.

A secondary objective of the investigation was to prioritize the cleanup of residential yards. Considering census information obtained from each property access agreement and the XRF data, special attention was directed to the following two scenarios:

- Residences with toddlers and children under 7 years of age; and
- Residences with soil lead levels greater than 2,500 ppm.

The first scenario was selected based on the findings of the MDOH study. Fourteen percent of children under the age of 7 years in the study area had elevated blood-lead concentrations resulting primarily from residential yard soils (1). The second scenario was chosen because those residential yard soils exceeding 2,500 ppm were considered to pose an excessive risk to residents of all ages coming into contact with them.

Figure 1. Site and Designated Area Location Map

To prevent any further exposure to the elevated lead levels, EPA desired the cleanup activities to be conducted at the same time as the investigation. The flexibility of the EDMS and the desktop GIS allowed this to occur by providing the USACE with timely information to focus their cleanup activities on the areas of concern.

<u>The GIS Approach</u>

The use and ease of applying GIS were supported by the following activities:

- Availability of the background database and mapping resources;
- Using ArcView GIS™, a desktop GIS, to maintain the sample location information; and
- Field reconnaissance during the investigation to confirm and catalog sample locations not previously listed in the background information.

BVSPC, with assistance from EPA, requested existing database coordinate files from Jasper County officials. These files were acquired in two databases. The first database was the City of Joplin 911 database, and the second database was a developing rural Jasper County 911 database. The databases received by BVSPC were a modified subset of the original database, containing address and coordinate information only. This subset served to maintain the privacy of each individual resident.

In addition, electronic US Geological Survey Maps of Jasper County were obtained from the USACE during their time-critical cleanup activities at the site. Images of the areal extent of the main smelter zone and the designated mine waste areas were imported from AutoCAD into ArcView. These AutoCAD files were created during previous project activities at the site. The USGS maps and the AutoCAD images provided the base map for the sampling areas of concern.

With the use of ArcView, the combination of the 911 databases and the base map enabled BVSPC to identify the residential properties within the areas of concern. Those properties falling within the areas of concern shown on the base map were selected to separate them from the main databases that contained properties for the entire County. ArcView allowed the selected subset of properties to be exported to Microsoft Excel™, where the new database of geographical data was sorted by various fields for ease of reference. This subset initially contained 8,000

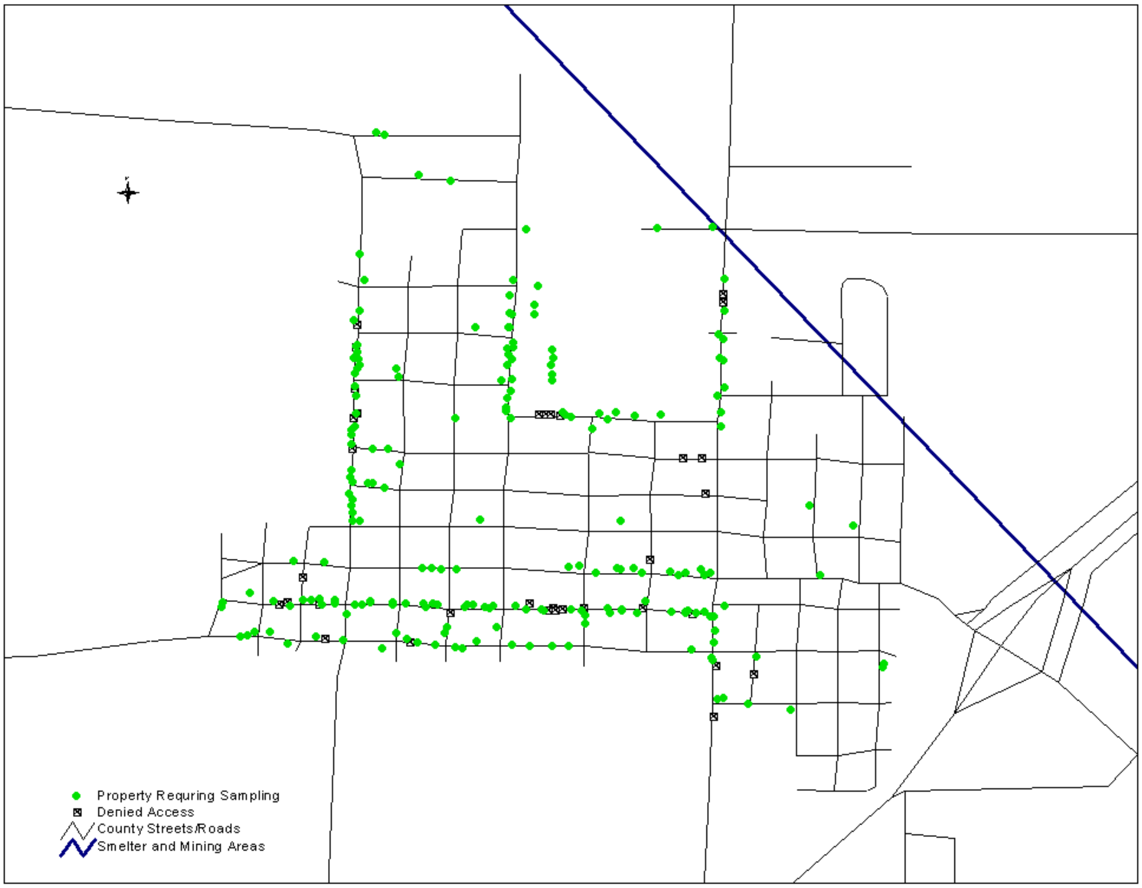


Figure 2. Residences "To Be Sampled" Map

properties requiring sampling within the Site for determination of surface soil lead levels. Figure 2 shows a portion of the properties to be sampled.

Field reconnaissance of these initial 8,000 properties was conducted as part of the access agreement process. Access agreements were required from the owner of each property to be sampled and included obtaining the owner's permission and taking a census of the residents living at the property. The EDMS significantly reduced the effort for field reconnaissance by incorporating the database information into a GIS presentable format, allowing BVSPC to easily identify the boundaries of the investigation. During reconnaissance, approximately 2,900 properties were recorded as being commercial properties, previously remediated properties, or properties where the owner denied access for sampling. These properties were omitted from the initial subset of 8,000. In addition to the remaining 5,100 properties in the database that required sampling, approximately 1,400 other residences were identified. These additional residences were mainly identified outside the city of Joplin, in smaller rural cities and towns. The county 911 database did not include the properties of homes within “city limits” since it was established as a rural property identification system. The rural 911 database was in the process of being developed and was incomplete in certain areas. The additional 1,400 identified residences would also require sampling and surveying to add them into the GIS database.

Portable GPS units, provided by EPA, were used to survey the 1,400 individual properties. EPA provided BVSPC with the coordinate location of the BVSPC Project Office in Joplin to ensure that the data points obtained at each property could be within $\pm$ 3 feet. Some interference was encountered during the GPS survey from various “line of sight” obstructions, including trees and electrical lines in the neighborhoods. GPS verification was obtained after several survey attempts for all additional locations requiring data. The GPS data was added to the existing geographical database for the properties that were sampled.

GIS Data Manipulation

The EDMS for this project, as demonstrated in Figure 3, was a comprehensive data management package that allowed BVSPC to manage large quantities of data. The data types included analytical, hydrogeological, and geographical information related to specific sampling locations. It served as the focal point of a “hands off” approach for input and output of environmental and GPS data. This “hands off” approach was an essential element during the project due to the extensive volume of data collected during the investigation. On average, three to five samples were collected per residential yard during the investigation. The data for each sample location, including XRF, laboratory confirmation, and GPS, were associated with the corresponding residential property. In essence, the EDMS eliminated the potential of transcription errors and data entry errors for 35,000 environmental data records and 6,500 coordinate and GPS records, while diminishing the duplication efforts of BVSPC personnel to manually review data and prioritize contaminated residential properties.

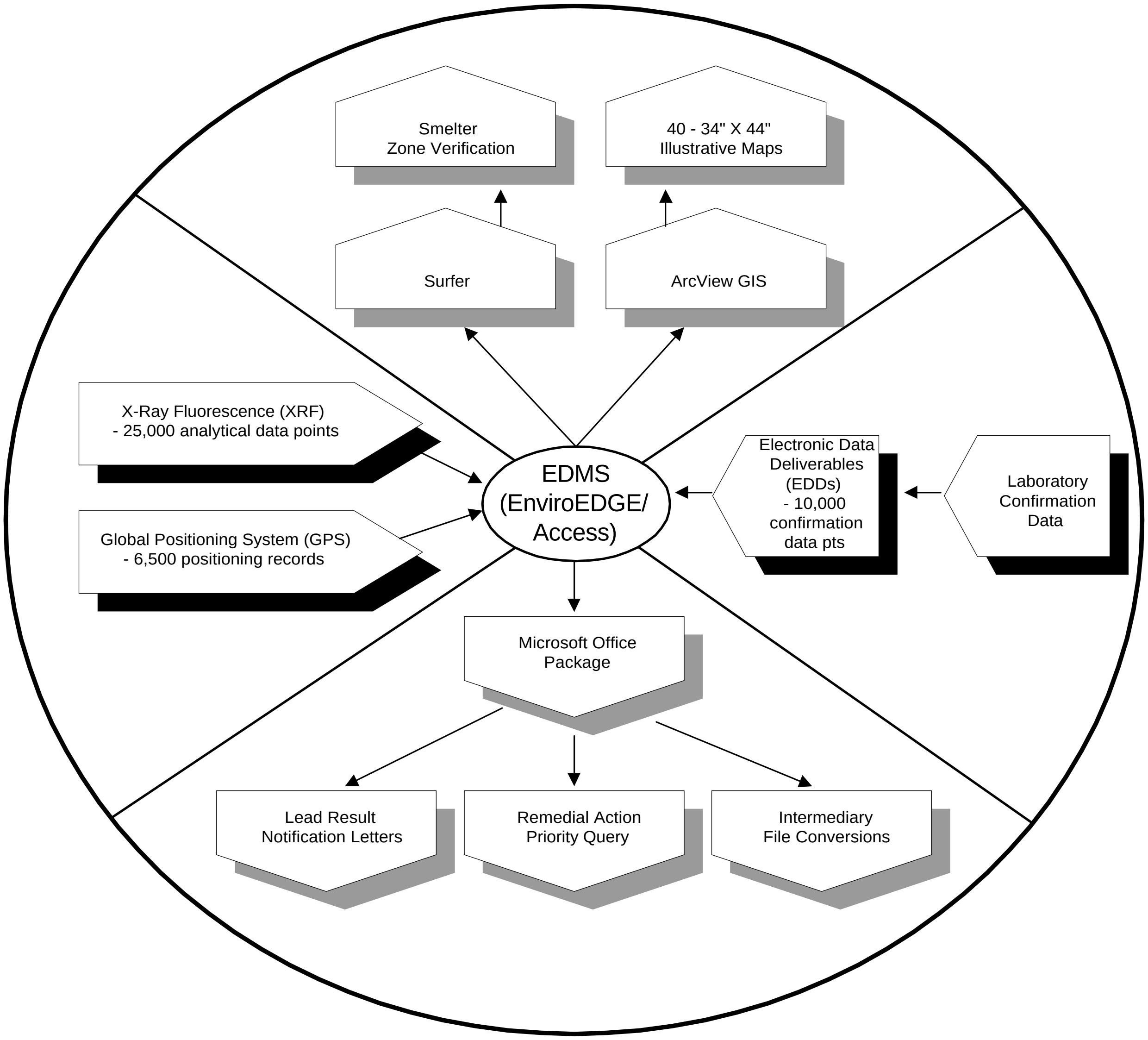


Figure 3. Electronic Data Management Flow Diagram

During this investigation, the EDMS allowed the analytical data to be sorted through a querying process. The following list includes examples of the types of queries run during the project:

- Lead concentrations above and below the action level;
- Lead concentrations indicating necessary remediation in ascending or descending order; and
- Areal sorts according to street name, city, or zip code, to further pinpoint the areas of greatest remediation need.

Each of these queries was performed using Access and exported to Excel as a spreadsheet file. Throughout the project, various specific preliminary queries were conducted at the request of the EPA. Representative queries included the following:

- Specific residence queries;
- Continuous updates of those residences having soil lead concentrations exceeding action levels;
- Updates of residences with children under age 7 and high lead levels; and
- Updates on homes with soil lead concentrations above 2,500 ppm.

XRF Data. Field environmental data was downloaded from XRF equipment into an electronic file that was transferred from the field office via e-mail to the main office. In the main office, the XRF file was formatted by an Excel macro to allow uploading into the EDMS. The data upload was simplified with the help of another macro written in Access that requested information from the database manager before searching for a specific data file to incorporate it into the EDMS. This brief process of transferring and uploading approximately 25,000 XRF data points allowed access to the data within three days of sampling. All data was immediately referenced to the corresponding geographical data to evaluate the existence of the “hot zones”.

Laboratory Data. Laboratory environmental confirmation data was uploaded directly into the EDMS using a pre-designed electronic data deliverable (EDD) package provided by the laboratory. Prior to receipt of the EDD, the laboratory data was validated and qualified by an outside contractor. During the upload process into the EDMS, an analytical verification process was accomplished that noted missing data. Approximately 10,000 confirmation data points were entered into the EDMS through this process.

GPS Data. GPS data was downloaded directly from the field GPS units into an electronic file that was transferred from the field office via e-mail to the main office. The GPS data underwent a post-processing procedure to correlate latitude and longitude with each residential property. BVSPC converted the latitude and longitude data to the State Plane-83 projection format used in ArcView. Once these procedures were completed, the 6,500 coordinate and GPS records were incorporated into the EDMS and the final presentation of data could begin.

GIS Presentation of Data

Presentation of the data was achieved through the availability of the created geographical database and several additional software packages, including Surfer, Microsoft Office™, and ArcView.

Surfer. Surfer, a 2-dimensional gridding and contouring package, was utilized to verify the main "smelter zone". After approximately 15,000 analytical samples were collected, data was queried from the EDMS and converted to a format able to be gridded and contoured in Surfer. With the additional sampling data obtained during the investigation, a contour map of soil lead concentrations was produced. The contour of the additional coverage area was exported from Surfer as an AutoCAD file that could be imported into ArcView for comparison with the original smelter zone. Upon review of the new contours, changes in the coverage of the original estimated smelter zone were identified (Figure 4). The same identification process, as described in previous paragraphs, was used to define the additional properties requiring sampling. The smelter zone was expanded in several areas and reduced in others, optimizing investigation efforts and maximizing the protection of public health.

Microsoft Office Package. The Microsoft Office Package gave BVSPC the flexibility of converting the various types of data into acceptable formats. These formats were used to prepare notification letters to property owners, to present requested queried information to EPA, and to query information for use in Surfer and ArcView for final geographical information.

Property owner notification letters were prepared with the help of Excel files and Word mail-merge capabilities. The letters informed the owners of the XRF lead sampling results for the high and low concentrations in their soils and gave a tentative timetable in which the remediation efforts would be conducted, if deemed necessary. These notices were sent to property owners within one month or less of the sampling date for their property. Without these electronic capabilities, each of the thousands of letters would have been completed individually. The effort for individual letters would have required much more personnel time and would not have been an expedient response to concerned residents.

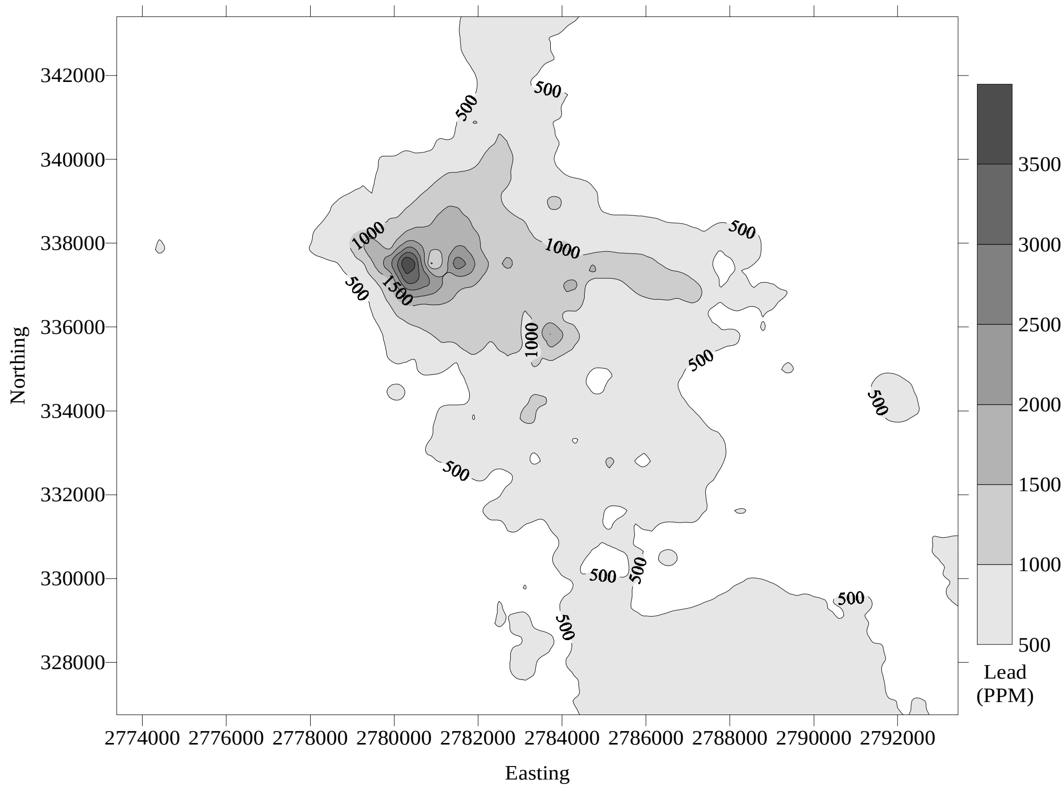


Figure 4. Smelter Zone Verification demonstrated in Surfer

The preliminary queries described earlier and the quick turnaround of the property owner notification letters allowed reassurances to concerned property owners and allowed the remediation prioritization for homes with children and those posing an excessive danger to residents.

ArcView GIS. ArcView was the final step in the presentation of all information to be used by EPA and the remediation contractor. It was used to create over 40 full-size maps of the remediation area (Figure 5) for the final design. ArcView easily illustrated the individual residences that exceeded the established lead soil cleanup level for this Site as well as those residences with lead concentrations below the action level. In addition, each residence identified on the map was linked to a table containing pertinent information, including property owner information, ages of any children living at the residence, and values of each XRF lead sample analysis (Figure 6). This procedure allowed a visualization of the extent and location of residential yards requiring soil remediation. Commercial locations, previously remediated properties, and those properties where sampling access was denied were also illustrated on the investigation summary maps. The maps allowed BVSPC personnel to confirm, track, and suggest the future progress of the soil investigation. These maps were a great asset to USACE, who used them to track the yards requiring remediation and those yards that had already been remediated. EPA also used the maps as illustrations at public meetings held in the area to keep residents informed of the progress.

Conclusions

The investigation was completed effectively and efficiently due to the flexibility and adaptability of the desktop GIS and our EDMS. As demonstrated above, the benefits of using ArcView in conjunction with the EDMS included the following:

⇒ Prioritized cleanup activities based on investigation results.
⇒ Maintained concurrent investigation and remediation efforts.
⇒ Provided a visual geographical tool for the investigation personnel.
⇒ Provided a visual geographical tool for the USACE cleanup crews.

Cost requirements were minimized due to the flexibility of ArcView GIS and the EDMS resources. The efforts for obtaining the geographical distribution of the contaminated properties would have required a much greater cost consideration during the implementation of the investigation and continuation of remedial efforts had a GIS system not been implemented during the project.

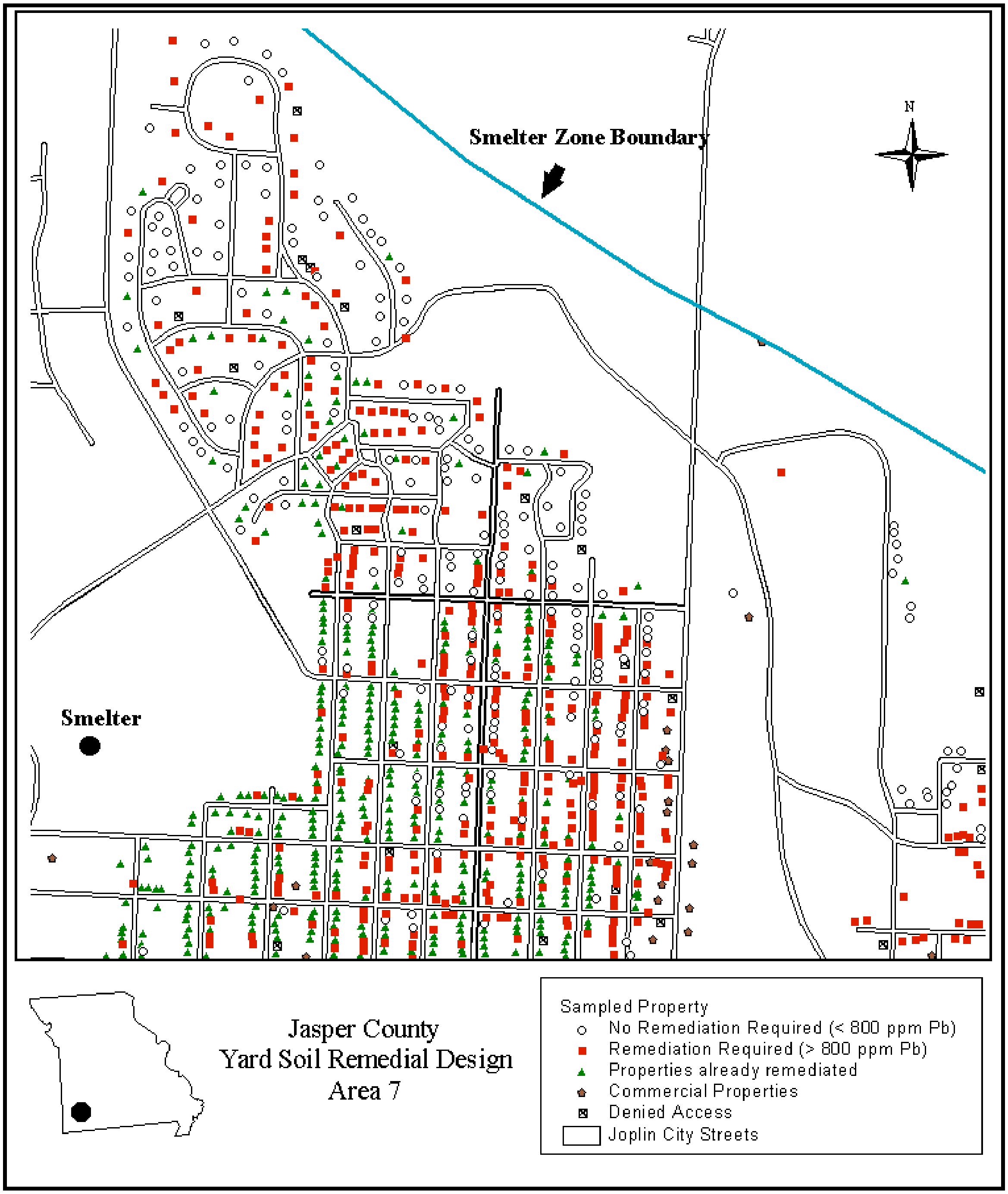


Figure 5. ArcView Presentation of Data

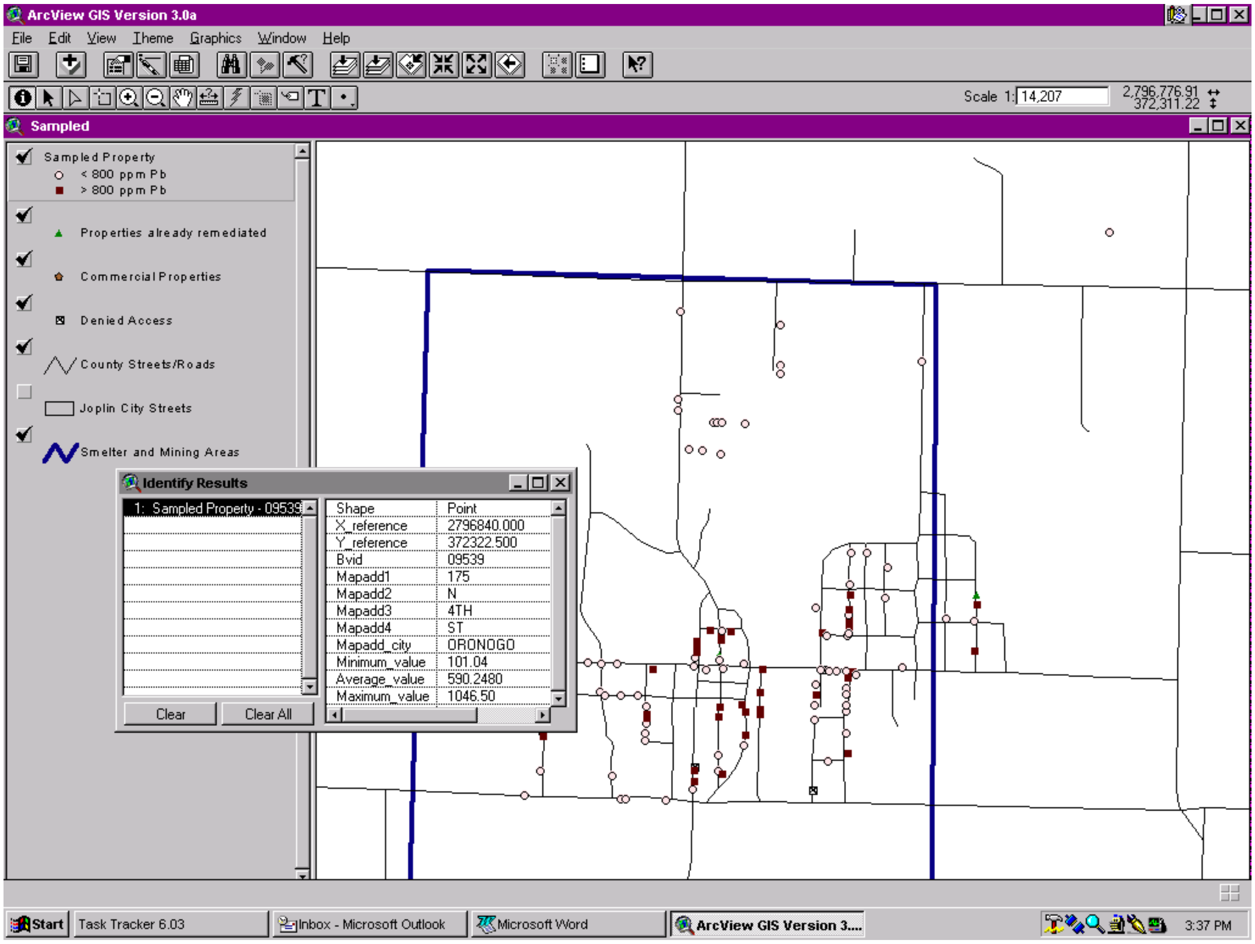


Figure 6. Linked Attributes Table to ArcView Property Location

BIBLIOGRAPHY

1. Missouri Department of Health (MDOH), 1994*, Jasper County, Missouri Superfund Site Lead and Cadmium Exposure Study,* May 1994.

2. Dames & Moore, 1994, *Residential Yard Assessment Report, Seven Designated Areas, Jasper County Site*, September 1994.

3. U.S. Environmental Protection Agency, 1996, *Record of Decision Declaration for the Oronogo-Duenweg Mining Belt Site, Operable Units 2 and 3, Jasper County, Missouri*, USEPA Region VII, August 1996.

4. Smith, Robert A., W. Todd Dudley, and Thomas L. Rutherford, 1995, *GIS Applications in Hazardous Waste Remediation*, Black & Veatch, 1995.

5. Black & Veatch Special Projects Corp., 1997, *Oronogo-Duenweg Mining Belt Site, Remedial Design, Statement of Work, Residential Yard Soils,* February 1997.

Decision Support System for Multiobjective Riparian/Wetland Corridor Planning

Margaret A. Fast and Tina K. Rajala
Kansas Water Office, Topeka, Kansas

Kansas has numerous programs that affect riparian corridors and associated wetlands. These programs include planning, monitoring, assistance, research, and regulatory activities. Although administration of these programs often overlaps, integration of program objectives into a holistic, multiobjective approach to resource planning and management has been lacking. A large amount of resource data was routinely collected and compiled, but no effective way had been developed to integrate these data into the decision-making process.

The Kansas Water Office (KWO) was awarded a grant in September 1992 from the U.S. Environmental Protection Agency (EPA) to develop a geographic information system (GIS) decision support system (DSS) that would enable the state to augment its ability to manage riparian/wetland corridors. The project used GIS to differentiate between reaches of a stream corridor to evaluate their environmental sensitivity. The Neosho River basin, one of 12 major hydrologic basins in Kansas, was used as a pilot to demonstrate the feasibility of the concept.

The KWO will use the DSS to help target sensitive areas in the Neosho basin for further planning activities. The project will also benefit other state agencies in their riparian/wetland corridor efforts. The implementation of planning objectives may involve local units of government and, ultimately, private landowners.

Major phases of the project included:

- A needs assessment study
- A feasibility analysis
- A system design
- Construction of the DSS for the Neosho River basin
- A final evaluation of the DSS capabilities

An interagency project advisory group (IPAG), consisting of representatives from eight agencies directly or indirectly involved in riparian and wetland protection activities, was formed to assist in project design and evaluation.

Major steps involved in designing the DSS included:

- Selection and GIS development of databases used for riparian corridor evaluation.
- Creation of riparian corridor segments.
- Development of an analysis methodology to apply to corridor segments.
- Evaluation of the DSS.

Databases Selected for Decision Support System Development

Many types of data were reviewed for the DSS. Several were not used due to the costs associated with geographically referencing the data, given the current data format.

The databases listed in Table 1 are available in the DSS.

During the system design phase of the project, the IPAG identified the need to develop a pilot study area for the DSS. The IPAG had difficulty understanding how a DSS would use geographically referenced data sets (coverages). Before committing to a design for the development of a basinwide system, the IPAG decided first to develop a pilot study area, with a specific focus (application), that could be on-line and demonstrated early. This would allow time for further refinement of the scope of work and identification of coverages to be developed prior to basinwide development of the DSS. For the pilot study application, the IPAG chose to assess the value and vulnerability of the riparian areas in two 11-digit hydrologic unit code (HUC11) watersheds to allow the user to evaluate a corridor segment and compare between segments and to prioritize or target segments for further planning activities.

As development of data layers progressed for the pilot, the IPAG quickly determined that the DSS project

Table 1. DSS Database List

DSS Name	Data Description	Source
Boundary	Neosho River basin boundary	Soil Conservation Service (SCS) HUC11 drainage basins; 1:100,000-scale[a]
Buffer	Riparian corridor	Original buffer on mainstem Neosho and Cottonwood; 147 corridor segments split on tributary confluences
Channels	Stream channelization	Division of Water Resources (DWR) legal description of locations
Con_ease	Conservation easements	Locations of important natural resources that could be purchased by the state from willing landowners for conservation protection
Contam	Water contamination	Kansas Department of Health and Environment (KDHE) contamination locations[a]
Corridor	Riparian corridor	Final riparian corridor; 63 corridor segments developed from HUC11 boundaries
County	County boundaries	Kansas Geological Survey (KGS) cartographic database; 1:24,000 scale[a]
Dams	Dam structures	DWR legal descriptions of locations
Dwrapp	Water appropriations	DWR legal descriptions of locations[a]
Gages	United States Geological Survey (USGS) stream gaging stations	USGS latitude-longitude descriptions; GIS cover developed by USGS
Geology	Surface geology	KGS 1:500,000-scale[a]
Huc11	11-digit hydrologic unit boundaries	SCS HUC11 drainage basins; 1:100,000-scale[a]
Hydr100k	Hydrology	USGS 1:100,000-scale digital hydrology[a]
Kats	Kansas water quality action targeting system	KDHE target valuable and vulnerable scores by HUC11 drainage basin
Landc	Land cover	1:100,000-scale developed from satellite imagery by the Kansas Applied Remote Sensing Program, University of Kansas[a]
Lc_stats	Land cover statistics	Summary statistics on land cover by corridor segment
MDS	Minimum desirable stream flow monitoring gages	Subset of USGS gaging stations
NPS	Nonpoint source pollution	Target watersheds identified in the Kansas Water Plan
Perenial	Perennial hydrology	Reselected perennial streams from 1:100,000 USGS digital hydrology
Pop	Population	Urban land cover (from landc) with 1980 and 1990 Census population data
PPL	Populated places	Geographic names information system (GNIS) entries for Kansas; GIS cover developed by USGS
Publand	Public lands	State and federally owned land digitized from 1:100,000-scale USGS quad maps
Roads	Roads	USGS 1:100,000-scale digital roads[a]
Sections	Section corners	KGS cartographic database; 1:24,000-scale[a]
Streamev	1981 stream evaluation	U.S. Fish and Wildlife stream evaluation study; Kansas Department of Wildlife and Parks (KDWP) provided data on paper maps
T_and_e	Threatened and endangered species	Stream locations of state and federal identified threatened and endangered species; KDWP provided data on paper maps
Temussel	Threatened and endangered species	Locations of state endangered floater mussels; KDWP provided data on paper maps
Tigrcity	City boundaries	U.S. Census 1:100,000-scale TIGER line data; boundaries only, areas not named (use with PPL)
Twp	Townships	KGS cartographic database; 1:24,000-scale[a]
Watrfowl	Water fowl locations	KDWP locations and counts of annual waterfowl migration; data developed from paper maps (Restrict public distribution of data per KDWP request.)
Wq_eff	Water quality: effluent	KDHE sampling sites; GIS cover developed by KDHE
Wq_grnd	Water quality: ground water	KDHE sampling sites; GIS cover developed by KDHE
Wq_lake	Water quality: lake	KDHE sampling sites; GIS cover developed by KDHE
Wq_strm	Water quality: stream	KDHE sampling sites; GIS cover developed by KDHE

[a] Data available at the Kansas Data Access and Support Center (DASC).

parameters would have to be limited to the riparian corridor along the mainstem of the Neosho and Cottonwood Rivers. The costs associated with developing riparian corridor segments for all perennial waters in the Neosho basin was far greater than the available funding.

Creation of Riparian Corridor Segments

A buffer width of one-half mile (one-quarter mile from each stream bank) for the mainstem of the Neosho and Cottonwood Rivers was used to produce the riparian corridor. If more time and funding had been available, riparian corridors for all perennial streams in the Neosho basin could have been developed. The development of this second view of data, organized by the HUC11 watershed, would then have been useful for individual watershed analysis because all perennial streams in the watershed could be analyzed.

The intersection of the HUC11 basin boundaries segmented the corridor. In several instances, small sliver polygons were produced where the HUC11 boundary paralleled the river within the 1/4-mile corridor. The sliver polygons were dissolved into the majority HUC11. In other words, this project assumed that the 1/4-mile corridor buffer was more accurate and useful than the 1:100,000-scale HUC11 boundary.

Many of the HUC11 boundaries that the Soil Conservation Service (SCS) developed actually follow the course of the Kansas streams, rather than intersect them. When this occurred along the Neosho and Cottonwood Rivers, we found that the resulting opposing corridor segments did not always balance with an equivalent length. Also, some HUC11 boundaries would first follow the river, then cross the river. This resulted in corridor segments that encompass both sides of the river for a portion of the segment and follow only one side of the river for another portion of the segment. To address these situations, the KWO arbitrarily added intersections to create equivalent left and right bank corridor segments and to create corridor segments that encompassed either one side of the river or both sides of the river.

Once the corridor segments were finalized and numbered, the corridor segment identification number (corrseg-id) was attached to the other GIS covers. This allows the reselection of data for a given corridor segment, using Boolean expressions in the DSS.

Development of an Analysis Methodology: Land Use

The IPAG determined that one of the most significant factors associated with the quality of the riparian corridor is land cover. Land cover was analyzed for the riparian corridor segments; the GIS cover lc_stats contains summary statistics for each corridor segment. The calculations discussed in the following paragraphs identify the data found in the lc_stats cover. Due to the size of the land cover data set in the Neosho River basin, the DSS includes only the land cover within the riparian corridor.

One way of identifying corridor segments in need of protection or remedial action is to determine the ratio of the number of acres in the corridor segment that contain the preferred riparian land cover types (grasses, woods, and water) to the number of acres that contain the least preferred types of land cover (crops and urban areas). The corridor segments can then be ranked according to that ratio.

Other calculations are useful:

- *bad_pct:* percentage of the corridor segment that contains crop and urban land cover types.
- *bad_tbad:* percentage of all crop and urban land cover for the entire riparian corridor that resides in the corridor segment.
- *'type'_pct:* percentage of the corridor segment that is crop, grass, wood, water, and urban. 'Type' refers to each of the five land cover types; lc_statsuses: a separate value for each (e.g., crop_pct).
- *'type'_t'type':* percentage of each type of land cover for the entire riparian corridor that resides in the corridor segment (e.g., crop_tcrop).
- *'type'_acres:* total acreage of each type of land cover in the corridor segment (e.g., crop_acres).
- *good_acres:* total acreage of grass, wood, and water in the corridor segment.
- *bad_acres:* total acreage of crop and urban in the corridor segment.

Another significant benefit of the DSS is the ability to see where the land cover types are in relation to the river. As an example, the ability to identify corridor segments that have crop land extending to the river on both banks is useful because they are the segments most vulnerable to bank erosion. Those segments can then be targeted for further remedial activities planning.

Decision Support System Requirements

The DSS data sets were developed and analyzed using ARC/INFO on a UNIX-based workstation. The final covers were then exported and transferred to a microcomputer for use in ARC/VIEW. Hardcopy prints are printed to a Tektronix Phaser III color wax printer with 18 Mb of RAM, running in Postscript mode.

The DSS data sets total 26 Mb. ARC/VIEW version 1 requires 8 Mb of RAM to load the program. To run the DSS efficiently, a 486DX-66 with 16 Mb of RAM is preferred. The DSS is slower on a 486DX-33 with 8 Mb of RAM. It was not tested on any other PC configuration, so a configuration in between the two may be satisfactory.

Processing GIS Data

Reselecting the perennial streams in the Neosho basin and further identifying the mainstem of the Neosho and Cottonwood Rivers using United States Geological Survey (USGS) 1:100,000-scale hydrography can be time consuming. Perhaps the River Reach III covers should replace that data in the future.

Attaching census data to the urban land cover polygons, as was done for the pop cover, is not recommended. Use of the TIGER line files and cover would give a more accurate distribution of the population. Because less than 5 percent of the riparian corridor had urban land cover, the KWO did not use the pop cover in its evaluation. Several summary covers of TIGER and census data will soon be available from DASC.

Clipping the other ARC/INFO covers to the Neosho basin and attaching the corrseg-id, using the identity command, was unremarkable.

Processing Non-GIS Digital Data

Channels and dams were in digital format but were not in ARC/INFO format. The files were processed using the LeoBase conversion software from the KGS, then generated into ARC/INFO covers. Some records were lost in the conversion. The LeoBase program fails to convert, or incorrectly converts, legal descriptions for sections that do not have four section corners (e.g., northeastern Kansas). The Division of Water Resources is in the process of attaching latitude-longitude to the point locations. Processing these data should take only a few hours at most.

Processing Nondigital Data

Several covers were developed on contract from paper maps or legal descriptions. They were: conservation easements (con_ease), public lands (publand), stream evaluation (streamev), threatened and endangered species (t_and_e and temussel), and water fowl (watrfowl). Most of the data for these covers were drafted on a 1:100,000-scale USGS quad map and digitized. The stream evaluation data were developed using a scanned paper map of the coded streams as a backdrop for the 1:100,000-scale hydrography; the digital streams were reselected and coded.

In summary, KWO's GIS personnel needed approximately 275 hours to develop the riparian corridor segments, process the land cover data and summary statistics, export the covers, transfer and import the covers for ARC/VIEW, and assist in the development and presentation of the DSS demo. Contract personnel spent approximately 183 hours developing GIS covers for the DSS. This does not include the time spent identifying the perennial and mainstem hydrology in the USGS 1:100,000-scale hydrology.

Final Evaluation of the Decision Support System

In its final evaluation of the system, the IPAG determined the system to be useful and an excellent start at consolidating a variety of data that have application for riparian corridor/wetland issues. Many IPAG members found ways to use the DSS in their own programs. Additional comments on the system evaluation are as follows:

- Concern about the lack of complete wetland data. The land cover data available could not identify wetland areas.
- Need for more detailed woodland data. Again, the resolution of the land cover data precluded detailed identification of woodland areas. The Kansas Biological Survey (KBS), the KWO, and EPA are now pursuing options to develop more detailed land cover data, including wetlands and woodlands.
- The lack of information on the tributaries did not allow full basin analysis, which would be desirable. This issue is addressed in the "construction" discussion above.
- Desirability of expanding the project with elevation and temporal data.
- Lack of definition of the floodplain. Federal Emergency Management Agency (FEMA) floodplain data are not easily incorporated into a GIS. Other options, including satellite imagery of the flood of 1993, will be evaluated.
- Project development requires extensive communication between program people and GIS technicians. This can be a daunting task due to the technical vocabularies involved and the many other ongoing activities of the participants.
- Consideration of the requirements for transferring the project to other potential users. GIS applications generally use large databases. User microcomputers may not have the CPU, RAM, and storage capacity necessary for the DSS application and often have a limited number of options for data transfer.
- Concern about costs and time associated with the expansion of the DSS to other basins in the state. This project was focused on one of the 12 major hydrologic regions in Kansas. Funding options, project scope, and system refinements based on the physical characteristics of the other basins need to be pursued.

The KWO learned that clearly defining a single DSS application at the outset of the project is critical. The KWO originally believed that the DSS could be developed with general descriptions of the broad range of program applications, utilized by multiple agencies, that could benefit from the DSS. Each participating agency could bring its programs and needs to the IPAG for discussion; the resulting DSS would then serve those multiple programs and needs. Instead, the ambiguity of the objective confused the IPAG. Once the IPAG chose to focus on a single application, the assessment of riparian corridor value and vulnerability to target priority areas for further planning activities, the IPAG became more confident in its advisory role. Upon completion of the project, the IPAG members could readily identify how the DSS could be enhanced, modified, or directly used in their own programs.

Integration of GIS and Hydrologic Models for Nutrient Management Planning

Clyde W. Fraisse, Kenneth L. Campbell, James W. Jones,
William G. Boggess, and Babak Negahban
Agricultural Engineering Department, University of Florida, Gainesville, Florida

Introduction

Recent evidence that agriculture in general, and animal waste in particular, may be an important factor in surface and ground-water quality degradation has increased the interest in nutrient management research. The presence of nitrogen and phosphorus in surface water bodies and ground water is a significant water quality problem in many parts of the world. Some forms of nitrogen and phosphorus, such as nitrate N and soluble P, are readily available to plants. If these forms are released into surface waters, eutrophic conditions that severely impair water quality may result. Advanced eutrophication (pH variations, oxygen fluctuations or lack of it in lower zones, organic substance accumulation) can cause physical and chemical changes that may interfere with recreational use and aesthetic appreciation of water. In addition, possible taste and odor problems caused by algae can make water less suitable or desirable for water supply and human consumption (1).

Increases in nutrient loadings to water resources have recently been observed in the southeastern United States, where well-drained sandy soils with low nutrient retention capacity and high water table conditions are found in most coastal areas. Those increases were associated statistically with nutrient sources such as agricultural fertilizers and dense animal populations (2, 3). Repetitive occurrences of extensive blooms of blue-green algae that threatened the overall health of Lake Okeechobee, located in southern Florida, were attributed to an increase in nutrient loadings to the lake. The South Florida Water Management District (SFWMD) reported an increase of phosphorus concentrations in the lake water from an annual average of 0.049 milligrams per liter in 1973 and 1974 to a peak of 0.122 milligrams per liter in 1988 (4).

Most water quality problems concerning phosphorus result from transport with sediment in surface runoff into receiving waters. Continuous high loadings from animal waste on sandy soils with low retention capacity, however, may contribute significant quantities of labile phosphorus to subsurface drainage. Ground-water aquifers may also become polluted due to recharge of high loadings of nitrogen. Drinking water with nitrate N concentrations higher than 10 milligrams per liter may lead to methemoglobinemia in infants. Ground-water monitoring of the Middle Suwannee River area in Florida has shown high concentrations of nitrate nitrogen near intensive agricultural operations. The U.S. Geological Survey has intensively monitored dairy and poultry farms and has found high nitrate levels below these operations compared with nearby control wells (5).

Animal waste management has always been a part of farming, but historically has been relatively easy due to the buffering capacity of the land. In fact, land application of animal waste at acceptable rates can provide crops with an adequate level of nutrients, help reduce soil erosion, and improve water holding capacity. As the animal industry attempts to meet the food requirements of a growing population, however, it applies new technologies that reduce the number of producers, but create larger, more concentrated operations. That, in addition to the decreasing amount of land available for waste application, has increased the potential for water quality degradation.

Successful planning of an animal waste management system requires the ability to simulate the impact of waste production, storage, treatment, and use on water resources. It must address the overall nutrient management for the operation, including other nutrient sources such as supplemental fertilizer applications. Livestock operations are highly variable in their physical facilities, management systems, and the soil, drainage, and climatic conditions that affect the risk of water pollution from animal wastes (6). Linkage between geographic information systems (GIS) and hydrologic models offers an excellent way to represent spatial features of the-fields being simulated and to improve results. In addition, a GIS containing a relational database is an excellent way to store, retrieve, and format the spatial and tabular data required to run a simulation model.

This paper examines some of the issues related to the integration of hydrologic/water quality models and GIS programs. In addition, the paper discusses the approaches used in the Lake Okeechobee Agricultural Decision Support System (LOADSS), which was recently developed to evaluate the effectiveness of different phosphorus control practices (PCPs) in the Lake Okeechobee basin. The paper also details a dairy model, designed to simulate and evaluate the impacts of alternative waste management policies for dairy operations, that is currently under development.

Hydrologic Models and GIS

By using models, we can better understand or explain natural phenomena and, under some conditions, make predictions in a deterministic or probabilistic sense (7).

A hydrologic model is a mathematical representation of the transport of water and its constituents on some part of the land surface or subsurface environment. Hydrologic models can be used as planning tools for determining management practices that minimize nutrient loadings from an agricultural activity to water resources. The results obtained depend on an accurate representation of the environment through which water flows and of the spatial distribution of rainfall characteristics. These models have successfully dealt with time, but they are often spatially aggregated or lumped-parameter models.

Recently, hydrologists have turned their attention to GIS for assistance in studying the movement of water and its constituents in the hydrologic cycle. GIS programs are computer-based tools to capture, manipulate, process, and display spatial or georeferenced data. They contain geometry data (coordinates and topological information) and attribute data (i.e., information describing the properties of geometrical objects such as points, lines, and areas) (8). A GIS can represent the spatial variation of a given field property by using a cell grid structure in which the area is partitioned into regular grid cells (raster GIS) or by using a set of points, lines, and polygons (vector GIS).

A close connection obviously exists between GIS and hydrologic models, and integrating them produces tremendous benefits. Parameter determination is currently one of the most active hydrology-related areas in GIS. Parameters such as land surface slope, channel length, land use, and soil properties of a watershed are being extracted from both raster and vector GIS programs, with a focus on raster-based systems. The spatial nature of GIS also provides an ideal structure for modeling. A GIS can be a substantial time saver that allows different modeling approaches to be tried, sparing manual encoding of parameters. Further, it can provide a tool for examining the spatial information from various user-defined perspectives (9). It enables the user to selectively analyze the data pertinent to the situation and try alternative approaches to analysis. GIS has been particularly successful in addressing environmental problems.

Approaches for Integrating GIS and Models

A significant amount of work has been done to integrate raster and vector GIS with hydrologic/water quality models. Several strategies and approaches for the integration have been tried. Initial work tended to use simpler models such as DRASTIC (10) and the Agricultural Pollution Potential Index (11). In these cases, the models were implemented within the GIS themselves. These studies attempted to develop GIS-based screening methods to rank nonpoint pollution potential. The use of more complex models requires that the GIS be used to retrieve, and possibly format, the model data. The model itself is implemented separately and communicates with GIS via data files. Goodchild (12) refers to this mode as "loose coupling," implying that the GIS and modeling software are coupled sufficiently to allow the transfer of data and perhaps also of results, in the reverse direction. Fedra (8) refers to this level of integration as "shallow coupling" (see Figure 1). Only the file formats and the corresponding input and output routines, usually of the model, must be adapted. Liao and Tim (13) describe an application of this type, in which an interface was developed to generate topographic data automatically and simplify the data input process for the Agricultural Nonpoint Source (AGNPS) Pollution Model (14), a water quality model.

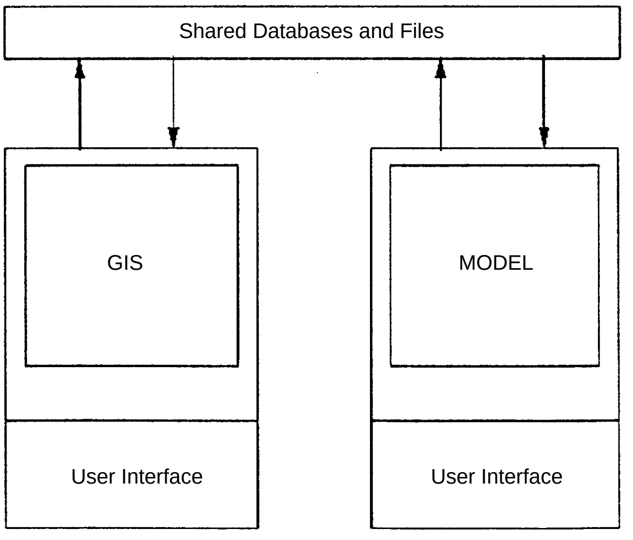


Figure 1. Loose or shallow coupling through common files (8).

Higher forms of connection use a common interface and transparent file or information sharing and transfer between the respective components (see Figure 2). The dairy model, currently under development, is an application of this kind. It will link the Ground-Water Loading Effects of Agricultural Management Systems (GLEAMS) (15) model and GIS to evaluate potential leaching and runoff of both nitrogen and phosphorus.

LOADSS is an extension of this type of application because it includes an optimization module that enables the system to select the best PCPs at the regional scale, based on goals and constraints defined by the user.

Both applications use ARC/INFO's arc macro language (AML), a high-level application language built into the GIS. A subset of functions of a full-featured GIS, such as creation of maps (including model output) and tabular reports, as well as model-related analysis, are embedded in the applications, giving the system great flexibility and performance. Fedra (8) describes a deeper level of integration that would merge the two previous approaches, such that the model becomes one of the analytical functions of a GIS, or the GIS becomes yet another option to generate and manipulate parameters, input and state variables, and model output, and to provide additional display options. In this case, software components would share memory rather than files.

The choice between integrating a water quality model with a raster or vector GIS depends on the importance of spatial interactions in the process being studied and the nature of the model itself. Some water quality models, such as GLEAMS, are field-scale models that provide edge-of-the-field values for surface runoff and erosion as well as deep percolation of water and its constituents. In this case, spatial interactions between adjacent fields are ignored and a vector GIS can be used to describe the system. Moreover, important factors in the simulation process, such as land use and management practices, are normally field attributes and thus, are better represented in a vector structure.

Other factors playing an important role in the hydrologic process, such as field slope, aspect, and specific catchment area, are hard to estimate in vector systems, however. A raster-based GIS is better suited for handling watershed models in which the routing process is important and spatial interactions are considered. For those, several algorithms for estimating important terrain attributes are often incorporated in commercially available raster-based GIS programs.

LOADSS

LOADSS was developed to help address problems created by phosphorus runoff into Lake Okeechobee. It was designed to allow regional planners to alter land uses and management practices in the Lake Okeechobee basin, then view the environmental and economic effects resulting from the changes. The Lake Okeechobee basin coverage incorporates information about land uses, soil associations, weather regions, management practices, hydrologic features, and political boundaries for approximately 1.5 million acres of land and consists of close to 8,000 polygons.

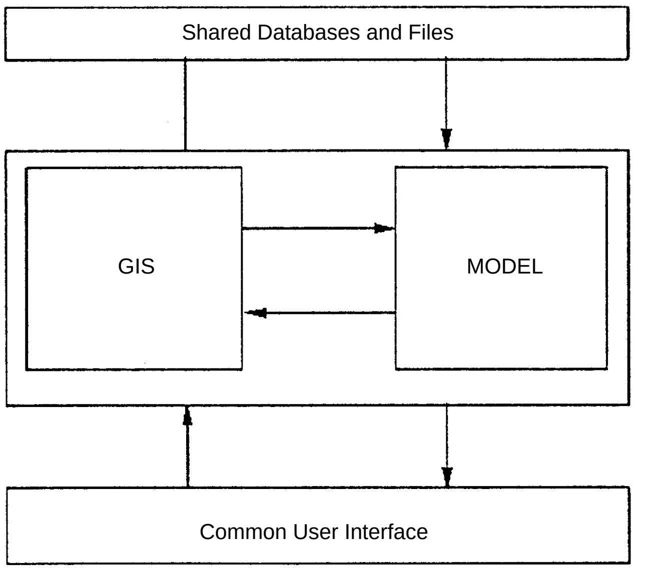


Figure 2. Deep coupling in a common framework (18).

The SFWMD, responsible for managing Lake Okeechobee, has initiated numerous projects to develop effective control practices for reducing the level of phosphorus in agricultural runoff as part of the Lake Okeechobee Surface Water Improvement and Management (SWIM) Plan. These projects, numbering more than 30, were designed to develop information on the control and management of phosphorus within the lake basin and to determine the costs and effectiveness of selected management options. Three types of control options are being studied:

- Nonpoint source controls, such as pasture management.
- Point source controls, such as sewage treatment.
- Basin-scale controls, such as aquifer storage and retrieval.

After completing most of these research efforts, the need arose for a comprehensive management tool that could integrate the results for all three classes of PCPs. In response to these needs, design and implementation of a decision support system was initiated with the following objectives (16):

- Organize spatial and nonspatial knowledge about soils, weather, land use, hydrography of the lake basin, and PCPs under a GIS environment.

- Develop and implement algorithms for modeling non-point source, point source, and basin-scale PCPs.
- Develop and implement mechanisms for evaluating the performance of the entire Lake Okeechobee basin under different combinations of PCPs applied to the basin.
- Design and develop a user interface that would facilitate use of the system by noncomputer experts.

The goal in developing LOADSS was to create an information system that would integrate available information to help regional planners make decisions. LOADSS can generate reports and maps concerning regional land attributes, call external hydrologic simulation models, and display actual water quality and quantity sampling station data. LOADSS is a collection of different components:

- The regional scale GIS-based model used to develop and manipulate regional plans for reducing phosphorus loading to Lake Okeechobee.
- The Interactive Dairy Model (IDM) used to develop field-level management plans for dairies and run the Field Hydrologic and Nutrient Transport Model (FHANTM) simulation model for nutrient transport modeling.
- An optimization module that enables the system to select the best PCPs at the regional scale (currently under development).

Although these components can run independently, they are fully integrated in the LOADSS package and can exchange information where necessary. A design schematic of LOADSS is given in Figure 3.

Regional-Scale GIS-Based Model

LOADSS serves both as a decision support system for regional planning and as a graphic user interface for controlling the different components. One consideration in the design of LOADSS was the size of the database that was being manipulated. Because the land use database consisted of nearly 8,000 polygons, running the simulation models interactively would not be a feasible option. Thus, the CREAMS-WT (17) runoff model was prerun for different levels of inputs and management for each land use, soil association, and weather region (18).

Depending on the land use and its relative importance as a contributor of phosphorus to the lake, anywhere from one (background levels of inputs to land uses like barren land) to 25 (dairies, beef pastures) levels of inputs were selected. Each set of inputs to a particular land use was given a separate PCP identification code. A CREAMS-WT simulation was performed for each PCP, on each soil association and weather region. This resulted in approximately 2,600 simulation runs. Annual average results were computed for use in LOADSS. CREAMS-WT provides an average annual estimate of phosphorus runoff from each polygon. Phosphorus assimilation along flow paths to Lake Okeechobee are estimated as an exponential decay function of distance traveled through canals and wetlands (4).

The imports, exports, and economics of each PCP are based on a per production unit basis. Depending on the type of polygon, the production unit can be acres (e.g., pastures, forests), number of cows (dairies), or millions of gallons of effluent (waste treatment plants and sugar mills). Developing a regional plan in LOADSS involves assigning a PCP identification code to each one of the polygons in the Lake Okeechobee basin. Accessing the results of a regional plan involves multiplying the production unit of each polygon by its appropriate database import, export, or economic attribute and summing the resulting values over all polygons in the Lake Okeechobee basin. LOADSS runs in the ARC/INFO Version 6.1.1 GIS software on SUN SPARC stations.

Interactive Dairy Model

Although the LOADSS level of detail is adequate for regional planning, a more detailed model was necessary to analyze individual dairies in the Lake Okeechobee basin, as dairies were one of the large, concentrated sources of phosphorus runoff into the lake. Thus, the IDM was developed and incorporated into LOADSS. IDM utilizes FHANTM to simulate phosphorus movement in dairy fields. FHANTM is a modification of DRAINMOD (19) with added functions to handle overland flow routing, dynamic seepage boundary, and soluble phosphorus algorithms for P input, mass balance, and transport (20).

Unlike in LOADSS, FHANTM is run interactively, as IDM requires. Furthermore, in LOADSS, the user can only select from a number of predefined PCPs, while in IDM, the user has access to more than 100 input and management variables, all of which can take a range of values. This allows for the development and evaluation of detailed dairy management plans that otherwise would be impossible at a regional scale. While LOADSS only provides average annual results, IDM displays daily time series simulation results. IDM uses the same assimilation algorithm and can produce the same phosphorus budget maps and reports as LOADSS.

Optimization Module

A variety of factors must be considered in planning nutrient management programs. Production and environmental goals need to be balanced, and these goals are often incompatible. Performing this exercise on a regional scale, comprising many fields for which a variety of land uses and management options can be assigned, is a tremendously time-consuming, if not impossible,

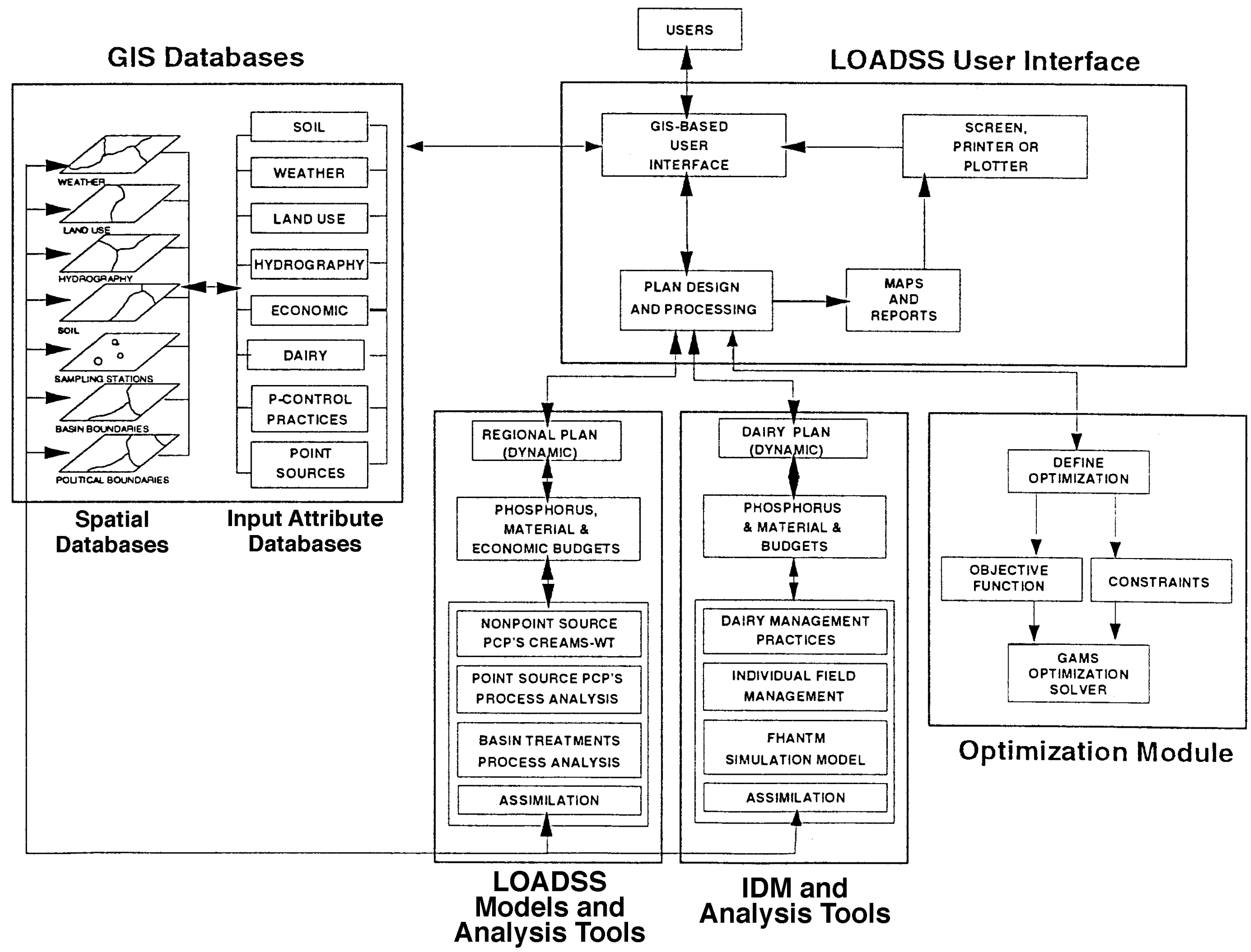


Figure 3. LOADSS design schematic (16).

task. The optimization component of LOADSS, currently under development, will determine the best combination of agricultural, environmental, and regulatory practices that protects and maintains the health of Lake Okeechobee and also maintains the economic viability of the region. The optimization process will provide another method for modifying the PCPs assigned to individual fields. Different optimization solution methods, such as linear programming and integer linear programming, will be available for solving the optimization problem that the user defines.

Dairy Simulation Model

The dairy model was expected to be fully functional by the end of 1994. It is designed to be an additional tool for answering questions related to the environmental costs and impacts of dairy operations. A design schematic of the dairy model is given in Figure 4. It differs from the LOADSS/IDM model in the following aspects:

- It is designed to be generic so that any dairy represented by a coverage for which relevant data, such as topography, soil characteristics, weather, and field boundaries, are available can be simulated.
- The GLEAMS water quality model will be used for simulating nutrient transport of nitrogen and phosphorus.
- The user will be able to assign a larger variety of crops and crop management practices to the individual fields, including crop rotation.

GLEAMS (15) is a field-scale water quality model that includes hydrology, erosion/sediment yield, pesticide, and nutrient transport submodels. GLEAMS was developed to use the management oriented CREAMS (21) model and incorporate more descriptive pesticide subroutines and more extensive treatment of the flow of water and chemicals in the root zone layers. The water is routed through computational soil layers to simulate the percolation through the root zone, but the volume of percolation in each layer is saved for later routing in the pesticide component. A minimum of three and a maximum of 12 layers with variable thickness may be used. Soil parameter values are provided by soil horizon, and

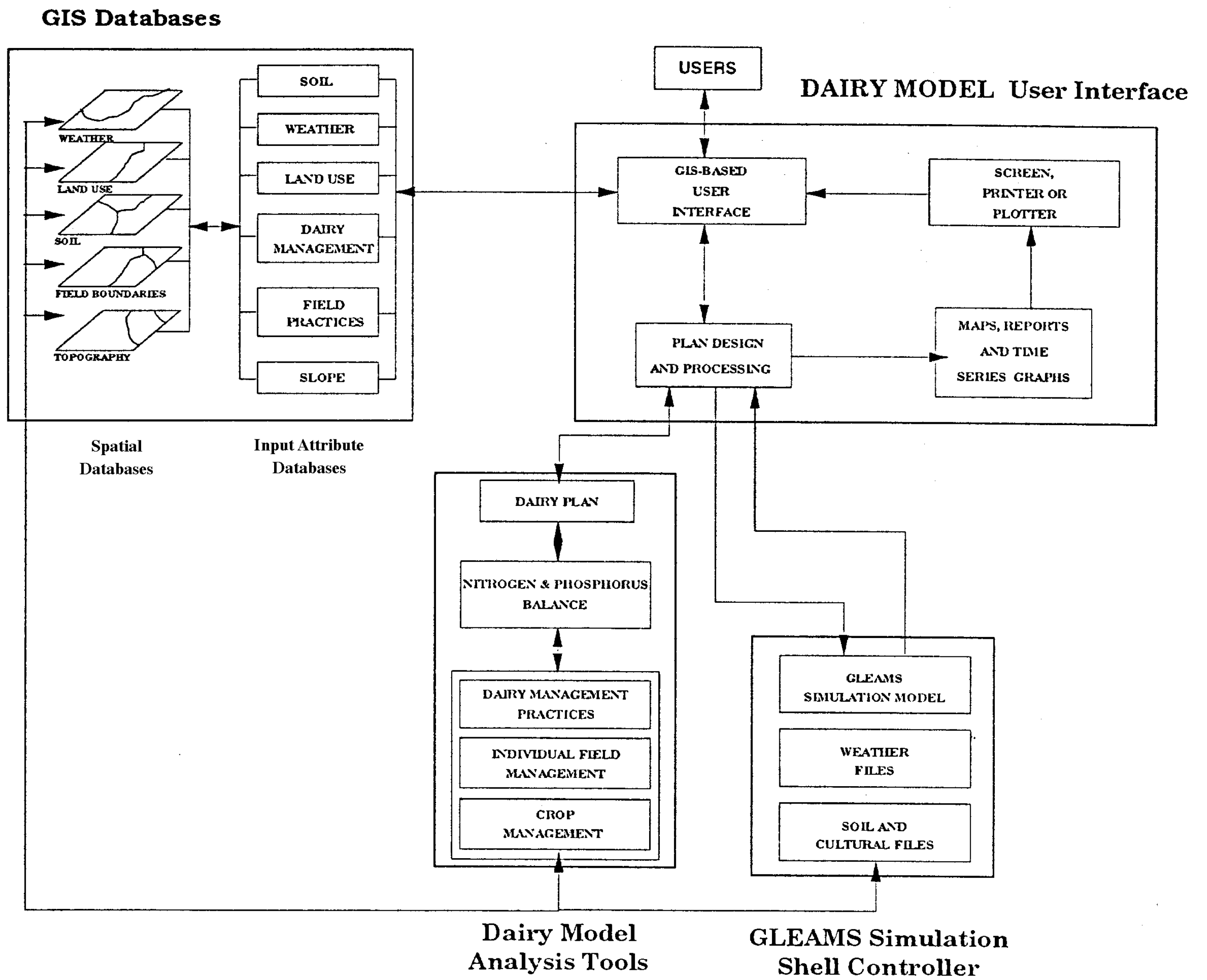


Figure 4. Dairy model design schematic.

the crop root zone may have up to five horizons. The values for parameters, such as porosity, water retention properties, and organic content, are automatically fitted into the proper computational layers. Two options are provided in the model to estimate potential evapotranspiration, the Priestly-Taylor method (22) and the Penman-Monteith method (23). The nutrient component of the model simulates land application of animal wastes by creating appropriate nitrogen and phosphorus pools for mineralization. It considers ammonia volatilization from surface-applied animal waste by using a relationship developed by Reddy et al. (24).

The graphic interface is designed to help the user plan a balanced nutrient management program for the dairy being simulated. First, total nutrient production and accounting are estimated, based on information related to the dairy management such as herd size, confinement system, waste characterization, and handling. Figure 5 shows the general structure of the graphic interface and a first version of the menu used to estimate the total amount of nitrogen and phosphorus available for assignment to the various fields. Nutrient losses during waste storage and treatment vary widely depending on the method of collection, storage, and treatment. Climate can also have a great effect on the losses. Covering all possible methods of storage and treatment is practically impossible, especially in an application that is designed to be generic and applied in any part of the country. A simplification was made to overcome this problem: the user must provide the percentage of original nitrogen and phosphorus that is retained after waste storage and treatment. The menus designed to enter information related to the management of fields and crops are given in Figure 6.

For each field, a sequence of crops can be defined in the Field Management Table, and for each crop, the sequence of practices or field operations is defined in the Crop Management Table. Every time a waste application operation is defined or a field is used as pasture for a certain period, the corresponding amount of nutri-

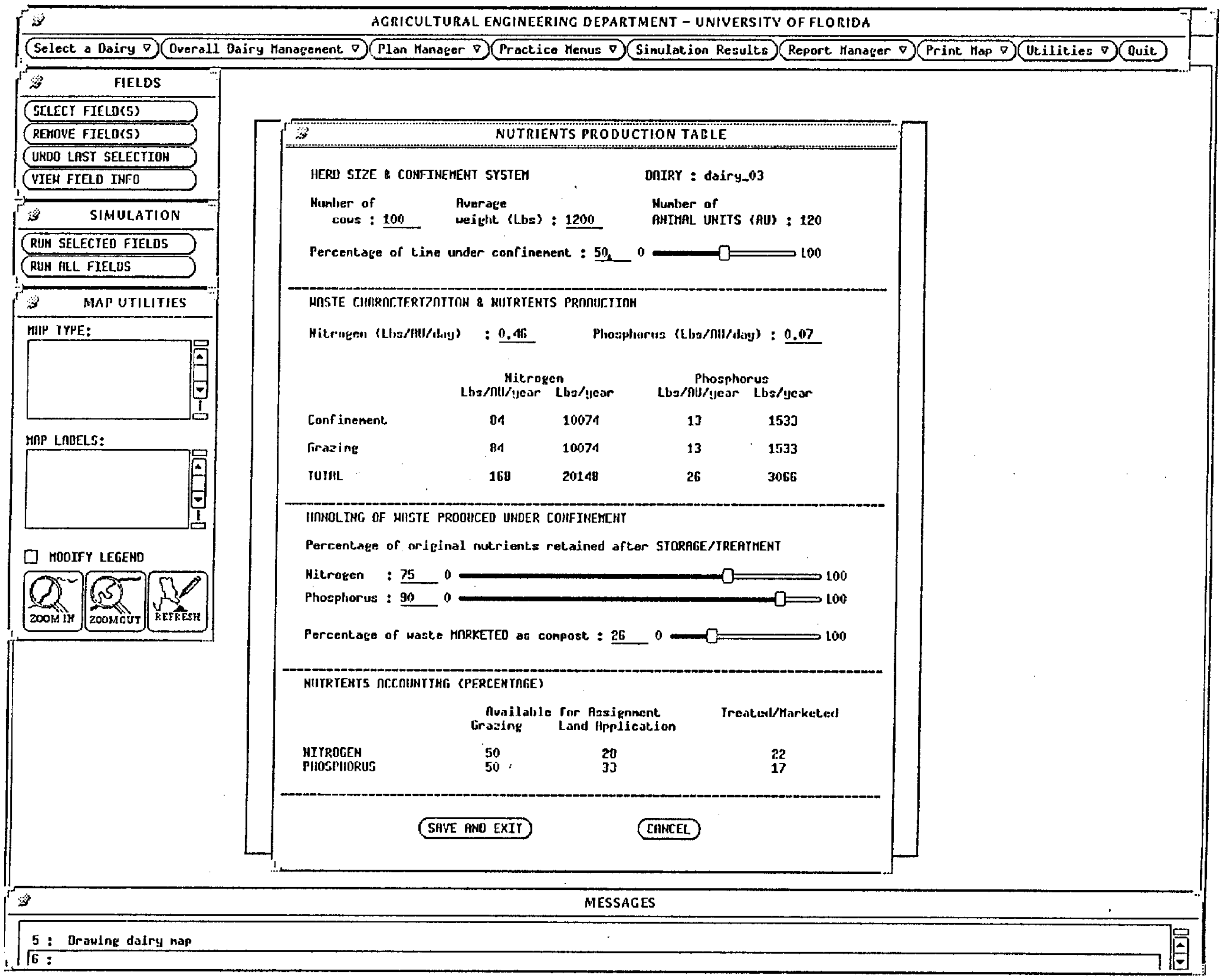


Figure 5. The nutrient production table of the dairy model interface.

ents will decrease from the amount available for assignment and the total available for future assignment will be updated. Once the total amount of nutrients is assigned, the model can be run for the several fields in the dairy and the results evaluated in terms of nutrient loadings to the edge of fields and ground water. Alternative plans can be designed and saved for comparison and selection of best management options. The best solutions in terms of reducing nutrient loadings to surface and ground water must also consider economic aspects. A producer's decision about competing waste management practices is ultimately economically motivated. Thus, the system will eventually include a tool for economic analysis of alternative management options.

Summary and Conclusions

The search for solutions to the many problems concerning nutrient management that affect water resources implies a continued demand for the development of modeling systems that can be used to analyze, in a holistic approach, the impact of alternative management policies.

The development of LOADSS exemplifies how the integration of hydrologic models and a GIS can be used for analyzing nutrient control practices at different scales. The addition of optimization algorithms further enhances the ability of policy- and decision-makers to analyze the impact of alternative management practices and land uses at the regional level.

The first part of LOADSS (Version 2.2) that includes the CREAMS-WT regional-scale model and the IDM components is fully functional and currently available at the SFWMD. Preliminary results show that LOADSS behaves consistently with measured data at the lake basin scale. Some of this, however, is due to offsetting errors in model behavior at the subbasin scale, particularly in subbasins that are adjacent to or very far from the lake. Currently, projects are underway to further verify and calibrate the model at the subbasin level to improve its

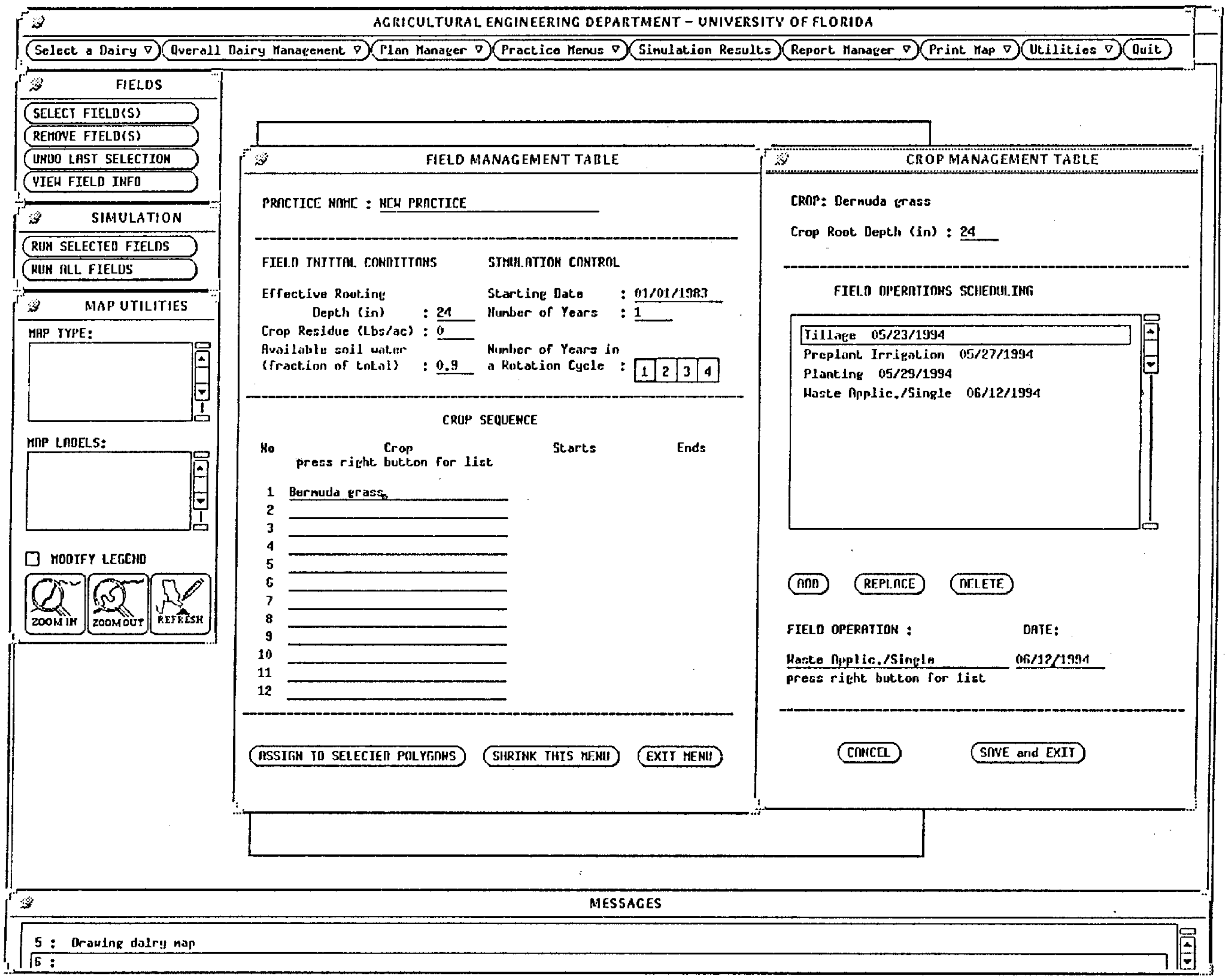


Figure 6. Field and crop management tables of the dairy model interface.

performance at smaller scales (16). Initial results of the optimization component are currently being evaluated.

The dairy model represents a different approach in integrating water quality models and GIS in that it is designed to be generic and focused mainly on the farm level. It is primarily designed to help policy- and decision-makers analyze the effects of alternative dairy waste management practices on the farm level. The framework can easily be adapted to handle different types of animal wastes (such as beef cattle and poultry) and to simulate the impact of other crop management practices such as pesticide applications.

References

1. Novotny, V., and G. Chesters. 1981. Handbook of nonpoint pollution: Sources and management. New York, NY: Van Nostrand Reinhold Company.

2. Council on Environmental Quality (CEQ). 1989. Environmental trends. Washington, DC: U.S. Government Printing Office.

3. U.S. Geological Survey (USGS). 1987. Analysis and interpretation of water quality trends in major U.S. rivers, 1974-81. Water Supply Paper 2307. Washington, DC: U.S. Government Printing Office.

4. SFWMD. 1989. Interim Surface Water Improvement and Management (SWIM) plan for Lake Okeechobee. West Palm Beach, FL: South Florida Water Management District.

5. Reck, W.R. 1994. GLEAMS modeling of BMPs to reduce nitrate leaching in Middle Suwannee River area. In: Proceedings of Second Conference on Environmentally Sound Agriculture, Orlando, FL. American Society of Agricultural Engineers. pp. 361-367.

6. Fedkiw, J. 1992. Impacts of animal wastes on water quality: A perspective from the USDA. In: Proceedings of the National Workshop on Livestock, Poultry, and Aquaculture Waste Management, Kansas City, MI (1991). American Society of Agricultural Engineers (ASAE) Pub. 03-92. pp. 52-62.

7. Woolhiser, D.A., and D.L. Brakensiek. 1982. Hydrologic system synthesis. In: Haan, C.T., H.P. Johnson, and D.L. Brakensiek, eds. Hydrologic modeling of small watersheds. ASAE Monograph No. 5. St. Joseph, MI: American Society of Agricultural Engineers. pp. 3-16.

8. Fedra, K. 1993. GIS and environmental modeling. In: Goodchild, M., B. Parks, and L. Steyaert, eds. Environmental modeling with GIS. New York, NY: Oxford University Press. pp. 35-50.

9. Tim, U.S., M. Milner, and J. Majure. 1992. Geographic information systems/simulation model linkage: Processes, problems, and opportunities. ASAE Paper No. 92-3610. St. Joseph, MI: American Society of Agricultural Engineers.

10. Whittemore, D.O., J.W. Merchant, J. Whistler, C.E. McElwee, and J.J. Woods. 1987. Ground-water protection planning using the ERDAS geographic information system: Automation of DRASTIC and time-related capture zones. In: Proceedings of the National Water Well Association (NWWA FOCUS) Conference on Midwestern Ground-Water Issues, Dublin, OH. pp. 359-374.

11. Petersen, G.W., J.M. Hamlett, G.M. Baumer, D.A. Miller, R.L. Day, and J.M. Russo. 1991. Evaluation of agricultural nonpoint pollution potential in Pennsylvania using a geographic information system. ER9105. University Park, PA: Environmental Resources Research Institute.

12. Goodchild, M. 1993. The state of GIS for environmental problem-solving. In: Goodchild, M., B. Parks, and L. Steyaert, eds. Environmental modeling with GIS. New York, NY: Oxford University Press. pp. 8-15.

13. Liao, H.H., and U.S. Tim. 1992. Integration of geographic information system (GIS) and hydrologic/water quality modeling: An interface. ASAE Paper No. 92-3612. St. Joseph, MI: American Society of Agricultural Engineers.

14. Young, R.A., C.A. Onstad, D.D. Bosch, and W.P. Anderson. 1989. AGNPS: A nonpoint source pollution model for evaluation of agricultural watersheds. J. Soil and Water Conserv. 44(2):168-173.

15. Leonard, R.A., W.G. Knisel, and D.A. Still. 1987. GLEAMS: Ground-water loading effects of agricultural management systems. Trans. of the Amer. Soc. of Agricul. Eng. (ASAE) 30(5):1,403-1,418.

16. Negahban, B., C. Fonyo, W. Boggess, K. Campbell, J. Jones, G. Kiker, E. Hamouda, E. Flaig, and H. Lal. 1993. A GIS-based decision support system for regional environmental planning. In: Proceedings of the Conference on Application of Advanced Information Technologies: Effective Management of Natural Resources, Spokane, WA. American Society of Agricultural Engineers (ASAE). pp. 169-178.

17. Heatwole, C.D., K.L. Campbell, and A.B. Bottcher. 1987. Modified CREAMS hydrology model for coastal plain flatwoods. Trans. of the Amer. Soc. Agricul. Eng. (ASAE) 30(4):1,014-1,022.

18. Kiker, G.A., K.L. Campbell, and J. Zhang. 1992. CREAMS-WT linked with GIS to simulate phosphorus loading. ASAE Paper No. 92-9016. St. Joseph, MI: American Society of Agricultural Engineers.

19. Skaggs, R.W. 1980. DRAINMOD reference report. U.S. Department of Agriculture, Soil Conservation Service. Fort Worth, TX: South National Technical Center.

20. Tremwel, T.K., and K.L. Campbell. 1992. FHANTM, a modified drainmod: Sensitivity and verification results. ASAE Paper No. 92-2045. St. Joseph, MI: American Society of Agricultural Engineers.

21. Knisel, W.G. 1980. CREAMS: A field-scale model for chemicals, runoff, and erosion from agricultural management systems. Conserv. Res. Rep. No. 26. Washington, DC: U.S. Department of Agriculture.

22. Priestly, C.H.B., and R.J. Taylor. 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Monthly Weather Rev. 100:81-92.

23. Jensen, M.E., R.D. Burman, and R.G. Allen, eds. 1990. Evapotranspiration and irrigation water requirements. Manuals and reports on engineering practice. Amer. Soc. Civil Eng. No. 70.

24. Reddy, K.R., R. Khaleel, M.R. Overcash, and P.W. Westerman. 1979. A nonpoint source model for land areas receiving animal waste: II. Ammonia volatilization. Trans. of the Amer. Soc. Agric. Eng. (ASAE) 22(6):1,398-1,405.

Using GIS To Rank Environmentally Sensitive Land In Orange County, Florida

Michael J. Gilbrook, HDR Engineering, Inc.[1]

ABSTRACT

In 1994, the Orange County (Florida) Planning Department selected HDR Engineering to conduct a GIS-based environmental constraints and development suitability study to support a new, proactive planning initiative. The goal of the project was to conduct a McHargian overlay analysis which would identify environmental constraints to development, and rank all non-urban lands within Orange County according to their environmental value. The project required development of a database consisting of eleven environmental constraint factors including vegetative cover, wetlands, wildlife habitat preferences, habitat corridors, floodplains, aquifer recharge potential and septic tank use suitability. A raster-based spatial model prepared in ARC/INFO GRID was used to identify and rank environmental constraints to future urban development based on the distribution and coincidence of the various environmental constraint factors. The HDR environmental science staff worked closely with the County Planning Department and the County's Environmental Mapping Committee (comprised of scientists, regulatory personnel, environmental activists, land owners, and planners) to prepare the GIS model and weighting scenarios. The final products included two baseline maps (one for ecological constraints, and one for physicochemical factors), and several maps generated by alternative weighting strategies.

INTRODUCTION

In December 1993, Orange County (Florida) hosted a Growth Management Exposition as a forum for unveiling the County's proposed Development Framework. The County's *Planning & Development Division Newsletter* released concurrently with the Exposition noted that:

> "The Orange County Comprehensive Policy Plan commits the County to the preparation of a Development Framework Element as a vision statement to guide the pattern of future development. ...The Comprehensive Policy Plan should contribute to, and be an embodiment of, the common vision that represents the

shared values and beliefs of the community for guiding future growth and development."

A "environmental suitability analysis" was among the tools cited by Orange County as necessary to the preparation and implementation of the Development Framework:

> "...[t]he addition of the concept of environmental suitability analysis into the Development Framework will strengthen the local planning process, facilitate efforts to obtain public ownership of environmentally sensitive lands, and increase the effectiveness of existing regulatory processes and procedures. It will also help separate and differentiate between planning for the future and regulating development."

That last point is crucial to understanding the importance of the County's environmental constraints analysis project. A typical regulatory approach to environmental planning addresses potential impacts piecemeal, as development is proposed, and attempts to minimize or mitigate ecological impacts. In contrast, Orange County wanted to develop a pro-active planning approach driven by an appreciation of the "carrying capacity" of the environment through composite mapping of environmental constraints to development. The environmental constraints mapping would then provide a long term guide for shaping future development patterns. Maps of environmental constraints can also be construed, inversely, as maps of development suitability (Twiss, 1975). Of course, a complete evaluation of development suitability would require combining the findings of the environmental constraints analysis with information on transportation systems, district plans, urban services, and proposed capital improvements.

In summary, the purpose of this study was to advance the County's movement towards a more holistic, ecosystem oriented approach to environmental planning. Specifically, HDR was tasked to "[d]evelop an environmental suitability index and map for Orange County that will compare and rank areas for environmental compatibility and development suitability according to the ecological characteristics of the region."

[1] HDR Engineering, Inc., 201 S. Orange Avenue, Suite 925, Orlando, FL 32801-3413; Voice (407) 872-7801, ext. 230; FAX (407) 872-0603; Email mgilbroo@hdrinc.com

OVERVIEW OF OVERLAY ANALYSIS

In his book, *Design With Nature*, planner Ian McHarg popularized the notion of map overlay analysis to evaluate land development suitability (McHarg, 1971). His method involved mapping landscape features which imposed some kind of limitation on urban development (such as wetlands, floodplains, or erosive soils) onto transparencies. For each map, the relative development suitability (or, conversely, environmental sensitivity) of landscape features would be ranked in order from high to low. Each map would be drawn at the same scale and registered to the same geographic area. The unsuitable areas on each map would be shaded in such a way that when the maps were overlaid one on another and viewed simultaneously on a light-table, areas which were mapped as unsuitable on all the contributing maps would appear very dark. Land areas which were unsuitable for development on fewer maps would appear proportionally lighter, providing a graduated scale from completely suitable to entirely unsuitable.

McHargian overlay analysis, as it came to be known, is a powerful planning technique. However, it was also inefficient and often impractical to perform by manual means. In the 25 years since the publication of McHarg's book, his technique has been refined though the use of computerized Geographic Information Systems (GIS). Examples of McHargian-type analyses using GIS are now quite common; the author had personally conducted two (Gilbrook, 1989a; Gilbrook, 1989b) prior to this study. The first step in conducting a McHargian analysis for Orange County was determining what environmental values would be incorporated into the evaluation process. The County Planning Department identified a tentative list of environmental criteria for the study. Working from that list as a starting point, HDR identified the following environmental issues which formed the basis for the environmental constraints analysis in Orange County:

- <u>Floodplain Protection</u> Development in floodplains places people and property at risk. Cumulative flood storage losses can exacerbate flooding elsewhere in the watershed. Alteration of floodplains adversely affects the hydrological, biological and physicochemical relationship between surface waters, wetlands, and uplands.

- <u>Protection of Wetlands</u> Wetlands provide critical habitat to a number of native central Florida species. Wetlands provide flood storage, hydrologic attenuation and erosion control functions.

- Maintaining Viable Wildlife Populations Wildlife populations must maintain minimum numbers in order to resist extinction from genetic inbreeding depression, disease, or climatic catastrophes. Viable populations require a minimum amount of suitable habitat to support them. Without protection (and maintenance) of adequate habitat, wildlife populations will likely become extinct over time.

- Preserving Biodiversity The natural environment is comprised of a complex assemblage of interdependent and co-evolved species. The effects of urbanization (e.g., forest fragmentation, enhancement of edge effects, fire suppression, invasive exotic species) serve to simplify and eliminate this diversity. Simplified ecosystems support fewer species and are less resilient to environmental challenges (e.g., storms, disease).

- Conservation Lands and Insular Ecology Conservation lands may maintain intact flora and faunal assemblages representative of the ecosystems present within those preserves as long as they enjoy contiguity with similar (or at least compatible) adjacent landscapes. Once isolated from similar habitat, conservation lands become essentially terrestrial islands subject to the species-area effects described by the science of insular ecology; the preserves lose species in geometric proportion to their final, isolated size.

- Groundwater Protection Drinking water supplies in central Florida are drawn from the Floridan aquifer, a deep, water-bearing geological stratum. The aquifer is recharged only by rainfall on relatively rare areas which have the necessary geologic conditions conducive to movement of water from the surface to the deep aquifer below. However, these areas are also conduits for potential contamination of the water supply for the same reason that they are good recharge sites. Similarly, the potential for contamination at the locations of drinking water wellheads and drainage wells (i.e., stormwater disposal wells to the aquifer) were also concerns for groundwater protection.

- Surface Water Protection Surface waters may be contaminated by pollutants borne by stormwater runoff from artificial impervious surfaces, made turbid by eroding sediment, or affected biologically by inappropriate management of shoreline wetlands or littoral vegetation. Septic tanks sited in areas of unsuitable soils may cause contamination of surface waters, presenting both an environmental hazard and a human health threat.

To conduct the study, each of these themes were translated into specific layers of digital data for analysis with GIS.

GIS DATA SOURCES

The environmental constraints analysis tapped a number of existing GIS data sources. Orange County and the two Water Management Districts which have jurisdiction over the County (St. Johns River WMD and South Florida WMD) were the source for most of the GIS data sets. Other important sources included the East Central Florida Regional Planning Council (Existing Land Use), the Florida Game and Fresh Water Fish Commission (Strategic Habitat Conservation Areas and Biodiversity "Hot Spots"), and the Florida Natural Areas Inventory (element occurrence records for rare plants and animals).

One of the most important data themes required was a current digital map of land cover to supply the information about wetlands, rare vegetative communities, and vegetative biodiversity needed to perform the environmental constraints analysis. A detailed digital Existing Land Use (ELU) map of Orange County, last updated in 1989, was available. Before these data could be applied to the environmental constraints analysis they had to be updated to reflect urban development which had taken place over the previous five years. To conduct the 1994 update, HDR scanned and georeferenced recent aerial photography for use in performing "heads up" digitizing of new urban land uses not reflected in the ELU 1989 data. The final Existing Land Use map appears in Figure 1; an inset of the map showing some of the features in the Econlockhatchee River basin of east Orange County appears in Figure 2.

Many of the GIS data sets were prepared from map sources compiled at a scale of 1:24,000 (i.e., 1" = 2,000') or smaller. This placed constraints on how finely the data could be interpreted. For example, the Florida Game and Fresh Water Fish Commission's (GFC's) Strategic Habitat Conservation Areas were developed from Landsat satellite imagery which had a minimum pixel resolution of 30 meters (about 98 feet, or 0.22 acres per pixel). In contrast, the Existing Land Use 1994 (ELU 94) data developed by HDR were captured at 1:24,000 scale, but had a minimum mapping unit of 2 acres; polygons of homogenous ground cover smaller than 2 acres were not necessarily mapped. Although these relatively low resolution input data sets limited their applicability for examining environmental constraints on small, individual parcels of land, their scale was appropriate for the regional analysis required by Orange County for the purpose of guiding comprehensive land planning.

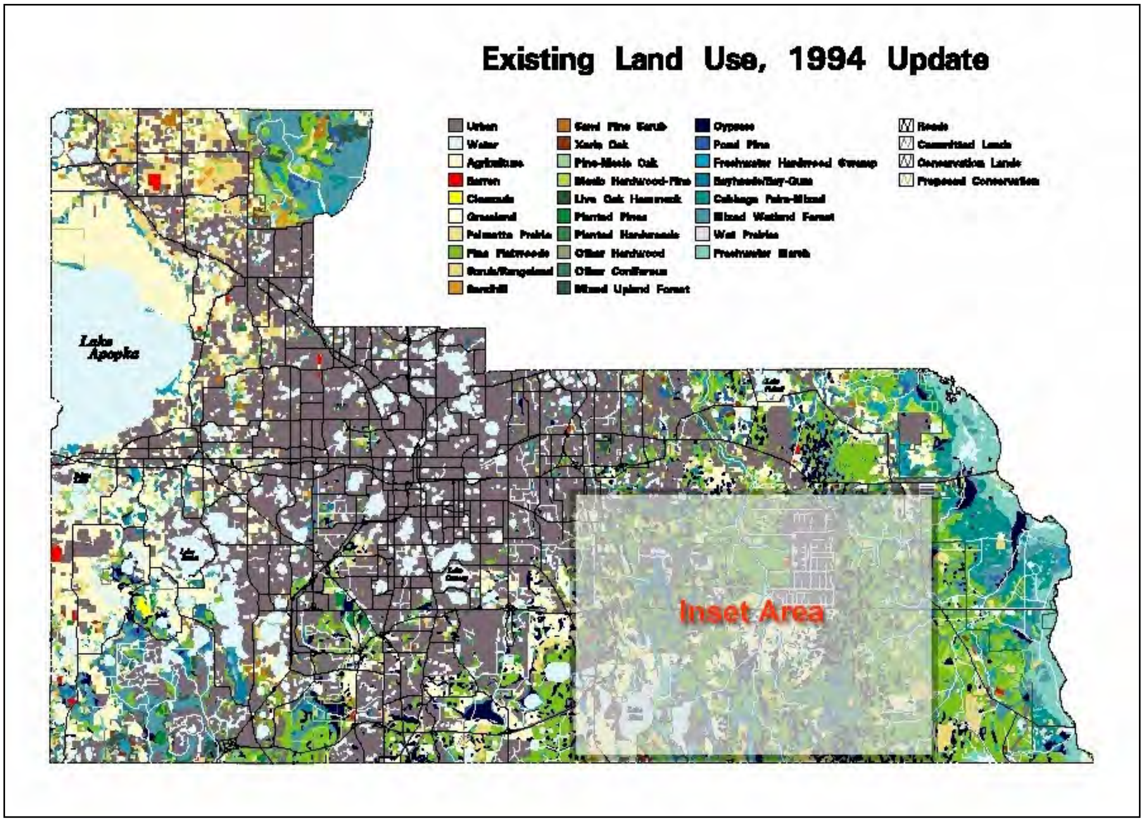


Figure 1. Existing Land Use 1994 Map. Dark gray areas are urbanized, other color represent undeveloped land cover. Inset area of Econlockhatchee River headwaters area appears in Figure 2 and subsequent map figures.

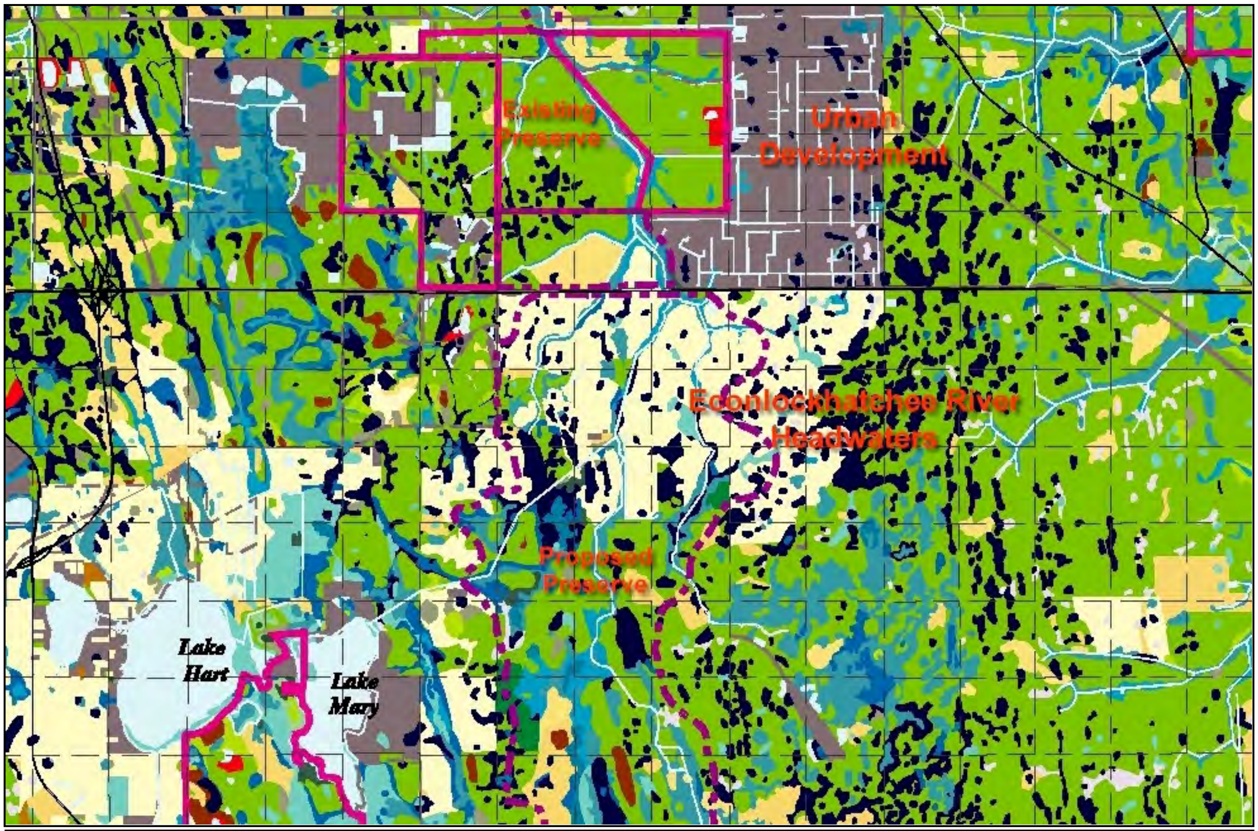


Figure 2. Inset of Existing Land Use 1994 in the Econlockhatchee River Headwaters Area, Southeastern Orange County.

GIS MODELING METHODOLOGY

This study employed ARC/INFO 6.1.1. GIS software manufactured by Environmental Systems Research Institute (ESRI). The software ran on a Sun Microsystems SPARC 10 Model 40 workstation under the Solaris 2.2 operating system (a variant of Unix System V).

ARC/INFO is principally a vector GIS system in which linear features (lines or polygons) appear as smooth lines connecting the many vertices whose coordinates define the shape of those features. In this respect an ARC/INFO visual display looks little different from a map prepared using Computer Aided Design (CAD) software. Nearly all of the preliminary GIS manipulation needed to produce the input data themes used in the study was conducted within the ARC/INFO vector environment. However, one drawback to the vector GIS environment is that it is a computationally intensive process. For ARC/INFO to combine two sets of countywide polygon data to create a third using a vector overlay process (e.g., UNION or INTERSECT) would have required considerable time. Overlays involving multiple input themes would have had to be processed in a series of pairs, until the final desired product was achieved. Even then, the GIS analyst would have had to use tabular database functions to resolve the meaning of the multiple layers of polygon data which had been combined.

Although the environmental constraints analysis could have been performed this way, a better approach was available through use of ARC/INFO's GRID module. Combining and manipulating GIS data in GRID was computationally much easier (and faster) than the topologically complex process of resolving the output from overlays involving many arbitrarily shaped vector polygons. Furthermore, the GRID process allowed for direct mathematical modeling of the input data sets to generate the desired composite output map; for example, differential weighting of the input data was as easy as multiplying all cells in an individual data layer by a numerical constant. Given its clear advantages, HDR used GRID for modeling in this study. Each of the vector polygon GIS input data themes was converted to a grid with a cell size of 98 feet. The 98 feet corresponded to the 30 meter resolution of the Landsat data which formed the basis for the GFC’s Strategic Habitat Conservation Areas, Biodiversity "Hot Spot" areas, and the SJRWMD's Regionally Significant Habitat areas. A cell resolution of 98 ft resulted in a minimum mapping of 0.22 acres, which we deemed small enough to adequately represent details in the various input coverages collected at 1:24,000 scale.

Following examination of the available GIS data sets, consideration of the study's objectives, discussions with the Orange County Planning Department staff, and input from the Environmental Mapping Advisory Committee, HDR determined that the most appropriate end product would not be a single map of environmental constraints, but *two* composite maps: an "Ecological Constraints Map" and a "Physical Constraints Map." The Ecological Constraints Map was derived from those factors whose attributes (mostly biological) denoted resources that were sensitive to *any* land alterations. For example, a forested area that was important for the maintenance of wildlife populations was likely to be adversely affected by substantial clearing. The type of urban development (e.g., residential, commercial or industrial) makes no difference; it is the loss of forest habitat, the reduction in forest interior, or the increase in fragmentation and edge effects which are important. Protection of such areas would require attention to development *intensities.*

The Physical Constraints Map was generated from input data themes which denoted constraints based on *physical* factors. Protection of these resources would require attention to the *type* of land use or the *development standards* imposed on that development. Aquifer recharge areas are the best example. The presence of a high recharge area does not necessarily preclude urban development. Instead, a recharge area may influence the type of land use considered suitable (e.g., no chemical industries), the enactment of special development standards (e.g., higher on-site runoff retention requirements), or both. This paper will only address the methodology used to prepare the Environmental Constraints Map; the process for generating the Physical Constraints Map was identical, differing only in the types of inputs involved. Several GIS layers contributed to each of the two composite maps of environmental constraints. In both models, we used a five-point scale or index to rank environmental sensitivity for a given GIS data layer. In all cases, a "1" indicated presence of "Very Low" environmental constraints for urban development, a "3" a "Moderate" level of environmental sensitivity, and a "5" indicated the presence of "Very High" environmental constraints. In the GIS analysis, the various contributing themes were combined or overlaid in such a way that each spot on the map represented an average of all the environmental constraint scores contributing to that map. Areas which had many high environmental sensitivity scores from contributing input data themes received a high total score, while areas with few or no constraints received a low composite score.

Initially, all factors were given equal weight in the overlay procedure to produce two "baseline" maps. After the "baseline" maps were produced, weighting factors were assigned to each of the GIS input layers to reflect the relative importance of their contribution to the ranking of environmentally sensitive lands in the County. The value of the weighting analysis was two-fold. First, application of weighting to a particularly important input data layer would permit it to "shine through" the muddle which might otherwise result from the combination of so many disparate input data sets. Second, the use of weighting tested the robustness of the findings of the GIS analysis. That is, the environmental sensitivity of areas which ranked similarly on several maps despite alternative weighting schemes could be interpreted with confidence. For instance, one would conclude that an area which showed up as having "Very High" environmental constraints in every map, regardless of weightings, was clearly dominated by environmentally sensitive factors. The weighting factors were developed in a workshop meeting of the Environmental Mapping Advisory Committee in concert with HDR and the Orange County Planning Department staff. The particulars of the weighting process are described below, following the discussion of the input data themes.

INPUT DATA FOR THE ECOLOGICAL CONSTRAINTS MAP

We identified nine different GIS input data layers, or themes, for use in generating the Ecological Constraints Map (Table 1). With two exceptions, all the input data layers were "derived" by combining two or more of the original GIS data sources in various ways, or by re-casting the original data set in a more useful form. Each input data layer is described below.

TABLE 1

Input Data Layers For Ecological Constraints Map GIS Model

	Environmental Constraint Factor				
GIS Input Data Layer	**Very Low (1)**	**Low (2)**	**Moderate (3)**	**High (4)**	**Very High (5)**
Floodplain Areas	Non-Floodplain	N/A	N/A	100 Year Floodplain	Regulatory Floodways
Wetland Areas	Uplands	N/A	Wet Prairie	Non-Forested Wetlands	Forested Wetlands
Ecological Integrity	All Other Land	SJRMWD RSH Areas	5 - 7+ Focal Species	GFC SHCAs	FNAI & GFC E & T Occurrence Point Locations
Vegetative Community Rareness	All Other Land	Area > 2%, < 10%	FNAI S3 or Area > 1%, < 2%	FNAI S2 or Area > 0.5%, < 1%	FNAI S1 or Area < 0.5%
Habitat Corridors/Biological Connectivity	Very Low Proximity Or Urban	Low Proximity, Agricultural lands	Medium Proximity, Grassland	High Proximity, Forested	Very High Proximity, Wetland or Water
Wetland Dominance	< 20 percentile, Urban or Water	20 – 40 percentile	41 – 60 percentile	61 – 80 percentile	> 80 percentile
Floodplain Dominance	< 20 percentile, Urban or Water	20 – 40 percentile	41 – 60 percentile	61 – 80 percentile	> 80 percentile
Vegetative Biodiversity	< 20 percentile, Urban or Water	20 – 40 percentile	41 – 60 percentile	61 – 80 percentile	> 80 percentile

Notes: (1) Class 5 reserved for Regulatory Floodways in Floodplain Dominance theme, but not used in this study; (2) SJRWMD = St. Johns River Water Management District; (3) GFC = Florida Game & Freshwater Fish Commission; (4) FNAI = Florida Natural Areas Inventory; (5) SHCA = Strategic Habitat Conservation Area; (6) RSH = Regionally Significant Habitat; (7) The Environmental Sensitivity score for Habitat Corridors/Biological Connectivity depended on both habitat type and proximity to the most direct path between two preserves (see Table 2).

Floodplain Areas. This was one of the two "non-derived" data themes. Floodplains were a "dichotomous" factor: areas above the 100-year floodplain boundary were ranked "1," whereas all lands identified as being within the 100-year floodplain (i.e., those designated "Zone A" on FEMA Flood Insurance Maps) were assigned a environmental constraint factor of "4." There were no intermediate classes. Class "5" was reserved for regulatory floodways, which would have received the greatest protection from alteration. Regulatory floodway boundaries were not available for this study, but their spot was reserved so that they could be inserted into the analysis and the model run again at a later time.

Wetland Areas. This is the other non-derived theme. Wetlands were ranked in order of their relative difficulty to maintain or re-create. Forested wetlands as a rule take a considerable time to grow to maturity, and are not easily re-created by humans; consequently, they were ranked "5." Non-forested wetlands (i.e., marshes) are extremely productive wetland environments, but are somewhat more readily replaced than forested wetlands; hence, they were assigned a value of “4.” Since vegetative communities similar to natural wet prairie areas are commonly created on pastures, such areas were assigned a “3.” All other (upland) areas were designated as "1."

Scientists, planners and developers often disagree on whether wetland size is a valid criterion for evaluating wetland value. While it is true that small, isolated wetlands in an urban setting may have little or no wildlife habitat value, this is emphatically *not* true of such wetlands immersed in a matrix of natural upland vegetation. Small, ephemeral wetlands are essential to the life cycles of many amphibians which cannot survive predation by fish in larger ponds. Furthermore, small wetlands are essential to the feeding (and nesting) success of wading birds, particularly the endangered wood stork. And although large wetlands might seem to dominate surface water hydrology, the depressional storage capacity of many small wetlands may be considerable. Since there was no good way to rank the ecological importance of wetlands based on size, size did not contribute to the ranking of the Wetland Areas theme for this study.

Ecological Integrity. This is a term which has gained in popularity to describe the process of protecting natural diversity, at scales ranging from populations to entire ecosystems (Minasian, 1994). Consistent with that concept, this input data layer was comprised of several data sources which had been combined to represent areas of hierarchically greater or lesser importance to the maintenance of natural floral and faunal populations in the County. Most sensitive on this scale (i.e., a ranking of "5") were the reported locations of species listed as endangered, threatened or species of special concern by the U.S. Fish & Wildlife Service, Florida Game and Fresh Water Fish Commission, or by the Florida Natural Areas Inventory. These point occurrences were obtained from the FNAI Element Occurrence database, and were represented by a 1,000 ft. radius circular area generated as a buffer around each point location. These circles, which are comparable to those used by Cox, et al. (1994) to highlight point occurrences of listed species or other significant wildlife, encompassed an area of about 72 acres each.

The map data used to identify lands ranked "4" for this data theme were obtained from digital maps of Strategic Habitat Conservation Areas (SHCAs) prepared by the Florida Game and Fresh Water Fish Commission (Cox, et al., 1994). Using a statewide map of vegetative cover derived by interpretation of Landsat satellite imagery collected during the mid- to late 1980's, the GFC identified polygons of vegetative cover which represented the habitats of 30 "focal species," most of which were listed as endangered or threatened. (Seventeen of the GFC's focal species occurred in Orange County, including the red-cockaded woodpecker, Florida scrub jay and gopher tortoise.) By mathematically modeling the minimum viable population sizes needed to ensure survival of the selected focal species for 200 years, the GFC then estimated the minimum necessary habitat required to maintain such populations in perpetuity. Following a GIS analysis of the Landsat-derived vegetative cover maps, the GFC located SHCAs which, if preserved, would secure the long term survival of the 30 focal species evaluated.

Using the same Landsat-derived vegetative cover maps, the GFC prepared a map of biodiversity "hot spots." To identify areas which might jointly serve a number of important wildlife species, the GFC overlaid the individual habitat maps of their focal species and identified areas whose habitat could support seven or more species, five to six species, or less than five species. Those areas identified as supporting either 5-6 or 7+ species were combined and ranked "3" in the Ecological Integrity data theme.

The St. Johns River Water Management District used the SHCA data as the basis for its identification of Regionally Significant Habitat (RSH). Since the method the GFC employed to identify SHCAs was tied to the habitat requirements of individual species, their process sometimes identified polygons of vegetative cover which were surrounded by other native vegetation *not* included in the SHCA. While preparing their RSH maps, the SJRWMD recognized the need to identify areas which could be either effectively regulated, or publicly acquired and managed; disjunct SHCAs embedded in other natural land cover wouldn't qualify. Consequently, the SJRWMD used the Cox (1994) GIS data in combination with vegetative cover from Orange County's 1989 Existing Land Use to "extend" the SHCA boundaries to the limits of immediately adjacent natural vegetative communities. The SJRWMD RSH areas were assigned a ranked of "2, ” and all other lands not included in one of the above classes were ranked as "1." A small part of the Ecological Integrity theme appears in Figure 3.

FNAI were assigned to ranks 5, 4, and 3, respectively. All other community types were ranked as "1." To identify rare community types in Orange County, HDR used the 1994 Existing Land Use (ELU ’94) data to calculate the total acreage for each non-urban, non-agricultural cover type. Based on these data, we assigned a rank of 5 to those communities comprising 0.5% or less of the total natural vegetative cover of approximately 291,000 acres. Rank 4 was assigned to communities representing between 0.5% and 1.0% of the total natural land cover, while those communities with 1% to 2% were assigned a rank of "3," and a rank of "2" attributed to communities falling between 2% and 10%. Natural land cover types with greater acreage were all assigned ranks of "1." Where the FNAI and Orange County ELU '94 derived ranks disagree, the vegetative community was assigned the higher (i.e., more sensitive) rank. Figure 4 illustrates a portion of the Vegetative Community Rareness theme.

Habitat Corridors/Biological Connectivity. This GIS data layer identified areas which may be important to the maintenance of biological connectivity between managed conservation areas. Put another way, this layer identified areas which should be protected to prevent the adverse effects of forest fragmentation and biological isolation in natural preserves. The idea of considering conservation lands as the anchor points for habitat corridors has a strong following in Florida (Noss, 1991; Harris and Atkins, 1991). We constructed this data theme from a SJRWMD coverage of existing and proposed conservation lands (hereafter called "preserves"), and habitat value as derived from the 1994 Existing Land Use data. Using the ARC/INFO GRID functions COSTDISTANCE and CORRIDOR, we identified areas representing the most direct connections between pairwise sets of proposed or existing preserves. We eliminated from consideration any preserve pairs for which there were either interposing urban or preserve areas, or which were more than 12 miles apart, and rated areas within the corridor in descending order according to their proximity to the shortest distance path: 1,000 feet, 0.5 miles, 0.75 miles and 1 mile. The corridors were overlaid with the habitat rank grid derived from

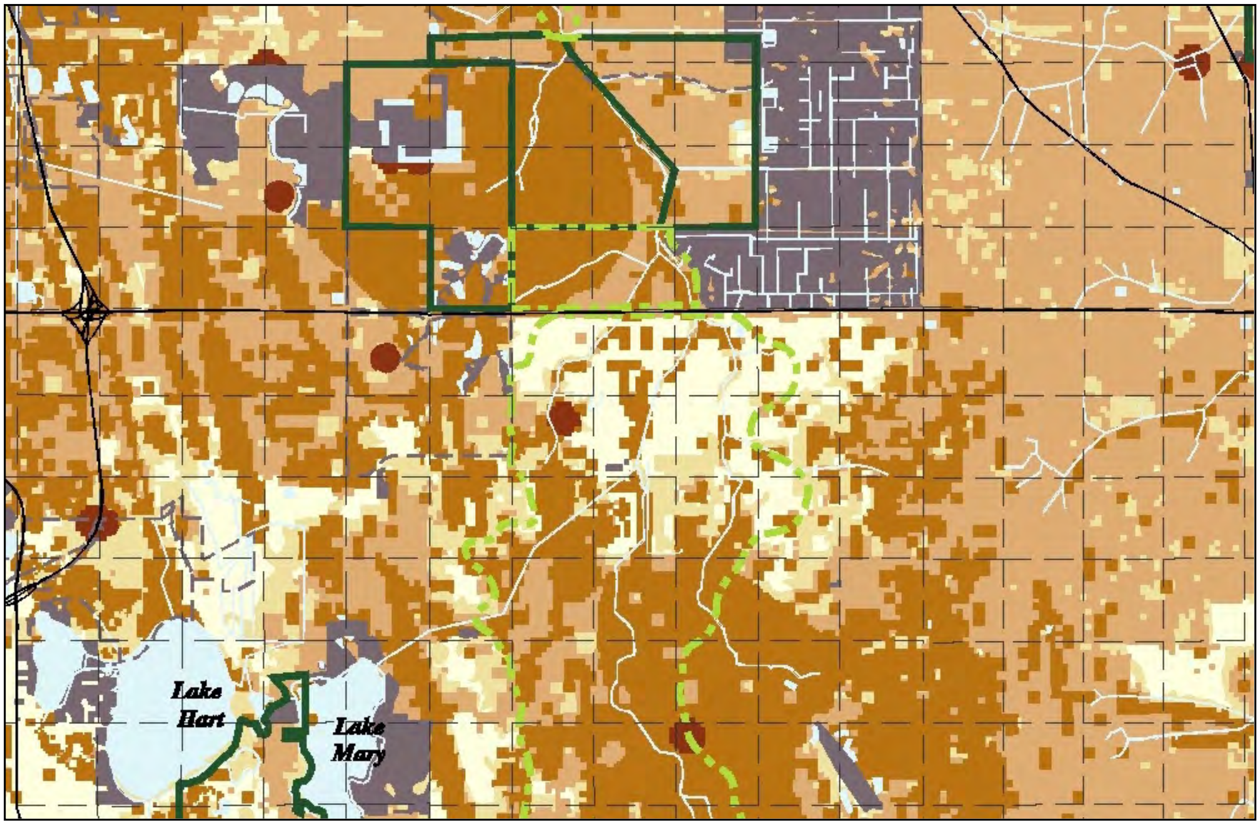


Figure 3. Ecological Integrity Theme. Dark colors represent areas of highest environmental constraints, light areas lowest. Dark circles are the 1000-foot radius circles around the observed locations of endangered and threatened species.

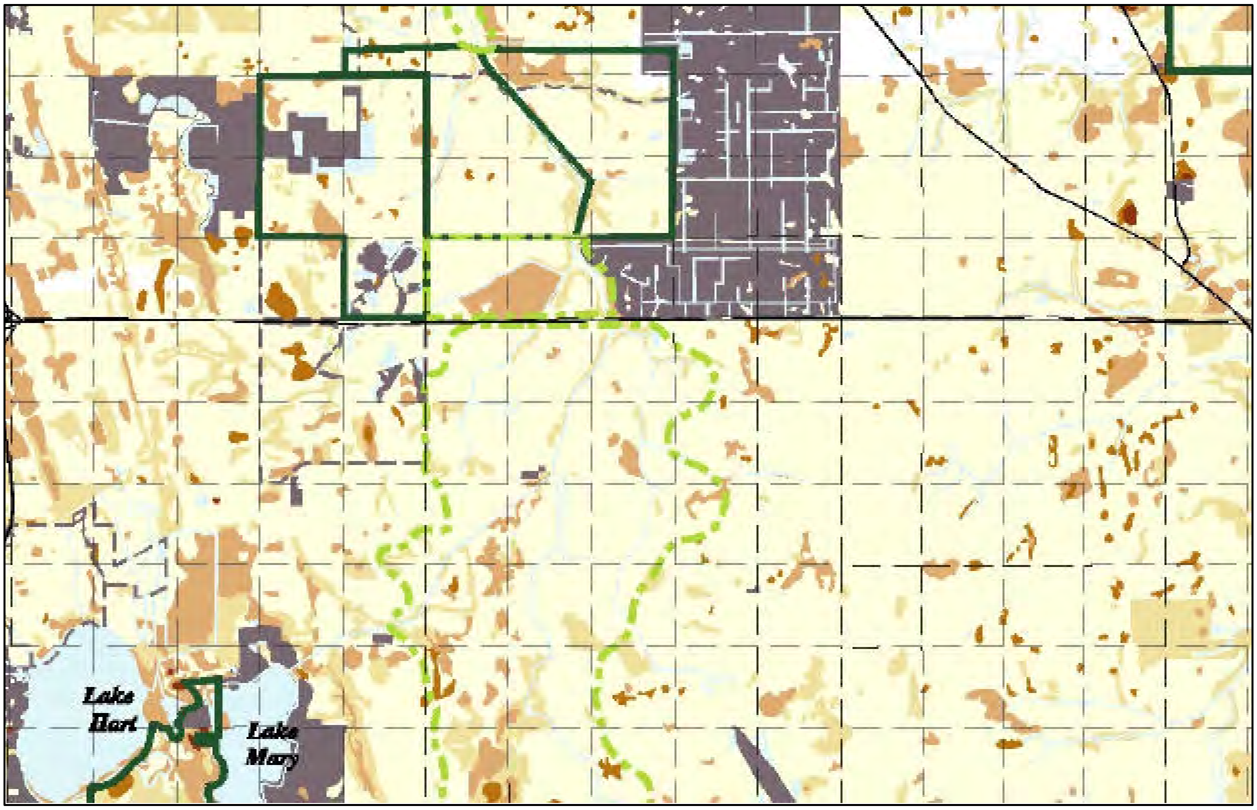


Figure 4. Vegetative Community Rareness Theme. Common natural communities appear in light colors. Darker shades of brown indicate communities of increasing rarity.

TABLE 2

Environmental Constraint Values Assigned To The Habitat Corridor Input Data Theme Based On Habitat Value And Corridor Centerline Proximity Scores

	Habitat Value Rank				
Proximity Rank	**1 (Urban)**	**2 (Agricultural)**	**3 (Non-Forested)**	**4 (Forested Uplands)**	**5 (Wetlands)**
1 (> 1.0 mi.)	1	1	1	1	1
2 (£ 1.0 mi.)	1	2	2	2	2
3 (£ 0.75 mi.)	1	2	3	3	3
4 (£ 0.5 mi.)	1	2	3	4	4
5 (£ 1000 ft.)	1	2	3	4	5

the ELU '94 land cover data to differentiate among corridor alternatives based on preferred vegetative cover. The proximity/land cover matrix used to assign final environmental constraint rank values to grid cells for this theme appears in Table 2, and a part of the Habitat Corridors map appears in Figure 5.

Wetland Dominance. The Wetland Areas GIS data layer addressed those parts of the landscape actually occupied by a wetland community type. However, in planning for the protection of natural resources in Orange County, it will not be possible to delineate every individual wetland for protection. Nonetheless, large upland areas characterized by many small wetlands are themselves ecologically valuable. Faunal and floral biodiversity of many small wetlands interspersed throughout an upland matrix will likely be greater than that of a single large wetland of equivalent acreage. Furthermore, the life cycles of many species are dependent upon smaller wetlands. Amphibians prefer small, ephemeral wetlands for breeding because such wetlands do not support a large population of predatory fish. Wading birds (notably the wood stork) benefit from the concentration of fish and amphibian prey within small, ephemeral wetlands. To address these issues, we used ARC/INFO to calculate the ratio of wetlands to non-urban uplands within each of the 1,000 one-square mile sections of the Public Land Survey (PLS). We assigned each PLS section a rank of 1 to 5 based on its placement in a quintile distribution.

Floodplain Dominance. As with wetlands, isolated floodplains may collectively have great hydrological or ecological value if they are sufficiently numerous or collectively large enough. To identify those areas which have a large portion of their land area within the 100-year floodplain, we generated this GIS data theme using exactly the same procedure as that for Wetlands Dominance, only using the Floodprone Areas GIS layer as the starting point.

Vegetative Biodiversity. All other factors being equal, an area with many different vegetative cover types (i.e., a high biodiversity) will typically be more ecologically valuable than an equal sized area with fewer different vegetative communities. To evaluate biodiversity, we could have the GIS simply count the number of different kinds of vegetative cover types in a PLS section, then use the quintile procedure outlined above. However, the mere *number* of different kinds of community types does not tell the whole story. For example, two sections might have the same number of different vegetative community types, but one might have six equally sized polygons, while the other section has one very large polygon and five very small ones. Clearly, the section

Figure 5. Habitat Corridors/Biological Connectivity Map. The darker colors represent areas of high habitat value near to the most direct path linking two existing or proposed preserves. Dark green solid lines border existing preserves, light green dashed lines surround proposed preserves.

with more equally sized polygons is more diverse, since the other section is almost entirely dominated by one cover type. In order to quantify landscape diversity, we applied the Simpson C' Index, which is generally accepted as a measure of population diversity (Krebs, 1989). Using the acreage information for vegetative communities within the ELU '94 GIS data, we calculated Simpson C' indices for all PLS sections, then subjected those values to a quintile analysis like that used for Wetlands and Floodplains Dominance, above. PLS sections were assigned the appropriate 1 through 5 ranks, with "5" representing the most diversity (highest Simpson's C' indices), and "1" the least diverse assemblages of community types. Barren or agricultural areas that did not contribute to the biodiversity measurement were assigned a value of "1," thereby ensuring that an entire PLS section did not receive a high biodiversity score based on a small natural remnant on an otherwise barren landscape. A sample of the Vegetative Biodiversity Map appears in Figure 6.

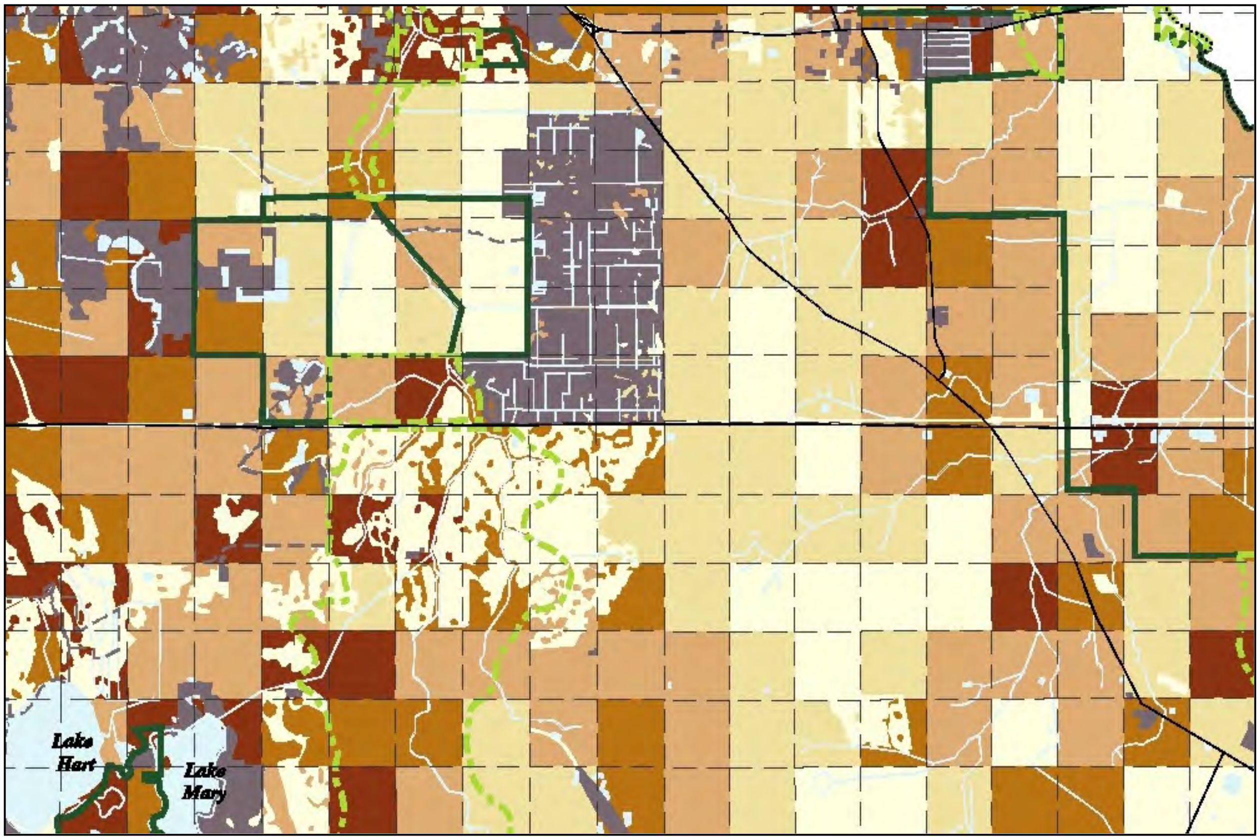
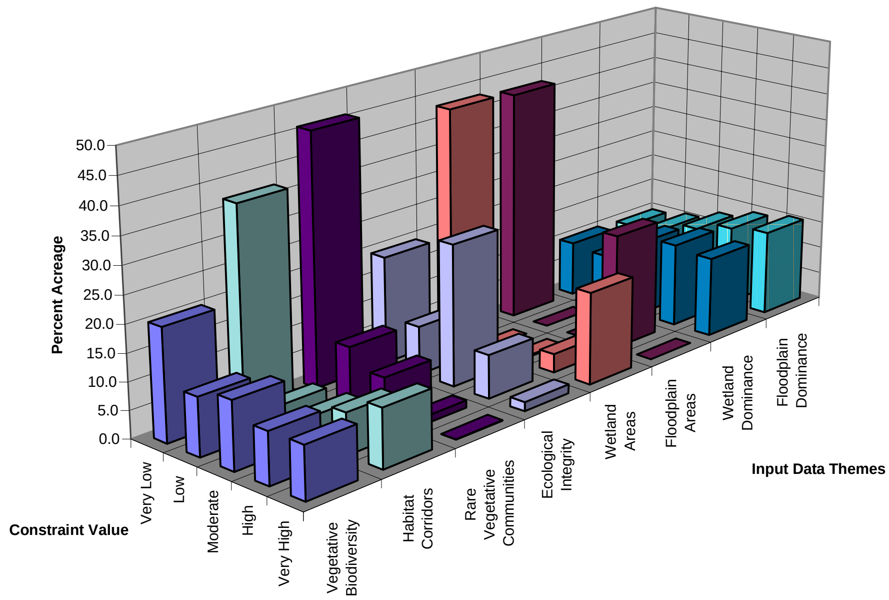


Figure 6. Vegetative Biodiversity Theme. Dashed lines represent section lines from the Public Land Survey. Each one-square mile section is color coded, from light to dark, based on its quintile rank among the distribution of all Simpson's C' indices calculated for the 1000 square miles in Orange County.

Areal Analysis Of Environmentally Sensitive Input Data. Figure 7 shows the relative amounts of undeveloped Orange County which were assigned to each of the five classes of environmental sensitivity for each of the input data layers contributing to the Ecological Constraints Maps. For most of the input data themes, the majority of County land was rated "Very Low," with a much lower percentage assigned to each of the more environmentally sensitive classes. The Wetland Dominance, Floodplain Dominance and Vegetative Biodiversity themes illustrated a nearly equitable distribution of land in each of the five constraints classes, owing to the quintile assignment methodology employed for those layers.

Figure 7
Comparison of Acreage Assigned to Environmental Sensitivity Classes for Each of the Ecological Constraints Input Data Themes



ALTERNATIVE WEIGHTING ANALYSES

Table 3 depicts the three alternative weighting schemes used for the Ecological Constraints composite maps. The first was dubbed the "Ecological Integrity" weighting option, since it weighted those factors most closely associated with that concept (i.e., the Ecological Integrity layer and the Habitat Corridor layer). The "Habitat Diversity" model weighted the Vegetative Biodiversity and Vegetative Community Rareness layers. Finally, the "Wetlands" model weighted the wetland boundary and wetland dominance layers.

In all cases, a maximum weight of "2" was used. This was the minimum whole-integer weight that can be applied, yet have a demonstrably visible effect on the outcome of the composite maps. To evaluate the appropriateness of the "2" factor, we conducted a sensitivity analysis on the Ecological Constraints map by running the Ecological Integrity model with weights of 2, 3, 4 and 5. We found that increases in the weight beyond 2 merely had the effect of making the output composite map look more like the weighted layers, and diminished the value of all other inputs. In contrast, the weight of 2 left most composite map effects intact, while emphasizing certain features from the weighted layers.

TABLE 3

Ecological Constraints Map Alternative Weighting Factors

GIS Input Data Layer	Alternative Weighting Schemes: Ecological Integrity	Habitat Diversity	Wetland Protection
Floodplain Areas	1	1	1
Wetland Areas	1	1	2
Ecological Integrity	2	1	1
Vegetative Community Rareness	1	2	1
Habitat Corridors/Biological Connectivity	2	1	1
Wetland Dominance	1	1	2
Floodplain Dominance	1	1	1
Vegetative Biodiversity	1	2	1

RESULTS & DISCUSSION

The areas ranked as most environmentally sensitive by the Baseline Ecological Constraints model (Figure 8) encompassed many locations already recognized by conservation planners as having high ecological value, including: Wekiwa Springs State Park, Kelly Park, Moss Park, Split Oak Mitigation Park, the various public lands associated with the St. Johns River, and the various areas proposed for acquisition within the headwaters of Reedy Creek, Shingle Creek and the Econlockhatchee River. The fact that the Baseline Ecological Constraints model identified existing and proposed conservation areas as among the most environmentally important lands in Orange County validated the model's credibility.

The three composite maps produced using weighting schemes were mostly very similar to the Baseline map, while exhibiting some minor differences attributable to the weighting scenario. The similarity of the weighted composite maps further reassured us that the basic premise of the environmental constraints map was sound, since the general pattern of environmentally sensitive land was relatively insensitive to changes in the input weights. However, the variations in the weighted maps did afford an opportunity for County planners to further evaluate those areas which appeared different and determine if they required more (or, perhaps, less) consideration in the development of land use plans or environmental protection ordinances. Some brief comments about the three weighted maps follow:

Ecological Integrity Weighted Map. The pattern on this map looked very much like that on the Baseline Model map. Areas along the potential habitat corridor between the Econlockhatchee River and the St. Johns River conservation lands ranked slightly higher than those same areas in the Baseline model, illustrating the effect contributed by the Habitat Corridors/Biological Connectivity theme to this map.

Habitat Diversity Weighted Map. As with the Ecological Integrity Weighted Map, the pattern of this map appeared very similar to the Baseline. It showed much less area designated as having "Very High" environmental constraints, but the loss occurred mostly within the limits of existing and proposed public lands. This effect reflects the fact that many public lands were dominated by large wetlands, whereas the Habitat Diversity Weighted model was more sensitive to rare upland communities, or diverse mixes of uplands and wetlands.

Wetland Importance Weighted Map. This map had much greater areas of "Very Low" and "Low" environmental constraints, especially in eastern Orange County. In other words, by emphasizing the importance of wetlands, the ecological value of upland areas was diminished. Most of the areas identified as "High" or "Very High" in the Baseline Map appeared similarly ranked here, but a greater amount of these areas followed wetland boundary lines. Furthermore, this map showed an increased tendency for "Moderate" ranked areas to conform to PLS section lines, an effect no doubt produced by the section-based wetland dominance layer. The headwaters areas of Reedy Creek, Shingle Creek and Lake Sheen still ranked very high.

Relative Acreages For Ecological Constraints Composite Maps. Figure 9 compares the percent Orange County acreage which fell to each of the class ranks for each model type. The overall pattern for all maps appears almost identical: Maximum acreage was associated with the "Low" category, with gradually diminishing acreage values for the higher ranks. The most visible difference in the amount of acreage in each constraint class appeared in the "Very High" level, in which the Wetland Importance model ranked nearly 8% of the County's non-urban, non-water area, compared to just under 5% for the Baseline and Ecological Integrity models, and 2.4% for the Habitat Diversity model.

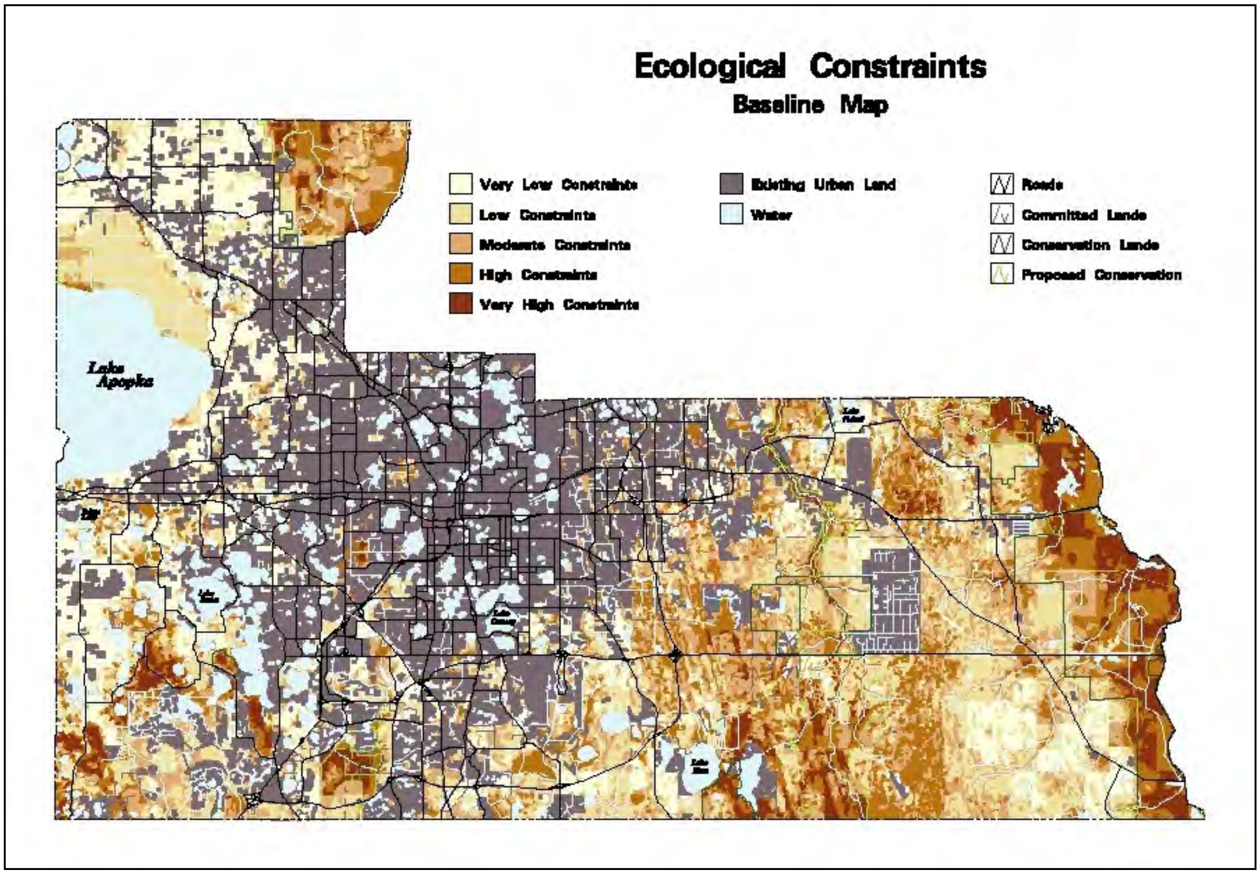
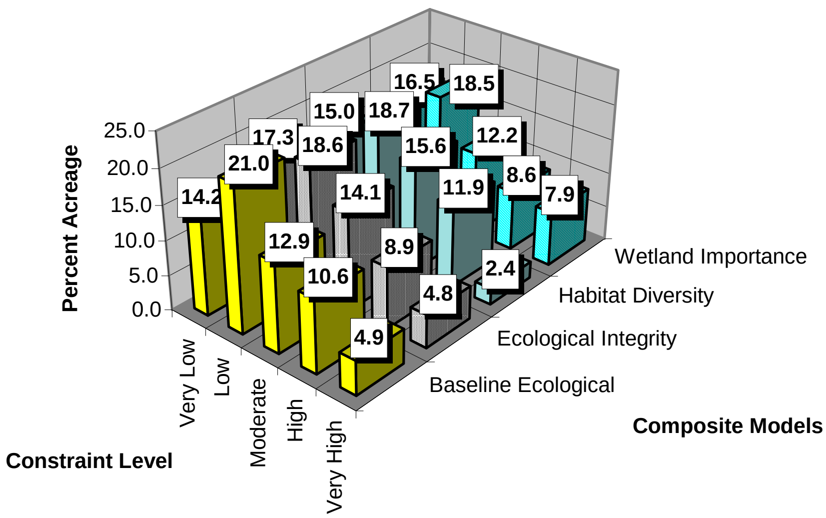


Figure 8. Baseline Ecological Constraints Composite Map. Gray areas represent existing urbanization. Environmental constraints in undeveloped areas are represented by shades of brown, from lightest ("Very Low Constraints") to darkest ("Very High Constraints").

Figure 9
Comparison of Percent of Undeveloped Acreage Assigned to Each Environmental Sensitivity Rank for the Ecological Constraints Models



CONCLUSIONS

The ARC/INFO GRID based environmental constraints model proved to be an efficient and effective way to prepare a composite map of environmental suitability as envisioned in Orange County's Development Framework. The modular nature of the GIS model provides for relatively easy updating of the input data layers and production of updated constraints maps as new data become available.

The maps produced by the alternative weighting models for the Ecological Constraints Map produced results which were very similar to that of the Baseline Map. The constancy of certain core areas on all weighting models bolsters the conclusion that these areas are truly environmentally sensitive. The areas which were affected by changes in weighting parameters provide an opportunity to evaluate which areas, outside of the obvious core areas, should be included in future public lands acquisition plans, special future land use planning, or other appropriate protection mechanisms.

ACKNOWLEDGEMENTS

The author wishes to recognize the members of the Environmental Mapping Advisory Committee for donating their valuable time and input to the successful completion of this study: Jack Amon, Lester Austin, III, Jim Bradner, Vera Carter (Chair), Nancy Christman, Michael Dennis, Barbara Durkin, Jeff Jones, Nancy Prine, John Richardson, Jim Sellen, P.K, Sharma, Bill Stimmell, Jack Stout, Renee Thomas, Jim Thomas (Vice Chair), Rick Walker, and John Winfree. The insightful comments and suggestions of this group frequently found their way into the final product.

The staff of the St. Johns River Water Management District provided invaluable assistance in supplying digital data, particularly Jim Cameron, Stuart Dary, Mike Kleinman and Linda McGrail. John Stys of the Florida Game and Fresh Water Fish Commission quickly supplied us with the digital raster files of the Commission’s gap analysis study. Thanks also to Scott Taylor of the Florida Natural Areas Inventory for his assistance in preparing copies of FNAI's element occurrence data.

Finally, I’d like to recognize the assistance and cooperation afforded by the Orange County Planning Department staff, especially that of Andre Anderson and Danielle Justice, who many times facilitated the progress of the project. This project was funded by Orange County Contract #Y4-635.

LITERATURE CITED

Cox, J., R. Kautz, M. MacLaughlin, and T. Gilbert. 1994. *Closing the Gaps In Florida's Wildlife Habitat Conservation System.* Office of Environmental Services, Florida Game and Fresh Water Fish Commission. Tallahassee, Florida.

Gilbrook, M.J. 1989a. *Marina Siting Suitability in the Coastal Estuaries of East Central Florida*. Florida Department of Environmental Regulation, Coastal Zone Management Program, Contract CZM-200. East Central Florida Regional Planning Council. Winter Park, Florida.

Gilbrook, M.J. 1989b. *Wekiva River Basin Acquisition Study*. Final report to the St. Johns River Water Management District. East Central Florida Regional Planning Council. Winter Park, Florida.

Harris, L.D. and K. Atkins. 1991. Faunal movement corridors in Florida. In: Hudson, W.E. (ed.). *Landscape Linkages and Biodiversity*. Island Press. Washington, D.C.

Krebs, C.J. 1989. *Ecological Methodology.* Harper & Row, Publishers, Inc. New York.

McHarg, I. 1971. *Design With Nature*. Doubleday/Natural History Press. Doubleday & Company, Inc. Garden City, New York.

Minasian, L. 1994. Interpreting and applying ecosystem management principles. *Environmental Exchange Point*. Florida Department of Environmental Protection, 4(3):17-23.

Noss, R.F. 1991. Landscape connectivity: Different functions at different scales. In: Hudson, W.E. (ed.). *Landscape Linkages and Biodiversity*. Island Press. Washington, D.C.

Twiss, R.H. 1975. Commentary - Nine approaches to environmental planning. In: Burchell, R.W. and D. Listokin (eds.). *Future Land Use: Energy, Environmental and Legal Constraints.* Center For Urban Policy Research, Rutgers University. New Brunswick, New Jersey

GIS Watershed Delineation Tools

James A. Goodrich[1], Lucille Garner[1], Jill Neal[1], Lee Bice[2],
Rick Van Remortel[2], Ramon Olivero[3]

[1]USEPA, NRMRL, WSWRD, Cincinnati,
[2]Lockheed Martin, Las Vegas, NV
[3]Lockheed Martin, RTP, NC

BACKGROUND

The 1996 amendments to Section 1453 of the Safe Drinking Water Act require the states to establish and implement a Source Water Assessment Program (SWAP). Source water is the water taken from rivers, reservoirs, or wells for use as public drinking water. Source water assessment is intended to provide a strong basis for developing, implementing, and improving a state=s source water protection plan. This program requires individual states to delineate protection areas for drinking water intakes, identify and inventory significant contaminants in the protection areas, and determine the susceptibility of public water supply systems to the contaminants released within the protection areas. SWAP can be used to focus environmental public health programs developed by federal, state, and local governments, as well as efforts of public water utilities and citizens, into a hydrologically defined geographic area.

INTRODUCTION

The Environmental Protection Agency is assisting the states in conducting source water assessment by identifying potential sources of data and pointing to methods for assessing source waters. This presentation provides guidance to states, municipalities, and public water utilities for assessing source waters using geographic information system (GIS) technology. The GIS platforms used to organize, analyze, and manipulate available data and generate new data for source water protection areas, as well as provide capabilities for presenting the data to the public in various forms, including maps and tables are also discussed. In addition, the National Risk Management Research Laboratory is developing a coordinated research approach for watersheds that include contaminated sediments, urban watersheds, ecological restoration, and source water protection. Included in the full report, as appendices, are three case studies demonstrating the use of selected GIS-based software and hydrologic models to conduct hypothetical source water evaluations. Contamination of water supplies may be responsible for more human sickness than any other anthropogenic activity (Anderman and Martin, 1986). Since limited water resources are increasingly shared by competing consumers, there is a growing concern about the quality of source waters. This concern has led to the establishment of laws and programs designed to help protect drinking water sources. Frequent evaluation and identification of sources of contamination are required by federal and state rules. A successful SWAP reduces the cost of water treatments and disinfections required. Following enactment of the SDWA, a number of programs were developed for public water supply protection and supervision, including watershed protection and control, sanitary surveys, and WHPPs. An EPA document titled AStates Source Water Assessment and Protection Programs Final Guidance@ (1997a) discusses how a SWAP can use information provided by the current water programs. Some of the programs include:

- a watershed control program (WCP) under the Surface Water Treatment Rule (EPA, 1990),
- sanitary surveys (EPA, 1995), and
- wellhead protection programs (EPA, 1995).

METHOD FOR ASSESSING SOURCE WATERS

Using a GIS for any application involves following some basic steps including:

- designing the GIS database,
- building the GIS database,
- using the GIS to analyze the data and show results.

For assessing source waters, elements of the design of a GIS database include:

- establishing the study area,
- delineating the watershed,
- determining data needs,
- inventorying data sources,
- determining coordinate system and scale, and
- deciding on the GIS infrastructure.

Building the database requires collecting data to characterize the study area and inventorying sources of contamination. Analyzing the data entails assessing potential sources of contamination, delineating source water protection areas, and producing display products.

CHARACTERIZE THE STUDY AREA

After deciding on the data requirements of the GIS database, the data should be obtained and converted to the chosen projection and units (feet, meters). The data types include descriptions of physical watersheds and contamination sources and types. To understand how contamination from a source reaches a drinking water intake, the factors that affect its flow should be described. These factors include, but are not limited to terrain, soils, hydrography, land use and land cover, and contaminant characteristics. For example, after a precipitation event, the type(s) of contamination resulting from surface runoff into a stream depends on the land use and land cover interactions (e.g., pesticide and fertilizer from agriculture, salts and grease from parking lots). The directional flow of surface runoff depends on the topography, and soil infiltration properties affect how much surface water reaches the groundwater. The following sections provide information about some of the data sets needed for assessing source waters.

Watershed Boundaries

Watershed or HUC boundaries are available from the USGS. The HUC boundaries are available at 1:2,000,000 scale and 1:250,000 scale. The USGS also provides information describing the hydrologic unit coding scheme. A watershed boundary data set can be created by delineating the boundary on large-scale maps that have elevation contour lines; the boundary can then be digitized.

Terrain

Terrain data can be derived from Digital Elevation Models (DEMs). DEMs are digital records of terrain elevations for ground positions that are horizontally spaced at regular intervals. The SPOT Image Corporation provides DEMs at 10-meter spacing created by digital autocorrelation of SPOT satellite image stereopairs which are stored in a format known as Terrain Access Made Easy (TAME) (ESRI, 1992). The USGS also provides 30-meter spaced DEMs at four scales: 7.5-minute, 15-minute, 2-arc-second, and 1-degree. The 7.5-minute (large-scale) data are produced in 7.5- by 7.5-minute blocks from digitized cartographic map contour overlays or from scanned National Aerial Photography Program (NAPP) photographs. The DEM data are stored as profiles in which the elevations are spaced 30 meters apart. The number of elevations between each profile will differ because of the variable angle between the quadrangle's true north and the grid north of the Universal Transverse Mercator (UTM) projection coordinate system. The DEM data for 7.5-minute units correspond to the USGS 7.5-minute topographic quadrangle map series for all of the United States and its territories, except Alaska. The 15-minute DEM (large-scale) data correspond to the USGS 15-minute topographic quadrangle map series of Alaska. The unit size changes with the latitude. The 15-minute DEM data are referenced horizontally to NAD27. The elevations along profiles are spaced 2 arc-seconds of latitude by 3 arc-seconds of longitude. The first and last data points along a profile are at the integer degrees of latitude.

Soils

The U.S. Department of Agriculture (USDA) Natural Resource Conservation Service (NRCS), formerly the Soil Conservation Service (SCS), has three soil geographic databases of varying scales. The data include physical and chemical soil properties for approximately 18,000 soil types. Each database has three categories: soil properties (particle size, bulk density, available water capacity, organic matter, salinity, and soil recreation), locational properties (flooding, water table depth, bedrock depth, and soil subsidence), and use and management properties (sanitary facilities, building site development, recreational development, rangeland potential, construction material, crops, woodland suitability, and wildlife habitat suitability). The most detailed level of information is provided by the Soil Survey Geographic data (SSURGO), which is available in 7.5-minute topographic quadrangle units (1:24,000) and is distributed as coverages for soil survey areas, usually containing over ten quadrangle units. State Soil Geographic data (STATSGO) is a coarser database designed for regional, multi state, river basin, state, and multi county resource planning, monitoring, and management. The STATSGO database is at 1:250,000 scale (1- by 2-degree quadrangle) and is distributed as statewide coverages. National Soil Geographic data (NATSGO) is a database which is suitable for national or regional resources assessment and planning. With a scale of 1:5,000,000, the NATSGO database has information about the major land resource areas.

Hydrography

Hydrography is available from several federal sources at a 1:24,000 scale and may be available in greater detail from state and local government agencies. The USGS digital line graphs (DLGs) are readily available and provide information on 5 main types of data categories: boundaries, public land survey, transportation (including pipelines and power lines), hydrography (streams and water bodies) and hypsography (elevation contours). The DLG data can be converted into other formats compatible with GIS software. The EPA Reach File system has a series of hydrologic databases that uniquely identify and interconnect stream segments (reaches) for the nation. RF3-Alpha is the latest and most detailed version of the reach file system, containing more reaches than the previous versions, RF1 and RF2. Stream segments have unique reach codes for determining the upstream and downstream reaches and identifying the stream name for each reach. River Reach data can be obtained from the STORET User Assistance Group in the EPA Office of Water.

Land Use and Land Cover

Land use and land cover data are available from several federal sources. In many cases, the federal data will be either out-of-date or not detailed enough. More detailed (large-scale) land use data may also be obtained from county assessor maps, which are available at various scales (e.g., 1:200, 1:2,400, 1:4,800). County assessor maps may provide better detail for inventorying contamination sources in urban areas. The various departments of highways and transportation can provide maps for city streets and other local and regional road maps.

Inventory Potential Sources of Contamination

Potential sources of contamination, also known as sanitary defects, are conditions that may result in contamination of a water supply. These may be point and nonpoint source pollutants, connections to unsafe water supplies, raw water bypasses in treatment plants, improperly designed or installed plumbing fixtures, or water and sewer pipes leaking into the same ditch. All known and potential sources of contamination should be included in the GIS database. Pollutants may be classified into categories depending on the likelihood of their introduction into the water supply and the level and significance of contamination that can result from them. A contaminant inventory can include records of operation, discharge, disposal, construction, and other permitted activities, as well as zoning and health records obtained from local government agencies. All relevant information should be gathered while focusing the search for contamination sources at sites of particular concern. These include, but are not limited to (EPA, 1991a):

- discharge sites: septic tanks, irrigation pipes
- storage, treatment, or disposal sites: landfills, underground tanks, mine tailings
- substance transporting sites: pipelines
- activities that result in discharges: highway construction, fertilizer application
- natural sources impacted by anthropogenic activities

Further information on contaminant inventory activities is provided in the EPA *Guide for* Conducting Contamination Source Inventories for Public Drinking Water Supply Protection Programs (EPA, 1991a). Some of these data, such as Toxic Chemical Release Inventory (TRI) data can be obtained from the EPA. Other data may need to be obtained through field surveys. Table 1 lists federal data sets that can be accessed for much of the above mentioned data.

Table 1. Federal Spatial Data Set Sources

U. S. Environmental Protection Agency (EPA)

Web Page: http://www.epa.gov/enviro/html/nsdi/spatial_extent.html
Description: The EPA Envirofacts Warehouse - Geospatial Data Clearinghouse

Web Page: http://www.epa.gov/OWOW/watershed/landcover/lulcmap.html
Description: This EPA Office of Water site contains land cover digital data

Web Page: http://earth1.epa.gov/oppe/spatial.html
Description: This EPA Office of Policy, Planning and Evaluation site contains access to GIS spatial data sites at the federal, state, and local levels.

Web Page: http://www.epa.gov/OWOW/NPS/rf/rfindex.html
Description: This EPA Office of Water site contains information on the EPA river reach files.

U.S. Geological Survey

Web Page: http://nhd.fgdc.gov/
Description: U.S. Geological Survey site containing information on the Digital Line Graphs (DLG) hydrography files and the EPA Reach File Version 3.0 (RF3).

Web Page: http://edcwww.cr.usgs.gov/doc/edchome/ndcdb/ndcdb.html
Description: U.S. Geological Survey site containing FTP file access to Digital Elevation Models (DEM), Digital Line Graphs (DLGs), and Land Use and Land Cover (LULC).

Web Page: http://mcmcweb.er.usgs.gov/
Description: U.S. Geological Survey site for the Mid-Continent Mapping Center in Rolla, Missouri, containing information on Digital Raster Graphics (DRG) as well as other products.

Web Page: http://water.usgs.gov/public/GIS/background.html
Description: U.S. Geological Survey Water Resources site containing metadata and FTP file access to numerous national coverages commonly used in water resources studies.

Web Page: http://edcwww.cr.usgs.gov/webglis/
Description: The U.S. Geological Survey Global Land Information System (GLIS) site provides descriptions and prices for geospatial data available from the USGS.

U.S. Fish and Wildlife Service

Web Page: http://www.nwi.fws.gov/nwi.htmDescription:
Description: The U.S. Fish and Wildlife Service National Wetland Inventory (NWI) site provides access to NWI data.

U.S. Department of Agriculture

Web Page: http://www.ftw.nrcs.usda.gov/nsdi_node.html
Description: U.S. Department of Agriculture (USDA) National Resources Conservation Service site containing FTP access to soils and other USDA data.

U.S. Bureau of the Census

Web Page: http://www.census.gov/ftp/pub/mp/www/rom/msrom12i.html
Description: The U.S. Census Bureau site provides brief descriptions of the TIGER/Line files, 1997 version. The data is available for the entire U.S. on 6 CD-ROMs for $1,500 or $250/CD-ROM for different sections of the country. Data is in TIGER/Line format.

Contamination Source Risk Analysis

After the GIS database has been built, the data can be analyzed to assess the risk associated with potential sources of contamination, delineate protection areas, and develop display products.

Classify the contaminant data into risk groups depending on the threat of contamination they pose to the source water. A method for prioritizing and weighing the level of risk from various forms of contamination is described in an EPA document (EPA, 1991b). Similar approaches may be adopted for surface water sources. The tasks in this phase may reveal the need for a new inquiry or a more thorough data gathering effort with respect to particular sites or contaminants. For more information, see Managing Groundwater Contamination Sources in Wellhead Protection Areas: A Priority Setting Approach (EPA, 1991b). Susceptibility analysis identifies the location, frequency, and significance of potential contaminants in the source water protection area and determines the likelihood the PWS will be contaminated by these sources. Water quality models may be used for estimating contamination levels and determining the significance of selected contaminants in the protection area or in the watershed.

Proximity Analysis and Delineation of Protection Areas

After potential sources of contamination are identified, their proximity to the water supply intakes can be mapped. A set of maps at various scales can be produced from the GIS database illustrating the proximity of potential pollutants to the water supply system. With data documenting geographic locations of actual and potential contaminants, a source water protection area can be delineated. Surface water sources used for drinking water supplies may be protected by delineating a protection area around or upstream from the source intake. Three approaches for delineating a protection area for surface water systems are topographic area, buffer distance, and stream-flow time of travel (TOT) (EPA,1997b). For systems using groundwater sources, approaches for delineating a WHP area are based on fixed-radius, hydrogeologic/geomorphic characteristics, and modeling, which includes analytical, semi-analytical, numerical flow and solute transport models (EPA, 1993). The appropriate method for a particular system is chosen as a balance between ease of use, level of detail needed, and available resources. The PWS systems using a combination of groundwater and surface water sources may consider conjunctive delineation of source water protection areas. Conjunctive delineation is the integrated delineation of the zone of groundwater contribution and the area of surface water contribution to a PWS. Further information on this subject can be found in

Delineation of Source Water Protection Areas, a Discussion for Managers; Part 1: A Conjunctive Approach for Groundwater and Surface Water (EPA, 1997c).

Topographic Area

Topographic area is defined as the watershed for the surface water feature. Watersheds are delineated by drawing a line along the highest elevation around the surface water feature. In this case, the study area itself is the source water protection area.

Buffer Zone

A buffer zone may be delineated for the purpose of protecting drinking water intake and is typically dependent on the hydrogeology, topography, and stream hydrology. A protection buffer for a source surface water intake is an upstream strip of vegetated land along the shore of the stream or lake. Buffer widths vary from 15 to 60 meters (approximately 50 to 200 ft) depending on topographic, land use, political, and legal factors (EPA, 1997b). Buffer zones reduce water quality impacts from runoff, increase wildlife habitat and improve stream-bank integrity. Systems with groundwater sources may use a fixed-radius protection area (buffer) around source wells depending on aquifer properties. The radius could be fixed arbitrarily or based on TOT (EPA, 1993).

Time of Travel

Water supply systems tapping rivers that are designated for commercial transportation or for industrial and municipal wastewater discharge may use TOT for source water intake protection. The time it takes a pollutant introduced into an upstream section of a river to travel to a source water intake is estimated using the stream-flow TOT. The contamination level of the pollutant at the intake can be evaluated using various water quality models. The TOT method provides tools for predicting impacts from spills or discharges at sections upstream of a drinking water intake, thereby enhancing protection strategies for emergency spills. A TOT is also used for delineating protection areas for groundwater-based systems by estimating contaminant transport into drinking water wells. Groundwater flow is significantly slower than that of surface water (e.g., years versus hours or days, respectively), allowing more response time for controlling or remediating spills and other plumes. The EPA (1993) provides comparisons of TOT-based methods used for delineating WHP areas.

Modeling

Surface runoff and groundwater models can be used for delineating a source water protection area. Analytical, semi-analytical, and numerical flow and solute transport models can estimate the potential water quality impacts from one or more pollution sources upstream of a drinking water intake. With knowledge of land uses (e.g., agricultural, industrial, residential), soil properties, and precipitation rates in an area, potential contaminant loadings from runoff or infiltration can be estimated. Modeling provides analytical tools for assessing water quality impacts resulting from land use changes, and may be used to identify effective water quality protection strategies. Some models need site-specific data which may, in turn, require field surveys.

Stream Network Analysis: Water Quality Study

Stream network analysis provides tools for studying how contaminants are transported in streams. Distributions of contaminate concentrations along a stream can be studied using the physical and chemical properties of the contaminant as well as the hydraulics of the stream. Most GIS software packages, such as ARC/INFO=s network analysis, have capability for modeling linear processes. More complex analyses can be performed by linking appropriate water quality models in ARC/INFO (e.g., Grayman et al.,1993).

Generate Display Products

Maps are graphic representations of geographic information, and, as such, provide powerful visual communication of ideas. The Surface Water Assessment Program requires strong public participation in all processes involving development of methods for, and implementation of, source water assessment. State agencies proposing or conducting a SWAP may use sets of maps for displaying the geographic extent of the SWAP program. For example, maps for public presentation can show stream segments with highlighted buffer areas and marked with potential pollution sites. A GIS provides the capability for generating such maps at various scales with selected sets of themes.

HARDWARE

The GIS hardware includes the computer on which the GIS operates and the peripherals used for data entry, transfer, and output. A wide range of hardware types are used, from centralized computer servers to desktop computers used as stand-alone stations or in networked configurations. The type and number of components in a system is dependent on the needs of the organization. Software vendors can help in recommending appropriate system configurations. The input and output devices (e.g., digitizers and plotters) are usually shared within an organization with more than one GIS user. Centralized computer servers and networking software can be used to enable multiple users to share GIS hardware and software. Hardware costs are not provided because costs are constantly changing, usually in favor of the buyer. Examples of GIS hardware components are listed below.

GIS Workstation

A GIS workstation should at a minimum include a high-speed central processing unit, keyboard, mouse, disk space, high-resolution color monitor for graphics display, and a compact disk read-only memory (CD-ROM). An external disk drive may be used for additional disk space. The GIS workstation can be either an IBM-compatible personal computer (PC) using a Windows operating system or a high-end graphics workstation using a Unix operating system. The Unix systems provide a more powerful environment for GIS than PCs. Unix workstations are usually faster than PCs in the analysis and display of complex digital data. However, they also cost more ($ 15,000+ vs. < $ 10,000 for a PC). A set of workstations loaded with GIS software may use a common server with a large amount of disk storage space. Also, data input and output devices may be attached to the server so all users can share them.

Data Transfer and Backup Devices

A GIS should include one or more data transfer and backup devices such as a compact disk writer, tape drive, or disk drive. These devices allow the user to transfer GIS data to a compact medium that can be easily stored or physically transferred. These devices are useful for performing data backups or transferring data between workstations or organizations that are not networked.

Data Output Devices

Output devices allow the user to print data and displays from the GIS. Printouts of GIS data are useful for data quality assurance and quality control (QA/QC) checks and for displaying results. The common GIS output devices are printers and plotters. These devices are available in a variety of sizes, produce output in color or black and white, and can vary widely in price. Most organizations will want at least a standard laser jet printer as well as a large-format color output device for plotting color maps for display and presentations.

Data Input Devices

GIS data input devices include digitizing tables, scanners, and GPS receivers. These devices enable a user to capture geographic information in digital form. A digitizing table is used for generating vector-based coordinate information directly from hard copy maps or photographs. A scanner is used to generate raster-based data from hard copy maps or photographs. A GPS receiver enables the user to capture coordinate data for features in the field. Once captured, GPS data must be post-processed on a workstation with specialized software to generate real-world coordinates.

SOFTWARE

Three categories of information processing software are used to assess source waters when using GIS technology: GIS, image processing, and relational database management. Examples of software for each of these categories are listed below. A software package listed in one category may also be capable of performing functions in another category. For example, a GIS package such as GRASS can be used for image processing. Similarly, some of the image processing software packages can be used as GIS tools. The names of the software are listed for informational purposes only and do not indicate endorsement. The PC-based software packages such as GRASS and ArcView can range in cost from free or low-cost ($200-$300) to several thousand dollars. High-end software packages such as ArcInfo, ERDAS Imagine, or Intergraph GIS will cost $10,000-$20,000. Prices for all software packages depend on current market value, whether the purchaser is eligible for discounts, and what additional modules are purchased in addition to the baseline package.

GIS Software

The GIS software is used for storing, analyzing, and displaying geographic data. The main components of a GIS software are the tools for data input and manipulation, database management, geographic query and analysis, and visualization and output. Several GIS packages are presented below for information.

Arc/Info

Arc/Info is a commercial software package developed by the Environmental Systems Research Institute (ESRI) and Henco Software, Inc. (Henco). Arc/Info provides tools for automation, management, display, and output of geographic and associated data. Arc/Info is a vector-based

GIS software that runs on Unix and Windows NT workstations. Arc/Info costs between $10,000 - $20,000. For more information contact ESRI at http://www.esri.com.

ArcView

ArcView is also produced by ESRI and is a menu-driven GIS with a subset of the functionality provided by Arc/Info. What ArcView lacks in functionality, it makes up for in a less steep learning curve and an easy-to-use graphical user interface (GUI). ArcView is a vector-based GIS software that runs on Unix or PC workstations. ArcView costs approximately $1,000. For more information contact ESRI at http://www.esri.com.

GRASS

The Geographic Resources Analysis Support System (GRASS) is a public-domain, raster-based GIS software used for geographic data management, image processing, graphics production, spatial modeling, and data visualization. GRASS was written by the U.S. Army Construction Engineering Research Laboratories (USA-CERL) branch of the U.S. Army Corps of Engineers and is currently maintained at the Department of Geology at Baylor University. GRASS runs on Unix and PC workstations. More information on GRASS can be found at http://www.baylor.edu/~grass. Additional information on some of the hydrology models that have been integrated into the GRASS GIS is available on http://soils.ecn.purdue.edu/~aggrass/models/ hydrology.html.

IDRISI

IDRISI is a raster-based GIS software that provides GIS, image processing, and spatial statistics analytical capabilities on DOS and Windows-based PCs. IDRISI provides analytical functionality of GIS, remote sensing, and databases for resources management. IDRISI was developed and is maintained by Clark Labs, a non-profit research organization within the Graduate School of Geography at Clark University. A commercial/private single-user license for IDRISI costs $990. Licenses for non-profit, government, and academic institutions cost less. For more details see http://www.clarklabs.org.

Intergraph GIS

Intergraph provides Windows-based software and a range of computing services for engineering, design, modeling, analysis, mapping, information technology, and creative graphics. The GIS MGE package provides data collection and editing, data import, image display and analysis, advanced spatial query and analysis, and cartographic quality maps. MGE costs approximately $10,000-$20,000. More information on Integraph GIS is available at http://www.intergraph.com.

Image Processing Software

Image processing software is used to process raster data, particularly remote sensing imagery data such as satellite imagery.

ENVI

The Environment for Visualizing Images (ENVI) is an image processing system which provides analysis and visualization of single-band, multispectral, hyperspectral, and radar remote sensing data. ENVI can process large spatial and spectral images, and runs on Unix; LINUX; Windows 3.1, NT, 95; the Macintosh; and the Power Mac. For more details contact ENVI at http://www.envi-sw.com/index.htm.

ERDAS Imagine

The ERDAS Imagine software is an image processing and raster GIS package that has a variety of applications ranging from simple image mapping to advanced remote sensing. Imagine provides tools for geometric correction, image analysis, visualization, map output, orthorectification, radar analysis, advanced classification tools, and graphical spatial data modeling. Imagine runs on Unix workstations and Windows platforms. ERDAS Image costs approximately $10,000-$20,000. More information on Imagine is available at http://www.erdas.com.

ER Mapper

ER Mapper provides integrated mapping software featuring image processing, map production, 3-D presentations, and GIS integration for Windows 95/NT and Unix. The ER Mapper software uses a concept that separates data from the image processing steps allowing the user to apply and view results from a single enhancement procedure in real time. The PC version of ER Mapper costs $4,300; the Unix version of ER Mapper costs $18,300. See http://www.ermapper.com for more information.

PCI

EASI/PACE image processing provides a variety of applications including image processing, geometric correction, vector utilities, and multilayer modeling. PCI implements the Generic Database (GDB) concept, which allows PCI programs to access image and other external data files without import and export. Contact PCI for more details at http://www.pci.on.ca.

TNTmips

TNTmips is a map and image processing system that contains fully featured GIS, CAD, and spatial database management systems. TNTmips has tools that interactively integrate elements of on-screen image processing and photo interpretation, and provides a diverse set of tools for registering, rectifying and stitching imagery, which are particularly useful for low-altitude aerial photography and videography. More information on TNTmips is available at http://www.sgi.com/Products/appsdirectory.dir/Applications/GIS_Defense_Imaging/ApplicationNumber7857.html.

Relational Database Management Software

Relational database management system (RDBMS) software enables large amounts of data to be entered, updated, related, viewed, queried and, otherwise, managed in an efficient manner. The data in an RDBMS is stored in a series of related tables which are designed to optimize the effort required for data entry, maintenance, and retrieval. RDBMS software is available for use on PCs, Unix workstations, networked systems, and mainframe computers. Most GIS software packages use an RDBMS to manage data such as maintaining topology and providing ways to efficiently enter, update, and query attribute data. Major RDBMS software includes Info, dBASE, MS Access, Ingres, Informix, Oracle, and Sybase.

SOFTWARE SUPPORT TOOLS

There are numerous software support tools available for use in assessing source waters. These tools operate within specific operating and software system environments. The information presented here is not an endorsement of any of these products. New products and improvements to existing products are continuously being introduced; therefore, users should conduct their own investigation of software tools to ensure they are getting the latest information. The selections are considered some of the more promising and potentially useful that were encountered during this GIS evaluation. It should be noted that there are hundreds of available hydrologic models described in the scientific literature, but many of these will probably not be suitable for use in a source water assessment. Principal purveyors of other downloadable software and hydrologic models not listed here include the EPA Center for Exposure Assessment Modeling (http://ftp.epa.gov/epa-ceam/wwwhtml/softwdos.htm), the USGS Water Resources Division (http://water.usgs.gov/software/), and the U.S. Army Corps of Engineers= Hydrologic Engineering Center (http://www.hec.usace.army.mil/). A selection of the software support systems to consider include:

- **Better Assessment Science Integrating Point and Nonpoint Sources** (BASINS) from the EPA Office of Science and Technology (OST). Full documentation of BASINS Version 2.0 is available in detail at http://www.epa.gov/ostwater/BASINS/***.***

- Riverine Emergency Management Model (REMM) was originally developed for hydrologic modeling in the upper Mississippi River in Minnesota. REMM is public-domain software and is freely available by the U.S. Army Corps of Engineers office in St. Paul, Minnesota (e-mail: webmaster@mvp-wc.usace.army.mil).

- Watershed Modeling System (WMS) Model was developed by the Environmental Modeling Research Laboratory of Brigham Young University in cooperation with the U.S. Army Corps of Engineers Waterways Experiment Station. The WMS is proprietary software and is available via the Engineering Computer Graphics Laboratory at Brigham Young University in Provo, Utah (http://www.ecgl.byu.edu). The software cost ranges from $500 to $2,600 depending on the desired modules.

- Underground Storage Tank (UST)-Access Software was developed using the Microsoft Access 2.0 relational data base management system. All UST-Access installation files are stored as self-executable archive files on the Cleanup Information (CLU-IN) Bulletin Board System of the EPA Office of Solid Waste and Emergency Response.

- Spatially Referenced Regressions on Watersheds (SPARROW) Model is an extension from the Hydrologic Simulation Program - Fortran (HMPF) modeling framework. The HMPF and SPARROW models are public-domain software freely available through USGS (http://www.usgs.gov) for the cost of pressing a CD (about $35).

- Hydrology Extension for the ArcView Spatial Analyst Software is a new Hydrology Extension for ArcView=s Spatial Analyst 1.1. ArcView Spatial Analyst 1.1 is proprietary software distributed by ESRI (http://www.esri.com). Price varies widely depending on the user=s affiliation, such as with government or industry.

- MassGIS Watershed Tools for the ArcView Spatial Analyst Software was developed by the State of Massachusetts, Division of Watershed Management, GIS Division (MassGIS) for use with the ArcView Spatial Analyst. ArcView Spatial Analyst is proprietary software distributed by ESRI (internet: www.esri.com). Price varies widely depending on the user=s affiliation with government or industry. The MassGIS watershed tools are public-domain software (john.rader@state.ma.us).

PERSONNEL REQUIREMENTS

To use a GIS effectively in any project, it is important to have personnel with a variety of specific skills. All of the software mentioned above (GIS, image processing, and RDBMS) require lengthy learning curves to be used effectively.

Data Entry Technician

Data entry includes automation or digitizing of maps, creating attribute tables, and importing databases. The data entry technician should have some knowledge of spatial concepts and experience in basic GIS use for creating thematic layers, and attribute data entry. Depending on the amount of data entry required, one or more technicians may be needed.

Spatial Data Analyst

The spatial data analyst is skilled in manipulating geographic data to retrieve pertinent, project-specific information such as mapping sources of contamination and their proximity to source waters, and delineating protection areas. This person must have a thorough understanding of the concepts presented in this Chapter and be experienced in using GIS and image processing technology. The spatial data analyst should also have some experience in working with utilities, hydrogeology, soils, environmental engineering, or sanitary engineering.

Field Surveyor

A field surveyor may be required if geographic or attribute data is not available and must be gathered in the field. The surveyor should be skilled in field survey management, GPS technology, and database development and have knowledge of sanitary or environmental engineering, soil science, or hydrogeology. Depending on the amount of field surveying required and the size of the area being surveyed, the field surveyor may require a support staff to assist with gathering information.

Soil Scientist

A soil scientist may be needed to evaluate the condition and physical properties of soils in the survey area. The Natural Resources Conservation Service (NRCS) formerly called the Soil Conservation Service may be contacted for technical assistance in this area.

System Administrator

A system administrator may be needed to administer the GIS and its peripherals such as digitizers, printers, and plotters. This is especially true for systems that require a network and have multiple users. A system administrator can help with hardware and software maintenance and replacement, network maintenance, system backups, and other administrative duties.

REFERENCES

Anderman, W.H. and G.Martin. 1986. Effect of public sewers on watershed contamination, *Journal of Environmental Health* 4(2):81-84.

EPA. 1997a. States source water assessment and protection programs final guidance. EPA 816-R-97-009. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

EPA. 1997b. State methods for delineating source water protection areas for surface water supplied sources of drinking water. EPA 816-R-97-008. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

EPA. 1997c. Delineation of Source Water Protection Areas, a Discussion for Managers; Part 1: A Conjunctive Approach for Ground Water and Surface Water, U.S. Environmental Protection Agency, Office of Water, Washington, DC. (Expected August 1997)

EPA. 1995. EPA/State Joint Guidance on Sanitary Surveys. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

EPA. 1994. The quality of our nation=s waters. EPA 841-S-95-004. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

EPA. 1993. Guidelines for delineation of wellhead protection areas. EPA 4405-93-001. U.S. Environmental Protection Agency, Office of Water, Office of Groundwater Protection, Washington, DC.

EPA. 1991a. Guide for Conducting Contamination Source Inventories for Public Drinking Water Supply Protection Programs. EPA 570/9-91-033. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

EPA. 1991b. Managing Groundwater Contamination Sources in Wellhead Protection Areas: A Priority Setting Approach. EPA 570/9-91-023. U.S. Environmental Protection Agency, Office of Groundwater and Drinking Water.

EPA. 1990. Guidance manual for compliance with the filtration and disinfection requirements for public water systems using surface water sources, U.S. Environmental Protection Agency, Office of Drinking Water, Washington, D.C.

ESRI, Inc. 1992. Understanding GIS: The ARC/INFO Way. Environmental Systems Research Institute, Redlands, CA.

Grayman, W.M., S.R. Kshirsagar, and R.M. Males. 1993. A Geographic Information System for the Ohio River Basin, Risk Reduction Engineering Laboratory, Office of Research and Development, US Environmental Protection Agency, Cincinnati, Ohio.

A GIS for the Ohio River Basin

Walter M. Grayman
W.M. Grayman Consulting Engineers, Cincinnati, Ohio

Sudhir R. Kshirsagar
Global Quality Corporation, Cincinnati, Ohio

Richard M. Males
RMM Technical Services, Inc., Cincinnati, Ohio

James A. Goodrich
Risk Reduction Engineering Laboratory, Office of Research and Development,
U.S. Environmental Protection Agency, Cincinnati, Ohio

Jason P. Heath
Ohio River Valley Water Sanitation Commission, Cincinnati, Ohio

Abstract

Much of the information used in the management of water quality in a river basin has a geographic or spatial component associated with it. As a result, spatially based computer models and database systems can be part of an effective water quality management and evaluation process. The Ohio River Valley Water Sanitation Commission (ORSANCO) is an interstate water pollution control agency serving the Ohio River and its eight member states. The U.S. Environmental Protection Agency (EPA) entered into a cooperative agreement with ORSANCO to develop and apply spatially based computer models and database systems in the Ohio River basin.

Three computer-based technologies have been applied and integrated: geographic information systems (GIS), water quality/hydraulic modeling, and database management.

GIS serves as a mechanism for storing, using, and displaying spatial data. The ARC/INFO GIS, EPA's agencywide standard, was used in the study, which assembled databases of land and stream information for the Ohio River basin. GIS represented streams in hydrologic catalog units along the Ohio River mainstem using EPA's new, detailed RF3-level Reach File System. The full Ohio River basin was represented using the less detailed RF1-level reach file. Modeling provides a way to examine the impacts of human-induced and natural events within the basin and to explore alternative strategies for mitigating these events.

Hydraulic information from the U.S. Army Corps of Engineers' FLOWSED model enabled EPA's WASP4 water quality model to be embedded in a menu-driven spill management system to facilitate modeling of the Ohio River mainstem under emergency spill conditions. A steady-state water quality modeling component was also developed under the ARC/INFO GIS to trace the movement and degradation of pollutants through any reaches in the RF1 representation of the full Ohio River basin.

Database management technology relates to the storage, analysis, and display of data. A detailed database of information on dischargers to the Ohio River mainstem was assembled under the PARADOX database management system using EPA's permit compliance system as the primary data source. Though these three technologies have been widely used in the field of water quality management, integration of these tools into a holistic mechanism provided the primary challenge of this study.

EPA's Risk Reduction Engineering Laboratory in Cincinnati, Ohio, developed this project summary to announce key findings of the research project, which is fully documented in a separate report of the same title.

Introduction

During the past 25 years, computers have been actively used in water quality management, demonstrating their potential to assist in a wide range of analysis and display tasks. Technologies such as geographic information systems (GIS), database management systems (DBMS), and mathematical modeling have been applied in the water quality management field and have proven to be effective tools. For computers to achieve their full potential, however, they must become integrated into the normal programmatic efforts of agencies and organizations in the planning, regulation, and operational areas of water quality management.

Recognizing this need for routine use of computer-based tools, the Ohio River Valley Water Sanitation Commission (ORSANCO) and the Risk Reduction Engineering Laboratory (RREL) of the U.S. Environmental Protection Agency (EPA) commenced a study in 1990. The goals of the study included the adaptation, development, and application of modeling and spatial database management (DBM) tools that could assist ORSANCO in its prescribed water quality management objectives. These goals were consistent with EPA's ongoing programs involving the use of GIS and modeling technology. The study's goals also coincided with EPA's Drinking Water Research Division's work over the past decade, which applied similar technology to study the vulnerability of water supplies on the Ohio and Mississippi Rivers to upstream discharges.

Methodology Overview

To address the goals of this project, three basic technologies have been applied and integrated: GIS, water quality/hydraulic modeling, and DBM. GIS serves as a mechanism for storing, using, and displaying spatial data. Modeling provides a way to examine the impacts of human-induced and natural events within the basin and to explore alternative strategies for mitigating these events. DBM technology relates to the storage, analysis, and display of data. Though these three technologies have been widely used in the field of water quality management, integration of these tools into a holistic mechanism provided the primary challenge of this study.

GIS Technology

The guiding principle in developing the GIS capability was to maximize the use of existing GIS technology and spatial databases. The study used ARC/INFO GIS, EPA's agencywide standard. Remote access of ARC/INFO on a VAX minicomputer facilitated the initial work. Subsequently, both PC ARC/INFO and a workstation-based system were obtained.

EPA has developed an extensive spatial database related to water quality and demographic parameters. This served as the primary source of spatial data for the study. Following is a summary of spatial data used in this study:

- State and county boundaries.
- City locations and characteristics.
- Water supply locations and characteristics.
- Locations and characteristics of dischargers to water bodies.
- Toxic loadings to air, water, and land.
- Dam locations and characteristics.
- Stream reaches and characteristics.

The primary organizing concept for the water-related information was EPA's Reach File System (1). This system provides a common mechanism within EPA and other agencies for identifying surface water segments, relating water resources data, and traversing the nation's surface water in hydrologic order within a computer environment. A hierarchical hydrologic code uniquely identifies each reach. Information available on each reach includes topological identification of adjacent reaches, characteristic information such as length and stream name, and stream flow and velocity estimates. The original reach file (designated as RF1) was developed in the early 1980s and included approximately 70,000 reaches nationwide. The most recent version (RF3) includes over 3,000,000 reaches nationwide.

As part of this project, an RF1-level database was established for the entire Ohio River basin. The RF3 reach file was implemented for the Ohio River mainstem and lower portions of tributaries. River miles along the Ohio River were digitized and established as an ARC/INFO coverage to provide a linkage between the reach file and river mile indexing used by ORSANCO and other agencies along the river. Figure 1 shows the RF1 reach file representation of the Ohio River basin along with state boundaries.

The study incorporated several EPA sources of information on dischargers to water bodies. The industrial facility discharger (IFD) file contains locational and characteristic data for National Pollutant Discharge Elimination System (NPDES) permitted discharges. Detailed permit limits and monitoring information was accessed from the permit compliance system (PCS). The toxic release inventory (TRI) system includes annual loading of selected chemicals to water, land, air, and sewer for selected industries based on quantity discharged. All water data are referenced to the NPDES permit number, which is spatially located by reach and river mile, and by latitude and longitude.

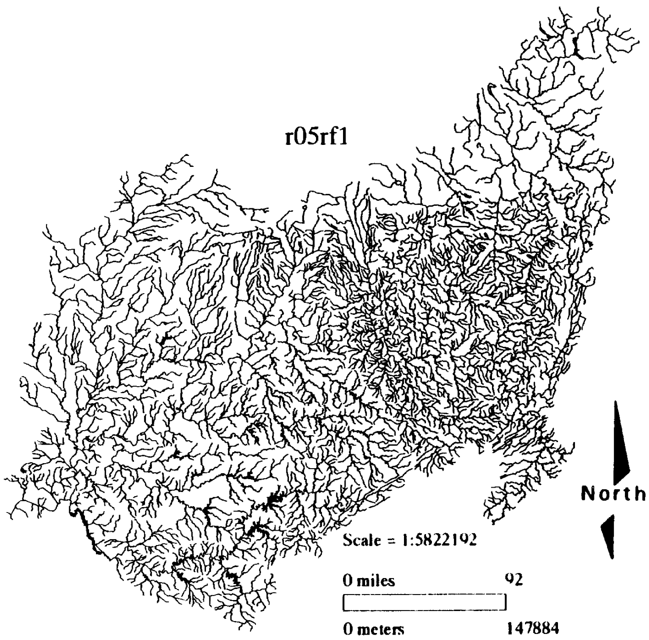


Figure 1. RF1 reaches in the Ohio River basin.

Spill Modeling

An important role that ORSANCO fills on the Ohio River relates to the monitoring and prediction of the fate of pollutant spills. Typically, ORSANCO serves as the overall communications link between states during such emergency conditions. ORSANCO coordinates and participates in monitoring and serves as the information center in gathering data and issuing predictions about the movement of spills in the river. In the past, a series of time-of-travel nomographs, based on National Weather Service flow forecasts, Corps of Engineers flow-velocity relationships, and previous experience, were used to predict the movement of spills. This project combined a hydraulic model with a water quality model to serve as a more robust method for making such predictions.

The U.S. Army Corps of Engineers' FLOWSED model was selected as the means of predicting daily flow quantities and water levels along the mainstem and portions of major tributaries near their confluence with the Ohio River (2). The Ohio River Division of the Corps of Engineers applies FLOWSED daily as part of its reservoir operations program. The Corps can generate 5-day forecasts of stage and flow for 400 mainstem and tributary segments, and ORSANCO can access the results via phone lines.

EPA's WASP4 water quality model was selected for use in the project (3). WASP4 is a dynamic compartment model that can be used to analyze a variety of water quality problems in a diverse set of water bodies. Because the primary use of the model in this project is quick response under emergency situations, only the toxic chemical portion of the model with first order decay is being used. The FLOWSED and WASP4 models have been combined into a user-friendly spatial decision support system framework described later in this project summary.

Discharger Database Management System

EPA's PCS and historical records maintained by ORSANCO furnish a rich source of data on discharge information for the Ohio River. To organize these data and make them available for analysis, a database was developed using the PARADOX DBM system.

The database was established using a relational structure with a series of related tables (two-dimensional flat files). Individual tables contain information on facilities, outfalls, permit limits, monitoring data, and codes used in the other tables. The NPDES permit number is used as the primary key in each data table. A mechanism for downloading and reformatting data from the national PCS database has been developed along with custom forms for viewing and editing data, and custom reports for preparing hard copy summaries. Latitude and longitude values for each facility can provide the locational mechanism for use of this data in conjunction with GIS.

Integration of GIS/Modeling/Database Technologies

A major objective of this study was the integration of GIS, modeling, and DBMS technologies into a holistic tool for use by ORSANCO. Several integration mechanisms were implemented as summarized below.

Steady-State Spill Tracing

The NETWORK component of the ARC/INFO GIS provides a steady-state, transportation-oriented routing capability. This capability was used in an arc macro language (AML) program to construct a routing procedure for determining downstream concentrations and travel times. The pollutant may be treated as a conservative element or represented by a first order exponential decay function. This capability has been implemented for use with the RF1 reach file representation of the full Ohio River basin. The user may select from six flow regimens: average flow, low flow, and four multiples of average flow ranging from one-tenth to 10 times average flow. This system gives ORSANCO the ability to estimate the arrival time of a spill from any RF1 tributary to the Ohio River mainstem.

Spill Management System

A PC-based spatial decision support system (SDSS) was built as a spill management system to be a quick response tool for analyzing and displaying the results of pollutant spills into the Ohio River. The schematic in Figure 2 illustrates the components in this computerized spill management system. The system is implemented in the C language using a commercial menuing system and a series of graphic display routines developed at EPA. Custom, written routines have been used to read the output from the U.S. Army Corps of Engineers' FLOWSED model, to generate input files for EPA's WASP4 model, to create output reports and output plots, and to provide an animated representation of the concentration profiles moving down the river. Figure 3 presents an example of a graphic output the system generated. Additionally, the system generates a file in

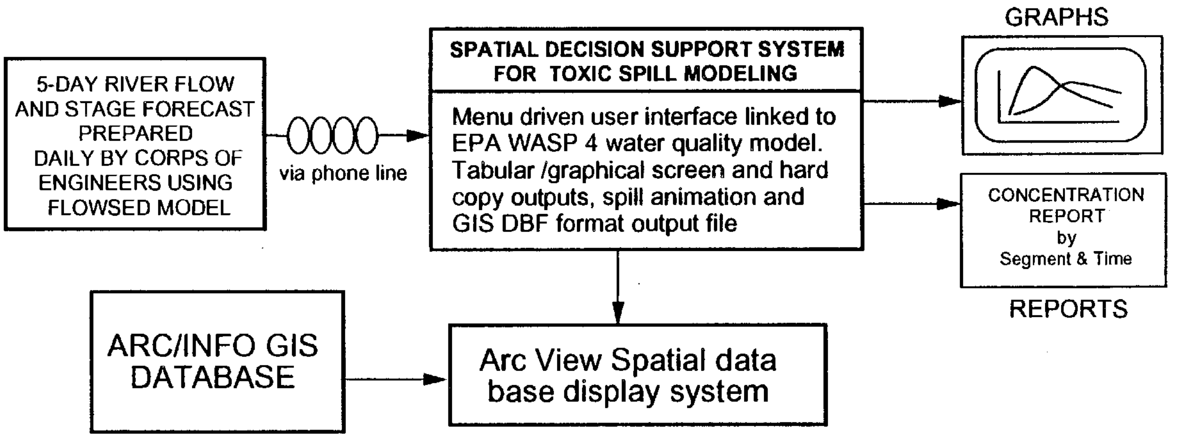


Figure 2. Schematic representation of spill modeling system process.

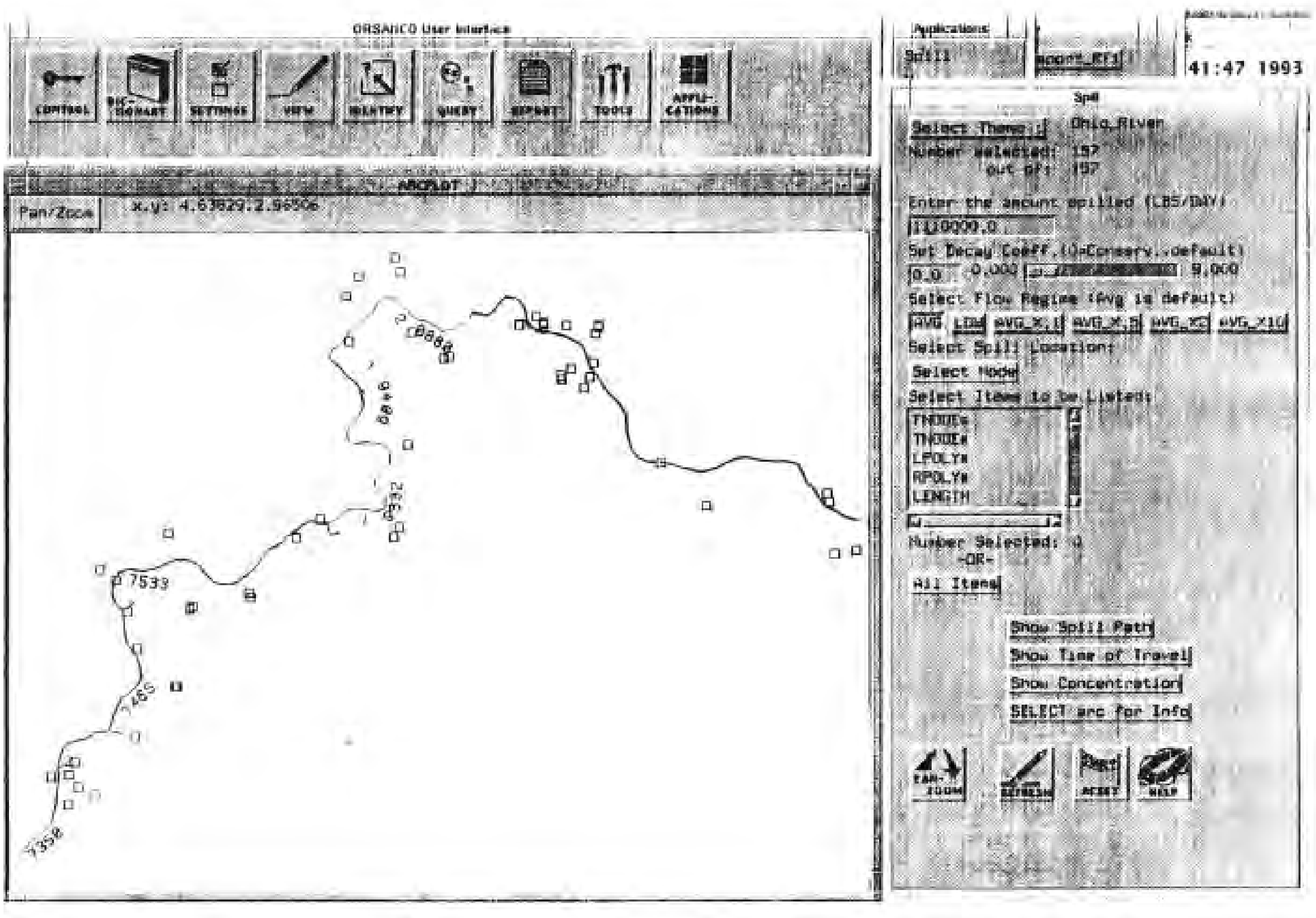


Figure 3. Graphic output from the basinwide network spill model.

DBF format that may be read by ARC/VIEW (the companion software to ARC/INFO for user-friendly viewing of spatial data).

Hardware Platform

Within the study, the initial hardware platform was a combination of local PCs (in Cincinnati) and a remote access terminal to a VAX computer located at EPA's National Computer Center in Research Triangle Park, North Carolina. The final platform, and the one on which the completed system was installed, comprised a UNIX-based Data General workstation and a PC workstation. The full hardware configuration is shown schematically in Figure 4.

Conclusions

The application of computer-based display, analysis, and modeling tools in conjunction with GIS technology proved to be an effective strategy for water quality management. This study used an existing GIS package and DBMS in conjunction with existing water quality and hydraulic models. The study focused primarily on assembling available spatial and relational databases and integrating the systems to provide a usable, effective tool.

References

1. Horn, R.C., and W.M. Grayman. 1993. Water-quality modeling with EPA reach file system. J. Water Res. Planning and Mgmt. 119(2):262-274.
2. Johnson, B.H. 1982. Development of a numerical modeling capability for the computation of unsteady flow on the Ohio River and its major tributaries. Vicksburg, MS: U.S. Army Engineer WES.
3. Ambrose, R.B., T.A. Wool, J.P. Connolly, and R.W. Schanz. 1988. WASP4, a hydrodynamic and water quality model: Model theory, users' manual, and programmers' guide. Athens, GA: EPA Environmental Research Laboratory. NTIS PB88185095.

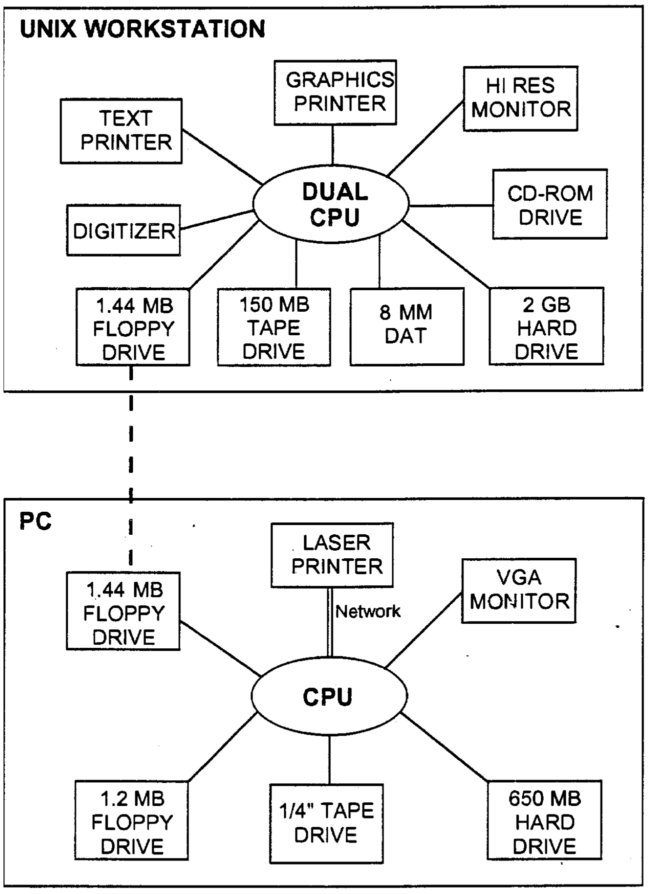


Figure 4. Hardware configuration.

Data Quality Issues Affecting GIS Use for Environmental Problem-Solving

Carol B. Griffin
Henry's Fork Foundation, Island Park, Idaho

"Abandon hope, all ye who enter here." Dante's quote might well be the advice that experienced geographic information system (GIS) users give to nonusers about to confront data quality issues associated with GIS use. Indeed, after reading this paper, some decision-makers might abandon attempts to use a GIS because of the error associated with it. Others may want to spend an inordinate amount of time and money trying to eliminate all error associated with GIS use. Neither option is prudent.

Data quality is important because it affects how reliable GIS-generated information is in the decision-making process. Too often, the availability of inexpensive digital data overshadows data quality concerns; people frequently use digital data because they are available, not because they have the necessary accuracy.

A GIS can help decision-makers use spatial information more fully than manual methods allow, but sometimes data quality issues cause concern about using GIS-generated outputs. Making environmental decisions without adequate consideration to data quality may lead to an erroneous decision, erode public confidence, or cause an agency to incur liability. This paper attempts to encourage decision-makers to become more aware of data quality issues, including the sources and magnitude of error.

GIS error research has necessarily progressed in a linear fashion, beginning with identifying and classifying sources of error. This paper discusses both inherent (source) error and the error that GIS operations introduce (operational error) during data input, storage, analysis/manipulation, and output (1). Strategies for coping with error and research into error reduction techniques have only recently received attention. Unfortunately, the answers to error management questions such as, "How will the error affect decision-making?" are not clear. The end of this paper covers several error management suggestions and anticipated software improvements designed to reduce errors, however.

Data Quality Concepts and Their Importance

Data quality is a major issue for GIS-generated maps, much more so than it is for paper maps. In part, this is because a GIS can perform operations on spatial data that would be nearly impossible without a GIS because of scale, complexity, and generalization issues (2). Cartographers adjust for these problems when they manually manipulate and instantly combine paper maps by adhering to long-standing cartographic principles, but GIS personnel may not be fully trained in these principles. A GIS enables an analyst, whether trained in cartographic principles or not, to combine or manipulate data in appropriate or in inappropriate, illogical, and erroneous ways. Lack of training coupled with the speed of spatial data manipulation can have serious consequences for an agency whose personnel produce and use GIS-generated maps.

Limited scientific understanding, limited ability to measure data, sampling error, inherent variability, and inadequacy of mathematical representations all contribute to uncertainties associated with spatial data. Uncertainty about spatial data consists of two parts: ignorance and variability. Ignorance means that variables have a "true" value, but it is unknown to us, whereas variability means one value cannot represent the variables.

Data quality defies a simple definition. For this paper, data quality can roughly mean how "good" the data are for a given purpose. People usually think of data quality in terms of error, but the term is broader and encompasses the six components outlined in the next section. Error can mean the difference between the observed values and the "true" value. The "true" value of a variable is usually unknown and unknowable, but for this paper's purposes, "true" could be the known value or the value one would obtain from field measurements (the discussion of data collection tries to dispel the notion that there is one "true" value for many variables, such as soil type in a given area or water temperature in a lake). Imperfect equipment or observers and environmental effects cause spatial error. According to Thapa and Bossler (3), errors fall into three categories:

- Gross errors and blunders (people or equipment).
- Systematic errors (which introduce bias).
- Random errors (due to imperfect instruments and observers).

In addition, another view divides spatial error into two different components: accuracy and precision. Accuracy means how close a value is to the “true” value or a known standard (absence of bias). Precision can have two definitions: it can be a measure of dispersion (standard deviation) of observations about a mean, or it can refer to the number of decimal digits used to represent a value (4). In the first definition of precision, a measurement of 6 feet plus or minus 1 foot is more precise than one of 6 feet plus or minus 3 feet. In the second definition, a value of 6.1794 feet is more precise than one of 6.1 feet. Figure 1 provides a graphic explanation of the difference between error, accuracy, and precision.

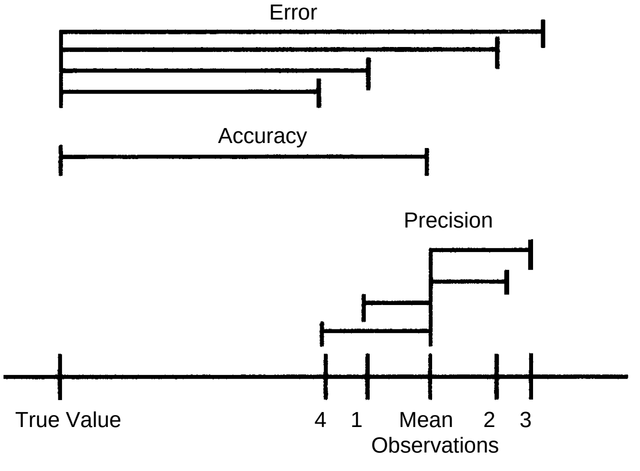


Figure 1. Relationship between error, accuracy, and precision.

Data are not accurate or inaccurate. Instead, data accuracy exists on a continuum, ranging from low to high accuracy. Although people strive for accurate (error-free) data, obtaining 100-percent accurate data is impractical. The list below provides some of the reasons why total accuracy is not obtainable (5):

- Objects to be measured are often vaguely defined.
- Some phenomena are variable in nature.
- Classification schemes are imprecise.
- Measurements are inherently imprecise.
- Gross errors of a nonstatistical nature can occur during measurement.
- Attributes encoded on an ordinal scale (high, medium, low) are approximate.
- Data represent a past state of reality.

Users of geographic data should strive for data that are only as accurate as they need. A variety of factors, of course, can determine need:

- Intended use of the data
- Budget constraints
- Time constraints
- Data storage considerations
- Potential liability

The main barrier to highly accurate data is lack of funds. Nale (6) suggests that rather than abandoning a GIS project because funds are not sufficient to achieve the desired accuracy, an agency should collect data at the desired accuracy from smaller areas, such as areas being developed or redeveloped. Over time, data collection at the desired accuracy can expand to include areas that lacked data due to budgetary constraints. Smith and Honeycutt (7) outline the use of a value of information approach in determining the need for more data (or more accurate data) based on the expected costs and benefits associated with data collection. If the benefits of increased data accuracy are greater than the expected costs, additional funds should be allocated to obtain more accurate data.

The intended use of data affects the type of data, as well as the data quality needed. Beard (8) divides GIS applications into six types (see Table 1). The specific type of data quality one needs (e.g., positional accuracy, attribute accuracy) also varies with the intended application. Analysts with inventory applications such as agricultural production are less concerned about positional accuracy than with an accurate assessment of anticipated crop yields (attribute accuracy). Decision-makers must

Table 1. Types of GIS Applications (8)

Application	Example
Siting	Finding optimal location (fire station, waste site)
Logistic	Movement or distribution through space (emergency response, military movement)
Routing	Optimal movement through a known network (mail, school bus)
Navigation	Way finding; may or may not involve a known network (ground, sea, air)
Inventory	Count and location of objects for a given time (census, tax rolls)
Monitoring/ Analysis	Examining processes over space and time (ecological, zoological, geological, epidemiological studies)

decide which data quality component is the most important for their use because optimizing all six components can be very expensive (9). An obvious conflict arises when local and state governments must meet multiple application needs simultaneously and thus feel forced to try to optimize several data quality components.

The nature of the decision may also help decision-makers determine the data quality they need. Beard (8) lists several of these factors (see Table 2). A political, high-risk decision requires higher quality data than a nonpolitical, low-risk decision because more public attention focuses on the former decision.

Table 2. Factors That May Affect the Data Quality Needed for Decision-Making (8)

Lower Data Quality Possibly Needed	Higher Data Quality Possibly Needed
Routine	Nonroutine
Nonpolitical	Political
Minimal risk	High risk
Noncontroversial	Controversial
Indefinite	Immediate
Local implication	Global implication

Components of Data Quality

The National Committee for Digital Cartographic Data Standards (9) identifies six components of digital cartographic data quality. This section discusses each of these components:

- Lineage
- Positional accuracy
- Attribute accuracy
- Logical consistency
- Completeness
- Temporal accuracy

Most components of data quality apply to both source and operational error.

Lineage

Because uses and users of data change, those at the national level have noted a recent push to include documentation when disseminating spatial data. Data lineage, also known as metadata or a data dictionary, is data about data. Metadata consists of information about the source data such as:

- Date of collection
- Short definition
- Data type, field length, and format
- Control points used
- Collection method, field notes, and maps
- Data processing steps
- Assessment of the reliability of source data
- Data quality reports

Access to this information can help GIS personnel determine if the data are appropriate for their use, thereby minimizing risks associated with using the wrong data or using data inappropriately. According to Chrisman (10), the only ethical and probably best legal strategy for those who produce spatial data is to reveal more information about the data (metadata) so that users can make informed decisions. Eagan and Ventura's article (11) contains a sample of a generic environmental data lineage report. The U.S. Environmental Protection Agency's (EPA's) new locational data policy requires contractors to estimate data accuracy and provide information about the lineage of the data (12).

Positional Accuracy

Anyone who has used a map has probably come across features that are not located where the map says they should be located and has experienced low positional accuracy. (Undoubtedly, they have also detected features that were not on the map, but that is a different issue.) Positional accuracy, frequently referred to as horizontal error, is how close a location on a map is to its "true" ground position. Features may be located inaccurately on maps for many reasons, including (13):

- Poor field work.
- Distortion of the original paper map (temperature, humidity).
- Poor conversion from raster to vector or vector to raster data.
- Data layers are collected at different times.
- Natural variability in data (tides, vegetation, soil).
- Human-induced changes (altering reservoir water levels).
- Movement of features (due to scale of the map and printing constraints) so they can be easily discerned by the map reader.
- Combining maps with different scales.
- Combining maps with different projection and coordinate systems.
- Different national horizontal datum in source materials.
- Different minimum mapping units.

Positional accuracy has two components: bias and precision. Bias reflects the average positional error of the sample points and indicates a systematic discrepancy (e.g., all locations are 7 feet east of where they should be). Estimating precision entails calculating the standard deviation of the dispersion of the positional errors. Usually, root mean square error (RMSE) is reported as the measure of positional accuracy, but it does not distinguish bias from precision (14). RMSE is frequently monitored during digitizing to minimize the introduction of additional positional error into the GIS.

To determine positional accuracy, one must compare the location of spatial data with an independent source of higher accuracy. Federal agencies that collect data and produce maps adhere to National Map Accuracy Standards (NMAS) for positional accuracy. Maps such as United States Geological Survey (USGS) topographic maps that conform to NMAS carry an explicit statement on them. Other groups also have developed standards for large-scale mapping (15).

NMAS for positional accuracy require that not more than 10 percent of well-defined points can be in error by more than one-thirtieth of an inch for maps at a scale of 1:20,000 or larger. For smaller scale maps, not more than 10 percent of well-defined points can be in error by more than one-fiftieth of an inch (16). Thus, less than 10 percent of the well-defined locations on a USGS 1:24,000 map can stand more than 40 feet from their "true" location; the other 90 percent of the well-defined points must stand less than 40 feet from their "true" location. Table 3 shows the acceptable positional accuracy for commonly used maps. Note that as scale decreases from 1:1,200 to 1:100,000, positional accuracy decreases.

Several important issues relate to NMAS. First, not all maps adhere to NMAS, which means their positional accuracy may be lower than NMAS or may be unknown. Second, NMAS do not indicate the location of points in error. Third, 10 percent of the well-defined points can have a positional error greater than the standards allow, but neither the location nor the magnitude of these errors are known. Fourth, NMAS apply to well-defined points; therefore, areas that are not well defined may have even lower positional accuracy. The implication of these errors in location is that users should use caution in making decisions that require high positional accuracy. Positional accuracy issues are particularly troublesome for GIS operations on small-scale maps or when combining large-scale maps (1:1,200) with small-scale maps (1:100,000).

Table 3. NMAS Horizontal (Positional) Accuracy

Scale	1 Inch = x Feet	Horizontal Accuracy +/- Feet
1:1,200	100	3.33
1:2,400	200	6.67
1:4,800	400	13.33
1:12,000	1,000	33.33
1:24,000	2,000	40.00
1:63,360	5,280	105.60
1:100,000	8,333	166.67

Recently, global positioning systems (GPS), which the U.S. military developed, have helped to obtain more accurate feature locations. GPS is not without error, however. The list below notes some of the possible sources of error associated with GPS use, some of which can be controlled while others cannot (17):

- Errors in orbital information.
- Errors in the satellite clocks.
- Errors in the receiver clocks.
- Ionospheric or tropospheric refraction.
- Deliberate degrading of the satellite signal.
- Obstructions that block the signal.
- Reflection of the GPS signal off buildings, water, or metal.
- Human error.

The importance of positional accuracy depends on the intended use of the data. In an urban area, a positional error of 1 foot on a tax map may be unacceptable because 1 foot may be worth millions of dollars. In a rural area, however, tax boundaries mapped within 10 feet of their surveyed location may be accurate enough (6). Somers (18) reports that positional accuracy of 10 to 20 feet may be sufficient for environmental analysis. She says the cost of increasing accuracy to 5 feet could increase the cost of data collection by a factor of 10. The decision-maker must determine the needed positional accuracy.

Attribute Accuracy

Attribute accuracy refers to how well the description of a characteristic of spatial data matches what actually exists on the ground. For some spatial data, the location does not change over time, but the value of the attribute does (e.g., the location of a census tract does not change, but the population within a census tract changes). Attribute accuracy is reported differently for continuous data (i.e., elevation, which has an infinite number of values) or discrete data (i.e., gender, which has a finite number of values).

NMAS exist for elevation contour lines on topographic maps. NMAS for vertical accuracy state that not more than 10 percent of the points tested shall be in error by more than one-half of the contour interval (16). A well-defined point on a USGS topographic map with a 10-foot contour interval could vary by 10 feet because the actual elevation could be 5 feet higher or lower than the map indicates. The implications of these errors are similar to the ones for positional accuracy. In addition, errors in elevation are important because small changes in elevation may significantly affect some GIS analysis operations such as the determination of aspect, slope, viewshed, and watershed boundaries.

NMAS do not exist for discrete variables such as land use derived from satellite imagery. Instead, a classification matrix reports attribute accuracy. Field checking or checking a portion of the classified image against a map of higher accuracy determines the accuracy of the land use classification. The result of the comparison is a table from which to calculate overall, producer's, and user's accuracy. Table 4 is an example of a classification accuracy matrix.

Table 4. Example of a Classification Accuracy Matrix (19)

Classified Data (Satellite Image) Number of Cells	Reference Data ("Ground Truth") Number of Cells: Forest	Water	Urban	Total
Forest	28	14	15	57
Water	1	15	5	21
Urban	1	1	20	22
Total	30	30	40	100

Overall Accuracy (sum of the main diagonal)

$$\frac{63}{100} = 63\%$$

Producer's Accuracy (column total)			User's Accuracy (row total)		
Forest	$= \frac{28}{30}$	= 93%	Forest	$= \frac{28}{57}$	= 49%
Water	$= \frac{15}{30}$	= 50%	Water	$= \frac{15}{21}$	= 71%
Urban	$= \frac{20}{40}$	= 50%	Urban	$= \frac{20}{22}$	= 91%

Overall accuracy is the percentage of correctly classified cells calculated as the sum of the main diagonal (19). Producer's accuracy is the total number of correct pixels in a category divided by the total number of pixels of that category as derived from the reference data (column total). It corresponds to how well the person classifying the image (the "producer") can correctly classify or map an area on the earth. In this example, the producer most accurately classified forested land (93 percent). User's accuracy describes the probability that a sample from the classified area actually represents that category on the ground. The map "user" is concerned about the map's reliability. In this example, the most accurately classified land use from the user's perspective is urban (91 percent).

The significance of overall, producer's, and user's accuracy depends on the intended use of the data. As an example, Chrisman (20) says that the error in distinguishing wetland from pasture may not matter to someone estimating open space, but the difference is critical if the person is estimating the amount of wildlife habitat available. Story and Congalton (19) provide an example of how to interpret a classification matrix. A forester looks at the classification matrix and sees that forest classification is 93 percent accurate (producer's accuracy); therefore, the analyst did not identify only 7 percent of the forest on the ground. Once the forester field checks the supposed forested area, she finds that only 49 percent (28 cells) of the sites mapped as forest are actually forest; the rest are water (14 cells) or urban (15 cells) areas.

A report of overall, producer's, and user's accuracy can help decision-makers determine the appropriateness of the classified image for their use by identifying potential errors in classification. This can help direct field work, which can improve the classification of the image and perhaps subsequent images. Because GIS analysis frequently uses land use, decision-makers need to know that significant variability can result when several analysts classify the same image. Bell and Pucherelli (21) found that consistency in classification can improve by having one person classify the entire image. McGwire (22) even found significant differences between analysts in unsupervised classification of Landsat imagery. Computers primarily perform unsupervised classification, which implies that different analysts would classify the same image in the same way.

Logical Consistency

Logical consistency focuses on flaws in the logical relationships among data elements. For example, a vector GIS should label all polygons with only one label per polygon, and all polygons should be closed. Logical inconsistency can also occur by collecting data layers at different times or from different scale maps with different positional accuracies. For example, the edge of a lake on the hydrology data layer should coincide with the edge of land in the land use data layer. If data on the lake were collected during a wet year rather than a dry year, the lake's volume would be higher than normal, affecting its location on the map. If land use data for the same area were collected during a dry year, the boundary of the lake on the two layers would not be the same.

Logical inconsistencies usually do not appear until the two maps are overlaid and the boundaries do not coincide (see Figure 2). The user must determine the "correct" location of the feature that appears misaligned on one or more data layers. The inconsistency between the

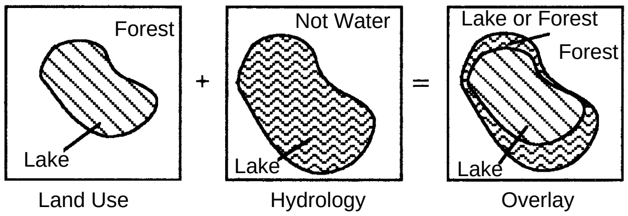


Figure 2. Logical inconsistency in lake and forest location.

location of the two layers resolves through a process called conflation. All maps are adjusted so that the feature on each data layer lines up with the same feature on the base map.

Completeness

Completeness focuses on the adequacy of data collection procedures. Robinson and Frank (5) discuss two kinds of uncertainty associated with collecting spatial data that can lead to error. One type of uncertainty is the inability to measure or predict an inherently exact characteristic or event with certainty. Examples of this are blunders in data collection or measurement error, neither of which can be accurately predicted. The other kind of uncertainty is associated with concepts that are inherently ambiguous. Crisp data sets, such as property boundaries, have little ambiguity; the only issue related to error is the positional accuracy in measuring the boundary. Because land use data are not crisp data sets, the challenge is to accurately represent an inherently inexact concept.

Although we know spatial data are variable, our classification systems generally ignore the second type of uncertainty. Analysts map data as though all variables had exact boundaries and all polygons consisted of homogeneous data. Burrough (4) reports that spatial variation of natural phenomena is "not just a local noise function or inaccuracy that can be removed by collecting more data or by increasing the precision of measurement, but is often a fundamental aspect of nature that occurs at all scales. . . ."

Mapping spatial data is a function of how humans aggregate and disaggregate data either in space, categories, quantities, or time; spatial data seldom exist in nature the way maps depict them (23). Data and relationships between data are sensitive to the scale and the zoning system in which the data are reported (24, 25). The modifiable area unit problem occurs because an analyst can recombine a given set of units or zones into the same total number of units producing very different results (see Figure 3).

The scale problem occurs because an analyst can combine a set of small units into a smaller number of larger units, which can change the inferences that can be

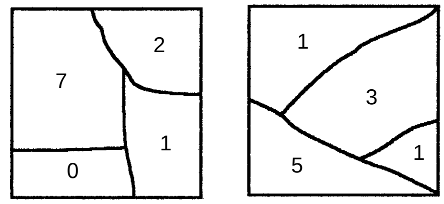


Figure 3. Modifiable area unit. (Number of units is constant; location of units changes.)

made from the data. In Figure 4, the area containing the highest values changes from the southwest corner in the first picture to the northern half in the second picture. For example, water quality data are scale-dependent because they vary based on the size and location of the collection area (e.g., adjacent to a point source discharge, a stream segment, the entire river, or the lake the river discharges into).

Kennedy (25) reports on a similar problem known as the small number problem. This problem occurs when calculations use a percentage, ratio, or rate for a geographic area for which the population of interest (denominator) is sparse or the numerator is a rare event (1 case of cancer per 1 million people). GIS-generated maps may highlight a statistically insignificant change in rare events. Small, random fluctuations in the numerator may cause large fluctuations in the resulting percentage, ratio, or rate. If policy-makers use these maps, priorities for public health policy may change because of the erroneous belief that an area is experiencing more unwanted rare events.

Data can be collected using a tag- or count-based system, which affects their usefulness. The tag approach categorizes items based on the dominant or average attribute and is ideal for planners who want only one value for each area. For example, each polygon in a county soil survey is tagged with one soil type. According to soil taxonomy rules, however, only about 35 percent of a delimited area on a soil survey must match its classification, and up to 10 percent may be a radically different soil (26). Although the text in the soil survey sets limits on data accuracy by listing major impurities found with each soil type, the GIS seldom carries that information because analysts only digitize soil boundaries and label data with the dominant attribute. This

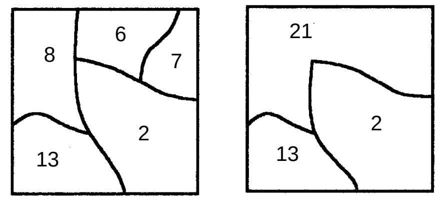


Figure 4. Scale problem (number of units changes).

leads to the depiction of apparently homogeneous soil units although the text specifies that the data are not homogeneous (27).

Some soil or land cover phenomena, even though present in small quantities and thus not mapped, may have great significance for hydrologic models, which makes the tag approach to data collection troublesome. Data collected using the count system allow the analyst to tabulate the frequency of occurrence or areal extent of a particular phenomenon. Environmental modelers prefer count data but are usually forced to use tag data, which can introduce error into their models (26). The new digital soils databases, STATSGO and SSURGO, are collected and depicted using a count format, which will help experienced analysts use the data more fully. Figure 5 shows the difference between tag and count methods of data collection.

Data are seldom complete because analysts use classification rules to indicate how homogeneous an area must be before it is classified a particular way (e.g., more than 50 percent, more than 75 percent). Another decision an analyst must make is where to draw the boundary between two different areas; it is seldom clear where a forest leaves off and a rural development begins. Analysts must also decide how or if to show inclusions (e.g., a forested area in the middle of agricultural land uses).

Temporal Accuracy

Collecting data at different times introduces error because the variable may have changed since data collection. The effect of time, reported as the date of the source material, depends on the intended use of the data. Some natural resource data have daily, weekly, seasonal, or annual cycles that are important to consider. For example, obtaining land use data from remotely sensed imagery in November for North Dakota produces a very different land use map than data analysts obtain during the July growing season.

In addition, demographic and land use information changes quickly in a rapidly urbanizing area. Data collected at several times can produce logical inconsistency between data layers, forcing the analyst to adjust the location of features to coincide with the base map. Another problem with collecting data at different times is that data may be collected using different standards, which may not be apparent to the user (4).

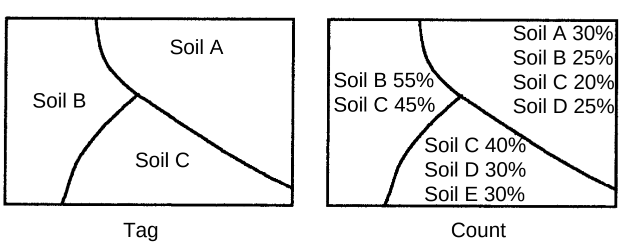


Figure 5. Tag and count methods of data collection.

Source Errors in a GIS

Source (or inherent) error derives from errors in data collection. The amount of error present in collected data is a function of the assumptions, methods, and procedures used to create the source map (28). Primary data refers to data collected from field sampling or remote sensing. Causes of the errors associated with this data are (3, 4, 8, 14, 29):

- Environmental conditions (e.g., temperature, humidity).
- Sampling system (e.g., incomplete or biased data collection).
- Time constraints.
- Map projection.
- Map construction techniques.
- Map design specifications.
- Symbolization of data.
- Natural variability.
- Imprecision due to vagueness (e.g., classifying a forest).
- Measurement error from unreliable, inaccurate, or biased observers.
- Measurement error from unreliable, inaccurate, or biased equipment.
- Lab errors (e.g., reproducibility between lab procedures and between labs).

The process of converting primary data to secondary data (usually a map) introduces additional error. Many of the data layers that a GIS analyst acquires are secondary data. Some of the errors associated with mapmaking are (3):

- Error in plotting control points.
- Compilation error.
- Error introduced in drawing.
- Error due to map generalization.
- Error in map reproduction.
- Error in color registration.
- Deformation of the material (temperature, humidity).
- Error introduced due to using a uniform scale.
- Uncertainty in the definition of a feature (boundary between two land uses).
- Error due to feature exaggeration.
- Error in digitization or scanning.

Converting paper maps to digital data for entry into a GIS (tertiary data) introduces still more error (the errors generated from converting paper maps into a digital format are discussed in the section on input error), in part because the purpose for which the data was collected differs from the intended use of the data.

Many types of error are associated with data collection:

- Data for the entire area may be incomplete.
- Data may be collected and mapped at inappropriate scales.
- Data may not be relevant for the intended application.
- Data may not be accessible because use is restricted.
- Resolution of the data may not be sufficient.
- Density of observations may not be sufficient.

The following discussion explains these types of errors.

Data for the Entire Area May Be Incomplete

An incomplete data record may be due to mechanical problems that interrupt recording devices, cloud cover or other types of interference, or financial constraints. Possible solutions to this problem include collecting additional data for the incomplete area, using information from a similar area, generalizing existing large-scale maps to match the less detailed data needed, or converting existing small-scale maps to large-scale maps to obtain data at the desired scale. Collecting additional data may not be a feasible solution because of time or money constraints. Extrapolating data from the surrogate area to the desired area can cause problems because the areas are not identical and the scale, accuracy, or resolution of the surrogate area data may be inappropriate for the intended use. The section on analysis/manipulation of data within a GIS covers the effect of generalization on data quality as well as the effect of converting small-scale maps to large-scale maps.

Data May Be Collected and Mapped at a Scale That Is Inappropriate for the Application

A variety of guidelines suggest the appropriate map scale to use for various applications (see Table 5). Also,

Table 5. Relationship Between Map Scale and Map Use (6)

Map Scale	Map Use
1:600 or larger	Engineering design
1:720 to 1:1,200	Engineering planning
1:2,400 to 1:4,800	General planning
1:6,000 and smaller	Regional planning

some maps and digital databases suggest the type of application for which they are appropriate (e.g., the STATSGO digital soil database is suitable for state and regional planning, whereas SSURGO is suitable for local level planning). Tosta (30) cites an example of combining wetland data with parcel boundaries to determine ownership of the land containing a wetland. If wetland mapping was done to plus or minus 100 feet positional accuracy and parcels are 40 feet wide, then the scale of the wetland map is inappropriate for determining if a wetland is located on a specific parcel.

Identifying the optimal scale of the necessary data is crucial because at some point, the cost of collection and storage exceeds the benefits of increasing the map scale. Lewis Carroll (1893) summed up the quest for data mapped at an ever larger scale and the problems associated with large-scale maps:

> “What do you consider the *largest* map that would be really useful?”
>
> “About six inches to the mile.”
>
> “Only *six inches!”* exclaimed Mein Herr. “We very soon got to six *yards* to the mile. Then we tried a *hundred* yards to the mile. And then came the grandest idea of all! We actually made a map of the country, on the scale of *a mile to the mile!”*
>
> “Have you used it much?” I enquired.
>
> “It has never been spread out, yet,” said Mein Herr. “The farmers objected: they said it would cover the whole country, and shut out the sunlight! So now we use the country itself, as its own map, and I assure you it does nearly as well.”

Data Collected May Not Be Relevant for the Intended Application

Frequently, using surrogate data is quicker or cheaper than collecting needed data (e.g., Landsat imagery rather than data field collection used to determine land use) (4). The accuracy and classification scheme used in collecting the data depends on the intended use of the data, which may not coincide with the analyst’s purpose. For instance, soil maps were developed to aid farmers in determining what crops they should plant and for estimating crop yield. Soil maps, however, see wide use for very different purposes (e.g., hydrologic and other environmental models). In addition, STORET data, collected at points, are typically extrapolated to represent water quality in an entire stream stretch.

Data May Not Be Accessible Because Use Is Restricted

An example of restricted data is Census data on individual households. An agency may not want to release data that reveal the location of endangered species. Another example is that people may not even want the information mapped. For example, some cavers do not want to reveal the location of caves to the U.S. Forest Service, which is charged under the federal Cave Resources Protection Act with protecting caves, because they think the best way to protect the caves is to not map them (31). The National Park Service is putting the location of petroglyphs in the Petroglyph National Monument into a GIS. Making their location known to the public, however, is troublesome because this may, in fact, encourage their vandalism (32). Other problems in obtaining data include difficulty in acquisition even if access is not restricted, expensive collection or input, or unsuitable format (4, 14).

Resolution of the Available Data May Not Be Sufficient

Spatial resolution is the minimum distance needed between two objects for the equipment to record the objects as two entities; that is, resolution is the smallest unit a map represents. To obtain an approximation of a map's resolution, divide the denominator of the map scale by 2,000 to get resolution in meters; for instance, a 1:24,000-scale map has a resolution of approximately 12 meters (33).

Resolution relates to accuracy in that different map scales conform to different accuracy standards. Two air photos shot from the same camera at the same distance above the ground have the same scale. If one photo has finer grain film, however, smaller details are evident on it, and this photo produces a map with higher resolution (34). According to Csillag (33), analysts cannot simultaneously optimize attribute accuracy and spatial resolution. As spatial resolution increases, attribute complexity increases (35). Also, the finer the spatial resolution, the greater the probability that random error significantly affects a data value.

Resolution of the data is not necessarily the same as the size of a raster cell in a database. Statistical sampling theory suggests using a raster cell size that is half the length (one-fourth of the area) of the smallest feature an analyst wishes to record. Raster data have a fixed spatial resolution that depends on the size of the cell employed, but a GIS analyst can divide or aggregate cells to achieve a different cell size. Frequently, an analyst transforms data collected at one level of resolution to a higher level of resolution than existed in the original source material. According to Everett and Simonett (23), "Geographic analysis, however, can be no better than that of the smallest bit of data which the system is capable of detecting." Vector data are limited by the resolution of input/output devices, limits on data storage, and the accuracy of the digitized location for individual points (36). The spatial resolution of the database and the processes that operate on it should be reduced to a level consistent with the data's accuracy.

The spatial resolution needed depends on the intended use of the data, cost, and data storage considerations. As resolution increases, so does the cost of collection and storage. Resolution sufficient to detect an object means that an analyst can reveal the presence of something. Identification, the ability to identify the object or feature, requires three times the spatial resolution of detection. Analysis, a finer level of identification, requires 10 to 100 times the resolution that identification needs (23). Increasing resolution increases the amount of data for storage, with storage requirements increasing by the square of the resolution of the data. For example, if the resolution of the data needs to change from 10-meter to 1-meter pixels, file size increases by 10^2 or 100 times (14).

Density of Observations May Be Insufficient

The density of observations serves as a general indicator of data reliability (4). Users need to know if sampling was done at the optimum density to resolve the pattern. Burrough determined that boulder clay in The Netherlands could be resolved by sampling at 20-meter intervals or less, whereas coversand showed little variation in sampling from 20- to 200-meter intervals.

Some strategies for reducing data collection errors are to:

- Adhere to professional standards
- Allocate enough time and money
- Use a rigorous sampling design
- Standardize data collection procedures
- Document data collection procedures
- Calibrate data collection instruments
- Use more accurate instruments
- Perform blunder checks to detect gross errors

Documenting data collection procedures and distributing them along with data allows potential users to determine if the data are suitable for their purposes. By not documenting procedures, errors in the source material are essentially "lost" by inputting the data to a GIS, and the errors become largely undetectable in subsequent GIS procedures. The result is that agencies that make decisions based on the GIS-generated map assume the source data are accurate, only to discover later that the map contains substantial errors in part due to errors in the source material.

Operational Errors in a GIS

Data input, storage, analysis/manipulation, and output can introduce operational errors. Digital maps, unlike paper maps, can accumulate new operational errors through GIS operations (8). Even if the input data were totally error-free, which the last section demonstrated is not the case, GIS operations can produce positional and attribute errors. The GIS operation itself determines to a large extent the types of errors that result.

Input Errors

The process of inputting spatial and attribute data can introduce error. The major sources of input error are manual entry of attribute features and scanning or digitizing spatial features. Manual entry errors include incomplete entry of attribute data, entering the wrong attribute data, or entering the right attribute data at the wrong location. Digitizing errors originate from equipment, personnel, or the source material (see Table 6).

Digitizing errors, such as under- and overshoot of lines and polygons that are not closed, can introduce error (see Figure 6). GIS software can "snap" lines together that really do not connect. Depending on the tolerance selected, this can result in the movement of both lines, which can decrease the accuracy of the resultant map.

Table 6. Types of Digitizing Errors (4, 14, 37)

Personnel Errors
- Changes in the origin
- Incorrect registration of the map on the digitizing table
- Creation of over- and undershoots
- Creation of polygons that are not closed
- Incomplete spatial data when data are not entered
- Duplication of spatial data when lines are digitized twice
- Line-following error (inability to trace map lines perfectly with the cursor)
- Line-sampling error (selection of points used to represent the map)
- Physiological error (involuntary muscle spasms)

Equipment Errors
- Digitizing table (center has higher positional accuracy than the edges)
- Resolution of the digitizer
- Differential accuracy depending on cursor orientation

Errors in Source Material
- Distortion because source maps have not been scale-corrected
- Distortion due to changes in temperature and humidity
- Necessity of digitizing sharp boundary lines when they are gradual transitions
- Width of map boundaries (0.4 mm) digitized with a 0.02-mm accuracy digitizer

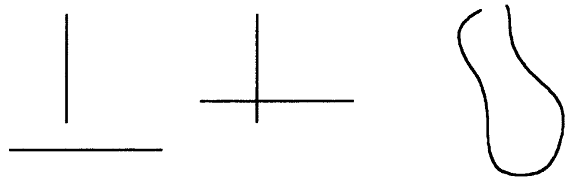


Figure 6. Common digitizing errors.

Despite the long list of personnel errors associated with digitizing, a good operator probably contributes the least error in the entire digitizing process (38). Giovachino discusses methods that can help determine equipment accuracy, including checking the repeatability, stability, and effect of cursor rotation. Digitizing accuracy varies based on the width, complexity, and density of the feature being digitized but typically varies from 0.01 to 0.003 (3).

One problem with digitized data is that the data can imply a false sense of precision. Boundaries on paper maps are frequently 0.4 mm wide but are digitized with 0.02-mm accuracy. The result is that the lines are stored with 0.02-mm accuracy, implying a level of precision that far exceeds the original data.

Minimizing digitizing errors is important because the errors can affect subsequent GIS analysis. Campbell and Mortenson (39) provide a list of procedures they used to reduce errors associated with digitizing and labeling:

- Use log sheets to ensure consistency and accountability, and to provide documentation.
- Check for completeness in digitizing all lines and polygons.
- Check for complete and accurate polygon labeling.
- Set an acceptable RMSE term for digitizing (usually 0.003).
- Always overshoot rather than undershoot when digitizing.
- Overlay a plot of the digitized data with the source map to check lines and polygons. If light passes between the digitized line segment and source map, redigitize it.
- Check digitized work immediately to provide feedback to the digitizer operator and to help identify and correct systematic errors.
- Limit digitizing to less than 4 hours a day.
- Involve people in doing GIS-related jobs other than digitizing to decrease turnover and increase the level of experience.

Storage Errors

Data storage in a GIS usually involves two main types of errors. First, many GIS systems have insufficient numerical precision, which can introduce error due to rounding. Integers are stored as 16 or 32 bits, which have four significant figures. Real numbers are stored as floating point numbers either in single precision (32 bit, 7 significant figures) or double precision (64 bit, 15 or 16 significant figures). If the data in a GIS range from fractions of a meter to full UTM coordinates, typical 32-bit GIS systems cannot store all the numbers. Using double precision (64 bits) reduces this problem but increases storage requirements.

Second, GIS processing and storage usually ignore significant digits (data precision). As a result, the precision of GIS processing frequently exceeds the accuracy of the data (40). When a GIS converts a temperature recorded and entered as 70 degrees Fahrenheit (nearest degree) to centigrade, the GIS stores the temperature as 21.111 degrees rather than 21 degrees, which the significant figures in the original temperature measurement would dictate. Using the accuracy of the data, not the precision of floating point arithmetic, partially resolves this but requires the user to make a special effort because the GIS does not automatically track significant figures.

Analysis/Manipulation Errors

GIS analysis/manipulation functions, designed to transform or combine data sets, also can introduce errors. These errors originate from the measurement scale used or during data conversion (vector to raster and raster to vector), map overlay, generalization, converting small-scale to large-scale maps, slope, viewshed, and other analysis functions. One of the biggest problems associated with GIS use is that data in digital form are subject to different uses than data in paper form because the user has access to multiple data layers.

Measurement Scale

Four measurement scales can depict spatial data: nominal, ordinal, interval, or ratio scales. A name or letter describes nominal data (e.g., land use type, hydrologic soil group C). Performing mathematical operations such as addition and subtraction on nominal data is meaningless. Ordinal or ranked data have an order to them such as low, medium, and high. Interval data have a known distance between the intervals such as 0, 1 to 5, 6 to 9, more than 9. Ratio data are similar to interval data except ratio data have a meaningful zero (e.g., temperature on the Kelvin scale).

Often during GIS operations, analysts convert interval or ratio data into nominal data (e.g., low slope is 0 to 3 percent, medium slope is 4 to 10 percent), resulting in a loss of information. Analysts should preserve the original slope values in the GIS in case the user later wants to modify the classification scheme. Robinson and Frank (5) describe the tradeoff between information content and the meaning that can be derived from it, which partly helps explain why interval data are frequently converted to nominal data. The authors identify a continuum progressing from nominal data at one end that is highly subjective, has low information content, and high meaning (low slope means something to the average user) to ratio data that has low subjectivity, high information content, and low meaning (a slope of 7 percent may not mean much to the average user).

Data Conversion

Errors can occur in converting a vector map to a raster map or a raster map to a vector map. For instance, remotely sensed data are collected using a raster-based system. Using a vector GIS, however, requires conversion from raster to vector data. The size of the error depends on the conversion algorithm, complexity of features, and grid cell size and orientation (13).

A line on a vector map converted to a raster map has lower accuracy in the raster representation because vector data structures store data more accurately than raster ones. When polygons in a vector GIS are converted to a raster GIS, the coding rule usually used assigns the value that covers the largest area within the cell of a categorical map to the entire cell (see Figure 7). For example, when placing a grid over a vector map with an urban land polygon adjacent to an agricultural polygon, the cell placement can include part of both polygons. If the resultant cell comprises 51 percent urban and 49 percent agricultural land, the cell is assigned 100 percent urban. Converting a numerical map between raster and vector systems requires spatial interpolation procedures. GIS software packages use different interpolation methods that can produce a different output even when using the same input data.

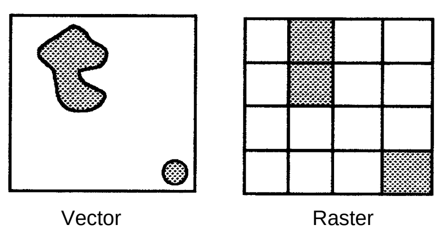


Figure 7. Polygon conversion from vector to raster data.

Map Overlay

Map overlay, used extensively in planning and natural resource management, is the combining of two or more data layers to create new information. In a vector GIS, slivers or spurious polygons can result from overlaying two data layers to produce a new map (slivers cannot be formed in a raster-based GIS). When combining the data layers, lines do not coincide, resulting in the creation of a new polygon or sliver that did not exist on either layer (see Figure 8). Unfortunately, as accuracy in digitizing increases, so does the number of slivers (41). Positional error in the boundaries can occur because of

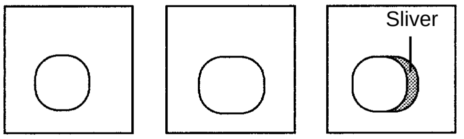


Figure 8. Sliver example.

mistakes in measuring or converting the data to digital form, incremental expansion or recession of a real world boundary over time, or the fact that certain boundaries are difficult to determine and thus are generalized differently (42).

The number of map layers, accuracy of each map layer, and the coincidence of errors at the same position from several map layers all determine the accuracy of the map overlay procedures (43). Using probability theory, Newcomer and Szajgin determined that the highest accuracy to expect from a map overlay is equal to the accuracy of the least accurate map layer. The lowest accuracy in map overlay occurs when errors in each map occur at unique points.

In the quest for more accurate results, GIS modelers have increased the complexity of their models and therefore have increased the number of data layers needed. Guptill (44) states, “Conventional wisdom would say that as you add more data to the solution of a problem, the likelihood of getting an accurate solution increases. However, if each additional data layer degrades the quality of the combined data set, and hence the accuracy of the solution, then additional data sets may be counterproductive.”

Generalization

Monmonier (45) provides an extensive discussion of geometric and content generalization procedures used in map-making. Table 7 lists common types of generalization. Generalizing data on a map helps to focus the user’s attention on one or two types of information and to filter out irrelevant details. Generalizing is performed by reducing the scale of the data; a 1:24,000-scale map can be generalized to a 1:100,000-scale map so that all data layers have the same scale. With generalizing, areas on a large-scale map become point or line features on a small-scale map (35). Obtaining some measurements from small-scale maps, however, requires caution. For example, a map may depict a 40-foot wide road as a single line one-fiftieth of an inch wide. On a 1:100,000 map, one-fiftieth of an inch translates into a 160-foot wide road—four times the actual width of the road.

Several studies have pointed to errors that can result from generalization. Wehde (46) compared soil maps generated from 0.017-acre grid cells and 11 progressively increasing grid cell sizes. He found that as grid cell size increased, map accuracy decreased. More recently, Stoms (47) found that generalizing a habitat map from 1 to 25, 100, and 500 hectares decreased the number of habitat types and the number of species predicted.

Table 7. Generalization Operations (45)

Geometric Generalization
Generalizing a Line
Simplification
Displacement
Smoothing
Enhancement
Selection
Generalizing a Point
Selection
Displacement
Graphic association
Abbreviation
Aggregation
Area conversion
Generalizing an Area
Selection
Simplification
Displacement
Smoothing
Enhancement
Aggregation
Dissolution
Segmentation
Point conversion
Line conversion
Content Generalization
Selection
Classification

Transforming Small-Scale Maps to Large-Scale Maps

Converting small-scale maps (1:250,000) to large-scale maps (1:24,000) is advisable only if the analyst fully appreciates the effect of this procedure on map quality. Data mapped at a small scale are subject to different accuracy standards than data mapped at a large scale. Connin (48) reports, “Problems with accuracy arise when positions are reported to decimal parts of a foot or meter, but the method of data capture may cause the positional error to be as much as hundreds of feet or meters.” Yet when converting the data from small- to large-scale, the data appear to have the accuracy of the large-scale map. Theoretically, data should not be transformed and used at a scale larger than the scale of the document from which the data are derived (3).

Slope and Viewshed

GIS software packages use a variety of algorithms to calculate slope and viewsheds and can produce very different results. Algorithms are an unambiguous set of rules or a finite sequence of operations used to carry out a procedure. Smith, Prisley, and Weih (49) used six different GIS algorithms to determine slope on 5,905 acres of land in order to calculate the amount of land deemed unsuitable for timber harvesting. They found that unsuitable land varied from 175 to 1,735 acres, indicating that different algorithms produce very different results. Felleman and Griffin (50) found that GIS packages with different algorithms generate alternate viewsheds (the area that can be seen from a point).

Output Errors

A variety of errors are associated with data output:

- Output devices create error.
- Paper shrinks and swells.
- Line implies certainty that may not exist because boundaries are gradual.
- A cell or polygon implies homogeneity.
- Scale can be modified to imply higher accuracy than exists in the source data.
- Precision can be modified to imply higher precision than exists in the source data.
- Depiction of symbols and colors may not follow convention.

An important problem associated with GIS-generated maps is that users make informal assessments about data quality, partially based on how they perceive the quality of the output. A hand-drawn map connotes a lower level of accuracy than a five-color, GIS-produced map complete with scale and agency logo. Another problem with output is that distinguishing highly accurate data from less accurate data is impossible on a GIS-generated map. Users want the output from a GIS to look like maps they usually see, perpetuating the notion that lines mark exact boundaries and that polygons or cells are homogeneous. Maps that federal mapping agencies produce frequently follow NMAS, but GIS-generated maps seldom adhere to published map accuracy standards. An agency could require that GIS map products meet NMAS, which would establish and maintain data standards from data collection to output.

A pen stroke of one-fiftieth of an inch on an output device translates to an error of 40 feet on the ground for a 1:24,000-scale map (6). Small changes in paper maps due to changes in temperature and humidity can represent several feet on the ground. As previously noted, analysts can modify the scale of GIS maps to whatever they desire. The basic rule of informational integrity is that the implied precision of data output should not exceed the precision (spatial, temporal, or mathematical) of the least precise input variable (26).

GIS-generated maps probably do not differ significantly from paper maps in their implication that lines and polygons on the map represent certainty and homogeneity. GIS-generated maps, however, may not depict standard symbols, sizes, shapes, colors, and orientation. For example, paper geological maps use dashed lines to show inferred, rather than actual, field collected data, but geological maps in a GIS may not follow the same convention (27). Cartographers conventionally use blue lines to indicate water, but a GIS map-maker can show water as red rather than blue.

Even more troublesome are the color schemes that some analysts use in depicting model output. Analysts often give little thought to assigning the colors to model results depicted as ordinal rankings. For example, areas of high erosion might be blue, medium erosion might be red, and low erosion might be green. This selection of colors ignores the intuitive meaning that people assign to colors. It has been suggested that the color ordering used in stop lights might provide a better option. In that case, areas of high erosion would be red, medium erosion would be yellow, and low erosion would be green.

Error Reduction Techniques

Although GIS users and researchers develop error reduction strategies, ultimately users must rely on GIS software developers to implement new error reduction techniques in GIS packages. Error reduction techniques range from simple software warnings to prohibiting a user from performing selected GIS procedures. Dutton (51) predicts that future GIS programs will automate data manipulation (i.e., size, format, and placement of feature labels on maps) in keeping with standard cartographic principles. Dutton (51) and Beard (8) also predict that future GIS packages will enforce metadata-based constraints such as operations that are illegal or illogical (e.g., determining the average value of nominal data such as land use), or are inadvisable (e.g., overlaying maps with widely different scales).

Another change Dutton anticipates is that software vendors will include information in manuals that explains how executing a specific command may affect the database. Graphic techniques to depict error are being developed for nonexpert users while experts tend to use spatial statistics. Felleman (52) and Berry (53) present an interesting graphic portrayal of an error map that may indicate the future of error maps. Additional research must determine what effect errors will have on decision-making.

Error Management

Ultimately, the decision-maker must determine what to do with the information in this paper. A decision-maker has a variety of possible courses of action, ranging from prudent steps that attempt to minimize error and the effect it has on decisions, to other less useful options. Possible actions are to:

- Abandon use of a GIS.
- Ignore the error associated with GIS use.
- Attempt to collect “error-free” data.
- Determine if the data are accurate enough for the intended purpose.
- Develop and use data quality procedures.
- Obtain and use an error report with GIS-generated output.
- Ask that GIS-generated maps show potential errors.
- Continually educate users about the appropriate use of spatial data.

First, the decision-maker could abandon any attempt to use a GIS because of the errors associated with its use. At times, this may be the appropriate strategy, but this approach ignores the potential benefits associated with GIS use.

Second, the decision-maker could ignore the error associated with GIS use and continue to use the GIS for decision-making. This type of “head in the sand” approach is not advisable because of the potential liability associated with making decisions based on inaccurate data.

Third, the decision-maker could engage in an expensive and time-consuming effort to collect highly accurate error in hopes that error becomes a nonissue. Depending on the intended use of the data, the cost of collecting more accurate data may exceed the benefit.

Fourth, the decision-maker could assess whether the information available is accurate enough for the intended purpose. If data quality is too low, the decision-maker may opt to collect new data at the desired quality. If collecting additional data is not possible, the decision-maker can explore what types of decisions are possible given the attainable data quality. For instance, Hunter and Goodchild (54) found that the data they were using were suitable only for initial screening rather than for regulatory and land-purchasing decisions.

Fifth, procedures to ensure high quality data could be developed and used in the data collection, input, and manipulation stages of building a GIS database.

Sixth, the decision-maker could require a quantitative or at least a qualitative report on the sources, magnitude, and effects of errors. The absence of an error report does not mean the map is error-free (36). Dutton (51) predicts that in the near future users of geographic data will demand error reports, confidence limits, and sensitivity analyses with GIS-generated output.

Seventh, the decision-maker could ask for GIS-generated maps that adequately portray the error in the final map. For example, areas where the uncertainty is high could appear in red on maps. Another option is to place a buffer around lines to indicate the relative positional accuracy of a line or to show transition zones. Finally, an analyst can present the output in ways other than a dichotomous yes or no; instead, the analyst may use yes, maybe, or no depictions or even more gradations.

Finally, Beard (8) introduced the concept of directing efforts toward educating users about use error. She defines use error as the misinterpretation of maps or misapplication of maps to tasks for which they are not appropriate. “We can’t assume that GIS will automatically be less susceptible to misuse than traditional maps, and it may, in fact, exacerbate the problem by expanding access to mapped information.” Beard argues that money directed to reducing source and operational error, while important, may not matter if use error is large.

Conclusions

GIS is a powerful tool for analyzing spatial data. Everyone who uses GIS-generated output, however, must be aware of source errors and operational errors introduced during data input, storage, analysis/manipulation, and output. Increased awareness of the sources and magnitude of error can help decision-makers determine if data are appropriate for their use. Decision-makers cannot leave data quality concerns to GIS analysts because efforts to improve data quality are not without cost, and the decision-makers typically control funding.

Decision-makers must not get caught up in the glamour of the spatial analyses and outputs that a GIS can produce. These attributes may lead decision-makers to ignore issues associated with uncertainty, error, accuracy, and precision. Inexpensive digital data can make analysts and decision-makers ignore data quality. If subsequent management decisions are made based on poor quality data, the resultant decisions may turn out wrong. This would give decision-makers a jaded view of the usefulness of GIS. An adequate understanding of data quality issues can help decision-makers ask the right questions of analysts and avoid making decisions that are inappropriate given the data quality.

References

1. Walsh, S.J., D.R. Lightfoot, and D.R. Butler. 1987. Recognition and assessment of error in geographic information systems. Photogrammetric Eng. and Remote Sensing 53(10):1,423-1,430.

2. Openshaw, S. 1989. Learning to live with errors in spatial databases. In: Goodchild, M., and S. Gopal, eds. The accuracy of spatial databases. Bristol, PA: Taylor and Francis. pp. 263-276.

3. Thapa, K., and J. Bossler. 1992. Accuracy of spatial data used in geographic information systems. Photogrammetric Eng. and Remote Sensing 58(6):835-841.

4. Burrough, P.A. 1986. Principles of geographical information systems for land resources assessment. New York, NY: Oxford University Press.

5. Robinson, V.B., and A.U. Frank. 1985. About different kinds of uncertainty in collections of spatial data. In: Proceedings of Digital Representations of Spatial Knowledge, Washington, DC, Auto-Carto 7. Falls Church, VA: ASPRS and ACSM. pp. 440-449.

6. Nale, D.K. 1992. Do traditional map standards conflict with a GIS landbase? GIS World 5(7):50-53.

7. Smith, G.R., and D.M. Honeycutt. 1987. Geographic data uncertainty, decision-making, and the value of information. In: Proceedings of GIS '87, San Francisco, CA (October 26-30). Falls Church, VA: ASPRS and ACSM. pp. 300-312.

8. Beard, K. 1989. Dimensions of use and value of geographic information. In: Onsrud, H., H.W. Calkins, and N.J. Obermeyer, eds. In: Proceedings of use and value of geographic information: Initiative four specialist meeting, Tenants Harbor, ME (August). Vol. 89-7. Santa Barbara, CA: NCGIA.

9. National Committee for Digital Cartographic Data Standards. 1988. The proposed standards for digital cartographic data. Amer. Cartographer 15(1):11-140.

10. Chrisman, N. 1994. A vision of digital libraries for geographic information, or how I stopped trying to find the on-ramps to the information superhighway. Geo. Info. Sys. 4(4):21-24.

11. Eagan, P.D., and S.J. Ventura. 1993. Enhancing value of environmental data: Data lineage reporting. J. Environ. Eng. 119(1):5-16.

12. U.S. EPA. 1990. Information resources management policy manual. Washington, DC.

13. Lunetta, R.S., R.C. Congalton, L.K. Fenstermaker, J.R. Jenson, K.C. McGwire, and L.R. Tinney. 1991. Remote sensing and geographic information system data integration: Error sources and research issues. Photogrammetric Eng. and Remote Sensing 57(6):677-688.

14. Aronoff, S. 1989. Geographic information systems: A management perspective. Ottawa, Ontario: WDL Publications.

15. ASPRS. 1990. ASPRS accuracy standards for large-scale maps. Photogrammetric Eng. and Remote Sensing 56(7):1,068-1,070.

16. Thompson, M.M. 1987. Maps for America: Cartographic products of the U.S. Geological Survey and others, 3rd ed. Washington, DC: U.S. Geological Survey.

17. Ferguson, J. 1993. What is it we obtain from GPS? Professional Surveyor 13(4):3-4.

18. Somers, R. 1994. Is a GIS department the answer? Geo. Info. Sys. 4(6):29-32.

19. Story, M., and R.G. Congalton. 1986. Accuracy assessment: A user's perspective. Photogrammetric Eng. and Remote Sensing 52(3):397-399.

20. Chrisman, N.R. 1987. The accuracy of map overlays: A reassessment. Landscape and Urban Planning 14:427-439.

21. Bell, A.D., and M.J. Pucherelli. 1992. Audubon Lake island erosion study using aerial photography and geographic information systems. Report No. R-92-12. Denver, CO: U.S. Department of the Interior, Bureau of Reclamation.

22. McGwire, K.C. 1992. Analyst variability in labeling of unsupervised classifications. Photogrammetric Eng. and Remote Sensing 58(12):1,673-1,677.

23. Everett, J., and D.S. Simonett. 1976. Principles, concepts, and philosophical problems in remote sensing. In: Lintz, J., and D.S. Simonett, eds. Remote sensing of the environment. Reading, MA: Addison-Wesley.

24. Fortheringham, A.S. 1989. Scale-independent spatial analysis. In: Goodchild, M., and S. Gopal, eds. The accuracy of spatial databases. Bristol, PA: Taylor and Francis. pp. 221-228.

25. Kennedy, S. 1989. The small number problem and the accuracy of spatial databases. In: Goodchild, M., and S. Gopal, eds. The accuracy of spatial databases. Bristol, PA: Taylor and Francis. pp. 187-196.

26. Brown, D.A., and P.J. Gersmehl. 1987. Maintaining relational accuracy of geocoded data in environmental modeling. In: Proceedings of GIS '87, San Francisco, CA (October 26-30). Falls Church, VA: ASPRS and ACSM. pp. 266-275.

27. Fisher, P.F. 1987. The nature of soil data in GIS—error or uncertainty. In: Proceedings of the International Geographic Information Systems (IGIS) Symposium, Arlington, VA (November 15-18). Vol. III, pp. 307-318. Arlington, VA: Association of American Geographers.

28. Walsh, S.J. 1989. User considerations in landscape characterization. In: Goodchild, M., and S. Gopal, eds. The accuracy of spatial databases. Bristol, PA: Taylor and Francis. pp. 35-43.

29. Stoms, D. 1987. Reasoning with uncertainty in intelligent geographic information systems. In: Proceedings of GIS '87, San Francisco, CA (October 26-30). Falls Church, VA: ASPRS and ACSM. pp. 693-700.

30. Tosta. 1993. The data wars: Part II. Geo. Info. Sys. 3(2):22-26.

31. Anonymous. 1994. Between a rock and a hard place. Backpacker 22(8):16-18.

32. Fletcher, M., and D. Sanchez. 1994. Etched in stone: Recovering Native American rock art. GPS World 5(10):20-29.

33. Csillag, F. 1991. Resolution revisited. In: 1991 ACSM-ASPRS Annual Convention, Baltimore, MD (March 25-28). Auto-Carto 10, Volume 6. Bethesda, MD: ACSM and ASPRS. pp. 15-28.

34. Star, J.L., and J.E. Estes. 1990. Geographic information systems: An introduction. Englewood Cliffs, NJ: Prentice Hall.

35. Goodchild, M. 1993. The state of GIS for environmental problem-solving. In: Goodchild, M.R., B.O. Parks, and L.T. Steyaert, eds. Environmental modeling with GIS. New York, NY: Oxford University Press. pp. 8-15.

36. Veregin, H. 1987. Error modeling in geographic information systems: A review. Unpublished manuscript.

37. Keefer, B.J., J.L. Smith, and T.G. Gregoire. 1991. Modeling and evaluating the effects of stream mode digitizing errors on map variables. Photogrammetric Eng. and Remote Sensing 57(7):957-963.

38. Giovachino, D. 1993. How to determine the accuracy of your graphic digitizer. Geo. Info. Sys. 3(3):50-53.

39. Campbell, W.G., and D.C. Mortenson. 1989. Ensuring the quality of geographic information system data: A practical application of quality control. Photogrammetric Eng. and Remote Sensing 55(11):1,613-1,618.

40. Goodchild, M., and S. Gopal. 1989. Preface. In: Goodchild, M., and S. Gopal, eds. The accuracy of spatial databases. Bristol, PA: Taylor and Francis.

41. Goodchild, M. 1993. Data models and data quality: Problems and prospects. In: Goodchild, M.F., B.O. Parks, and L.T. Steyaert, eds. Environmental modeling with GIS. New York, NY: Oxford University Press. pp. 94-103.

42. Lester, M., and N. Chrisman. 1991. Not all slivers are skinny: A comparison of two methods for detecting positional error in categorical maps. In: Proceedings of GIS/LIS '91, Atlanta, GA (October 28-November 1). Bethesda, MD: ASPRS. pp. 648-658.

43. Newcomer, J.A., and J. Szajgin. 1984. Accumulation of thematic map errors in digital overlay analysis. Amer. Cartographer 11(1):58-62.

44. Guptill, S.C. 1989. Inclusion of accuracy data in a feature-based, object-oriented data model. In: Goodchild, M., and S. Gopal, eds. The accuracy of spatial databases. Bristol, PA: Taylor and Francis. pp. 91-98.

45. Monmonier, M. 1991. How to lie with maps. Chicago, IL: University of Chicago Press.

46. Wehde, M. 1982. Grid cell size in relation to errors in maps and inventories produced by computerized map processing. Photogrammetric Eng. and Remote Sensing 48(8):1,289-1,298.

47. Stoms, D.M. 1992. Effects of habitat map generalization in biodiversity assessment. Photogrammetric Eng. and Remote Sensing 58(11):1,587-1,591.

48. Connin, R. 1994. The trouble with data. Professional Surveyor 14(1):24-25.

49. Smith, J.L., S.P. Prisley, and R.C. Weih. 1991. Considering the effect of spatial data variability on the outcomes of forest management decisions. In: Proceedings of GIS/LIS '91, Atlanta, GA (October 28-November 1). Bethesda, MD: ASPRS. pp. 286-292.

50. Felleman, J., and C.B. Griffin. 1990. The role of error in GIS-based viewshed determination: A problem analysis. Report No. IEPP-90-2. Syracuse, NY: Institute of Environmental Policy and Planning, State University of New York, College of Environmental Science and Forestry.

51. Dutton, G. 1991. Improving spatial analysis in GIS environments. In: 1991 ACSM-ASPRS Annual Convention, Baltimore, MD (March 25-28). Auto-Carto 10, Volume 6. Bethesda, MD: ACSM and ASPRS. pp. 168-185.

52. Felleman, J. 1992. Letter to the editor. Geo. Info. Sys. 2(8):12-13.

53. Berry, J.K. 1994. The this, that, there rule. GIS World 7(7):22.

54. Hunter, G.J., and M.F. Goodchild. 1993. Managing uncertainty in spatial databases: Putting theory into practice. In: Proceedings of the annual meeting of the Urban and Regional Information Systems Association, Atlanta, GA (July 25-29). Washington, DC: Urban and Regional Information Systems Association. pp. 1-15.

Expedition of Water-Surface-Profile Computations Using GIS

Ralph J. Haefner, K. Scott Jackson, and James M. Sherwood
Water Resources Division, U.S. Geological Survey, Columbus, Ohio

Abstract

Water-surface profiles computed by use of a step-backwater model such as Water Surface PROfile (WSPRO) are frequently used in insurance studies, highway design, and development planning to delineate flood boundaries. The WSPRO model requires input of horizontal and vertical coordinate data that define cross-sectional river-channel geometry. Cross-sectional and other hydraulic data are manually coded into the WSPRO model, a labor-intensive procedure. For each cross section, output from the model assists in approximating the flood boundaries and high-water elevations of floods with specific recurrence intervals (for example, 100-year or 500-year). The flood-boundary locations along a series of cross sections are connected to delineate the flood-prone areas for the selected recurrence intervals.

To expedite the data collection and coding tasks required for modeling, the geographic information system (GIS), ARC/INFO, was used to manipulate and process digital data supplied in AutoCAD drawing interchange file (DXF) format. The DXF files, which were derived from aerial photographs, included 2-foot elevation data along topographic contours with +0.5-foot resolution and the outlines of stream channels. Cross-section lines, located according to standard step-backwater criteria, were digitized across the valleys. A three-dimensional surface was generated from the 2-foot contours by use of the GIS software, and the digitized section lines were overlain on this surface. GIS calculated the intersections of contour lines and cross-section lines, which provided most of the required cross-sectional geometry data for input to the WSPRO model.

Most of the data collection and coding processes were automated, significantly reducing labor costs and human error. In addition, maps at various scales can be easily produced as needed after digitizing the flood-prone areas from the WSPRO model into GIS.

Introduction and Problem Statement

Losses due to flood damage generally cost the American public hundreds of millions of dollars annually. In 1968, the National Flood Insurance Act established the National Flood Insurance Program (NFIP) to help reduce the cost to the public and provide a framework to help reduce future losses. The Federal Emergency Management Agency (FEMA) administers the NFIP. As listed in Mrazik and Kinberg (1), the major objectives of the NFIP are to:

- Make nationwide flood insurance available to all communities subject to periodic flooding.
- Guide future development, where practical, away from flood-prone areas.
- Encourage state and local governments to make appropriate land use adjustments to restrict development of land that is subject to flood damage.
- Establish a cooperative program involving the federal government and the private insurance industry.
- Encourage lending institutions, as a matter of national policy, to assist in furthering program objectives.
- Authorize the continuing studies of flood hazards.

Studies of flood-prone areas typically involve using step-backwater computer algorithms (digital models) to estimate river water-surface profile elevations and flood-inundation patterns along the topography of the river and its overbanks. FEMA recognizes the U.S. Geological Survey's (USGS's) step-backwater model, Water Surface PROfile (WSPRO), as a suitable computer model for use in flood insurance studies (2, 3). Basic data input for step-backwater models includes:

- Estimates of flood discharge and initial water-surface elevations.
- Stream cross-sectional geometry.
- Roughness coefficients for cross sections.

- Contracted opening geometry if bridges or culverts are located along the study reach.

Obtaining meaningful model results typically requires numerous stream cross sections referenced to a common elevation datum along a stream reach. The data-collection efforts to obtain these cross-sectional data require costly, labor-intensive fieldwork. Study efforts along lengthy stream reaches may, however, involve the generation of a contour map using aerial photogrammetric mapping techniques. Processing the spatial data may still require extensive labor to extract the cross-sectional data needed for the WSPRO model.

The development of geographic information systems (GIS) technology has greatly enhanced analyses of spatial data such as topography. In an effort to improve the quality of mapping and delineation of flood-prone areas in Summit County, Ohio, the USGS developed a method of using a GIS as a pre- and postprocessor of the input and output data for the WSPRO model. This paper describes the steps the USGS used to develop this interface and discusses some difficulties encountered during the process.

Approach

Several steps were taken that resulted in the delineation of a flood-prone area in Summit County, Ohio. These steps are shown in a flow chart (see Figure 1) and described below.

Data were obtained for this study via aerial photography during April 1990. These data include mappable features at the given scale including topography at 2-foot contour intervals, stream boundaries, roads, and buildings. The data are estimated to be vertically accurate to +0.5 feet. The data were put into AutoCAD and were prepared for delivery to the USGS on 3.5-inch floppy disks in AutoCAD drawing interchange file (DXF) ASCII format. ARC/INFO was used to convert the DXF file into two separate data layers containing only the topography and traces of stream banks within the study area.

A three-dimensional surface was generated from the topographic data using the ARC/INFO software package Triangulated Irregular Network (TIN). Cross-section lines were digitized over the topography data layer. The cross sections were placed according to standard step-backwater criteria (4) and were generally:

- Perpendicular to stream flow
- At major breaks in streambed profiles
- At minimum and maximum cross-sectional areas
- At major changes in stream conveyance
- Spaced about one cross-section width apart

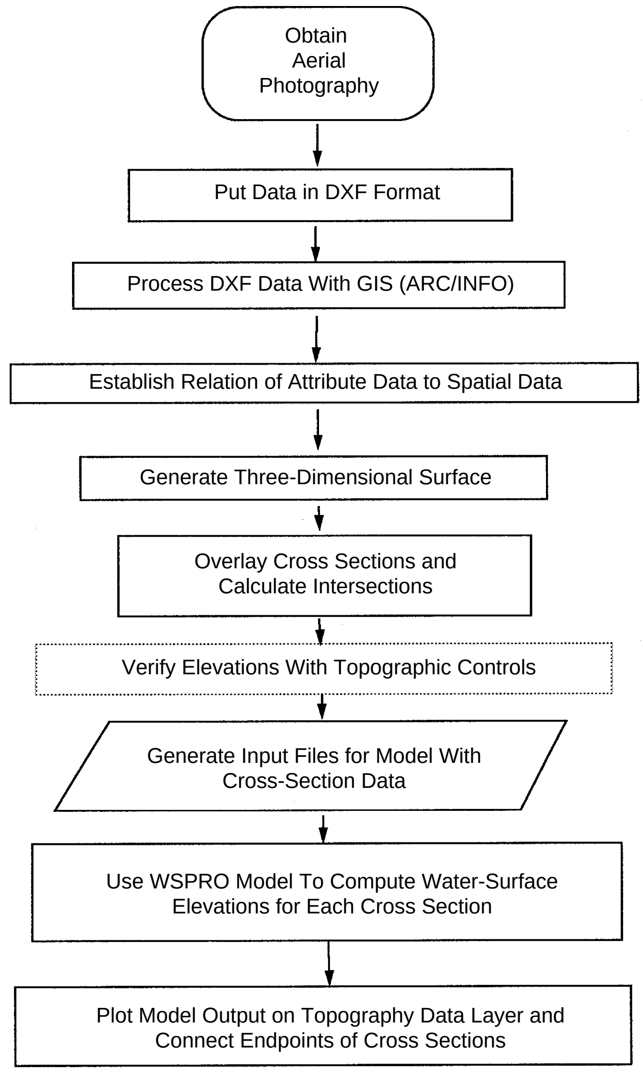


Figure 1. Flowchart of data conversion and processing for use in the Water Surface PROfile.

The cross-section lines were then overlaid on the three-dimensional surface of topography, and GIS calculated the intersections of the contour lines and cross sections. The locations and elevations of these intersections were output as an ASCII file and slightly modified for input into the WSPRO model.

These GIS data were used along with the aforementioned required data as input to the WSPRO model. Input for the model included estimates of the 100-year flood discharge (5), stream cross-sectional geometry (supplied by this work), and estimates of roughness coefficients for cross sections. The WSPRO model was then run, providing output in the form of water-surface elevations at specific distances along section lines corresponding to the simulated elevation of a 100-year flood.

Points corresponding to the flood elevations along the cross-section lines were plotted on the topography data layer and were connected manually (to delineate flood boundaries) by interpolating the elevations with respect to adjacent contours. A polygon of the flood "surface" was generated and drawn on a map (see Figure 2).

Results

The supplied topographic data were of sufficient quality and resolution to substitute for field-surveyed elevations; however, field surveys to verify the elevations along the cross sections would augment this quality control process (see Figure 1). Typically, a crew of two individuals may take up to 4 days to survey and reduce the field data for the study area chosen for this study. Because aerial photography is commonly substituted for land surveying, the most significant effort and source of error may come from manually extracting elevations and distances along cross sections for input into the WSPRO

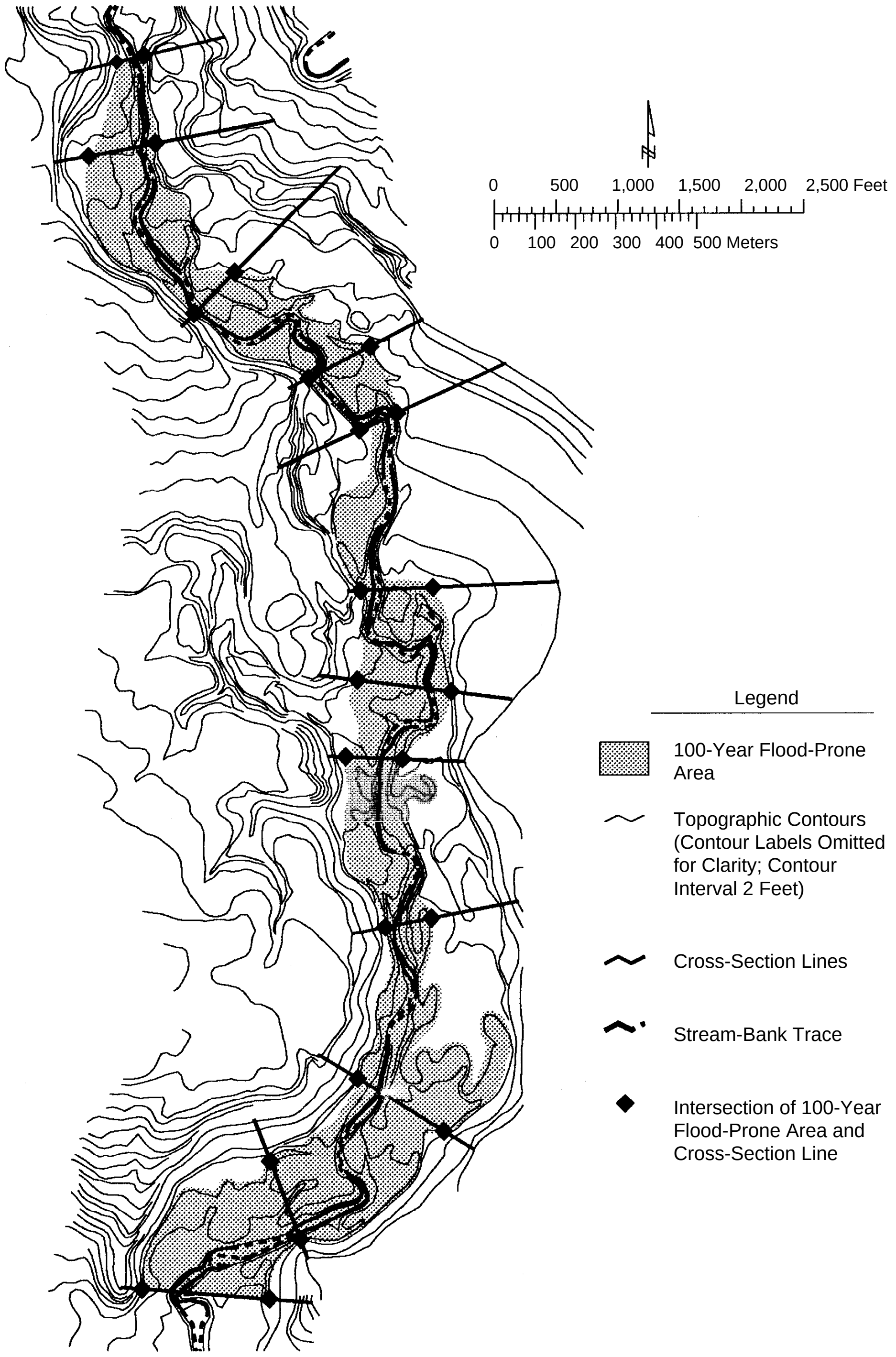


Figure 2. Watershed showing delineation of 100-year flood-prone area.

model. Initial development of the method to use GIS for this analysis took approximately 1 week to refine; however, future analyses would probably only take one person 1 day to perform. This represents a significant cost savings. Additionally, reducing the amount of human-induced error can substantially improve the reliability and accuracy of the computer-generated flood-prone area data.

Because topography, stream traces, and other features are supplied in the DXF file, these data can easily be brought into GIS. Maps can be made that show these features in relation to the predicted flood-prone area. Maps showing a variety of features can be produced at any scale, with accuracy limited only by the accuracy of the source scale. Additionally, GIS can calculate the intersection of map features that may lie within the flood-prone area, such as buildings that may contain hazardous materials. GIS can also overlay land use data layers within the flood-prone area to define areas that should not be developed or that have already been overdeveloped in accordance with the aforementioned NFIP objectives.

FEMA now requests that future flood-study mapping be completed using GIS format, a common goal that both the USGS and FEMA are working toward. These data are important to land planners, flood-plain regulators, and insurance companies that rely on accurate estimates of flood-prone areas. By increasing the accessibility of the data by using GIS, we can substantially improve our ability to analyze spatial data efficiently.

Problems Encountered

Problems using data supplied in DXF format in conjunction with GIS resulted primarily from the fact that the DXF data were prepared for the purpose of making a topographic map, not a GIS data layer. The contour lines were segmented; that is, where ends of segments met, they were not physically connected to form a topologically viable data layer. The data layer needed to be edited because GIS requires topology for spatial-data processing. Additionally, in areas where the topographic gradient was particularly steep, contour lines were omitted. In both cases, an attempt was made to allow GIS to establish a physical connection of contour lines, but subsequent manual interpolation was also required. This may have introduced error into the data set. If future work requires the use of DXF data, the request for data should specifically state that all topographic contours be continuous.

AutoCAD stores data differently from GIS, so a relationship needed to be established between the data file containing elevations and the data file associated with the lines that make up the topography data layer. Several lines from the DXF file did not have any data associated with them, thus necessitating the addition of contour elevation data by context with the adjacent contours that did have data. This step may also have introduced errors, but quality-control measures to verify the topographic contours and contour elevations could help to minimize these errors.

Output from the WSPRO model is in the form of a series of points along cross sections that were connected by manual interpolation. This step also may introduce some error, but the same process must be performed when not using GIS.

Conclusions

This report documents an example of how GIS can be used to facilitate step-backwater modeling of flood-prone areas. The results of the study show that significant savings may be expected in the form of reduced labor requirements. Furthermore, FEMA now requires the use of GIS to conduct flood-study mapping, thus providing a means to conduct additional spatial analyses more efficiently. As aerial photography and GIS technology improve, although additional sources of error may arise, the overall accuracy, reliability, and reproducibility of the model input and results should also improve.

References

1. Mrazik, B.R., and H.A. Kinberg. 1991. National flood insurance program: Twenty years of progress toward decreasing nationwide flood losses. Water Supply Paper 2375. U.S. Geological Survey.

2. Shearman, J.O., W.H. Kirby, V.R. Schneider, and H.N. Flippo. 1986. Bridge waterways analysis model. Research Report FHWA/RD-86/108. Federal Highway Administration.

3. Shearman, J.O. 1990. Users manual for WSPRO, a computer model for water-surface profile computations. FHWA-IP-89-027. Federal Highway Administration.

4. Davidian, J. 1984. Computation of water-surface profiles in open channels: Techniques of water-resources investigations of the United States Geological Survey. In: Applications of Hydraulics, Vol. 3.

5. Koltun, G.F., and J.W. Roberts. 1990. Techniques for estimating flood-peak discharges of rural unregulated streams in Ohio. Investigations Report 89-4126. U.S. Geological Survey, Water Resources.

Reporting on the Development of an Environmental GIS Application – Wetlands Restoration in the Central Valley of California

David T. Hansen[1]

Introduction

Appropriate documentation of information used in addressing environmental problems is a common issue. This paper focuses on information supporting geographic information system (GIS) applications on wetlands in the Central Valley of California. It identifies the role GIS played in the preparation of a report to Congress on wetlands and water supply in the Central Valley. It also describes the role that GIS is playing in the dissemination of information on wetlands and potential wetland habitat development to communities and groups in the Central Valley. Information or data documentation are major elements supporting this data and these GIS applications. For GIS data, this information is commonly referred to as metadata in conformance with the "Content Standards for Digital Geospatial Metadata (FGDC, 1994 and 1998).

GIS is an integrative technology. It typically has involved specialists from a variety of disciplines who associate and integrate different data sets into a spatially referenced system. The development of GIS data typically follows a series of steps:

1. Based on a conceptual model of the environmental issue, geographic features or data of interest are identified.
2. These features are mapped or the data is spatially referenced in a map coordinate system.
3. The mapped features or data are digitally captured and processed in a software system to create a GIS data theme containing a graphical representation of the feature and associated information as attributes.
4. This digital data are reviewed and assessed as to fitness for the application.
5. The digital data is then ready for display, query, and analysis with other digital data for that geographic area.

[1] U.S. Bureau of Reclamation, Mid Pacific Region, Dept. of the Interior, 2800 Cottage Way, Sacramento, CA 95825-1898; (916) 978-5268; (916) 978-5290 FAX; dhansen@mp.usbr.gov

6. The resulting data output from GIS analysis then is again reviewed and assessed against the requirements for the application.
7.

Increased speed and capacity of computer systems and the development of graphical user interfaces (GUI) have brought GIS and geospatial analysis to the computer desktop. Technologies such as global positioning systems (GPS), remote sensing, and scanning technologies have combined some of these steps and shortened the time required for digital data development. Many of these steps which were essentially GIS back office operations can now be performed by managers and increasingly the public. This has enabled managers and the public to access and use data in a map or geographic format to address environmental problems.

In the case of the Central Valley Joint Venture, desktop GIS has permitted the direct involvement of managers and partners in data development and GIS analysis. Desktop GIS applications developed as part of this program permit analysis of the data at the local level with community groups. Many GIS applications are robust enough to permit the loading and use of locally developed data in modeling environmental systems. Providing information supporting this data and the GIS application are critical in the dissemination of the data and application down to the local level. This case study will follow these steps identifying the information needed by the Central Valley Joint Venture partners for evaluating the data for application and use. This information is metadata and terms used will follow the elements of the "Content Standards for Digital Geospatial Metadata" (FGDC, 1994 revised 1998). However, this represents only a subset of the elements in these standards. The relationship between FGDC metadata elements and GIS data development are further described in a draft guides for documenting the development of a GIS data theme (Hansen, 1998) and reporting metadata for data management, data catalogs, and data transfer (Hansen, 1998).

Conceptual Model – Identification of Data Requirements

The Central Valley Joint Venture was established as part of the North American Waterfowl Management Plan signed by Canada and the United States in 1986. In 1990, the Central Valley Joint Venture issued an implementation plan (CDFG, 1990) for wetland habitat restoration and enhancement. Wetland and adjacent upland habitat are important wintering areas for waterfowl in the Pacific Flyway. By the mid 1980's, waterfowl populations were approaching 30 percent of long term averages. Much of this decline is associated with the loss of wetland habitat since the turn of the century. The Central Valley is a semi arid area. Successful wetland habitat restoration in the valley requires a dependable and adequate water supply. Water supply for wetlands must be balanced against environmental requirements as well as water requirements use for agriculture and municipalities. The Central Valley is a major agricultural producing area for California and the nation. Surface water flowing primarily from snow melt in the Sierra Nevada mountains provides approximately 70 percent of the water used for agriculture and municipalities in the State. The plan attempts to balance water requirements for wetlands, agriculture and municipal use without impairing the supply for other aquatic and terrestrial species.

In 1995, the Central Valley Joint Venture partners prepared a statement of work for a report to Congress based on the implementation plan (USFWS, 1995). Main topics for this report are the identification of methods for improving the reliability of water supply for existing private wetlands and identification of water requirements for an additional 120,000 acres to be restored to wetland habitat in the Central Valley. As part of the statement of work, the cooperator in the report was to incorporate any appropriate data into a desktop GIS.

The statement of work identified a variety of information needed for the report. Much of this was geographic in nature such as the location of existing wetlands, lands under wetland or conservation easements, and lands suitable for wetland habitat development. Information was also identified that was not strictly geographic in nature such as the identification of constraints affecting the protection or restoration of wetlands or the reliability of water supply for wetlands. The statement of work recognized that not all information required to address issues in the report were suitable or could be developed in time for analysis as GIS data themes. It recognized that other tools would be needed to address some issues for the report to Congress.

As part of the statement of work, metadata on the collected data was to be provided by the contractor. The statement of work, itself, provided some of the initial information called for in the “Content Standards for Digital Geospatial Metadata” (FGDC, 1994 and 1998). The statement of work and the implementation plan prepared in 1990 form the basis of the conceptual model for identifying data and information requirements. Data themes were identified and the purpose for collecting this information. The time period of data content, target source scale (1:24,000), and GIS data format and system were identified. Information not explicitly identified in the statement of work included actual data sources, map coordinate system, and method of coordinate control.

The statement of work was a working document which was revisited as information was developed to adjust for changes in the availability of data for the report.

Mapping, Digital Capture, and Database Development

Mapping, digital capture, and database development are distinct process steps. Increasing with new technologies, these steps occur concurrently. Such a theme for the partners in the Joint Venture was the identification of existing wetland habitat on the valley floor. The program participated in a joint project with other organizations which identified wetlands, riparian habitat, and other land use (DU, 1997). This GIS theme achieved a minimum resolution for wetlands at about 0.8 hectares (2 acres) using remote sensing techniques. Many of the other GIS data themes were already mapped independently of the program requirements. The cooperator following the statement of work digitally captured the features represented on these maps, merged separate sources for a particular theme together, and constructed databases of attributes for the digital features.

Focus areas or areas for analysis is another GIS theme that required definition, mapping, and digital capture for the Joint Venture partners. The floor of the Central Valley covers approximately 4 million hectares (10 million acres). This was too large an area for data development within the time constraints of the program. One of the initial tasks of Joint Venture was to narrow the focus of data collection efforts to smaller areas within the Valley floor. The wetland habitat GIS theme as it was being developed and the personal knowledge of the Joint Venture Partners assisted in identifying focus areas for intensive data collection efforts. The resulting focus areas represent approximately 0.8 million hectares (2 million acres) of the valley floor. These areas include virtually all the areas of existing Public and private wetlands.

For the partners in the Joint Venture Program, digital capture and database development represented the black box phase of GIS data development. Key information from this stage for the Joint Venture partners included the sources of data for the GIS themes, definition of criteria used in digitally capturing the features, database definitions, and criteria or rules for classifying the attributes of those features. Since the area of interest or focus areas for Joint Venture cover such a broad area, multiple sources of data were required for each GIS theme. Often, these sources represented different time periods for mapping. Different sources also raised issues of consistency in mapping criteria and in attributes identified for the mapped features.

Evaluation of Digital Data for Application and Use

Using the desktop GIS, information collected and digitally captured as a GIS data theme could be reviewed directly by the Joint Venture partners. This evaluation occurred repeatedly during data development. This was helpful to the Joint Venture partners as well as the GIS data developer. Managers as well as staff could directly evaluate the data against the issues identified in the statement of work and their own knowledge of the area. GIS analysis could be interactively performed to evaluate the various data themes for addressing issues required in the report. This review identified GIS themes that were not useful in addressing the issues. The Joint Venture could then focus attention on other methods for developing information to address those issues. Surrogates to represent information that could not be directly represented in GIS could be addressed and evaluated. The Joint Venture partners could visually review:

- Extent of coverage of a particular GIS theme for the areas of interest,
- Data gaps between GIS themes for the same area,
- Attributes carried by the GIS themes and the definitions for those attributes, and
- Attributes relationship to the issues identified for the report.

Information that could not be directly displayed in GIS were:

- Sources used to construct each GIS theme,
- Time periods represented by the GIS theme,
- Consistency of a GIS theme for all areas of the Valley, and
- Criteria used in classifying the attributes of the GIS themes.

This metadata was not available at this stage to the Joint Venture partners.

This information represents a subset of information identified in the “Content Standards”. Although available to the data developer, this information was not in a form easily provided to the data users. While the data provider had some experience in spatial data capture, the provider had little experience in working directly with data users and in recognizing information that they might need. The data developer was deferring metadata compilation until the end of the data development. The “Content Standards” had been recently adopted by FGDC and the data provider had little experience in addressing and applying the standards. This hindered the evaluation of the GIS data by the Joint Venture partners.

GIS Display, Query, and Analysis and Evaluation of Analysis

For the Joint Venture program, GIS analysis and evaluation occurred concurrently. With the GIS desktop application, the partners in the Joint Venture program were involved directly in applying GIS to address some of the issues for the report. This included display, query and reporting of the following information for the report:

- Location and extent of public managed wetlands,
- Private lands under easements for wetland habitat and conservation,
- Private lands managed for waterfowl or duck clubs, and
- Major water supply agencies for those lands.

The Joint Venture partners ran a variety of different scenarios using the desktop GIS to identify lands suitable for wetland habitat restoration. These scenarios were based on criteria defined and run by the Joint Venture partners at their meetings. These scenarios were primarily based on the following GIS data themes:

- Soil characteristics suitable for wetland habitat development,
- Land use,
- Existing Publicly managed wetlands,
- Lands with easements for wetland habitat or wildlife conservation, and
- Private wetlands.

A variety of other GIS data themes were available for display with these themes for review with the results of the scenarios.

The Joint Venture partners could evaluate the results of the scenarios for issues required in the report. Criteria for the scenarios could be evaluated and adjusted to meet specific needs for different areas of the Central Valley. At the time of these meetings, the desktop GIS was not exploited to the full in documenting the various scenarios that the participants posed. Documentation of the final scenarios do form the basis for the some of information reported to Congress.

Application Development – Public Outreach

Several of these GIS data themes and other data developed in concert with the Joint Venture program led to the development of a stand alone application. This application contains several GIS data themes and software for assessing the suitability of areas for wetland habitat restoration (Ducks Unlimited, 1998). The application is basically a modeling tool for ranking and weighting various GIS themes as to their suitability for developing and maintaining wetland habitat. The user selects the themes that they want to use in the analysis, weights or assigns values contained in the attribute table, and ranks the theme on basis of its importance for wetland habitat. The application converts the themes into a raster data structure and combines the raster data sets into new data set or surface whose cell values represent suitability of that cell for wetland habitat. The user selects the cell size for processing. This application has been issued on a CD-ROM containing GIS data sets, metadata, and some sample scenarios. The data on the CD requires desktop GIS software, but the application is open in the sense that other spatial data can be loaded and used in the analysis.

This tool was developed not only to assist offices of the individual partners in the Joint Venture program but also as an outreach tool to local community groups and for education. This application provides the opportunity to address issues at the local level. Locally developed as well as other data can be used in the application. Metadata describes the GIS data themes and a user guide has been prepared for using the application. While metadata is included with the GIS data themes, it is expected that this information will not address all questions posed by the users of the data. Describing the GIS application following the “Content Standards of Digital Geospatial Metadata” is beyond the scope of these standards. As is true for most guides, the guide for the application focuses on the mechanics of operating the application. The guide probably does not address all questions on how the various themes could be ranked or weighted.

The local user is responsible for evaluating the results of running the analysis. It is up to the user to decide on the cell size for appropriate for the selected GIS themes. The user evaluates the resulting output surface against the cell size used, themes selected, and the values assigned to the themes. The guide can not answer all questions on what spatial analysis is, how the model process actually runs, or what uncertainty can be assigned to the results of analysis. To some users, it will appear to be another black box. Access to additional information can be provided by citations to other documents and providing contact information for individuals and organizations involved with the data or the application.

Summary

This has been a review of the information needed at different stages in the development of GIS data and applications for the Central Valley Joint Venture Program of California. Partners in Joint Venture were able to evaluate the data, to run spatial analysis, and to review results from multiple scenarios. Issues were identified that could not be adequately addressed as GIS data themes. Resources could then be focused on other methods to develop this information. Some of these GIS themes along with other GIS themes supported the development of an independent GIS application. With this application, community groups and the public can develop their own scenarios for identifying wetlands at the desk top level.

The development of relatively easy to use desktop GIS software provides the opportunity for final data users to be directly involved in data evaluation and spatial data analysis. New and improved technologies have compressed the steps and time required for spatial data development and the incorporation of that data into GIS applications. Environmental GIS applications are increasingly available for use by local communities and the public. Many are robust enough to permit the loading of locally developed data. This represents current trends as described in “The Future of Spatial Data and Society” (NRC, 1997). Under this trend, spatial data proliferate rapidly and the tools to use this data are widely available. Under this scenario, key issues are education or training in using spatial data and access to information to evaluate and apply the data. Lack of adequate information can to lead to increased uncertainty in the use of GIS data and increased litigation.

Joint Venture partners were hindered in their evaluation and application of the GIS data themes because some information was not available at their meetings. This information was not available because the data developer lacked experience in working concurrently with users in developing and applying data. As these applications are reaching down to community groups and the local level, the GIS users are further removed from data producers and application developers. To effectively run these applications, GIS data users can be expected to need more information than can be easily contained in our existing metadata descriptions or in a user guide.

Documentation during this development is critical for the effective application and use of the data or the application. The “Content Standards for Digital Geospatial Metadata” (FGDC, 1994, revised 1998) provides a lexicon of commonly accepted terms for describing GIS data. It identifies what metadata should accompany the digital data when it has been completed and transferred. It does not address how metadata is to be developed or presented during data development. These standards are a comprehensive list of elements of metadata for geospatial data but they are not exhaustive. These standards do not address the development of GIS applications for use at the local level.

We as GIS data or application developers need to be directly involved in the application and development of these standards. The involvement of data developers and application developers is needed for guides on the development and use of GIS data and applications. The Environmental Protection Agency has been an active participant in the development of many standards related to environmental quality (Johnson, 1996). For environmental data, there are a variety of standards and guides often specific to a particular discipline supporting the data collection and modeling efforts. The “Content Standards for National Biological Information Infrastructure Metadata” (USGS-BRD, 1995) has been issued for describing biological data. The process of developing of GIS environmental applications is similar to the development of a ground-water model to a site specific application (ASTM D5979) and the steps followed in environmental site characterization (ASTM D5730). For water quality and monitoring, there are a host of commonly adopted standards. This paper is offered to continue discussion and to encourage involvement in the development of guides for describing environmental data and applications using GIS.

Bibliography

ASTM, 1996, "D 5979 – *Guide for Conceptualization and Characterization of Ground –Water Flow Systems*" ASTM, 100 Barr Harbor Drive, West Conshohocken PA 19428-2959.

ASTM, 1996, "D 5730 – *Guide for Site Characterization for Environmental Purposes with Emphasis on Soil, Rock, the Vadose Zone and Ground Water*" ASTM, 100 Barr Harbor Drive, West Conshohocken PA 19428-2959.

ASTM, 1996, "D 5714 – *Specification for Content of Digital Geospatial Metadata*" ASTM, 100 Barr Harbor Drive, West Conshohocken PA 19428-2959.

California Department of Fish and Game, California Waterfowl Association, Ducks Unlimited, February, 1990, "*Central Valley Habitat Joint Venture Implementation Plan*", Sacramento California (CDFG, 1990).

Ducks Unlimited, January, 1997, "*California Wetland and Riparian Geographic Information System Project*", Prepared for California Department of Fish and Game, California Wildlife Conservation Board, U.S. Bureau of Reclamation; Sacramento, California (DU, 1997).

Ducks Unlimited, November 1998, "*Central Valley Project Improvement Act GIS Model Version 1.0, User's Guide*", Sacramento, California (DU, 1998).

Federal Geographic Data Committee, June 1994 (revised 1998*), "Content Standards for Digital Geospatial Metadata"*, Washington, D.C. (FGDC 1994 and 1998).

Hansen, David T., 1998, *"Guide for Documenting the Development of Geospatial Data for Site Investigations*", Draft for review by ASTM Technical Sub Committees of D18, 100 Barr Harbor Drive, West Conshohocken PA 19428-2959 (Hansen, 1998)

Hansen, David T., 1998, *"Guide on Reporting Geospatial Metadata for On-line Data Management, Data Catalogs, and for Data Transfers*", Draft for review by ASTM Technical Sub Committees of D18, 100 Barr Harbor Drive, West Conshohocken PA 19428-2959 (Hansen, 1998)

Johnson, A. I., 1996, "*The Accelerated Development of Standards for Environmental Data Collection*", Sampling and Environmental Nedia, ASTM STP 1282, James Howard Morgan Ed., ASTM, 100 Barr Harbor Drive, West Conshohocken, PA 19428-2959 (Johnson, 1998).

Mapping Sciences Committee, National Research Council, 1997, "*The Future of Spatial Data and Society: Summary of a Workshop*", National Academy Press, Washington, D.C. (NRC, 1997).

U.S. Fish and Wildlife Service, 1995, "*Scope of Investigations and Plan of Work under Cooperative Agreement for Wetland Water Supply Investigations*", Sacramento California. (USFWS, 1995).

U.S. Geological Survey, Biological Resources Division, 1995, "*Content Standard for National Biological Information Infrastructure Metadata*", formerly National Biological Service, Department of Interior, Washington, D.C. (USBS-BRD, 1995).

Using a Geographic Information Systems Application to Implement Risk Based Decisions in Corrective Action

Lesley Hay Wilson, P.E., Andrew P. Romanek, and David R. Maidment, Ph.D., P.E.,
The University of Texas at Austin
James R. Rocco, Sage Risk Solutions LLC

Abstract

The implementation of site-wide corrective action using risk-based decision making at large and complex industrial, energy or defense facilities presents a number of challenges. To address these challenges, a spatial environmental risk assessment methodology has been developed by connecting Geographic Information Systems (GIS), relational databases and spreadsheets. The methodology is based on a description of the facility and a site conceptual model. A case study site with multiple potential sources, transport mechanisms and receptors has been used to evaluate the methodology. The case study facility is an approximately 300-acre crude oil refinery and petroleum products terminal that has operated since the early 1900's. This paper will discuss the development of two key elements of the spatial environmental risk assessment, the Digital Facility Description and the Spatial Site Conceptual Model using the GIS application.

Introduction

Risk-based decision making provides a mechanism to determine the necessary and cost-efficient strategies for protection of human health and the environment. It is an iterative process that begins with a planning phase that incorporates risk management decisions with a site conceptual model. The process then proceeds to an evaluation phase where data collection, and fate and transport analysis, provide the basis for evaluation of the exposure pathways identified in the site conceptual model. Finally, the process ends with a decision phase in which the plan is compared to the results of the analysis to determine whether further evaluation or remedial action is warranted or if "no further action" is appropriate. Risk-based decision making requires the interrelationship of three basic processes: risk assessment, risk management and risk communication.

Risk Assessment

Risk assessment is a process that quantifies the potential for adverse effects to human health and the environment caused by an exposure to a chemical of concern released to the environment. Where there are no current or potential exposures to a chemical of concern, or where the concentration of a chemical of concern is not harmful to human health or the environment then, the risk assessment will conclude that there is no unacceptable risk. Risk assessment is accomplished by collecting information to construct exposure hypotheses for a chemical of concern, or a group of chemicals of concern, and evaluating those hypotheses to determine the potential for adverse effects from human or ecological exposures to chemicals of concern in the environment. This process is based on the National Academy of Sciences Risk Assessment paradigm (NAS, 1983).

Effective risk assessment is based on several important activities. First, a comprehensive site conceptual model is needed to provide the working hypothesis for a site. The site conceptual model is the understanding of the potential exposure pathways based on the chemical characteristics and the physical setting of the site. In order for the site conceptual model to be comprehensive, it needs to identify all of the potential exposure pathways and be updated, as new information becomes available. It provides the mechanism for determining the necessity and scope of data collection, and a template for evaluating exposure pathway completeness. Second, effective data collection is needed to evaluate the exposure pathways identified in the site conceptual model. Data collection can be both qualitative (e.g., location of source areas, historical release information) or quantitative (e.g., concentration of chemicals of concern in environmental media, hydrogeological characteristics). It requires spatially defined and relationally organized property characteristics such as physical features, and information related to chemical and media characteristics such as analytical results and hydrogeological information. Data collection needs to be focused on developing and updating the site conceptual model, evaluating exposure pathways, determining appropriate initial response actions and comparing site conditions to the corrective action goals. Third, the value of the information that is being collected must be considered. Only the quantity and quality of data necessary to provide a sound basis for the decisions to be made should be collected. Collecting data for the sake of more data is not necessary or cost-effective. Data collection must consider the use of the data to be collected and the potential for that additional data to change the decisions that will be made. Fourth, the fate and transport of chemicals of concern in the environment needs to be considered when evaluating exposure pathways. However, the results of fate and transport analysis must be confirmed through the collection of empirical data. Finally, "no further action" is not always the appropriate result of a risk assessment. Interim remedial action, remedial action as well as further evaluation are alternatives to be considered.

Risk Management

Risk management decisions are necessary to support site-specific determinations that are protective of human health and the environment and to provide a means for similar decisions to be made for similar circumstances. Many risk management decisions rely on the application of scientific methodologies such as the determination of the chemicals of concern to be considered in the risk assessment; the appropriate toxicity factors for the chemicals of concern; and the appropriate data quality and quantity. Other risk management decisions rely on non-scientific factors for definition such as the determination of a process for stakeholder involvement; an approach for ground water resource and use; and a consistent set of criteria for the objective comparison of alternatives. Clearly, a risk-based decision cannot be made without having defined the appropriate risk management decisions.

Risk Communication

Risk-based decision making requires the identification and involvement of the individuals, organizations and other entities that are directly affected by the corrective action process, commonly referred to as the stakeholders. Since risk management decisions are necessary to support determinations that are protective of human health and the environment and require the consideration of a combination of scientific, social, political, personal and economic factors, it is critical that the risk management decisions are acceptable to most if not all of the stakeholders. In addition, the application of the risk management decisions within the risk assessment process must be clearly understood and accepted by all of the stakeholders. Therefore, early, effective and regular risk communication is critical to the successful implementation of risk-based decision making.

Risk communication, however, has been the most overlooked component of risk-based decision making and all too often undertaken after the evaluation has been completed and the decisions have been made rather than as part of the process. When applying risk-based decision making there will typically be a number of alternatives for solving or addressing the environmental condition of a property. Each alternative will have characteristics that define its benefits and its risks to the stakeholders. There will also be stakeholders with differing interests, objectives, and levels of knowledge. However, the perception of these benefits and risks may vary among the stakeholders and it cannot be assumed that all stakeholders will have the knowledge or background to effectively participate in the process. Therefore, effective risk communication must make information concerning a corrective action easily accessible to the stakeholders, provide a mechanism for the stakeholders to visualize the results and be interactive to allow for participation in the process by the stakeholders.

Spatial Environmental Risk Assessment

In recent years, the use of risk assessment and risk-based decision making in environmental management has gained increasing attention (Rocco and Hay Wilson, 1998, Washburn and Edelmann, 1998). However, there are a number of significant challenges to the practical application of risk-based decision making at large, complex industrial, energy or defense facilities. For these types of facilities, the implementation of a risk-based approach has been difficult and a practical methodology has not been demonstrated (Hay Wilson *et al.,* 1998).

The most significant challenges arise from the multiple potential sources, multiple chemicals of concern and multiple potential receptors. In these large facilities, there can literally be hundreds of potential exposure pathways to analyze. In practice, at complex facilities, sources are evaluated individually or in small groups. As an example, the corrective action program under the Resource Conservation and Recovery Act (RCRA) encourages this piecemeal evaluation through its focus on individual solid waste management units (SWMU). In general, the implementation of program-specific (e.g., air, water, waste) regulations by the Environmental Protection Agency (EPA) has also perpetuated this non-holistic approach. The same is true for the regulation of Department of Defense and Department of Energy facilities. Individual areas of a facility are studied using a straightforward process to analyze exposure pathways for each source-receptor pair. However, there is typically no attempt to understand the interaction of all of the sources and pathways on facility-wide risks or the affects of these multiple sources and pathways on the environmental management decisions. It has also been the case in the past that the goals for corrective action projects were based on very low to non-detect concentrations of chemicals of concern, so the need to understand all of the potential exposures was not as great. In addition, often many individuals are involved, over a number of years, at a significant cost, in the calculations of the risks for each of the different areas for a facility. Many of these facilities are also regulated under different regulatory programs, with different regulatory agencies and no one investigator examines all of the results (Hay Wilson, 1998).

The availability of cost-effective computing power and information systems applications can provide the foundation for the development of computer-based systems to manage information and perform calculations for large numbers of pathways. In particular, the information processing capabilities of Geographic Information Systems (GIS), relational databases, spreadsheets and other computer code-based models can provide a mechanism to construct a methodology to implement risk assessment at large, complex facilities (Hay Wilson, 1998). In the spatial environmental risk assessment methodology the ESRI software ArcView® is being used for the GIS functions (ESRI, 1998). The Microsoft Office products Excel® and Access® are being used for the spreadsheet and relational database components, respectively (Microsoft, 1997).

Building the Digital Facility Description

In order to conduct the analyses required for making risk-based decisions and to provide a common point of reference for all of the stakeholders, an understanding of the features of the facility and surrounding area must be developed. This understanding can best be represented in digital files so that analyses may be conducted, data managed and calculations performed efficiently. These digital files make up the ***digital facility description***, that is, the description of natural and man-made features and information about the physical, geological, hydrological and chemical characteristics of the facility and the region. The digital facility description is the foundation upon which the site conceptual model is developed and all of the exposure analyses conducted.

The digital facility description consists of two major components; a spatial database and a tabular database. The spatial database contains the GIS shape files and coverages of geographic features related to regional information (e.g., surface water flow, geology) and facility information (e.g., locations of current and former process areas and storage tanks). It must include the coverages necessary to conduct the exposure pathway analyses and should include coverages that define the physical setting of the facility and provide contextual information (Romanek, 1999). The regional information can be obtained from various websites, including the United States Geological Survey (USGS) the Environmental Protection Agency (EPA), and includes topography, land use, and digital elevation data coverages. The facility information can be obtained from digital aerial orthophotographs or can be developed based on CAD files that have been converted to GIS files. The tabular database contains the facility feature descriptions and chemical and physical data (e.g., soil property measurements, chemical of concern concentration data) linked to each other through common fields, thus forming a relational database. The relational database provides an easily manageable format for storing, adding, and retrieving different types of information (Romanek, 1999).

The information included in the digital facility description for the case study is a comprehensive compilation of available information for the facility from environmental, geo-technical and other investigations or activities that have been conducted at the facility over the past ten to fifteen years and the available regional spatial data. The regional information for the case study digital facility description includes data from regional data sources, primarily state agency web sites and EPA web sites. The facility information for the case study digital facility description includes facility data sources, an environmental measurements database developed by the facility and the facility geographic coverages developed from a 1997 aerial survey. In addition, facility information has been collected from site investigation reports, historical maps, engineering drawings and older aerial photographs.

The case study facility is an approximately 300-acre crude oil refinery and petroleum products terminal that has operated since the early 1900's. For this case study site, an aerial survey was conducted and the facility coverages of operating units, tanks, waste areas, surface waters, etc., were digitized from the resulting orthophotographs. Figure 1 includes six facility scale coverages of the physical site features. When displayed together these depict a typical site plan. The types of analyses that are expected to be conducted will dictate the resolution needed for the mapping. As an example, for the case study site it was important to understand the potential for surface water runoff to the adjacent creeks and the river, so a digital terrain model (DTM) was developed from the orthophotographs to provide a detailed description of the land surface.

Developing the Spatial Site Conceptual Model

The ***site conceptual model*** is a critical component of the risk assessment process and the focal point of risk-based decision making. It provides the working hypothesis of all of the potential exposure pathways associated with the chemicals of concern identified at the many potential sources, their movement in the environment and their relationship to potential receptors. The site conceptual model is a synthesis of spatial and observational data. The challenge in assessing environmental risk at a large, complex facility lies in capturing the complexity of multiple sources and receptors. Typically, simplified site conceptual models are used to represent the relationships between the sources and receptors at facilities. However, using a simplified site conceptual model can potentially lead to an inaccurate understanding about the effects on receptors and expected results from implementing a remedial action alternative (Koerner *et al.,* 1998). This is particularly true when assessing the environmental risk to the many potential receptors existing on and off of a facility resulting from multiple sources and receptors, as is the case at most large industrial facilities. It is also often the case that the site conceptual model is developed at the beginning of the risk assessment project as a static display of the pathways thought to be of importance to the investigation at that time. The typical representation is a series of flow charts (ASTM, 1995, ASTM 1998, BP, 1997). For a facility with multiple sources of many chemicals of concern and various potential receptors, the presentation of flow charts is not very informative.

Figure 1. Compilation of Six Facility Coverages to Display a Site Plan

The public has received the process of risk assessment with skepticism, largely because it is not generally perceived to be an open and transparent calculation process. It has been viewed by many as a "black-box" approach. However, through the use of spatial and tabular databases, a spatial site conceptual model can be developed to describe the working hypothesis of the potential exposure pathways for a facility and provide a mechanism to communicate what is known and what is not known about the site and the exposure scenarios among the engineers, decision-makers and other stakeholders. The spatial site conceptual model can, therefore, be a means for different stakeholders to identify the exposure scenarios for which they have the greatest concerns and quickly identify the scenarios that have already been analyzed. To accomplish this the spatial site conceptual model is linked to the environmental modeling process and to the calculations of risk or corrective action goals. The Risk Assessment Data Model shown in Figure 2 depicts an exposure pathway evaluation that is tailored to the digital and spatial processes (Hay Wilson, 1998).

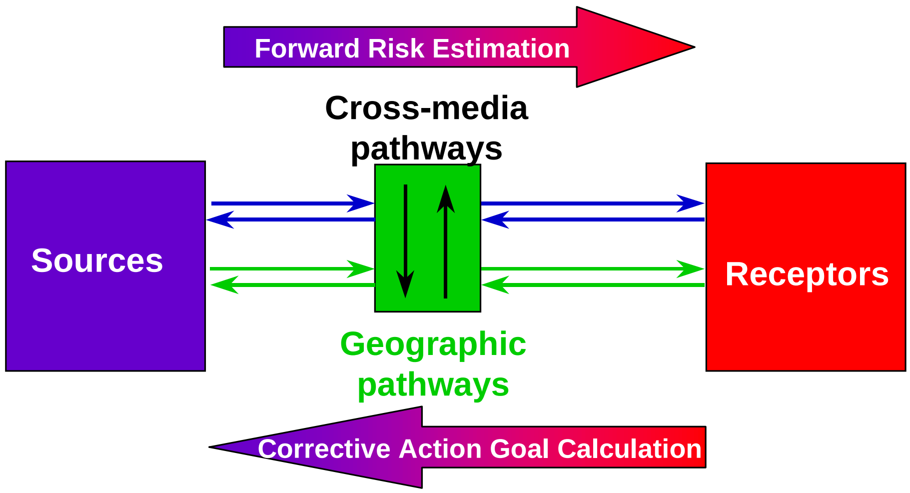


Figure 2. Risk Assessment Data Model

The spatial site conceptual model is developed by identifying each complete or potentially complete exposure pathway for a site. The exposure pathways are segmented into the functional elements of sources, cross-media transfer elements, geographic transport and receptors. Segmenting the pathways in this way allows for the connection of the elements to their computational counter parts. The descriptions of the sources are linked to the tabular database that describes the release history for each of the sources, or the sources can be linked to the spreadsheet that is used to calculate the representative concentration of the chemical of concern at a particular source area given the analytical results from environmental media sampling. Transfer and transport segments are linked to the spreadsheet calculations describing the environmental processes.

The spatial site conceptual model is an evergreen description of the current understanding of the site. Developing it using PC-based software makes the application accessible to all of the members of the project team (e.g., state regulators, responsible party project managers, consultants) and provides a mechanism to update the entire exposure pathway evaluation as new data is added. This saves time during modeling runs and reduces errors in data transfers. The fact that the exposure pathway evaluation results are stored in a tabular database means that the outcomes can be catalogued and tracked. This provides the stakeholders with the documentation they seek to identify the pathways analyzed and the outcomes. This process makes the risk assessment more reliable and repeatable.

Itemizing the exposure pathways in a database and segmenting the pathways facilitates the tracking of many multiple sources using conservative analytical models (Koerner, 1998). The sources that affect an individual receptor area can be identified and remedial action scenarios can be evaluated through an iterative calculation procedure. Cataloguing all of the potential impacts of the many multiple potential sources at a facility will provide the data and understanding of the site to support a site-wide analysis of exposures and ultimately of risks.

The spatial site conceptual model includes a spatial representation of the elements for multiple sources, pathways and receptors. The components of the spatial site conceptual model (e.g., sources and receptors) are represented as individual themes or data layers in vector representation. The user identifies the potential sources, transport mechanisms and potential receptors within the GIS application. Themes for potential sources are constructed based on the digital facility description data, the historical information and the environmental measurements. The sources are defined as point coverages. These coverages may include individual points for releases (e.g., the location of an emissions stack) or they may be the center points for source areas (e.g., the center of a defined area of soils containing chemicals of concern). Receptor locations are identified based on the spatial data and the current and potential future activities on the facility and surrounding properties. The receptor locations are defined by points, represented by points of exposure (e.g., the location of a drinking water well) or areas, represented by potentially affected areas (e.g., the residential neighborhood near a facility). The areas for which receptor identification is needed can be identified in the GIS application based on land use, census data or digital ecological habitat data. The transport mechanisms that link the sources to the receptors are grouped in themes based on environmental media. Lines define the lateral transport mechanisms such as groundwater flow. All of the elements are drawn in the GIS application. Scripts are used to assign distances, coordinates and areas to the spatial objects. Figure 3 includes one source area, six potentially affected areas, six points of exposure, and four geographic transport mechanisms.

The modeling of the environmental process is implemented using a tiered approach. For simple evaluations the single point to point relationships are analyzed using vector operations and one-dimensional algorithms. Simple fate and transport algorithms have been assembled in a group of spreadsheets. In higher level evaluations the modeling is accomplished as raster or grid functions.

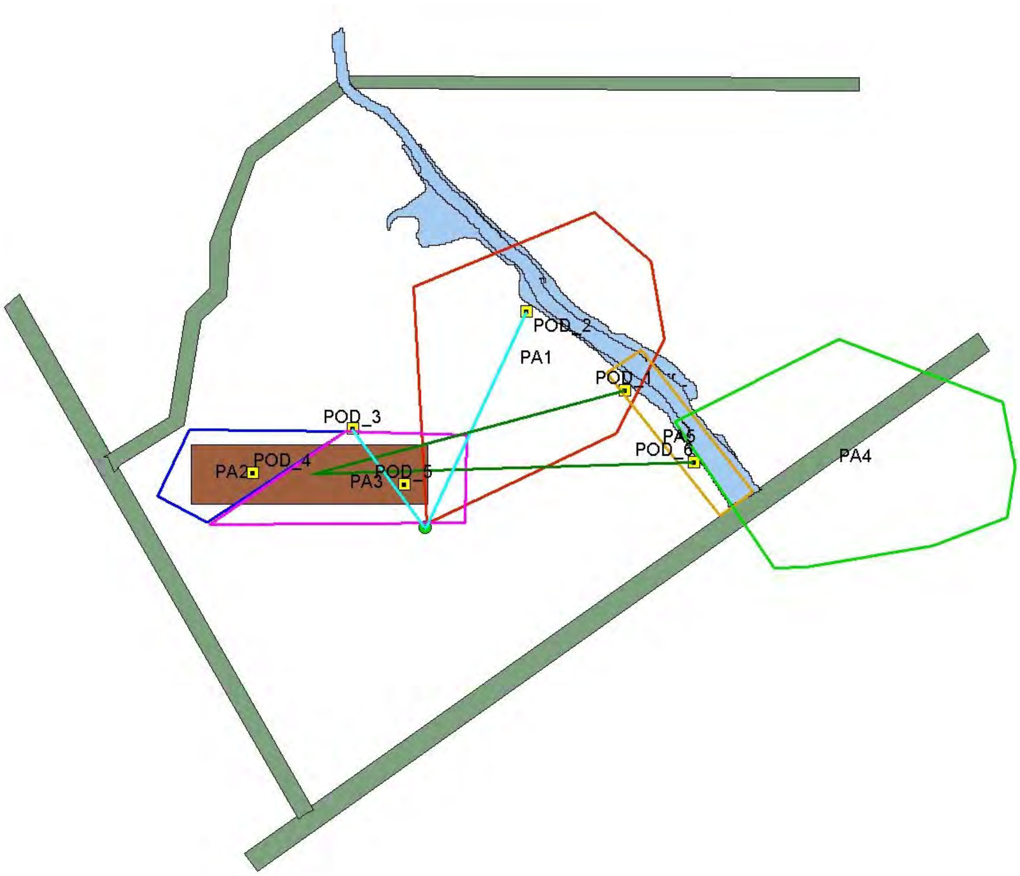


Figure 3. Example Spatial Representation of the Site Conceptual Model Components

Grid-based groundwater and surface water models have been developed using the methodologies presented by Maidment, 1996 and Romanek, 1999.

Conclusions

Risk-based decision making is an iterative process that provides a mechanism to determine the necessary and cost-efficient strategy for protection of human health and the environment. It requires the interrelationship and interoperation of risk assessment, risk management and risk communication. Risk Assessment is accomplished by collecting information to construct exposure hypotheses for a chemical of concern or group of chemicals of concern and evaluating those hypotheses to determine the current and future potential for adverse effects to human or ecological exposures from chemicals of concern in the environment. The incorporation of risk assessment as an integral component of an overall decision making process provides a mechanism to determine the necessary and cost-efficient strategy for protection of human health and the environment. In addition, improved methods for engineers and scientists conducting risk assessments and greater acceptance of risk assessments by the stakeholders will support the general environmental regulatory agency and community goals of environmental protection and sustainable economic development.

The contribution of this research is a new spatial risk assessment methodology using GIS to practically implement a holistic environmental risk assessment that accounts for the multiple exposure pathways at large, complex facilities and support risk-based environmental management decisions. This methodology provides an information processing system that more clearly ties the data and information for a study area to the risk-based decisions that are made for an environmental management project. In addition, the application of a spatial site conceptual model methodology helps automate the selection of exposure pathways to be considered, provides a mechanism to document the pathway completeness evaluation and provides connections to the transport calculation components and to the site conceptual model tabular database. In this manner, the process of evaluation and calculation can be more clearly understood and presented.

Acknowledgements

BP Amoco provided funding to support this research. Dr. Robert Gilbert, Dr. Randall Charbeneau and Ms. Susan Sharp provided insights and support.

References

American Society for Testing and Materials (ASTM). 1995. Standard Guide for Developing Conceptual Site Models for Contaminated Sites*.* E 1689-95.

American Society for Testing and Materials (ASTM). 1998. Standard Provisional Guide for Risk-Based Corrective Action. PS104-98.

BP Exploration & Oil Inc. (BP). 1997. "Risk-Based Decision Process Guidance Manual." BP Exploration & Oil Inc. Cleveland Ohio.

Environmental Systems Research Institute (ESRI). 1998. ArcView GIS, Version 3.1, Redlands, CA.

Hay Wilson, L. 1998. "A Spatial Risk Assessment Methodology for Environmental Risk-Based Decision-Making at Large, Complex Facilities," Dissertation Proposal, The University of Texas at Austin, Department of Civil Engineering, Austin, Texas.

Hay Wilson, L. L. N. Koerner, A. P. Romanek, J. R. Rocco, R. B. Gilbert. 1998. "Critical Success Factors for Implementing Risk-Based Decision Making at a Large Refinery Site," *Contaminated Sites Remediation Conference: Challenges Posed by Urban & Industrial Sites Proceedings*, March 23-25, 1999, Fremantle Western Australia.

Koerner, L. N., L. Hay Wilson, A. P. Romanek, J. R. Rocco, S. L. Sharp, R. B. Gilbert. 1998. "Maximizing the Value of Information in Risk-Based Decision-Making: Challenges and Solutions," American Nuclear Society Conference, April 3-5, 1998, Pasco WA.

Koerner, L. N. 1998. "Development of a Site Conceptual Model Using a Relational Database," Masters Thesis. University of Texas at Austin, Department of Civil Engineering, Austin, Texas.

Maidment, D. R. 1996. "Environmental Modeling within GIS," in *GIS and Environmental Modeling: Progress and Research Issues*. Goodchild, M. F. *et al., eds*., GIS World, Inc. Fort Collins, Colorado, p 315(9).

Microsoft Corporation 1997. Microsoft Office Professional Edition 97, Seattle, WA.

National Academy of Sciences (NAS). 1983. Risk Assessment in the Federal Government. Managing the Process. National Academy Press, Washington, D.C.

Rocco, J. R. and L. Hay Wilson. 1998. "The Evolution of Risk-Based Corrective Action," Published in proceedings: *Geo-Congress 98, Risk-Based Corrective Action and Brownfields Restoration*, ASCE, Boston, Massachusetts, p44(11).

Romanek, A. P. 1999. "Building the Foundation for Environmental Risk Assessment at the Marcus Hook Refinery using Geographic Information Systems," Masters Thesis. University of Texas at Austin, Department of Civil Engineering, Austin, Texas.

Washburn, S. T. and K.G. Edelmann. 1998. "Development of Risk-Based Remediation Strategies," Published in proceedings: *Geo-Congress 98, Risk-Based Corrective Action and Brownfields Restoration*, ASCE, Boston, Massachusetts, p30 (14).

Characterizing the Hydrogeology of Acid Mine Discharges from the Kempton Mine Complex, West Virginia and Maryland

Benjamin R. Hayes
Department of Geological and Environmental Sciences
Susquehanna University

Edgar W. Meiser, Meiser and Earl, Inc.

Constance Lyons, Maryland Department of the Environment, Bureau of Mines

INTRODUCTION

The Kempton Mine Complex consists of an area of interconnected, abandoned underground coal mine workings in the Upper Freeport Coal seam in southwestern Maryland and northeastern West Virginia (Figure 1). The mine complex was operated by the Davis Coal and Coke Company from 1885 to 1950. The workings encompass an area of 31.3 km^2. Groundwater discharges from the abandoned mine contribute on the order of six million gallons per day of acid mine drainage (AMD) to the North Branch of the Potomac River and the North Fork of the Blackwater River.

OBJECTIVES

The Maryland Bureau of Mines is undertaking a comprehensive investigation of the Kempton Mine Complex to develop remedial measures to reduce AMD by (1) decreasing the quantity of recharge to the deep mine and/or (2) improving the quality of the discharge from the mine pool. Before specific remedial measures are identified and evaluated, a GIS was developed to refine our understanding of the mine *structure* and *hydrology*. The visualization and computational capabilities of GIS enhance our ability to conceptualize the geometry and hydrology of the deep mine complex.

GIS DEVELOPMENT STRATEGY

Sources and Types of Data. Data incorporated into this study are diverse and of varying quality. The watershed characteristics and mine information included published geologic maps and boring data from the West Virginia Geologic Survey, Maryland Geologic Survey, U. S. Geological Survey, and U.S. Bureau of Mines, as well as records from the coal company including mine inspection reports, coal production reports, and water management reports. These data were synthesized, converted into digital form, and archived into Microsoft Access™

databases. Selected figures and important pages from old reports were scanned and archived as raster images in the database. Geologic maps, structural contours, hypsography, hydrology, and transportation features were digitized from the Blackwater, Davis, Lead Mine, and Mozark Mountain USGS 7.5 minute quandrangles. Detailed maps of underground mine workings at a scale of 1 inch equals 100 feet that were originally surveyed by mine engineers with the Davis Coal and Coke Company were digitized in AutoCAD and incorporated into ArcView GIS coverages of the mine region. Boring locations and coal depths were incorporated into the coverages. Details of each coverage were documented in metadata files in HTML format that can be read and written by any web browser.

Historical aerial photographs from the 1950s through 1980s were scanned and combined in the GIS with recent (March, 1999) low-altitude orthophotographs of the mine region. The recent orthophotos were georeferenced and rectified to control points surveyed on the ground using precision GPS survey techniques.

Hydrologic and water quality data collected at field monitoring stations were organized in Microsoft Excel™ spreadsheets (flat files). These spreadsheets were used to analyze time series of mine discharges, stream flow hydrographs, and water quality samples.

<u>Data Analysis</u>. The GIS proved valuable at integrating the historical boring data and mine maps with the ongoing mine discharge and water quality data. Three-dimensional stratigraphic cross-sections and isopach maps help identify hydrogeologic features such as stratigraphic pinch outs, changes in dip, and mine pool barriers. Large sets of spatial data can be integrated and analyzed to generate digital terrain models, correlate stratigraphic units, and compute the locations of outcrops of selected strata at the ground surface. Algorithms are currently being developed to generate new maps of flooded mine areas, mine pool shorelines, and directions of groundwater flow.

MINE STRUCTURE

In order to characterize the hydrogeology of the mine complex, one must first understand the details of the structure of the mine. The structural features include the location, orientations, and dimensions of the Upper Freeport coal and the underground workings within it.
Specific questions about the mine *structure* include:

1. regional geologic structure of the coal seam;
2. precise locations of underground mine workings;
3. degree of coal extraction;
4. locations of coal outcrops; and
5. proximity to overlying coal seams that have been mined.

Regional geologic structure of the coal seam. The Upper Freeport coal seam is found in a broad, north-east plunging syncline in Upper Potomac coal basin. The fold axis trends about 5 degrees and plunges 2 to 3 degrees to the northeast. The dip of the syncline limbs varies from 2 to over 16 degrees, with minor variations. Over most of the Kempton Mine, the Upper Freeport coal seam is 4 to 4.5 feet thick and pinches out on the eastern limb of the syncline (Figure 2). No major faults are mapped on the regional geologic maps, but numerous faults and offsets are described in the mine reports.

Figure 2. Kempton Mine and elevation contours of the base of the Upper Freeport coal seam. A detail structure contour map of the base of the Upper Freeport coal was prepared by the West Virginia Geological Survey from over 230 borings compiled from multiple sources. The boring logs provide top and bottom Upper Freeport coal elevations which were incorporated into the GIS database. We used the GIS to generate trend surface and three-dimensional boring diagrams from the boring database.

Precise locations of underground mine workings. The ability to generate maps showing the precise locations of underground mine workings in relation to present-day features is one of the most valuable contributions of the GIS in this study. The original (1885-1950) mine maps at a scale of 1 inch to 100 feet were carefully digitized by the current property owner, Western Pocahontas Properties in AutoCAD format. The coverages were georeferenced in West Virginia State Planer Coordinates and verified using a network of 12 monuments surveyed using precision GPS technology. The CAD coverages were imported into ArcView GIS and draped on top of digital orthophotographs and (DRG) maps of the region (Figure 3).

Degree of coal extraction. Headings and pillars are clearly visible on the old mine maps. The digital maps of the old mine workings were extensively cleaned and edited to generate closed polygon coverages of the subsurface mines (Figure 4).

Secondary mining, i.e. removal of coal pillars, occurred over most of the mine area. The remaining sections of the mine apparently remain intact as a labyrinth of tunnels and pillars. A noteworthy exception to this is the unmined coal barrier that trends east-west, located 3,000 feet north of Coketon (Figure 4). This barrier has a width of approximately 400 feet and begins near Thomas and extends 5700 feet to the west. The western tip of this barrier lies at elevation 2870 feet MSL according to the coal structure elevation contours.

<u>Locations of coal outcrops</u>. The outcrop of the Upper Freeport coal has been surface-mined along its entire length. Because of the surface disturbance of the coal at its outcrop, the actual location of the coal crop line is difficult to define and locate in the field. By projecting the structural contour surface of the coal seam onto the ground surface topography, the GIS enables us to generate maps showing actual position of coal crop lines. The crop lines, corresponding to the edge of the coal underclay, or pavement, are needed to calculate watershed areas contributing to the deep mine.

<u>Proximity to overlying coal seams that have been mined</u>. The Bakerstown coal, which lies approximately 180 to 200 feet above the Freeport coal, was extensively deep-mined and surface-mined. Because the Bakerstown coal is relatively close to the surface, and because it has been surface-mined in many areas above the Kempton Complex, it is very effective in capturing recharge from precipitation. Once the water has "sumped" into the Bakerstown coal seam, it inevitably drains into the underlying deep mines. We suspect however, that in some areas the opposite situation is true, where flow from the Kempton complex may leak upward into and discharge from the Bakerstown coal mines.

MINE HYDROLOGY

Once the deep mine maps were refined accurately using GIS, our ability to *visualize* the relationships between the Kempton Mine and overlying aquifers and areas of deep mining in the Bakerstown coal was improved greatly. The visualization capability of GIS enables us to develop the *fundamental concepts* of the deep mine's role in the regional groundwater hydrology.

Questions about the mine *hydrology* include:

1. the extent of mine flooding and elevations of mine pools;
2. interconnection between deep mine to surface water runoff;

3. leakage from streams and wetlands into the deep mines;
4. quantities of recharge and discharge; and
5. hydraulic gradients and groundwater flow directions in strata overlying the mine.

Extent of mine flooding and elevations of mine pools. Flooding in the abandoned mine workings is divided into two distinct and separate mine pools. The northern mine pool has a surveyed elevation of 2656 feet MSL and discharges from an 8-foot diameter air shaft and 12-inch borehole into Laurel Run about one mile north of Kempton, Maryland. These discharges are located approximately 4.5 miles upstream of the North Branch Potomac River. The Kempton mine pool elevation is constant and controlled by the elevation of the air shaft.
The southern mine pool has an approximated elevation of 2870 feet MSL and discharges from several deep mine entries immediately south of the village of Coketon, West Virginia into the North Fork of the Blackwater River. The elevation of this pool must be controlled by the intact coal barrier west of Thomas, West Virginia described earlier. This barrier is relatively impermeable, considering that in the vicinity of Rose Hill Cemetery northwest of Thomas, it narrows to 200 feet in width and impounds hydraulic heads exceeding 200 feet.

Interconnection between deep mine to surface water runoff. Mine discharge records show significant increases in response to large rainfall events, suggesting direct inflow of surface water into the mine. We used the GIS to delineate areas of shallow cover above the mines (Figure 5). These areas will be examined in future field efforts to identify any sources of direct inflow.

Leakage from streams and wetlands into the deep mines. Leakage from wetlands along the Potomac River has been observed in abandoned mine shafts at the town of Kempton, Maryland. Streamflow losses are predictable in streams draining the east slope of Backbone Mountain, as they cross the outcrop of Upper Freeport coal where the Kempton deep mines are closest to the surface. We used the GIS to overlay stream coverages with coal outcrop maps in order to delineate contributing watersheds and to identify locations of potential leakage. At these locations, predicted streamflows can be compared with measured streamflows to determine whether significant leakage is occurring.

Quantities of recharge and discharge. In October and November, 1998, eleven weirs were installed to gage discharges from the Kempton Mine Complex. Two weirs were installed on discharges from the Kempton pool to Laurel Run; one on the air shaft and the other on the borehole. Nine weirs were installed on discharges from the Coketon pool to North Fork Blackwater River.

We are using the GIS to analyze time series of discharge records from the weirs and precipitation records from nearby gaging stations. We are developing ArcView AVENUE™ scripts to plot hydrographs for each weir and to compute the total volume of discharge for any specified period of time. The total discharge volumes are then compared with the respective watershed areas contributing to the deep mine, delineated by the GIS, to compute the mine recharge rates (gpm/acre). The computed mine recharge rates can be compared to measured precipitation, as well to as average regional recharge, to determine where recharge to the Kempton Mine is anomalously high.

Based on preliminary weir measurements, recharge rates in the Coketon mine pool watershed appear to be greater than recharge to the Kempton mine pool watershed. The Kempton mine pool discharges up to 6 millions gallons per day while the Coketon discharges a maximum of approximately 4.4 million gallons per day – 75 percent of the northern mine pool. However, the Coketon pool has a recharge area of only 2,900 acres - 63 percent of the northern mine area of 4,600 acres.

<u>Hydraulic gradients and groundwater flow directions</u>. In late summer/early fall 1999, piezometers will be installed in aquifers overlying the Upper Freeport coal. Comparison of hydraulic heads in these aquifers to mine pool levels will dictate vertical directions of groundwater flow and hydraulic gradients. The head data will be incorporated into the GIS and gradients computed and cross-sections generated showing groundwater flow directions.

FUTURE WORK

As field investigations continue, the GIS is being used to locate new test borings and monitoring wells, delineate areas for subsidence monitoring, integrate historical and current mine discharge water quality data, and characterize groundwater recharge rates and flow directions. By analyzing the spatial variability of the hydrogeologic data, we plan to use the GIS to discover relationships among various parameters. Temporal variations in certain parameters can also be analyzed to discern patterns of change not apparent with static, two-dimensional maps. For example, areas of mine pool fluctuations can be combined with coverages of coal structure to delineate areas of AMD generation.

ACKNOWLEDGEMENTS

This project is being accomplished and funded through the cooperation and team effort of numerous state, federal, and private entities: Maryland Department of the Environment, Maryland Department of Natural Resources, University of Maryland, West Virginia Department of Environmental Protection, Susquehanna University, U.S. Environmental Protection Agency, Region 3, U.S. Department of Energy, U.S. Office of Surface Mining, Anker Coal Company, Buffalo Coal Company, Mettiki Coal Corporation, Western Pocahontas Properties, Meiser and Earl, Inc.

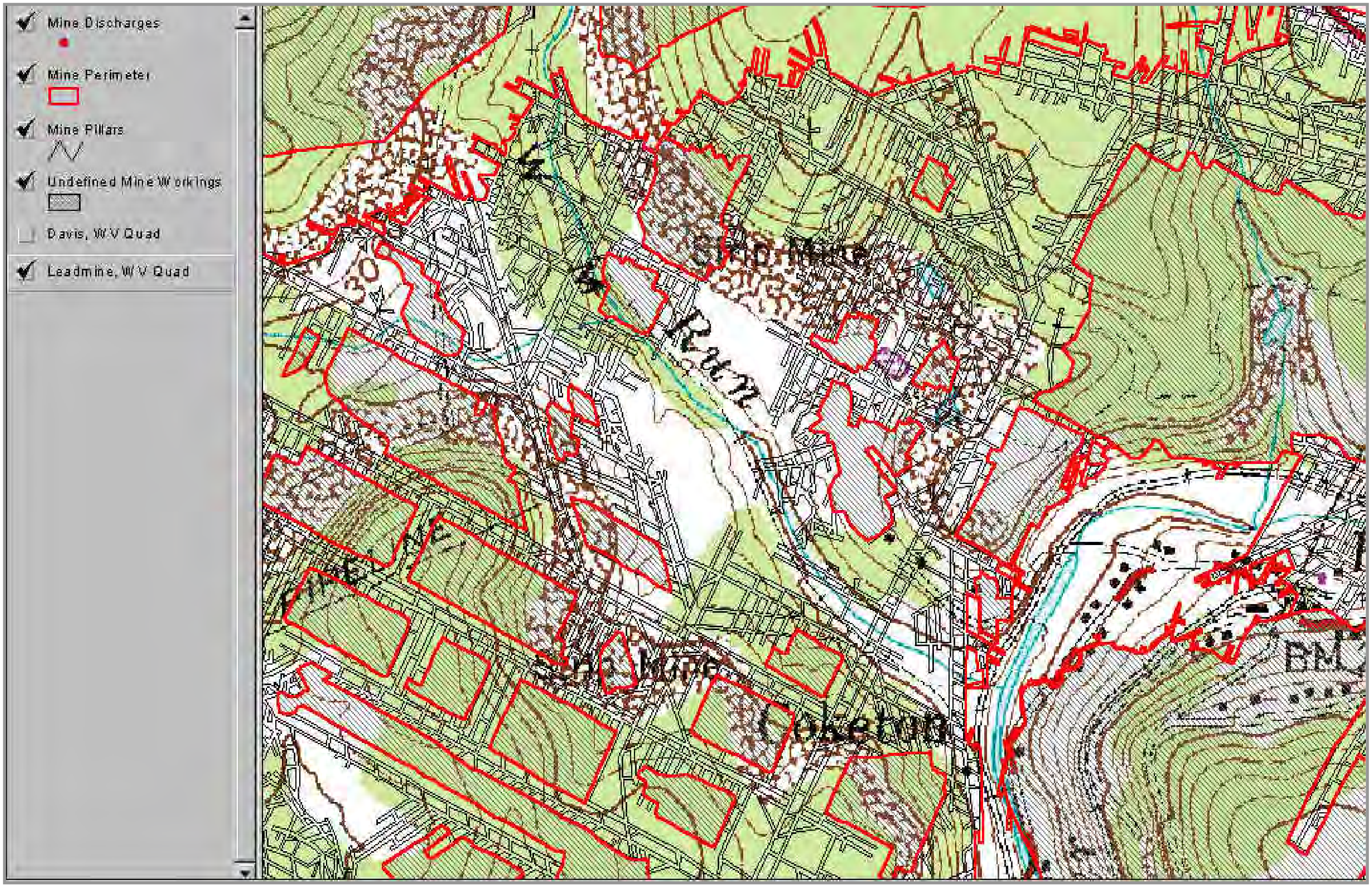


Figure 1. Study location map. Inset shows GIS coverage of the mine pillars and haulways draped over USGS quadrangle maps.

Figure 2. Kempton Mine and elevation contours of the base of the Upper Freeport coal seam.

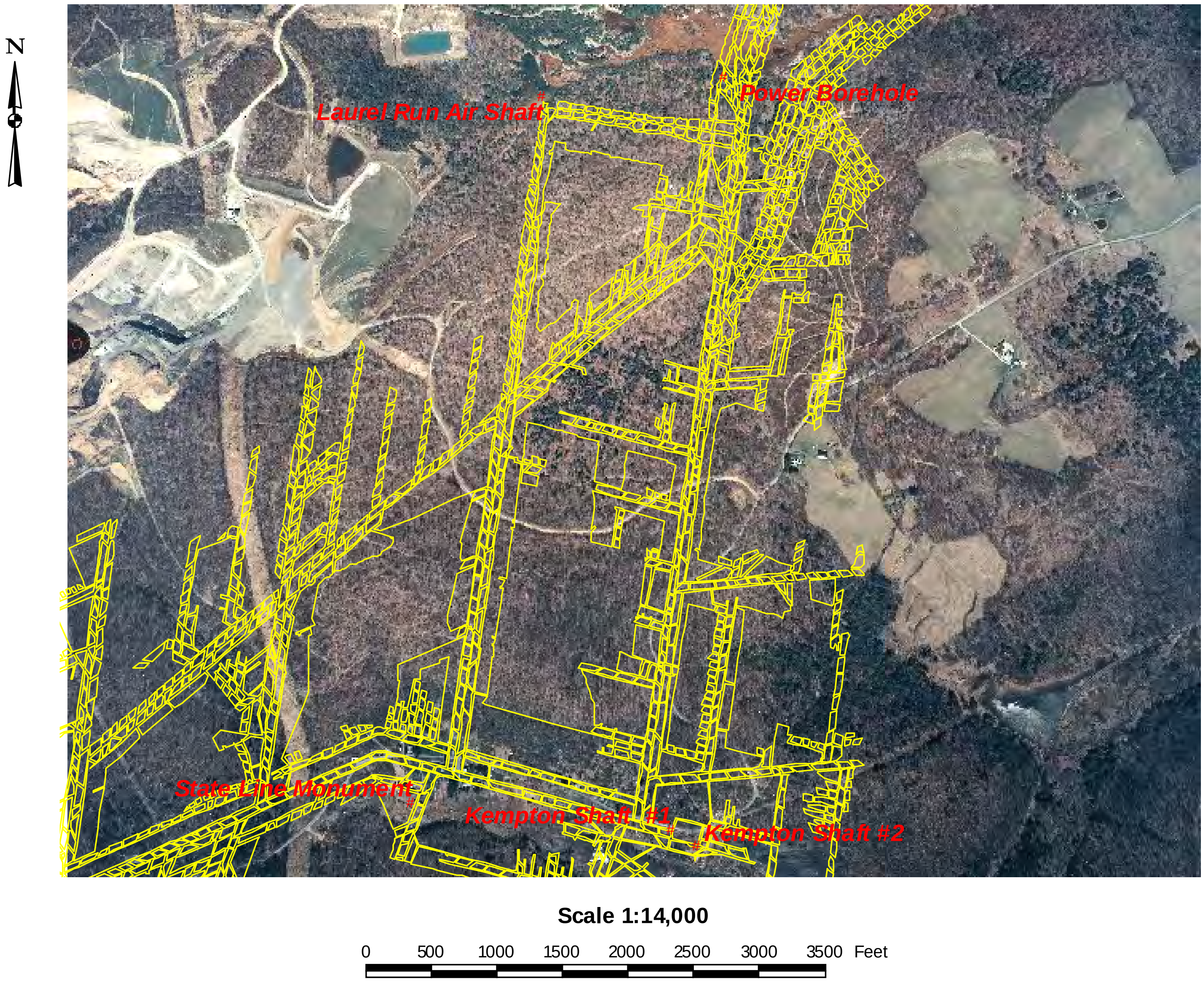


Figure 3. GIS coverages of northern section of Kempton Mine complex showing detailed mine working maps draped on top of color NAPP aerial photograph.

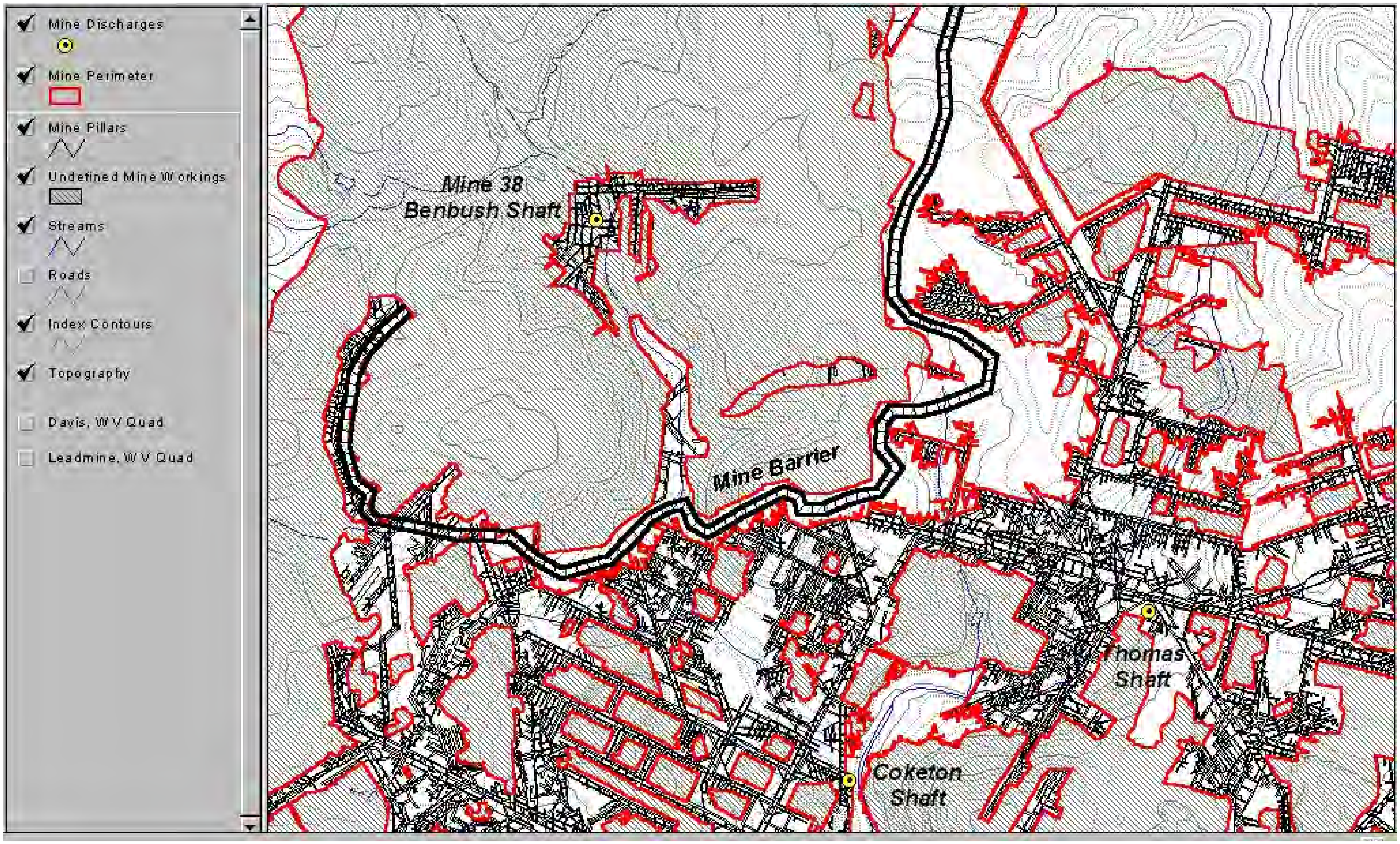


Figure 4. Southern portion of the Kempton Mine complex showing location of intact mine barrier.

Figure 5. Overburden thickness map of Kempton Mine Complex.

Use of GIS Tools for Conducting Community On-site Septic Management Planning

David Healy, Chief, GIS Services & Bruce F. Douglas, Senior Geoscientist

Stone Environmental, Inc.
58 East State Street
Montpelier, Vermont, USA
802 229-1879
dhealy@stone-env.com

ABSTRACT

This paper presents a summary of the authors' experiences in using GIS tools in support of Community Septic Management. This experience examines a number of Massachusetts and Vermont case studies in which the use of GIS was a primary tool in the planning and implementation of community septic management. These are the towns of Duxbury and Tisbury, Massachusetts, and the towns of Jericho, Vermont. Based on these case studies lessons can be drawn on how to further evolve the use of the technology in aiding community environmental management. This effort is based on the State of Massachusetts' nation-leading efforts in addressing the issue of management of on-site septic systems. The state's efforts are designed to address current failures, the protection of environmentally sensitive areas, and to assist communities develop tools and capabilities for insuring that problems are addressed and systems remain functional at all times. The Vermont community experiences are driven by a desire at the local level to avoid environmental pollution and costly centralized sewer systems.

The paper will describe all aspects of acquiring and using state and local GIS data to carry out the plans in four communities. Issues surrounding availability, quality, and data formats will be described. Each of the case study communities approached the problem slightly differently based on the available data. Each community used GIS to different levels for decision making and management. From each of these experiences, a summary of the successes, pitfalls and challenges will be described.

Overview

Decentralized wastewater management is a growing trend in the United States. Faced with soaring costs for large centralized treatment systems, communities are increasingly turning to the smarter management philosophy associated with the decentralized approach. Traditionally, one expensive solution has been available to communities that have outgrown or outlived their on-site septic systems — a sewer system and an expensive treatment facility. With federal funds becoming increasingly scarce, most small communities can not afford this type of conventional centralized approach.

By definition, decentralized wastewater management employs all available treatment and disposal technologies. The appropriate technologies, in a measure that meets current needs and takes into consideration future growth, are matched with the treatment and disposal requirements that have been identified. The end result is a unique municipal wastewater management solution that includes a program of preventive maintenance designed to identify weaknesses or potential failures before they become a problem.

This approach also simplifies future maintenance and planning. Community-wide decentralized wastewater management offers an opportunity to track the condition of individual systems, the relationship of those systems to other community infrastructure, like drinking water sources, and to the environment.

The Role of Geographic Information Systems

Geographic Information Systems (GIS) are a powerful tool for identifying and examining problem areas to display information for public understanding. GIS played a significant role in supporting a number of elements in each of the following case histories. The overlay capability of GIS was used to show environmentally sensitive areas, soil suitability for on-site systems, and areas of the community using subsurface disposal for prioritizing replacement systems. GIS was used for mapping soil types to determine suitability for siting systems, mapping community infrastructure and existing systems, tracking plans for future development, mapping environmentally sensitive areas, and tracking maintenance, repair and upgrade information for an entire community. Maps can be readily generated from the GIS to make decision making more timely and public education more effective. The maps effectively communicate complex information at-a-glance in a graphical format that is easily understood.

Database Applications

Databases are considered essential to enable the efficient management of on-site systems at the local government level. Databases can be used to track permits, inspections, pump-outs and failures of on-site systems, maintain on-site system records, and prepare mailings for necessary system maintenance. The database should be capable of storing, retrieving and reporting all data pertinent to all of the on-site systems. The foundation of the database should consist of the following key data areas:

1. System related information including components, septic tank size, application and permit numbers, maintenance (septic tank pump-out dates, inspection dates and inspection results), plans, and images.
2. Parcel map and lot designation, related information including soil test data and structures, number of bedrooms and design flow.
3. People related information. Names and addresses of people used for lookup in various sections of the database.
4. Easy linkage to GIS. Any database characteristic can be displayed graphically through GIS.

Jericho, Vermont

The Town of Jericho is a suburban and rural community located in the Lake Champlain Valley area of Northwestern Vermont. The town includes 92 square kilometers (35 square miles) of land. The 5,000 residents rely on individual and cluster type on-site wastewater treatment and disposal systems (septic systems).

Town officials recognized the potential for treatment and disposal problems to develop as the town grows and existing septic systems age. The Jericho Board of Selectmen hired consultants to determine the most appropriate means of addressing short-term and long-term wastewater disposal needs for the town. The process involves assessing existing systems, characterizing environmental conditions, and analyzing the community's future options.

The study focused on 3 areas:

Jericho Center — a traditional 19th century New England village.
Jericho Village — a traditional village with some commercial properties.
Route 15 district — a recently developed commercial zone.

Data Availability: Data for Jericho came from two primary sources the town planning office and the Vermont Center for Geographic Information. The primary databases used for the analysis include:

SSURGO Digital Soils
Parcels
Roads
Surface Waters
Well Locations
Source Protection Areas
Swimming Areas
Wetlands
Floodways

Quality: This data was primarily digitized from 1:5000 Orthophotos and 1:24000 USGS Quad Sheets. FGDC type metadata did not exist for town developed data. The later issue poses long term problems for understanding the source parameters used. For this community mixing the two scales of data did not pose a major problem. For certain the FEMA flood data is highly inaccurate, but was not essential for this study.

Data Formats: Local data was in MapInfo format in Vermont State Plane Feet. State provided data was provided in ArcInfo format. Local data was converted to ArcInfo format and put into Vermont State Plane Meters, NAD 83.

The study team conducted a thorough assessment of existing systems and conditions. They identified environmentally sensitive areas and parcels with special conditions or limitations and conducted analyses and developed maps.

The GIS analysis revealed that soils in the study areas are generally suitable for on-site systems on terraces in the major stream valleys where the densest residential and commercial development is typically located. In individual neighborhoods, a 2 to 30 percent rate of on-site system failure in a 10-year period was evaluated. The failures were predominantly due to age, design and construction of systems. Most of these failures have been effectively resolved by on-site replacement of the septic tank and/or soil absorption system with conventional technologies. Due to the low density of development, centralized collection systems do not appear to be cost-effective. Due to existing use of the river nearest the most densely developed areas in town for swimming, direct discharge of treated wastewater was not an option either.

GIS Integration: The town has received the consolidated set of reprojected data and various ArcView projects developed for this project on CD-ROM. It has acquired ArcView and is using is for various planning and zoning/regulatory functions.

The consultants identified specific options for community wastewater management based on the town's current needs and plans for future growth. They concluded that existing septic systems and a locally developed septic system management program would be the most cost-effective option. To implement this as a sustainable option, Jericho has established a local Wastewater Planning Committee with three objectives:

1. Develop a decentralized wastewater management plan based on a town wide assessment of need.
2. Develop a community homeowner and student education program to increase the local understanding of on-site systems.
3. Demonstrate the effectiveness of septic tank risers and septic tank effluent filters in facilitating the inspection and maintenance of on-site systems.

Jericho's forward-thinking approach to decentralized wastewater management is saving the town money by using existing wastewater infrastructure, protecting homeowners' investments in their current systems, and avoiding the high cost of developing a centralized sewer and wastewater treatment facility.

Duxbury, Massachusetts

The coastal community of Duxbury, is located 35 miles south of Boston. Duxbury has a population of approximately 14,000 and a land area of approximately 61 square kilometers (24 square miles). Greater than 95 percent of the residents rely on individual on-site sewage systems and the community is committed to using on-site systems as a long-term solution to wastewater management. The town relies on an extensive sand and gravel aquifer for community drinking water supplies. In addition to aesthetic and recreational value of the freshwater and saltwater resources of the community, there are numerous cranberry bogs in commercial cultivation and a significant potential for shellfish harvesting in the coastal embayments.

Over the past three decades, Duxbury has been making substantial efforts to protect their groundwater and surface water quality with a permitting program for on-site wastewater disposal systems. In 1996, the town dedicated two shared soil absorption systems designed to address severe problem areas along Duxbury Bay. These shared systems (each with less than 37,800 litres per day design flows) handled wastewater from three parcels along the Bluefish River and 18 parcels in the Snug Harbor area by conveying the wastewater to sites located inland that are suitable for soil absorption systems. The establishment of these systems has enabled the opening of shellfish harvesting areas due to a decrease in bacteria in the coastal embayments. The town has recently voted to design and build another 37,800 litres per day cluster on-site system to serve approximately 30 households in a residential area along Kingston Bay to improve water quality in a historic shellfish harvesting area. To continue their on-going efforts in this area, the town has recently completed a Community Septic Management Plan (CSMP) to provide a clear process for decentralized wastewater management.

Duxbury's CSMP consists of the following components:

1. Comprehensive inventory of existing systems.
2. Parcel information permit records and septic tank pump-out records stored in a database to track permitting and maintenance.
3. Development of a local GIS system and training of Town personnel to identify:
    a. Drinking water source protection areas
    b. Freshwater wetlands, ponds, vernal pools, rivers and streams
    c. Saltwater wetlands, coastal resource zones
    d. Buffer zones around these areas
    e. Parcel maps to determine environmental sensitivity of particular parcels.
4. Public Education and Information development including a brochure explaining the existing situation and the community septic management plan.
5. Betterment Loan/Upgrade Program.
6. Ranking environmental sensitivity by category to determine priority of parcels for loan program to assist homeowners in upgrading failed systems.
7. Voluntary Maintenance program that recommends setting routine system maintenance in conjunction with the following requirements in local ordinance: detailed reporting of septic tank condition, liquid and solid levels at septic tank pumping.

Data Availability: Data for Duxbury came from two primary sources the University of Massachusetts, Boston which had completed an Open Space Plan for the town and from MassGIS office. The primary databases used for the analysis include:

Digital Soils
Parcels
Land Cover
Roads
Surface Waters
Well Locations
Source Protection Areas
Swimming Areas
Wetlands
Floodways
Title 5 Setbacks

Quality: This data came from highly mixed sources. MassGIS is primarily 1:24000 scale data digitized from USGS Quad Sheets and 1:20000. FGDC type metadata did not exist for UMass data.

Data Formats: MassGIS data was in Massachusetts State Plane Meters Feet North American Datum (NAD) 1983. On the other hand, Umass' GIS data was in an unknown and undocumented format. It turned out after much trial and error that it was in Massachusetts State Plane Feet, NAD 1927.

Specific GIS Processing: The ranking system devise for the betterment loan program was based on a assignment of weighted values to six factors. In this community the intersecting of coverages with the weighted values was competed to derive a composite weighting. A "grid-based" math algebra operation could have just as easily been used.

GIS Integration: Duxbury has received a complete copy of all the data and ArcView projects developed for the project. Two days of training were provided to staff members in the Planning and Health Departments. They are currently using GIS in many diverse planning and management functions.

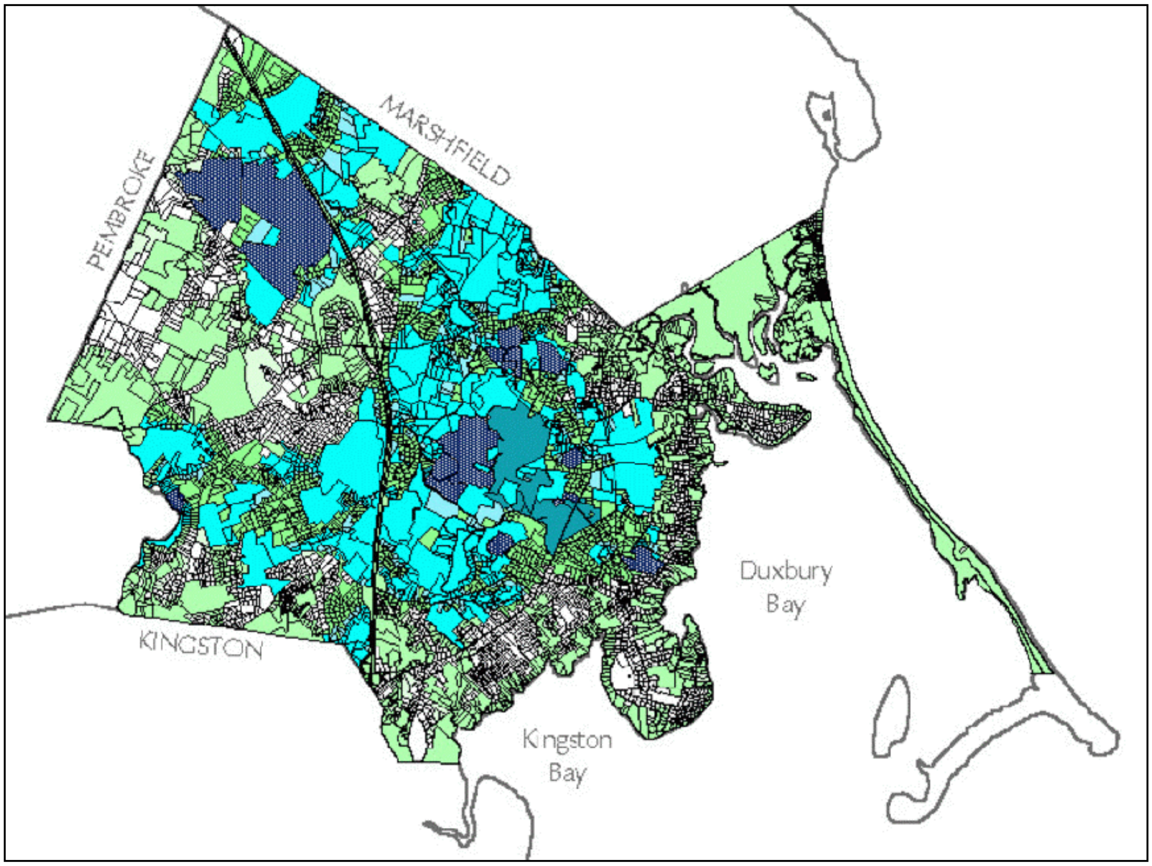


Figure 1: Priority Ranking Results

Tisbury, Massachusetts

The Town of Tisbury is a coastal resort community located on the island of Martha's Vineyard located 8 km (5 miles) off the southwestern coast of Massachusetts, USA. The Town covers an area of 54 square kilometers (21 square miles). The population of Tisbury is approximately 3,000 year round residents and 10,000 seasonal residents. Approximately three-quarters of Tisbury's residents rely on groundwater from municipal wells that tap a glacial sand and gravel deposit underlying the western half of the island. The remaining residents utilize individual wells, generally tapping the same aquifer. Currently all properties in town are served by on-site sewage disposal systems. The Tisbury Board of Health estimates that there are currently approximately 2,500 individual on-site systems.

The intent of Tisbury's wastewater management program is to provide an institutional and regulatory framework enabling the long-term viability of the on-site wastewater treatment and disposal facilities in the Town. The program includes a Community Wastewater Management Plan (CWMP), Watershed Management Strategy, Public Outreach and Education, Institutional and Regulatory Requirements, and a Program Implementation Strategy. The CWMP is in the final planning stage and has not been adopted by the community.

The town is taking a very pro-active approach to on-site system management. The local Board of Health and Wastewater Planning Committee is developing a management program with required inspections for every system and requiring pump-outs of septic tanks based on expected rates of solids accumulation for the specific tank size and system usage. GIS has been used to delineate environmentally sensitive areas and parcels; a ground water flow model has been used to determine water table contribution to surface waters; and a database is being developed to track the management program and notify residents of compliance requirements. This approach enables the residents to easily see the relationship between the areas that utilized on-site systems and environmentally sensitive areas such as aquifers, streams, coastal ponds and wetlands using a risk-based management strategy.

Environmentally sensitive areas in Tisbury will be used to identify high priority areas are for the management of on-site systems. Initial wastewater management districts have been defined to address the downtown Vineyard Haven area, and the low elevation (less than 3 meters (10 feet) above mean sea level) areas with the potential for systems to have the least vertical separation to groundwater.

Periodic maintenance is critical to ensuring the long-term success of even the most basic septic system. An inspection and maintenance program has been designed to assess current infrastructure conditions, ensure proper use and maintenance of on-site systems, and reduce future failures. A relational computer database, the Septic Information Management System (SIMS), will be established to maintain an up-to-date inventory of all onsite systems in town and to track the permitting of new systems, upgrades of existing systems, and inspection and maintenance program. The database will also be useful to ensure that all owners are adequately addressing the unique needs of their system. One of the primary goals of the management plan is to provide for better septage (solids and liquid pumped out of septic tanks) management by creating a predictable and manageable production of septage. This plan pertains to those parcels outside of the area serviced by the centralized wastewater collection and treatment system. While the central service area will have management requirements, they will likely differ from the rest of town and will be addressed in a separate management program.

On-site systems are designed to treat domestic wastewater before reaching the groundwater and down-gradient surface waters. The residual components of this treatment process, such as nitrates, have an impact on the environment. However the degree of impact is relative to the sensitivity of the groundwater beneath the on-site systems and the sensitivity of the surface waters where the groundwater discharges. The watershed management strategy is designed to protect the environmental resources of Tisbury at an appropriate level to the sensitivity of the different areas in town.

A key element of the watershed management strategy is to use a risk assessment/risk management approach. During the risk assessment process, the town will develop rankings regarding the value and vulnerability of local receiving environments to impacts from on-site systems, and to define the areas in town which contribute flow to the receiving environments. A steering committee of stakeholders will be established to address the needs of the community in this process. The rankings will be used to determine appropriate levels of treatment and develop a risk management program, in order to protect public health and the environment. For example, specific areas may be delineated where nitrogen removal is required for upgrades and new on-site systems to reduce nitrate loading to a particularly sensitive and valuable receiving environment.

The second program in the watershed management strategy is the development of a long-term groundwater monitoring program. A network of surface and groundwater sampling stations will be established up to monitor trends in water quality.

Environmental professionals, municipal departments, and community environmental groups will conduct the risk assessment/risk management process. The water quality monitoring program will be run in conjunction with the Martha’s Vineyard Commission, the University of Massachusetts Extension program, the Town of Tisbury, and local professionals. Funding for the watershed management program is will be provided through a combination of local and federal sources.

Data Availability: Data for Tisbury came from two primary sources the MassGIS office and a previous consultant who had worked for the town. The primary databases used for the analysis include:

Parcels
Building Footprints
Title 5 Buffered Footprints
FEMA Flood Data
Land Cover
Surface Waters
Roads
Well Locations
Source Protection Areas
Wetlands
Possible Discharge Areas
Color Infrared Orthophoto

Quality: This data came from two sources. MassGIS is primarily 1:24000 scale data digitized from USGS Quad Sheets and 1:20000. The consultant's data came primarily from the town's CAD-based tax maps for which were assembled and rubber sheeted to fit with the MassGIS data. FGDC type metadata did not exist for the consultant data. The FEMA data was the least accurate of the data used. The focus of the study was for the village of Vineyard Haven. Given the coastal nature of the data, the town boundaries varied widely depending on the original source.

Data Formats: Most Data was acquired in State Plane Meters, Mainland Zone, NAD 1983.The one exception was the digital CIR Orthophoto which came from the Massachusetts Coastal Zone Office in State Plane Meters, Island Zone, NAD 1983. Since we derived contours for the study area from this data, it had to be reprojected to State Plane Meters, Mainland Zone. These kinds of issues working with available data always presents problems in all studies to date.

GIS Integration: The town has recently embarked on an effort to develop a town-wide GIS program. This has begun with an evaluation of existing data and assessment of the many department needs.

Geographic Information and Tools for Informed Decisions: The Lake Superior Decision Support Project

George E. Host, Lucinda B. Johnson, Carl Richards
Natural Resources Research Institute

Pat Collins
MN Department of Natural Resources

Introduction

Recent trends in land use, such as increased population movement from urban to rural areas, conversion of forest land, and increased development have emerged as key issues affecting natural resource management in the Lake States. As units of government ranging from local townships to the federal governments of the US and Canada plan for the future, the need for data and tools for sound decision-making has become critical. Nonetheless, at the scale of the Lake Superior Basin, we lack synoptic information on even the most fundamental data layers required for sound planning. Among these are comprehensive coverages of land use/land cover, transportation infrastructure, hydrography, demography, and even the bathymetry and shorelines of Lake Superior itself.

In addition to the lack of spatial data, smaller units of government often embark on land use planning exercises with few tools at their disposal. While computer simulation models, draft ordinances, and decision support tools are receiving wider use in planning, these tools are often out of reach of local governments who lack equipment and expertise required for their use.

Finally, in natural resource management, the general public is often faced with information that has been slanted in favor of the perspectives of industrial or environmental advocacy groups. There is a critical need for a source of neutral data to allow the public to develop informed opinions on current issues.

To this end, we have begun a project to help resolve some the issues of data accessibility and interpretation with respect to land use planning in the Lake Superior Basin. We have three key objectives:

1) To develop synoptic databases across the Lake Superior Basin, and to make these data available through the Internet and other data distribution formats.

2) To concurrently develop decision support tools to assist local units of government in land use planning activities. These tools will consist of CD-based resources, including spatial data, prototype planning documents, and flowcharts to guide users through the planning process.

3) To develop and deploy information kiosks in Visitor's Centers and other publicly accessible locations around the Lake Superior Basin.

The following paper charts our experiences and progress to date, and provides some insights into issues related to integrating information and decision support tools into the planning process.

Motivation

This project was motivated by several key factors:

1. Forestry and forest products industries dominate the basin's economy.

2. Demand for tourism and leisure activities is replacing mining and other previous contributors to the region's economy, with an associated increase in development to leverage this demand. The Lake Superior Basin's wealth of natural features is a major factor driving these changes.

3. The Basin is a unique harbor of biodiversity, on a world-wide basis (TNC). A considerable portion of the Basin's biodiversity resources are located along coastal shores and wetlands.

4. The Basin's natural and biodiversity resources are increasingly in conflict with both forestry and development. And increasingly, the three are in conflict at a given place simultaneously. People want to visit and live, by and large, where natural resources are rich and biodiversity is highest.

It stands to reason that the Basin's future is directly dependent on the resolution of these land and resource use conflicts. But the resolution of these conflicts will not occur as a result of a small number of large government agencies or private land holders developing and implementing policies that address land use issues. The real impacts of development result from the cumulative effects of a large number of small land use decisions, spread across time and space. When integrated across time and space, these land use decisions have major impacts on the Basin's resources.

The unique role of the Lake Superior Decision Support project is to develop and foster the use of an infrastructure, based on GIS and computer models, upon which the Basin's incremental land use decisions can be made more effectively. Our target users are local governments, resource management agencies, commercial interests, citizen volunteers, advocacy groups, aboriginal groups, and educational organizations. There are two fundamental missions of the project, both intended to sustain the Basin's resource in a truly ecosystem-based approach (i.e., sustaining humans as a part of the natural world). The first mission is to provide practical, useable tools that can be used by people involved in day-to-day land use decisions. The second mission is to provide a context and a demand for these tools by providing educational and interpretive information for their use.

Beneficiaries

The direct beneficiaries of this project are those organizations and individuals that will directly use the GIS applications and databases developed by this project. But there will also be many indirect benefits that are difficult, if not impossible, to quantify. For example:

1. Local governments will gain tools to more effectively create and implement zoning ordinances that mold day-to-day land-use decisions.

2. Educational and interpretive organizations will gain resources that will help them inform the public about these resources and energize their target audiences to encourage their use in the planning process.

3. Regional agencies (including state/provincial and federal agencies) will, over time, gain access to a consistent GIS database for assessing regional and basin scale issues. In this process, data gaps are identified.

4. All parties to local and regional land use decisions and policies will gain from a collective knowledge base and shared and unbiased information base.

5. Ultimately the Basin and its people benefit from more effective land use decisions that help to sustain and steward the Basin's resources.

Data development

Data development began by prioritizing a list of key data layers critical for land use planning and capable of being mapped synoptically across the Lake Superior Basin. The geographic scope of the data was the Lake Superior watershed boundary plus a 50 km buffer (Figure 1). The buffer allows us to access impacts to the basin that may be due to activities outside the watershed boundary. The watershed boundary was constructed by merging a number of independently-derived boundaries developed at state or provincial scales – it was informative that neither the watershed boundary nor the shoreline of Lake Superior previously existed in a continuous, fine-resolution GIS coverage.

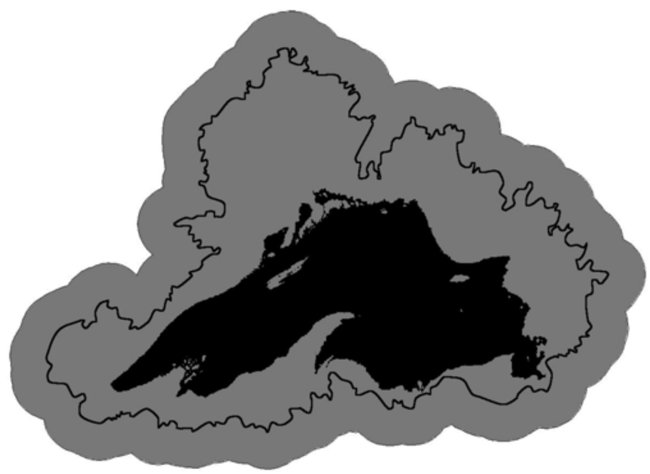


Figure 1. Lake Superior basin boundary with 50 km buffer.

The base scale of the data aggregated across the whole basin was set at 1:250,000; a compromise between the desired accuracy of the data and the ability to easily store, manipulate and visualize the data. We then conducted a broad survey of agencies involved in development of spatial data for natural resource management to identify the source, scope, and resolution of key data layers.

A critical step in the identification and acquisition of data was the development of Memoranda of Understanding and Cooperative Agreements with collaborative groups, such as management agencies and industry. A key issue is that many agencies operate on a cost-recovery basis, in which GIS products are sold to recover costs of development. In many cases, licensing agreements require that data be made available only in image format (e.g. GIF or TIF), rather than as spatially-referenced GIS databases. In addition, providing appropriate acknowledgments, respecting publication rights, and developing well-defined dissemination criteria was essential for this effort. To date, we have assembled approximately 30 databases across all or part of the Basin, as shown in Table 1.

Table 1.
Major spatial data themes, geographic extent and key attributes for data compiled for this project.

Data theme	Geographic extent	Attributes	Scale	Resolution
Bathymetric Model	Lake Superior	Interpolated depth values	1:547,000	1 km
Census data	MN, WI, MI, Ontario	Population, demography, housing		
Civil divisions, minor	MN, WI, MI	Townships	1:100,000	
Climate	MN, WI, MI, ONT	Interpolated data, monthly temp, rainfall		5 km
DEM	Lake Superior Basin	Digital Elevation Model		1 km
Ecological classification system	MN, WI, MI	ECS, Province to Subsection		
Ecological classification system	N MN, N WI	Land Type Associations		
Ecological classification system	ONT	Hierarchy: ecozone, region and district		
Forest inventory (FIA)	MN, WI, MI	Forest inventory data by sample point, 1990s		
Forest inventory (presettlement)	MN	GLO witness trees, section corner description	1:100,000	
Geology	MI, MN, WI, ONT	Surficial geology, landforms	1:1,000,000	
Habitat megasites	NE MN	Wilderness areas, ecoregions, parks etc.		
Habitat projects	Lake Superior basin	Habitat projects, location, description		
Habitat sites	Lake Superior basin	Habitat sites, location, description		
Hydrography	MN, WI, MI, ONT	Lakes, streams, rivers	1:100,000	
Land Cover: Forest vegetation	MN, WI, MI, S. CAN	24 classes of forest vegetation		1 km
Land Cover: General	Great Lakes Basin	Anderson II, Land cover/use		200 m
Land Cover: Original vegetation	MI,MN,WI	40 classes derived from GLO survey notes		
Land ownership	MI,MN,WI	Public/private ownership		40 acre
Pollutants-point source	Lake Superior basin	Industrial point source locations		
Pollutants-point source	Lake Superior basin	Municipal point source locations		
Public Land Survey	MN, WI, MI, ONT	Township, range and section boundaries		
Satellite imagery	Lake Superior	Mosaic of Landsat MSS, mid		200 m

Table 1.
Major spatial data themes, geographic extent and key attributes for data compiled for this project.

	Basin	1980's		
Satellite imagery	Lake Superior Basin	AVHRR NDVI Composite		1 km
Satellite imagery	Lake Superior Basin	Night image of city lights		1 km
Soils	ONT	Drainage, surface material, deposition, etc		
Soils STATSGO	MN, WI	Many attributes: texture, drainage, depth		
Transportation	MN, WI, MI, ONT	Classification of roads, railways, others	1:100,000	
Watersheds, Major	Lake Superior basin	Major Watershed boundaries		
Watersheds, Major	MN, WI	Major Watershed boundaries		
Watersheds, Minor	MN	Minor watershed boundaries		

Metadata

Documenting the source, accuracy, resolution, and other key attributes of spatial data is critical for its effective use. We are using the Minnesota Geographic Metadata Guidelines, developed by the Minnesota Land Management Information System, a state agency that serves as a clearinghouse for spatial information. The Geographic Metadata Guidelines are a simplified set of the Federal Geographic Data Committee’s Content Standards for Digital Spatial Metadata. They consist of 99 elements divided into seven categories describing the following data characteristics:

1. What kind of a database is this?
2. What is its quality?
3. How is it organized?
4. Where is the database located?
5. What features does this database describe, and how are they characterized?
6. Is this database distributed? If so, how?
7. Who documented this data?

We are implementing the metadata with a software tool called DataLogr, developed by the State of Michigan. The metadata for each data layer will be available on-line on the LSGIS web site.

Delivering Information: Web sites, Kiosks, and Exhibits

The LSGIS web site can be found at www.nrri.umn.edu/lsgis. It was developed and maintained with Microsoft Frontpage@ and allows for the interactive querying of maps using ESRI's Internet Map Server@ (IMS). IMS provides a number of tools for retrieving information from maps, and includes the ability to provide progressively more information as a user zooms in to finer scales of resolution.

In parallel with the Internet-based information, we are developing a touchscreen computer display that provides a similar suite of information. These touchscreen computers will be deployed in kiosks at several visitors' centers around the Lake States. We are also currently working with the Great Lakes Aquarium at the Lake Superior Center in Duluth, MN to integrate information collected in this project into a public "town meeting" forum that allows participants to assess data and make decisions on a series of problems facing the Lake Superior watershed. The touch screens use Site Explorer@, a map-oriented information utility developed by cooperator Mike Koutnik.

Decision Support Projects

A key part of this project is the implementation of the data and tools to address land use planning issues. We have two ongoing projects in this area, as described below:

Hydrologic Modeling of the Miller Creek Watershed

Miller Creek is a relatively high quality trout stream that runs through the cities of Duluth and Hermantown, MN. The stream originates in wetlands and undeveloped shrublands and woodlands northwest of Duluth, flows through a heavily commercialized area around Miller Hill Mall, through residential areas as it drops rapidly towards Lake Superior, and ends in an industrial area near its mouth. The upper part of the watershed is fairly flat, and the stream has a moderate gradient. Near the mall it is channelized for several hundred meters and starts to receive stormwater discharge from storm drains. Further downstream, in the more heavily residential areas, multiple stormwater channels enter Miller Creek. The stream develops a high gradient near its mouth. The stream is routed underneath the city in stormwater conduits and ditches until it enters the St. Louis River estuary of Lake Superior

The Miller Creek watershed contains several new strip malls, and further commercial development is under construction. This has generated a wide public concern over issues over environmental quality, particularly with respect to degradation of trout habitat resulting from increased stream temperatures, sediment inputs, and degraded water quality. To inform the decision-making process, we have assembled detailed land use maps of the Miller Creek watershed, and employed the EPA's PC-SWMM, a well-documented hydrologic model, to predict flow and pollution loads to the creek.

Our first goal was to model, as accurately as possible, the volume and temporal distribution of flow in Miller Creek, based on overland, ground water, storm sewer conduit, and natural channel flow using PC-SWMM. The second objective was to estimate pollution loads and concentrations in runoff based on land use, road densities, local street cleaning programs, and locally calculated and published land use-specific pollution loading rates. The modeling of these parameters will allow planners and others to determine the likely effects of proposed changes in the basin and evaluate options for reducing pollution loads to Miller Creek.

The model predicts the hydrograph associated with rain events fairly well (Figure 2). Since a large part of the pollution concentration model is based on flow rates, this accurate flow model was needed before proceeding to the pollutant modeling stage. Hydrologic modeling was complicated by the large amount of wetlands in the headwaters of the Miller Creek basin, but results to date show significant differences in water yield related to land use. For example, the effects of the Miller Hill Mall commercialized area on flow in Miller Creek is shown in Figure 3: the impervious surfaces around the mall result in a large increase in flow volume from a relatively small area, in spite of flow mitigation from detention ponds. If the flow volume needed to be reduced, the model could be used to determine the size of pond required to reduce the volume to the desired level. The pollution model is nearly complete as well, and should provide similar comparative data and scenario analysis capabilities.

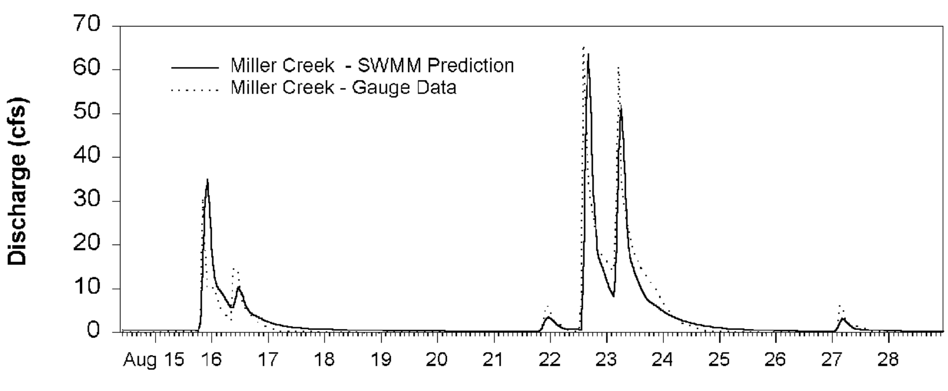


Figure 2. Predicted and observed discharge from middle reach of Miller Creek.

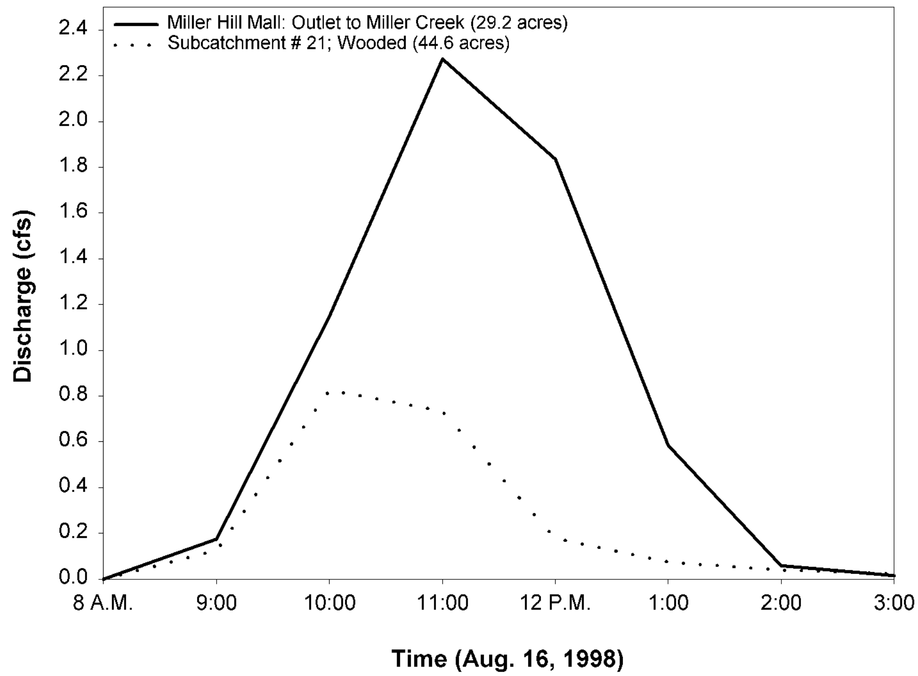


Figure 3. Comparison of hydrographs above and below mall development area on Miller Creek.

Land use planning resources for northern Wisconsin

Our second prototype effort involves compiling spatial data and other resources to support local units of government in citizen-based land use planning exercises. The objective of this exercise is to compile data and tools relevant to local-scale land use planning onto a CD-ROM for use by local units of government. Within the watershed, these are primarily towns and townships. The CD will contain the following elements:

- basic information describing the processes involved in planning
- spatial data – land use, transportation, rivers and lakes, natural features, political boundaries, and other data layers relevant to local-scale planning.
- planning tools – example surveys and ordinances language, zoning policies, development/preservation strategies, and other instruments that a local government could tailor to its specific needs
- landscape graphics - air photo or line drawing examples of different scenarios illustrating housing density and patterns (clustered or dispersed), riparian buffers, and other land use strategies. This would provide information on what a future landscape might look like under particular management strategies
- Example landscape plans developed by cooperating governments

The CD will be written in standard HTML code to allow access with available web browsers. Publicly available map viewing tools such as Arc Explorer@ will also be provide to provide basic querying of map products. The CD will be designed to stand alone, but will also contain links to web sites to access a broader range of information.

Education and citizen involvement are central to understanding the data and tools available on the CD (Figure 4). To assist in issue identification, example citizen surveys will be available from the Land-use Issues menu. Data will be available at both township and county levels, the latter allowing for contextual analysis of issues. Tools for assessing the suitability of planning elements will also be available. Example landscape plans will be available, along with information on implementing and administrating landscape plans once they are developed.

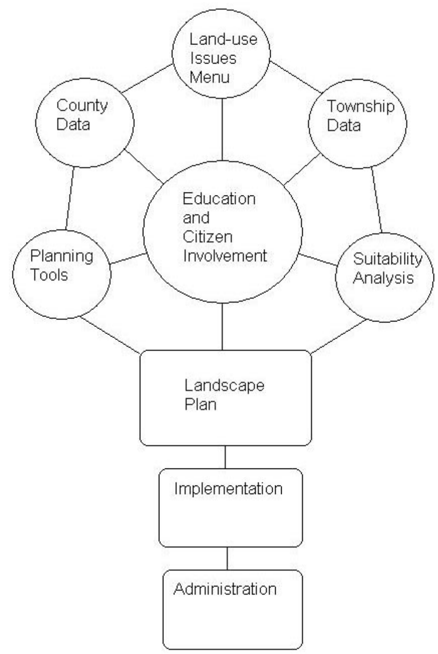


Figure 4. Schematic of LSDSS planning support CD architecture.

This CD-based resource will be developed and tested in cooperation with a number of partners from local and state governments. The anticipated completion and delivery date of the CD is June 2000. The information on the CD will also be available through the LSGIS web site.

Summary

The Lake Superior Decision Support project is an integrated effort to provide data and planning tools to citizens and local units of government to assist in land use planning efforts. This work should provide a means of reducing the cumulative impacts to the Lake Superior Basin through informed decision-making at the local level.

Acknowledgements

This project was funded by the US Environmental Protection Agency through the Minnesota Department of Natural Resources. We acknowledge the encouragement and support of the Lake Superior Ecosystem Cooperative, where the key ideas of the project originated. Mark White, Gerry Sjerven, Jesse Schomberg and Amy Trauger have been integral parts of this effort. A special acknowledgement goes to cooperator Mike Koutnik of Environmental Systems Research Institute for his support and contributions to data and kiosk development. Sandy Shultz of Ashland-Bayfield-Douglas-Iron Land Conservation Department and David Lonsdale of the Great Lakes Aquarium have provided valuable support and advice to the project.

A GIS-based Approach to Predicting Wetland Drainage and Wildlife Habitat Loss in the Prairie Pothole Region of South-central Canada

David Howerter
Institute for Wetland and Waterfowl Research
Ducks Unlimited Canada
P.O. Box 1160
Stonewall, MB, Canada R0C 2Z0 and

Lee Moats
Ducks Unlimited Canada
1606 4th Avenue Regina
Saskatchewan, Canada S4P 3W7

Background

The prairie pothole region covers approximately 870,000 km^2 of the north-central United States and south-central Canada (Batt 1996). The area extends from the tall grass prairies of northwestern Iowa west to the short grass prairie of southern Alberta and north to the boreal transition aspen (*Populus tremuloides*) parklands of central Saskatchewan (Figure 1). Approximately 13,000 years ago, retreating glacial ice revealed a rolling terrain comprised of end, ground and lateral moraines; glacial drift and till; and eskers (Winter 1989, Batt 1996). Remnant ice remained within the till after the main body of the glaciers retreated. As this ice melted, kettles were created. Subsequently, these kettles filled with water and became the potholes of modern vernacular. (Pielou 1991, Batt 1996). The majority of potholes are not connected by a natur

Figure 1. Prairie Pothole Ecoregion

Prairie wetlands serve a variety of hydrological/ecological functions including storage and control of surface water (Winter 1989, Miller and Nudds 1996, Murkin 1998), recharge of groundwater supplies (Winter 1989, LaBaugh et al. 1998, Murkin 1998), filters for sediments and agricultural chemicals (Grue et al. 1989, Neely and Baker 1989, Gleason and Euliss 1998, Goldsborough and Crumpton 1998, Murkin 1998), sinks for excess nutrients (Crumpton and Goldsborough 1998, Murkin 1998), and habitat for a wide array of invertebrate, fish and wildlife species (e.g., Batt et al. 1989, Fritzell 1989, Peterka 1989). This area is particularly important to waterfowl where > 50% of the continental duck populations are annually produced (Batt et al. 1989).

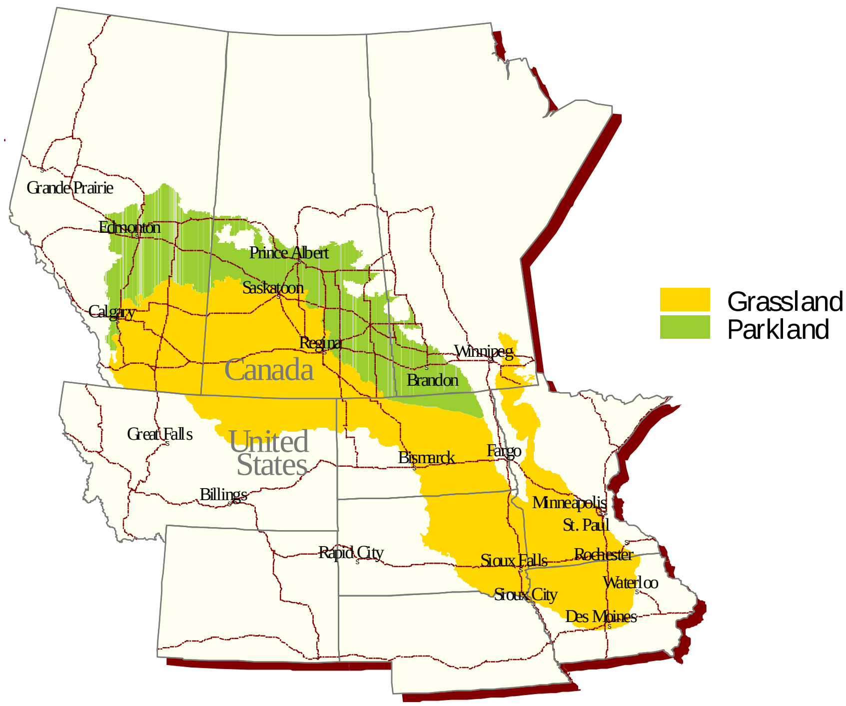


Figure 1. Prairie Pothole Ecoregion

European settlement of the PPR began around 1878. Farming was the primary motivation for settling the area and agriculture remains a dominant economic force in the region (Leitch 1989, Leitch and Fridgen 1998). Accompanying this expansion of agriculture was a dramatic transformation of the landscape. Samson and Knopf (1994) estimate that as little as 23% of the pre-settlement prairie remains in Canada. Conversion in the parklands (Figure 1) has been even more extensive (Turner et al. 1987). Similarly, > 70% of the wetlands in the PPR have been drained or severely degraded (Turner et al. 1987, Dahl 1990, Batt 1996).

These changes to the landscape have had consequences for a variety of ecosystem functions throughout the region. For example, Miller and Nudds (1996) attributed increased flooding within the Mississippi River valley over the past decades to reduced natural upland vegetation and wetland drainage Other hydrologic functions such as groundwater recharge and the ability to filter agrochemicals are similarly reduced by drainage. The focus of this paper, however, is the effect these changes have had on the PPR's ability to support wildlife populations--specifically upland-nesting duck species.

In 1986, in response to declining duck populations attributed to land use changes throughout North America and exacerbated by prolonged drought on the prairie nesting grounds, the governments of the United States and Canada signed into law the North American Waterfowl Management Plan (NAWMP; joined in 1994 by Mexico). The goal of the Plan was to return waterfowl populations to the levels of the mid-1970's—a period with abundant waterfowl populations. The Plan was structured as a number of joint ventures targeting either species of concern or regions of special importance to waterfowl populations. Two of these joint ventures (Prairie Habitat - Canada; Prairie Pothole - U.S.) focus on restoring nesting habitats within the PPR.

During the planning stages for the Prairie Habitat Joint Venture (PHJV), one of the underlying tenets was that policy changes would ensure that existing habitats remained intact. Ultimately, this has not occurred. Draining wetlands and conversion of lands for agriculture continues unabated—particularly in Canada. For instance, between 1971 and 1996 natural land has declined by 2.4 million hectares within the prairie pothole region of Canada (Figure 2) while there has been a corresponding increase in cultivation (Figure 3). Likewise, wetlands continue to be lost. For example, in a single region near Wadena, Saskatchewan, the number of wetland hectares declined from 3,019 to 563 between 1949 and 1998. Between 1980 and 1998, 41% of the remaining ha were drained (Ducks Unlimited, unpublished data). In the U.S. portion of the PPR, nearly 15,000 wetland ha continue to be degraded annually, despite provisions that have linked wetland protection to agricultural subsidies (Tiner 1984). If, as currently proposed, conservation provisions are de-coupled from agricultural subsidies, the incentive to retain wetlands will be further weakened (J. K. Ringelman, Ducks Unlimited, Inc., Bismarck, ND).

Exacerbating the problem of continued loss of existing upland and wetland habitats, recent assessments of PHJV habitat programs indicate that restored grasslands may be less productive than existing native/naturalized cover for nesting ducks (Institute for Wetland and Waterfowl Research, unpublished data). Therefore, a paradigm shift was in order for habitat managers faced with a constantly eroding habitat base coupled with restoration efforts that may be less effective than envisioned. Instead of primarily converting cropland back to grassland vegetation, it has become increasingly clear that preserving existing habitats may be a more effective strategy. However, limited conservation resources preclude the possibility of securing all existing cover, and conversion rates vary spatially. Therefore, to prioritize which habitat

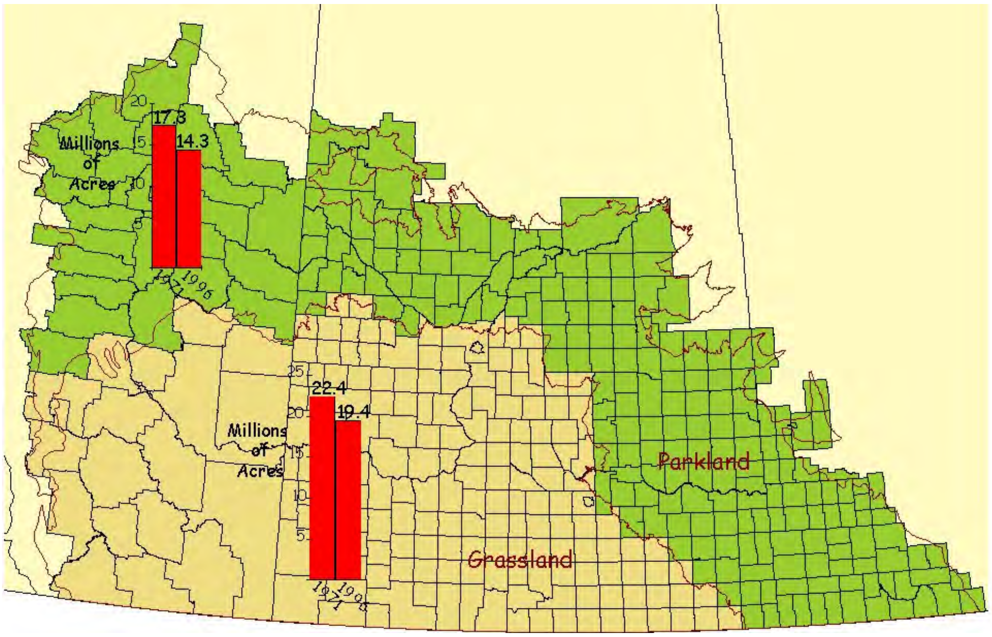


Figure 2. Decline of natural land in the Canadian prairie pothole region 1971-1996. (Source: Statistics Canada 1996)

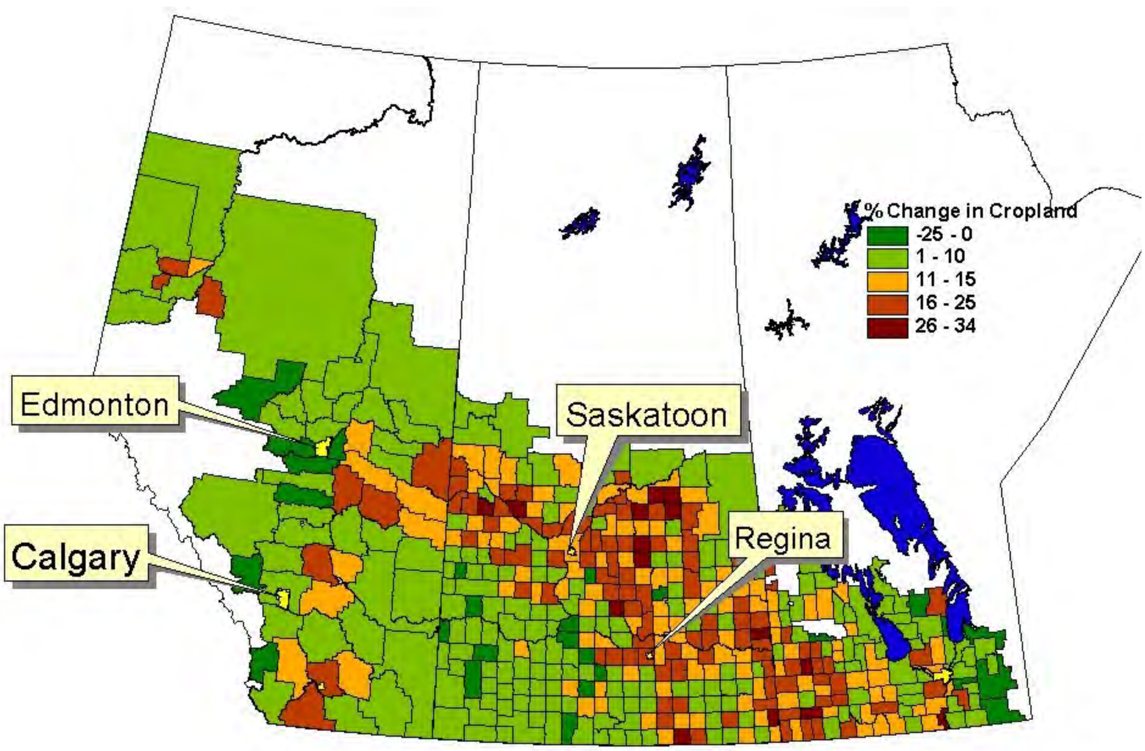


Figure 3. Change in cropped areas 1971-1996 (Source: Statistics Canada Agriculture Census 1996)

parcels to secure, habitat managers need to identify the areas with the highest potential for waterfowl production that are also at high risk of conversion to agriculture. Our objective, then, is to develop a spatially-explicit model to project where land conversion is most likely. This model ideally will be will be hierarchical allowing identification of "high risk" areas at both the regional and local scales. For this paper, we discuss our plan for the regional model in some detail and touch only briefly on additional considerations for the local model.

Hypothesized relationships

For conservation dollar investment decisions to be made wisely, resource managers require a spatially-explicit tool to project which areas are most likely to be converted to cropland in the near future. To allow quantification of the relationships that will allow optimization of resource expenditures, risk of conversion to agriculture can be thought of as an expected rate of habitat loss. In essence, the decision about where to expend resources on securement of existing habitats within the foreseeable planning horizon takes the form:

$\boldsymbol{\delta P = f\,(P, Cs, R)}$, where:

- $\boldsymbol{\delta P}$ = change in duck production/habitat expenditure
- **P** = current duck production (***f*** **[Amount of suitable upland cover, wetland abundance]**)
- $\mathbf{C_s}$ = site-specific cost of conservation, and
- **R** = risk = (**E[loss of habitat]**).

Our hypothesized relationships are depicted in Figure 4. In Figure 4a, we've speculated that $\mathbf{C_s}$ for a piece of land is likely correlated with its inherent fertility. Therefore, the least expensive land most likely has low productivity for agriculture **and** ducks (e.g., Manitoba's Interlake). Alternatively, the most costly (productive) land likely has been already largely converted to agricultural and/or residential purposes. As a result, the areas that are currently most productive for duck production likely occur on land with marginal suitability for farming and at intermediate land costs.

Following similar reasoning, we have hypothesized that **R**, too, is highest for lands with intermediate costs (Figure 4b); poor quality land (low cost) likely isn't worth converting to cropland for economic reasons, and high cost land likely already has been converted.

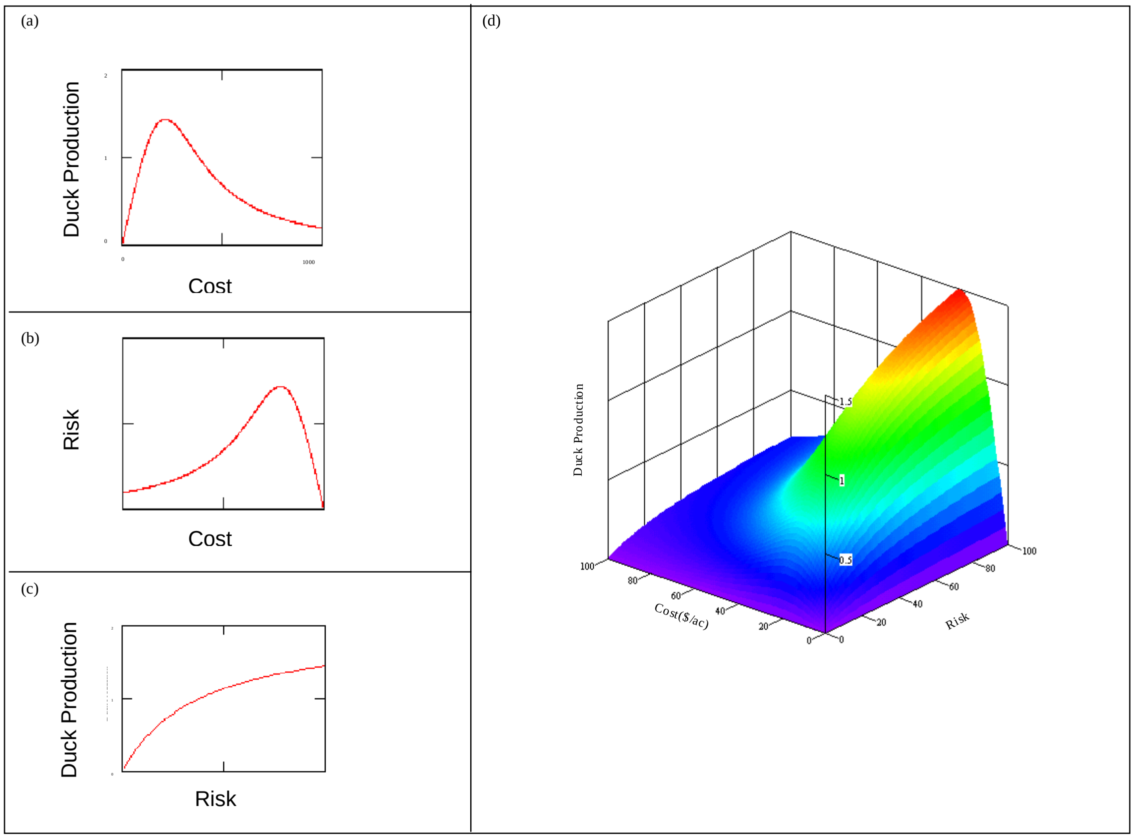


Figure 4. Hypothesized relationships between: (a) cost of securing a given piece of habitat and potential duck production, (b) the risk of habitat being converted to agriculture and cost of securing the habitat, (c) potential duck production and risk of conversion, and (d) all three components.

Because we have hypothesized that both **P** and **R** are unimodal in relation to **Cs**, **P** and **R** are positively associated when plotted against each other (Figure 4c). Combining these relationships into a single model results in a relationship with one realization that may resemble Figure 4d. Figure 4d depicts several interesting points. First, the best land for duck production is also the land at highest risk of conversion. Second, happily, this land likely is of only intermediate cost.

Model Development

Several techniques have been proposed for predicting land use change. The simplest of which is to use historical trends to project forward in time making the assumption that past trends will continue into the future. Empirical models based on posited driving forces may, however, provide better predictions (Robinson et al. 1994) and allow for exploration of causative factors.

To develop our model we plan to use relatively recent (e.g., since 1971) conversion data (source: Statistics Canada – Census of Agriculture) and a number of candidate explanatory variables (Table 1). An information theoretic approach (e.g. Akaike's Information Criterion; Burnham and Anderson [1998]) will be used with regression analysis to select the most parsimonious set of explanatory variables.

Table 1. List of candidate variables useful for predicting land conversion to agriculture in the prairie pothole region.

Candidate variable	**Hypothesized relationship**	**Data Source**
Topography	Areas with relatively steep topography are difficult to cultivate and/or drain	Canada Surveys and Mapping
Landowner demographics	Conversion most likely when land changes ownership	Statistics Canada
Soil type	Related to cropping practices	Canada Soil Information System
Crop profitability	Input costs and crop prices influence an individual's decisions on how best to use their land	Provincial Agriculture Departments
Proximity to an existing water conveyance	Proximity to water conveyance makes drainage less costly	Provincial water boards

Topography and proximity to an existing water conveyance both relate to the cost of conversion, while soil type and crop profitability both relate to the benefit received by the manager for converting. Crop profitability will require a separate "sub-model" with spatially-explicit crop prices and input and transportation costs. Alternatively, land prices may serve as a suitable and less data-intensive proxy for crop profitability for model development.

Evaluation/monitoring

The model we have described will be developed using (recent) historical data. We realize, however, that extrapolating forward based on past relationships is inherently dangerous. Therefore, a series of monitoring stations will be established throughout the areas targeted as important for nesting ducks. Through an adaptive process uncertainty will be reduced in time.

Using empirical Bayesian techniques, these stations will allow us to iteratively update model parameters. This process will yield immediate results with increasing confidence over time.

Local

Because much of the data used to build the regional risk model will be at a fairly coarse scale (50-200 km^2), additional information likely will be needed at the finest geographic scale to accurately predict the risk of an existing patch of habitat being converted to cropland. Factors such as landowner demographics and topography will likely still be important predictors, but added factors such as size of the habitat patch and surrounding land use will also likely be significant. Also, while true crop profitability might be predictable through entirely empirical information, it almost certainly will be much more difficult to predict the vagaries of individual landowner decisions. To attempt to explain how decisions are made, new data will need to be gathered to determine how producers determine profitability when clearing, breaking or draining land.

Integrating the Model: A Decision Support System

The model that we have described in this paper represents one component of a decision support system designed to optimize expenditures of conservation dollars. Simultaneously, spatially-explicit models to predict duck production and the cost of securing important habitat areas also will be developed (Figure 5). The combined output from these 3 components should allow habitat managers to make informed decisions about how best to prioritize expenditures.

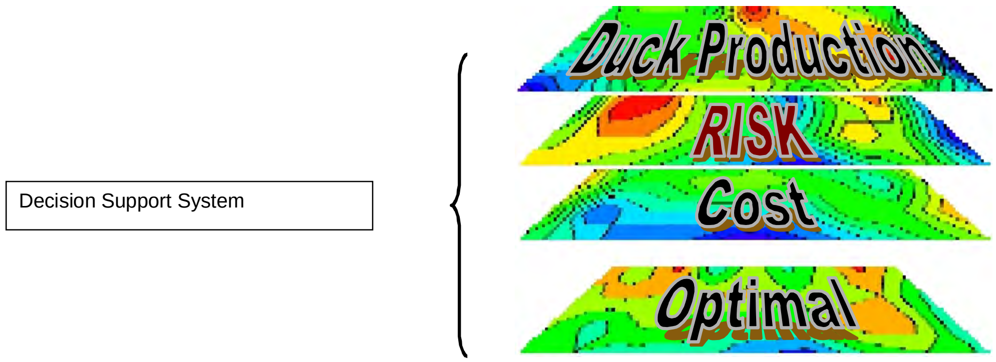


Figure 5. Conceptual model of Decision Support System

Acknowledgements

We thank K. Guyn for helpful comments on a previous draft of this paper. J. Holland and B. Kazmerick provided technical assistance and assisted with preparation of the figures. K. LePoudre provided unpublished data.

References

Batt, B. D. J. 1996. Prairie ecology—prairie wetlands. Pages 77-88 *in* F. B. Samson and F. L. Knopf, editors. Prairie conservation: preserving North America's most endangered ecosystem. Island Press, Washington, D.C. 339pp.

_____, M. G. Anderson, C. D. Anderson, and F. D. Caswell. 1989. The use of prairie potholes by North American ducks. Pages 204-227 *in* A. van der Valk, editor. Northern prairie wetlands. Iowa State University Press, Ames. 400pp.

Burnham, K. P. and . R. Anderson. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York. 353pp.

Crumpton, W. G. and L. G. Goldsborough. 1998. Nitrogen transformation and fate in prairie wetlands. Great Plains Research 8:57-72.

Dahl, Thomas E. 1990. Wetlands losses in the United States 1780's to 1980's.U.S. Department of the Interior, Fish and Wildlife Service, Washington,D.C. Jamestown, ND: Northern Prairie Wildlife Research Center Home Page. http://www.npwrc.usgs.gov/resource/othrdata/wetloss/wetloss.htm (Version 16JUL97).

Fritzell, E. K. 1989. Mammals in prairie wetlands. Pages 268-301 *in* A. van der Valk, editor. Northern prairie wetlands. Iowa State University Press, Ames. 400pp.

Gleason, R. A., and N. H. Euliss, Jr. 1998. Sedimentation and prairie wetlands. Great Plains Research 8:97-112.

Grue, C. E., M. W. Tome, T. A. Messmer, D. B. Henry, G. A. Swanson, and L. R. DeWeese. 1989. Agricultural chemicals and prairie pothole wetlands: meeting the needs of the resource and the farmer—U.S. perspective. Transactions of the North American Wildlife and Natural resources Conference 54:43-58

Goldsborough, L. G., and W. G. Crumpton. 1998. Distribution and environmental fate of pesticides in prairie wetlands. Great Plains Research 8:73-95.

LaBaugh, J. W., T. C. Winter, and D. O. Rosenbery. 1998. Hydrologic functions of prairie wetlands. Great Plains Research 8:17-37.

Leitch, J. A. 1989. Politicoeconomic overview of the prairie potholes. Pages 3-14 *in* A. van der Valk, editor. Northern prairie wetlands. Iowa State University Press, Ames. 400pp.

_____, and P. Fridgen. 1998. Functions and values of prairie wetlands: economic realities. Great Plains Research 8:157-168.

Miller, M. W., and T. D. Nudds. 1996. Prairie landscape change and flooding in the Mississippi River valley. Conservation Biology 10:847-853.

Murkin, H. R. 1998. Freshwater functions and values of prairie wetlands. Great Plains Research 8:3-15.

Neely, R. K. and J. L. Baker. 1989. Nitrogen and phosphorus dynamics and the fate of agricultural runoff. Pages 92-131 *in* A. van der Valk, editor. Northern prairie wetlands. Iowa State University Press, Ames. 400pp.

Peterka, J. J. 1989. Fishes in northern prairie wetlands. Pages 302-315 *in* A. van der Valk, editor. Northern prairie wetlands. Iowa State University Press, Ames. 400pp.

Pielou, E. C. 1991. After the ice age: the return of life to glaciated North America. University of Chicago Press, Chicago. 366p.

Robinson, J., S. Brush,, I. Douglas, T.E. Graedel, D. Graetz, W. Hodge, D. Liverman, J. Melillo, R. Moss, A. Naumov, G. Njiru, J. Penner, P. Rogers, V. Ruttan, and J. Sturdevant. 1994. Land-use and land-cover projections: report from working group C. Pages 73-92 *in* W. B. Meyer and B. L. Turner, II, editors. Changes in land use and land cover: a global perspective. Cambridge University Press, Cambridge, U.K. 537pp.

Samson, F. and F. Knopf. 1994. Prairie conservation in North America. Bioscience 44:418-421.

Tiner, R. W., Jr. 1984. Wetlands of the United States: current status and recent trends. U.S. Dept. of the Interior, Fish and Wildlife service, Washington, D.C. 59pp

Turner, B. C., G. S. Hochbaum, F. D. Caswell, and D. J. Nieman. 1987. Agricultural impacts on wetland habitats on the Canadian Prairies, 1981-1985. Transactions of the North American Wildlife and Natural Resources Conference 52:206-215.

Winter, T. C. 1989. Hydrologic studies of wetlands in the northern prairie. Pages 16-54 *in* A. van der Valk, editor. Northern prairie wetlands. Iowa State University Press, Ames. 400pp.

Onsite Wastewater Management Program in Hamilton County, Ohio – An Integrated Approach to Improving Water Quality and Preventing Disease

Tim Ingram, Terry Hull, Travis Goodman, and Staff
Division of Water Quality, Hamilton County, Ohio

Background

Postwar economic and population growth of the 1950s launched Hamilton County, Ohio, families into suburbia. Ohio's most southwestern county and home of Cincinnati expanded typically. Construction of suburban residential housing outstripped the development of traditional urban infrastructure. But extended aeration technology made it possible to put a miniature wastewater treatment plant in the yard of every new homeowner. Home aeration units were marketed as virtually maintenance-free devices capable of producing effluent with the quality of drinking water. These fallacies, in concert with the natural limitations of the sites being developed and a community willing to believe that which was too good to be true, brought public health consequences.

Since most of Hamilton County escaped the advance of the last (Wisconsinan) glacier, its steep shale ridges remain covered with thin top soils and slowly-permeable silty-clay subsoils. Only small areas in the valleys of the three principal streams are underlain by the more-permeable glacial deposits. Seasonal rain and snow melt runoff cut the ridges with innumerable erosion channels that join to form intermittent streams. Conditions that limited the use of leaching devices appeared to be optimal for use of aeration units with surface discharges. Supported by exaggerated claims of operational efficiency, public officials were quick to accept the technology as a way to support new tax-generating development without the cost of sewer construction. Developers enthusiastically supported a public policy that reduced their capital costs and raised profits. Homeowners *believed* that they had an effective, low cost, state-of-the-art sewage treatment device. With discharges running to any downhill drainage way, out of sight/out of mind became the prevailing attitude, and it took a long time to change.

By the mid-1980's, sewage contamination of the West Fork of the Mill Creek elicited strong public reaction. When a hepatitis A outbreak occurred in 1989, the public clamor over lack of public sewers, ineffective sewer system operation, and poorly functioning home aeration units intensified. Public alarm further increased when a child playing in a stream polluted by aeration systems caught a rare protozoan infection and was hospitalized for days.

By the early 1990's, the total number of units was unknown, but estimates ranged from 20,000 to 40,000 units. Public perception regarding the public health impact was markedly changed. An irate homeowner persisted until the Hamilton County General Health District's Board of Health officially declared her neighborhood a public health nuisance and requested the State to order construction of a public sewer. Public concern peaked when the Board of Health authorized use of aeration units in a proposed subdivision where the effluent would drain directly into a downstream homeowner's recreation lake. The resulting legal action brought Ohio Environmental Protection Agency and Ohio Department of Health sanctions against the Board of Health. A political response occurred in 1993 as Board members were replaced and a new Health Commissioner (pubic health officer) was employed. The new Board determined that water quality improvements and waterborne disease prevention were dependent on the effective use of household sewage disposal systems. Thus, the Board made a commitment to:

- Update its household sewage regulation, which had been in force since 1959;
- Reconsider its previous approval actions regarding the usage of aeration sewage systems in residential subdivisions;
- Inventory and evaluate existing home sewage systems;
- Build community partnerships and educate homeowners about system operation and maintenance requirements;
- Expedite complaint investigations; and
- Strengthen enforcement when sewage system repairs are necessary.

This paper discusses the establishment of an operation permit program which was developed to achieve several of these commitments. The purpose of this paper is to describe the components of the operation permit program and the steps that have been taken to improve water quality in Hamilton County streams.

Building the Operation Permit Program

<u>Policy Development</u>

In Ohio, local health districts have sewage system permitting authority for one-, two-, and three-family dwellings. These districts enforce either the minimum state code or more stringent regulations adopted by the local board of health. Under the minimum code, a newly constructed system is automatically permitted for operation upon final construction approval. This operation permit remains in force until it is revoked by board action. Such permits are rarely revoked.

Some health districts issue operation permits for existing sewage systems. In 1994, the Hamilton County General Health District (HCGHD) initiated a comprehensive operation permit program. Under this program operating permits are issued for both new and existing household sewage disposal systems.

On December 13, 1993, the Board of Health adopted a new household sewage disposal code and created a new division to enforce it. Previously, the sewage permitting functions were a part of the Health District's plumbing division. To implement the Board's commitments, the Division of Water Quality and Waste Management was established. The main function of this division was to implement *Regulation 529, the Household Sewage Disposal Code*.

To guide the new program, the Hamilton County Board of Health established the following policies:

- No permit fee will be billed until an inspection of the household sewage disposal system(HSDS) is completed and the owner is provided with a written inspection report.
- All household sewage disposal systems (HSDS) with electrical components (aeration units, etc.) are subject to an annual inspection and, if the system is found to be operating properly, a one-year operation permit will be issued.
- All non-mechanical or non-electrical household sewage disposal systems (HSDS) will be inspected once every three years and, if found to be operating, a three-year operation permit will be issued (effective on Sept. 11, 1995).
- All newly-constructed dwellings with newly-installed household sewage disposal systems (HSDS), permitted and approved by the HCGHD, will be exempted from the operation permit program for a period of two years following construction approval.

Critical to the success of the operation permit program was the establishment of inspection criteria for determining proper operation. Considering the varying ages of the systems in use and the different standards under which they were manufactured and installed, a balanced approach was needed that would achieve water quality objectives while maintaining public and political support. The Health Commissioner and staff were charged with the development of the inspection criteria.

It was decided that system performance would be evaluated by observation rather than effluent sampling. This decision took into account the fact that many HSDS had no separate access for sampling, as well as the unlikelihood that most older systems could meet effluent quality standards established by the Board of Health. Also, considered was the cost to sample as many as 20,000 sewage systems and that many individual systems were connected to common collector lines. Collector lines function as private sewers that transport effluent from several homes. The ten observational criteria selected to establish improper performance for mechanical HSDS, thus resulting in the designation “disapproved” are:

1) Motor missing,
2) Motor inoperable (cold),
3) Motor not drawing air or insufficiently drawing air,
4) Broken lid(s), i.e., piece missing or broken to the extent that it allows entrance of surface water, or lid cannot be lifted without collapse; decayed metal grating,
5) Flooded filter,
6) Visual evidence of septic sewage, i.e., black, odorous,
7) Visual evidence of electric service problem,
8) Components are not functioning in accordance with design standard,
9) Discharge creates a public health nuisance, and
10) An access riser has not been brought to grade over each compartment requiring maintenance.

Fee Structure

Policies concerning program fees have changed as program activities evolved and knowledge was gained regarding operating costs and the willingness of HSDS owners to pay. At the outset, the Board of Health established a single operation permit fee of $40.00. As program implementation progressed, two problems became apparent. First, there was no way to recover the cost of the additional inspections required to achieve and confirm system repairs when the initial inspection resulted in disapproval. Second, no permit was issued to HSDS owners who failed to pay their permit fee. Consequently, those who were delinquent soon forgot about the need for an operation permit as well as their indebtedness. The Health District then entered a new era-marked by the need for dunning letters and small claims court appearances-in which it acted somewhat like a small utility company.

In July of 1995, the fee structure was changed based on a report issued to the Board of Health entitled, *Aeration Sewage Disposal Systems in Hamilton County*. This report evaluated the progress made to date. Based on community input, the Board implemented a variable fee structure. The annual fee covering the initial inspection and permit was reduced to $30.00. A reinspection fee of $30.00 was established for second and subsequent re-inspections. As an inducement for HSDS owners to provide regular maintenance, the second reinspection fee was set at $15.00 for owners holding a maintenance contract with a registered and bonded service company.

Delinquent permit fees rose as high as 62% but annually averaged around 13% for the 1994-1995 time period. This amounted to $32,320 of uncollected permit fees at the end of 1995. In 1996 a collection agency was retained by the Health District at a cost of $4.95 per outstanding account. Also, a $10 late fee was assessed against each delinquent account. Delinquent fees dropped to an average of 9% in subsequent years. Many homeowners were more threatened by a bad credit rating than criminal prosecution. Nonetheless, the percentage of unpaid accounts was still too high. In 1998, the HCGHD worked with local Ohio legislators, and legislation was passed allowing for unpaid operation permit fees to be certified by the Health Commissioner to the County Auditor for placement as a lien on the property. With the passage of this legislation, the "playing field" has been leveled so that all HSDS owners have a financial stake in the operation permit program.

Staff Development

The development and implementation of policies regarding Health District staff and staff activity has been crucial to successfully carrying out the operation permit program. These policies have been established largely by management and include issues such as, personnel qualifications, inspection protocol, and training.

Field work was done initially by registered sanitarians and sanitarians-in-training. During the first year staff turnover was unacceptably high. Newly-graduated environmentalist quickly gained valuable field experience, but were just as quickly tired by the large volume of inspections and the repetitive, sometimes confrontational nature of the work. Experienced sanitarians soon found the work not challenging. Management found the solution rested in the employment of technicians as inspectors. With specialized training and sanitarian oversight, these staff have maintained their enthusiasm and perform their inspections in a highly competent and efficient manner. The water quality technicians have worked well with HSDS owners and the repair contractors, too.

The effectiveness of the inspection staff can be attributed to the development of operating rules which emphasize thorough training, clear identity, and consistency in performance of duties. Each inspector receives detailed training regarding the operation of each brand of aeration unit that he/she will inspect in the field. Contact with repair contractors is encouraged, and where possible, attendance at manufacturer’s training school is supported. Inspector training in communication and public relations is also provided, while membership in professional environmental associations is encouraged.

Inspectors adhere to a dress code wearing distinctively colored shirts and jackets. Clearly visible photo identification badges are worn. Each inspection begins with a stop at the front door to explain the purpose of the visit to the citizen, followed by a brand-specific standardized inspection of the sewage system, and, if the owner is present informing him of the inspection findings.

The Inventory – What Was Found

Aeration sewage systems are a high maintenance type of household sewage disposal system. These systems contain electrical components and filters which require maintenance by the homeowners. Homeowners were unaware, for the most part, of these maintenance needs.

The media reported that 40,000 aeration sewage systems discharged untreated waste water into homeowners' backyards. However, no one knew for sure the exact number nor their operating condition. Installation permit records were not complete. While citizens worried about disease and pollution, an Ohio EPA connection ban and class action lawsuit put the Hamilton County Board of Health in a reactive mode. There was a rush to inventory and evaluate the operating status of all existing aeration sewage systems first. The Board of Health directed the Health Commissioner to address the following questions:

1) How many aeration sewage systems exist in Hamilton County?
2) What are the manufactured types?
3) Where are the aeration sewage systems located?
4) What is the operational status of these systems?
5) How much water pollution and how many health nuisances have been created from these aeration sewage systems?

From January 1994 through July 1995, sanitarians and water quality technicians blanketed the County inspecting home aeration systems. The staff visited over 10,000 properties – many on several occasions – to find aeration systems, locate collector lines, consult with homeowners, or reinspect systems for compliance. Inspection sheets were filled out and information entered into a custom-designed database for ease of tracking.

Within eighteen months 9,145 home aeration sewage systems were located and inspected. Six manufactured-types of home aeration sewage systems were found. (See Table 1). The high percentage of the Cavitette brand in use was a significant concern because there was no active manufacturer or replacement parts available. The Cavitette brand was manufactured locally in the late 1950's through the 1960's. The design of this system was based on less stringent standards than the National Sanitation Foundation Standard 40.

Table 1
Percent of Total Home Aeration Units by Manufacturer

Manufacturer	**Percent**
Cavitette	22.4
Coate	10.6
Jet	36.6
Multiflo	2.4
Norweco	0.1
Oldham	27.9

Of the 9,145 aeration systems inspected, 34 percent or 3,077 were disapproved during this time period. Table 2 summarizes the number of that had been repaired by July 1995 and the average number of reinspections to obtain compliance.

Table 2
Disapproved Systems with Completed Repairs

	Number of systems	Number of reinspects	Mean reinspects to compliance
Totals	2,741	3,861	1.41

Many of the home aeration sewage systems connected to a common yet private sewer line known as a collector line. These collector lines served as few as two homes or as many as 50 homes. The effluent from the collector lines discharged into storm water sewers, ephemeral streams, and onto the ground surface. None of the collector lines had National Pollutant Discharge Elimination Systems (NPDES) permits. It had been observed and reported that the discharges from these collector lines had deteriorated the water quality in the county's streams and waterways. Prior to 1994, there was no documented wastewater monitoring of these collector lines. However, in the summer of 1994 as a part of the operation permit program, Project CLEAN (Collector Line Evaluation and Assessment of Needs), was implemented. Wastewater samples were taken from 197 collector lines. The samples were analyzed for biochemical oxygen demand (BOD), suspended solids (SS), and fecal coliform bacteria. The data results from the initial round of sampling are shown in Table 3.

Table 3
Project CLEAN Effluent Sample Data

Township/ Municipality	Number of samples	Fecal under 5,000	Fecal 5,000 to 50,000	Fecal over 50,000	Fecal TNTC	Mean SS	Mean BOD	No. meeting standards
Anderson	7	2	2	0	3	273	11	0
Colerain	47	13	13	7	14	100	37	12
Crosby	1	0	0	1	0	230	74	0
Delhi	6	2	2	0	2	44	10	0
Green	104	34	29	29	12	83	32	17
Harrison	4	2	0	0	2	208	30	0
Madeira	13	1	6	3	3	67	50	1
Miami	7	0	0	5	2	30	29	0
Springfield	4	0	1	1	2	60	32	0
Sycamore	1	1	0	0	0	10	8	1
Symmes	2	0	0	2	0	26	604	0
Whitewater	1	0	0	0	1	148	40	0
Total Samples	197	55	53	48	41	197	197	31

Water Quality Standards	**HCGHD**	**ODH**	**Ohio EPA**
Fecal Coliform (colonies/100 ml)	5,000	none	1,000
Suspended Solids (mg/l)	40	40	12
Biochemical Oxygen Demand (mg/l)	20	20	10
Ammonia Nitrogen (NH3) (mg/l)	none	none	1 (summer), 3 (winter)
Dissolved Oxygen (mg/l)	none	none	6

When the Board of Health adopted a new sewage code in late 1993, they also codified discharge standards for suspended solids, biochemical oxygen demand, and fecal coliform bacteria, which are:

- BOD: 20 mg / L
- Suspended Solids: 40 mg / L
- Fecal Coliform Bacteria: $\leq$ 5,000 colony forming units / 100 ml

Table 3 reveals that 31 or 16% of the collector lines met the Board of Health discharge standards. Nearly three-fourths (72%) failed to meet the standard for fecal coliform bacteria. This was not surprising since disinfection devices had not been installed on many of the individual aeration sewage systems. The initial inventory and evaluation of aeration sewage systems was completed.

Non-mechanical Sewage Systems

A contention by County citizens early in the program was that the Health District ignored the pollution created by non-mechanical household sewage disposal systems, like leach lines and dry wells. Once the initial inventory and evaluation of aeration (or mechanical) sewage systems was completed, the Health District staff began the inventory and evaluation of non-mechanical sewage systems in 1996. Table 4 shows the number and the type of non-mechanical systems inspected in 1996 and 1997, and the number of total units that failed the initial inspection.

Table 4
Non-mechanical Sewage Systems

1996		**1997**	
Type	**Number**	**Type**	**Number**
Drywell	783	Drywell	162
Leach Lines	370	Leach Lines	476
Subsurface Sandfilter	181	Subsurface Sandfilter	1
Privy	1	Privy	212
Other	84	Other	27
Total	1419	Total	878

Total initial failure 59 (4.2%)

Total initial failure 117 (13.3%)

The inventory and evaluation of non-mechanical sewage systems is a work-in-progress. By May 17, 1998, a total of 2605 non-mechanical sewage systems had been assessed.

The following inspection criteria are used to approve or disapprove a non-mechanical household sewage disposal system.

Approve if:

a) System components can be located and
b) The system uses in-soil dissipation and there is no surface seepage of gray, malodorous effluent (minor seasonal wetness is acceptable).
c) The system uses surface discharge, the discharging effluent is clear, and not malodorous, does not pond on the inspected property or on adjacent property, and the discharge pipe is accessible with sufficient freeboard to allow collection of an effluent sample.

Disapprove if:

a) Gray malodorous sewage is seeping to the ground surface creating a nuisance.
b) Gray, malodorous sewage is discharging from the sewage system to an adjacent property or roadside drainage way.
c) Any similar condition is occurring which is creating public health nuisance.

Upon the initial inspection, the percentage of non-mechanical systems disapproved overall averaged 9%.

The inventory and evaluation of these sewage systems continues. The District did not meet its goal of inspecting and inventorying all non-mechanical sewage systems by December 31, 1998. On March 8, 1999, the Board of Health changed the non-mechanical permit from a three-year permit to a five-year permit.

<u>Building Community Partnerships</u>

All the media attention over the lack of governmental oversight and widespread pollution did not convince the vast majority of citizens with home sewage systems that the operation permit program was necessary. Media and citizenry criticism for the Health District's historic lack of oversight was replaced by concern about Health District persistence in requiring sewage system repairs and payment of permit fees. During the first eighteen months of the program, the Health

District staff handled 20 to 40 telephone calls per week. About 265 citizens felt strongly enough to write letters protesting the program. The theme of the letters ranged from why inspect septic systems to a governmental invasion of their property rights and privacy.

The Health District initiated a community education program in order to gain public acceptance of the operation permit program, to improve political support, and to teach homeowners about HSDS operation and maintenance. The education program utilized a variety of strategies. Brochures were developed and mailed to homeowners and handed out at neighborhood festivals and community halls. Numerous press releases and editorials in the local newspapers about the benefits of the operation permit program to the community were published. Health District staff conducted presentations at neighborhood gatherings or backyard barbecues about the importance of HSDS.

The Metropolitan Sewer District (MSD) and the Health District forged a partnership and collaborated at several public meetings. The two agencies worked in concert to extend public sewers into those watersheds where sewage nuisances were prevalent and HSDS upgrades were not feasible.

A public sewer assessment credit was established by the Board of Hamilton County Commissioners. This credit was the brainstorm of MSD and Health District officials as well as a group of western Hamilton County business leaders and elected officials. The credit program stated that all single family dwellings with HSDS in existence on or prior to September 20, 1995, were eligible for $5,000.00 credit towards their public sewer assessment. The public sewer assessment credit helped convince many homeowners to sewer their neighborhoods.

Using GIS

In 1996, the Health District began using a geographical information system (GIS) known as CAGIS (Cincinnati Area Geographic Information System). CAGIS technology allows the Health District to place all home sewage systems, stream quality data, and communicable disease data on computer generated maps. CAGIS technology allows layers of information to be overlaid on top of each other in order to carry out analyses. For instance, the public sewer system layer can be overlaid with the County home sewage system layer. This allows the user to quickly observe the proximity of public sewers to home sewage systems. Other layers of data, such as stream data and communicable disease information, are overlaid to look for patterns or clusters of disease or pollution associated with home sewage systems. The CAGIS technology is a powerful public health surveillance tool for targeting resources.

Results – Has the Operation Permit Program Made a Difference?

Data comparison between two time periods, 1994 - 1995 and 1996-1997, will be used to reveal program successes and failures. Table 5 compares pass / fail inspections of home aeration systems for the two time periods.

Table 5
Pass / Fail Inspections of Home Aeration Sewage Systems

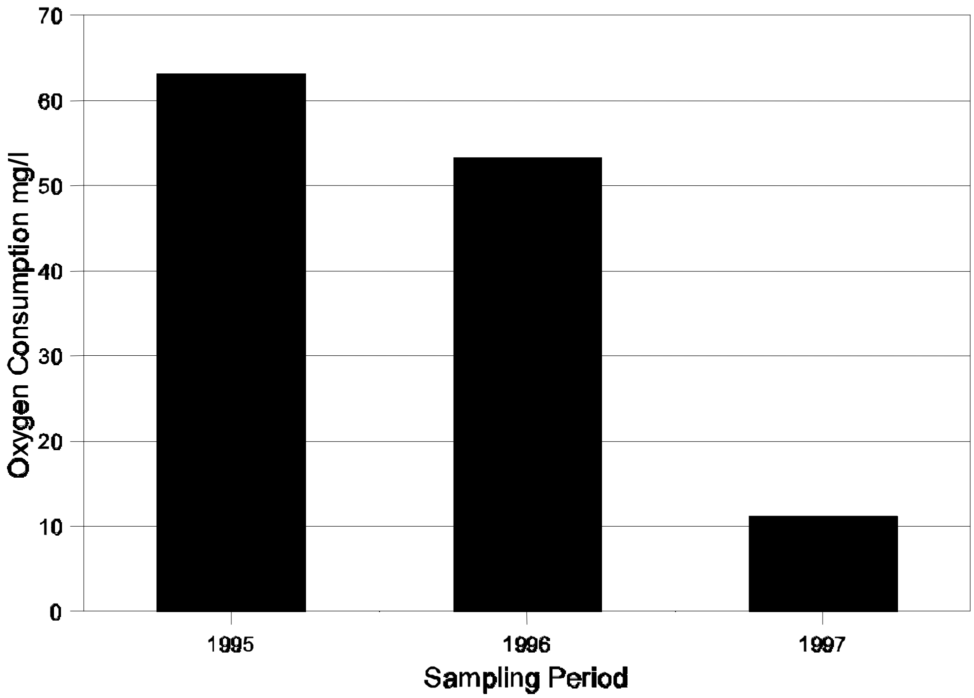
1994 - 1995		**1996 - 1997**	
Number inspected	**Percent failed first inspection**	**Number inspected**	**Percent failed first inspection**
14,992	33.1% (4,962 no.)	17,685	6% (1,061 no.)

As of January 1, 1998, the total number of aeration systems located by Health District staff had been 9,515. The staff continues to find additional aeration systems in remote areas of the County.

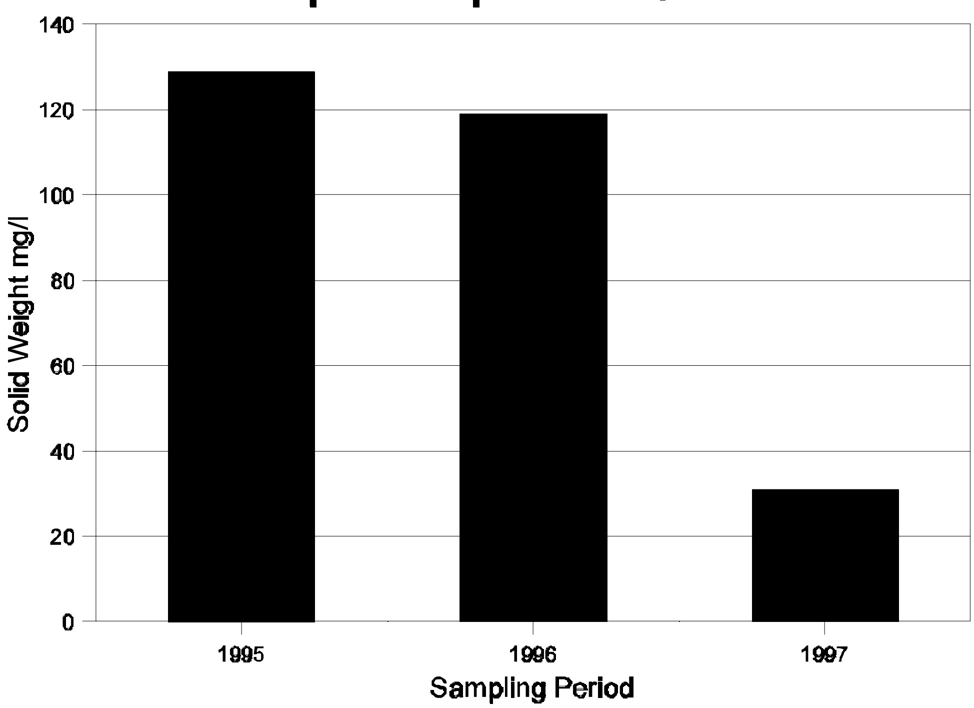
When the data between the two time periods is compared, clearly there is a large reduction in number of systems failing the first inspection. Homeowners are assuring their aeration systems are maintained. The number of homeowners with private maintenance contracts increased from 1,623 in 1995 to 2,274 in 1996. However, in 1997 the number of homeowners with private maintenance contracts decreased slightly to 2220.

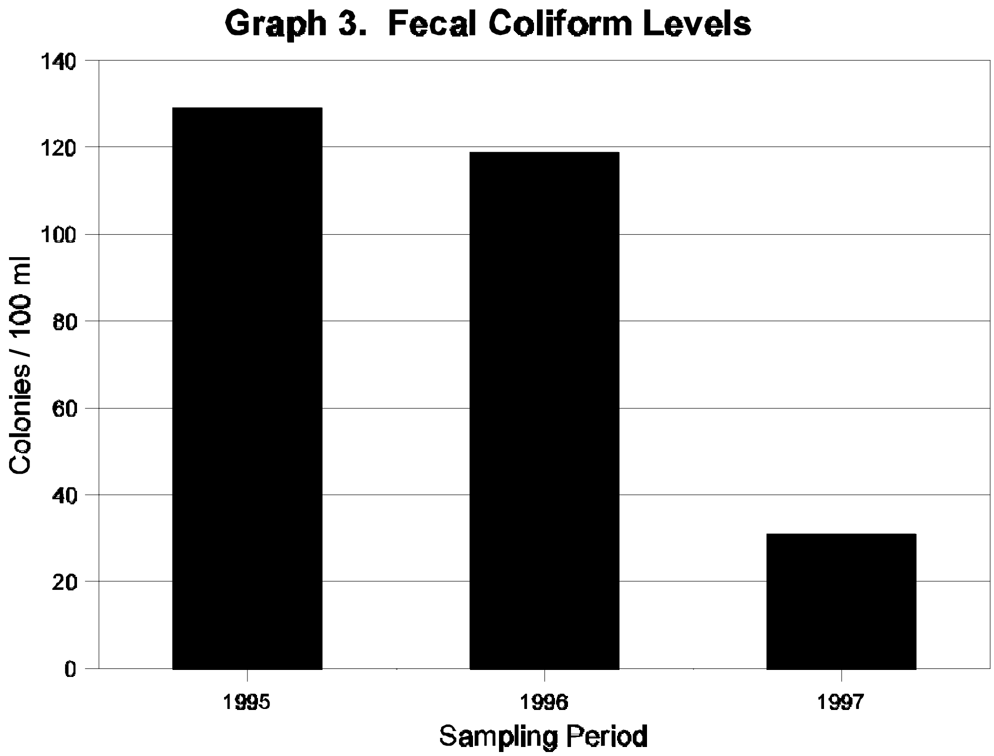
Additional sampling of collector lines to determine program effectiveness was carried out. A 12% randomly selected subset (23 no.) of the original Project CLEAN sampling locations were sampled and analyzed for BOD, SS, and fecal coliform bacteria. The collector line discharges were sampled once each year over the course of a three-year period. (See Graphs, pages 14-15).

Graph 1. Biochemical Oxygen Demand



Graph 2. Suspended Solids





Graphs 1, 2, and 3 reveal the reduction of BOD, suspended solids, and fecal coliform bacteria loading over the three-year period. The average BOD levels for all samples taken in 1997 show a fourfold decrease in BOD when compared to the samples taken the first two years of the program. Similar decreases in suspended solids and fecal coliform bacteria occurred. These graphs demonstrate that the operation permit program for home aeration systems connected to collector lines has reduced the potential for disease conditions and reduced pollution in those neighborhoods.

The stream sampling program was initiated in 1997 as an outgrowth of Project CLEAN. Forty sampling locations were selected. In 1997, quarterly stream samples were taken from these locations in an effort to observe differences in water quality over time. During sample collection, stream velocity, air temperature, and overall weather condition are noted. The goal of the stream monitoring program is to assess the water quality impact within the watershed. Table 6 shows the stream sampling data. But, there are not enough data points to establish any trends other than random variation.

Table 6
Stream Sampling Mean Values

Minimum n = 40

Sampling data	BOD	Fecal	NH3	SS
3/12/97	1.63	435.98	0.07	7.34
6/19/97	6.26	11541.46	0.13	26.29
9/25/97	4.45	1637.5	0.27	6.9
12/17/97	5.22	5007.73	0.6	6.91
4/2/98	5.18	4546.17	0.5	14.59
5/6/98	5.91	31992.93	0.54	113.69

Citizen Opinion Survey

During the spring of 1998, the Health District mailed out 1,000 customer surveys to homeowners with aeration sewage systems. The sample selection was produced using the number on the cash register receipt in the home aeration system database. Questions asked on the survey form included the categories of program effectiveness, program personnel, and program components. The response has been good. As of June 5, 1998, 314 survey questionnaires were completed and returned. Table 7 shows the results.

Table 7
Citizen Survey Results

Program Effectiveness:	Mean of Responses		Standard Deviation
Odors have improved	2.92		1.10
Steam clarity improved	2.96		0.95
Visible sewage decreased	2.95		0.90
Service good value	2.83		1.18
Fees affordable	2.91		1.19
Program Personnel:			
Inspectors identifiable	2.59		0.92
Inspectors courteous	2.17		0.76
Inspectors helpful	2.41		0.85
Office courteous	2.56		0.8l
Office helpful	2.61		0.80
Program Components:			
Information adequate	2.35		
Information helpful	2.35		
Information understandable	2.26		
Process understood		207 yes, 37 no	
Information received		166 yes, 65 no	

Program Effectiveness and Program Personnel: Respondents were given six choices that were rated on a scale of one to six: (1) strongly agree, (2) agree, (3) neutral, (4) disagree, (5) strongly disagree, (6) N/A.

Program Components – information adequate, information helpful, information understandable: Respondents given same six choices listed above.

Program Components – process understood, information received: respondents given two choices, yes / no.

Under "program effectiveness," slightly more people agree than disagree that the operation permit program has reduced odors, improved stream clarity, and provides a valuable service. Written comments from respondents included nine people wanting sewers, seven providing compliments, nine noting fewer odors, six citizens wanted lower fees, five questioned fairness of the program relative to pollution from semi-public sewage plants, three people resented governmental intrusion and eleven citizens considered the operation permit program redundant because of private maintenance contracts with service companies.

Under the categories, "program personnel" and "program components," a larger number of people agreed than disagreed that the staff were helpful and the information received was helpful. Written comments received from citizen survey respondents were similar to the comments mentioned above.

Enteric Diseases and Sewage

With the advent of GIS technology at the Health District, the staff are also monitoring for any association between communicable disease reports and their proximity to home sewage systems. In 1997 an elevated number of cases of Giardia were reported in Hamilton County. A grouping of giardia cases appeared to be located in and around collector line discharges. However, upon further investigation by public health nurses and sanitarians, the exposure source was determined to be outside of Hamilton County. The GIS technology will continue to be used at the Health District as a public health surveillance tool.

What's Next?

The operation permit program has reduced pollution and disease conditions. Odors and stream clarity have improved in many neighborhoods. The Health District plans to continue advancing water quality improvements, improving existing programs and adding new ones. For example, work has begun on a program that will index stream conditions so that children may play and explore their backyard creeks and not worry about the threat of disease.

Some program initiatives will be changed. Instead of a assessing home sewage systems and water quality by political jurisdiction, a new approach currently underway is to use a watershed-based approach for improving water quality. A big help to accomplishing this will be the new automated land and development permitting system. The major building and development permitting agencies of Hamilton County and the City of Cincinnati are installing an automated land and infrastructure management system known as Permits Plus. This new permitting process will coordinate all permitting agency activity by placing all data regarding a parcel of land into a computer file. This will then allow a specific project to be monitored from concept to completion and thereafter for maintenance support. The integrated approach will provide up-to-date land use and infrastructure changes in a watershed and allow for the Health District to assess these impacts on water quality.

Another proposed project includes a review by the Health Commissioner of the contract between the Ohio EPA and HCGHD for the Health District to regularly inspect semi-public sewage disposal systems. It is estimated that between 250 to 300 semi-public sewage disposal systems exist in Hamilton County. Semi-public sewage systems serve small businesses. Neither the Ohio EPA nor the HCGHD are inspecting these systems.

The Health District will use comments received from the citizen surveys to form new local coalitions. These coalitions will assist the Health District in its endeavor to educate home owners about the necessity of maintenance as an important part of owning and operating a sewage system. If these local coalitions could be assigned to a sub-watershed in which they reside, the coalition could become an environmental neighborhood watch group. They could report not only stream impacts from malfunctioning sewage systems but also illegal dumping.

The Health Dstrict should begin educating zoning commissions and the elected officials about the importance of considering home sewage systems as a part of the local utility infrastructure. Citizens in these decision making roles tend to default land use decisions in unsewered areas to local Boards of Health. Future zoning decisions need to take into consideration the watershed impact created by changes in land use and the choice of sewage utility to serve that use.

Last Words, for Now

No one imagined forty-five years ago the impact that the indiscriminate usage of household sewage disposal systems, especially aeration systems, would have on the neighborhoods and backyard streams of Hamilton County. Improper siting and the disregard of maintenance needs has created a costly cleanup for many citizens today. The passage of a new sewage code and the implementation of a multifaceted operation permit program has slowed further water quality degradation and facilitated needed remediation efforts. However, it has taken a lot of effort and community support to reach this point. The Health District will strive to meet citizen expectations.

GIS and GPS in Environmental Remediation Oversight at Federal Facilities in Ohio

James L. Coon, Kelly Kaletsky, William G. Lohner, and Thomas A. Schneider
Ohio EPA, Office of Federal Facilities Oversight
Dayton, Ohio

Abstract

This paper presents some of the experiences and plans of the Ohio EPA Office of Federal Facilities Oversight's (OFFO) use of GIS and GPS for environmental remediation oversight at the U.S. Department of Energy's (DOE) Fernald, Mound, and associated sites. The Fernald site is a former uranium metal production facility within DOE's nuclear weapons complex. Mound's mission included production, development, and research in support of DOE's weapons and energy related programs. Contamination of soil and groundwater at both sites is being remediated under binding agreements with the State of Ohio and USEPA. The primary contaminants of concern at the sites include several radionuclides and chlorinated solvents.

OFFO uses GIS/GPS to enhance environmental monitoring and remediation oversight. These technologies are utilized within OFFO's environmental monitoring program for sample location and parameter selection, data interpretation, and presentation. GPS is used to integrate sample data into OFFO's GIS and for permanently linking precise and accurate geographic data to samples and waste units. It is important to spatially identify contamination because as remediation progresses visual references (e.g., buildings, infrastructure) are being removed at both sites.

Availability of the GIS allows OFFO to perform independent analysis and review of DOE contractor generated data, models, maps, and designs. This ability helps alleviate concerns associated with "black box" models and data interpretation. OFFO's independent analysis has increased regulatory confidence and the efficiency of design reviews. GIS/GPS technology allows OFFO to record and present complex data in a visual format aiding in stakeholder education and awareness.

GIS and GPS are used in planning and evaluating natural resource restoration projects at Fernald. These systems can be used to map wetland mitigation projects and evaluate functional changes including floristic and hydrological parameters. For restoration planning, post

excavation topography models are used to develop appropriate landscape plans. OFFO oversight, including GIS support, is an important element ensuring DOE’s fulfillment of commitments to wetland mitigation and Natural Resource Damage restoration.

This presentation is intended to highlight some of OFFO’s activities covered above and some lessons learned in implementing the GIS/GPS program. OFFO’s five years of GIS/GPS development have resulted in numerous lessons learned and ideas for increasing effectiveness through the use of GIS/GPS.

Background

One of the responsibilities of the Ohio Environmental Protection Agency's Office of Federal Facilities Oversight is to conduct regulatory oversight of the environmental restoration activities at the United States Department of Energy's (DOE) Fernald and Mound sites as well as several Department of Defense (DOD) sites. The Fernald site, formerly known as the Feed Materials Production Center, is a 1050-acre facility located in a rural residential area, 8 miles northwest of Cincinnati, Ohio. Production of uranium metal, including slightly enriched, depleted, and natural uranium began in 1953.Small amounts of thorium metal were also produced. Production stopped in July 1989 to focus resources on environmental restoration. In December 1989, the site was added to the United States Environmental Protection Agency's (USEPA) National Priorities List.

The Fernald site is located over the Great Miami Aquifer (GMA), which is a valued natural resource and has been designated a sole source aquifer by the USEPA. Ground water has been contaminated with uranium, various other radionuclides, inorganic, and organic contaminants. A plume of uranium contaminated ground water extends approximately 1 mile down gradient of the facility. Six waste pits, used during production, contain approximately 430,910 metric tons of waste, including uranium, thorium, and other radioactive and chemical contaminants. Four concrete silos were constructed at Fernald to store radioactive materials. Two of them, the K-65 silos, contain high radium-bearing residues, one contains lower-level dried uranium-bearing residue, and one has not been used. In addition to the ground water contamination, surface water, sediment and soil have been contaminated by past activities at the facility.

The Mound Plant is located in Montgomery County within the city of Miamisburg, Ohio, approximately 10 miles southwest of Dayton, Ohio. The Miamisburg downtown area, five schools, six parks, and many city and township residences are located in the immediate vicinity of the plant. The Mound Plant is also located within 2000 feet of the Great Miami River (GMR), which flows for 170 miles through southwest Ohio to its confluence with the Ohio River. A portion of the Mound Plant overlies the GMR Buried Valley Aquifer (BVA), which is a sole source aquifer providing potable water for a major portion of Miami Valley residents and industries.

The Mound Plant was established in 1946 as a facility to support atomic weapons research. Its primary mission was the process development, production engineering, manufacturing, surveillance, and evaluation of explosive components for the United States nuclear defense stockpile. Secondary missions over the next 40 years included nuclear material safeguards, radioactive waste management and recovery, building and testing of nuclear generators for NASA spacecraft, and purification of non-radioactive isotopes.

OFFO was created in the spring of 1994 by the Ohio Environmental Protection Agency. It was created to coordinate and conduct regulatory oversight of the investigation and remediation activities at federal facility cleanup sites within Ohio. OFFO facilitates environmental monitoring, emergency response, regulatory oversight, and public outreach at DOE and Department of Defense sites in Ohio. Funding for Ohio EPA's oversight of the DOE Fernald and Mound sites is provided in Cost Recovery Grants between the State of Ohio and DOE. Additional information on OFFO is available on the Internet (http://offo2.epa.state.oh.us).

The Fernald group within OFFO has been employing Geographic Information Systems (GIS) and a Global Positioning System (GPS) since early 1995 for regulatory oversight and environmental monitoring at the site. As our experience and confidence with the GIS has increased, our reliance on it for informed decision making has increased. The GIS provides timely and accurate information in various formats for problem solving and decision making. The GPS has been employed to geographically reference new data as it is collected as well as to acquire positional information on existing structures and waste units.

The Mound group, also within OFFO, has also been using GIS/GPS for regulatory oversight and environmental monitoring at the Mound plant. The GIS has allowed for verification of sampling results within Potential Release Sites (PRS) as well as providing independent modeling of potential contaminant plumes within a PRS. The GPS at Mound and related sites has been used for data collection similar to the Fernald site.

GIS/GPS Program

OFFO uses Intergraph's Modular GIS Environment (MGE) and Environmental Resource Management Applications (ERMA) software. The ERMA software suite includes modules for groundwater modeling, data management, geologic and subsurface analysis. MicroStation is used for graphical manipulation of resulting images. The attribute data for this project is stored in the Oracle relational database management software. Oracle provides the database tools necessary to store, organize, and manipulate the large amount of environmental data collected at the Fernald and Mound sites over the past decade. Oracle also works in conjunction with the MGE and ERMA software from Intergraph. The Intergraph system was selected in order to remain consistent with the system used by DOE's operating contractors at both sites.

An Ashtech Reliance Husky GPS field unit, Reliance 12-channel GPS base station and associated Ashtech software are used to complete field mapping projects. Specifically, GPS is used to support OFFO's GIS, monitoring, and oversight programs. Air, water, soil, and biota samples are collected by members of the environmental monitoring team. GPS coordinators accompany monitoring team members into the field and gather GPS data for specific sampling locations. Following data collection, post-processing occurs at Ohio EPA using Ashtech Reliance software. Dependent upon the quality of the data and location environment, the resulting sampling points may achieve an accuracy within 0.1 meters.

After post processing, the geographic information which was recorded in the field is loaded into the GIS where it is used for sample location mapping, natural resource projects, data interpretation and presentations. The information is also used to determine future sampling locations based on existing contaminant data, select analytical parameters including matrix and contaminants for future sampling, and the creation of contours, maps, and models. Geographically referenced data is easily integrated within the existing Oracle database for timely evaluations of temporal data changes. These maps and data analyses are readily converted into presentations and demonstrations for public education concerning ongoing remediation and monitoring. The ability to have a graphical representation of environmental sampling locations has proven beneficial in educating the public regarding OFFO's environmental monitoring program.

Currently, remediation activities are taking place at the Fernald site. During remediation, all buildings will be demolished and the site will become a series of excavations, thus eliminating most visual references. Waste units will be remediated to risk based cleanup levels leaving above-background concentrations of contaminants in place. Having the ability to navigate back to the former waste units will allow OFFO to determine if cleanup levels have been attained and if any remaining contaminants are migrating at a rate which was not predicted by site modeling.

Before OFFO had GIS/GPS capabilities, environmental models and maps presented in DOE submittals were reviewed on paper. Document reviews required significant staff time reviewing binders of data, or an assumption that results and interpretations provided by DOE were correct. The review process has been significantly enhanced with the acquisition and implementation of OFFO's GIS/GPS program. These new tools have provided OFFO the capability to interactively manipulate and review environmental models and maps produced by DOE contractors and to independently verify conclusions presented in the models.

A series of environmental contaminant visual images have been initiated by OFFO utilizing bubble or graduated circle mapping. The use of bubble maps allows for quick and relatively easy visual representation of contamination of a specified area. These projects involved querying and posting data from OFFO’s soil database in order to create bubble maps containing contaminant ranges that break down according to appropriate regulatory levels. The total uranium contaminant concentrations of importance are those greater than or equal to the Final Remediation Level (FRL), which is 80 mg/kg in soil for much of the site, and those exceeding the Waste Acceptance Criteria (WAC) for Fernald’s onsite disposal facility, 1030 mg/kg. Larger symbols on the map indicate higher corresponding total uranium contaminant values for soil samples. OFFO utilized bubble maps at Fernald to determine the extent of total uranium in the soil for a railroad expansion just north of the former production area. This technique was also used in the Operable Unit 4 silos area to help visualize radium contamination in soils that had eroded into Paddys Run and separately to assess the presence of RCRA metals in the Paddys Run corridor. These visualizations were used in evaluating an appropriate response actions and waste disposition options.

Much of the work at the Mound site is proceeding based on administrative release blocks or parcels. Within each release block or parcel there are more potential release sites (PRS) which represent discrete areas of potential contamination. Contamination within each PRS is being evaluated and remediated, as necessary, prior to release of the PRS to the Miamisburg Mound Community Improvement Corporation (MMCIC) for industrial redevelopment. MMCIC is a nonprofit corporation funded in part by DOE to foster the prompt industrial reuse of the Mound facility as DOE’s work at the site concludes.

To evaluate each PRS, OFFO uses the GIS to generate graphics that display all sampling locations in and near each PRS. This enables staff to evaluate not only what contaminants are known to exist within the PRS but also contaminants nearby that could impact the PRS. An example of this is approach is illustrated in work surrounding PRS 420.PRS 420 is a relatively small area where initial document submittals to the agency showed only isolated contamination. The GIS was then used to plot contamination in and around the PRS revealing significant contamination had been present nearby. Upon more detailed review it became apparent that a minor area of contamination in PRS 420 was actually part of a larger area of contamination that was mostly outside of the PRS 420 boundaries. The GIS enabled a rapid areal review surrounding PRS 420 that otherwise could have only been accomplished after days of pouring through stacks of documents and drawings.

Identification of this larger area lead to improvements in the process used by DOE and its contractor to screen data relevant to each PRS being evaluated. Another outcome of this work was the determination that prior cleanup occurred in the area and that no means existed to determine what data represented previously remediated areas. Subsequent work has gone into establishing an appropriate data tracking system that ties data to the location and flags data from remedial projects so that proper historical and current status can more easily be determined. This data tracking improvement will save money and reduce the work put into resampling areas where cleanups have been completed in the past. Now OFFO staff and citizens reviewing PRS submittals have increased confidence that they are looking at the full extent of information available about the PRS.

OFFO has also used Voxel Analyst to create three-dimensional models depicting potential contamination at Mound.PRS266 is a hillside where thorium contaminated soils were dumped many years ago. Initially, as part of the typical regulatory regime, site generated models were reviewed. To increase our understanding of the waste disposal areas nature and extent, OFFO independently generated a three-dimensional model based upon sampling results obtained in and around the hillside. This modeling resulted in similar results to the site generated models and improved regulatory confidence in project decisions.

Our most recent project involving GIS and GPS is the monitoring of construction and success of six acres of mitigation wetland at the Fernald site. Ohio EPA is using GPS to survey monitoring points for ongoing restoration research. The research project evaluates the impacts of using donor site soils in the constructed wetland basins. In addition to surveying monitoring points the GPS has been used in a roving mode to delineate the boundaries of the constructed wetland. As data are collected from the treatment and control plots they are entered into the GIS which allows for geographic referencing and visual analysis of the data. Data such as percent canopy cover, soil moisture, microbial activity, soil source and treatment type are all collected in the GIS. These data can then be used to evaluate possible correlations and differences in the treatment regime and success monitoring. GPS and GIS can also be used to evaluate overall success of the restoration project including plant survival rates, percent cover in seeded areas, water levels, etc. The wetland mitigation project is an early phase of what will be a much larger natural resource restoration effort at Fernald. As a part of a proposed settlement of a Natural Resource Damages claim, DOE will be completing natural resource restoration on more than 800 acres of the Fernald site. OFFO continues to evaluate methods for using GIS and GPS in our oversight and monitoring of DOE’s natural resource restoration activities.

Remediation of the Fernald site will require large excavations of varying depths across large portions of the site. Remediation of soil contaminated with uranium will require excavations in excess of 25 feet deep in some areas. OFFO has used estimated excavation depths and areas generated by Fluor Daniel Fernald along with existing topographic maps to develop visualizations of the post remediation topography of the site. Developing this post excavation visualizations have been useful in evaluating future land uses for the site. When viewing the extreme variations in post excavation topography and considering the volume of soils that would be required to reestablish the original site topography, decisions regarding final land use are much easier. OFFO has used these visualizations of current and future topography in meetings with DOE and their contractor as well as with the public. Using the information from these visualizations and input from stakeholders DOE has developed a conceptual site model that involves the creation of large open water bodies, interconnecting wetlands, expanded riparian zones and areas of upland forest.

Conclusion

The Ohio EPA has successfully developed and implemented a GIS/GPS program which is being integrated into the environmental remediation oversight process. GIS tools have allowed the Ohio EPA to perform advanced data analysis and gain spatial insight on the problems encountered at the Fernald site. We are improving the data analysis process by creating maps and models with environmental data that has been generated at the Fernald and Mound sites. The GIS allows us to view the massive amounts of data that have been collected at the Fernald site in a spatial context. The ability to view, manipulate and analyze these data sets allows us to draw conclusions about spatial correlations and distribution across the three-dimensional landscape. This information, which is generated using GIS and environmental modeling tools, is then incorporated into the decision making process in the remediation of the Fernald and Mound sites. It is the Ohio EPA’s belief that the system has helped the agency provide better oversight for the State of Ohio and assisted DOE in cleaning up the sites in a efficient and cost-effective manner.

Environmental Justice in Kentucky: Examining the Relationships Between Low-Income and Minority Communities and the Location of Landfills, and TSD Facilities

Larisa J. Keith
Northern Kentucky Area Planning Commission

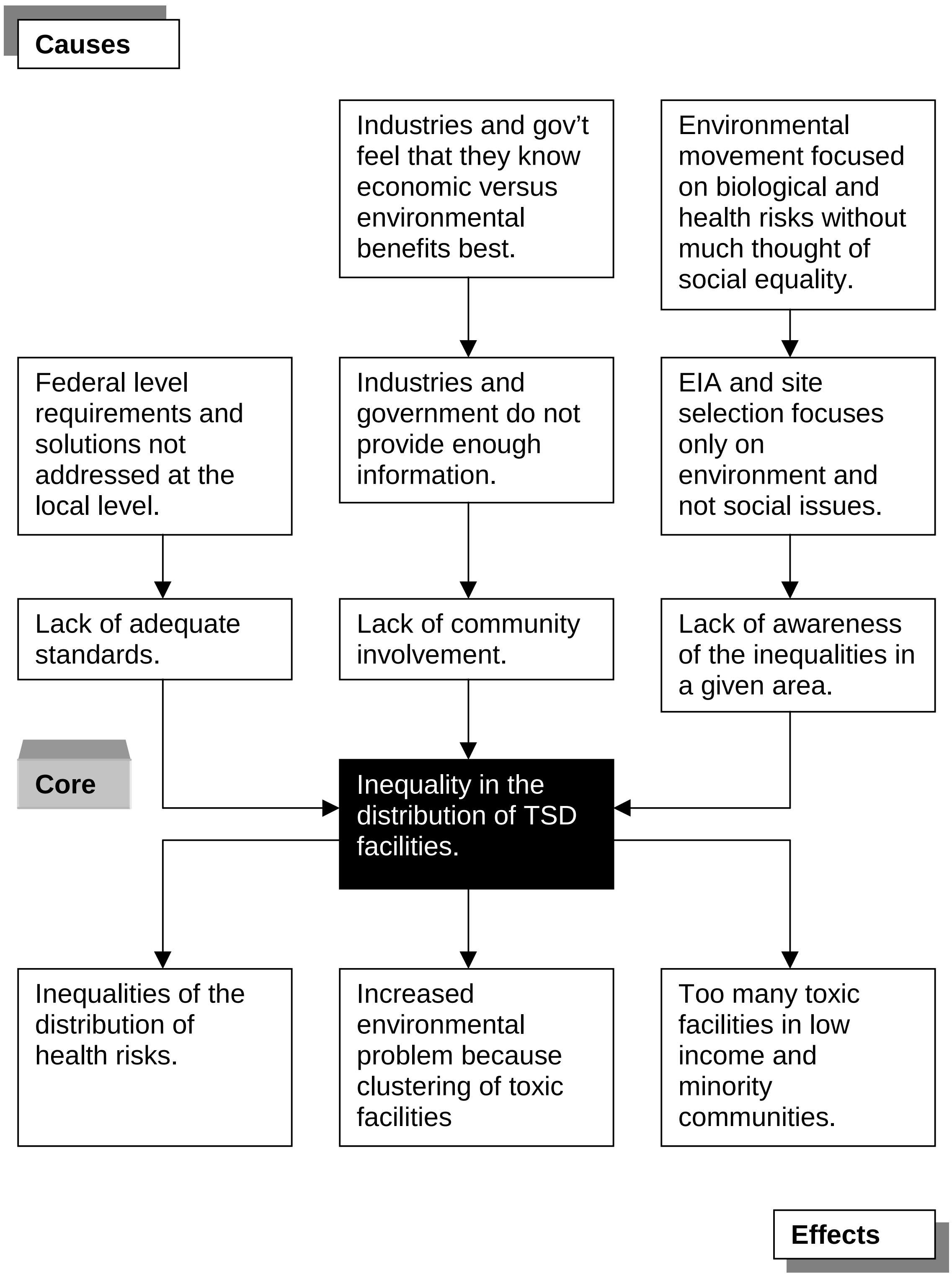
1.1 Problem

Environmental justice has been a topic of discussion among environmentalists since around the early part of the century, although it was not termed such at that time. This topic reaches from lead based paint to nuclear waste dumps. These toxins and all of the hazards in between that may cause human health problems have been the cause of controversy based on their location and their remediation. Studies have been conducted to try to determine the severity of inequality in the siting of such facilities in low income or minority communities. This project attempts to identify relationships between selected socio-economic variables and treatment, storage and disposal facilities, within the Commonwealth of Kentucky in the United States, and to define some of the problems associated with the possible inequality of siting such facilities. These cause and effect relationships are outlined in problems in Figure 1.

Kentucky has several landfills, hazardous waste management facilities, and incinerators located throughout the state. In fact, a report on environmental justice cites Kentucky, along with Tennessee and Mississippi as states where minorities face some of the most disproportionate risks when its comes to hazardous waste siting. It was found that the percentage of minority residents who live near waste sites in these two states was more than twice that in other statewide populations (Charlier, 1994). A community activist in Harlan County, Kentucky, stated that the reason for this might be that poor communities often don’t have the financial or technical resources to deal with the problem (Melnykovych, 1993). Often the complaints of low income or minorities are not heard. However, in one case in Kentucky, a $20,000 grant was awarded to Louisville’s Justice Resource Center, to aid in empowering citizens to fight against the locating of polluters in the West End (First Comes Awareness, 1996). In response to this another $312,000 federal grant was awarded to the city to reduce emissions from the Rubbertown chemical plant complex and other industry (Melnykovych, 1996). These examples illustrate that environmental justice is perceived as a problem in cities and counties in Kentucky.

Figure 1

Problem Tree



1.2 Objective

The purpose of this study is to examine the relationships between the location of treatment, storage and disposal facilities and low income or minority communities in Kentucky. This project will locate the facilities and attempt to determine if there are any significant relationships between their concentration and the racial make up, economic status and type of householder. Also, this project seeks to determine if distance from these facilities has any affect on the characteristics of the surrounding population. For example, are more minority communities close to the facilities than there are at a further distance? The project does not attempt to determine any historical relationship, such as whether or not the facility located itself in the communities before or after the demographic characteristics were in place. This study attempts to find any basic statistical relationships that may be present. This study will begin to develop some initial evidence of any cause and effect relationship, but will not address these cause and effects in detail. In the future, this study may be used as the basis of further studies of historic themes and why the hypothesized relationships may exist. Also, if these relationships are found to suggest causality then policies may need to be implemented to address inequalities in TSD facility siting. The ultimate objective is to learn whether or not these relationships exist so that further studies may be suggested, successful siting methods highlighted, and possible policy solutions created.

1.3 Methods

This study comes in two parts, the first of which is devoted to the relationships between the number of treatment, storage, and disposal facilities (TSDFs) and population characteristics. This section seeks to answer if there is a relationship between all of the variables combined, including the percentage African American, Hispanic, and other minority groups, the percentage of persons below the state poverty level, the percentage of female-headed households, and the per capita income, and the number of treatment, storage and disposal facilities at the county and zip code level. Also, are the relationships between each variable and the number of treatment, storage and disposal facilities positive or negative? The use of these questions will help to determine if there is a direct spatial relationship between these variables and the TSDFs.

The second section of the project will be dedicated to the effect of distance from treatment, disposal, and waste facilities on population characteristics. Do the percentages of minorities increase or decrease with distance from a facility? Do percentages below the state poverty level increase or decrease with distance from a facility? Do the percentages of female-headed households increase or decrease with distance from a facility? Do the per capita incomes increase or decrease with distance from a facility? This section will only utilize zip code data. Again, these questions will aid in determining if there are any spatial relationships between variables and waste management sites, and whether distance has a relationship with the distribution of socio-economic characteristics of the communities surrounding facilities.

1.4 Hypotheses

The following expected correlations are under the assumption that the facilities were in place after the population characteristics were created. The hypotheses for the first section of the project, assuming that there are relationships between the dependent socioeconomic variables and the independent facility variable (number of sites) at both the county and zip code level, are as follows:

1) The greater the number of facilities, the greater the percentage of African Americans (positive relationship);
2) The greater the number of facilities, the greater the percentage of Hispanics (positive relationship);
3) The greater the number of facilities, the greater the percentage of other minority groups (positive relationship);
4) The greater the number of facilities, the greater the percentage of all persons below the state poverty level (positive relationship);
5) The greater the number of facilities, the greater the percentage of female headed households (positive relationship); and
6) The greater the number of facilities, the lower the per capita income (negative relationship).

The hypotheses for the second section of the project, assuming that there are relationships between the dependent population variables and the independent variable (distance from the sites) at the zip code level, are as follows:

1) As the distance from each facility gets greater, the percentage of the population that is African American gets lower;
2) As the distance from each facility gets greater, the percentage of the population that is Hispanic gets lower;
3) As the distance from each facility gets greater, the percentage of the population labeled as other minority groups gets lower;

4) As the distance from each facility gets greater, the percentage of all persons that is below the state poverty level gets lower;
5) As the distance from each facility gets greater, the percentage of the female headed households; and
6) As the distance from each facility gets greater, the per capita income gets higher.

Literature indicates that these hypotheses could be accepted. These studies are discussed further in the literature review.

This project does not attempt to make a causal connection between socio-economic variables and TSDF sites. The objective is to find any relationship that may indicate that further studies are needed to connect causes to effects. This is not a historical review determining who came first, the minority and poverty communities or the facilities; it is a review of the conditions at one moment in time.

The significance of this study is that it will increase knowledge on the level of environmental injustices in Kentucky, related to the location of TSD facilities. This will aid in creating a basis for both residents of the state and decision-makers within industry and government to address the issue of environmental justice with a background of the current situation through policy making. Also this study will create awareness of some possible solutions to any recognized problems with the methods of siting in regard to social characteristics.

This research problem can be addressed by beginning with a background analysis of the subject of environmental justice, its causes and effects, and some of the actions taken to avoid it. The following chapter applies to these topics and others to give a framework of the situation of environmental justice at the current time and how it evolved to be such.

2.1 Background of Environmental Justice

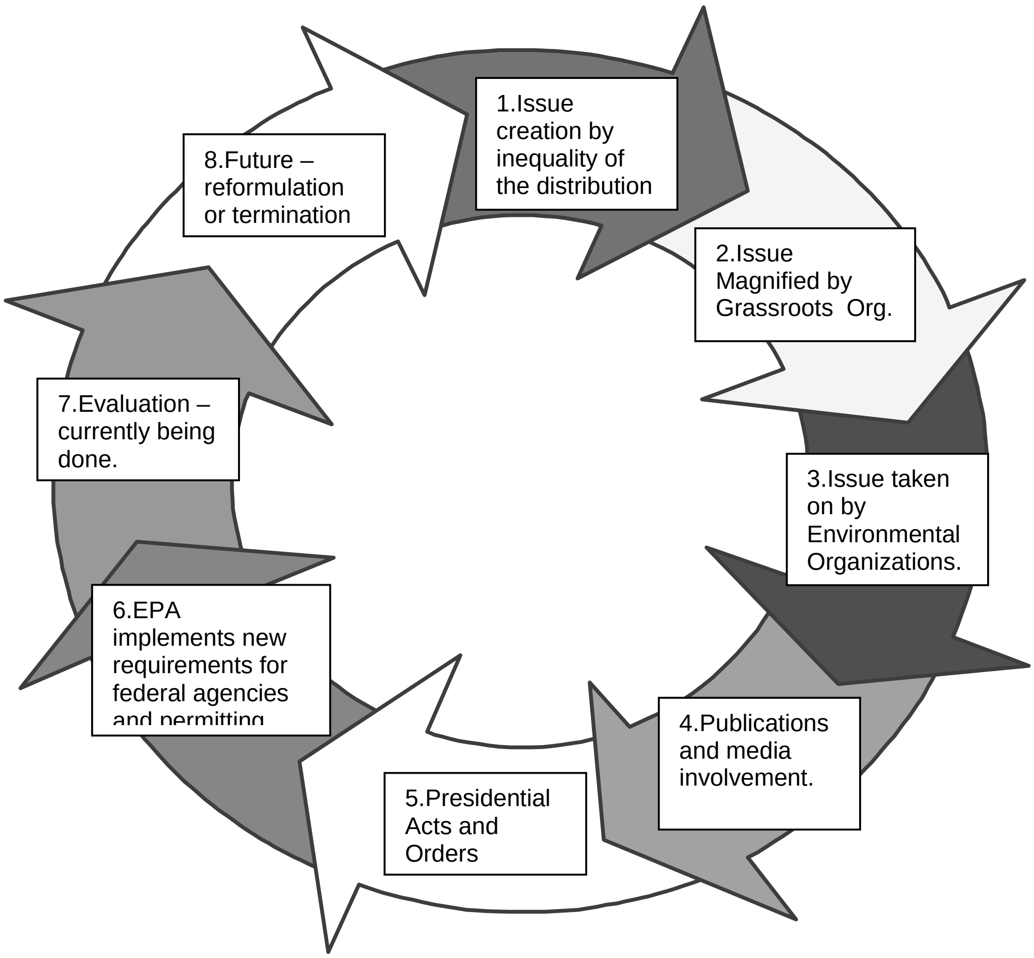
To better understand this project it is necessary to provide a background and the definitions of used terms. Understanding the background of environmental justice and some of the key proponents will aid in clearly defining the scope of this project. As with any issue, environmental justice has taken several levels within its lifecycle (Figure 2). The arrival of the issue, the popularization of the issue in government, and the implementation of the solutions to the issue given by government have all been steps leading up to the current situation of environmental justice. This life cycle of the environmental justice movement is discussed first in this chapter with the issue of toxic waste and health risks, which are related to the cause for the initial uprising of the issue. Next, the grassroots organizations that first brought the issue to light within their own communities and throughout the nation are discussed. The issue being placed on the national agenda for political approval is also discussed through a review of the acts and orders created by the Office of the President. Implementation is also discussed here through EPA regulations toward siting and treatment, storage and disposal facility requirements, which is more specifically related to this study. The categories outlined in the following sections give a background to the lifecycle of environmental justice to date. The EPA is currently doing evaluation of the effectiveness of these policies and reformulation may or may not occur after this is released.

2.1.1 Environmental Justice and the Environmental Movement

In the United States there have been essentially three major themes in the environmental movement according to Ortolano (1996). These themes include the anthropocentric point of view, which is the dominant theme, the environmental justice issue, and the biocentrism point of view. The biocentric view is simply the realization that all species have an intrinsic value. This view was popular in the early 1900’s and reemerged in the 1970’s with environmental ethics.

The anthropocentric theme includes several motives for protecting the environment. The first of which is public health. Early in the nineteenth century environmental NGOs were stressing the importance of public health related to the environment. This has also been the case inside the environmental justice movement itself. These health risks are later discussed in this chapter. Other motives included: 1) technological and scientific management of resources and the acceptance of waste as an unavoidable use of the environment; 2) the effects of human interference on biological systems, especially after WWII and the problem with DDT in the 1960’s; 3) preservation of wilderness in the late nineteenth century; and 4) the resurgence of transcendentalism in the 1970‘s (Ortolano, 1996). However, each of these problems was usually raised by those individuals with sufficient income to allow them to delve into such issues.

Figure 2
Life Cycle of Environmental Justice



The environmental justice theme was first raised in raised in the 1980's when it was seen that environmental programs often imposed unequal costs on people and groups that they effect, which raised further questions about the fairness of distribution (Weale, 1992). The environmental justice events that are further discussed in this chapter fall into the time period extending from the 1960's and 1970's when health issues were highly recognized throughout the anthropocentric environmental realm, through the uprising of grassroots organizations and environmental groups. Hazardous waste became popular with the public in the 1970's, just as other issues dealing with the environment. The upwelling of interest in these issues coincided with the passage of several acts in legislation. The environmental justice issue was developed at the same time as pollution regulation and policymaking was expanded to toxic chemicals. The Toxic Substances Control Act of 1976 was the first legislation to raise the issue of toxic materials. After this came the Resource Conservation and Recovery Act of 1976, which was amended in 1984 to give specific regulations to TSDs. In 1980, The Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) was passed, which helped to deal with toxins after they had been deposited in the environment.

The recognition of the dangers of toxic waste seems to have occurred in both the political and public communities. However, in the public's view, these legislations and the programs were not being applied in an evenly distributed manner. This led to the conflicts in several areas across the nation, where communities got involved in the siting of such facilities. The increase in knowledge about the dangers of toxic materials and the recognition of this by legislation was enough to spur residents to take an active part in an issue that could have negative impacts on their lives. Following these reactions from the public, further legislation (discussed further in this chapter) was enacted to try and address these problems of inequality. This did not occur until the mid 1990's and the effects of these programs and legislations are still being evaluated in 1999.

2.1.2 Toxic Waste and Health Risks

There are many health risks associated with toxic materials. Although it is often difficult to account for the precise amount of risk involved, it is important to realize the possible consequences to human life. In the United States the definition of risk includes several aspects. First, it includes the probability of the occurrence of an event, such as a discharge or spill of toxic materials. Second, it includes the probability that toxins will be released by an event. Third, it addresses the probable quantity, concentrations, transport and fate of toxins released into the environment, as determined by conditions at the time of the event. Risk also means the probability of exposure of individuals, populations, or ecosystems to toxic substances released into the environment. Finally, risk involves the probability of adverse human health or environmental effects from exposure to toxins released into the environment (Cohrssen and Covello, 1989).

It is understandable that no numerical level of risk will be acceptable to everyone, and also that eliminating all risk to all people is impossible. Therefore it is important to balance the potential risks to individuals from toxins against the broader benefits to the community. This is an element of the concept of environmental justice. No one should be unnecessarily subjected to potential risks more than others should.

The risks of exposure to toxins can occur through the manufacturing process, transportation to market, use in consumer goods, and disposal as waste (Heath, 1989). Each of these can result in toxic releases into the groundwater, atmosphere, the food chain, surface water and soil, which can lead to increased, and often unrealized, human exposure. These exposures can result in several health problems. Acute biological abnormalities may include skin rashes, acute neurological effects, acute hepatic or renal disease, and acute gastrointestinal or cardiovascular illness. Other sub-clinical acute effects include abnormalities in liver enzyme function, nerve condition sloping and chromosomal damage (Heath, 1989). These effects, of course, depend on the type and amount of the chemical. Creating these causal relationships is often difficult and depends on the accuracy of information regarding exposure and abnormalities. However, any possibility of negative health impacts that may or may not create a decrease in the quality of life of human beings is a significant issue that needs to be addressed with the utmost caution.

2.1.3 Grass Roots Organizations

Historically, the location of locally unwanted land uses has followed the path of least resistance; those associated with health and environmental risks, such as toxic waste sites, included. However, throughout the years, although low income and minorities have been subjected to a disproportionate amount of pollution in their neighborhoods and workplaces, they have only been slightly involved in the environmental movement. Low income and minority communities have had few advocates at the national level and within the environmental movement (Bullard, 1990). It was not consistently large toxic corporations that created opposition at the start of the environmental justice movement but also the smaller, home based dangers. Lead-based paint illnesses were, in the early part of the century, one of the first signs that minorities were getting more than their fair share of health risks due to toxins. Lead was a middle class issue in the broad population in the inner city from the 1940's through the 1970's. The first major effort to deal with it was by a black community in Chicago in 1965. While lead was a big issue to society for a while, failure came within policy to address relationships between housing community health and the inner city environmental where many people of color lived. This reinforced the notion that environmental issues were white and middle class. Lead was a continuing problem in the 1980's. In substandard housing, particularly, contamination was high (Gottlieb, 1993).

Another distinctive difference in the distribution of health risks was in the realm of low income and low skilled workers. Low-income industrial workers, who, in the early 1900's, were the majority black, were also subject to the dangers of non-controlled environmental risks associated with industry. This was a time when recruiting on the basis of race was not an unusual practice for particularly hazardous jobs. There were asbestos cover-ups from the 1930's through the 1950's, and coal mining cover-ups in the 1960's. In 1967 the Association of Disabled Miners and Widows was formed as the first black lung protection group. In West Virginia the Black Lung Association in 1969 spurred legislation such as the Coal Mining Health and Safety and the Occupation Safety and Health Act (Gottlieb, 1993).

These sorts of communities in which environmental injustices have occurred sometimes did so because the cities had to make a choice between environmental conservation and economic development. Cities in need of employment opportunities, although they may have tried to attract clean industry, often had to settle for dirty industry. The view of the communities was often not to oppose anything that would bring in jobs (Bullard, 1990).

Also in the 1960's and 1970's, came the emergence of concern for pesticides due to the publication of *Silent Spring,* not only in conjunction with wildlife hazards, but also farm worker health and safety. This was also spurred by minority groups, specifically Chicano and Mexican born workers. However, as of 1990, farm workers were still excluded from any pesticide policy debates.

During the period of the Cold War, uranium activities in the southwest were booming, mostly in Native American areas. Soon after, concerns emerged about health and environmental impacts produced an "environmental legacy" of radiation that was linked to illnesses and deaths. Native Americans were now unrecognized victims as well (Gottlieb, 1993).

The environmental movement itself though the sixties and seventies, which was concerned with booming population growth, was said to have racial implications. This was related to the usurpation of resources and increased pollution (Gottlieb, 1993). In response, since the 1970s, grass roots organizations have been involved by challenging local governments on the control of such things as factories and landfills (Brown, 1994). The Natural Resource Defense Council, the Lawyer's Committee for Civil Rights Under Law, the Sierra Club and the Committee for Racial Justice of the United Church of Christ all became involved by endorsing legislation and by cooperating to gain Senate approval of the environmental justice legislation (Coyle and Carmody, 1993). In 1979, a suburban African American neighborhood in Houston filed the first lawsuit to charge environmental discrimination. This lawsuit charged Browning Ferris Industries with locating a municipal solid waste landfill in their community. All five of the city's other sanitary landfills and six of the eight municipal solid waste incinerators were in mostly African American neighborhoods. However only 28% of the city's population was African American itself (Bullard, 1994).

In the 1970's an organization called the Urban Environmental Conference (UEC) began as a legislative and lobbying environmental counterpart to the Leadership Conference of Civil Rights. In 1983 this organization funded a conference on toxics and minorities in New Orleans which was the first major attempt to identify toxic issues as issues of discrimination and social justice. Unfortunately, the UEC lost its funding and failed to adequately create a connection between civil rights and the environment (Gottlieb, 1993).

In the 1980's, new groups were formed raising issues of risk discrimination in environmental terms now known as environmental justice. "Primary actors in these new campaigns were alternative environmental groups for which the ethnicity factor remained prominent" (Gottlieb, 1993). Environmental justice gained national attention in 1982 when a protest broke out in a predominantly African American community in North Carolina against the siting of a burial site for soil contaminated with PCBs (Knollenberg, 1998). This site was a minority community, and was found to be scientifically unsuitable, but was still developed (Bullard, 1990). Prompted by these demonstrations, in 1983 the General Accounting Offices (GAO) studied hazardous waste landfill siting in the District of Columbia region and found a relationship with black percentages (Bullard, 1990). Due to the upwelling of public protest, the interest of many other organizations consequently turned to the topic of environmental justice. For example, the Commission for Racial Justice of the United Church of Christ did one of the most well known studies on the topic in 1989. This study concluded that race was a likely determinant of the location of commercial waste facilities and uncontrolled toxic waste sites (Brown, 1994). Around the same time Bullard released his study entitled *Dumping in Dixie* (1990). Both proved to be crucial documents in situating environmental racism.

After these studies on environmental justice, public attention continued to grow. Since the Church of Christ study, several other grassroots organizations have been involved in the fight for environmental justice. For example, the Los Angeles City Energy Recovery Project, The South West Organization Committee in Albuquerque who held the "International Hearing of Toxics in Minority Communities" in 1989, the Louisiana Toxics Project, and the United Farm Workers in California. These organizations, along with the study by the Church of Christ, helped in recognizing the "disparate distribution of environmental contaminants in their communities." The Church of Christ study also pushed the topic to academia. It suggested that universities give assistance to minorities seeking training in the environmental fields and that the universities should develop curricula in environmental sociology (Lee, 1992). In response, the EPA cosponsored the conference on the Environment, Minorities and Women, designed internship program for educators at black colleges and universities, and created an Indian Task Force to be sure they were aware of, and took advantage of, EPA programs (Lee, 1992).

During the 1990's new groups and coalitions were being formed to combat injustice in the environmental movement. Louisiana based non-Anglo activists and the New Mexico based South West Organizing Project began stressing the problem with the CEO's of several major companies. This involvement created significant concern among mainstream groups such as The Sierra Club, The Audubon Society, and Friends of the Earth (Gottlieb, 1993). In 1991 the "People of Color Environmental Summit" was held in California. This summit resulted in a public policy agenda to address environmental injustices. It recommended that Congress and the President address it through legislation, executive orders, and other governmental policy (Bass 1998).

2.1.4 Acts/Government Orders/ EPA Regulations

When any issue gets large enough to draw attention from the public then it is probable that the government will get involved. This has been the case with environmental justice. In response to the fore mentioned public outcry, the government became more concerned. The Environmental Justice Act of 1993 was a bill directing the Environmental Protection Agency (EPA) and other federal agencies to cooperate in the clean up of the 100 most toxic counties in the nation. This bill would trigger additional governmental assistance to these counties in most need (Coyle and Carmody, 1993).

Finally, in February of 1994, President William Clinton gave Executive Order 12898 that would lend *Federal Actions to Address Environmental Justice in Minority Populations and Low-income Populations* (Bass, 1998). This policy established several things. It gave federal agencies the mission to achieve environmental justice by identifying and addressing any disproportionately high and adverse human health or environmental effects of its programs, policies and activities on minority populations and low income populations of the United States and its territories and possessions (U.S. Congress, 1998). The Executive Order created the Interagency Working Group of Environmental Justice that included the heads of several executive agencies and offices such as the Department of Defense (U.S. Congress, 1998). The Executive Order required that each federal agency develop an agency wide environmental justice strategy listing all programs, policies, planning and public participation practices, enforcement, and rule makings related to human health and the environment (Bass, 1998). The Working Group had to submit to the President a report describing the implementation of the order and the strategies (U.S. Congress, 1998). The Executive Order created the development of agencies and the Working Group to implement the policy for environmental justice. However, the Executive Order did not create new opportunities for legal challenges against federal agencies for violations of the order (Bass, 1998).

The next step of the President was to put out a presidential memorandum. This directed federal agencies to do five things: 1) to analyze the environmental effects of federal actions including effects on minority and low income communities when required by NEPA; 2) to address adverse effects of any mitigation measures outlined or analyzed in an Environmental Assessment, and Environmental Impact Statement or a record of decision on minority and low income committees; 3) to provide opportunities for community input in the NEPA process, to include and identify potential effects and mitigation measures in consultation with the affected communities, improving the accessibility of meeting and providing access to crucial documents and notices; 4) that EPA, when reviewing NEPA documents, had to ensure that the involved agency had fully analyzed the environmental effects on minority and low income communities; and 5) to ensure that the public, including minorities and low income communities, had access to public information relating to human health or environmental planning, regulations, and enforcement when required under the Freedom of Information Act, The Sunshine Act, or the Emergency Planning and Community Right to Know Act (Bass, 1998).

More recently, the EPA proposed a new policy that will guide the investigation of any complaint that state or local permitting decisions has violated federal civil rights laws (Johnson, 1998). It is an interim guidance policy for environmental justice cases (Information Access Company, 1998). It will provide a framework for EPA’s Office of Civil Rights to handle complaints under

Title VI of the Civil Rights Act of 1964 that prohibits discrimination (Johnson 1998). If the EPA finds discrimination and the permitting agency can not eliminate it, then EPA will deny, annul, suspend or terminate its funding. It may also refer the matter to the Department of Justice for litigation (Williams et al, 1998). This process is applicable to all that receive federal funds (Williams et al, 1998). Some say that the Title VI guidance from the EPA fails to define many important terms, including how they will determine whether a particular community has suffered "disproportionately high adverse human health or environmental effects" (Johnson, 1998). The EPA's Guidance has initiated a lot of controversy because of its vagueness and is being evaluated and may be changed based on suggestions that have been made. A report that will evaluate the interim guidance policy is scheduled to come out in 1999 (Information Access Company, 1998).

While those given legislations may seem vague when considering where to locate a toxic waste facility, there are more rigid standards that focus on treatment, storage and disposal facilities specifically. These standards are based on The Resource Conservation and Recovery Act (RCRA) of 1976. For example, if the generator is not notified by TSDFs that waste has been received within a certain period of time specified by the EPA, the generator files a report to indicate that a departure from the required procedure has occurred. In addition, if the waste received by a TSD facility is different than the waste detailed on the manifest given by the generator, then a report must be again filed describing the discrepancy (Ortolano, 1997).

All TSD facilities that receive solid hazardous waste must also obtain a permit. There are several requirements that must be satisfied to acquire this. These requirements include records and reports that notify authorities of releases, provisions that will ensure that unauthorized persons do not enter the facility, technical specifications such as liners in landfills that prevent hazardous material from entering groundwater, monitoring to detect releases of hazardous materials, corrective action for any release that may have human health and environmental risks, detailed plans for closing the facility and maintenance of monitors, and financial responsibility in the form of bonds or insurance that can sufficiently pay for any damages incurred. The EPA also issues requirements on treatment of hazardous materials so that they must use the best available technology to clean out contaminants (Ortolano, 1997).

More relevant to this study, there are specific standards for siting of facilities. Most of which involve topography and natural environment. For example, location standards specific to

Kentucky include consideration of seismic conditions and flood plains. However, there is no mention of any standards based on population characteristics (KAR, 1997).

It has been stated that typically the process of site selection is based upon technical engineering, environmental and economic suitability, and public and political acceptability. Site suitability is dependent on the characteristics of the facility, characteristics of the site, and economic and legal feasibility. It involves regulatory approval by a federal agency or a federally approved state agency, and site approval by state or local government, as well as the previously mentioned permits under RCRA. These local government regulatory influences include land use controls, nuisance and construction controls, and taxes and fees implied on facilities. Site acceptability to the host community is also a concern. This concern stems from the possibility of negative consequences, due to the siting, and not usually the probability of such. Therefore, in order to obtain acceptability, information on impacts must be made readily available and open communication and negotiation must be continued throughout site selection. Environmental impacts are often assessed through environmental impact statements and socioeconomic impacts are addressed through proof of economic need for the facility, compensation, preventive measures and mitigation (Andrews and Lynn, 1989).

The regulations and legislation presented throughout this section have begun to create some support for the topic environmental justice at the national and federal level. Local powers may have a hand in siting decisions, however no equality issues seem to be addressed at that level either. It has yet to be seen how the federal regulations will specifically address the issue of siting without inclusion of explicit regulations within state or local statutes.

2.1.5 Specific Literature Related to the Study

The books and articles presented in this section provide a background to this project that has leant several methodologies, limitations, and information on the subject of environmental justice. There have been a number of studies done that support the objectives of this project. This type of study has been done on several areas, including Texas, Florida, Michigan, and Los Angeles County, as well as on a national scale (Yandle and Burton, 1996, Pollock and Vittas, 1995, Hockman and Morris, 1998, Boer, Pastor, Sadd and Snyder, 1997, and Oakes, Anderton, and Anderson, 1996). These studies were key in determining the methodology and hypotheses for this project. Several of the cited studies utilize similar variables, and methodologies, and have similar sets of conclusion.

There seems to be some disagreement on how to deal with measuring the impacts of waste facilities on surrounding communities. Should a study use census tracts, census block groups, or zip codes? Most of the studies reviewed for this project utilize census tracts for the study area. However, in Michigan, zip codes were used (Hockman and Morris, 1998) and in Florida, census block groups were used (Pollock and Vittas, 1995). Although research indicates that census tracts will give the best indicators of the concentration of waste sites in low income or minority communities, this study will utilize zip codes due to the limited data availability on such a small scale. While the zip codes may not be as appropriate to the study as census tracts, as discussed in section 3.1 and 5.2, there are acceptable because they are relatively small areas, in this study, ranging from 0.45 to 454.65 square miles. The population ranges within these zip codes are from 27 to 46,612 persons. This yields a density range of 1.4 to 7985.8 persons per square mile within the zip codes. These numbers give a relatively good variance that will provide for a adequate analysis.

Variables that are best for use as indicators of disproportionate siting are also not identical in each of these studies. What are common are the percentage of the population that is minority, and some form of income indicator. The type of waste facility used varies among each as well. In both the studies on Los Angeles County and at the national level, treatment, storage and disposal facilities were used (Boer, Pastor, Sadd, and Snyder, 1997, and Oakes, Anderton, and Anderson, 1996) and in Texas hazardous waste landfills were used.

Several different methods of statistical analysis have been used in studying this topic. One of the most widely used statistical methods that were used in the previous studies is regression. Others that were cited include t-test and Wilcoxon nonparametric Z-test (Oakes, Anderton, and Anderson, 1996, and Boer, Pastor, Sadd and Snyder, 1997) as well as chi-square and Cramer's V (Yandle and Burton, 1996).

The national tract level study indicates that commercial TSDFs are on an average sited in communities that are neither disproportionately poor nor minority. The study also notes that this finding does not rule out specific local bias in TSDF siting or the impacts of siting on one specific community (Oakes, Anderton, and Anderson, 1996). At the conclusion of the analyses in the Florida, Texas, Michigan, and Los Angeles County some sort of relationship was found with at lease one of the variable used in each study. More specifically, the study by Pollock and Vittas (1995) found that African Americans and Hispanics reside closer to potentially hazardous sources in Florida. The study by Yandle and Burton (1996) found that there was a statistically significant relationship between relative poverty and hazardous waste landfill siting in metropolitan Texas. It was found that both race and income combined were strong predictors of hazardous waste facilities siting in Michigan, and that separately, race was a more potent predictor (Hockman and Morris, 1998). In Los Angeles County, Boer, Pastor, Sadd, and Snyder (1997) found that most communities affected by TSDFs are working class communities of color located near industrial areas. *The Commercial Appeal* studied census tracts in which landfills, sewage-treatment plants, hazardous-waste processors, medical waste incinerators, leading toxic polluters and abandoned hazardous waste sites were located in Shelby County, Kentucky. This study found that of the 53 facilities that were studied, 26 were in census tracts where Blacks were the majority, and 25 were in tracts where whites were the majority (1994). These smaller scale examples give the foundation for the need for further projects focused on more specific areas such as counties and census tracts in the Commonwealth of Kentucky.

The studies just cited are the most helpful in determining what things to look for in the first section of this project. Methodologies, sources of information, breadth of study area, and time periods, as well as the pros and cons of each are seen. This is useful information in determining the course of action used for this project.

The second section of this project, which concentrates on how socio-economic variables are influenced by distance from the facilities, is modeled after a study done by Wang and Auffrey (1998) at the University of Cincinnati, in Ohio. The variables used in this study included mortality rates and related socio-economic factors and information from the Master Sites List and the Toxic Release Inventory from the Ohio Environmental Protection Agency. The study area was the City of Cincinnati and the units of study used were census block groups. For this study of Kentucky the variables and study area will be different but the methods of analysis will be basically the same. These methods will be discussed further in the methodology section of this proposal.

Other authors have made comments on the Yandle and Burton article that are also of use to this project. One such article is by Robert D. Bullard (1996). Bullard offers a list of assumptions that have been unjustly made by Yandle and Burton in their article. Assumptions given in this article provide for some points of thought that may be applied to this project. These assumptions include the following: 1) environmental justice is limited to waste facility siting, 2) leaving out some facilities and communities will not affect study results, 3) all census tracts are created equal with respect to population density and proximity of the population to facilities, 4) census tracts represent homogeneous neighborhoods, 5) all hazardous waste landfills are created equal and, 6) environmental justice and environmental racism are the same.

This study realized that environmental justice includes several factors, not only the location of TSDFs. It is not an assumption of this project that generalities may be made to all facilities based on the findings. Leaving out facilities and communities may limit the study, however, the information available often dictates such things. Although census tracts are not used in this project, the same limitation may be stated for zip codes. These limitations are recognized and are discussed further in the analysis. Environmental justice has been previously defined in this paper, so not to confuse the reader with different terminology. This project does not attempt to create new or to redefine definitions of environmental racism or environmental justice.

2.1.6 Definitions

Environmental justice is defined by the EPA as “equitable treatment of all people, regardless of race, income, culture or social class with respect to the development, implementation and enforcement of environmental laws, regulations, and policies” (Heaton, 1999). Equitable treatment, in this case, means that no group should bare a disproportionate share of negative environmental impacts because of government actions (Bass 1998).

Minority and low income communities in this project are defined by following census variables: 1) the percentage of Blacks, Hispanics, and other nonwhites; 2) the percentage of all persons that are below the state poverty level; 3) the percentage of female-headed households; and 4) per capita income. The definitions of each of these separate variables came from the US Bureau of the Census (1990). Black is defined as persons who indicated their race as “Black or Negro” or reported entries such as African American, Afro-American, Black Puerto Rican, Jamaican, Nigerian, West Indian or Haitian. Hispanic is defined by the 1990 Census by Hispanic Origin. This group includes those who classified themselves as Mexican, Puerto Rican, Cuban, or Other Spanish/Hispanic origin such as those from Spain, the Spanish-speaking countries of Central or South America, or the Dominican Republic or those who generalized themselves in categories such as Spanish, Spanish-American, Hispanic, Hispano, and Latino. Origin is defined by the census as the ancestry, nationality group, lineage, or country of birth of the person or the person’s parents or ancestors before their arrival in the United States. Persons of

Hispanic origin may be of any race. In this study "other" is a combination of the remaining nonwhite races. Within this group are included American Indian, Eskimo, Aleut, Asian and Pacific Islander. American Indian is defined as persons who indicated their race as "American Indian," entered the name of an Indian tribe, or reported such entries as Canadian Indian, French American Indian, or Spanish American Indian. Eskimo is also in this category, this label includes persons who indicated their race as "Eskimo," or reported entries such as Arctic Slope, Inupiat, and Yupik. Aleut is included in this category and is composed of persons who indicated their race as "Aleut," or reported entries such as Alutiiq, Egegik, and Pribilovian. Asian and Pacific Islander categories are comprised of two different groups. Asian is composed of persons who indicated their race as Chinese or who identified themselves as Cantonese, Tibetan, or Chinese American, Taiwanese or Formosan, Filipino, Philipino, Philipine, or Philipino American, Japanese, Nipponese or Japanese American, Asian Indian, Bengalese, Bharat, Dravidian, East Indian, or Goanese, Korean and Korean American, Vietnamese and Vietnamese American, Cambodian, Hmong, Laohmong, or Mong, Laotian, Laos or Lao, Thai, Thailand or Siamese, Bangladeshi, Burmese, Indonesian, Pakistani, Sri Lankan, Ameriasian, or Eurasian. Pacific Islander within this category includes persons who indicated their race as Hawaiian, part Hawaiian or Native Hawaiian, Samoan, American Samoan, or Western Samoan, Guamanian, Chamorro or Guam, and others such as Tahitian, Northern Mariana Islander, Palauan, Fijan, Polynesian, Micronesian or Melanesian (U.S. Bureau of the Census, 1990).

Female-headed households are defined as families with a female householder and no spouse of householder present (U.S. Bureau of the Census, 1990).

Persons categorized as being below the poverty level are all determined by the Social Security Administration definition modified by the Federal interagency committees and prescribed by the Office of Management and Budget. The core of the definition is the 1961 economy food plan, the least costly of four nutritionally adequate food plans designed by the Department of Agriculture. This is based on the determination that families of three or more persons spend approximately one-third of their income on food, hence, the poverty level for these families was set at three times the cost of the economy food plan. For smaller families and persons living alone, the cost of the economy food plan was multiplied by factors that were slightly higher to compensate for the relatively larger fixed expenses for these smaller households. Thresholds were set by the Census and if the total income of each family or unrelated individual in the sample was less than the corresponding threshold, it was classified as “below the poverty level”. The poverty thresholds were applied on a national basis (U.S. Bureau of the Census, 1990).

Per capita income is defined by the 1990 Census as the mean income, which is calculated by dividing the total income of a particular statistical universe, in this case counties and zip codes, by the number of units in that universe. It needs to be understood that this number is usually recorded by memory and are only estimates and often therefore underreported or overestimated. Procedures were also used by the Census to estimate appropriate values for unreported numbers (U.S. Bureau of the Census, 1990).

The hazardous facility that is to be used in this study is labeled as a treatment, storage, and disposal facility (TSDF). TSD facilities include treatment units such as incinerators, dewatering facilities, and waste solidification facilities, landfills, surface impoundments, and underground injection wells. Treatment, storage and disposal facilities are the last stop for hazardous waste in the system of tracking hazardous waste under The Resource Conservation and Recovery Act (RCRA) of 1976 (Ortolano, 1997). Title 401 of the Kentucky Administrative Regulations (KAR), which lays out the standards for TSDF’s, gives specific definitions for each type of facility. A treatment facility is termed as using any method, technique or process including neutralization, designed to change the physical chemical or biological character or composition of any hazardous waste so as to neutralize such waste, to recover energy or material resources from the waste, or to render such waste either nonhazardous or less hazardous. This type of facility makes the waste safer to transport, store, or dispose of; or amenable for recovery, amenable for storage, or reduced in volume. A storage facility is simply a place in which hazardous waste is held for a temporary period, at the end of which the hazardous waste is treated, disposed of, or stored elsewhere. A facility at which hazardous waste is intentionally placed into or on any land or water, and at which waste will remain after closure is termed a disposal facility. These are the definitions accepted by the Natural Resources and Environmental Protection Cabinet and the Department for Environmental Protection Division of Waste Management, and so they are accepted for this study. The specific TSD facilities that are used in this project were obtained from the Kentucky Department of Waste Management’s Notifier’s list. This list is compiled of all hazardous waste generators (large, small, and conditionally exempt generators), treatment, storage, and disposal facilities, recyclers, burner/blenders and transporters of hazardous waste in the Commonwealth of Kentucky. All of the above facilities (except conditionally exempt generators) are required by law to notify the State of their hazardous waste generation or transportation. Conditionally exempt generators (CEG'S) generate less than 100 pounds of hazardous waste a month and therefore are not required to register, but Kentucky encourages notification of the CEG's.

This background provides the needed information to understand the problem of environmental justice and to see the need for studies such as this one. These definitions also lead to the next chapter, which applies particular methodologies to the variables mentioned in this chapter, and which will allow for the previous hypotheses to be tested.

3.1 Methodology and Data

The purpose of this project is to analyze the relationships between the following variables and the number of treatment, storage and disposal facilities in the State of Kentucky. Data on the percentage of African Americans, Hispanics and other nonwhite groups, the percentage of all persons that are below the state poverty level, the percentage of female headed households, and the per capita income were obtained from the U.S. Bureau of the Census. The 1990 data will be used at both the county and the zip code level. There are 120 counties in the State of Kentucky within its 39,732.3 square miles (Figure 3). This gives rise to a rational comparison with zip codes. This rationality will be important in evaluating the change of effect on areas of different sizes. The particular population variables that have been chosen typically reflect the characteristics of the population as being minority or low income. These variables have often been sited as key in determining if disproportionately high and adverse health or environmental effects are imposed on minority and low-income populations.

The location of treatment, storage and disposal facilities was obtained from the Kentucky Division of Waste Management. The data retrieved on these facilities include the name of the facilities, both the facility and the mailing addresses, the type of facility, and a contact person for each facility. This material was manipulated to determine the number of sites in each county and in each zip code. Some facilities were not used in the zip code analyses that were used in the county analysis because of lack of data from the census. Also, for GIS purposes, the address information created a means to locate the facilities on a map for graphic representation of the relationships to the population variables and for distance analysis.

The waste management facility sites and the population, poverty and income variables are used to determine any relationships that may imply disproportionate siting of facilities. These variables are used in statistical analyses to show the strength of the relationship between all of the variables combined and the number of TSDFs. The method of analysis to be used for the first part of the project is regression and bivariate correlation coefficient.

For the second part of the project, concentrating on distance, as before mentioned, the methods of analysis somewhat follows those in the study by Wang and Auffrey (1998). Here, GIS was used for geocoding and buffering analysis. Statistical analysis, including the independent t-test for equality of means and bivariate correlation coefficients, are calculated to determine the relationships and their direction.

Close attention is paid to any limitations created by the statistical analysis used, the study area used, the types of data collected, and other specifics of the project, and what influences these decisions have on the outcomes.

The data and statistical methodology, along with GIS, give the information on the relationship of population characteristics and the siting of treatment, storage and disposal facilities in Kentucky. This will begin to determine if there are disproportionately distributed facilities in Kentucky that may in turn create a higher amount of adverse health or environmental effects on low income or minority populations.

3.1.1 Comparison Analysis

For the first section of this analysis, a simple creation and comparison of maps is done. This entails making use of Geographic Information Systems, specifically in the form of ArcView. The information that was given on counties and zip codes in the form of shape files from ArcView is utilized for creation of the images. Attribute files listing the characteristics of both counties and zip codes were then attached to these images. The information within these attribute files was, again, obtained originally from the U.S. Bureau of the Census (1990), and was manipulated to fit the form and content required in ArcView. The addresses for the treatment, storage and disposal facilities that were obtained from the Kentucky Division of Waste Treatment were also inputted into ArcView for the geocoding and comparison.

3.1.1.1 Geocoding

The addresses obtained from the Division of Waste Treatment were geocoded in ArcView by zip code. This function entailed comparison between the zip code given for each facility and the data within the zip code file in ArcView. This function gives each match a score from zero to one hundred. Each facility used for the creation of this map received a score of 100, although not all facilities listed were matched. This non-matching occurred because of the missing data with the ArcView zip code file. This resulted in locating the precise location of most of the facilities on the zip code level. The locations were then translated into an attribute form that was more easily comparable with the other variables. The maps that were created from this information show the number of facilities in each county (Figure 4) and in each zip code (Figure 5). The number of facilities for the county map increased from the number in the zip code map because each facility was matched to a county that was present in the ArcView shape file of counties and not all zip codes containing facilities were present in that file.

3.1.1.2 Boundary and Attribute Files

The data obtained from the U.S. Bureau of the Census (1990) was also translated into map form that made comparison possible. County municipal lines and zip code lines within the Commonwealth of Kentucky were used as the boundaries for this project. As with the facility information, the variables used were in an attribute file that was joined to the county and zip code shapes. These files, along with the additional facility information, were used to create graduated color schemes for each variable. The patterns of the color schemes for each variable are compared to determine similarities.

3.1.1.3 Multiple Linear Regression and Bivariate Correlation

A regression analysis was done at both the county and zip code level. In order to determine the relationship between multiple variables and the number of facilities, the socioeconomic variables were plotted as the independent variables and the number of facilities as the dependent variable. This analysis gives the nature of the relationship between a group of variables. It is a mechanism of forecasting (Levin and Fox, 1988). The statistic given is $r^{2,}$ which is the coefficient of determination. It yields the proportion of variance in the dependent variables that is explained by the independent variable. It gives this proportion as a percentage, which is r^2 multiplied by one hundred. The significance level used for this study was 95 percent.

Bivariate correlation was also conducted with Microsoft Excel to test the significance of the relationship between the facilities and the socioeconomic variables. This statistic gives the degree and type relationship between two variables. This analysis results in a number from positive one to negative one, positive one being a perfect positive relationship and negative one being a perfect negative relationship. This number is defined as “the ratio of the covariance between X and Y to the product of their respective standard deviations” (Levin and Fox, 1988). Numbers other than one will show different strengths in the relationships. Positive or negative 0.6 is a strong relationship, positive or negative 0.3 is a moderate relationship, positive or negative 0.1 is a weak relationship and 0.0 is no relationship (Levin and Fox, 1988). For this analysis, weak relationships will range specifically from 0.1 to 0.30, moderate relationships will range from 0.31 to 0.6 and strong relationships from 0.6 and up. This statistic will create the basis of acceptance or rejection of the hypothesis.

3.1.2 Distance Analysis

In order to account for the variable of space it was necessary in this project to deal with distance from each of the facilities and the spatial effect on the populations. Since zip codes are the smallest unit used, these are the focus of the distance analysis. The use of distance is an important factor of understanding the possible relationships between the location of facilities and socioeconomic characteristics. If some pattern is found to exist for one or more of the variables with distance from the facilities, then it may suggest a spatial connection.

3.1.2.1 Buffering Analysis of Zip Code Data

The buffering was again made possible by ArcView GIS. As the use of zip code data made it difficult to assign precise rings around each of the facilities, a buffer of zip codes immediately adjacent to zip codes with facilities was used. The act of buffering entailed selecting from the zip code theme all of the zip codes that contained one or more facilities. A separate shape file was created from these. Next, the zip codes designated to be immediately adjacent to these zip codes with facilities were selected from the original zip code theme. This was done by the select by theme function. Then, zip codes that were found to be adjacent to the first ring were selected in the same manner. Deletion of the overlapping zip codes from the previous selection was done. The same was done again for adjacent zip codes two more times.

3.1.2.2 Reclassification

The buffering method used to reclassify zip codes into five categories was basic adjacency to zip codes with facilities present. The zip codes with facilities were grouped into the first category and labeled with the number five to denote the closest characteristics to the facilities. The first ring around the zip codes with facilities was created from all of the zip codes that were immediately adjacent to those zip codes. These were labeled with the number to four to denote that they were at a different distance from the facilities. This continued on to create three more categories labeled from three to one; with one representing the zip codes that were furthest away from those zip codes with facilities.

3.1.2.3 Bivariate Correlation and T-test

As done in the previous analyses, the bivariate correlation is used in this section of the study. The analysis will be a bit different, however, in that it utilizes different data indicating zip codes with and without facilities. By this it will address the distance factor by assigning numbers to the zip codes based on their distances from those zip codes with facilities. This created a sort of ranking of the zip codes.

Also in this section of the analysis, a t-test for the difference in means is conducted. This analysis utilizes the adjustment for unequal variances since one sample variance is more than twice the other is. This statistic gives a way to determine the differences between two means of the same variable, those with facilities, and those without. This is done by calculating the mean for each group, finding the standard error of the difference between means and the determining the critical value for t. Then the critical value from the t table is compared to the actual t value. If the actual t value is greater than or less than the positive and negative t table critical value then one can say there is a difference in the means (Levin and Fox, 1988). This test will find any of the following relationships: 1) $mean_{with} \geq mean_{without}$; 2) $mean_{with} \leq mean_{without}$; 3) $mean_{with} = mean_{without.}$ This will be done using the zip code data only. The test done will be two tailed and the level of significance is 0.05. This analysis will be done using Microsoft Excel.

These methodologies and data lead to the following chapter, in which the results will be shown of the tests and mapping laid out in this section. Some things mentioned in this chapter will be further explained and analyzed in the following chapters.

4.1 Comparison Analysis

This analysis begins by comparing maps created in ArcView for each socioeconomic variable with the maps of the number of facilities. Each variable was spit into four categories. This was done by the quantile method. This method involves splitting the number of data points by four so that each category will contain approximately the same number of entries. This was used for both county and zip code maps Using this type of methodology makes it easy to compare each map. The map analysis done here is a basic description of the patterns of each characteristic within the counties and the zip codes. These patterns may or may not show evident likenesses between each variable and the number of TSDs. Patterns are compared and more detailed information is given on significant numbers for each variable. This analysis is followed by the statistical results from the regression analyses and the bivariate correlation analyses for both county and zip code information.

4.1.1 County Map Analysis

When comparing the maps showing the number of treatment, storage and disposal facilities (Figure 4) and the percentage of the population that is black (Figure 6) in each county it is noticeable that both are somewhat concentrated more in the western and central section of the state than in eastern Kentucky. Of the counties that contain three to twenty-five facilities the percentage of the population that is black ranges from 0.0 to 17.0. Counties with the three highest percentages of Blacks include Jefferson (17.0), Christian (18.7), and Fulton (24.5), each of which contains facilities. Christian and Fulton County contain one facility, and Jefferson has twenty-five. Those counties containing facilities and that have the three lowest percentages of Blacks include Marshall (0.0), Martin (0.1), and Grayson (0.4).

Counties that are in the highest category for both variables include Jefferson, Fayette, Shelby and Hardin. Those in the second highest category of percent Blacks and the second highest category of facilities include Anderson, Mercer, and Muhlenburg. Harrison is also included in this category with a percentage between 2.3 and 5.5 and has three facilities, which places it in the top category for number of facilities. Those counties in the third highest category of Blacks that contain facilities include Boone, Campbell, Carroll, and Hancock each with one facility, Bullitt, Ohio, and Pulaski with two facilities, and Boyd, with ten facilities. Of the forty counties with facilities, three are in the lowest category of black percentages.

In comparing the number of facilities and the percentage of the population that has Hispanic origin, it can still be seen, though not to such an extent that both are more concentrated in the western and central areas of Kentucky (Figure 7). Of the counties that contain three to twenty-five facilities the percentage of the population that has Hispanic origin ranges from 0.1 to 2.7. Counties with the highest percentages of Hispanic include Meade (2.0), Hardin (2.7), and Christian (3.4). These counties all include one TSD facility, Hardin with 3. Henry County has the lowest percentage of Hispanics, at zero. There are several other counties containing facilities that have 0.1 and 0.2 percent Hispanic. Counties that are in the highest category for both variables include Jefferson, Boyd, Fayette, and Marshall. Those in the second highest category of percent Hispanic and the second highest category of facilities include Mercer, Hopkins, and Pulaski. Those counties in the third highest category of Hispanic that are also in the third category of facilities include Campbell, Bell and Calloway, each with one facility. Of the forty counties with facilities, eight are in the lowest category of Hispanic percentages.

In comparing the number of facilities and the percentage of the population that has Hispanic origin, it can still be seen, though not to such an extent that both are more concentrated in the western and central areas of Kentucky (Figure 7). Of the counties that contain three to twenty-five facilities the percentage of the population that has Hispanic origin ranges from 0.1 to 2.7. Counties with the highest percentages of Hispanic include Meade (2.0), Hardin (2.7), and Christian (3.4). These counties all include one TSD facility, Hardin with 3. Henry County has the lowest percentage of Hispanics, at zero. There are several other counties containing facilities that have 0.1 and 0.2 percent Hispanic. Counties that are in the highest category for both variables include Jefferson, Boyd, Fayette, and Marshall. Those in the second highest category of percent Hispanic and the second highest category of facilities include Mercer, Hopkins, and Pulaski. Those counties in the third highest category of Hispanic that are also in the third category of facilities include Campbell, Bell and Calloway, each with one facility. Of the forty counties with facilities, eight are in the lowest category of Hispanic percentages.

The map of the percentage of the population that is other nonwhite is more evenly dispersed than the two previous variables, however the highest percentages do appear to still be slightly concentrated in western and central areas (Figure 8). There are counties throughout each region that contain some small percentages of other races. Of the counties that contain three to twenty-five facilities the percentage of the population that is other nonwhite ranges from 0.1 to 3.9 Counties with the three highest percentages of other nonwhite include Hardin (3.9), Christian (3.7), and Meade (2.4). Both Christian and Meade contain one facility, Hardin with three. The lowest percentage of the population of the counties that is other nonwhite is zero.

Counties that are in the highest category for both variables include Jefferson, Fayette, Boyd, and Hardin. Those in the second highest category of percent other nonwhite and the second highest category of facilities include Anderson, Bullitt, Mercer, Boyle, and Pulaski. Those counties in the third highest category of other nonwhite that contain facilities include Bath, Hancock, Daviess, Bell, and Todd, each with one facility, Muhlenburg and McCracken with two facilities, Harrison with three and Marshall with five. Of the forty counties with facilities, four are in the lowest category of other nonwhite percentages.

The map showing the percentage of female-headed households has a random pattern dispersed throughout the state (Figure 9). Of the counties that contain three to twenty-five facilities the percentage of the population that is black ranges from 11.6 to 20.8 Counties with the three highest percentages of female-headed households include Fulton (22.9), Fayette (20.8), and Calloway (20.1) Fulton and Calloway each have one facility, and Fayette has six. Those counties containing facilities and that have the three lowest percentages of female-headed households include Bullitt (7.4), Oldham (8.2) and Meade (8.4).

Counties that are in the highest category for both variables include Jefferson, Fayette, and Boyd. Those in the second highest category of percent female-headed households and the second highest category of facilities include Mercer, Hopkins, and Muhlenburg. Also included in the second highest category of female-headed households is Harrison with three facilities and Marshall with five. Those counties in the third highest category of female-headed households and facilities include Grayson, and Christian, each with one facility. Of the forty counties with facilities, six are in the lowest category of female-headed household percentages.

The map showing the percentage of the population below poverty level has a pattern increasing from northwest to southeast (Figure 10). The highest percentages are found almost solely in the eastern part of the state. Of the counties that contain three to twenty-five facilities the percentage of the population that is below poverty ranges from 12.3 to 16.6 percent. Counties with the three highest percentages of Blacks include Owsley (51.0), McCreary (45.3), and Wolfe (43.6). None of these counties contain facilities. Those counties with the three lowest percentages of persons below poverty include Oldham (5.9), Boone (7.3), and Woodford (7.7). Oldham and Boone both contain one facility.

There are no counties that are in the highest category for both variables. Counties that are in the highest category for the percentage of persons below poverty level include Martin, Bell and Cumberland, each of which have one facility. Those in the second highest category of percent below poverty include Carroll, Bath, Union, Grayson and Fulton, each with one facility, and Ohio, Muhlenburg, and Pulaski, each with two. There is a wide range of the number of facilities in the counties included in the third highest category of persons below poverty. Those counties that are also in the third category in the number of facilities include Mason, Henry, Hancock, Ballard, Warren and Todd. Of the forty counties with facilities, eighteen are in the lowest category of below poverty percentages.

The map of the per capita incomes for each county shows the highest amounts in the northern Kentucky area and the Louisville to Lexington corridor (Figure 11). The lowest seems to coincide, of course, to the higher percentages of those below poverty level in the southeastern part of the state. Since the relationship hypothesis is negative, this part of the map analysis will discuss it in this manner as well. Of the counties that contain three to twenty-five facilities the per capita income ranges from ranges from $10,271 to $14,962. Counties with the three highest per capita incomes include Woodford ($14,151), Fayette ($14,692) and Oldham (15,510). Those counties with the three lowest per capita incomes are McCreary ($5,153), Owsley ($5,791), and Wolfe ($5,998).

Counties with the lowest per capita incomes with facilities present include Bell ($6,858) and Cumberland ($7,037). Both of these counties have only one TSD facility. There are no counties that are in the lowest category for both variables. Those in the second lowest categories for per capita income with facilities present include Bath, Meade, Martin, Ohio, Grayson, Pulaski and Todd. Each of these counties contains either one or two facilities. The counties contained in the third lowest category of per capita incomes that contain facilities include Carroll, Mason, Henry, Madison, Ballard, Christian, Calloway and Fulton, with one facility, Muhlenburg with two, and Harrison, with three. Of the forty counties with facilities, twenty-one are in the highest category of per capita incomes.

It is interesting to note that those counties listed as the highest percentages with facilities present were often sited more than once. For example, Christian County was in the top three percentages with facilities for three variables. Hardin, Meade, Fulton, Bell and Cumberland were listed in the top for two. Also noticeable for some variables is that the highest overall percentages were the same as the highest percentages for counties with facilities. For example, female-headed households, Blacks, other nonwhites and Hispanics all listed the same respective counties for both the top three overall percentages and the top three for those with facilities present. Finally, it is interesting to note that Jefferson County, which had the highest number of facilities (25), was listed in the highest percentage categories for all of the variables except the percentage of persons below poverty level. These kinds of occurrences could point toward more specific siting issues with respect to a combination of socioeconomic variables.

4.1.2 Zip Code Map Analysis

Given that zip codes are simply smaller sections of the counties, when combined they should create approximately the same numbers given for each variable in the county data. Due to this many of the same patterns are repeated in the zip code maps concerning the location of concentrations of each category.

It can be seen by looking at the map of the percentage of the population that is Black that the concentrations are again in the western and central section of the state (Figure 12). Of the highest category, which ranges from 5.1 to 16.0, there are only three zip codes that are also in the highest category for the number of facilities. These zip codes are 40213, 40214, and 40216, which are all located in Louisville and all contain four facilities. There are twenty-eight other zip codes included in this category of percent Blacks that contain facilities, as well. In the second highest category of percentage Black, there are three zip codes that are also in the second highest category of number of facilities. The zip codes are located in Louisville (40272), Lawrenceburg (40342), and Harrodsburg (40330). Also in this category are two zip codes with three facilities, these include Cynthiana and Ashland. There are seven other zip codes that contain facilities included in this category. There are twelve zip codes with facilities in the third category of percentages and five in the lowest category.

The zip codes that contain the most facilities range in percentages of Blacks from 0.0 to 13.3. The highest three percentages of those zip codes with facilities are 40211with 96.0 percent, 40203 with 54.3 percent, and 40212 with 49.6 percent, all of which are within the city of Louisville. The lowest percentage found in zip codes with facilities is 0.0. The highest percentages of Blacks found in all of the zip codes include 40211 with 96 percent.

Concentrations of the highest percentages of Hispanic are also in the western and central part of the state (Figure 13). Of the highest category, which ranges from 0.6 to 21.6, there are only two zip codes that are also in the highest category for the number of facilities. These zip codes are 40214 (Louisville) with four facilities, and 41101 (Lexington) with three facilities. There are fifteen other zip codes included in this category of percent Hispanic that contain facilities, as well. In the second highest category of percentage Hispanic, there are two zip codes that are also in the second highest category of number of facilities. The zip codes are located in Louisville (40213), and Harrodsburg (40330). Also in this category are two zip codes with more than two facilities, these include Louisville (40212) with four and Calvert City (42029) with five. There are ten other zip codes that contain facilities included in this category. There are twenty-two zip codes in the third category of Hispanics that contain facilities. Seven zip codes with facilities are within the lowest percentages.

The zip codes that contain the most facilities range in percentages of Hispanics from 0.0 to 2.3. The highest three percentages of those zip codes with facilities are 40121 (Fort Knox) with 7.8 percent, 40511 (Lexington) with 4.4 percent, and 41101 (Ashland) with 2.3 percent. The lowest percentage found in zip codes with facilities is 0.0. The highest percentages of Hispanics found in all of the zip codes include 42523 (Belcher) with 21.8 percent.

Once more the highest concentration of the other nonwhite percentages are located in the western and central part of the state, although this variable is a bit more evenly dispersed than the previous two (Figure 14). Of the highest category, which ranges from 0.7 to 25.9, there are only three zip codes that are also in the highest category for the number of facilities. These zip codes are 41101 in Ashland containing three facilities, and 40214 and 40216 in Louisville each containing four facilities. There are twenty-seven other zip codes included in this category of percent other nonwhites that contain facilities, as well. In the second highest category of percentage other nonwhite, there are two zip codes that are also in the second highest category of number of facilities. These zip codes are located in Lawrenceburg (40342), and Harrodsburg (40330). Also in this category are two zip codes with more than two facilities, these include Louisville (40213) with four facilities, and Calvert City with five. There are twelve other zip codes that contain facilities included in this category. There are fourteen zip codes in the third highest category of percent other nonwhite and no zip codes with facilities in the lowest category.

The zip codes that contain the most facilities range in percentages of other nonwhites from 0.1 to 2.0 percent. The highest three percentages of those zip codes with facilities are 40121 (Fort Knox) with 7.4 percent, 40509 (Lexington) with 3.3 percent, and 40510 (Lexington) and 40208 (Louisville) both with 2.5 percent. The lowest percentage found in zip codes with facilities is 0.1. The highest percentages of other nonwhites found in all of the zip codes include 40061 (Saint Catharine) with 25.9 percent.

The map of the percentage of female-headed households is evenly dispersed throughout the state (Figure 15). Of the highest category, which ranges from 16.0 to 78.0. There are only three zip codes that are also in the highest category for the number of facilities. These zip codes are 40213, 40214 (both in Louisville) with four facilities, and 41031 (Cynthiana) with three. There are twenty-eight other zip codes included in this category of percent female-headed households that contain facilities, as well. In the second highest category of percentage female headed households, there is only one zip code that is also in the second highest category of number of facilities. This zip code is located in Lawrenceburg (40342). Also in this category are three zip codes with more than two facilities, these include 40216 (Louisville) with four facilities, 42029 (Calvert City) with five, and 41129 (Catlettsburg) with six. There are thirteen other zip codes that contain facilities included in this category. There are eight zip codes in the third highest category of percent female-headed households and four zip codes with facilities in the lowest category.

The zip codes that contain the most facilities range in percentages of female-headed households from 10.4 to 19.5. The highest three percentages of those zip codes with facilities are 40203 with 34.1 percent, 40206 with 28.3 percent and 40208 with 26.0 percent, all of which are within the city of Louisville. The lowest percentages of female-headed households that contain facilities are 42058 (Ledbetter) with 12.6 percent, 40109 (Brooks) with 12.4 percent and 40121 (Fort Knox) with 0.1 percent. The lowest percentage found in all zip codes is 0.0. The highest percentages of female-headed households found in all of the zip codes include 41640 (Hueysville) with 58.0 percent.

The map of the percentage of persons below poverty shows that the highest percentages are mostly concentrated in the eastern portion of the state (Figure 16). Of the highest category, which ranges from 33.8 to 100.0 percent, there are only three zip codes that contain facilities. These zip codes are 40203 and 40211, both in Louisville, and 41267 in Warfield. In the second highest category of persons below poverty, there is only one zip code that is also in the second highest category of number of facilities. This zip code is located in Louisville (40212). There are ten other zip codes that contain facilities included in this category. There are twenty-one zip codes in the third highest category of persons below the poverty level and twenty-five zip codes with facilities in the lowest category.

The zip codes that contain the most facilities range in percentages of persons below poverty from 10.9 to 16.4. The highest three percentages of those zip codes with facilities are 40203

(Louisville) with 46.4 percent, 41267 (Warfield) with 42.2 percent and 40211 (Louisville) with 34.5 percent. The lowest percentages of female-headed households that contain facilities are 40059 (Prospect) and 41017 ((Fort Mitchell) with 2.9 percent, 41042 (Florence) with 7.1 and 40121 (Fort Knox) with 7.2. The lowest percentage found in all zip codes is 0.0. The highest percentages of persons below poverty found in all of the zip codes include 41351 (Mistletoe), 42250 (Huff), and 40825 (Dizney), each with 100 percent.

The map of the per capita incomes for zip codes, just as it was for the counties, shows the highest amounts in the northern Kentucky area and the Louisville to Lexington corridor (Figure 17). Again, since the relationship hypothesis is negative, this part of the map analysis will discuss it in this manner as well. Of the category with the lowest per capita incomes, ranging from $0 to $7,030, there are only two zip codes that contain facilities. These include 40203 (Louisville) and 42717 (Burkesville). In the category with the second lowest per capita incomes there are seven counties containing facilities, two of them with two. These include 40211 and 40212 in Louisville. There are sixteen zip codes in the category with the third lowest per capita incomes and fifteen zip codes with facilities in the highest per capita income category.

The zip codes that contain the most facilities range in per capita incomes from $10,630 to $12,246. The highest three incomes of those zip codes with facilities are 40059 (Prospect) with $31,530, 40510 (Lexington) with $23,349, and 41017 (Fort Mitchell) with $18,535. The lowest per capita incomes that contain facilities are 42343 (Fordsville) with $7,393, 42717 (Burkesville) with $6,992, and 40203 (Louisville) with $6,660. The lowest income found in all zip codes $930 in Dizney (40825). The highest per capita income is found in Prospect.

It is interesting to note that those counties listed as the highest percentages with facilities present were often sited more than once. For example, 40203 (Louisville) was in the top three percentages with facilities for four variables. 40211 (Louisville), 40208 (Louisville), and 40121 (Fort Knox) were listed in the top for two. Also noticeable for some zip codes is that they occur in the top categories for several variables. For example, 40214 (Louisville) is listed in four top categories. It has four facilities. 41101, 40213, 40216 and 41031 are all listed in two variables' highest category. These kinds of occurrences could point toward more specific siting issues with respect to a combination of socioeconomic variables.

4.1.3 Statistical Analysis

The regression statistics for the county analysis gives a r^2 value of 0.25 (Table 1). This number means that approximately 25 percent of the variance in the number of facilities is explained by all of the combined socioeconomic variables combined. This means that 75 percent of the variation in the number of facilities in a county is explained by variables other than those used in this study. The regression statistics for the zip code analysis give a r^2 value of 0.08. This number means that approximately 8 percent of the variance in the number of facilities is explained by the socioeconomic variables combined. This means that 92 percent of the variation is explained by other external variables. This shows that there is not a very strong cause and effect relationship between the variables that can be seen from this type of statistic. The coefficients noted in Table 2 show that the parameters are not significantly different from zero, which means that separately, there are no significant contributions from the independent variable towards the variation of the dependent variable.

Table 1

Regression Statistics

	County	Zip Code
R Squared	.25	.08

The bivariate correlation analysis for the county information shows that most of the variables used had a weak relationship with the number of facilities. The strongest relationship, still only moderate, is seen with per capita income, at 0.39, although it is not the hypothesized relationship. The correlations, as can be seen in Table 3 are all positive, with the exception of the percentage of persons below poverty at –0.24. The weakest correlation can be seen with the percentage of Hispanic origin at 0.21. Each of the correlations did occur, although moderately, in the hypothesized direction except the per capita income and the percentage persons below poverty. The positive relationships that did occur in the hypothesized direction included all of the racial information and the female-headed household data. This type of correlation means that as the number of TSD facilities increases, the percentage of these variables increases as well. The two variables that did not have the hypothesized direction of relation indicate that TSD facilities may be located in higher income or more prominent counties.

Table 2

Regression Coefficients

Coefficients		
Variable	***County***	***Zip Codes***
% Black	0.046	0.014
% Hispanic Origin	0.17	0.0042
% Other Nonwhite	0.60	0.0056
% Female Headed Households	0.11	0.0035
% Persons Below Poverty	0.15	-0.0017
Per Capita Income	0.000934	0.000013

Significance level - .05

Table 3

Bivariate Correlation Coefficient

Variable	***County***	***Zip Code***
	Correlation Coefficient With # of Facilities	***Correlation Coefficient With # of Facilities***
% Black	0.34	0.25
% Hispanic	0.21	0.06
% Other Nonwhite	0.26	0.08
% Female Headed Households	0.24	0.12
% Below Poverty	-0.24	-0.14
Per Capita Income	0.39	0.17

Significance level = .05

Weak = ±.1 to ±.3

Moderate = ±.3 to ±.6

Strong = ±.6 to ±1

The bivariate correlation coefficients for the zip code data show relationships that are not as strong as at the county level (3). The variable with the strongest relationship is the percentage Black with 0.25. The other variables only show weaker relationships ranging from 0.06 to 0.17. Again, the racial variables and the female headed household variable do have the hypothesized direction of correlation, and the per capita income and the percentage of persons below poverty have the opposite. The weakest correlation is again with the Hispanic origin percentages.

4.2 Distance Analysis

This analysis begins by creating buffers around those zip codes with facilities. As previously mentioned, each zip code was assigned a number to coincide with the distance from the zip codes containing facilities. The rings that were created by these assignments were separated and displayed on an ArcView map (Figure 18). These rings were created by the select by theme tool within ArcView. The zip codes immediately adjacent to the zip codes with facilities were selected for the first ring. The zip codes immediately adjacent to those zip codes in the first ring, not including those with facilities were selected for the second ring, and so on. The map analysis done here is a basic description of the patterns and zip codes within each ring, and some of the characteristics from the socioeconomic patterns within each. This analysis is followed by the statistical results from the bivariate correlation analysis, which was created in Excel from the transferring of the coinciding reclassification numbers for each ring, from ArcView. Also included in the statistical analysis is a t-test for equality of means. The same numbers were used here as in the bivariate coefficient, but manipulated into two groups, one with facilities and one without. The means of each variable are calculated and the t-value is analyzed in this chapter.

4.2.1 Zip Code Map Analysis

There are 731 zip codes shown on the map exhibiting the assigned buffers. Of those, 60 zip codes have facilities, 225 zip codes are immediately adjacent to the zip codes with facilities and categorized in the first ring, 181 zip codes are included in the second ring, 92 zip codes in the third and 173 zip codes in the fourth. The majority of the zip codes (30.8 percent) are within the first ring. The lowest percentage of zip codes is found in the third ring at 12.6 percent. The averages for each variable within each ring are included in Table 4. The highest averages for each variable is found in the zip codes with TSD facilities. These averages decrease, except for the percentage of other nonwhite persons, with distance from the zip codes with facilities. Although some decreases are not extremely large, they can still be seen in the numbers. In general, it would seem that the mapped patterns might be somewhat the same for the variables, as are seen for the map with buffers around the facilities.

4.2.2 Statistical Analysis

The bivariate correlation analysis for the distance information shows that most of the variables used had a moderate to weak relationship with the distance from facilities. Moderate relationships included the percentage Black, the percentage below poverty and the per capita income. The strongest relationship is seen with the percent of persons below poverty, at -0.44, although it is not the hypothesized direction of the relationship. The correlations, as can be seen in Table 5 are all positive, with the exception of the percentage of persons below the poverty level. The weakest correlation can be seen with the percentage of Hispanic origin at 0.06. Each of the correlations did occur in the hypothesized direction except the per capita income and the percentage persons below poverty. The positive relationships that did occur in the hypothesized direction included all of the racial information and the female-headed household data. This type of correlation means that as the assigned number denoting distance from the zip codes with the TSD facilities increases, the percentage of these variables increases as well. In other words, the distance from the zip codes is associated with the percentages of the variables; as the distance from the zip codes with facilities increases, the percentage of these variables decreases. The two variables that did not have the hypothesized direction of relation indicate that higher income or more prominent neighborhoods are closer to the facilities than low income**.**

Table 4

Averages for Zip Codes with Facilities and Each Surrounding Buffer

Variable	Zip Codes With Facilities	Ring 1	Ring 2	Ring 3	Ring 4
% Black	10.3	3.7	2.5	0.9	0.4
% Hispanic	0.7	0.4	0.4	0.3	0.3
% Other Nonwhite	1.0	0.5	0.3	0.5	0.3
% Female headed Households	16.2	13.0	12.6	11.1	10.9
% Below Poverty	17.8	18.6	23.7	31.6	33.6
Per Capita Income	$11,309	$10,533	$9,195	$7,505	$7,245

Table 5

Bivariate Correlation Coefficient for Distance

	Zip Code
Variable	***Correlation Coefficient with Distance Classification***
% Black	0.29
% Hispanic	0.06
% Other Nonwhite	0.11
% Female Headed Households	0.20
% Persons Below Poverty	-0.44
Per Capita Income	0.43

Significance level = .05

Weak = $\pm$.1 to $\pm$.3

Moderate = $\pm$.3 to $\pm$.6

Strong = $\pm$.6 to $\pm$1

The t-test statistical analysis for the difference in means showed that there were significant differences in the means for those zip codes with facilities and those without (Table 6). Variables that were found to have significantly lower means for zip codes with out facilities than for those with include the percentage Black, percentage Hispanic, percentage other nonwhite, and per capita income. For these variables the critical value was found to be approximately 2.00 and the t statistic was found to be around -4.00. This indicates that there was indeed a significant difference in the means for these percentages between those zip codes with facilities and those without. However, it indicates that the percentages in non-facility areas are lower than in areas with facilities. This means that higher amounts of these variables are in areas with no facilities. Therefore, in these cases the null hypothesis could not be accepted at the significance level 0.05, and indications are that environmental justice is taking place in siting of TSDFs. Concerning the percentage of persons below poverty, again there was a significant difference in the means, this time the mean is higher in zip codes without facilities than in those with facilities present. In this case the critical value is again 2.00 and the t-statistic is 6.48. Again the null hypothesis could not be accepted, as the t-value exceeded the critical value.

Table 6

t-Test for Difference in Means

Variable	***Critical Value***	***t-Statistic***
% Black	+2.00, -2.00	-3.82
% Hispanic	+2.00, -2.00	-4.70
% Other Nonwhite	+1.99, -1.99	-4.03
% Female Headed Households	+1.99, -1.99	-4.03
% Below Poverty	+1.99, -1.99	6.48
Per Capita Income	+2.00, -1.99	-4.76

These results have shown the relationships in the case of Kentucky between the socioeconomic variables and the number of TSD facilities. The following chapter addresses these analysis in discussion and conclusions and gives some recommendations based on these findings as well as for future references. The limitation of this analysis and its results is also discussed.

5.1 Discussion

In discussing this study the first section to look at is the map analysis. This analysis found some interesting likenesses in some of the specific zip codes and counties with facilities in that several of the counties listed in the highest percentage of one variable was often listed in others. This occurred in both the county and zip code comparisons. Most notably, Jefferson County and zip codes within Jefferson County were at the top of the list on several occasions. This could point to the need for a more detailed analysis to determine if there are disproportionate risks in this specific area.

The results of found in this study indicate that while there may be some slight relationships with the amount of minority populations and the number of facilities in the hypothesized positive direction (as the number of facilities goes up, so do the variable percentages), the hypothesized relationships for the number of persons below poverty and the per capita income relationships were exactly opposite. In fact the relationships with these two variables and the number of TSDs was among the strongest, especially within the distance analysis. This relationship may be explained by the reasoning that these facilities may pay high wages and that their workers live near the facility. It also may be caused by the fact that these facilities may be located in an industrial area that also has high wages and employ surrounding residents. This may be seen as compensation for the risks. Income in these areas could also be skewed by a few extremely high incomes that would drive the average up and not show precisely the characteristics of the neighborhoods immediately adjacent to the facilities. The difference in incomes could also be a function of the urban and rural separation.

Since the highest correlation for both zip codes and counties, excluding the distance analysis, was found to be with the percentage of Blacks, this would seem to be the strongest relationship and should be the one concentrated on in future studies. This relationship may be moderately attributed to environmental injustices, or could be a factor of inner city versus rural population characteristics. The fact that the other variables had a higher correlation with the county data than the zip code data may be a factor of distance and population dispersal.

The percentage of female-headed households is an interesting variable. Although it looks on the map as if it has no specific pattern, in the statistical analysis it has a weak relationship at the county level and is among the strongest (although still weak) at the zip code level.

For the distance analysis, as before mentioned, the strongest correlations were found with the percentage below poverty and the per capita income. The relationships with the minority variables indicate weak correlations with each variable. The female-headed household correlation was again between weak and moderate. The distance analysis did show some moderate relationships with three of the variables. This indicates that the distance from a TSD facility is a factor that needs to be included in determining any relationship of this sort. Proximity is key in determining spatial relationships, i.e. the areas in close contact with high health risks and their population characteristics.

Finally, the hypotheses findings for the county data are as follows:

1) Accepted at a moderate level - The greater the number of facilities, the greater the percentage of African Americans (positive relationship);
2) Accepted at a weak level - The greater the number of facilities, the greater the percentage of Hispanics (positive relationship);
3) Accepted at a weak level - The greater the number of facilities, the greater the percentage of other minority groups (positive relationship);
4) Accepted at a weak level - The greater the number of facilities, the greater the percentage of female-headed households (positive relationship);
5) Rejected (the direction of the relationship was opposite of the hypothesized) – The greater the number of facilities, the lower the percentage of persons below the state poverty level (negative relationship); and
6) Rejected (the direction of the relationship was opposite of the hypothesized) – The greater the number of facilities, the higher the per capita income (positive relationship).

The hypotheses findings for the zip code data are as follows:

1) Accepted at a weak level - The greater the number of facilities, the greater the percentage of African Americans (positive relationship);
2) Accepted at a weak level - The greater the number of facilities, the greater the percentage of Hispanics (positive relationship);
3) Accepted at a weak level - The greater the number of facilities, the greater the percentage of other minority groups (positive relationship);
4) Accepted at a weak level - The greater the number of facilities, the greater the percentage of female headed households (positive relationship);
5) Rejected (the direction of the relationship was opposite of the hypothesized) – The greater the number of facilities, the lower the percentage of persons below the state poverty level (negative relationship); and
6) Rejected (the direction of the relationship was opposite of the hypothesized) – The greater the number of facilities, the higher the per capita income (positive relationship).

The hypotheses findings for the second section of the project which included distance are the following:

1) Accepted at a moderate level - As the distance from each facility gets greater, the percentage of the population that is African American gets lower;
2) Accepted at a weak level - As the distance from each facility gets greater, the percentage of the population that is Hispanic gets lower;
3) Accepted at a weak level - As the distance from each facility gets greater, the percentage of the population labeled as other minority groups gets lower;
4) Accepted at a weak level - As the distance from each facility gets greater, the percentage of female-headed households gets lower;
5) Rejected (the direction of the relationship was opposite of the hypothesized) - As the distance from each facility gets greater, the percentage of the persons below poverty increases; and
6) Rejected (the direction of the relationship was opposite of the hypothesized) - As the distance from each facility gets greater, the per capita income gets higher.

The t-test statistical analysis resulted in the rejecting of the null hypothesis for all of the variables. This indicates that the relationship between the averages of each variable, were significantly different in zip codes with and without TSD facilities. However, these differences are opposite of the expected. The statistic shows that the zip codes lacking facilities have the highest per capita income and the least percentages of minorities, female-headed households and persons below poverty. Based on this statistic alone it would seem that there are no prejudices being produced from the siting of the facilities used in this study. However, the other tests done within this study should also be considered. Taking into account that the bivariate correlation coefficient found some weak relationships between the variables, it can be seen that the variables used in this study are not a likely determinant of TSD facility siting in Kentucky.

5.2 Conclusion

In conclusion, this project found minimal relationships between treatment, storage and disposal facilities and the percentages of low income and minorities. The strongest relationships were found with the percentage of Blacks, the percentage of persons below poverty and the per capita income. With both the percent Black and the per capita income, it was found that to some extent, the percentage and dollar amount increased with the number of facilities. The percentage of persons below poverty was found to decrease with the number of facilities. These findings implicate that there is only a slight bit of inequality in the siting of TSD facilities in Kentucky. Most likely, other factors are a higher determinant to siting that race or poverty.

This study shows that while not on a case by case basis, overall, there are very little environmental injustices occurring with resect to TSDF siting. This implies that the current policies and methods for site selection have been somewhat sufficient. However, since the tendencies could be there, policies could be strengthened and siting methodology and evaluation could be done more comprehensively, utilizing social issues and community participation.

5.3 Limitations

As with all projects this study has several limitations, from the data to the analysis. First the data for the socioeconomic variables was collected in 1990 and the data for the treatment, storage and disposal facilities was collected in 1998. This is a problem if several of the facilities were opened after 1990. In this case, the socioeconomic variables may have greatly changed since those facilities were created. This study assumes that the socioeconomic variables have not changed significantly since the 1990 Census was taken. This limits the analysis in such a way that more information will need to be collected and updated as it becomes available. For example, when the 2000 Census is taken, these numbers should be substituted in the statistical and map analysis in order to update the information and make for a more accurate analysis.

Also, some data was not given in the census for some zip codes. This created a problem again for the analysis in that some facilities were not included in the statistical and map analyses because data was not available. Also in the ArcView shape files for Kentucky, not all of the zip codes were present as well. This affects the data in such a way as to not give a complete picture of where the facilities are located and what the surrounding characteristics are. These facilities and socioeconomic data may have either weakened or strengthened the analyses.

As for the socioeconomic characteristics themselves, specifically the poverty and income data could be questionable in this study. The definition of poverty given by the census may not always be reflective in the neighborhood opinions. What may be a low quality of life to some may not be to others. This may affect the actual perception of what a poverty stricken or low-income area may be. As well, it is worth noting that in some areas the incomes of a few wealthy persons may have skewed the average so that it does not accurately depict the characteristics of the area.

Since zip codes and counties were utilized for this analysis, it is necessary to discuss the limitations of using data in such a manner. Although zip codes are relatively small units, they often vary in size and population characteristics. It would have been better to utilize census tract information since theoretically they are supposed to be somewhat homogeneous in characteristics. However, due to data availability, zip codes were chosen. When looking at the results, one needs to understand how the units that were used affect the outcome. As for the county data, it also is varied in size and composition. Several counties differ in basic characteristics such as number of persons. Again, this needs to be considered when consulting the results.

Also, as before mentioned, this is not a historical review of TSD facility siting and the characteristics of the surrounding communities before and after the siting. This fact creates an unknown circumstance of which came first, and thus time is the limitation in this case.

This study focuses on only one type of toxic waste facility. Therefore, it needs to be understood that the results in this study can not be generalized to other types of facilities. This means that for any other type of facility, while the methodology may be the same, the data collection and use of the information may need to different. As well, generalities may not be made, based on the results in this study, to specific sites. The outcomes of this project are for an overall view of TSD sites in Kentucky.

Finally, as with any other statistical analysis, the results of the methods used in this project do not represent fact.

5.4 Further Studies/Recommendations

The findings in this study should not be taken as a final step in the analysis of the location of treatment, storage and disposal facilities in Kentucky. Further studies need to be conducted in order to come to an overall conclusion on the situation of this issue. These studies include a historical analysis detailing the socioeconomic conditions of the surroundings before and after siting of the facilities. This analysis would help to determine if possibly a large composition of minorities or low income persons preceded the location of the facility, or if the composition was created after the facility was sited in that specific area. As well, a case study on one specific facility or one specific zip code or county with several facilities could be done. This analysis could entail an in depth look at the actual siting of the facility or facilities, such as specifics on why the site was chosen and what the alternatives were. Here the economic and the environmental issues that played a part in the siting decision could be discussed, along with any social issues. Also within this type of analysis, it would be important to understand what the opinion of the community was, at the time of siting, and is presently, as well as the demographic characteristics before and after siting. Yet another interesting twist to this topic would be to delve into a policy analysis of exactly how the national environmental justice policies and regulations trickled down to the local level to effect the specific siting of a facility in Kentucky.

Although obviously no community will welcome health risks into their back yard, if incentives and compensation are given in addition to possibly much needed jobs for the low skilled residents of the community, such facilities may not always be seen in a negative light. For example, if the general public is invited to be involved in the early stages of planning for a toxic facility then they will be less likely to be suspicious of the company. If the community residents are not involved until halfway through the process when someone complains, it often seems like the company is trying to hide something. One good way to actively involve the community is to initially present them with a list of criteria for siting the facility and allow them to add to and make comments on such. Finally, when an evaluation of sites has been done, the community participants should be provided with reasons why possible sites are or are not acceptable based on the requirements set forth previously, as well as any external pros or cons to allowing such a facility to site in the neighborhood. Of course, it is not always easy to get residents to participate in such events. This is the role of the environmental organizations. Their addition to the community awareness of the issues could be the driving force of enhanced participation in the siting decision making.

There can be several pros and cons of inviting such a toxic waste facility as TSDs to the community. As with any new industries there is potential for an increased tax base, more jobs, greater infrastructure and services, more diversity in economic activity and investment into the community, and more dollars being generated within the community via the multiplier effect. However, although less calculable, the health risks of such facilities are ever present. These risks come from the potential spillage of toxins into all aspects of the environment, from which it has been hypothesized, will increase mortality rates over time. These toxic facilities also create environmental degradation and can limit the future use of the land. All of these things are important factors in choosing a site for such a facility and should be presented to the public so they can be well informed to make the decision to support the facility or not. Other incentives may be included that may serve to sway doubting communities. These incentives from the industry may include the following: 1) higher tax payment, contributions to the community for schools or other worthy causes; 2) employment guarantees, such as guaranteeing that a certain percentage of the new jobs will go to the local workers; 3) high wages and benefits to workers; 4) payment for and creation of new infrastructure such as roadways; and 5) maintenance of special dedicated green space to make up for any environmental degradation that may occur due to the incoming industry.

There may also be a need for increased policy creation on the subject of siting toxic facilities that may prove hazardous to the health of surrounding persons. Again, compensation for risk could be written into the local legislation for siting, drawing from any development or economic needs that the community may have. Also policies passed down from the EPA could be enhanced. For example, the environmental impact statement could be utilized as a tool for social issues as well as environmental. While it is the role of all federal agencies to incorporate environmental justice measures in their work, it may be that this is not adequate enough to address the problem. EIA could be a tool not only to evaluate the soil, air, water and other environmental issues, but also to take into account the social characteristics of the surrounding community, what other toxic sites are located in the area, if any alternative sites are suitable that will offer more equality in distribution, and the community's awareness and approval or disapproval of the facility.

At the local level, comprehensive planning and zoning could be an excellent tool to guide development of such toxic facilities. In this case, each municipality should address the issue of siting from a proactive point of view. Having standards and areas zoned for industrial, specifically toxic industries, will make it easier for those industries to choose where to locate and easier for localities to have a say where they may or may not do so. If zoning is implemented in such a manner as to take into account the suitability of the land to the industry as well as the socioeconomic characteristics and community opinion, then facility siting will not be considered an issue of injustice.

This process would include several of the important factors mentioned in this study and would be easy to create and to comply with. From this standpoint, it should be the initiative of the government to choose satisfactory sites for such facilities through municipal planning and zoning ordinances, created with community participation and overall approval. Making the community feel as if they have a hand in choices concerning toxic waste, and their own health, will reduce the fear of such industries. However, this recommendation does depend on the area having a comprehensive plan and zoning ordinances in place.

Each of these recommendations has its strengths and weaknesses. What is common throughout is that, no matter through what medium, community participation is an important factor for creating consensus and reducing suspicion of these types of facilities. Although this study did not find outstanding significances in the relationships between socioeconomic factors and the number of TSDs in Kentucky, creating an easier and more agreeable site selection process for toxic waste facilities, including TSDs, through community participation, will be an asset to all parties involved.

Figure 3
Kentucky Counties,
1999

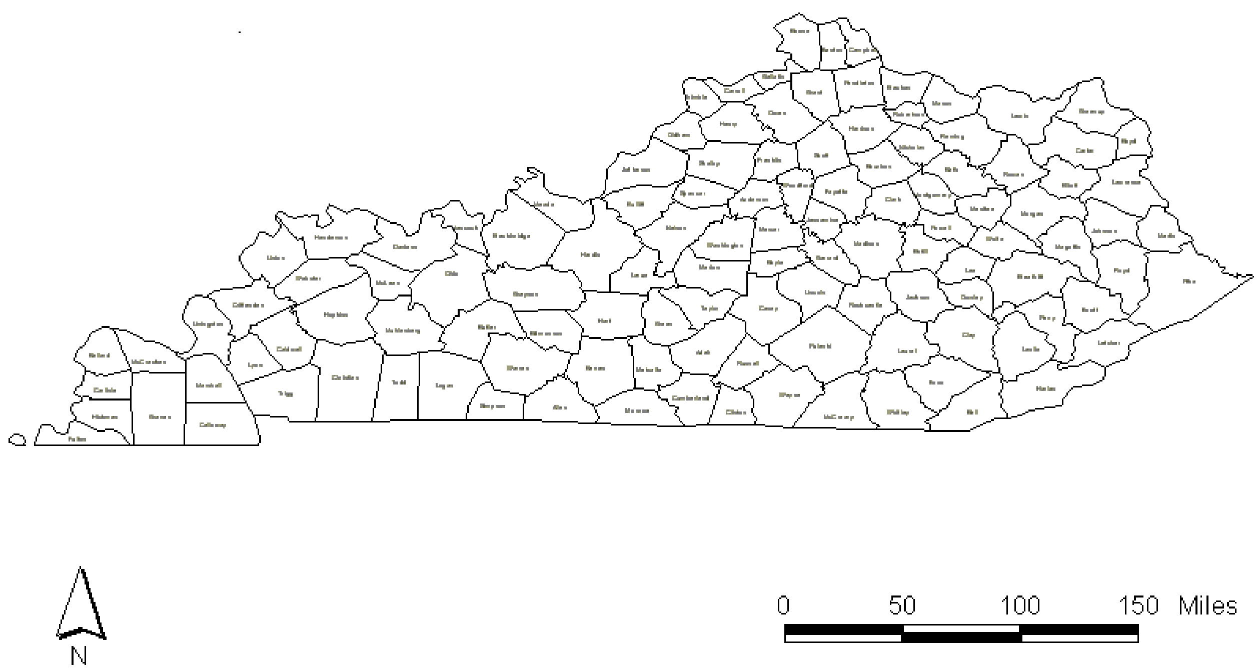


Figure 4
Number of Treatment, Storage and Disposal Facilities,
by County, Kentucky, 1998

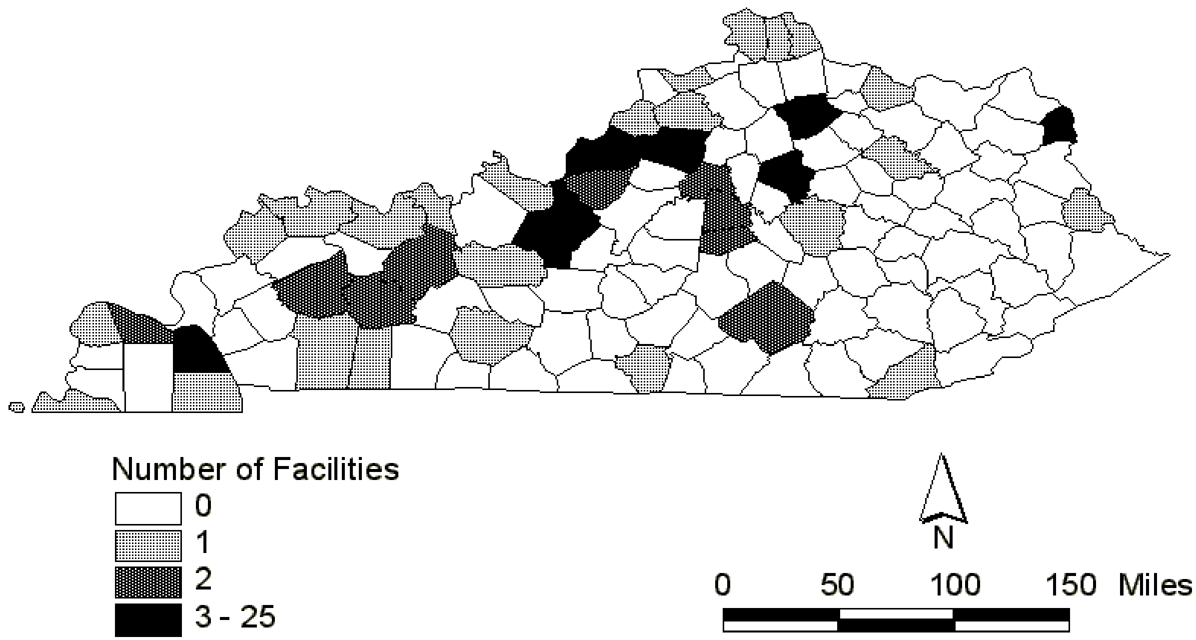


Figure 5
Number ofTreatment, Storage and Disposal Facilities by Zip Code, Kentucky, 1998

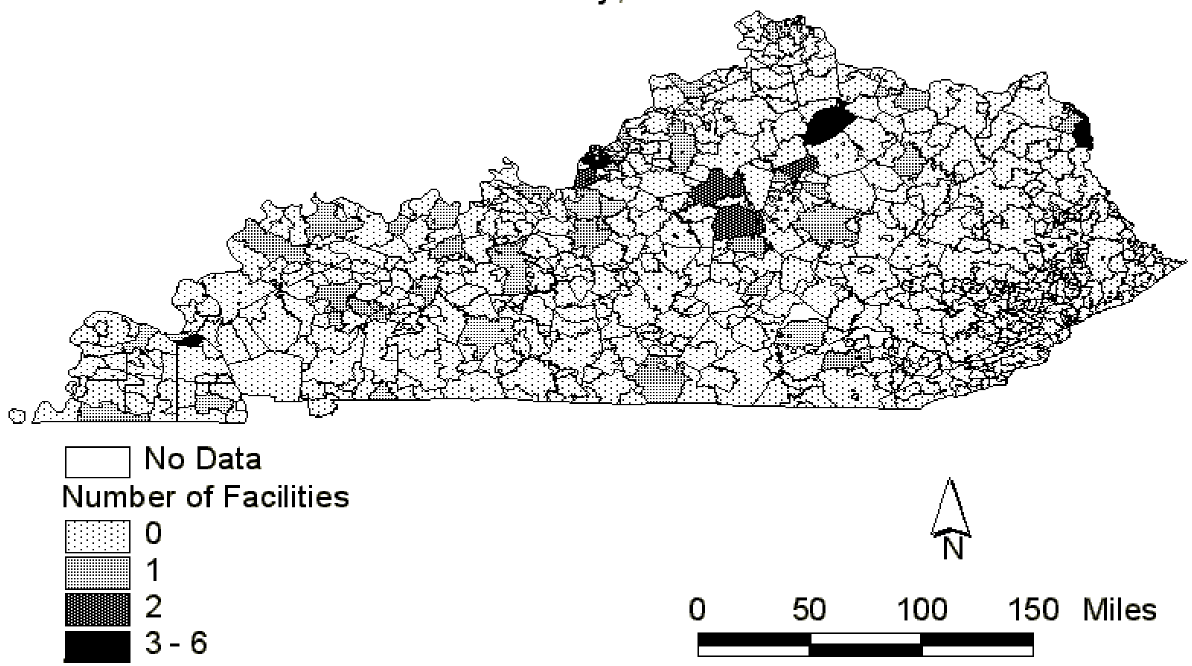


Figure 6
Percentage Black Population by County,
Kentucky, 1990

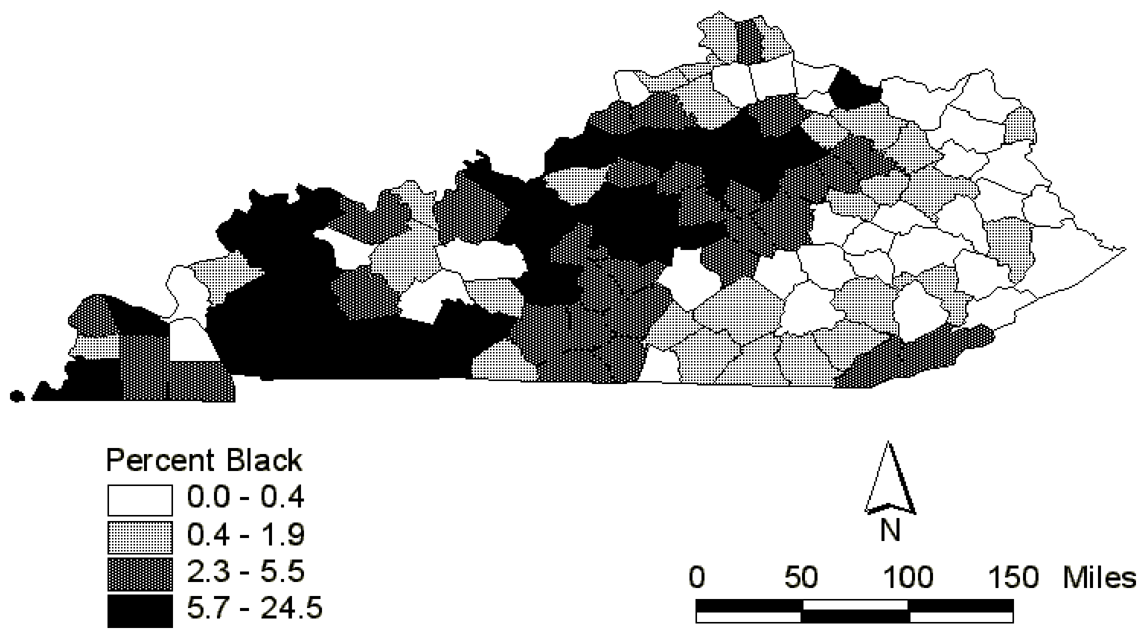


Figure 7
Percentage Hispanic Population by County,
Kentucky, 1990

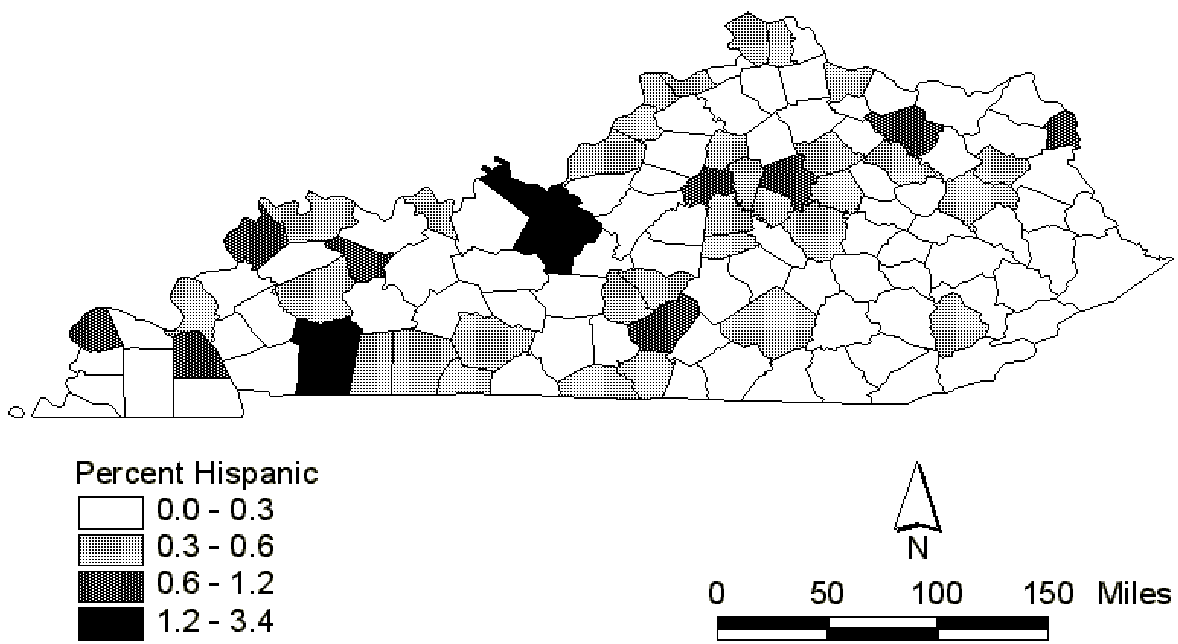


Figure 8
Percentage Other Nonwhite Population by County, Kentucky, 1990

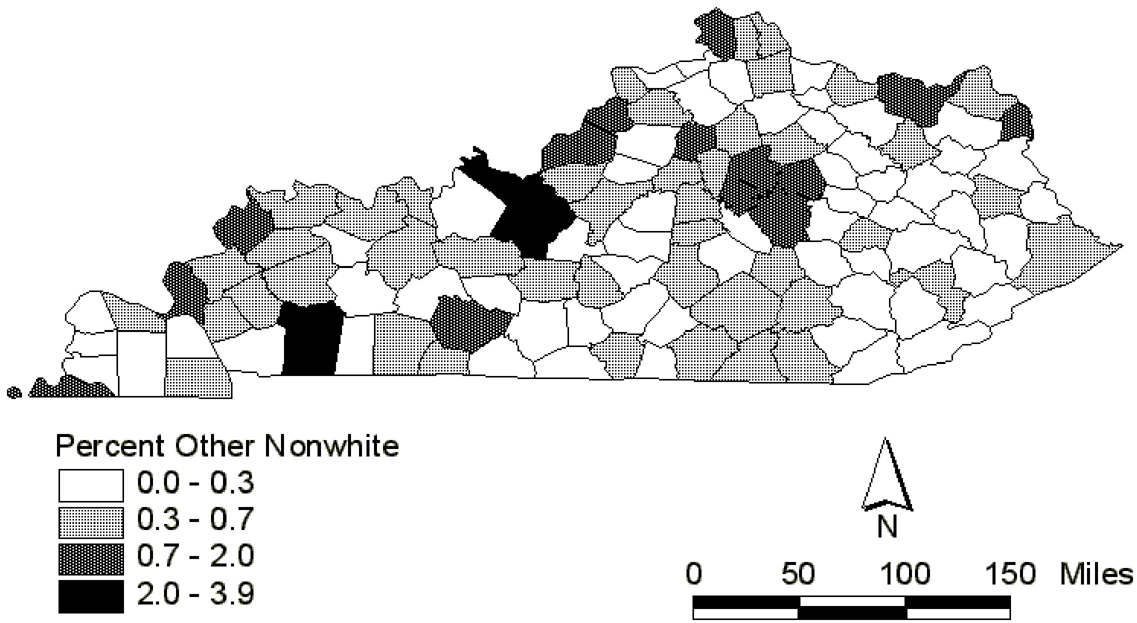


Figure 9
Percentage Female Headed Households by County,
Kentucky, 1990

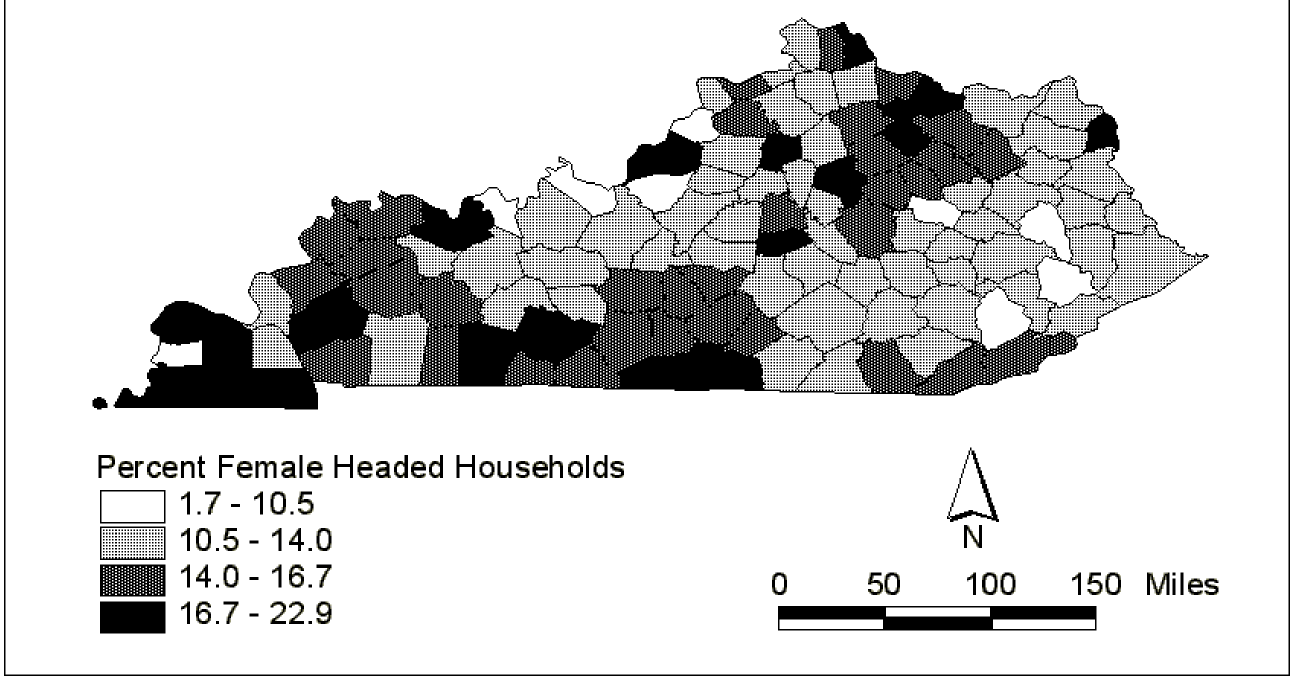
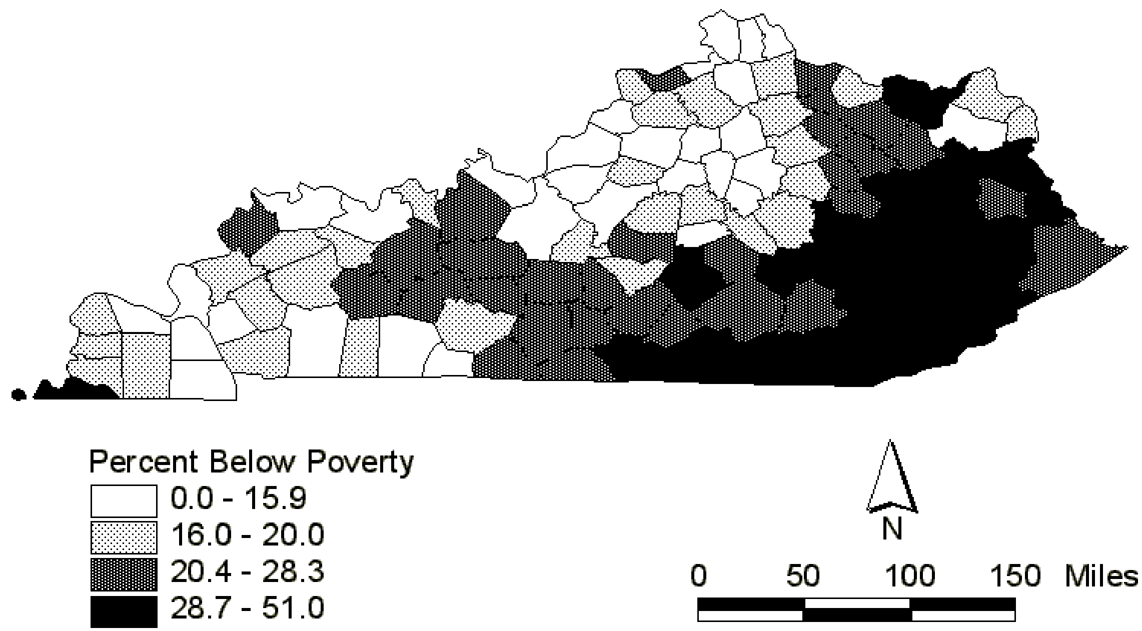


Figure 10
Percentage of the Population Below Poverty Level by County, Kentucky, 1990



GIS in Statewide Ground-Water Vulnerability Evaluation to Pollution Potential

Navulur Kumar and Bernard A. Engel
Department of Agricultural Engineering, Purdue University, West Lafayette, Indiana

Abstract

The ground-water vulnerability of Indiana to pollution potential was evaluated using a geographic information systems (GIS) environment. The Geographic Resources Analysis Support System (GRASS) and the GRID submodule of ARC/INFO were used to conduct the analysis and to identify and display the areas sensitive to ground-water pollution potential. The state soils geographic (STATSGO) database was employed to retrieve statewide soils information required for the analysis. The information from the STATSGO database was used in two models, DRASTIC (acronym representing the following hydrogeologic settings: **D**epth to water table, aquifer **R**echarge, **A**quifer media, **S**oil media, **T**opography, **I**mpact of vadose zone, and hydraulic **C**onductivity of the aquifer) and SEEPAGE (System for Early Evaluation of Pollution Potential of Agriculture Ground-Water Environments). These models employ a numerical ranking system and consider various hydrogeologic settings that affect the ground-water quality of a region. Ground-water vulnerability maps were prepared for the state of Indiana based on DRASTIC and SEEPAGE results. Continuing work is planned to determine the accuracy of the results by comparing the existing well-water quality data. The DRASTIC Index and SEEPAGE Index number (SIN) maps show great potential as screening tools for policy decision-making in ground-water management.

Introduction

Ground-water contamination due to fertilizer and pesticide use in agricultural management systems is of wide concern. In 1989, reports of ground-water contamination in New York wells led the U.S. Environmental Protection Agency (EPA) to conduct a nationwide survey on well contamination in the United States. These wells were tested for presence of nitrate, pesticides, and pesticide breakdown products (1). Statistically, the wells selected represent more than 94,600 wells in approximately 38,300 community water systems. Over 52 percent of the community water systems and 57 percent of the rural domestic wells tested contained nitrates (2).

Indiana has abundant ground-water systems providing drinking water for 60 percent of its population. A study on well-water quality detected pesticides in 4 percent of wells tested in Indiana. Also, 10 percent of private wells and 2 percent of noncommunity wells contained excessive nitrate levels (3).

Statewide maps showing the areas vulnerable to ground-water contamination have many potential uses such as implementation of ground-water management strategies to prevent degradation of ground-water quality and monitoring of ground-water systems. These maps will be helpful in evaluating the existing and potential policies for ground-water protection. Ground-water models such as SEEPAGE (System for Early Evaluation of Pollution Potential of Agriculture Ground-Water Environments) and DRASTIC (acronym representing the following hydrogeologic settings: **D**epth to water table, aquifer **R**echarge, **A**quifer media, **S**oil media, **T**opography, **I**mpact of vadose zone, and hydraulic **C**onductivity of the aquifer) can be applied on a regional scale to develop such maps.

The data layers required for these models are commonly available data such as pH and organic matter content. For most states, the statewide ground-water vulnerability maps generated using DRASTIC were produced from 1:2,000,000-scale data (4). EPA (2) found that these maps did not correlate well with the water quality analysis performed for the national survey of pesticides in drinking water wells. States need more detailed and accurate maps to implement ground-water management programs. The state soils geographic (STATSGO) database at the 1:250,000-scale might be useful for studies at a larger scale.

The geographic information systems (GIS) environment is widely applied for diverse applications in resources management and other areas. It offers the facilities to store, manipulate, and analyze data in different formats and at different scales. The DRASTIC and SEEPAGE models can be integrated within the GIS environment to produce the final ground-water vulnerability maps.

Objectives

The purpose of the study was to prepare maps showing areas in Indiana vulnerable to ground-water pollution. This goal was accomplished by considering hydrogeologic factors in each region that affect the mobility and leaching of the contaminant reaching the aquifer. The prime objectives of this research were to:

- Evaluate Indiana's ground-water vulnerability to pollution potential using the DRASTIC and SEEPAGE models:
  - Integrate and evaluate the models in a GIS environment (Geographic Resources Analysis Support System [GRASS] ARC/INFO).
  - Develop a graphic user interface (GUI) in ARC/INFO to conduct the analyses.
- Compare the pollution potential map from the DRASTIC model with the map developed using the SEEPAGE Index number (SIN).
- Validate the accuracy of the present approach by comparing the vulnerability maps with the existing well-water quality data sampled across the state.

DRASTIC

DRASTIC is a ground-water quality model for evaluating the pollution potential of large areas using the hydrogeologic settings of the region (4-6). EPA developed this model in the 1980s. DRASTIC includes different hydrogeologic settings that influence a region's pollution potential. A hydrogeologic setting is a mappable unit with common hydrogeologic characteristics. This model employs a numerical ranking system that assigns relative weights to parameters that help evaluate relative ground-water vulnerability to contamination.

The hydrogeologic settings that make up the acronym DRASTIC are:

- *[D] Depth to water table:* Compared with deep water tables, shallow water tables pose a greater chance for the contaminant to reach the ground-water surface.
- *[R] Recharge (net):* Net recharge is the amount of water per unit area of soil that percolates to the aquifer. This is the principal vehicle that transports the contaminant to the ground water. Higher recharges increase the chances of the contaminant being transported to the ground-water table.
- *[A] Aquifer media:* The material of the aquifer determines the mobility of the contaminant traveling through it. An increase in travel time of the pollutant through the aquifer increases contaminant attenuation.
- *[S] Soil media:* Soil media is the uppermost portion of the unsaturated zone/vadose zone characterized by significant biologic activity. This, in addition to the aquifer media, determines the amount of water percolating to the ground-water surface. Soils with clays and silts have larger water holding capacity and thus increase the travel time of the contaminant through the root zone.
- *[T] Topography (slope):* The higher the slope, the lower the pollution potential due to higher runoff and erosion rates, which include pollutants that infiltrate the soil.
- *[I] Impact of vadose zone:* The unsaturated zone above the water table is referred to as the vadose zone. The texture of the vadose zone determines the travel time of the contaminant. Authors of this model suggest using the layer that most restricts water flow.
- *[C] Conductivity (hydraulic):* Hydraulic conductivity of the soil media determines the amount of water percolating through the aquifer to the ground water. For highly permeable soils, the travel time of the pollutant is decreased within the aquifer.

The major assumptions outlined in DRASTIC are:

- The contaminant is introduced at the surface.
- The contaminant reaches ground water by precipitation.
- The contaminant has the mobility of water.
- The area of the study site is more than 100 acres.

DRASTIC evaluates pollution potential based on the seven hydrogeologic settings listed above. Each factor is assigned a weight based on its relative significance in affecting the pollution potential. Each factor is also assigned a rating for different ranges of the values. Typical ratings range from 1 to 10, and weights range from 1 to 5. The DRASTIC Index, a measure of pollution potential, is computed by summation of the products of rating and weights of each factor as follows:

DRASTIC Index =
$$D_rD_w + R_rR_w + A_rA_w + S_rS_w + T_rT_w + I_rI_w + C_rC_w$$

where:

Dr = Ratings for the depth to water table
Dw = Weights for the depth to water table
Rr = Ratings for different ranges of aquifer recharge
Rw = Weights for the aquifer recharge
Ar = Ratings for the aquifer media
Aw = Weights for the aquifer media
Sr = Ratings for soil media
Sw = Weights for soil media
Tr = Ratings for topography (slope)
Tw = Weights for topography
Ir = Ratings for the vadose zone

Iw = Weights for the vadose zone
Cr = Ratings for different rates of hydraulic conductivity
Cw = Weights for hydraulic conductivity

DRASTIC assigns two different weights depending upon the type of contaminant. Pesticides are given different weights than general contaminants. In assigning the weights, DRASTIC considers the different properties of pesticides as they travel through the vadose zone and root zone of the soil media.

The higher the DRASTIC Index, the greater the relative pollution potential. The DRASTIC Index is divided into four categories: low, moderate, high, and very high. The sites with high and very high categories are more vulnerable to contaminations and hence should be reviewed by the site specialist. These weights are relative, however. Low pollution potential does not necessarily indicate that a site is free from ground-water contamination. It indicates only that the site is less susceptible to contamination than sites with high or very high DRASTIC ratings.

SEEPAGE

The SEEPAGE model is a combination of three models adapted to meet the Soil Conservation Service's (SCS's) need to assist field personnel (7, 8). SEEPAGE considers hydrogeologic settings and physical properties of the soil that affect ground-water vulnerability to pollution potential. SEEPAGE is also a numerical ranking model that considers contamination from both concentrated and dispersed sources.

The SEEPAGE model considers the following parameters:

- Soil slope
- Depth to water table
- Vadose zone material
- Aquifer material
- Soil depth
- Attenuation potential

The attenuation potential further considers the following factors:

- Texture of surface soil
- Texture of subsoil
- Surface layer pH
- Organic matter content of the surface
- Soil drainage class
- Soil permeability (least permeable layer)

Each factor is assigned a numerical weight ranging from 1 to 50 based on its relative significance, with the parameter that has the most significant effect on water quality assigned a weight of 50 and the least significant assigned a weight of 1. The weights are different for concentrated or site-specific sources, and dispersed or nonspecific sources.

Similar to DRASTIC, each factor can be divided into ranges and ratings, varying from 1 to 50. The ratings of the aquifer media and vadose zone are subjective and can be changed for a particular region. Once the scores of the six factors are obtained, they are summed to obtain the SIN. These values represent pollution potential, where a high SIN implies relatively more vulnerability of the ground-water system to contamination. The SIN values are arranged into four categories of pollution potential: low, moderate, high, and very high. A high or very high SIN category indicates that the site has significant constraints for ground-water quality management (7).

GIS

GIS has been widely used for natural resources management and planning, primarily during the past decade. A GIS can be combined with a ground-water quality model to identify and rank the areas vulnerable to pollution potential for different scenarios and land use practices. Many GIS software packages are available. GRASS is a raster-based public domain software developed by the U.S. Army Construction Engineers Research Laboratory (9). This software can assign different weights to, or reclass, the data layers and combine map layers, and is suitable for implementing the DRASTIC and SEEPAGE models. ARC/INFO is a GIS software developed by Environmental Systems Research Institute (ESRI) in Redlands, California. The GRID submodule of ARC/INFO facilitates the handling of raster data. Also, the capability to develop a menu-based GUI helps users easily implement the models. The GRID submodule also can reclass and manipulate the map layers suitable for conducting the analyses.

Methodology

Developing the Data Layers in GRASS and ARC/INFO

The STATSGO database from SCS comes at a scale of 1:250,000 and is distributed in different data formats. This study used the STATSGO database in the ARC/INFO format. The database is organized into map units that have up to 21 components. These map components have information assigned to layers of soil horizons. Each layer is attributed various soil properties such as pH or organic matter content (10). Each property is assigned a high and a low value for a map unit. The STATSGO map for Indiana is available in the vector format. This map was exported to GRASS as a vector coverage (11) and was converted into a raster coverage within the GRASS GIS environment. This was used as the base map for the DRASTIC and SEEPAGE analyses. The hydrogeologic parameters required for the models were identified from the corresponding INFO data tables and were exported into an ASCII file. Code was developed to generate a GRASS reclass file assigning the weighted values of the parameters to the corresponding map units in the base map. The STATSGO base map imported into GRASS was reclassed for each hydrogeologic setting (e.g., topography, pH) to create the data layers required for DRASTIC and SEEPAGE analyses.

The map layers of the hydrogeologic parameters in GRASS were then exported to ARC/INFO as raster coverages. A GRASS command was developed that allows the output ASCII file from GRASS to be imported into ARC/INFO directly without further modifications to the header in the ASCII file.

Developing a Graphic User Interface in ARC/INFO

The dynamic form-menu option (12) was used to develop a GUI for both DRASTIC and SEEPAGE analyses (see Figures 1 and 2). Because ratings for some parameters are subjective, the GUI provided an option to change the weights assigned to hydrogeologic settings. The coverages must already be assigned ratings before using the interface, however. The interface also allows users to reclassify the final vulnerability maps qualitatively (13) into four categories (low, moderate, high, and very high) after viewing the range of DRASTIC Index or SIN values.

Conducting the Analyses

The data layers were developed separately for the high and low values of the hydrogeologic settings. Once all the data layers were compiled, the corresponding ratings and weights were assigned and the analyses were conducted using the GUI. The data layers aquifer recharge, aquifer media, and vadose zone media were not available, so the analyses were conducted without these base maps. The SEEPAGE analysis was performed for concentrated/point sources of pollution. The final vulnerability indexes from the analyses were classified into four categories (low, medium, high, and very high) (see Table 1) to generate the final statewide vulnerability maps.

Table 1. Pollution Potential Categories Using SEEPAGE and DRASTIC Indexes

	Range of DRASTIC/SEEPAGE Index			
Analysis	**Low**	**Moderate**	**High**	**Very High**
SEEPAGE	1-24	25-48	49-70	> 70
DRASTIC	30-70	71-100	101-110	> 110

Validating the Accuracy of the Vulnerability Maps

The ground-water vulnerability maps produced by DRASTIC analysis were compared with those generated using the SEEPAGE model. The final statewide ground-water vulnerability maps from either approach were compared with the well-water quality data sampled from over 2,500 wells (see Figure 3), and the number of wells falling into each vulnerability category was tabulated.

Results and Discussion

Statewide analysis of ground-water vulnerability to pollution potential was conducted using the DRASTIC and SEEPAGE analyses at a scale of 1:250,000 in the raster format. The analyses were conducted for both the high and low values of the hydrogeologic settings, and the final vulnerability maps were prepared for the state of Indiana (see Figures 4 and 5). The vulnerability maps from both approaches were compared in the GRASS environment.

In both analyses, the low values of hydrogeologic settings resulted in more areas being classified as high and very high categories, compared with the high values of hydrogeologic settings in a map unit. The DRASTIC analysis placed more areas in the very high vulnerability category, compared with the SEEPAGE analysis, which categorized the same areas as high vulnerability. The nitrate-nitrogen concentrations observed in the wells were compared with the final vulnerability maps, and the number of wells falling into each of the four vulnerability categories (low, moderate, high, and very high) were summarized (see Tables 2 and 3).

Approximately 80 percent of the wells with concentrations less than 5 parts per million are classified under the moderate vulnerability category in SEEPAGE analysis. Overall, the results from the analyses did not correlate satisfactorily with the observed well-water quality data. Unavailability of the data layers aquifer media, aquifer recharge, and vadose zone media may have caused these results. The well-water quality data of nitrate-nitrogen contaminations was considered only for testing map accuracy, whereas the analyses do not account for the type of contaminant, its severity, and its volume in the generation of vulnerability maps. Other limitations of these approaches, including that the factors influencing aquifer contamination (e.g., direction of water flow, land use, population at risk, point sources of pollution) are not considered in the ground-water vulnerability evaluation, might also have led to the observed results.

The DRASTIC and SEEPAGE analyses can be improved by incorporating data layers such as land use and nitrate loadings in computing the DRASTIC and SEEPAGE Indexes. The STATSGO database can be used for

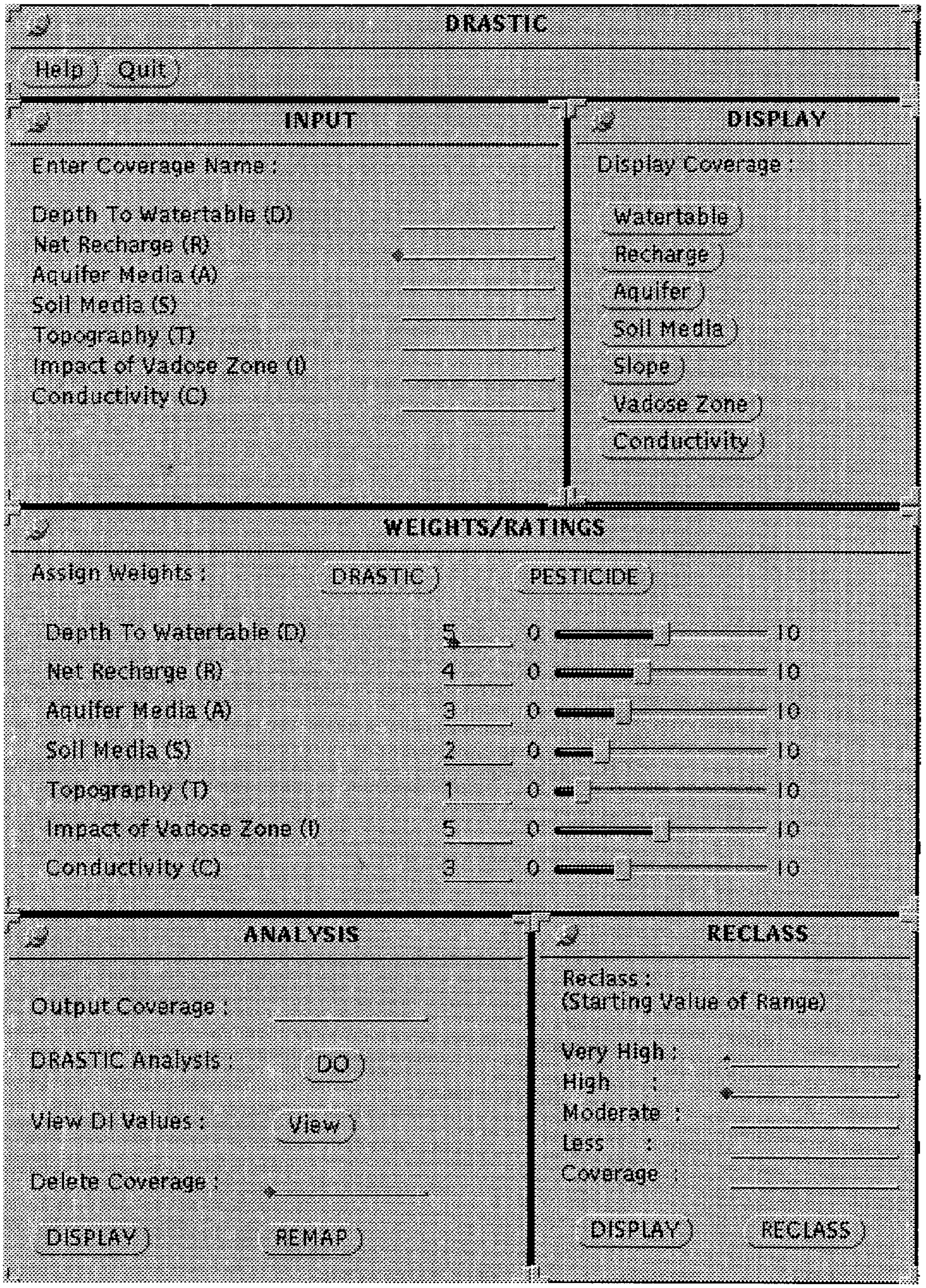


Figure 1. GUI for DRASTIC analysis.

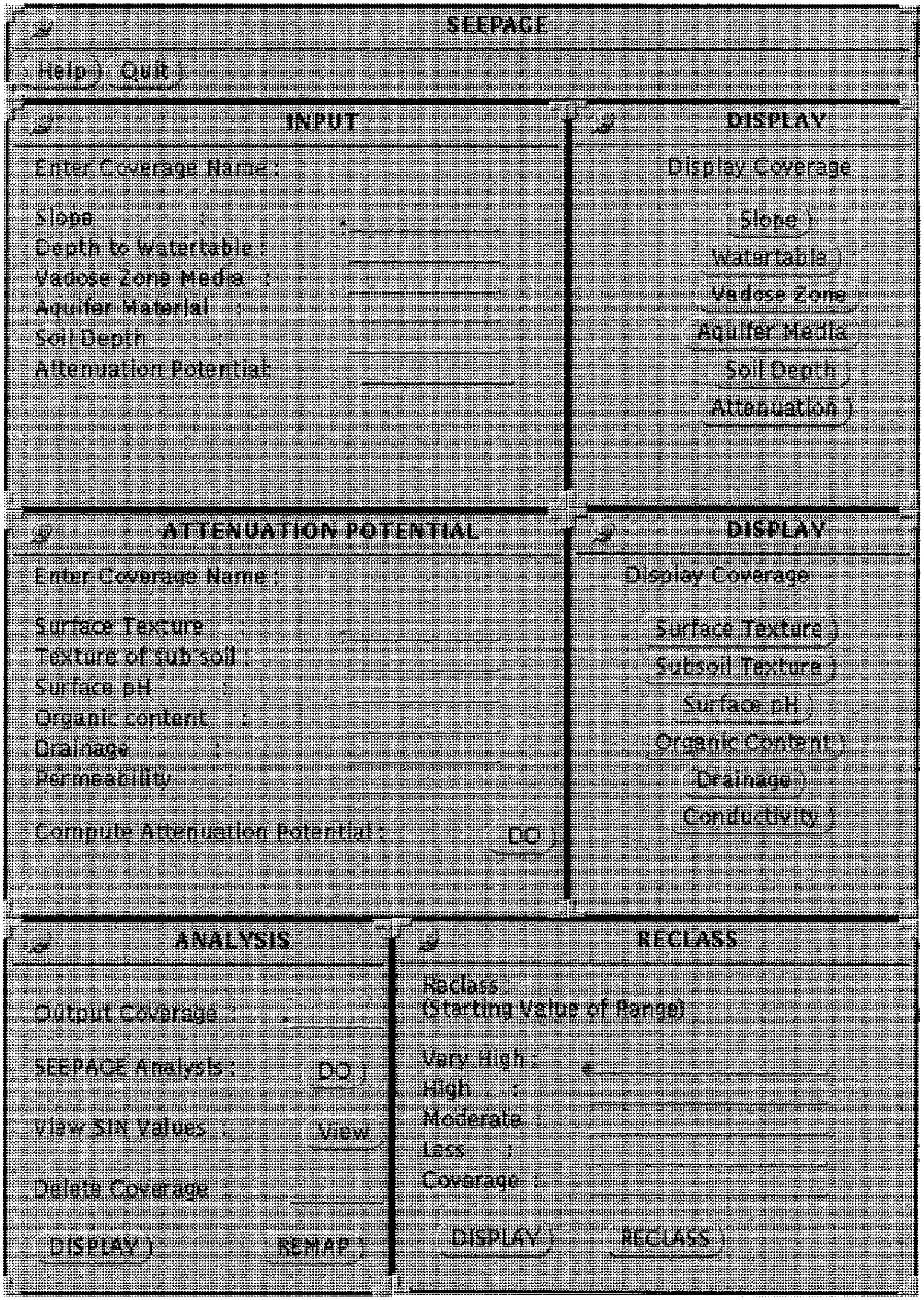


Figure 2. GUI for SEEPAGE analysis.

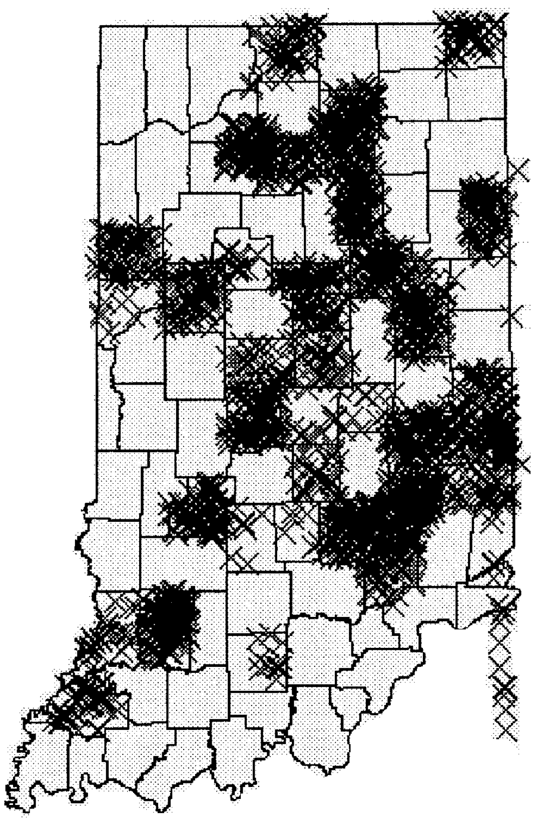


Figure 3. Sampling sites (wells) for water quality data.

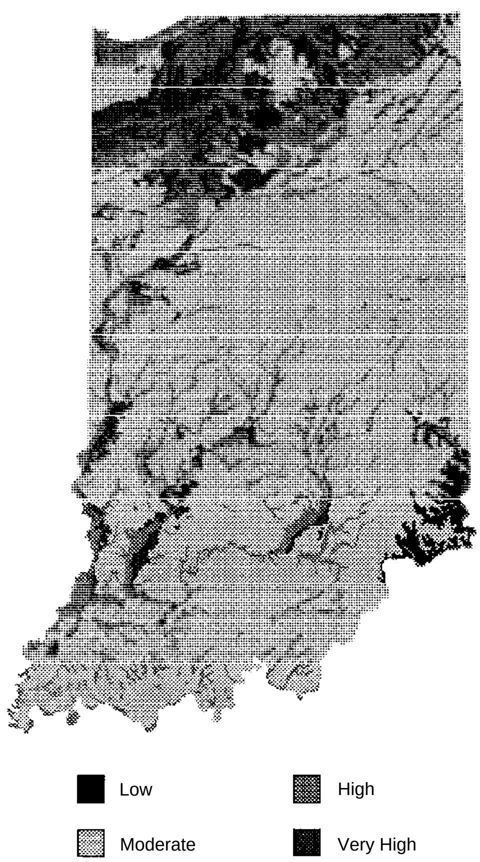


Figure 4. Ground-water vulnerability map using DRASTIC analysis.

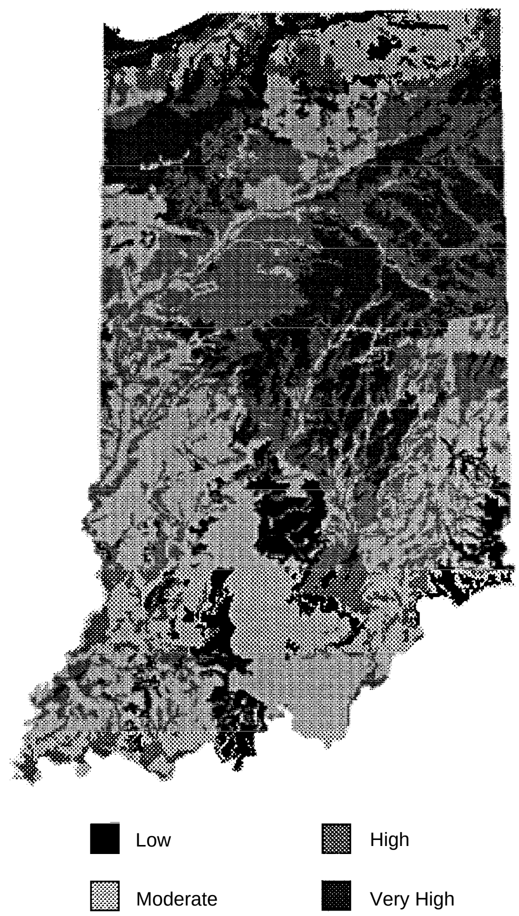


Figure 5. Ground-water vulnerability map using SEEPAGE analysis.

Table 2. Comparison of SEEPAGE Results With Observed Nitrate-Nitrogen Concentrations in Wells

	Nitrate-Nitrogen Levels		
Category	**0-5**	**5-10**	**> 10**
Low	7	2	8
Moderate	1,322	76	384
High	249	11	1
Very high	194	2	64

Table 3. Comparison of DRASTIC Results With Observed Nitrate-Nitrogen Concentrations in Wells

	Nitrate-Nitrogen Levels		
Category	**0-5**	**5-10**	**> 10**
Low	11	2	9
Moderate	541	41	290
High	720	39	213
Very high	500	9	65

developing most of the data layers required for the analyses. GIS is a useful tool for integrating ground-water quality models and facilitates testing the models for different scenarios. The GUI helps users easily conduct analyses and facilitates changing the weights for subjective hydrogeologic settings. DRASTIC and SEEPAGE approaches show great potential as screening tools for policy decision-making in ground-water management.

Summary

Ground-water pollution from agricultural management systems is of wide concern. Few models address ground-water vulnerability on a regional scale. The DRASTIC and SEEPAGE models are numerical ranking models that consider various hydrogeologic settings affecting the contamination of a region. The data required for these models are commonly available data, and the STATSGO database at 1:250,000-scale was used in this study. These models were integrated in the GIS environment of GRASS and ARC/INFO in the raster format. A menu-based GUI was developed in ARC/INFO for conducting the analyses. The vulnerability maps generated from DRASTIC and SEEPAGE analyses were compared. The statewide vulnerability maps also were compared with the well-water quality data to validate the accuracy of the models.

References

1. U.S. EPA. 1990. National pesticide survey: Phase I report. Washington, DC.
2. U.S. EPA. 1992. Another look: National survey of pesticides in drinking water wells, Phase II report. EPA/579/09-91/020. Washington, DC.
3. Department of Environmental Management, Groundwater Section. 1989. Indiana ground-water protection: A guide book (June).
4. U.S. EPA. 1987. DRASTIC: A standardized system for evaluating ground-water pollution potential using hydrogeologic settings. EPA/600/2-87/035. Washington, DC.
5. U.S. EPA. 1985. DRASTIC: A standardized system for evaluating ground-water pollution potential using hydrogeologic settings. EPA/600/2-85/0108. U.S. EPA, Robert S. Kerr Environmental Research Laboratory, Ada, OK.
6. Deichert, L.A., and J.M. Hamlet. 1992. Nonpoint ground-water pollution potential in Pennsylvania. ASAE Paper No. 922531. American Society of Agricultural Engineers International Winter Meeting, Nashville, TN.
7. Richert, S.E., S.E. Young, and C. Johnson. 1992. SEEPAGE: A GIS model for ground-water pollution potential. ASAE Paper No. 922592. American Society of Agricultural Engineers International Winter Meeting, Nashville, TN.
8. Engel, B.A., and D.D. Jones. 1992. Technique for developing ground-water vulnerability to nitrate maps for large areas: Research proposal 12-1992.
9. U.S. CERL. 1990. GRASS: Geographical Resources Analysis Supporting System user's manual. Champaign, IL: U.S. Army Construction Engineers Research Laboratory.
10. SCS. 1992. State Soils Geographic Database (STATSGO) user's guide. Soil Conservation Service Publication No. 1492.
11. ESRI. 1992. ARC/INFO command references: Arc command references. Redlands, CA: Environmental Systems Research Institute.
12. ESRI. 1992. ARC/INFO user's guide: AML user's guide. Redlands, CA: Environmental Systems Research Institute.
13. ESRI. 1992. ARC/INFO user's guide: Cell-based modeling with GRID. Redlands, CA: Environmental Systems Research Institute.

Using a Geographic Information System for Cost-Effective Reductions in Nonpoint Source Pollution: The Case of Conservation Buffers

Mark S. Landry, Darrell J. Bosch, and Tone M. Nordberg[1]

Abstract

Policy and management decisions to control agricultural nonpoint source (NPS) pollution can be based on different levels of information about the site-specific pollutant reductions and costs of control practices. While targeting practices based on site-specific criteria can reduce NPS pollution control costs, the best procedure for implementing targeting must be determined empirically. A geographic information system (GIS) can be used to evaluate and implement targeting procedures. The objective of this study is to determine the potential to reduce costs of controlling phosphorus (P) runoff by using a GIS to target the location of riparian buffer systems (RBS) based on full and partial information about site characteristics. Full information assumes knowledge of the potential P runoff reduction per dollar of RBS expenditure at each site. Partial information assumes knowledge of the animal numbers and animal density on each farm adjacent to a stream.

The study area is the Muddy Creek watershed in Rockingham County, Virginia, which is dominated by dairy and poultry operations. The watershed has been identified as a high priority watershed on the 1998 Virginia 303(d) Total Maximum Daily Loads list. A GIS was used to estimate the cost and potential P runoff reduction of each RBS and the animal density and animal numbers of each farm containing a potential site for locating a RBS.

Public and private costs of RBS installation and maintenance were included. Public costs include transaction costs and government cost share for RBS installation and maintenance. Private costs include the farmer's share of installation and maintenance costs and opportunity costs of land removed from production. Phosphorus runoff interception by the RBS was based on the Universal Soil Loss Equation (USLE), which was adapted to account for P runoff. The GIS was used to construct dairy, poultry, and beef farms which approximate the actual number of farms and livestock in the watershed.

---

[1] Graduate Research Assistant and Professor, Department of Agricultural and Applied Economics, and Graduate Research Assistant, Department of Biological Systems Engineering, Virginia Polytechnic Institute and State University, Blacksburg. The authors express appreciation to Conrad Heatwole for research assistance.

Uniform installation of RBS resulted in a 60% reduction of P runoff at an annualized cost of $45,332 or $26 per lb. Targeting with full information resulted in a 47% reduction in P runoff at an annualized cost of $24,535 or $18 per lb. Although P runoff was 377 lbs. more per year compared to uniform buffer installation, costs per lb. of P control were reduced by 30%. Total P runoff was reduced by 37% when RBS were targeted based on animal density. Costs of reductions were $21 per lb. Targeting based on animal numbers reduced P runoff by 30% at a cost of $25 per lb.

Targeting with the aid of a GIS can increase the cost effectiveness of RBS installation for controlling P runoff when full or partial information is available on RBS pollution reduction and costs at specific sites. An additional benefit of the GIS is the ability to visualize the locations of farms where RBS are located under alternative targeting criteria.

A GIS can be a useful tool for evaluating NPS pollution control options in watersheds. Further research is needed on ways to enhance the effectiveness of GIS in supporting NPS pollution control decisions.

Introduction

Over the past several decades, there has been increasing emphasis on the assessment and management of nonpoint source (NPS) pollution from agricultural lands. Agriculture accounts for more than 60% of the total NPS pollution of surface water bodies in the United States (U.S. EPA). The inherent variability of agricultural NPS pollution is a function of the heterogeneous landscape in a watershed, the complex interactions between ecological components, the total amount of pollutants, and the spatial distribution of land uses (National Research Council). The effectiveness of watershed policy and management decisions to control agricultural NPS pollution depend on these site-specific factors.

Potentially costs of NPS pollution control can be reduced by targeting sites with low-cost control opportunities. Targeting was examined in the early 1980's by the USDA for the purpose of allocating limited Soil Conservation Service funding and technical assistance to aid in conserving soil resources on highly erodible land in the U.S. (Nielson). A number of studies have identified the potential for reducing costs or increasing environmental benefits by targeting policies. Ribaudo (1986; 1989) showed the effects of considering off-site impacts of sediment erosion on the priority ranking of watersheds in the U.S. for erosion control efforts. Carpentier, Bosch and Batie found that targeting nitrogen (N) runoff reductions to farms with lowest costs reduced farm and taxpayer costs by 75 percent compared to the uniform policy. Studies by Fox, Umali and Dickinson; Dickinson, Rudra and Wall; and Setia and Magleby found that targeting resulted in large reductions in costs of controlling sediment delivery to streams.

The most effective method for targeting sites depends on watershed socioeconomic and physical characteristics and type of pollutant to be controlled. A systematic integration of site-specific, farm-level physical and economic models is necessary to capture the dynamic environmental processes and economic behavior of farmers in association with NPS pollution control policy (Antle and Just; Segerson and Wu). Empirical evaluation of targeting strategies can be done using a spatial decision support system (SDSS). Spatial decision support systems have been described as "flexible, integrated software for accessing, retrieving, and generating reports on database information plus simulation and decision models for conducting alternative testing, sensitivity analysis, and automated goal seeking." (Covington, et al., pp. 25-26) Spatial decision support systems can include programming models, simulation models, data manipulation systems as well as a user interface (Armstrong and Densham). The SDSS are distinguished by their ease of use and flexibility (Armstrong and Densham; Geoffrion). Examples of SDSS developed for evaluating NPS pollution control in agriculture include LOADSS (Negahban, et al.), MIKE SHE (Danish Hydraulic Institute), and WAMDSS (CARES).

The most critical component of an effective watershed SDSS is the use of a GIS (Fulcher, Prato, and Zhou). A GIS can be defined as the process of acquiring, processing, storing, managing, manipulating, and displaying spatial data that are associated with user-specified attributes (Aronoff). Interfacing a GIS in a decision support system is a logical approach to maximizing the capacity of a watershed assessment tool for targeting BMPs and/or other policy constraints (Bosch, Batie, and Carpentier).

In designing the GIS to evaluate alternative targeting strategies, a key question is what variables should be included. Detailed information about costs and effectiveness of control practices at each site might provide the most efficient targeting scheme. However some information may be very expensive or unavailable. Perhaps targeting based on partial information about costs and effectiveness of control practices can reduce pollution control costs.

The objective of this study is to determine the effectiveness of using information from a GIS for targeting riparian buffer strips (RBS) for the purpose of reducing total P in runoff from an agriculture-intensive watershed. The following three policies for installation of RBS's are compared in terms of private costs, public costs, and reductions in P runoff:

1. <u>Uniform design standard</u> Requiring all farms in the watershed to install a 60-foot wide RBS along streams and channels;

2. <u>Targeted design standard with full information</u> Targeting 60-foot wide RBS on specific farms throughout the watershed based on full information about buffer costs and pollution interception; and

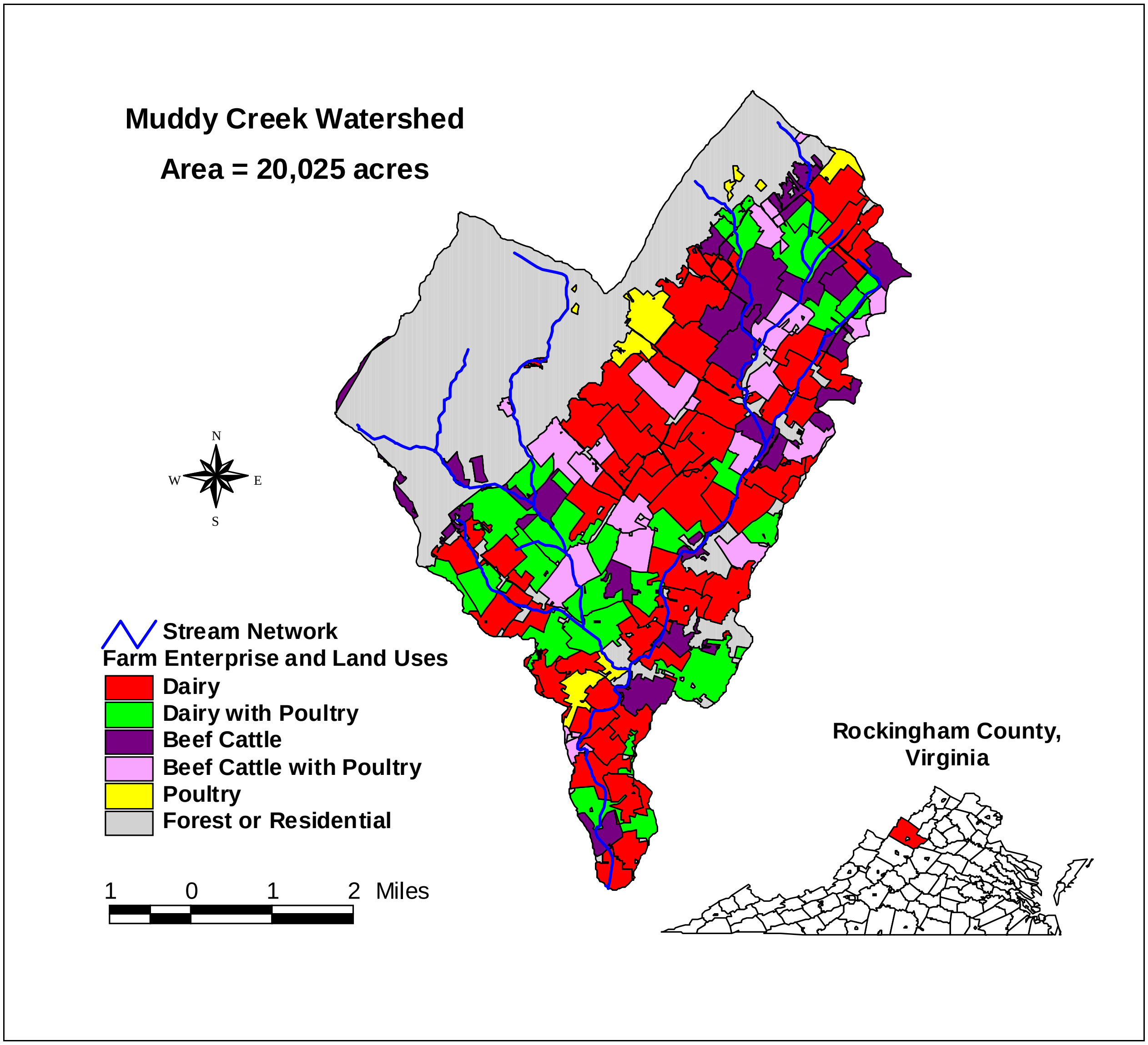
3. <u>Targeted design standard with partial information</u> Targeting 60-foot wide RBS on specific farms based on partial information about buffer costs and pollution interception.

Case Application: Targeting Riparian Buffer Strips in Muddy Creek Watershed, Virginia

Study Area: The Muddy Creek watershed is located in the Shenandoah Valley, approximately 10 miles to the west-northwest of Harrisonburg, Virginia in Rockingham County (see Figure 1). Muddy Creek flows south to its confluence with Dry River, which discharges into the North River then into the Shenandoah River and eventually into the Chesapeake Bay. The Muddy Creek watershed is 20,025 acres in size with approximately 100-110 farms. The watershed is listed on the 1998 Clean Water Act 303(d) Total Maximum Daily Load (TMDL) high priority list as an impaired water body in terms of fecal bacteria concentration (VA-DEQ, 1998).

About 65% of the farms are predominantly dairy, 30% are beef, and 5% are poultry. Dairy and beef operations also have poultry. Beef and dairy manure and poultry litter are typically spread on cropland and pastures. Most farms use a seven-year crop rotation, five consecutive years of corn/corn silage followed by two years of alfalfa. Corn is often rotated with rye, which is used either as a cover crop or harvested as silage. Pasture may be improved or native rangeland. Cattle are usually not restricted from surface water sources and most farms do not currently use field or riparian buffers to improve runoff water quality. While a nutrient management plan is required on farms with poultry (VA-DEQ, 1999), most farmers without poultry have not developed a nutrient management plan although a majority of this group have manure storage facilities.

Figure 1. Muddy Creek Watershed, Rockingham County, Virginia.



Riparian Buffer Strips: Riparian buffer strips (RBS) are areas that are managed to reduce the amount of pollution suspended in runoff. A RBS is installed downslope of a point source (e.g., feedlot) or nonpoint source (e.g., cropland) area on which pollution is generated. A RBS reduces pollution exiting the site via runoff by promoting infiltration of water and water soluble constituents, increasing the adsorption of pollutants onto the vegetation and soil, and increasing the absorption of nutrients into plants (Dillaha et al.; Landry and Thurow). The appeal of the RBS is that it can be easier and more economical to install and maintain than physical structures (Pritchard et al.), can be aesthetically pleasing and can provide a source of income as pasture, hay, timber, or wildlife refuge (Purvis et al.; Shabman and Smith).

The USDA launched the National Buffer Initiative in 1997 emphasizing the importance and effectiveness of riparian buffers for achieving water quality protection objectives (USDA-NRCS). The national goal is to install two million miles (up to seven million acres) of conservation buffers by the year 2002. The Buffer Initiative is sponsored by numerous government agencies and agricultural producers working together to promote the use of RBS and their eligibility in incentives programs such as cost-sharing through the Conservation Reserve Program (CRP).

Policy Scenarios

Uniform design standard: This policy assumes a 60-foot wide RBS is required along both sides of all streams and channels. Farms required to install RBS's are assumed to take advantage of all available government incentives. Two government assistance programs are considered 1) the Conservation Reserve Program (CRP) (10 year contract, 50% cost-share towards RBS establishment, and a $70/acre annual rental payment) and 2) the Virginia Water Quality Improvement Act (VWQIA) (25% cost-share towards BMP installation up to $17,500).

Targeted Design Standard - Full Information: Full information is defined as knowing the P runoff reduction per dollar of buffer expenditure. Phosphorus runoff reduction was calculated using the GIS and was based on soil type, infiltration rate, plant or crop nutrient uptake, manure and/or litter application rate, slope, weather characteristics, and RBS pollutant interception coefficients.

Targeted design standard - partial information: This scenario is used to determine whether limited, readily available information can be used to designate RBS installation sites in a manner that approaches the total costs and water quality improvement given full information. We targeted 60-foot wide RBS's under two partial information policy scenarios based on farms having the greatest number of cattle (dairy or beef) and by farms with the highest manure application rates per acre based on animal density (animal units per acre of land).

Empirical Procedures

Riparian Buffer Cost Specification: Private and public costs associated with RBS installation and maintenance are illustrated in Table 1. Private costs equal the annualized sum of opportunity costs plus establishment and maintenance costs minus government cost-share (the RBS is

Table 1. Costs and returns of riparian buffer strip.[a]

Item	Information or Cost Description		
Life of project (years)	10		
Vegetation composition	Fescue grass and Ladino Clover		
Farmer Costs (present value)	***Corn Silage ($/acre)***	***Hay Land ($/acre)***	***Pasture[b] ($/acre)***
Returns	0	0	0
Land opportunity costs[c]	1,105	772	116
Seed	61	61	15
Lime	54	54	14
Fertilizer	26	26	7
Labor and Equipment[d]	166	166	166
(a) Total	1,412	1,079	318
Cost-share			
CRP 50% cost-share	154	154	101
CRP rental payment ($70/ac)	541	541	541
VWQIA cost-share	77	77	51
(b) Total	772	772	693
Net farmer costs (a-b)	640	307	-375
Annualized farmer cost	83	40	-49
Public Costs (present value)	***($/acre)***		
CRP 50% cost-share	154		
CRP rental payment ($70/acre)	541		
VWQIA cost-share	77		
Net public costs	772		
Annualized public cost	100		
Transaction Costs[e]	***($/farm)***		
Information	28		
Contracting	1		
Enforcement	77		
Annualized transaction cost	106		

[a] Internal rate of return equals 5%.
[b] Fescue grass - Ladino Clover pasture with 25% additional seed, lime, and fertilizer added in buffer.
[c] We assumed a relatively high yield for corn silage and hay/pasture based on the likelihood of moderately sloping, high productivity soil in the bottomland adjacent to the stream. The land opportunity cost for corn silage was equal to the expected yield (16.5 tons/ac) x $25/ton – total variable costs ($254), for hay land was equal to the expected yield (3.3 tons/ac) x $70/ton - total variable costs ($131), and for pasture was equal to the expected yield (1.7 tons/ac) x $70/ton - total variable cost ($101) (Virginia Cooperative Extension).
[d] Taken from Leeds, Forster, and Brown.
[e] Values taken from Carpentier and Bosch estimates for strip cropping. Costs are annualized based on a 5% real interest rate.

unharvested and produces no revenue). Land opportunity costs represent the income given up when land is taken out of production and used for buffers. These costs varied based on soil productivity and land use. Cropland was used in corn silage production; hay land was used for orchard grass or timothy grass; and pasture was used for Fescue grass-Ladino clover pasture. For example, the income given up from corn silage minus variable costs of production is $254 per year (see footnote to Table 1). The present value of corn silage income given up over a 10-

year period is $1,105. After adding other buffer installation costs, the present value of total buffer costs is $1,412. Federal and state cost-share and rental payments are subtracted leaving a net present value of $640, or $83 on an annualized basis.

Public costs include federal and state cost share and rental payments as well as transaction costs. The present value of public cost share plus rental payments for a 10-year period is $772 or $100 on an annualized basis. Transaction costs include the costs of informing farmers about buffer requirements, contracting with farmers to install buffers, and enforcement to insure that buffers are installed and maintained. Annualized transaction costs are $106 per farm.

NPS/Water Quality Model: The Universal Soil Loss Equation (USLE) was used to determine the average annual soil loss. Average annual manure applied to the field was used to find the available N and P in the soil. Enrichment ratios were used to calculate the N and P concentrations in the runoff that reached surface waters.

ArcView was used to do the GIS modeling, with a USGS DEM, soils map, landuse and streams as input themes. The sinkholes of the DEM were filled and a slope grid and flow direction grid for the watershed were derived from the filled DEM. The slope of each field polygon from the landuse theme was found by summarizing the landuse by each field polygon with respect to the slope grid, hence an average slope for each field was created. The length factor was found based on the assumption that each polygon was a square, the perimeter was calculated and then divided by four equal sides. This simplification resulted in some fields exceeding the max length of 100 m proposed for the use of the USLE (Novotny and Olem). Then the slope-length factor of the USLE equation was calculated from equation (1) below (Novotny and Olem), where L equals length in meters and S equals slope as a decimal fraction.

$$LS = L^{1/2}(0.0138 + 0.00974 \times S + 0.00138 \times S^2) \tag{1}$$

The cropping management factor, C, and the erosion-control practice factor, P, were selected for each landuse from Novotny and Olem, Gupta, and Schwab et al. For the rainfall energy factor, R, a value of 175 was taken from Novotny and Olem. The soil erodibility factor, K, was found for each soil in the Rockingham County Soil Survey (USDA-SCS), an average value for K was used if different values were listed for different soil depths.

The amount of the eroded soil calculated with the USLE that would actually leave each field and reach surface waters was based on a sediment delivery ratio of 30% (Novotny and Olem). It was assumed that the only N and P available in the soil came from the manure applied to the fields. The concentration of sediment-bound P in the soil (fraction of a gram P/ gram soil) was based on the assumption of a 1.3 kg/m^3 bulk density for the soil and a 5 cm depth of interaction assumed for the surface applied manure. The sediment bound P loss was then calculated with an enrichment ratio of 1.6 (Edwards et al.). The available N in manure was assumed to be in organic form, with 83.3% TKN (Edwards et al.). The concentration of N in sediment was calculated in the same manner as was done for P. The enrichment ratio of N was assumed to be equal to the organic enrichment ratio, equation (2) (Novotny and Olem). The organic matter of the soils was assumed to be an average of 2.5 % for the watershed.

$$ER_{or} = \frac{0.3}{(\%\ organic\ matter)} + 1.08 \tag{2}$$

The runoff from each field was not routed through the watershed and attenuation of nutrients in stream was not considered. The reductions of N and P loadings by the buffers were assumed to be constant and not to change by field location with respect to the streams. This is a simplification, since the buffers would have a greater impact for fields located closer to the streams. Possible reductions in the long term effectiveness of buffers were not considered in the study.

The pollutant delivery calculations assumed all manure and litter were land applied on the producing farm and commercial fertilizer was applied at a rate that was equal to 70% of crop removal. Commercial fertilizer application rates were set at levels needed to make total nutrient applications (manure plus commercial fertilizer) result in excess nutrient availability at levels approaching mean values observed by Bosch et al. These commercial fertilizer rates may overstate excess nutrients given recent nutrient management planning efforts in the area.

Data in Table 2 were used to calculate N and P application rates assuming 70% of the N and 100% of the P are available to plants based on annual land application of all manure and litter produced. For example, an acre of corn silage on high productivity soil removes 217 lbs. of N and 82 lbs. of P annually. The N and P application rates are based on the amount of manure and litter produced as a function of animal density and the nutrient content. Manure and litter are applied to meet 30% of the vegetation requirements and excess N and P are available for export via runoff and leaching.

Table 2. Dry weight manure and litter nutrient content.[a]

	Unit	Percent Moisture	Dry Weight	Organic N[b]	P
		--- % ---	--- lbs. ---	-------- lbs./unit --------	
Dairy/Beef Cattle Manure	1,000 gal	94.5	457	12.3	5.3
Poultry Litter	ton	66.2	1,323	49.0	27.8

[a] Source: Pease, Parsons, and Kenyon.
[b] Inorganic N content of manure was not considered. Inorganic N is subject to heavy volatilization losses (25 to 75%) when manure is unincorporated as was assumed in this study.

Table 3 indicates the average expected NPS pollution reductions for the 60-foot wide RBS based on typical soil characteristics (silty loam), vegetation composition (Fescue grass and Ladino clover), and management activities (no harvesting) for the buffers in the watershed.

Table 3. Pollutant trapping reductions (%) used for a 60-foot riparian buffer strip in Muddy Creek Watershed, Rockingham County, Virginia.

	Percent Pollutant Reductions (% slope)[a]		
Constituent	1 – 5%	5 – 10%	10 – 20%
Total Suspended Solids[b] (TSS)	80	65	50
Total Phosphorus (TP)	60	50	40
Total Kjeldahl Nitrogen (TKN)	55	40	25

[a] Average percent slope of the original field or installed riparian buffer adjacent to the stream.
[b] Constituent used to measure sediment delivery.

GIS: A GIS was built for the study area with ArcView (ESRI) using data from federal and state agencies (see Table 4). Soil productivity and potential corn, hay, and pasture yields were based on the Virginia Agronomic Land Use Evaluation System (VALUES) (Simpson et al.).

Table 4. Data included in the GIS for the Muddy Creek watershed.

Data Type	Data Source
Digital elevation map (DEM)	U.S. Geological Survey (USGS)
Road network	U.S. Census Bureau – Tiger Files
Soil data	U.S. Department of Agriculture – Natural Resources Conservation Service (NRCS)
Stream network	NRCS, stream network delineator extension developed for ArcView GIS
Land use, land use proportions	Virginia Department of Conservation and Recreation (DCR)
Streamflow data, nutrient monitoring data	United States Geological Service (USGS)
Weather data	Virginia State Climatology Office
Watershed size	Virginia Hydrologic Unit Atlas (NRCS)

Farm-scale information was necessary to determine animal numbers, manure application rates, and numbers of acres eligible for buffers on each farm. Farm boundaries were unavailable, therefore, a field clustering routine developed in the GIS was used to allocate 71 diaries and 36 beef cattle farms based on criteria representative of the area (Schroeder; VA-DEQ, 1999; Parsons; US-Dept. of Commerce). Fields were assigned to farms based on the Thiessen method according to proximity and greatest area within polygons formed around the nearest point representing central farm nodes.

Farm enterprises consisting of three sizes of dairies (60, 100, and 150 cows) and three sizes of beef cattle operations (40, 70, and 150 cows) were distributed across the watershed according to farm size. For example, the smallest dairies were assigned 60 cows, medium sized dairies were assigned 100 cows, and the largest dairies were assigned 150 cows. The locations of broiler and turkey houses were included in the land use GIS data coverage, thus farms which included these structures were assigned 25,000 broilers per house or 16,000 turkeys per house, accordingly. The aggregate number of animals on the farms generated using the GIS approached the total number of animals in the watershed: 6,533 dairy cows; 3,134 beef cows; 500,000 broilers; and 350,000 turkeys (Schroeder; VA-DEQ, 1999).

Targeting RBS based on full information assumes that NPS pollution reduction is maximized for a given sum of costs (public and private). The full information targeted buffers scenario limited the amount of stream miles that could be buffered to one-half the distance considered in the uniform design policy. The farms adjacent to streams and channels were ranked according to those having fields which resulted in the highest P reductions per dollar of cost. Fields ranked in the top half of all fields in pounds of P reduction per dollar of cost were targeted for buffer installation. Costs include cash and non-cash costs, where the cash costs equal the government outlays for transaction costs and cost-share and farmers' outlays for buffer establishment. The non-cash costs are the farmers' opportunity costs for the land idled by the buffer.

Under the partial information scenario based on animal number, the farms were ranked according to the number of cows per farm. Farms with equal numbers of cows were ranked based on numbers of broilers or turkeys. The second partial information policy scenario required farm size to determine animal density in acres. With this added information, we calculated N and P application rates/acre based on the nutrient content of the manure and/or litter produced and ranked the farms accordingly. For both partial information scenarios all other data were considered unknown, including potential yields on lands used for buffers, RBS costs, public costs, and the NPS pollutant interception rates of the buffers. The RBS's were allocated across the watershed until 50% of the potential stream miles were buffered.

Results

Analysis using the GIS indicated that 69 farms were subject to the RBS policies based on stream proximity. Baseline annual pollutant loadings from fields adjacent to streams and channels for sediment; phosphorus; and nitrogen were 23,059 tons; 2,900 lbs.; and 1,653 lbs.; respectively. Approximately 279 acres of buffers were installed under the uniform design standard. Nineteen of the 32 stream-miles were considered eligible for RBS installation. The remainder were ineligible based on adjacent land uses such as forestland, farmsteads, and residences which do not support animal production. Table 5 illustrates variation in policy effectiveness across the four policy scenarios under analysis.

Table 5. Costs and effectiveness of riparian buffer strips in alternative policy scenarios.

Policy Scenario	Phosphorus Reduction[a] (lbs.)	Private Costs ($)	Public Costs ($)	Average Cost ($/lb.)	Number of Farms
Uniform	1,740 (60%)	10,098	35,234	26	69
Targeted: Full Information	1,363 (47%)	4,724	19,811	18	55
Targeted: Animal Number	870 (30%)	4,419	17,279	25	36
Targeted: Animal Density	1,073 (37%)	5,014	17,799	21	30

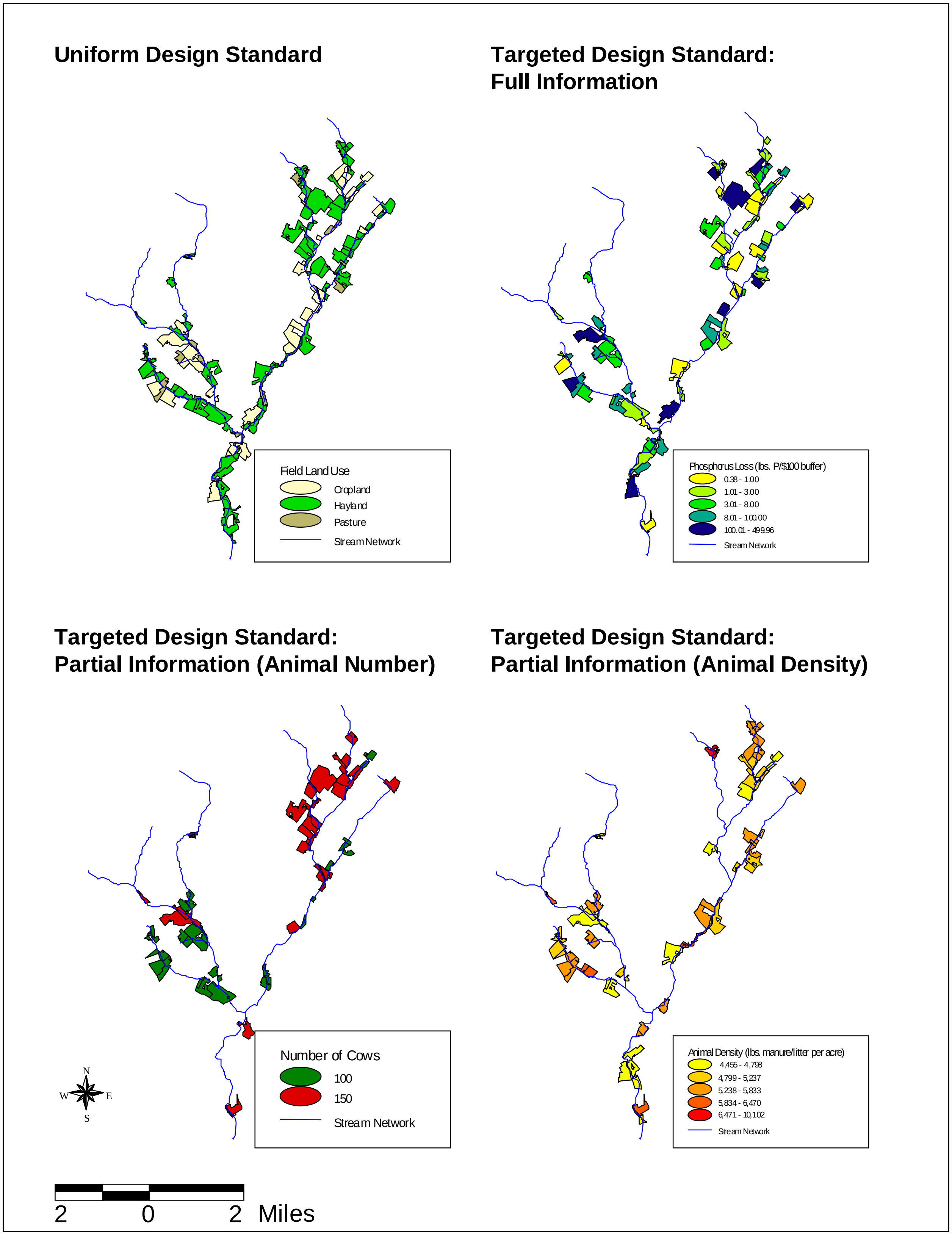
[a] Number in parentheses represent the percent reduction compared to the baseline. The baseline had 2,900 lbs. of phosphorus runoff.

Figure 2 compares the various NPS pollution control policies based on the targeting criteria used. The fields affected by each policy vary. The spatial impacts, which are displayed using the GIS, illustrate the distribution and relative effectiveness of each policy scenario.

The uniform design standard achieved a 60% reduction in P runoff (1,740 lbs.) at a total (public and private) annualized cost of $45,332 or $26 per lb. The levels of NPS pollutant reductions under the uniform design are consistent with the expected reductions for the 60-foot wide RBS. Requiring all eligible stream miles to be buffered implies that all 69 farms along the streams are affected by the policy.

Full information RBS targeting on 50% of the eligible stream miles affected 55 farms, which had at least one field bordering the stream requiring a buffer. Total (public and private) annualized costs were $24,535. Total P runoff reduction was 47% of the baseline (1,363 lbs.) or $18 per lb. The average cost per pound of P interception was reduced by 30% although total P interception was 377 lbs. less under this scenario.

Figure 2. Fields Affected by Targeting Criteria Associated with Each Policy Scenario.



Targeting based on animal numbers and animal density removed less P than full information. Total P interception was 37% (1,073 lbs.) for animal density targeting and 30% (870 lbs.) for animal number targeting. Total costs were $22,813 for animal density targeting ($21 per lb.) and $21,698 for animal number targeting ($25 per lb.). The number of farms affected was 30 for animal numbers targeting and 36 for animal density targeting. Partial information resulted in less efficient targeting than full information in terms of average cost per pound of P reduction, but was more efficient than a uniform standard. The partial information targeting scenarios suggest that average pollution control costs can be reduced by targeting control measures using basic information about farms' pollution potential.

Summary and Conclusions

The public is increasingly concerned to reduce NPS pollution. Given public ambivalence about increasing taxes or regulations, it is important that pollution control costs be lowered. Cost reduction will require that effective policies be designed both in terms of the type of policy instrument used and how the instrument is targeted. Policy selection and targeting can be enhanced by considering fixed spatial attributes of farms that affect their pollution potential and their costs of controlling pollution. A GIS can be used to assist in identifying and targeting farms with high potential for reducing pollution at low cost.

A case analysis of Muddy Creek watershed in Virginia demonstrates the potential use of GIS to assist in determining the costs and pollution reduction from installing RBS's within the watershed. The case study results suggest that using a GIS to target RBS's based on full or partial information about pollution control costs can be an effective tool for reducing pollution control costs. Targeting pollution control measures might be politically difficult due to farmers' concerns with fairness. One way to overcome this concern would be for governments to compensate farmers for all buffer installation and opportunity costs.

Because of their user accessibility, GIS and SDSS are likely to increase in importance for evaluating and implementing water quality protection policies. For example, they can be used to assess plans for achieving TMDLs in watersheds with impaired waters (U.S. EPA). Continued development of SDSS capabilities will be needed to expand their usefulness. The appropriate spatial resolution of the watershed which balances design and implementation costs with accuracy may change depending on watershed physical and economic conditions and the objectives of the users. How can SDSS's be designed to operate at multiple scales of resolution from very large watersheds (e.g. Chesapeake Bay drainage) to small local watersheds?

The physical and economic variables that are most influential in determining NPS pollution responses to changing policy and economic conditions should be identified. While these variables are likely to change depending on the type of problem being addressed (e.g. pesticide pollution vs. nutrient pollution) and the type of agricultural watershed, further research could identify some general relationships. For example, what is the relative importance of farmers' knowledge, attitudes, and goals compared to the fixed physical characteristics of their land in determining potential pollution? While farmers' attitudes and knowledge are often unknown to the researcher and subject to rapid change, a better understanding of their importance would help analysts better characterize the uncertainty from the SDSS.

Output from complex decision support models are subject to errors due to inherent randomness in the processes being simulated, errors and gaps in input data, and uncertainty about the relationships being modeled (Suter and Barnthouse). How can uncertainty be incorporated into the SDSS and communicated to users? Research is needed to extend the capability of SDSS to handle changes over time. This capability would facilitate tracing firms in space and time as they respond to economic trends, policy changes, or other exogenous shocks. For example, what are the watershed implications of continued expansion of intensive livestock industries? Similarly, SDSS need better capability to model physical changes (e.g. pollution) over time. Further development of dynamic capabilities within SDSS would allow them to be used for inter-temporal optimization decisions.

More study is needed of the most likely users of SDSS for NPS pollution control. What are the decisions they make and what types of information do they need to help make those decisions? Better knowledge of user needs can guide data collection strategies, enhance SDSS design, and show where further knowledge of underlying socioeconomic behavior and physical processes is needed.

References

Antle, J.M. and R.E. Just. "Conceptual and Empirical Foundations for Agricultural-Environmental Policy Analysis." *Journal of Environmental Quality* 21 (July-September 1992):307-316.

Armstrong, M.P. and P.J. Densham. "Database Organization Strategies for Spatial Decision Support Systems." *International Journal of Geographic Information Systems* 4(1990):3-20.

Aronoff, S. *Geographic Information Systems: A Management Perspective*. Ottawa, Ontario: WDL Publications, 1995.

Bosch, D. J., J. W. Pease, S. S. Batie, and V. O. Shanholtz. “Crop Selection, Tillage Practices, and Chemical and Nutrient Applications in Two Regions of the Chesapeake Bay Watershed.” Bulletin VPI-VWRRC-BULL 176. Virginia Water Resources Research Center, Virginia Polytechnic Institute and State University, 1992.

Bosch, D.J., S.S. Batie, and L.C. Carpentier. "The Value of Information for Targeting Water Quality Protection Programs within Watersheds." In *Economics of Nutrient Management Policy: Proceedings of a Regional Workshop*. (P.E. Norris and L..L.. Danielson (eds.)), SRIEG-10 No. 32, Southern Development Center No. 180, Mississippi State University, Mississippi State, Mississippi. March 1994. pp. 93-106.

Carpentier, C. L., and D. J. Bosch. "Regulatory Alternatives to Reduce Nitrogen Runoff: Case Study of Lower Susquehanna Dairy Farms." Department of Agricultural and Applied Economics, Virginia Polytechnic Institute and State University. Invited paper, Conference on Flexible Incentives for the Adoption of Environmental Technologies in Agriculture, June 8-10, 1997. Gainesville, Florida.

Carpentier, C.L., D.J. Bosch, and S.S. Batie. "Using Spatial Information to Reduce Costs of Controlling Agricultural Nonpoint Source Pollution." *Agricultural and Resource Economics Review* 20 (April 1998): 72-84.

Center for Agricultural, Resource, and Environmental Systems (CARES). "Projects: Watershed Management Decision Support System (WAMADSS)." http://www.cares.missouri.edu/cares/projects/WM.html . University of Missouri - Columbia, June 1997.

Covington, W.W., D.B. Wood, D.L. Young, D.P. Dykstra, and L.D. Garrett. "TEAMS: A Decision Support System for Multiresource Management." *Journal of Forestry* 86 (1988): 25-33.

Danish Hydraulic Institute. "MIKE SHE: An Integrated Hydroligical Modeling System." http://www.dhi.dk/mikeshe/index.htm . 1999.

Dickinson, W. T., R. P. Rudra, and G. J. Wall. “Targeting Remedial Measures to Control Nonpoint Source Pollution.” *Water Resources Bulletin* 26, no. 3(1990): 499-507.

Dillaha, T.A., R.B. Reneau, S. Mostaghimi, and D. Lee. "Vegetative Filter Strips for Agricultural Nonpoint Source Pollution Control." *Transactions of the ASAE* 32, no. 2 (1989):513-519.

Edwards, D.R., C.T. Haan, A.N. Sharpley, J.F. Murdoch, T.C. Daniel, P.A. Moore Jr. "Application of Simplified Phosphorus Transport Models to Pasture Fields in Northwest Arkansas." *Transactions of ASAE* 39, no. 2(1996): 489-496.

Environmental Systems Research Institute, Inc. (ESRI). *Using Arcview GIS*. Redlands, California: ESRI, 1996.

Fox, G., G. Umali, and T. Dickinson. "An Economic Analysis of Targeting Soil Conservation Measures with Respect to Off-Site Water Quality." *Canadian Journal of Agricultural Economics* 43 (1995): 105-118.

Fulcher, C., T. Prato, and Y. Zhou. "Watershed Management Decision Support System." Unpublished Working Paper, Center of Agricultural, Resource and Environmental Systems. University of Missouri, Columbia, Missouri, 1998.

Geoffrion, A. M. “Can OR/MS Evolve Fast Enough.” *Interfaces* 13(1983): 10.

Gupta, R. S. *Hydrology and Hydraulics Systems*. New Jersey: Prentice Hall, 1989.

Landry, M.S. and T.L. Thurow. *Function and Design of Vegetation Filter Strips: An Annotated Bibliography*. Bulletin No. 97-1, Texas State Soil and Water Conservation Board, Temple, Texas, 1997.

Leeds, R., D.L. Foster, and L.C. Brown. "Ohio State University Extension Factsheet - Vegetative Filter Strips: Economics." http://ohio-state.edu/aex-fact/468.html . Agricultural Engineering Department, The Ohio State University, Columbus, Ohio, 1994.

National Research Council. *Soil and Water Quality: An Agenda For Agriculture*. Edited by J. Overton. Washington D.C.: National Academy Press, 1993.

Negahban, B., C. Fonyo, W. G. Boggess, J. W. Jones, K. L. Campbell, G. Kiker, E. Flaig, and H. Lal. “LOADSS: A GIS-based Decision Support System for Regional Environmental Planning.” *Ecological Engineering* 5(1995): 391-404.

Nielson, J. "Conservation Targeting: Success or Failure?" *Journal of Soil and Water Conservation* 41 (March-April 1986): 70-76.

Novotny, V.,H. Olem. *Water Quality: Prevention Identification and Management of Diffuse Pollution.* New York, New York: Van Nostrand Reinhold, 1994.

Parsons, R. L. “Financial Costs and Economic Tradeoffs of Alternative Manure Management Policies on Dairy and Dairy/Poultry Farms in Rockingham County, Virginia.” Ph.D., thesis. Virginia Polytechnic Institute and State University, 1995.

Pease, J., R. Parsons, and D. Kenyon. “Economic and Environmental Impacts of Nutrient Loss Reductions on Dairy and Dairy/Poultry Farms.” REAP Bulletin 448-231. Virginia Polytechnic Institute and State University, 1998.

Pritchard, T.W., J.G. Lee, and B.A. Engel. Reducing Agricultural Sediment: An Economic Analysis of Filter Strips Versus Micro-Targeting. *Water Science Technology* 28, nos. 3-5(1993): 561-568.

Purvis, A., J.P. Hoehn, V.L. Sorenson, and F.J. Pierce. "Farmers' Response to a Filter Strip Program: Results from a Contingent Valuation Survey." *Journal of Soil and Water Conservation* 44, no. 5(1989): 501-503.

Ribaudo, M. O. “Consideration of Offsite Impacts in Targeting Soil Conservation Programs.” *Land Economics* 62, no. 4(1986): 402-411.

Ribaudo, M. O. “Targeting the Conservation Reserve Control Program to Maximize Water Quality Benefits.” *Land Economics* 65, no. 4(1989): 320-332.

Schwab, G.O., D.D. Fangmeier, W.J. Elliot, R.K. Frevert. *Soil and Water Conservation Engineering*. New York, New York: John Wiley & Sons, Inc., 1993.

Setia, P., and R. Magleby. "Economic Efficiency of Targeting Agricultural Nonpoint Pollution Controls." Economic Research Service, U.S. D. A. Selected Paper, Northeastern Agricultural and Resource Economics Association Annual Meeting, June 20-22 1988. Orono, Maine.

Schroeder, Phil. 1999. Personal communication. County Agent, Virginia Cooperative Extension Service, Harrisonburg, Virginia.

Shabman, L. and C. Smith. *Assessing Landowner Level Costs for Riparian Forest Buffer System Adoption in Virginia's Chesapeake Bay Watershed*. Report prepared for the U.S. Forest Service, Annapolis, Maryland, 1998. 76 pp.

Simpson, T.W., S.J. Donohue, G.W. Hawkins, M.M. Monnett, and J.C. Baker. "*The Development and Implementation of the Virginia Agronomic Land Use Evaluation System (VALUES)*." Department of Crop and Soil Environmental Sciences, Virginia Tech, Blacksburg, Virginia, 1992.

Suter, G. W., and L. Barnthouse. "Assessment Concepts." In *Ecological Risk Assessment.* G. W. Suter, ed. pp. 21-47. Boca Raton, Florida: Lewis Publishers, 1993.

U.S. Dept. of Agriculture - Natural Resources Conservation Service (USDA-NRCS). "Buffer Strips: Common Sense Conservation." http://www.nhq.usda.gov/CCS/Buffers.html . 1997.

U.S. Dept. of Agriculture - Soil Conservation Service (USDA-SCS). *Soil Survey of Rockingham County*. Washington, D.C., 1982.

U.S. Dept. of Commerce. *1992 Census of Agriculture, Part 46, Virginia State and County Data*. Bureau of the Census, Washington, D.C., 1994.

U.S. Environmental Protection Agency (U.S. EPA). "Guidance for Water Quality-Based Decisions: The TMDL Process." http://www.epa.gov/OWOW/tmdl/*decisions/dec1c.html*. April, 1991. Assessment and Watershed Protection Division, U.S. Environmental Protection Agency. Accessed: February 24 1999.

Virginia Department of Environmental Quality (VA-DEQ). *Virginia 303(d) Total Maximum Daily Load Priority List Report*. Richmond, Virginia, May 1998.

Virginia Department of Environmental Quality (VA-DEQ). *Draft Fecal Coliform TMDL Development for Muddy Creek, Virginia*. Richmond, Virginia, May 1999.

Wu, J. and K. Segerson. “On the Use of Aggregate Data to Evaluate Groundwater Protection Policies.” *Water Resources Bulletin* 31(1995): 1773-1780.

Merging Transportation and Environmental Planning Using Geographic Information Systems (GIS)

Elizabeth Lanzer, Washington State Department of Transportation

Abstract

The Washington State Department of Transportation annually plans over 400 highway construction projects ranging from repaving efforts to developing entirely new roads. The challenge brought to the Environmental Affairs Office at WSDOT is how to effectively and consistently make use of existing Geographic Information System (GIS) information in the transportation planning process. This paper describes the development of two GIS efforts that examine environmental data in relationship to highway construction projects.

The two examples in this paper describe how WSDOT is actively incorporating GIS as a tool for environmental assessments. The first example, Environmental GIS Workbench (Workbench), is an application that was developed to provide new GIS users with easy access to existing environmental data. The application shows environmental and transportation data on screen for visual examination by the users. The second example, Environmental Screening (Screening), is a GIS product that offers additional analysis. Highway projects are evaluated for proximity to environmental information (e.g. wetlands, rivers, etc.) and then ranked high, medium, low or none according to potential environmental impacts.

One of the challenges of both efforts is the complexity and volume of environmental data. Data sources range from private, local, state and federal sources. Many of which differed in format, scale, projection, data quality, documentation and maintenance requirements. The GIS products both utilized the same data, however the level of analysis done by the GIS products were different.

Introduction

The Washington State Department of Transportation (WSDOT) is responsible for efficiently building, maintaining, operating and promoting safe and coordinated transportation systems to serve the public[1]. These transportation systems include highways, ferries, railroads, airports and selected river systems that link people, freight and locations across the state. The highway system is the largest of these with over 7,000 miles of state roads. WSDOT annually plans over

400 highway construction projects. For the purposes of this paper, references to “projects” should be interpreted as highway construction projects ranging from re-paving to developing entirely new roadways. Unfortunately, environmental requirements are not always optimally considered. However, recent mandates, such as the listing of many salmonid species under the Endangered Species Act, have created a paradigm shift within WSDOT that emphasizes consideration of environmental factors and created the opportunity for these GIS efforts.

The challenge brought to the Environmental Affairs Office at WSDOT is how to effectively and consistently make use of existing Geographic Information System (GIS) information in the transportation planning process. Traditional environmental reports are generated through the labor intensive process of obtaining maps, data or reports from government entities and then reviewing that information on a project by project basis. This process is time consuming and inefficient when trying to meet deadlines. Limitations of comparing paper maps against each other, repetition of data collection and information inaccessibility hinder the environmental review of highway projects. GIS, on the other hand, integrates software, hardware, data and people allowing more efficient environmental assessments.

The two examples in this paper describe how WSDOT is actively incorporating GIS as a tool for environmental assessments. The first example, Environmental GIS Workbench (Workbench), is an application that was developed to provide users with easy access to environmental data, but provides relatively little analysis. The application shows environmental and transportation data on screen for visual examination by the users. The second example, Environmental Screening (Screening), offers additional analysis. Highway projects are evaluated for proximity to environmental information (e.g. wetlands, rivers, etc.) and then ranked high, medium, low or none according to potential environmental impact.

Background

The Washington State Department of Transportation’s (WSDOT) agency standard GIS software is ArcView 3.1 (NT) and ARC/INFO 7.2.1 (NT and UNIX) from ESRI (Redlands, CA). The Department has approximately ten ARC/INFO licenses and over 150 ArcView licenses statewide. Training and support is provided for GIS users through the WSDOT GIS Support Team which includes GIS professionals from throughout the department.

Agency available GIS data (vector and imagery) is stored on a central GIS server. Vector data is stored in geographic projection, decimal degrees, spheroid GRS80 and datum NAD83. Imagery data is generally stored as delivered by the source agency (e.g. stateplane south, NAD27, feet). All GIS data that is stored on the central server has Federal Geographic Data Committee (FGDC) standard metadata reports in both plain text and HTML (web page) format.

GIS data is administered centrally, but managed by a distributed group of business specialists called data stewards. A typical data steward is a GIS or technology specialist housed within a specific functional office within the agency. Working with their colleagues, they determine data requirements and standards to meet the needs of their function within the agency. As data meeting these needs are gathered or created, the data steward posts the data with the agency GIS data administrator and defines rules for accessing and updating the data they are responsible for.

GIS data that is available on the server is either generated internally in WSDOT or obtained from external data sources. WSDOT creates and maintains a dynamically segmented state highway model and a vast amount of associated tabular data. These tabular records are mapped within the GIS linear highway model as events. Event data are placed on the GIS system using milepost information. For example, construction projects can be located on the GIS coverage by knowing the state route number, beginning and ending mileposts. Milepost information is collected for environmental data such as fish passage barriers and deer kill locations. Partnerships have also been formed with the Washington Department of Health to create GIS coverages on drinking water wells.

A significant amount of WSDOT's environmental data is obtained from public and private sources. The information is formatted for use on WSDOT's systems and placed on the central GIS server when appropriate (See Attachment A). In some cases, data sharing agreements with other agencies are necessary to obtain the information, particularly when the data is considered sensitive (e.g. certain wildlife habitats). Data collected from outside sources are updated on an established basis (in some cases, every six months and in other cases every two to three years) so that reasonably current data is available to the WSDOT GIS user community at WSDOT. This data forms the foundation upon which WSDOT GIS products and analysis are based.

Environmental GIS Workbench Application (Workbench)

The Environmental GIS Workbench Application (Workbench) is a desktop GIS application that provides easy access to existing environmental data and environmental data processing tools needed in the development of environmental reviews for highway projects planned for construction within two years. This application includes a convenient method for accessing data by subject and location in order to analyze the data, and a means of capturing the results for documentation. The application is programmed using Avenue programming language and incorporated into an ArcView extension. The intention is to reduce the amount of training and the learning curve presently needed for new users to access the existing data thereby improving the efficiency and the quality of the review process. [2] The Environmental Review Summary form provides environmental documentation regarding air quality, fish habitat, wetlands, water quality and other environmental factors (See Attachment B). Data for the application are imported from various Washington State Department of Transportation application systems, or obtained from other agencies where appropriate.

Prior to developing the Workbench application, the environmental review forms took approximately two to eight hours to complete. Most of that time was spent gathering the necessary maps and documents. Much of that information was already available on the WSDOT GIS server, but was not easy to access or query. The Workbench utilizes GIS technology to create layers of maps that display available environmental data (See Attachment A).

A steering committee was convened to guide the application development process. This group defined the budget and general expectations of the product as well as provided the conduit for management and policy support. A smaller workgroup was also formed to define the specific data and tool requirements. The Environmental Review Summary form that is completed for all projects provided the foundation for determining the questions the Workbench would be expected to answer. The workgroup identified which existing data was appropriate to incorporate and created a list describing new data to be obtained. There were several datasets that were identified as critical but not currently available at WSDOT. The top seven were prioritized, obtained and added to the application.

Most of the environmental data included in the Workbench have statewide coverage, but there are occasional county specific coverages. County specific coverages are included in the

Workbench, but are clearly labeled as such to prevent confusion. Whenever possible, the coverages for the application are saved as shapefiles to improve application speed. The scale of the included data varies from 1:24,000 to 1:500,000. All data that is accessible through the Workbench has accompanying Federal Geographic Data Committee (FGDC) compliant metadata created by WSDOT in HTML format. Usually, the data source supplies metadata files. These files are hotlinked to the WSDOT HTML metadata reports. The WSDOT HTML metadata reports are accessible through a tool in the Workbench. By providing the users with metadata, the environmental specialists utilizing the Workbench are able to make decisions about whether a particular data set is appropriate to answer their questions.

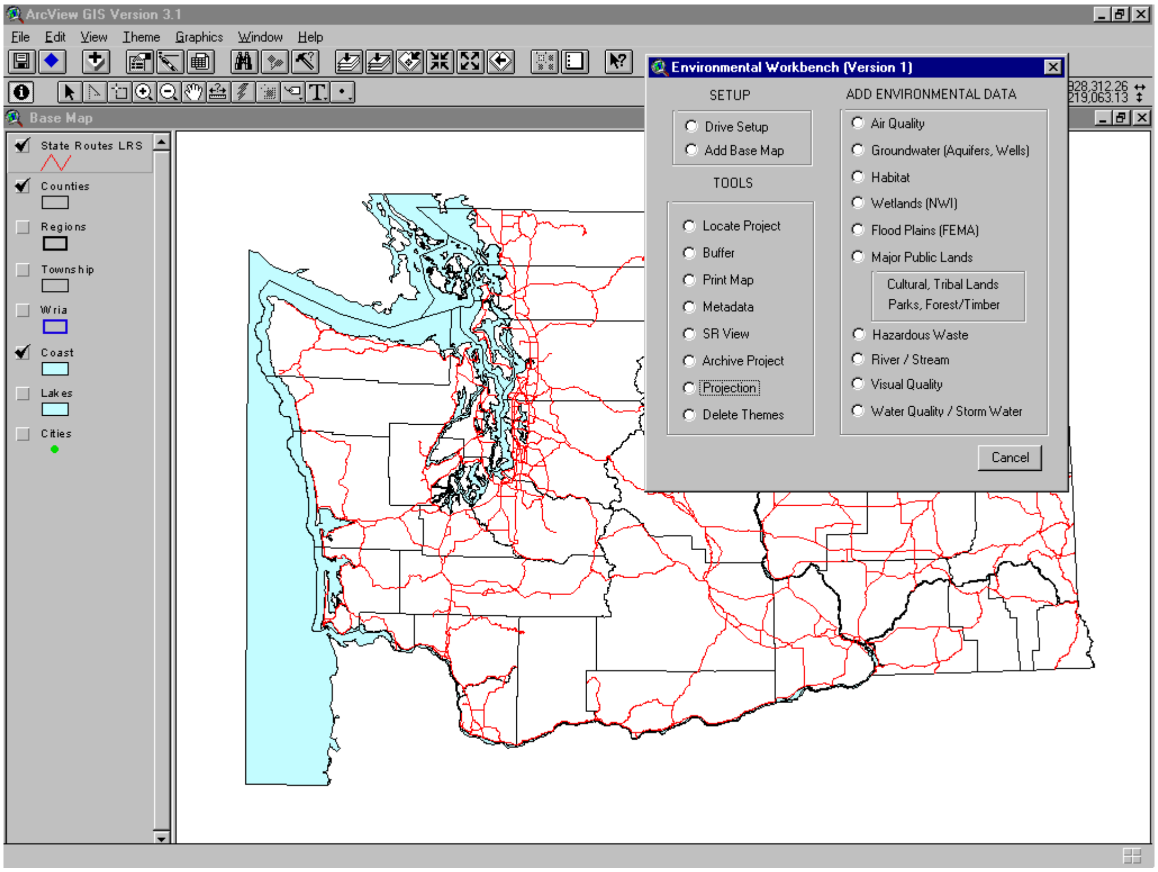
Training users is an essential part of success for the Workbench. Training is structured as a half day session on ArcView and the Workbench extension. The second half of the day is devoted to discussions on data quality, what assumptions can or cannot be made and how to understand and use the metadata. GIS data usually looks different than the maps and reports the specialists are accustomed to seeing, so this also provides an opportunity to introduce them to how to use GIS data and how to query it.

Maintenance of the Workbench is important so that users trust the application to have reasonably current data. Data that is already programmed into the Workbench are updated seamlessly by using the same coverage/shapefile naming conventions. Additional programming is required when adding entirely new data to the application. The Workbench is scheduled for deployment in July of 1999. Beta testing is complete and so far the feedback is positive. A feedback database is being maintained to consolidate comments from users and use these comments to make future improvements in the Workbench. Future versions are anticipated with the first comprehensive review of the application tentatively scheduled for Spring 2000.

Environmental Screening (Screening)

Environmental Screening was developed to evaluate highway construction projects for proximity to environmental information (e.g. wetlands, rivers, etc.) and then ranked high, medium, low or none according to potential environmental impact. This environmental flagging will be included in the long term planning and budgeting for the project.

Figure 1. Example Interface of the Washington Department of Transportation's Environmental GIS Workbench Application.



Environmental and GIS specialists analyzed highway projects and environmental data using ArcView software. Eleven environmental categories were originally chosen for investigation: air, cultural resources, flooding and wetlands, geologic hazards, habitat, hazardous materials, noise, recreation, environmental justice, visual impacts, and water quality. Upon further investigation, insufficient data or time existed to complete the visual impacts, geologic hazards, recreation, environmental justice and noise sections.

The air and cultural resources categories were visually analyzed. All relevant data was mapped. Subject area specialists visually examined the paper map for intersections between the proposed highway projects and the relevant data. Broad conclusions were drawn as to the extent of potential impacts.

The remaining categories: flooding and wetlands, habitat, hazardous materials and water quality were analyzed using a query based method. All available relevant data was mapped. Subject area specialists determined an appropriate buffer distance from the highway, within which, an intersection with an environmental data feature would constitute a “hit”. The environmental data was weighted (See Attachment C) and if the highway project received a “hit” on the environmental data, then that segment was assigned a weight, according to the severity of potential environmental impact. After all queries for a particular subject area were completed, all weights were totaled for each highway project, and a final rating of high, medium, low, or none, was assigned. Paper maps showing the ratings (See Figure 2) and the proportion of highway projects with environmental impacts were produced (See Figure 3). Again, broad conclusions were drawn as to the extent of potential impacts.

Figure 2. Example of Results for the Washington Department of Transportation’s Environmental Screening Project.

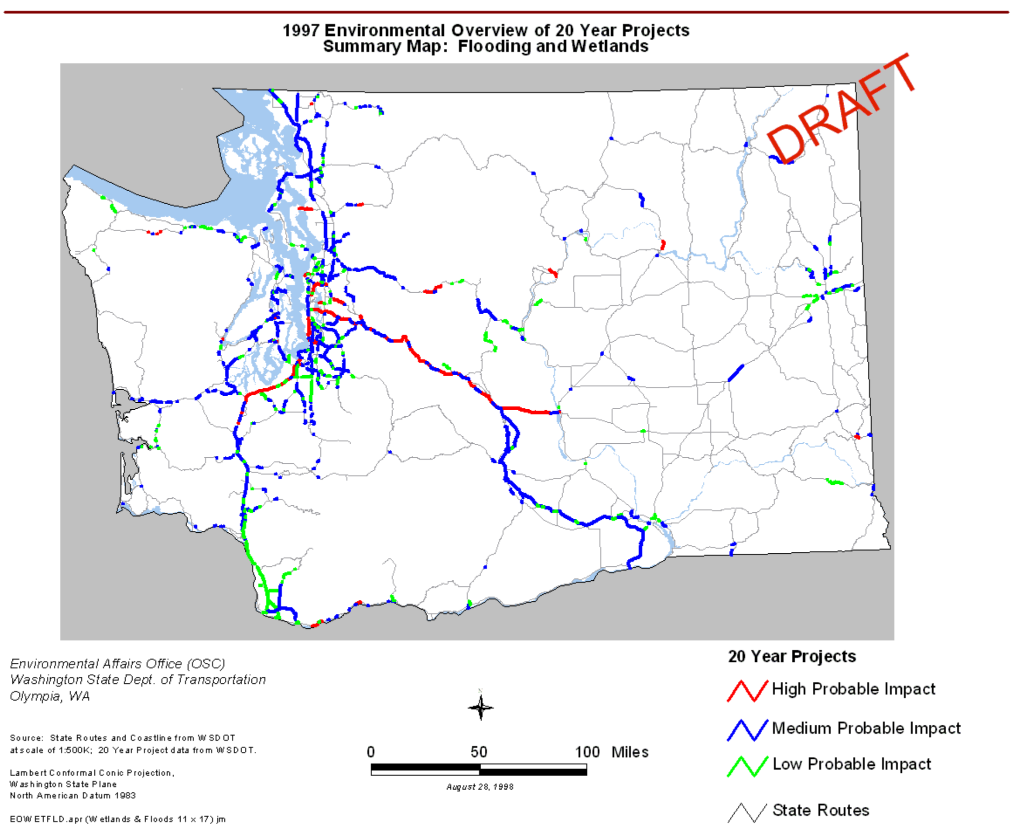
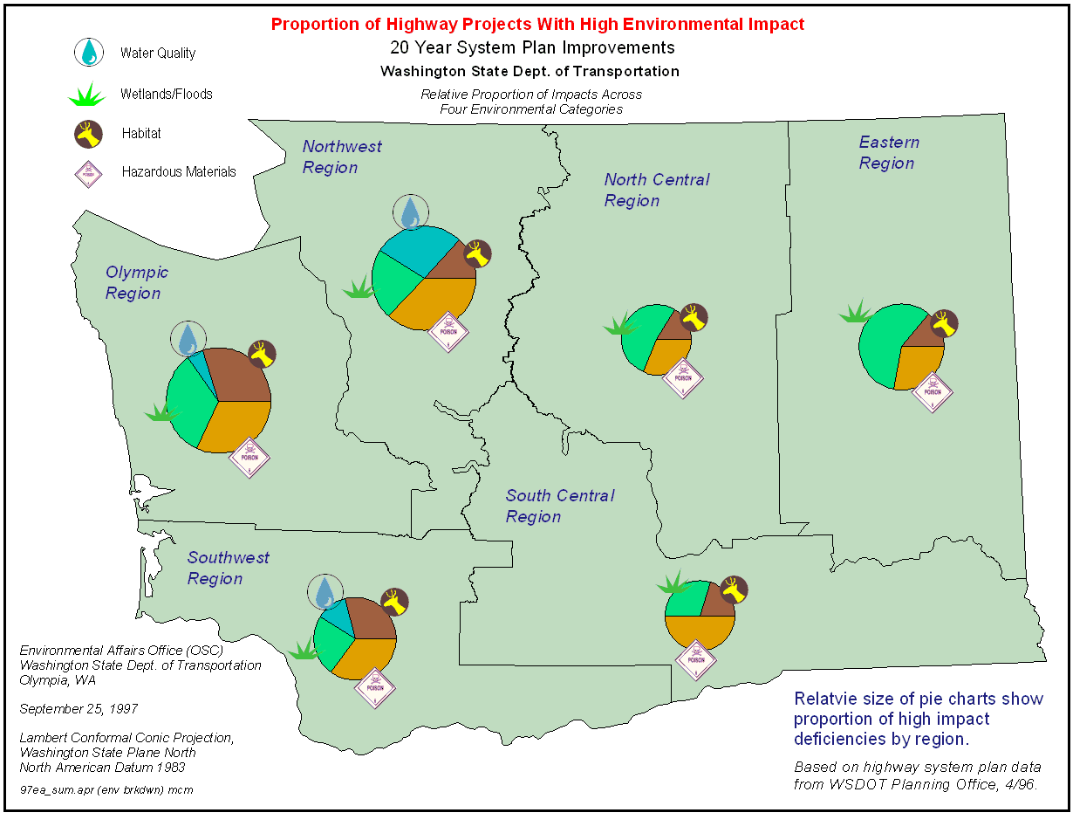


Figure 3. Proportion of Highway Projects with High Environmental Impacts by Washington Department of Transportation Regions.



Next Steps for Environmental Screening

This effort was originally conducted in 1996 - 1997 as a demonstration of how GIS could incorporate environmental information earlier in the transportation planning process. Limitations of these analyses included a lack of appropriate environmental data, that the results did not reflect multiple “hits” per project and that weighting and rating systems varied among the subject areas making cumulative analyses difficult. Since 1997 there have been improvements to the amount and quality of appropriate environmental data and enhancements to the software (ArcView) used for conducting the queries.

In May 1999, this project was re-evaluated by the Environmental Affairs Office and the Planning Office at WSDOT. It was decided to continue developing the Screening and improve the previous effort to support the year 2000 Washington Transportation Plan.

Several process changes are being made in the current Environmental Screening effort. More people are involved in the product/output definition phase with the potential for it to be programmed as a GIS application. The data weighting and ranking systems are also being re-evaluated. The current effort is expected to be applied to all modes of transportation in determining effects, not just highway projects. The long term vision of this activity is to use environmental data to help identify transportation alternatives (e.g. using transit or rail as a solution in urban areas with air quality problems instead of highway improvements which may increase air pollution problems). There are some limitations in the GIS coverages for other modes besides highways, but that is one of the issues to be addressed as work progresses. One of the goals for the current effort is to set up a clearly documented and repeatable process that allows new transportation projects to be promptly evaluated for potential environmental effects.

Discussion

The Workbench and Screening tools are fundamentally different in terms of how the data is applied. In the Workbench, the data is simply retrieved from the server and displayed. Most of the queries applied to the data are to aid in the visual display. In the Screening, the data is retrieved, criteria is applied to the data then an evaluation is made on the highway project. The Workbench relies on the user to complete the environmental analysis while the Screening depends on environmental criteria programmed into the application.

The Workbench requires a greater degree of confidence in the data even though there is less pre-programmed analysis than in the Screening. The purpose of the Workbench is to analyze specific sections of the highways. This may require zooming into one to two mile sections of a roadway and visually examining the available data. The better the information going into the Workbench, the greater the confidence in the results coming back out. The Screening, on the other hand, is used to raise warnings about potential future impacts. Since the projects are still 20 years from construction, data with less resolution is acceptable for the task. As a particular highway project progresses towards construction, better quality data will be needed to evaluate potential environmental impacts.

The famous “Garbage in-Garbage out” theory applies here. Using the GIS data steward concept at WSDOT is one data quality control point. Information is not stored on WSDOT’s central GIS server unless it meets certain minimum standards and has accompanying metadata. Business specific GIS applications expand the audience using GIS data. Complete and accurate metadata is essential for these products to be successful. Discerning users are skeptical of the data being shown on-screen so that easy to use documentation is important. Achieving user buy-in so that the system or application will be used is also a critical success factor.

Another advantage to the Washington Department of Transportation's central GIS server concept, is that both efforts draw upon the same information sets. Some efficiencies come from the ability to develop data that will be useful in both analyses. At this point, data that is more detailed than 1:24,000 is difficult to obtain consistently on a statewide basis. Local governments generally have more detailed environmental data but how to incorporate that information into WSDOT and associated GIS products is still a work in progress. As improvements are made, they can be integrated with associated GIS applications and products.

One of the challenges of both projects is the complexity and volume of available environmental information. Data sources range from private, local, state and federal entities. Even when data is available in GIS format, many differ in format, scale, projection, data quality, documentation and maintenance requirements. Some critical environmental data is simply not available in GIS format and will be expensive to convert. Additional data problems occur when several state agencies maintain similar but different GIS layers. For example, at least four different state agencies currently maintain a 1:100,000 stream layer each with different attributes important for environmental analysis. Coordinating and determining appropriate uses of each stream layer takes considerable time. The Washington Geographic Information Council (WAGIC) is addressing these data coordination issues at the state level. WSDOT, state, local, federal and tribal entities are all represented in this forum. During the 1999 State Legislative session, several bills and money appropriations were passed that will help state agencies tackle data coordination problems. Improvements made at the state level can eventually be integrated into GIS products at WSDOT.

Conclusions/ Summary

These two GIS products are anticipated to provide a solid foundation for WSDOT to develop GIS as a standard tool in the agency for conducting environmental analyses. Developing GIS solutions for environmental problems are inherently multi-disciplinary efforts that require teamwork and partnerships. WSDOT will be looking at partnering with other public and private entities to combine money, technology and ideas for incorporating environmental analysis into transportation planning in more efficient and effective ways.

Acknowledgments

The author wishes to acknowledge several organizations and individuals whose technical support helped make these GIS projects and this paper possible.

Marci Carte and Patty Lynch, Washington State Department of Transportation

Washington State Department of Transportation: Environmental Affairs Office, GeoServices, Program Management, Regional Environmental Offices

Key Data Providers:
Washington State Department of Health
Washington State Department of Ecology
Washington State Department of Natural Resources
Washington State Department of Fish and Wildlife

Attachment A : List of Data used in the Workbench or Screening GIS Products.
(The complete WSDOT Geospatial Catalog is at http://www.wsdot.wa.gov/gis/GeoDataCatalog).

Data Set Title	Source Scale	Originator	Included in Workbench?	Included in Screening?
General Reference Data:				
Transportation				
State Routes (mainlines) LRS	500K	WSDOT	Y	Y
Roadside Landscape Classifications	500K	WSDOT	Y	
Political and Admin. Boundaries				
County Boundaries, statewide	500K	WSDOT	Y	Y
City Limits of Washington State	24K	WSDOT		Y
Major Cities (points)	500K	WSDOT	Y	
DOT Regions	500K	WSDOT	Y	Y
Major Public Lands by WA Dept. of	100K	WADNR	Y	Y
Natural Resources:				
City Parks	100K	WADNR	Y	Y
County Parks	100K	WADNR	Y	Y
DNR Managed Lands	100K	WADNR	Y	
Experimental Forests	100K	WADNR	Y	
Federal/State Fish Hatcheries	100K	WADNR	Y	
Federal/State Medical Facilities	100K	WADNR	Y	Y
Military/Tribal Reservations (see also Tribal Lands, Military Lands)	100K	WADNR	Y	Y
Monuments	100K	WADNR	Y	Y
Municipal Watersheds	100K	WADNR	Y	
National Forests	100K	WADNR	Y	Y
National Historic Parks	100K	WADNR	Y	Y
National Parks	100K	WADNR	Y	Y
Public School Lands	100K	WADNR	Y	
Recreation	100K	WADNR	Y	
State Parks	100K	WADNR	Y	
Wilderness Areas	100K	WADNR	Y	
Wildlife Refuges	100K	WADNR	Y	Y
Geographic Reference				
Townships	500K	WADNR	Y	
TIGER-Census Bureau base maps	100K	USCB		Y
ENVIRONMENTAL DATA:				
Air Quality:				
Carbon Monoxide Non-Attainment Areas	500K	WADOE	Y	Y
Ozone Non-Attainment Areas	500K	WADOE	Y	Y
Particulates Non-Attainment Areas	500K	WADOE	Y	Y

Data Set Title	Source Scale	Originator	Included in Workbench?	Included in Screening?
Fish and Wildlife:				
Chinook Evolutionarily Significant Units	250K	NMFS	Y	
Chum Evolutionarily Significant Units	250K	NMFS	Y	
Coastal Cutthroat Trout Evolutionarily Significant Units	250K	NMFS	Y	
Coho Evolutionarily Significant Units	250K	NMFS	Y	
Endangered Species Act Listing Status for Salmon and Trout	none	WADOE	Y	
Fish (Salmonid) Passage Barriers	500K	WDFW	Y	
Marbled Murrelet Point Sitings	24K	WDFW	Y	
Marbled Murrelet Critical Habitat	100K	USFW	Y	Y
Seabird Colonies	none	WDFW		Y
Sockeye Evolutionarily Significant Units	250K	NMFS	Y	
Spotted Owl Critical Habitat	100K	USFW	Y	Y
Spotted Owl Special Emphasis Areas	none	WADNR	Y	Y
Steelhead Evolutionarily Significant Units	250K	NMFS	Y	
Sensitive Environmental Data		WDFW		
Priority Habitat and Species	24K	WDFW	Y	
Spotted Owl Nests	24K	WDFW	Y	Y
Wildlife Heritage Data	24K	WDFW	Y	Y
Groundwater and Wells:				
Critical Aquifer Recharge Areas, Clallam County	24K	WSDOT/ Clallam Co.	Y	Y
Critical Aquifer Recharge Areas, Clark County	24K	WSDOT/ Clark Co.	Y	Y
Critical Aquifer Recharge Areas, Franklin County	250K	WSDOT/ Franklin Co.	Y	Y
Critical Aquifer Recharge Areas, Island County	none	WSDOT/ Island Co.	Y	Y
Critical Aquifer Recharge Areas, King County	100K	WSDOT/ King Co.	Y	Y
Critical Aquifer Recharge Areas, Kitsap County	24K	WSDOT/ Kitsap Co.	Y	Y
Critical Aquifer Recharge Areas, Lincoln County	500K	WSDOT/ Lincoln Co.	Y	Y
Critical Aquifer Recharge Areas, Pend Oreille County	24K	WSDOT/ Pend Oreille Co.	Y	Y
Critical Aquifer Recharge Areas, Pierce County	none	WSDOT/ Pierce Co.	Y	Y

Data Set Title	Source Scale	Originator	Included in Workbench?	Included in Screening?
Critical Aquifer Recharge Areas, Thurston County	24K	WSDOT/ Thurston Co.	Y	Y
Critical Aquifer Recharge Areas, Whatcom County	500K	WSDOT/ Whatcom Co.	Y	Y
Sole Source Aquifers:	100K	EPA	Y	Y
Wellhead Protection Zones--statewide	24K	WSDOT	Y	
Wells, Group A, WA State	24K	WSDOT	Y	Y
Wells, Group B, WA State	24K	WSDOT	Y	
Hazardous Materials:				
CERCLIS--Comprehensive Environment Response Compensation and Liability Information System (Superfund sites)	none	WADOE	Y	Y
RCRA Facilities--generators, transporters, treaters, storers, and disposers of hazardous waste	none	EPA	Y	Y
Toxic Cleanup Program sites--confirmed and suspected hazardous materials sites	none	WADOE	Y	Y
Hydrography:				
Coastlines, Puget Sound and Columbia River (Major Shorelines)	500K	WSDOT	Y	Y
Floodzones (by county)--statewide	24K	FEMA	Y	Y
Major Lakes of Washington	500K	WSDOT	Y	Y
Major Rivers of Washington	100K	WADOE	Y	Y
National Wetlands Inventory (1,500+ individual quadrangle files)	24K	USFW	Y	
Salmonid Habitat Locations	100K	WDFW	Y	
WA County Series, Hydrography	24K	WSDOT	Y	
Washington Resource Inventory System (Watersheds)	100K	WADOE	Y	Y
Plants:				
Rare and Native Plants-WA state	24K	WADNR	Y	Y
Water Quality				
1994 303d listed water bodies--does not meet state water quality standards	100K	WADOE	Y	Y
National Pollutant Discharge Elimination System Permit Areas:				
National Pollutant Discharge Elimination System Sites: sites holding permit to discharge wastewater to surface water	none	WADOE	Y	Y

Data Set Title	Source Scale	Originator	Included in Workbench?	Included in Screening?
Clark County	500K	WSDOT	Y	Y
Island/Snohomish Co.	500K	WSDOT	Y	Y
South Puget Sound	500K	WSDOT	Y	Y
Spokane County	500K	WSDOT	Y	Y
Stormwater Outfalls along State Routes:	24K	WSDOT	Y	Y

DEFINITIONS—

Data Set Title:	Title or commonly used name of the data set.
File Location:	The path by which the data set is located on WSDOT's GIS servers.
Originator:	The creator or source of the data set.
Source scale:	The scale denominator of the data set's source material. For example, 24K indicates data derived from sources at 1:24,000 scale.

ABBREVIATIONS—

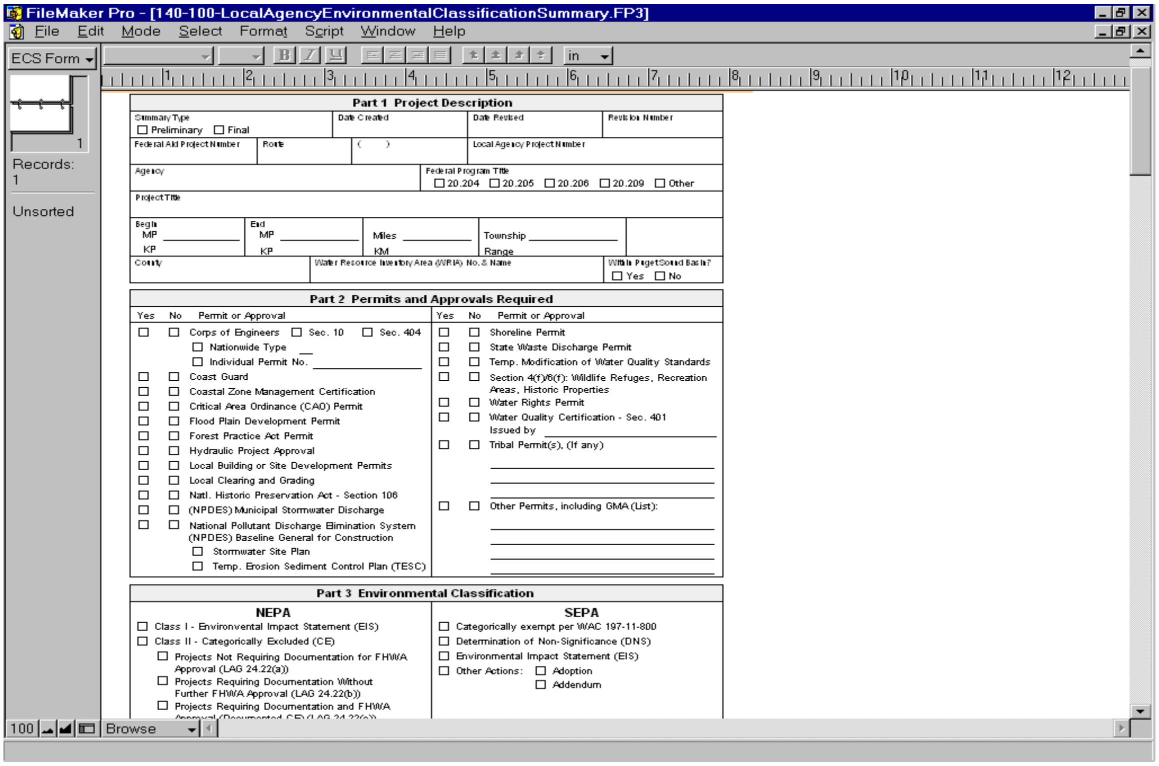
24K	1:24,000 scale—1 map inch represents 2,000 feet
100K	1:100,000 scale—1 map inch represents 1.58 miles
250K	1:250,000 scale—1 map inch represents 3.95 miles
500K	1:500,000 scale—1 map inch represents 7.89 miles
LRS	Linear Reference System
no scale	data is of mixed scales or scale not applicable

ORIGINATORS—

DCTED	Washington Department of Community, Trade and Economic Development
ESD	Washington Employment Security Department
FEMA	Federal Emergency Management Agency
NMFS	National Marine Fisheries Service
PSRC	Puget Sound Regional Council
USEPA	United States Environmental Protection Agency
USFW	United States Fish and Wildlife Service
USGS	United States Geological Survey
USCB	United States Census Bureau

IACOR	Interagency Committee on Outdoor Recreation
WADNR	Washington Department of Natural Resources
WADOH	Washington State Department of Health
WADOE	Washington Department of Ecology
WDFW	Washington Department of Fish and Wildlife
WSDOT	Washington State Department of Transportation

Attachment B: Environmental Classification Summary (Environmental Review Summary).



FileMaker Pro - [140-100-LocalAgencyEnvironmentalClassificationSummary.FP3]

File Edit Mode Select Format Script Window Help

ECS Form

Records: 1

Unsorted

Part 1 Project Description

Summary Type ☐ Preliminary ☐ Final | Date Created | Date Revised | Revision Number

Federal Aid Project Number | Route | ( ) | Local Agency Project Number

Agency | Federal Program Title ☐ 20.204 ☐ 20.205 ☐ 20.206 ☐ 20.209 ☐ Other

Project Title

Begin MP KP | End MP KP | Miles KM | Township Range

County | Water Resource Inventory Area (WRIA) No. & Name | Within Puget Sound Basin? ☐ Yes ☐ No

Part 2 Permits and Approvals Required

Yes No Permit or Approval

- ☐ ☐ Corps of Engineers ☐ Sec. 10 ☐ Sec. 404
  - ☐ Nationwide Type
  - ☐ Individual Permit No.
- ☐ ☐ Coast Guard
- ☐ ☐ Coastal Zone Management Certification
- ☐ ☐ Critical Area Ordinance (CAO) Permit
- ☐ ☐ Flood Plain Development Permit
- ☐ ☐ Forest Practice Act Permit
- ☐ ☐ Hydraulic Project Approval
- ☐ ☐ Local Building or Site Development Permits
- ☐ ☐ Local Clearing and Grading
- ☐ ☐ Natl. Historic Preservation Act - Section 106
- ☐ ☐ (NPDES) Municipal Stormwater Discharge
- ☐ ☐ National Pollutant Discharge Elimination System (NPDES) Baseline General for Construction
  - ☐ Stormwater Site Plan
  - ☐ Temp. Erosion Sediment Control Plan (TESC)

Yes No Permit or Approval

- ☐ ☐ Shoreline Permit
- ☐ ☐ State Waste Discharge Permit
- ☐ ☐ Temp. Modification of Water Quality Standards
- ☐ ☐ Section 4(f)/6(f): Wildlife Refuges, Recreation Areas, Historic Properties
- ☐ ☐ Water Rights Permit
- ☐ ☐ Water Quality Certification - Sec. 401 Issued by
- ☐ ☐ Tribal Permit(s), (If any)
- ☐ ☐ Other Permits, including GMA (List):

Part 3 Environmental Classification

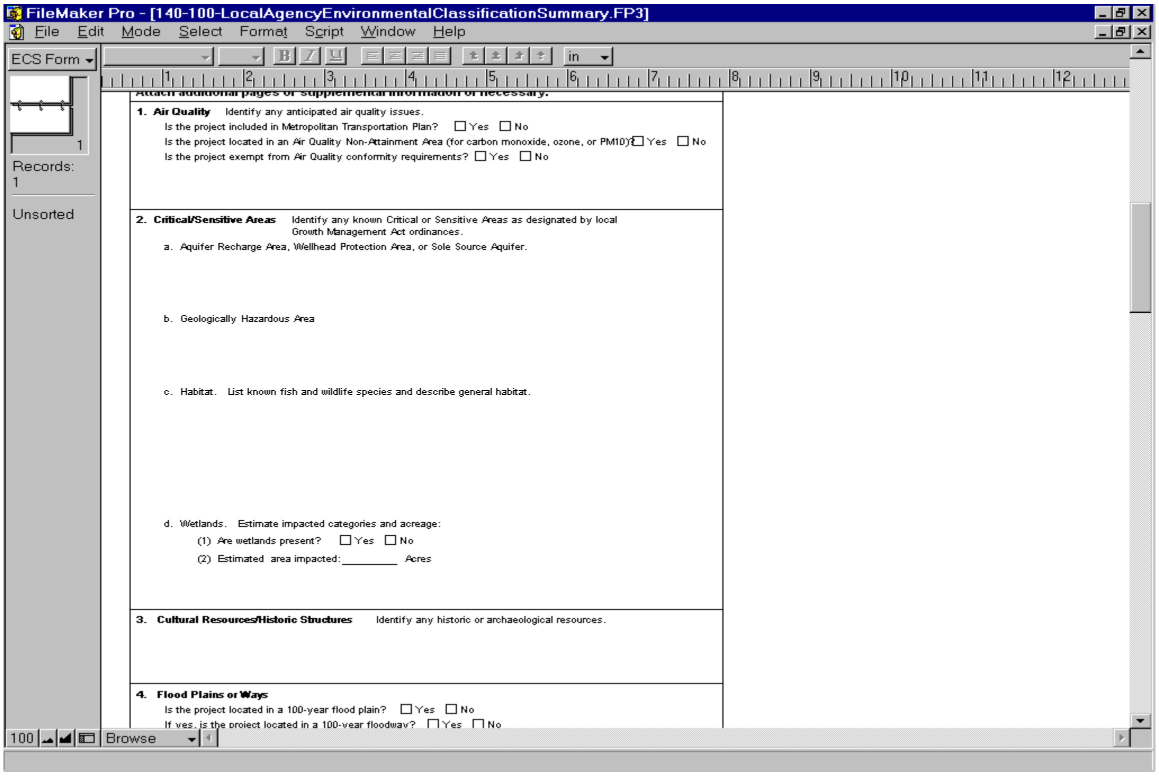
NEPA

- ☐ Class I - Environmental Impact Statement (EIS)
- ☐ Class II - Categorically Excluded (CE)
  - ☐ Projects Not Requiring Documentation for FHWA Approval (LAG 24.22(a))
  - ☐ Projects Requiring Documentation Without Further FHWA Approval (LAG 24.22(b))
  - ☐ Projects Requiring Documentation and FHWA Approval (Documented CE) (LAG 24.22(c))

SEPA

- ☐ Categorically exempt per WAC 197-11-800
- ☐ Determination of Non-Significance (DNS)
- ☐ Environmental Impact Statement (EIS)
- ☐ Other Actions: ☐ Adoption ☐ Addendum

100 Browse



FileMaker Pro - [140-100-LocalAgencyEnvironmentalClassificationSummary.FP3]

File Edit Mode Select Format Script Window Help

ECS Form

Records: 1

Unsorted

Attach additional pages of supplemental information of necessary.

1. Air Quality Identify any anticipated air quality issues.

Is the project included in Metropolitan Transportation Plan? ☐ Yes ☐ No

Is the project located in an Air Quality Non-Attainment Area (for carbon monoxide, ozone, or PM10)? ☐ Yes ☐ No

Is the project exempt from Air Quality conformity requirements? ☐ Yes ☐ No

2. Critical/Sensitive Areas Identify any known Critical or Sensitive Areas as designated by local Growth Management Act ordinances.

a. Aquifer Recharge Area, Wellhead Protection Area, or Sole Source Aquifer.

b. Geologically Hazardous Area

c. Habitat. List known fish and wildlife species and describe general habitat.

d. Wetlands. Estimate impacted categories and acreage:

(1) Are wetlands present? ☐ Yes ☐ No

(2) Estimated area impacted: ________ Acres

3. Cultural Resources/Historic Structures Identify any historic or archaeological resources.

4. Flood Plains or Ways

Is the project located in a 100-year flood plain? ☐ Yes ☐ No

If yes, is the project located in a 100-year floodway? ☐ Yes ☐ No

100 Browse

Attachment C: Example of Weighting and Ranking System for Environmental Screening

FLOODING & WETLANDS SUBJECT AREA

Subject Area Specialist: Gloria Skinner, Wetland Biologist, Washington Department of Transportation

Analysis: Query-Based

Buffer Distance: ½ mile either side of roadway

Data:	**Weight: (3 = most serious)**
Coastline	3
Floodways (not available for all counties)	1
HSP	0
Major Lakes	2
Major Rivers	2
Native Plant Wetland Subset	2
Streams	2
Wildlife Wetland Subset (includes threatened/endangered)	3
Wildlife Wetland Subset (no threatened/endangered)	2

Final Rating:	**Cumulative Score**
Low (1)	0 – 1
Medium (2)	2 – 6
High (3)	7 - 15 and any impact to a data category weighted 3

[1] 1998-2003 Strategic Plan. Washington Department of Transportation. Olympia, WA. August 1998.
[2] Environmental GIS Workbench Charter. Washington Department of Transportation. Olympia, WA. January 1999.

Assessing and Managing the Impacts of Urban Sprawl on Environmentally Critical Areas: A Case Study of Portage County, Ohio

Jay Lee[1] and Lynne J. Erickson[2]

Introduction

This project attempts to evaluate how the processes of uncontrolled growths of urbanized areas impact on environmentally critical areas and agricultural lands. Funded by Environmental Projection Agency, we have carried out this project using Portage County, Ohio as an example to test how GIS technology can be used to evaluate the effectiveness of various growth management strategies. GIS is also used to generate alternative build-out scenarios for policy makers to interact with the public. The results of this project provide valuable experience that can be applied in other regions for conserving environmentally critical areas and farmlands.

The uncontrolled urban growth is widely referred to as urban sprawl. It often implies negative impacts. Urban sprawl brings with it leapfrogging and low density developments of residential, commercial, and industrial land uses that are said to be undesirable and wasteful of lands. While the processes of urbanization and sub-urbanization are still active in northeast Ohio, the inevitable growth has greatly impacted on the adjacent areas. Even though planners have proposed and used various growth management strategies, the process as happened in reality take long time to see the out come and is unfortunately irreversible.

To assist local governments and the general public in understanding what uncontrolled urban growth may mean to their regions, what alternatives they may have to control the growth, or how the impacts may be if various strategies are used, GIS technology is used to develop simulation programs that incorporate various growth management strategies and tools for assessing the impacts on farmlands and environmentally critical areas (ECA) by urban growth.

In this project, a total of three scenarios have been implemented: (1) continued growth, (2) compact growth and (3) environmentally conservative growth.

[1] Applied Geography Laboratory, Department of Geography, Kent State University, Kent, OH 44242-0001. Email: jlee@kent.edu

[2] Portage County Regional Planning Commission, Ravenna, OH 44266. Email: lerickson@pcrpc.com

Portage County in Northeast Ohio

Affected by the expansion of the greater Cleveland-Akron-Canton metropolitan region, Portage County, Ohio has experienced a significant growth in recent years. Farmlands and ECA are being rapidly converted to residential, industrial and commercial developments. With the concern for the loss of ECA and farmlands and for a more efficient way of controlling the urban sprawl, the Portage County Regional Planning Commission has worked with Kent State University Applied Geography Laboratory to develop GIS tools for modeling and managing urban growth in the county.

Portage County is a semi-rural county that is 45 miles from Cleveland and 12 miles from Akron-Canton areas. While the growth of residential, commercial and industrial lands have been steady in the past decades, the most recent 10 years have seen especially fast growth.

GIS Technology

GIS technology is an effective tool to simulate the form and process of urban development on the landscape and to provide decision makers with the opportunity to evaluate alternative development policies on farmlands and ECA. The GIS tools developed for this study include: (1) the first stage of evaluating past trends of urban sprawl, (2) the second stage of simulating future growth with various growth management strategies, and (3) the third stage of assessing the impact by the simulated growth. Our study incorporated data sets from public and private agencies as well as recent Thematic Mapper remote sensing data. The GIS tools developed in this study prove to be efficient in simulating, mapping and measuring the impacts by urban sprawl. Using similar data sets, these GIS procedures can be easily adapted to carry out similar work in other regions.

At the core of the GIS tools developed for this project are a series of computer programs coded in Visual C++. These programs first process and combine GIS data layers to bitmaps to reduce the size of data storage needed. The simulations of future land use are implemented so that they are run according to specifications defined by the analysts. The simulated build-out scenarios are then imported to ArcInfo™ and ArcView™ for assessing the impacts by the simulated growths.

Assessment

To support the assessment of impacts by urban sprawl, a total of 13 layers of information were integrated into a GIS database. These layers include data for past and present land use patterns, soils, slopes, ground water availability, ground water pollution potential, wetland inventory, farmlands, and natural heritage data describing locations where special plants/animals/geologic structures are recorded. For ECA, areas of steep slopes, high ground water pollution potential, wetlands and those areas adjacent to special plants/animals were extracted from GIS layers and combined to form a new layer.

To simulate future urban growth, past and present land use patterns have been used to evaluate the spatial patterns and magnitude of urban growth in the county. A total of three management strategies were applied: (a) continued growth model - continuing the trend found from analyzing the past/present land use patterns, (b) compact growth model - applying moderate growth control such as open space development, growth centers, etc., and (c) environmentally conservative growth model – growing first in non-ECA and only when absolutely necessary in ECA.

In each growth model, projections of growth in residential units, commercial establishments and industrial establishments are estimated based on past trends and local master plans. The simulations can be performed for any given years as specified by the analysts. Furthermore, residential growths have been separated into residential units along frontage of roads and units in residential subdivisions with or without open space designs. For commercial and industrial establishments, areal sizes were estimated using average size of existing facilities in respective neighborhoods.

For continued growth model, trends were calculated from the land use patterns of 1975, 1985, 1995 and 1997. We calculated the proportion of frontage/subdivision residential developments, individual/conglomerated/striped commercial and industrial facilities as the basis for simulating future growth. In compact growth model, the open space design of residential units is used in neighborhoods where local zoning codes are appropriate. In addition, centers for future growth have been designated by local planners to guide the simulations so to control the growth to be within a half mile of radius from the centers. Finally, the environmentally conservative growth model requires future growth to occur first in areas that are not classified as ECA. The simulations place future development in ECA only when non-ECA are exhausted in the neighborhoods.

In the simulations, we did not have much problems of finding sites for new development of residential units. However, projected new development for commercial and industrial lands are limited by the amount of lands zoned for such uses in each township, city, or village. In several occasions we found some townships do not have enough land zoned for commercial and/or industrial to accommodate all the projected growth for these uses. This, of course, indicates that these townships need to re-consider their zoning districts and revise their master plan.

To allow more realistic simulations, we have structured the simulation programs to have the function of handling spill-over developments. If a particular township, city, or village does not have enough zoned commercial areas, for example, the commercial areas in the adjacent township is used to accommodate the growth. The simulation program has this function as a choice, not a mandatory step, so that analysts can generate a wide variety of scenarios to test their management strategies.

To assess the impact by urban sprawl, layers of simulated land use patterns are overlaid onto layers of information describing farmlands and other ECA. Tabulated results of loss of ECA and farmlands were correlated with changes of criteria used in simulations to observe their impacts.

Table 1: Acreage of projected new development, farmland loss, and loss of environmentally conservative areas (ECA).

	(In acres)	**Residential Lands**	**Commercial Lands**	**Industrial Lands**	**Farmlands Loss**	**ECA Loss**
Model 1	**cities**	*9,458*	*446*	*1,420*	*970*	*645*
	villages	*201*	*18*	*26*	*7*	*18*
	townships	*12,617*	*237*	*500*	*5,321*	*627*
Model 2	**cities**	*3,112*	*157*	*760*	*332*	*148*
	villages	*89*	*12*	*12*	*3*	*6*
	townships	*7,249*	*138*	*292*	*3,023*	*362*
Model 3	**cities**	*5,007*	*240*	*870*	*420*	*27*
	villages	*102*	*15*	*16*	*6*	*2*
	townships	*8,834*	*172*	*376*	*3,546*	*57*

Given the printing scale and the extent of the entire Portage County, it is not possible to include maps showing all the details of simulated land uses here. Alternatively, the following three maps show only the most northwest corner of Portage County as it is the are that is the closest to Cleveland. As the three maps show, the continued growth model consumes developable areas in a least efficient way. The compact model, on the other hand, provides the least waste of lands between simulated development. Finally, the environmentally conservative growth model shows developments in areas that are not environmentally sensitive.

Figure 1: Continued Growth Model. The simulated land use is scattered around the entire region, showing the least control and the most wasted land.

Given the printing scale and the extent of the entire Portage County, it is not possible to include maps showing all the details of simulated land uses here. Alternatively, the following three maps show only the most northwest corner of Portage County as it is the are that is the closest to Cleveland. As the three maps show, the continued growth model consumes developable areas in a least efficient way. The compact model, on the other hand, provides the least waste of lands between simulated development. Finally, the environmentally conservative growth model shows developments in areas that are not environmentally sensitive.

Figure 2: Compact Growth Model. The simulated land use shows a greater degree of conglomeration, avoiding wasted land between developments.

Figure 3: Environmentally Conservative Growth Model. The simulated land use is mostly located in areas that are not environmentally sensitive as defined by ECA.

Concluding Remarks

We have found that our approach of using a combination of simulation programs and GIS technology is able to provide realistic evaluation of how urban sprawl impacts ECA. The simulations can be guided by different projections for future growth under various growth management strategies. This provides great flexibility for planners and other decision makers to experiment what they think as appropriate growth scenarios. GIS technology is used in this project for data management, mapping, and overlaying analyses. This project has demonstrated the feasibility of such approach to efficient modeling of simulating future urban growth. Finally, the GIS procedures developed in this study can be easily adopted for other regions with similar data sets.

A High-Resolution Weather Data System (HRWxDS)

for Environmental Modeling and Monitoring

David R. Legates
University of Delaware
Newark, Delaware 19716 USA

Kenneth R. Nixon, Thomas Stockdale, and Geoffrey E. Quelch
Computational Geosciences Inc.
Norman, Oklahoma 73069 USA

Abstract

The High-Resolution Weather Data System (HRWxDS) is a real-time, site-specific, operational system that couples this new weather information with surface observations, hydrological modeling, and an interface to facilitate more informed decision-making tasks. Spatially- and temporally-distributed meteorological and hydrological fields produced by the HRWxDS include precipitation, wind velocity, air temperature, and atmospheric humidity and pressure. Derived fields include crop stress factors, soil moisture content, and runoff potential. Both digital and graphical products are produced that can be used for monitoring and analyzing meteorological and hydrological conditions for a particular location or region. This system is designed to improve site-specific water resource management for a variety of purposes including river management for optimal hydroelectric power generation, soil moisture monitoring for optimal irrigation scheduling or for wildfire prediction, wind estimation for pesticide drift applications, and rainfall/flood monitoring for enhanced emergency management.

1. Introduction

Surface observations are usually the only source of meteorological data for applications involving environmental modeling and monitoring. Advances in Geographic Information Systems (GIS) methodology and sophisticated spatial interpolation techniques, as well as the development and installation of the national network of Doppler weather radars, now has allowed for improved, high-resolution weather data to be used in the newer generation of distributed (*i.e.,* grid-based) environmental models. Our High-Resolution Weather Data System (HRWxDS) is designed to facilitate more informed decision-making tasks by coupling this new weather information with surface observations, hydrological modeling, and a graphical user interface in a GIS-based structure. The system operates in real-time to provide site-specific hydrometeorological information of spatially and temporally distributed meteorological and hydrological fields including precipitation (from raw radar estimates, surface observations, and gage-calibrated radar estimates), surface wind velocity (speed and direction), air temperature, atmospheric humidity (dewpoint and relative humidity), and atmospheric pressure. Modeling of the surface water hydrology using these observational inputs and land surface information (soils and vegetation) allows for estimation of crop stress factors (both temperature and moisture stress), soil moisture content and deficit, and runoff potential. The HRWxDS produces both digital and graphical products to allow for monitoring and further analysis of hydrometeorological conditions for a particular location or region.

The purpose of the HRWxDS is to provide improved site-specific hydrometeorological information to better manage our nations' water resources. Applications include river stage management for optimal hydroelectric power generation, soil moisture monitoring for wildfire potential or optimal irrigation scheduling, surface wind estimation for pesticide drift applications, air temperature monitoring for identification of electric power demands or freeze potential, and rainfall/flood monitoring for enhanced emergency management. Presently, the HRWxDS and its components are being used in a wide variety of applications (Nixon and Legates, 1994; Legates *et al.*, 1996; 1998) including operation of the front-end of a river management system to model the real-time water flow for the Catawba River Basin in North Carolina (by Duke Energy Corporation) and an assessment of the spatial and temporal distribution of rainfall for several flooding events in Texas. Irrigation scheduling and pesticide drift applications in southwestern Oklahoma also are under development.

Sophisticated software engineering techniques incorporating the latest in software development methods were used in the development of the HRWxDS. With its design for spatial analysis, the HRWxDS incorporates advanced GIS tools and its products easily can be input to commercial GIS packages for further analysis and presentation. Moreover, the design of the HRWxDS incorporates a modular, extendible architecture to facilitate easily the development of new products and the incorporation of new algorithms. Research on the HRWxDS extends directly from a successful technology transfer project first began in the Center for Computational Geosciences at the University of Oklahoma.

2. Overview of the High-Resolution Weather Data System

Five major subsystems comprise the HRWxDS (Figure 1). The *Setup and Processing Control* is a graphical user interface that allows the user to optimize the data acquisition and processing/analysis to the specific region/task at hand. It performs system initialization, process control, event timing, and system shutdown. The *Data Acquisition Module* is responsible for acquiring the necessary surface observation station data and radar products. As the HRWxDS is a real-time system, this module is responsible for data exception handling when, for example, internet connections fail or a radar site is down. The *Product Processing Component (PPC)* generates the gridded hydrometeorological products for each of the observational fields (*i.e.*, precipitation, air temperature, atmospheric humidity, wind velocity, and atmospheric pressure). The *Event Logging Facility (ELF)* provides the capability to log all significant processing events and notify someone (via e-mail, pager, *etc.*) when exceptional conditions -- extreme or unusual hydrometerological events or system failure -- warrant. Finally, the *Decision Support/Display Subsystem* provides display of gridded products generated by the PPC. This subsystem can analyze the hydrometeorological products or it can be tailored to form an application-specific user decision support system.

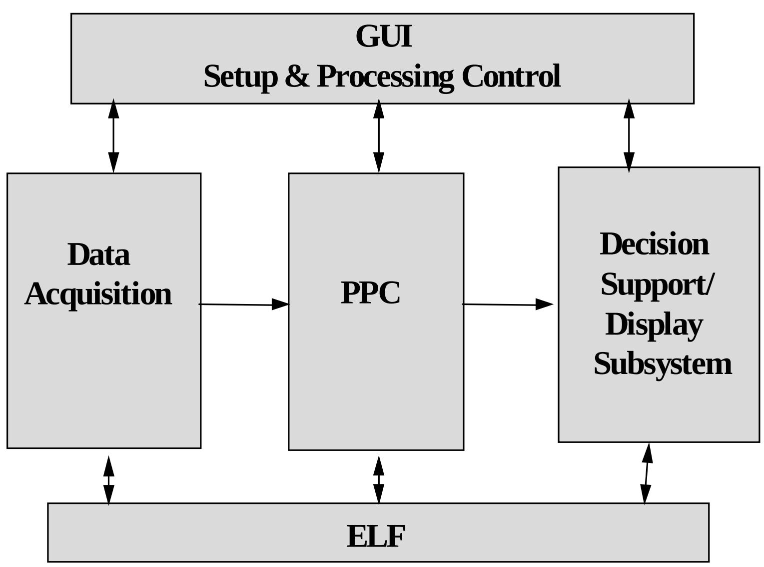


Figure 1: A Schematic Overview of The HRWxDS Subsystems.

3. Hydrometeorological Data Acquisition Requirements

One of the main advantages to the HRWxDS is that it uses generally available meteorological data from the National Weather Service (NWS) in addition to surface observation measurements assessable from local or private mesonetworks. The two primary inputs are surface station observations and the WSR-88D Digital Precipitation Array (DPA) -- both obtained generally on an hourly time-step (although the HRWxDS can be operated using virtually any time step).

Surface observation station data used by the HRWxDS may include data from the NWS First-Order Station Network, from local or private mesonetworks, or both. Applications of the HRWxDS have used the Oklahoma Mesonet data for analyses within the State of Oklahoma, IFLOWS data from the US Geological Survey in the Appalachian Mountains, and local rain gage networks in and around Houston, Texas -- in addition to the national NWS network. Hydrometeorological measurements utilized by the HRWxDS include air temperature, wind speed and direction, atmospheric pressure, relative humidity or dew point temperature, and precipitation. Although the NWS First-Order Station Network does not include solar radiation, the HRWxDS also can use this variable if a local network provides such information (*e.g.*, the Oklahoma Mesonet).

Biases -- most often underestimates -- in precipitation measurement with can-type rain gages can occur from several sources. These include the deleterious effect of the wind, wetting losses on the interior walls of the gage, evaporation from the gage between the end of precipitation and its measurement, splashing into and out of the gage, mechanical errors (*e.g.*, friction of pen plotters and the inability of tipping-bucket gages to accommodate heavy precipitation rates), and observational effects (Groisman and Legates, 1994; 1995). Using the surface wind speed measurements, the HRWxDS is able to estimate the biases that result from the most significant forms of the gage bias -- wind effects and wetting losses. These biases can amount to approximately six percent in summer for much of the United States to more than thirty percent in the northern tier of states in the winter (Legates and Deliberty, 1993a,b). Thus, the HRWxDS can reduce gage undercatch biases caused by the gage measurement process.

In addition to the surface observations, the HRWxDS also includes precipitation estimates from the NWS national network of weather radars. In the 1990s, the NWS began installation of the next generation of weather radars, now known as the WSR-88D (Weather Surveillance Radar --

1988 Doppler). These radars provide a variety of meteorological information, including several precipitation products. Among these include graphical images -- used primarily for visual analysis and television weather displays -- and a single digital precipitation product known as the Digital Precipitation Array (DPA). The DPA is a running hourly precipitation estimate developed from the radar reflectivities (microwave energy bouncing from hydrometeors) measured over the previous hour. It is updated every five or six minutes when the radar has detected precipitation somewhere in the radar umbrella. For more information on the DPA, see Fulton *et al*. (1998) or Legates (2000).

For some applications, a single WSR-88D radar may not cover the entire region of interest or a single site may be covered by more than one WSR-88D radar. The advantage of the HRWxDS is that it ingests data from multiple radars and can mosaic their DPAs to provide spatial coverage of virtually any geographic area of interest. Applications for the Catawba River Basin in North Carolina and over the State of Oklahoma require four and eight radars, respectively, for complete coverage. Indeed, in an assessment of precipitation for the Southern Great Plains, twenty-three radars were used. This ability to ingest, utilize, and mosaic multiple radars makes the HRWxDS an invaluable tool for spatial analysis of hydrometeorological conditions for large geographic regions.

4. Gridded Hydrometeorological Data Products

A suite of gridded hydrometeorological data products is produced by the HRWxDS. All products have a spatial resolution of approximately 4 km x 4 km (*i.e.*, at the nodes of the Hydrological Rainfall Analysis Project or HRAP grid) with individual estimates for each hydrometeorological variable for each HRAP cell. The suite of gridded products includes five precipitation products and seven hydrometeorological companion products. All products are provided in both a common GIF and netCDF format for maximum flexibility and portability.

With respect to precipitation, the HRWxDS provides five fields -- Gage-Based Precipitation, Radar-Based Precipitation, Gage-Calibrated Radar Precipitation, and 12- and 24-Hour Precipitation Accumulation. The Gage-Based product provides gridded precipitation estimates based solely on rain gage observations adjusted for gage measurement biases that include the effect of rain gage design (*e.g.*, shielded or unshielded gage, size of orifice opening, and height of gage orifice above the ground) as well as the meteorological conditions around the rain gage. The Radar-Based product presents precipitation estimates from the radar DPAs using an NWS-specified fixed reflectivity-to-rainfall conversion equation and is not calibrated using rain gage observations. Moreover, it is a mosaic of multiple radars using either an averaging or maximization (default) of radar overlap areas, depending upon the user's preference.

The most important component of the PPC is the Gage-calibrated Radar Precipitation product. Gage observations are paired with the hourly composite radar reflectivities to remove errors in the radar precipitation estimates. Such errors can arise from (1) uncertainties in obtaining accurate estimates of reflectivity, (2) inconsistencies in the reflectivity-to-rainfall conversion algorithms, and (3) effects occurring below the radar beam. The calibration procedure utilizes a variety of spatial statistical techniques using the bias-measurement-adjusted gage observations to estimate and remove the radar bias. Each radar is calibrated separately and the calibrated estimates then are combined to form a composite mosaic from multiple radars. In addition, adding these calibrated mosaics to the desired temporal resolution produces the 12-hour and 24-hour Precipitation Accumulation products. For more information on radar biases, please consult Wilson and Brandes (1979), Doviak and Zrnic (1984), and Fulton *et al*. (1998). Legates *et al.* (1999b) and Legates (2000) also provide a more detailed discussion of the HRWxDS calibration algorithm.

Air temperature, dew point temperature, relative humidity, wind speed and direction, solar radiation, and atmospheric pressure, in addition to precipitation, also are interpolated to the HRAP grid intersections. Spatial interpolation is accomplished through a modified version of Shepard's inverse distance weighting method (Willmott *et al*., 1984) that accounts for both the distance and distribution of the observations around the interpolate point. This method represents an exact pass through the data (the interpolation returns the observed value when an observation is coincident with the grid intersection). Air temperature is interpolated by reducing the observations to sea-level equivalent (potential temperature), interpolating potential temperature, and then returning the observations to the elevations of the HRAP grid cells using the environmental lapse rate and a digital elevation model (DEM) of the geographic area to be interpolated. Since dew point temperature is upward bounded by air temperature, the return of air temperature to the HRAP elevation follows the moist adiabatic lapse rate when saturation is reached. Dew point temperature then is set to air temperature in that case (*i.e.*, a relative humidity of 100%). Relative humidity is calculated directly for each grid cell from the dew point and air temperature estimates for each grid cell with no interpolation. For more information on the interpolation procedure, please consult Legates *et al.* (1999a).

5. The Graphical User Interface and User/Setup Options

A graphical user interface (GUI) is included in the HRWxDS to assist in customizing and tailoring the system to meet specific operational requirements and set processing control options. Several user-adaptable features provided through the GUI. For example, the HRWxDS can be easily setup to analyze and monitor one or more specific geographic regions from a small river basin to large regions of the country. Once a geographic region has been defined, the HRWxDS automatically generates the HRAP grid that covers this area. Then the user can select the specific region to be processed and the HRWxDS can switch from one region to another.

The HRWxDS can be run in either "Attended" or "Unattended" modes of operation. In "Unattended" mode, the system selects from predetermined responses to handle anticipated operational exceptions. This mode is most frequently used for real-time, operational monitoring of environmental conditions. In "Attended" mode, a user has the capability to dynamically interact with the processing and handle exception conditions as they occur. This mode is frequently used for historical/forensic analyses for regions or times when the HRWxDS was not running in real-time or to update missing data during power outages or other interruptions.

5.1 Data Initialization Specifications

To incorporate radar data for a specific geographic region, the radars that cover the area of interest can be dynamically added or deleted, depending upon interest and availability. In addition, the HRWxDS has the capability to incorporate surface observation measurements of national (*e.g.*, NWS), local/private mesonetworks, or both. As with radars, stations can be dynamically added or deleted. Each rain gage location is described in a rain gage metadata file which includes information about the rain gage design (shielded or unshielded gage), size of orifice opening, height above ground, and sheltering conditions. These data are used to provide estimates of the gage measurement bias adjustments.

5.2 Processing Control Setup and Knobs

The GUI also provides specifications for the processing start time and processing interval. To ensure that all pertinent data have been received, a delay in starting time can be included. For example, processing for the current hour (at the top of the hour) may be delayed until fifteen minutes past the hour to ensure that all station observations and radar DPAs have been transmitted. In addition, time limits also are set to specify which input data are to be considered valid for the current processing cycle.

A set of "processing control knobs" also is provided to adjust and fine-tune the processing. Although the PPC uses a large array of control knobs, only a few are described here. "Mosaic Knobs" control the numerical technique (averaging or maximization) used to develop the mosaics for both the Radar-Based and the Gage-Calibrated Radar products. "Hail Detection and Suppression Knobs" provide absolute (derived from reflectivity) and relative (grid cell-to-grid cell comparisons) rainfall rates for the detection and removal of hail contamination in the radar reflectivities. Additionally, the disposition of cells can be specified for which hail contamination has been identified (*i.e.*, cells can be set to missing, truncated to the maximum allowable precipitation rate, or interpolated from surrounding cell values). A "Tropical Rain Event Knob" identifies a precipitation event as tropical (*e.g.,* tropical storm, hurricane, tropical depression) which adjusts (upwards) the acceptable upper limit of possible rainfall rates and suppresses the identification of heavy rainfall as hail contamination. The "Missing Data Exception Handling Knob" is a processing control feature that automatically substitutes interpolated rain gage data in the Gage-Calibrated Radar product for areas where radar data are missing (a radar may be down or data retrieval has been delayed).

A number of other processing knobs controls the spatial interpolation procedure. For example, the "Radius of Influence Knobs" are used to specify the distance over which an observation will influence the interpolated values. A separate knob controls each weather variable (*e.g.*, rain gage measurements, air temperature, dewpoint temperature, and atmospheric pressure). The "Optimization/Calibration Knob" provides for the operational calibration and optimization of the spatial interpolant. It allows distance decay effects and elevational influences to be estimated from the observed data rather than assuming constant values. "Elevation Effect Knobs" allow for elevation effects to be either enabled or disabled in the interpolation of each weather variable. Some variables (*e.g.*, air temperature) are affected by elevation whereas others (*e.g.*, sea-level atmospheric pressure) are not.

The HRWxDS system also includes a feature that allows a user in "Attended" mode the opportunity to add temporary observations. These can be one-time field measurement or report (*e.g.*, a single hourly observation from a station that does not report regularly) or data to manually adjust the Gage-Calibrated Radar Precipitation product at the user's discretion. This

"Virtual Rain gage Facility" allows for inclusion of observations from either rainfall or snowfall data (conversion from snow to liquid water equivalent is adjusted by the "Snow Density Knob"). Rainfall also can be suppressed by the Virtual Rain gage Facility to provide adjustments for virga or ground clutter conditions.

5.3 General Operational Features

Other controls available through the GUI allow for the selection of measurement units, either English or Metric, a fixed or variable color bar, and display scale thresholds. The overlay feature provides the capability to display radar or rain gage locations, basin delineation, HRAP grids, political boundaries, rivers and streams, roads, or other overlay combinations on the GIF images that are produced. An animation capability is included to display a time-series of products to provide a visual representation of changing meteorological conditions. Individual cell values also can be interrogated by simply moving the cursor to the cell of interest. Latitude and longitude of the cell are displayed with the cell's value for the currently displayed product. In addition, the HRWxDS has the capability to display the contribution of each radar to the Radar-Based Precipitation and the Gage-Calibrated Precipitation products through the DPA Viewer.

6. The Event Logging Facility (ELF)

The HRWxDS contains a robust Event Logging Facility (ELF) that records system configuration information (*e.g.*, data paths, processing parameters, and region definitions), data retrieval status messages, statistics for input and output data, and processing status information. It is a central facility that tracks all events of which the user may need to be made aware. Destination for event logs may be specified to a disk file and/or a screen dialog box. Each message is assigned to an appropriate category -- status (memory and disk usage), processing information, statistics, operating system errors, minor error, major error, or fatal error. A "Log Viewer" provides the capability to view the ELF file/records by filtering the category of records to be displayed (*e.g.*, show only major/fatal error messages or show processing information and statistics).

7. An Example of Weather Information Produced by the HRWxDS

At present, the HRWxDS is continuously running in operational mode at Duke Energy Corporation in Charlotte, NC where it is being used to provide precipitation inputs to a River Management System for the Catawba River Basin in central North Carolina and northern South Carolina. In addition, it also provides real-time estimates for the State of Oklahoma and for a region covering much of the Southern Great Plains (see Figure 2). More information about the HRWxDS can be found on the Computational Geosciences Inc. Web Page at http://www.telepath.com/compgeo.

To illustrate the products generated by the HRWxDS, an example from 5 GMT (Midnight CDT) on June 5, 1999 for the Southern Great Plains Region (Oklahoma, northern and central Texas, and surrounding regions) will be shown. Station observations (small squares) and radar locations (large squares) are presented in Figure 2 along with the HRAP rectangle that was used for the analysis.

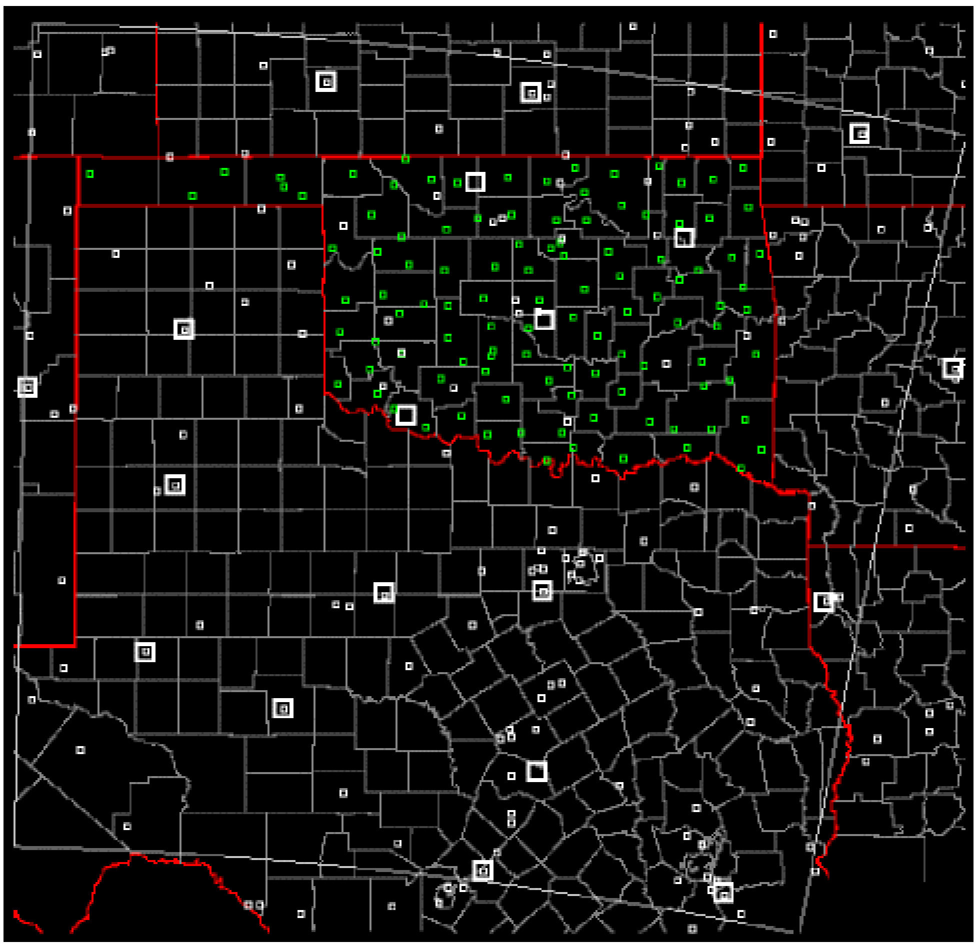


Figure 2: Locations of the radars (large squares) and surface observing stations (small squares) used over an area of the southern Great Plains (approximately 1000 km x 1000 km). Higher station densities over Oklahoma are attributable to use of the Oklahoma Mesonet. The white, four-sided border shows the boundary of the HRAP grid used for this region.

During the early morning hours, a line of convective showers with precipitation rates exceeding 25mm per hour moved into western Oklahoma and south-central Kansas. The HRWxDS was in operational, "unattended" mode during this event and provided real-time estimates of hourly surface weather conditions for this storm. Images of the precipitation for this hour are shown (Figure 3) for the Gage-Based, Radar-Based, Gage-Calibrated Radar, and 24-hour Precipitation products. Note that although these fields are illustrated here in graphical form, the main purpose of the HRWxDS is to produce high-resolution *digital* representations of the various meteorological variables. Thus, these fields from the HRWxDS can be readily input to a distributed hydrological model or any raster-based GIS. Although the HRWxDS does not utilize any commercial GIS package, it contains sophisticated spatial analysis and interpolation algorithms operating in a spatial (GIS) framework. Output from the HRWxDS is easily incorporated as ARC/INFO™ overlays.

A comparison of the Gage-Based (Figure 3, top) and Radar-Based Precipitation products (Figure 3, bottom) clearly illustrates the advantages of the radar precipitation estimates. Outside Oklahoma, the NWS first-order weather station network is used. Note that the sparse density of the NWS/Oklahoma Mesonet combination only is able to resolve two areas of relatively heavy precipitation in northwestern and west-central Oklahoma and a smaller area of lighter precipitation in southwestern Oklahoma. By contrast, the radar mosaic for this area illustrates a much larger region of precipitation with two embedded cells of heavy rainfall in central- and northwestern Oklahoma. Precipitation in southwestern Oklahoma can be seen to extend farther back into Texas where the gage network is much more sparse and is not supplemented by the Oklahoma Mesonet, a relatively dense network of 111 stations (see Figure 2 for a station distribution). The spatial fidelity resolved by the radars is much higher than even the Oklahoma Mesonet can provide and a better picture of the convection area over this region. Note that the area of light rain across much of the region is under-represented by the gage network and, since no gages were located beneath the smaller, embedded storm cells, the magnitude of rain falling from these cells also is considerably underestimated.

The NWS radar product (*i.e.*, the Level III DPA) tends to underestimate rainfall rates for a variety of reasons (see Wilson and Brandes, 1979; Fulton *et al*., 1998; Legates, 2000). This was particularly true for this event as well. Consequently, calibration of the radar reflectivities using the real-time gage measurements has inflated the radar estimates to be commensurate with the gage observations. Note that for the precipitation cell in west-central Oklahoma, the

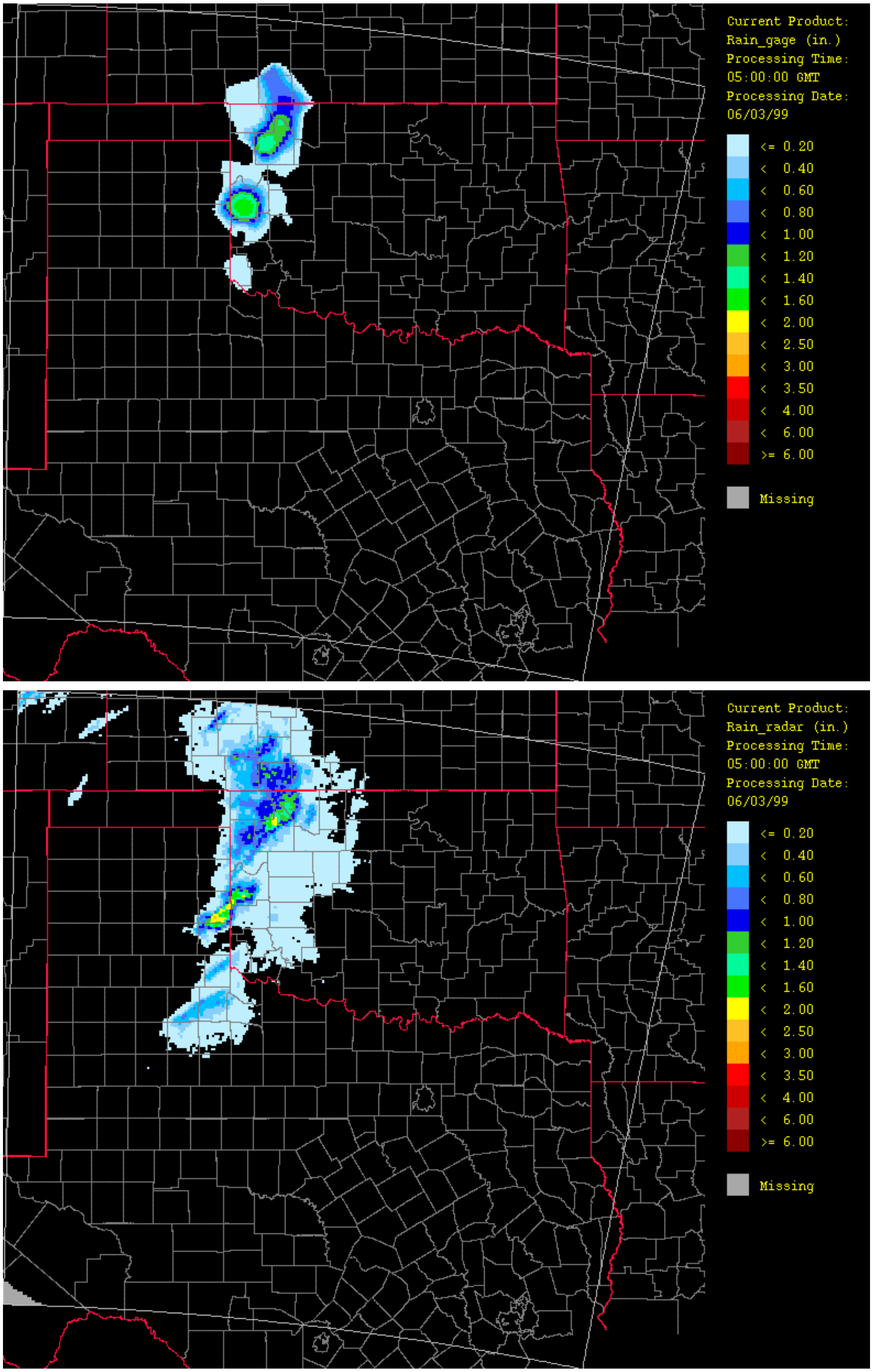


Figure 3: Estimates of precipitation from the Gage-Based Precipitation product (top) and the Radar-Based Precipitation product (bottom) for 5:00 GMT on June 3, 1999.

calibration resulted in a slight decrease in precipitation intensities. This shows the difference between the radars and the precipitation estimates obtained from the different radars that viewed different portions of the precipitation. The mean fit error for this hour was less than 10 mm indicating a good agreement between the gage observations and the calibrated radar product (Figure 4, top) was obtained. The 24-hour Precipitation product also is shown (Figure 4, bottom - Note the different color bars used for the 24-hour product).

Throughout much of the United States, rain gage densities are seldom as dense as the Oklahoma Mesonet and are usually as sparse as the NWS network is here (Figure 2). Precipitation, and especially convective precipitation, events are under-represented by existing gage networks. Weather radars help to provide a more complete picture of the true distribution of precipitation, although some exceptions exist. Most notably, mountains obscure the radar beam so that some remote and relatively inaccessible areas of the western United States are not covered adequately by the radar network. Given the sparse gage networks, use of the WSR-88D radar precipitation estimates, when calibrated properly by the HRWxDS, clearly provide a more accurate representation of the true precipitation over virtually all of the conterminous forty-eight states.

In addition to precipitation, shown here are the air temperature (Figure 5, top) and wind vector (Figure 5, bottom) fields for the same hour. Air behind the advancing convective system has been cooled by the precipitation by as much as 20°F (11°C) in some places and a strong horizontal temperature gradient lies just behind the area of convection. Airflow is predominantly from the south and southeast -- as is common for late spring -- with the largest area of winds exceeding 20 mph (10 ms^{-1}) feeding the convection region on the southern flank. Lighter and more easterly winds are found over the northeastern part of the region.

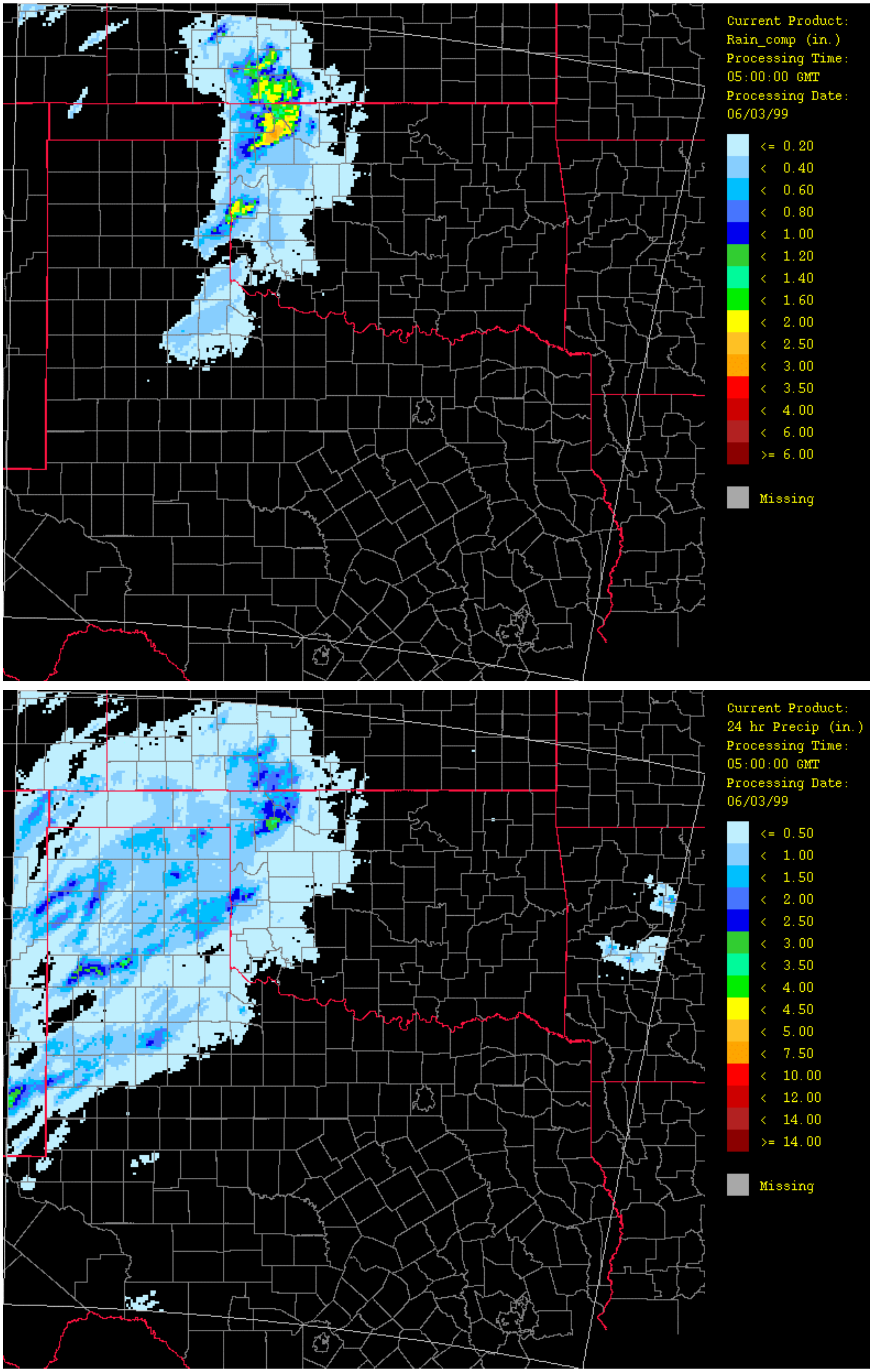
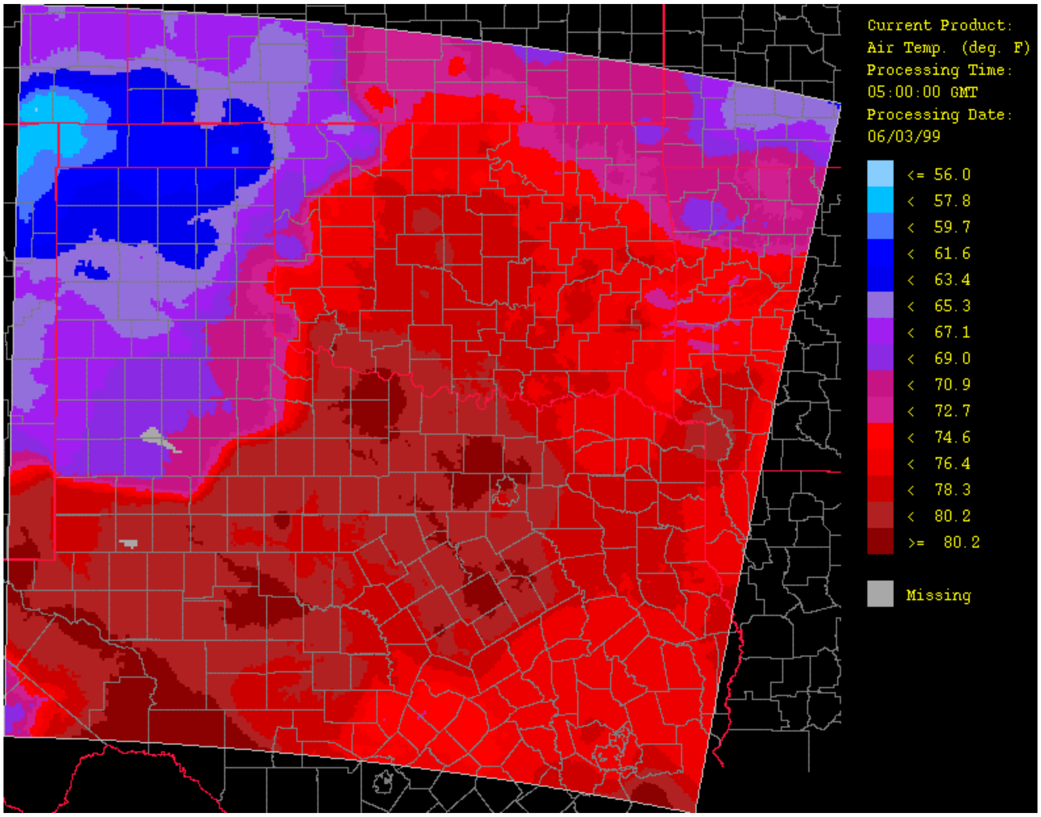


Figure 4: Estimates of precipitation from the Gage-Calibrated Radar product (top) and the 24-hour Precipitation product (bottom) for 5:00 GMT on June 3, 1999.


Current Product:
Air Temp. (deg. F)
Processing Time:
05:00:00 GMT
Processing Date:
06/03/99
<= 56.0
< 57.8
< 59.7
< 61.6
< 63.4
< 65.3
< 67.1
< 69.0
< 70.9
< 72.7
< 74.6
< 76.4
< 78.3
< 80.2
>= 80.2
Missing

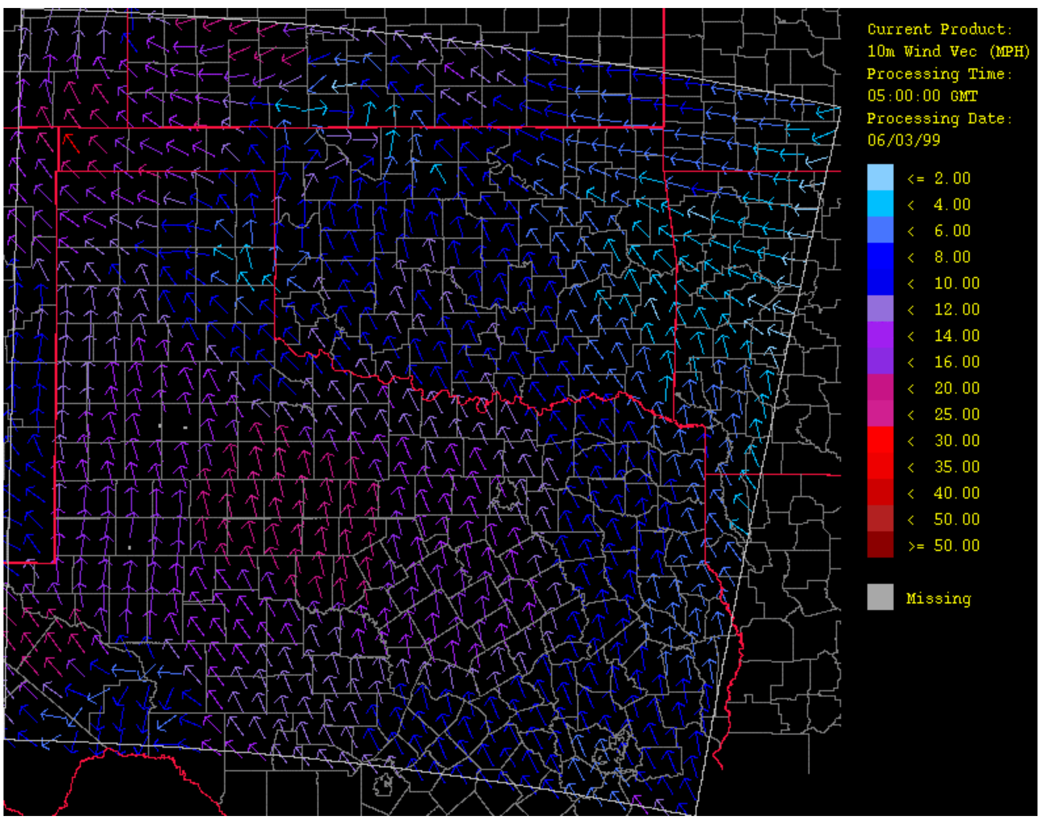

Current Product:
10m Wind Vec (MPH)
Processing Time:
05:00:00 GMT
Processing Date:
06/03/99
<= 2.00
< 4.00
< 6.00
< 8.00
< 10.00
< 12.00
< 14.00
< 16.00
< 20.00
< 25.00
< 30.00
< 35.00
< 40.00
< 50.00
>= 50.00
Missing

Figure 5: Estimates of air temperature (top) and wind vector (bottom) for 5:00 GMT on June 3, 1999.

8. The HRWxDS and Its Role within the Water Resource Decision Support System

The HRWxDs is but one component of a larger project, the Water Resource Decision Support System (WRDSS) which is designed to be a real-time tool for water resource decision makers (Legates *et al*., 1996; 1998). Using the real-time data ingest and weather product generation capabilities of the HRWxDS, Computational Geosciences Inc. has begun development of the WRDSS which couples soil moisture and runoff simulation models with the HRWxDS in a framework that allows development of application-specific decision support systems (Figure 6). Through the WRDSS, water resource managers can be supplied with readily available and easily understood information about current meteorological and hydrological conditions over a wide geographic area. In addition to the graphical representations, all weather and surface hydrological products are available in a digital format so the estimates can be imported to other hydrological models and applications.

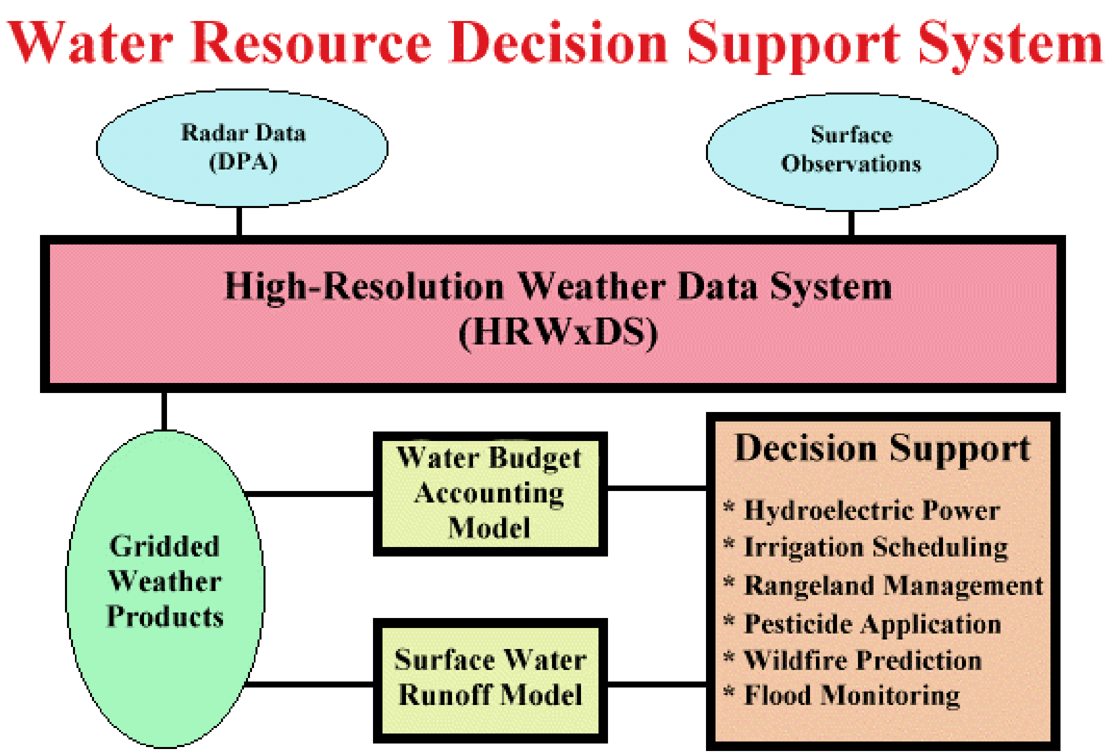


Figure 6: Schematic of the Water Resource Decision Support System (WRDSS).

9. Conclusion

Obtaining accurate, high-resolution, real-time estimates of surface meteorological variables has been a problem that has long faced hydrologists and hydrometeorologists. Development and implementation of the national network of WSR-88D weather radars and their ability to represent each storm's "precipitation footprint" has led to a revolution in hydrological modeling and monitoring. When properly calibrated with rain gage observations, as is accomplished by the High-Resolution Weather Data System (HRWxDS), these estimates have the potential to provide water resource decision managers with more reliable real-time precipitation assessments and, consequently, allow for more accurate spatial information to be included in the decision making process. The spatial interpolation procedures developed and contained within the HRWxDS provides the "state-of-the-art" in estimation of surface meteorological fields from irregularly-spaced station observation networks reporting in real time.

Development of the HRWxDS coupled with the Water Resource Decision Support System (WRDSS) provides decision-makers with an integrated tool that allows for real-time assessments of water resource and hydrological conditions. This tool, which converts weather and hydrological data into usable *information*, will prove invaluable by eliminating much of the assumptions, guesswork, and tedious monitoring and maintenance of inadequate gage networks that presently characterize water resource management. Both the HRWxDS and the WRDSS *system* will continue to evolve to meet the needs of decision-makers and to exploit advances in estimating precipitation by weather radars and obtaining high-resolution estimates of surface meteorological fields.

References

Doviak, R.J., and Zrnic, D.S., 1984: *Doppler Radar and Weather Observations.* Orlando, Florida: Academic Press, Inc.

Fulton, R.A., Breidenbach, J.P., Seo, D.-J., Miller, D.A., and O'Bannon, T., 1998: The WSR-88D Rainfall Algorithm. *Weather and Forecasting*, **13**:377–395.

Groisman, P.Ya., and Legates, D.R., 1994: Accuracy of Historical United States Precipitation Data. *Bulletin of the American Meteorological Society*, **75**:215–227.

Groisman, P.Ya., and Legates, D.R., 1995: Documenting and Detecting Long-Term Precipitation Trends: Where We Are and What Should Be Done. *Climatic Change*, **31**:601–622.

Legates, D.R., 2000: Real-Time Calibration of Precipitation Estimates. *The Professional Geographer*, forthcoming.

Legates, D.R., and DeLiberty, T.L., 1993a: Measurement Biases in the United States Rain gage Network. *Symposium on Geographic Information* Systems and Water Resources, American Water Resources Association, 547–557.

Legates, D.R., and DeLiberty, T.L., 1993b: Precipitation Measurement Biases in the United States. *Water Resources Bulletin*, **29**(5), 855–861.

Legates, D.R., Nixon, K.R., Stockdale, T.D., and Quelch, G.E., 1996: Soil Water Management Using a Water Resource Decision Support System and Calibrated WSR-88D Precipitation Estimates. *Proceedings, AWRA Symposium on GIS and Water Resources*, American Water Resources Association, 427-435.

Legates, D.R., Nixon, K.R., Stockdale, T.D., and Quelch, G.E., 1998: Use of the WSR-88D Weather Radars in Rangeland Management. *Specialty Conference on Rangeland Management and Water Resources*, American Water Resources Association, 55–64.

Legates, D.R., Nixon, K.R., Quelch, G.E., and Stockdale, T.D., 1999a: Environmental Modeling and Monitoring using a High-Resolution Weather Data System (HRWxDS). *Proceedings of the Fourth International Conference on GeoComputation*, Fredericksburg, VA, forthcoming.

Legates, D.R., Nixon, K.R., Stockdale, T.D., and Quelch, G.E., 1999b. Real-time and Historical Calibration of WSR-88D Precipitation Estimates. *Proceedings, Eleventh Conference on Applied Climatology*, American Meteorological Society, Dallas, TX, 76–77.

Nixon, K.R., and D.R. Legates, 1994: Water Resource Decision Support System SBIR Phase I Report for the US Department of Agriculture.

Wilson, J.W., and Brandes, E.A., 1979: Radar Measurement of Rainfall -- A Summary. *Bulletin of the American Meteorological Society*, **60**:1048–1058.

Fractal Dimension as an Indicator of Human Disturbance in Galveston Bay, Texas

Amy J. Liu[1,3], Guy N. Cameron[2*]
1. Department of Biology, University of Houston, Houston, TX 77204
2. Department of Biological Sciences, University of Cincinnati, Cincinnati, OH 45221-0006
3. Present address: National Risk Management Research Laboratory, MS 690, US EPA, Cincinnati, OH 45268

ABSTRACT

High productivity and accessibility to humans have made coastal wetlands attractive sites for human settlements. This study analyzed the wetland landscape patterns in Galveston Bay, Texas. The study described the relationships between the fractal dimension and factors which affect the wetland landscapes: land use, vegetation type, size, location, and level of human disturbance. The perimeter-area method was used to calculate the fractal dimension. There was a significant difference in the fractal dimension of wetlands when classified according to land use, vegetation type, size, and level of human disturbance. Furthermore, increasing the size of the road buffers did not have a significant effect on the fractal dimension of wetlands. These results will be important in determining how wetlands can be managed as natural resources and nature reserves.

INTRODUCTION

Wetlands are amongst the most important of the world's ecosystems. They play a significant role in supporting higher levels of biological diversity, as well as primary and secondary productivity, modulate flows of water, nutrients, and materials across the landscape, and provide wildlife habitat (Holland et al., 1991). If wetlands are to be properly used and protected, it is important that their spatial properties be well-understood and monitored.

Shaw and Fredine (1956) estimated that 35% of the wetlands in the United States had been lost by the 1950s, and Frayer et al. (1983) estimated a net loss of more than 3.7 million hectares, or 8.5%, between the 1950s and 1970s. Dahl and Johnson (1991) measured a net loss of 1.5%, or 28,400 hectares, between the 1970s and 1980s. These estimates of wetland losses provide an incomplete picture of the dynamics of change because they only describe the area of wetlands lost to agriculture, industry, and urban development and do not show the conversion of wetlands from one land use class to another by human activities. From 1970s to the mid-1980s, there was a net gain in freshwater marshes, despite the destruction of approximately 200,000 hectares, because 288,800 hectares of swamps were converted to freshwater marshes (Dahl and Johnson, 1991).

Human activities are the principal causes of loss of wetland (Craig et al., 1979). Wetland hydrology, such as drainage conditions and circulation patterns, makes wetlands more sensitive than woodlands to highway construction activities (McLeese and Whiteside, 1977). Land is lost in coastal areas through flood control measures, agricultural practices, and canal construction (Gagliano, 1973; Craig et al., 1979). Construction of canals and roads which establish permanent barriers to the growth of a particular land patch also will change the nature of the interaction with adjacent patches of land and will prevent the patch from becoming the dominant patch in the landscape (Forman et al., 1986). Therefore, because of the significant effect of land losses to wetland spatial dynamics, it is important to examine the space occupation properties (fractal properties) found in wetland ecosystems.

Fractals are based on the premise that a measure assigned to an object depends on the object's appropriate dimension. The fractal dimension describes the relationship between a quantity *Q*, and the length scale, *L*, over which *Q* is measured, where $Q(L)=L^{Dq}$, and D_q is the fractal dimension. D_q describes how the quantity *Q* varies with scale. Hence, for larger values of D_q, the length changes faster because the curve is more complex.

Having a useful knowledge of fractal dimensions has several implications for ecologists. First, the probability of an organism encountering a boundary increases as the fractal dimension of the patch increases (Weins, 1992). Second, the functionality of an ecotone is dependent on the surface area over which the ecotone extends. Since the fractal dimension is a more accurate measure of length (and thus, surface area), the functionality of a wetland patch is directly tied to its fractal dimension (Kent and Wong, 1982). Third, organisms requiring a particular amount of edge habitat may be restricted if the fractal dimension of a wetland patch falls below a critical threshold (Henderson et al., 1985). Fourth, an important feature of a fractal curve or surface is that its length or area becomes disproportionately large as the unit of measurement is decreased (Mandelbrot, 1977).

---

[*] To whom correspondence should be addressed.

Krummel et al. (1987) evaluated the fractal dimension (*D*) of patterns of deciduous forests in Mississippi using the perimeter/area method on aerial photographs. Their analysis revealed that small areas of forest tended to be smoother (D = 1.20), while larger areas had more complex boundaries (D = 1.52). This was interpreted to indicate that human disturbances predominated at smaller scales, allowing for smoother geometry and a lower fractal dimension, while natural processes tended to predominate at larger scales. Bradbury et al. (1984) investigated the possibility of hierarchical scaling in an Australian coral reef. They detected three ranges of scales, which corresponded with the scales of three major reef structures: individual coral colonies (D = 1.1), whole adult living colonies (D = 1.05), and groves and buttresses (D = 1.15). Since the major effect of low fractal dimension is to reduce the intimacy of contact of the living surface of the reef with the surrounding water, the authors speculated that corals attempted, at the adult colony level, to reduce the level of their contact with the medium.

As with any mathematical representation of nature, fractal geometry is an attempt to search for order in the complex patterns characterizing living systems. Life is composed of interactions and fluxes of matter, energy and information through interfaces, which suggest an “interpenetration volume” between two adjacent interacting elements. These interfaces, Frontier (1987) suggested, were neither surfaces or volumes, but fractals.

Fractals, then, seem to be an appropriate tool for investigating wetlands, since wetlands are contact zones between terrestrial and aquatic systems -- they are characterized by fluxes of energy and matter (nutrients and hydrology). Only recently has there been an appreciation for the ways in which ecotones frequently intensify or concentrate these activities. Johnston (1991) determined the net effect of wetlands on water quality, reporting that retention of nitrogen in natural freshwater wetlands ranged from 14 to 100%, and retention of phosphorus ranged from 4 to 80%. Peterjohn et al. (1984) found that riparian forests in the eastern United States exerted major influences on the flow of nitrogen and phosphorus as these chemicals moved from agricultural fields across forests and into streams. Much of the nitrogen and phosphorus was captured in the vegetation and soils of the riparian ecotone, with 60-75% of all nutrients captured occurring in the first 20 m of the riparian forest ecotone. The surface area of the contact zone, then, is important in the economy of the surrounding ecosystems and watersheds. The surface area of the contact zone depends on the length of the boundary, or, more precisely, on its fractal dimension (where length is not uniquely defined and depends on the scale at which

it is being measured).
Changes in landscape patterns may relate to the flows of material or energy across landscapes. For example, erosional processes, or sediment movement across landscapes and the abundance and distribution of wildlife might be predictable using indices of landscape pattern (Turner, 1987). Species that favor or require particular types of edges may decline if the amount of edge (which is directly related to the fractal dimension) declines, whereas species requiring extensive areas of land may benefit from the increasing size of patches (Henderson et al., 1985).

However, few studies have examined the importance of fractal dimensions in wetlands. The amount of wetland edge was closely related to the size of the harvest of offshore shrimp (Browder et al., 1989), and the rate of water loss from small sloughs varied directly with the length of shoreline per unit area (Millar, 1971). In addition, fractal dimension was significantly related to the level of human impact for riparian forests in Iowa (Rex and Malanson, 1990). Wickham et al.'s (1994) findings that the fractal dimension of wetlands increased with the number of agriculture and residential land cover components contrasted with the expected decrease in fractal dimension, with an increase in surrounding agricultural and human land use increased (Forman et al., 1986; Kummel et al., 1987).

While previous studies have linked size, presence of disturbance, and organizational structure, to shape complexity in ecological systems, the present study broadens the scope of fractal analysis by examining the relationships between ecological and anthropogenic processes and spatial patterns in coastal wetlands. Specific objectives included:

- Describe the relationship between the fractal dimension and specific factors which affect wetland landscapes, such as land use (wetland vs. non-wetland), vegetation type, size of wetland, location of wetland, and level of human disturbance or impact.

- Present a methodology to evaluate the effects of anthropogenic scaling which will be helpful in formulating hypotheses concerning the spatial scales of process-pattern interactions.

METHODS

Digitized maps of wetlands in the Galveston Bay were obtained from the United States National Wetlands Inventory (NHI) database. A wetland was defined as “land where saturation with water is the dominant factor determining the nature of soil development and the types of plant and animal communities living in the soil and on its surface” (Coward in et al., 1979). Descriptions of the NHI classifications (Coward in et al., 1979) were used to group wetlands down to class into different vegetation types, which constituted a wetland patch. Thirty-one 7.5 minute topographic quadrangles comprised the area surrounding Galveston Bay: Harris County, Chambers County, Galveston County, and Brazoria County. Quadrangles were edge matched and merged using ARC/INFO and ARCEDIT, and wetland sites were selected for analysis in the following categories: land use (wetland vs. non-wetland), vegetation type, size (large vs. small), location (palustrine vs. estuarine), and level of disturbance (Rex and Malanson, 1990). For the size comparison, all wetland patches in the study set of wetlands in the Galveston Bay system were used. For the level of disturbance studies, digitized road maps which detailed the transportation uses of land down to residential streets around the Galveston Bay system were obtained from the Texas Department of Transportation (TxDOT).

Perimeter and area information of each wetland patch or polygon were obtained from the digitized maps using ARC/INFO. The degree of complexity of a polygon was characterized by the fractal dimension D, such that the perimeter P of a patch was related to the area A of the same patch by $P = A^{(D/2)}$. The perimeter-area method for calculating fractal dimensions regressed log(P) on log(A) to evaluate D. The fractal dimension of each data set of wetland polygons (described below) was estimated by regressing the logarithm of polygon perimeter on the logarithm of polygon area (Mandelbrot 1977; Kummel et al. 1987). The fractal dimension was computed as twice the regression slope. Regression slopes were compared by testing for homogeneity of slope and homogeneity of elevation using an analysis of covariance (Snedecor et al., 1980). Each data set (for vegetation type) included at least 30 data points (wetlands or polygons) for inclusion in the data analysis (for statistical robustness). Tukey’s HSD test was used to compare which pairs of regression slopes were significantly different in the comparison of vegetation type.

<u>Vegetation type comparisons:</u> Wetlands were grouped according to vegetation types, which are described in the attribute table of each digitized map coverage: estuarine intertidal emergent,

estuarine intertidal scrub-shrub, palustrine emergent, palustrine scrub-shrub, and palustrine forested wetlands (Coward in et al., 1979). Estuarine intertidal emergent wetlands are salt marshes (inundated with water of salinity > 0.5 ‰) consisting of herbaceous plants, such as *Spartina alterniflora* (smooth cordgrass), *Distichlis spicata* (spike grass), and *Juncus roemerianus* (needlegrass rush). Estuarine intertidal scrub-shrub wetlands are salt marshes with woody vegetation that is < 6 m (20 feet) in height, such as *Iva frutescens* (big-leaf sumpweed) and *Baccharis halimifolia* (sea-myrtle). Palustrine emergent wetlands are freshwater marshes (inundated with water of salinity < 0.5 ‰) consisting of herbaceous plants, such *as Spartina patens* (saltmeadow cordgrass), *Scirpus californicus* (California bulrush), and *Phragmites australis* (common reed). Palustrine scrub-shrub wetlands are freshwater marshes with woody vegetation that is < 6 m (20 feet) in height, such as *Salix nigra* (black willow) and *Sapium sebiferum* (Chinese tallow). Palustrine forested wetlands are freshwater marshes with woody vegetation that is > 6 m (20 feet) in height, such as *Taxodium distichum* (bald cypress), *Fraxinus pennsylvania* (green ash), and *Acer rubrum* (red maple).

<u>Land use comparsions</u>: Wetland polygons consisted of polygons from the vegetation types described above. Non-wetland polygons consisted of all upland areas in the data set.

<u>Location comparisons:</u> Estuarine wetlands consist of salt and brackish marshes, including coastal wetlands, and palustrine wetlands are inland or freshwater wetlands (White et al., 1993; Mitsch and Gosselink, 1993; Moulton et al., 1997).

<u>Size comparisons</u>: All wetland patches (n=3101) were ranked according to area. The largest third (1000) was designated as large wetlands, and the smallest third (1000) was designated as small wetlands.

<u>Comparison of level of disturbance:</u> Human impact can be evidenced by the amount of transportation corridors (roads and railroads). The most recent NHI wetland coverage (1989) was overlaid with digital TxDOT road coverages. The roads coverage from TxDOT was buffered to explore the extent of road effects on wetland boundaries, using: a 10-m buffer, a 25-m buffer, a 50-m buffer, and a 100-m buffer. The wetland coverage was then overlaid with a buffered road coverage, resulting in a coverage containing segments of wetland boundary which fell into the road buffer. The level of disturbance surrounding a patch of wetland was defined as the percent of the wetland perimeter that had clear evidence of impact (a wetland boundary which fell into

the road/railroad buffer region). The percentage of impact was quantified as a ratio: the length of the impact divided by the total perimeter of the wetland polygon. This method of measuring the level of disturbance was similar to that described in Rex and Malanson (1990). Wetlands were then ranked according to the percentage of impact, divided into thirds, and classified into a low, medium, or high category, based on the ranking.

RESULTS

Fractal dimension of wetlands. There was a significant difference in the fractal dimensions of wetlands with different vegetation types (F=3.97, df=4, P=0.0032; Table 1). The palustrine forested wetlands had the highest fractal dimension (D=1.358), and the estuarine scrub-shrub wetlands had the lowest fractal dimension (D=1.224). Comparisons of the fractal dimensions of the wetland types divided the wetlands into three groups: estuarine intertidal emergent and palustrine emergent; palustrine scrub-shrub and palustrine forested; and estuarine scrub-shrub. The intercepts of the regressions (used in calculating the fractal dimensions of wetlands of different vegetation types) were significantly different (F=12.90, df=4, P=0.0001). There was no significant difference in the fractal dimensions of estuarine and palustrine wetlands (F=0.11, df=1, P=0.7431, Table 2). However, there was a significant difference in the intercepts of these two regressions (F=37.06, df=1, P=0.0001). The fractal dimension of the 1000 smallest wetland polygons (area=196 m^2 to 3782 m^2; D=1.136) was significantly smaller than the fractal dimension of the 1000 largest wetland polygons (area=17,016 m^2 to 37,609,893 m^2; D=1.344; F=46.32, df=1, P=0.0001; Table 3). There was no significant difference in the intercepts of the regressions (F=0.04, df=1, P=0.8381). Wetlands had a significantly higher fractal dimension (1.324) than non-wetlands (1.228; F=211.74, df=1, P=0.0001; Table 4). There was also a significant difference in the intercepts of these regressions (F=13.94, df=1, P=0.0002).

Wetlands with a high degree of impact had a significantly lower fractal dimension than wetlands with a low degree of impact; this trend was also present in wetlands when I used a 10, 25, 50, or 100-meter buffer (Tables 5-8; cf. 1.366 vs. 1.254 in Table 6). In addition, wetlands that were classified as having a low level of impact did not have a significantly different fractal dimension when the size of the road buffer changed (F=0.08, df=3, P=0.9728; Table 9). This was also true for wetlands classified as having a medium level of impact and a high level of impact (Tables 10-11).

DISCUSSION

Wetlands had a significantly higher fractal dimension than non-wetlands (Table 7). This difference may be related to the fact that wetlands are formed by natural processes which form complexly folded boundaries. Certain topological and hydrological patterns, such as tides and storm surges, may be generated by diffusion processes which result in a convoluted boundary. On the other hand, non-wetlands are more often bounded by residential areas, agricultural fields, road constructions, and drainages which shape the boundaries in a more linear fashion.

The fractal dimensions of different wetland types also differed (Table 1). One conclusion might relate this difference to the location (or salinity level) of the wetlands. However, the fractal dimension of coastal wetlands was not significantly different than the fractal dimension of inland wetlands (Table 2). Data on the vegetation types indicate that wetlands with woody vegetation had a significantly different fractal dimension than wetlands with herbaceous vegetation. Palustrine forested wetlands usually occur along rivers and streams, whose edges were formed by natural factors. Palustrine forested and palustrine scrub-shrub wetlands may require more nutrients to support their higher biomass than the herbaceous plants do; thus, a longer and more complex edge may facilitate the flow and intake of nutrients into their systems.

Fractal dimension differed for small patches compared with large patches of wetlands (Table 3), supporting the results of previous studies (Kummel et al., 1987; Bradbury et al., 1984) and suggesting that two separate processes -- natural and anthropogenic -- may have operated to create these patch shapes. These differences may imply changes in spatial scale in the underlying processes that control the shape complexity of wetland patches in the Galveston Bay (Kummel et al., 1987).

There was a significant difference between the intercepts of the regressions slopes used to calculate the fractal dimension for all comparisons except for those comparing size and road buffers (Tables 3, and 9-11, respectively). In creating fractal landscapes, Milne (1992) asserted that it was important to interpret the intercepts of the regression slopes, which “relates to the sheer preponderance of the pattern relative to the extent of the map” within which the pattern occurs. A small value for the intercept with a high fractal dimension indicates a relatively compact pattern occupying a small portion of the study region. A high value for the intercept with a low fractal dimension indicated a highly dispersed pattern spanning a majority of the study area. More research into the interplay between the intercept and the fractal dimension is needed to examine the topic of scale extrapolation.

Wetlands that were heavily impacted by human disturbance had a significantly lower fractal dimension than wetlands that were not as heavily impacted (Tables 5-8). This was explained by the fact that heavily impacted wetlands usually were bounded by more roads and canals which had a rectilinear edge, exhibiting a lower fractal dimension. Previous studies have shown that the presence of human disturbance can affect the fractal geometry of a patch (Forman et al., 1986; Kummel et al., 1987). The present study examined the effects of different levels of impact (roads and railroads) on wetland boundaries. Roads have a significant impact on wetland ecosystems because they may reduce regional biodiversity by impeding migration of small mammals between local populations (Merriam et al., 1989), modifying wetland hydrology and siltation patterns (Andrews, 1990), increasing the amount of edge in habitat patches (Soule, 1992), increasing mortality through roadkills (Fahrig et al., 1995), facilitating the invasion of exotic species (Lonsdale and Lane, 1994), and/or increasing human access to wildlife habitats (Young, 1994). Findlay et al. (1997) found that species richness of plants, birds, and herptiles decreased as the density of paved roads surrounding southern Ontario wetlands increased.

Wetlands affected by the same level of impact did not have significantly different fractal dimensions as the size of the road buffer increased (Tables 9-11). It seems that the effect of roads on wetland boundaries does not change as area between them increases to 100 meters. Further research should be conducted to examine how different levels of impaction affect wetland cycles, community composition, and productivity.

Since different levels of human disturbance (Tables 5-8), different vegetation types (Table 1), and different wetland sizes (Table 3) cause landscapes with different fractal dimensions, it would be interesting to examine whether different generative processes truly form landscape features with different fractal dimensions, and how these processes contribute to patch shape. A mathematical model describing the relative contribution of each factor to overall wetland patch shape may be helpful in identifying the principal components of wetland change. This information would be useful for developing studies to determine interrelationships between ecological, hydrological, and anthropogenic processes operating at different spatial scales. For instance, how could fractal dimension serve as an index of the overall health of wetlands? Or, should fractal dimension be an important parameter in the design of coastal wetlands? Minello et al. (1994) stated that "both marsh surface elevation and edge should be considered when designing salt marsh habitats," since edge habitat is used by high densities of nekton in coastal salt marshes. Data show that vegetation along the marsh edge is used to a greater extent than inner marsh habitat (Peterson and Turner, 1994). Since the fractal dimension is a more accurate description of edge length, perhaps fractal wetlands offer a practical solution to the design of new wetlands. Further research should define and resolve to these processes and form the basis of recommendations about how wetlands can be better managed as natural resources and nature reserves.

Table 1. Comparison of Fractal Dimension *D* of Wetlands by Vegetation Type (Different superscripts indicate significant difference)

Wetland Type	*D*	a	R-Squared	Residual MS	n
(a) Estuarine Intertidal Emergent	1.322^{1}	0.2141	0.97	0.05474	1164
(b) Estuarine Scrub-Shrub	1.224^{2}	0.5707	0.95	0.04177	108
(c) Palustrine Emergent	1.314^{1}	0.1934	0.96	0.05511	1267
(d) Palustrine Forested	1.358^{3}	0.0049	0.95	0.06328	243
(e) Palustrine Scrub-Shrub	1.340^{3}	0.0790	0.94	0.04634	314

Comparison of slopes: F=3.97, P=0.0032 (significant difference)
Comparison of intercept (a): F=12.90, P=0.0001 (significant difference)

Table 2. Comparison of Fractal Dimension *D* of Wetlands by Location

Location	*D*	a	R-Squared	Residual MS	n
Coastal	1.3204	0.2184	0.97	0.05452	1273
Inland	1.3234	0.1520	0.95	0.05480	1826

Comparison of slopes: F=0.11, P=0.7431 (no significant difference)
Comparison of intercept (a): F=37.06, P=0.0001 (significant difference)

Table 3. Comparison of Fractal Dimension *D* of Wetlands by Size

Size	*D*	a	R-squared	Residual MS	n
Small	1.136	0.9000	0.90	0.01096	1000
Large	1.344	0.0741	0.89	0.11244	1000

Comparison of slope: F=46.32, P=0.0001 (significant difference)
Comparison of intercept (a): F=0.04, P=0.8381 (no significant difference)

Table 4. Comparison of Fractal Dimension *D* of Wetlands vs. Non-Wetlands

Category	*D*	a	R-squared	Residual MS	n
Non-Wetlands	1.228	0.5617	0.97	0.2860	2881
Wetlands	1.324	0.1700	0.96	0.5531	3100

Comparison of slopes: F=211.74, P=0.0001 (significant difference)
Comparison of intercept (a): F=13.94, P=0.0002 (significant difference)

Table 5. Comparison of Fractal Dimension *D* of Wetlands Impacted by a 10-meter Road Buffer

Category	*D*	a	R-squared	Residual MS	n
Low impact	1.366	0.0244	0.97	0.08032	210
Medium impact	1.310	0.2500	0.95	0.04334	210
High impact	1.254	0.3740	0.95	0.02538	209

Comparison of slopes: F=7.38, P=0.0007 (significant difference)
Comparison of intercept (a): F=23.08, P=0.0001 (significant difference)

Table 6. Comparison of Fractal Dimension *D* of Wetlands Impacted by a 25-meter Road Buffer

Category	*D*	a	R-squared	Residual MS	n
Low impact	1.376	-0.0330	0.97	0.07891	220
Medium impact	1.316	0.2319	0.95	0.04320	220
High impact	1.254	0.3769	0.95	0.02475	220

Comparison of slopes: F=9.46, P=0.0001 (significant difference)
Comparison of intercept (a): F=21.73, P=0.0001 (significant difference)

Table 7. Comparison of Fractal Dimension *D* of Wetlands Impacted by a 50-meter Road Buffer

Category	*D*	a	R-squared	Residual MS	n
Low impact	1.376	-0.023	0.97	0.0795	235
Medium impact	1.318	0.234	0.96	0.0406	235
High impact	1.236	0.456	0.96	0.02313	234

Comparison of slopes: F=13.42, P=0.0001 (significant difference)
Comparison of intercept (a): F=28.05, P=0.0001 (significant difference)

Table 8. Comparison of Fractal Dimension *D* of Wetlands Impacted by a 100-meter Road Buffer

Category	*D*	a	R-squared	Residual MS	n
Low impact	1.374	-0.0159	0.96	0.09199	265
Medium impact	1.300	0.3161	0.95	0.03713	265
High impact	1.238	0.4477	0.95	0.02449	265

Comparison of slopes: F=13.54, P=0.0001 (significant difference)
Comparison of intercept (a): F=29.86, P=0.0001 (significant difference)

Table 9. Comparison of Fractal Dimension *D* of Wetlands That Are Low Impacted with Various Road Buffers

Category	*D*	a	R-squared	Residual MS	n
10-meter buffer	1.366	0.0244	0.97	0.08032	210
25-meter buffer	1.376	-0.0330	0.97	0.07891	220
50-meter buffer	1.376	-0.023	0.97	0.0795	235
100-meter buffer	1.374	-0.0159	0.96	0.09199	265

Comparison of slopes: F=0.08, P=0.9728 (no significant difference)
Comparison of intercept (a): F=0.08, P=0.9729 (no significant difference)

Table 10. Comparison of Fractal Dimension *D* of Wetlands That Are Medium Impacted with Various Road Buffers

Category	*D*	a	R-squared	Residual MS	n
10-meter buffer	1.310	0.2500	0.95	0.04334	210
25-meter buffer	1.316	0.2319	0.95	0.04320	220
50-meter buffer	1.318	0.2340	0.96	0.04060	235
100-meter buffer	1.300	0.3161	0.95	0.02449	265

Comparison of slopes: F=0.41, P=0.7441 (no significant difference)
Comparison of intercept (a): F=1.18, P=0.3180 (no significant difference)

Table 11. Comparison of Fractal Dimension *D* of Wetlands That Are High Impacted with Various Road Buffers

Category	*D*	a	R-squared	Residual MS	n
10-meter buffer	1.250	0.4451	0.96	0.02837	209
25-meter buffer	1.204	0.3769	0.96	0.02238	220
50-meter buffer	1.236	0.4566	0.96	0.02324	234
100-meter buffer	1.238	0.4487	0.95	0.02442	265

Comparison of slopes: F=1.17, P=0.3207 (no significant difference)
Comparison of intercept (a): F=3.30, P=0.0199 (no significant difference)

REFERENCES

Anderson, J.R., E. E. Hardy, J. T. Roach, and R. E. Witman. 1976. A land use and land cover classification system for use with remote sensor data. U. S. Geological Survey, Prof. Paper 964.

Andrews, A. 1990. Fragmentation of habitat by roads and utility corridors: a review. Australian Zoologist. 26: 130-141.

Atlas, R. M., and R. Bartha. 1981. Microbial Ecology: Fundamentals and Applications. Addison-Welsy, Reading, Mass.

Baker, W. L. 1989. Macro- and Micro-scale influences on riparian vegetation in Western Colorado. Annals of the Assoc. Amer. Geographers. 79: 65-78.

Benton, A. R., W. Sell, and C. A. Clark. 1979. Monitoring and mapping of Texas coastal wetlands: Galveston Bay and Sabine areas. Technical Report RBC-102.

Bradbury, R. H., R. E. Richest, and D. G. Green. 1984. Facials in ecology: methods and interpretation. Marine Ecology - Progress Series. 14: 295-295.

Brand, F., and J. Cohen. 1987. Environmental correlates of food chain length. Science. 238:956-960.

Broader, J. A., L. N. May, A. Rosenthal, and J. G. Gosselink. 1989. Modeling future trends in wetland loss and brown shrimp production in Louisiana using thematic mapper imagery. Remote Sensing of Environment. 56: 871-886.

Bruce, K. A., G. N. Cameron, P. A. Harcombe, G. Jubinsky. 1997. Introduction, Impact on Native Habitats, and Management of a Woody Invader, the Chinese Tallow Tree, Sapium sebiferum. Natural Areas Journal. 13: 255-260.

Burgess, R. L., and D. M. Sharpe. (eds) 1981. Forest Island Dynamics in Man-Dominated Landscapes. Springer, New York.

Burrough, P. A. 1981. Fractal dimensions of landscapes and other environmental data. Nature. 294: 240-242.

Coward in, L. M., V. Carter, F. C. Golet, and E. T. LaRoe. 1979. Classification of Wetlands and Deepwater Habitats of the United States. U.S. Fish & Wildlife Service Pub. FWS/OBS-79/31, Washington, D.C.

Craig, N. J., R. E. Turner, and J. W. Day Jr. 1979. Land loss in coastal Louisiana (U.S.A.). Environmental Management. 3:133-144.

Dahl, T. E. 1990. Wetland Losses in the United States, 1780s to 1980s. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D. C.

Dahl, T. E., and C. E. Johnson. 1991. Wetland Status and Trends in the Conterminous United States Mid-1970s to mid-1980s. U.S. Department of Interior, Fish and Wildlife Service, Washington, D.C.

Diamond, J. M. 1975. The island dilemma: lessons of modern biogeographic studies for the design of natural reserves. Biological Conservation. 7: 129-145.

Fahrig, L., J. H. Pedlar, S. E. Pope, P. D. Taylor, and J. F. Wegner. 1995. Effect of road traffic on amphibian density. Biological Conservation. 72: 1-6.

Findlay, C. S., and Jeff Houlahan. 1997. Anthropogenic correlates of species richness in southern Ontario wetlands. Conservation Biology. 11: 1000-1009.

Forman, R. T. T., and M. Godron. 1986. Landscape Ecology. John Wiley & Sons, New York.

FRAGSTATS. Version 2.0. 1994. McGarigal, K, and B. J. Marks. FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure.

Frayer, W. E., T. J. Monahan, D. C. Bowden, F. A. Graybill. 1983. Status and Trends of Wetlands and Deepwater Habitat in the Conterminous United States, 1950s to 1970s. Dept. of Forest and Wood Sciences, Colorado State University, Fort Collins.

Frissell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management. 10: 199-214.

Frontier, S. 1987. Applications of fractal theory to ecology. P. and L. Legendre, eds. Developments in Numerical Ecology. Springer-Verlag, Berlin. pp. 335-378.

Gagliano, S. M. 1973. Canals, dredging, and land reclamation in the Louisiana coastal zone. Hydrologic and Geologic Studies of Coastal Louisiana, Rept. No. 14. Coastal Resources Unit. Center for Wetland Resources, Louisiana State University, Baton, Rouge, LA.

Gagliano, S. M., K. J. Meyer-Arendt, and K. M. Wicker. 1981. Land loss in the Mississippi River deltaic plain. Trans. Gulf Coast Assoc. Geol. Soc. 31:295-300.

Gibbs, J. P. 1993. Importance of small wetlands for the persistence of local populations of wetland-associated animals. Wetlands. 13: 25-31.

Glaser, P. H. 1987. The Ecology of Patterned Boreal Peatlands of Northern Minnesota: A Community Profile. U. S. Fish and Wildlife Service Report. Report 85 (7.14) Washington, D.C.

Gosselink, J. G., and R. E. Turner. 1978. The role of hydrology in freshwater wetland ecosystems. Freshwater Wetlands: Ecological Processes and Management Potential. R. E. Good, D. F. Whigham, and R. L. Simpson, eds. Academic Press, New York. pp. 63-78.

Gosselink, J. G., M. M Brinson, L. C. Lee, and G. T. Auble. 1990a. Human activities and ecological processes in bottomland hardwood ecosystems: the report of the ecosystem workgroup in Ecological Processes and Cumulative Impacts: Illustrated by Bottomland Hardwood Wetland Ecosystems. J. G. Gosselink, L. C. Lee, and T. A Muir eds. Lewis Publishers, Chelsea, Mich. pp. 549-598.

Hanski, I., and M. E. Gilpin. 1991. Metapopulation dynamics: brief history and conceptual domain. Biological Journal of the Linnean Society. 42: 3-16.

Hastings, H. M., R. Pekelney, R. Monticciolo, D. Vun Kannon, and D. Del Monte. 1982. Time scales, persistence and patchiness. BioSystems. 15: 281-289.

Henderson, M. T., G. Merriam, and J. Wegner. 1985. Patchy environments and species survival: chipmunks in an agricultural mosaic. Biol. Conserv. 31: 95-105.

Holland, M. M, P. G. Risser, and R. J. Naiman. 1991. Ecotones: The Role of Landscape Boundaries in the Management and Restoration of Changing Environment. Chapman & Hall, New York.

Johnson, R. R., and C. H. Lowe. 1985. On the development of riparian ecology. Riparian Ecosystems and Their Management. R. R. Johnson et al., U. S. Department of Agriculture, Forest Service, General Tech. Report RM-120. Washington, D. C. pp. 112-116.

Johnston, C. A. 1991. Sediment and nutrient retention by freshwater wetlands: effects on surface water quality. Critical Reviews in Environmental Control. 21: 491-565.

Kent, C., and J. Wong. 1982. An index of littoral zone complexity and its measurement. Canadian Journal of Fisheries and Aquatic Science. 39: 847-853.

Korcak, J. 1938. Bull. Inst. Int. Stat. III. 295-299.

Kummel, J. R., R. H. Gardner, G. Sugihara, R. V. OÕNeill, and P. R. Coleman. 1987. Landscape patterns in a disturbed environment. Oikos. 48: 321-324.

Lam, N. S. N., and L. De Cola. 1993. Facials in Geography. Prentice-Hall, Inc., Englewood, New Jersey.

Laurance, W. F. 1991. Edge effects in tropical forest fragments: application of a model for the design of nature reserves. Biological Conservation. 57: 295-219.

Laurance, W. F. and E. Yensen. 1991. Predicting the impacts of edge effects in fragmented habitats. Biological Conservation. 55: 77-92.

Li, H. and Reynolds, J.F. A simulation experiment to quantify spatial heterogeneity in categorical maps. Ecology, 75(8): 2446-2455.

Loehle, C. 1983. The fractal dimension and ecology. Speculations in Science and Technology, 6: 131-142.

Lonsdale, W.M., and Lane, A.M. 1994. Tourist vehicles as vectors of weed seeds in Kakadu National Park, North Australia. Biological Conservation, 69: 277-283.

Lovejoy, S. 1982. Area-perimeter relation for rain and cloud areas. Science, 216: 185-187.

MacArthur, R.H., and Wilson, E.O. 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, New Jersey.

Mandelbrot, B.B. 1977. Facials: Form, Chance, and Dimension. Freeman, San Francisco.

McGarigal, K., and Marks, B. J. 1994. FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure. Version 2.0.

McLeese, R.L., and Whiteside, E.P. 1977. Ecological effects of highway construction upon Michigan woodlots soil relationships. Journal of Environmental Quality, 6: 467-471.

Meltzer, M.I., and Hastings, H.M. 1992. The use of facials to assess the ecological impact of increased cattle production: case study from the Runde Communal Land, Zimbabwe. Journal of Applied Ecology, 29: 635-646.

Merriam, G., M. Kozakiewicz, M., Tsuchiyia, E. and Hawley, K.. 1989. Barriers as boundaries for metapopulations and demes of *Peromyscus leucopus* in farm landscapes. Landscape Ecology, 2: 227-235.

Millar, J.B. 1971. Shoreline-area ratio as a factor in rate of water loss from small sloughs. Journal of Hydrology, 14: 259-284.

Milne, B.T. 1991. The utility of fractal geometry in landscape design. Landscape and Urban Planning, 21: 81-90.

Milne, B.T. 1992. Spatial aggregation and neutral models in fractal landscapes. The American Naturalist, 139: 32-57.

Minello, T.J., Zimmerman, R.J. and R. Medina, R. 1994. The importance of edge for natant macrofauna in a created salt marsh. Wetlands, 14:184-198.

Mitsch, W.J., and Gosselink, J.G. 1993. Wetlands. Van Nostrand Reinhold, New York.

Moulton, D.W., Dahl, T.E. and Dall, D. M. 1997. Texas Coastal Wetlands: Status and Trends, Mid-1950s to Early 1990s. U.S. Department of Interior, Fish and Wildlife Service, Southwestern Region, Albuquerque, New Mexico.

Neill, C., and Deegan, L. A.. 1986. The effect of Mississippi River delta lobe development on the habitat community and diversity of Louisiana coastal wetlands. American Midl. Naturalist, 116: 296-303.

O'Neill, R.V., Kummel, J.R., Gardner, R.H., Sugihara, G., Jackson, B., DeAngelis, D.L., Milne, B.T., Turner, M.G., Zygmunt, B., Christiensen, S.W., Dale, V.H., and Graham, R.L.. 1988. Indices of landscape pattern. Landscape Ecology, 1: 153-162.

Palmer, M.W. 1992. The coexistence of species in fractal landscapes. The American Naturalist. 139: 375-397.

Patterson, B.D. 1987. The principle of nested subsets and its implications for biological conservation. Conservation Biology, 1: 323-334.

Peterjohn, W.T. and Correll, D.L. 1984. Nutrient dynamics in an agricultural watershed: observation on the role of a riparian forest. Ecology, 65: 1466-1475.

Peterson, G.W., and Turner, R.E. 1994. The value of salt marsh edge vs. interior as a habitat for fish and decapod crustaceans in a Louisiana tidal marsh. Estuaries, 17:235-262.

Pinay, G., and Decamps, H. 1988. The role of riparian woods in regulating nitrogen fluxes between the alluvial aquifer and surface water: a conceptual model. Regul. Rivers Res. & Management, 2: 507-516.

Rex, K.D. and Malanson, G.P. 1990. The fractal shape of riparian forests. Landscape Ecology, 4: 249-258.

Radforth, N.W. 1962. Organic terrain and geomorphology. Can. Geogr., 6(3-4): 166-171.

Risser, P.G. 1995. The status of the science of examining ecotones. BioScience, 45: 318-325.

Rochefort, L., Vitt, D.H. and Bayley, S. E.. 1990. Growth, production, and decomposition dynamics of Spagnum under natural and experimentally acidified conditions. Ecology, 71: 1986-2000.

SAS. 1982. User's Guide: Statistics. Statistical Analysis Systems, Cary, N.C.

Scaife, W.W., Turner, R.E. and R. Costanza, R. 1983. Coastal Louisiana recent land loss and canal impacts. Environment Management, 7: 433-442.

Schlesinger, W.H. 1978. Community structure, dynamics and nutrient cycling in the Okefenokee cypress swamp-forest. Ecological Monographs, 48: 43-66.

Shaw, S.P., and Fredine, C.G. 1956. Wetlands of the United States, Their Extent, and Their Value for Waterfowl and Other Wildlife. U.S. Department of Interior, Fish and Wildlife Service, Circular 39, Washington, D.C.

Shipley, F.S., and Kiesling, R.W. (eds). 1994. The State of the Bay: A Characterization of the Galveston Bay Ecosystem. The Galveston Bay National Estuary Program. Publication GBNEP-44.

Simberloff, D.S. 1982. Big advantages of small refuges. Natural History, 91: 6-14.

Snedecor, G.W., and Cochran, W. G. 1980. Statistical Methods. The Iowa State University Press, Ames. pp. 384-389.

Soule, M.E., Alberts, A.C. and Bolger, D.T.. 1992. The effects of habitat fragmentation on chapparal plants and vertebrates. Oikos, 63: 39-47.

Southworth, A.D. 1989. Conserving southeastern coastal wetlands. In Chandler, W.J.(ed), Audubon Wildlife Report 1989/1990. pp. 223-257. Academic Press, New York.

Stone, J.H., Bayr, L.M. Jr., and Day, J.W. Jr. 1978. Effects of canals on freshwater marshes in coastal Louisiana and implications for management. *In*: Good, R.E., Whigham, D.F., and Simpson, R.L. (eds), Freshwater wetlands: ecological processes and management potential. pp. 299-320. Academic Press, New York.

Sugihara, G., and May, R.M.. 1990. Applications of fractals in ecology. Trends in Ecology and Evolution, 5: 79-86.

Turner, M.G. (ed). 1987. Landscape Heterogeneity and Disturbance. Springer-Verlag, New York.

Turner, M.G. 1989. Landscape ecology: the effects of pattern on process. *In*: Johnston, R.F., Frank, P.W., and Michener, C.D. (eds), Annual Review of Ecology and Systematics. pp. 171-198. Annual Reviews, Inc., Palo Alto.

Urban, D.L., O'Neill, R.F., and Shugart, H.H. Jr. 1987. Landscape Ecology. BioScience, 37: 119-127.

Walker, D. 1970. Direction and rate in some British post-glacial hydroseres. *In*: Walker, D., and West, R.G. (eds), Studies in the Vegetational History of the British Isles. pp. 117-139. Cambridge University Press, Cambridge.

Warrick, R. A., Barrow, E. M. and Wigley, T. M. L.. 1993. Climate and sea level change: observations, projections and implications. Cambridge University Press, Cambridge.

Weins, J. A. 1992. Ecological flows across landscape boundaries: a conceptual overview. *In:* Hansen, A.J., and Di Castri, F. (eds), Landscape Boundaries: Consequences for Biotic Diversity and Ecological Flows. Ecological Studies 92. pp. 217-235. Springer-Verlag, New York.

White, W.A., Tremblay, T. A., Wermund, E. G. Jr., and Handley, L. R. 1993. Trends and Status of Wetland and Aquatic Habitats in the Galveston Bay System, Texas. Publication GBNEP-31. The Galveston Bay National Estuary Program.

Wickham, J.D., and Norton, D.J. 1994. Mapping and analyzing landscape patterns. Landscape Ecology, 9: 7-23.

Wilson, E.O. and Willis, E.O. 1975. Applied biogeography. *In*:: Cody, M.L., and Diamonds, J.M. (eds), Ecology and the evolution of communities. pp. 522-534. Harvard University Press, Cambridge, MA.

Woodwell, G.M., Hall, C.A.S., Whitney, D.E., and Houghton, R.A. 1979. The Flax Pond ecosystem study: exchanges of inorganic nitrogen between an estuarine marsh and Long Island Sound. Ecology, 60:695-702.

Young, K.R. 1994. Roads and the environmental degradation of tropical montane forests. Conservation Biology, 8: 972-976.

Environmental Risk Assessment Using GIS
Issues of Scale, Resolution, Methodology, and Place

Robert B. McMaster, Eric Sheppard, Helga Leitner, Hongguo Tian, and Jeffrey Matson
Department of Geography
University of Minnesota

Abstract

In studying the relationship between geographic information science/systems (GIS) and society, many, or even most, of the key issues--access, democratization, privacy--play themselves out in the area of environmental risk assessment. In particular, the application of GIS--and spatial methodologies in general--in order to better understand a given population's exposure to technological hazards has increased rapidly over the past decade. Furthermore, the empirical geographical relationship between exposure to toxic chemicals, race and poverty in the US has been shown to vary dramatically depending on the definitions of risk and proximity used, the spatial scope or scale study area examined (neighborhood, inner city, metropolitan region, state or province, or nation state), and the geographical resolution of the data used in the analysis (whether the information analyzed is recorded for counties, municipalities, census tracts, block groups or individual blocks). This environmental risk assessment project focuses specifically on our findings related to issues of scale, resolution, methodology, and place, obtained from an analysis of the Twin Cities Metropolitan region.

Introduction

In looking at the relationship between geographic information science/systems (GIS) and society, many, or even most, of the key issues--access, democratization, privacy--play themselves out in the area of environmental risk assessment. In particular, the application of GIS--and spatial methodologies in general--in order to better understand a given population's exposure to technological hazards has increased rapidly over the past decade. One can point to a plethora of studies that have applied GIS to these human-produced hazards at a variety of scales and resolutions, and have applied significantly different spatial methodologies. Unfortunately, the effects of scale, resolution, and methodology are poorly understood. Additionally, the *place* itself, as constructed through a complicated set of economic, political, social, and economic forces and relationships, has rarely been used to develop a deeper and historical sense of the intermingling of the hazards and communities/neighborhoods in which they exist. Our multiyear project is attempting to unravel some of these vexing problems.

First we provide the context and definitions needed for our research. These definitions evolve from the growing literature on, and series of national-level events related to, environmental justice, including the 1979 environmental justice challenge to the siting of a waste facility by the City of Houston, Texas, the 1986 passage of the SARA Title III legislation, the 1987 environmental justice study by the United Church of Christ, the 1992 establishment of the Office of Environmental Equity, and the 1994 Presidential Executive Order 12898 that requires federal agencies to adopt the principle of environmental justice in programmatic decisions. A thorough review of the environmental justice movement may be found in *Toxics Watch 1995* (Inform, 1995). Since human environment hazards effect the air, water, and soil, this project, based in the Twin Cities metropolitan region, utilizes a series of *hazardous materials* sites, including TRI (Toxic Release Inventory), Superfund, Petrofund, and Land Recycling. However, most of the methodological work involves the *airborne toxic releases* produced by the TRI sites. In order to assess potential exposure to such sites (remembering that while exposure can be roughly calculated, risk itself is based on a complicated set of physiological, historical, and toxicological variables) a set of both geodemographic (based on census data) and institutional populations were used. Whereas *environmental risk assessment* is the more comprehensive term that applies to assessing the relationships between all types of environmental hazards and the humans impacted, *environmental justice or injustice* is the special case of assessing the unequal spatial distribution of hazardous materials sites, where the assumption is that disadvantaged populations (the poor, minorities, the elderly) are disproportionately affected. In our own work, these hazardous materials and population data have been analyzed at a variety of scales and resolutions, ranging from the regional seven-county level to the individual neighborhood level, using a variety of spatial methodologies. This paper focuses specifically on our findings related to scale, resolution, methodology, and place. The empirical geographical relationship between exposure to toxic chemicals, race and poverty in the US has been shown to vary dramatically depending on the definitions of risk and proximity used, the spatial scope or scale study area examined (neighborhood, inner city, metropolitan region, state or province, or nation state), and the geographical resolution of the data used in the analysis (whether the information analyzed is recorded for counties, municipalities, census tracts, block groups or individual blocks). For a fuller analysis of these differences see McMaster, Leitner and Sheppard (1997) and Mohai (1995).

Scale

Geographic scale is the extent of areal coverage used in an inquiry, and is applied in the opposite sense of cartographic scale, which is based on a strict mathematical relationship between map space and earth space using the representative fraction. Geographically, large scale represents larger areas; cartographically, large scale represents smaller areas (but with more detail). The misuse of these two terms causes significant confusion among the many interdisciplinary researchers that work in this area. Our work has looked at environmental justice issues at regional scales--the seven-county metropolitan area--to the individual neighborhood scale. Interpretations of environmental justice change as one looks at the problem from these differing scales.

Our work in Minneapolis suggests that spatial analysis of environmental equity at different scales can help reveal different patterns. In this section, we examine the effect of scale on patterns, and its implications for processes, holding both data resolution and definitions of proximity and vulnerable populations constant.

Figures 1-4 show, respectively, the relationships between race and proximity to Toxic Release Inventory sites for Phillips neighborhood, a racially diverse low income inner city neighborhood of Minneapolis; for the City of Minneapolis itself; for Hennepin County, the metropolitan-wide county within which Minneapolis is located but which also extends to the northern and western suburbs; and the entire seven-county metropolitan region. Although the use of the category "non-white" largely excludes the Hispanic population from our analysis, they represent only a very small presence in three districts of the Twin Cities metropolitan area. For Phillips neighborhood and Minneapolis (Maps 1 and 2), there is no clear relation of proximity between TRI sites and non-white residential concentrations. Indeed for Minneapolis, TRI sites are disproportionately concentrated in northeastern Minneapolis, whereas communities of color are concentrated in the near north side and in south Minneapolis, suggesting that TRI sites in fact are closer to white communities. At the scale of Hennepin County (Figure 3), however, the concentration of communities of color (within Minneapolis) overlaps with the largest cluster of TRI sites. More generally, this map shows that the concentration of both TRI sites and the non-white population decreases westward and northwestward from the City of Minneapolis. One could make the argument, therefore, that at this scale there is evidence of environmental injustice, as both TRI sites (nearly one-half) and populations of color are concentrated in the central city. The regional map provides yet another picture (Figure 4). Here one can detect a toxic “corridor” that tends to follow the Mississippi River as it flows form the northwest, through the City of Minneapolis, and southeastward through southern Ramsey County and St. Paul. This pattern tends to disperse itself as the distance from the river increases. The pattern undoubtedly illustrates the importance of the river in the early industrial history of the region.

Table 1. Percentage of persons living in poverty in Block Groups with and without a TRI site.

	All Metro counties	Anoka	Carver	Dakota	Hennepin	Ramsey	Scott	Washington
With TRI Site	10.54	6.95	4.43	6.96	12.62	12.90	3.67	7.96
Without TRI Site	7.73	5.08	4.83	4.09	8.75	10.91	4.16	4.05

In light of this, it would be erroneous to conclude from the analysis of Phillips neighborhood, or Minneapolis, that there is environmental justice. This would commit a form of ecological fallacy by presuming that processes affecting the spatial distribution of populations and industrial sites only operate at the neighborhood scale. Even with no environmental injustice within Phillips neighborhood, neighborhood residents as a whole may be disproportionately exposed to toxic chemicals compared to others in the central cities; and almost certainly are more exposed than most suburban communities. Tables 1 and 2, which utilize the block-group level census data for the entire seven-county metropolitan region, show comparisons between those block-groups that contain, and do not contain, a TRI site for two census variables: percentage of persons living in poverty and percentage of persons over age 16 unemployed. As would be expected, the most significant differences occur with the poverty comparison where, for instance, in Hennepin County 12.6% of the population that resides in a block-group with a TRI site lives in poverty compared with 8.7% in those block-groups without a TRI site.

Table 2. Percentage of persons over age 16 and unemployed in Block Groups with and without a TRI site.

	All Metro counties	Anoka	Carver	Dakota	Hennepin	Ramsey	Scott	Washington
With TRI Site	2.65	2.67	1.39	2.72	2.91	2.40	2.72	2.55
Without TRI Site	2.61	2.79	2.00	2.14	2.72	2.71	2.14	2.12

These contrasting patterns can be explained by paying attention to the geographical processes of urban development underlying them, and in particular, to which kinds of processes dominate intra-urban patterns at different scales. The positive association between TRI sites and communities of color at the metropolitan scale (i.e., Hennepin County) reflects the history of suburbanization of selected population groups. The central city emerged early as a center of manufacturing, with areas zoned for this purpose. In part in reaction to the environmental consequences of living close to industries and the pollution and congestion associated with them, and taking advantage of new transportation technologies and of Federal subsidization of highway construction and house purchases, better-off white residents began to move out of the central city rapidly after the Second World War. It is well documented that suburbanization was much easier and more attractive for white and middle-income people than for racial minorities and those in poverty. Indeed, suburbanization too often was a means to escape inner city neighborhoods with new African American and other minority inmigrants.

Once suburbs were established, a variety of exclusionary and discriminatory practices often prevented minority and low income groups from moving to the suburbs. Inner city communities became, in turn, places of protection and self-reinforcement for ethnic groups excluded from mainstream society. Although such discriminatory practices have been outlawed (if not eliminated), these processes, reinforced by the limited opportunities for low income households in the market for new housing, have been inscribed in the urban landscape; the presence of communities of color outside the central city remains minimal (Figure 3). The contrasting association between TRI sites and communities of color at different scales can thus be accounted for by the way in which, at a metropolitan scale, suburbanization enabled more wealthy and white households to escape the many ‘problems’ of the inner city, including those of industrial pollution. Industry has also been suburbanizing, as can be seen in the locations of TRI sites in Bloomington, Eden Prairie, Plymouth, St. Louis Park and Brooklyn Park (Figure 3 and 4). Nevertheless, it is harder for dirty or hazardous industries to gain acceptance in suburban communities, and the greatest concentration remains in the central city.

A focus on TRI sites only addresses one way in which the quality of the urban environment varies socially and spatially (Figure 5). It is important to add to this other environmental ‘bads’ such as other fixed sources of potentially toxic chemicals (e.g., Superfund or Petrofund sites), hazardous emissions from ambient sources (especially transportation), soil pollution, radon and lead in houses, and even social risks in certain environments (e.g., undesirable social institutions, and crime). At the same time, inequities in access to urban environmental 'goods' should also be accounted for, such as lakes, green space and bike paths. We have done no systematic analysis of these, but casual observation suggests a similar gradient from low income inner city neighborhoods, to suburban and higher income inner city communities, suggesting that the relationships described above, and the processes generating them are generalizable to a broader group of environmental goods and bads.

The resolution, or granularity, of the data also affects measures of environmental injustice. Our data sets have applied census data at the census tract, block-group, and block levels. We currently are working with parcel-level data, which allows for a household-level analysis. Interpretations can change significantly when address-specific institutional data are included, such as that on schools and day care centers. Further improvements will involve the addition of the 2000 census and approximations of day-time populations.

Geographical Methodology

Whereas most studies use relatively simple measures for establishing the relationship between hazardous materials and at-risk populations, it is clear that more robust methodologies are needed. For instance, relationships are often based on simplistic measures of proximity--hazardous materials sites falling within a block group with a high percentage of minorities--which can not deal with differential toxicity, actual distance from site, and exact meteorological conditions. While GIS-based buffer analysis provides a better measure of proximity, it is nonetheless problematic in that all buffers are normally generated at the same size. The generation of plumes ameliorates this problem, but assumes that exact meteorological conditions are known ahead of time. Additionally, one needs to think carefully about the statistical significance of the existing distribution. To address this concern we have ventured into the realm of Monte Carlo simulation, where, through a randomization process, we can compute the sampling distribution with respect to which the importance of observed statistics can be assessed.

Plume modeling

Let us first provide the results of some preliminary plume modeling. The model we worked with, ALOHA, is an example of a Gaussian Plume Disperson Model. Further details may be found in the ALOHA User's Manual (NOAA, 1996). As with any type of model, certain assumptions must be realized. For the ALOHA model, these include the following:

- All heavy gas releases originate at ground level.
- 95% of the time the wind will not blow the pollutant outside of the dashed exposure line generated by the program.
- The ground below a leaking tank or a puddle is flat.
- The average concentrations will be highest near the release point and along the centerline of any pollutant cloud and will drop off smoothly and gradually in the downwind and cross wind directions.
- A dispersing chemical cloud does not react with the gases that make up the atmosphere.
- The ground below a dispersing cloud is flat and free of obstacles.

Figure 6 and 7 illustrate a series of plumes generated for the City of Minneapolis using two different chemicals—Toluene and Nitric Acid. Figure 6 depicts four plumes representing four different months—January, April, June, and September. For each of the four months, atmospheric conditions representing average temperature, humidity, wind speed, and prevailing wind conditions were applied and the assumption was made that 20% of the chemical was released over a one-hour period. The thirty-eight TRI sites for Minneapolis are identified with the blue point symbols. Comparing the two plume maps shows a striking difference in the area impacted by a potential chemical release. Whereas the composite toxic footprint (based on the four different months) is somewhat localized for toluene (Figure 6), the nitric acid plume covers close to 40% of the City of Minneapolis (Figure 7). Thus one can see that plume modeling represents a significant improvement in measuring potential exposure as compared to a simplistic GIS-based buffer analysis. Our research in exploring the potential for plume modeling is ongoing, with the calculation of plumes for all chemicals within the City, with measuring the impact on various populations, and with computing plumes for the entire Hennepin County Region.

<u>Monte Carlo Simulation</u>

Increasingly, geographers are applying Monte Carlo simulation methods to enable a comparison between an existing spatial pattern, and a theoretical distribution derived from a repeated randomization—called bootstrapping--of the existing data. This allows researchers to apply what is equivalent to a test of significance in determining where an existing statistic falls in relation to a large number (n normally is well above 1000) of "theoretical" statistics. First, we present some basic information on what we define as proximity ratios for the City of Minneapolis.

Table 3. Poverty rates, and proximate, and non-proximate block groups for different races and young children.

Proximity Measure	*Whites*	*African Americans*	*American Indians*	*Asians*	*Hispanics*	*Children <5*	*Total*
Within TRIBG	22	43	64	52	47	47	30
Outside TRIBG	11	40	52	45	26	32	18
Proximity Ratio	1.98	1.07	1.22	1.15	1.80	1.46	1.72

Table 3 is based on a comparison of those census block groups that contain a TRI site (Within TRIBG) and those block groups that do not contain a TRI site (Outside TRIBG). The proximity ratio is the ratio of the within TRIBG poverty rate and the outside TRIBG poverty rate. This represents, of course, a simple measure of proximity—the assumption that a given population within the same enumeration unit as a hazardous site is at higher risk. This same measure—the proximity ratio—was also applied to comparing populations inside and outside of GIS-calculated buffers. We apply the buffering technique since many existing environmental risk studies also use a circular buffer as a measure of proximity. Table 4 compares the block group proximity ratios with those produced through 100-, 500, and 1000-yard buffers.

Table 4. Differences in proximity ratios for four different measures of proximity—block group comparison, and 100-, 500-, and 1000-yard buffers.

Proximity Measure	*Whites*	*African Americans*	*American Indians*	*Asians*	*Hispanics*	*Children <5*	*Total*
Block Groups	1.98	1.07	1.22	1.15	1.80	1.46	1.72
100 yard buffer	1.83	1.27	1.13	1.26	1.66	1.48	1.78
500 yard buffer	2.00	1.28	1.23	1.33	1.48	1.65	1.94
1000 yard buffer	2.22	1.39	1.37	1.71	1.70	2.21	2.21

The critical question, of course, is whether these proximity ratios represent a “significant” difference? To determine this, the Monte Carlo technique was implemented using two methods. For method 1, given that 38 block groups contain a TRI site, these 38 were randomized 1,500 times, with the census statistics—percent white in poverty, percent African American in poverty…) recomputed for each randomization. 1,500 statistics (means, variances) are thus generated to create the theoretical sampling distribution.

Method 2 involved 1,500 randomizations of the actual thirty-eight TRI sites using x-y coordinate pairs. For each randomization, 100-, 500-, and 1000-yard buffers were generated and census statistics inside and outside the buffers were calculated. Table 5 provides the observed poverty levels as percentiles of the simulated distribution for populations near TRI sites. In this table, a value of 99.9% means the observed value was greater than 1,500 of the simulations. One can see that, as the measure of proximity is increased from the block-group level to the 1000 yard buffer level, the percentiles increase. Whereas the “true” mean of Asians in poverty was only greater than 73.1% of the simulations, this same value was greater than 99.9% of the simulations when a 1000-yard buffer was applied. Further details may be found in Sheppard, et. al., 1999).

Table 5. Observed poverty levels as percentiles of the simulated distribution for populations near TRI sites, 1995.

Proximity Measure	*Whites*	*African Americans*	*American Indians*	*Asians*	*Hispanics*	*Children <5*	*Total*
Block Groups	99.9%	86.8%	87.1%	73.1%	80.1%	97.9%	99.9%
100 yard buffer	99.9	95.5	78.8	86.4	94.1	97.8	99.9
500 yard buffer	99.9	98.8	96.6	88.8	94.5	99.9	99.9
1000 yard buffer	99.9	99.9	99.9	99.9	99.9	99.9	99.9

<u>Proximal Spaces.</u>

A final method of assessing the relationship between TRI sites and geodemographics involves the calculation of “Proximal” spaces based on Thiessen Polygons. The characteristic of the Thiessen Polygons, of course, is that any block group within a polygon is closer to that polygon's TRI site than any other TRI site. The structure of the Thiessen Polygons also effectively provides a visualization of the density and distribution of the city's TRI sites.

Place

The neighborhoods in Minneapolis are intricate *places* with complex social and industrial histories. Questions of environmental risk assessment, and justice, have to carefully account for the historical sequencing of industrial toxic sites and migrations of populations in to and out of the neighborhood. This type of work involves a careful historical reconstruction using company records, census data, and historical land records, maps and atlases (Pulido, 1996). In the end, claims of environmental injustice must include a meticulous accounting of the place, not just a modern analysis of easy-to-obtain institutional data. Additionally, we have discovered that the standard institutional databases such as TRI are often insufficient for characterizing a neighborhood's environmental concerns, where many toxic sites must be determined through the acquisition of local neighborhood knowledge.

<u>Local Knowledge and Neighborhood Environmental Inventories</u>

When place-based analyses of environmental justice are carried out at local, neighborhood scales, the possibility exists of a form of geographical analysis that changes the relationship between the researcher and the community researched, with far reaching implications. A common characteristic of spatial analytic approaches, but also of many intensive analyses of places, is that the researcher defines the nature of the problem to be studied and the methods to be used. This restriction is clear in the case of the kinds of spatial analyses described above. The data used for such analyses, typically TRI sites, represent just one source of toxic risk. Local communities have other perspectives on risk that suggest the need for other data. Indeed, one of the most publicized cases of exposure to toxics, Love Canal, was based on neighborhood data collection in the absence of public data about the site in question. Neighborhood analysis revealed the existence of toxic waste that simply had been forgotten about and eventually had dropped out of public discourse over the years since the canal was filled with chemicals and covered over.

Figure 8 provides an example of the impact of including local knowledge. It shows the additional sites of risk identified in an environmental inventory completed by the Marcy Holmes neighborhood organization in Minneapolis in the summer of 1996. The map shows how the addition of locations identified by the community, as places where toxic chemicals may be stored and emitted, creates a denser pattern of potential sources of exposure than that derived from public databases. It also shows the proximity of schools, community centers and daycare centers to these locations, which the community wished to highlight in order to assess the proximity of vulnerable groups to hazardous facilities.

Research in four Minneapolis neighborhoods has revealed that different communities have very different perceptions of the environmental risks that they are exposed to (Table 6); perceptions which may have little to do with proximity to TRI sites (Leitner and Elwood, 1998). Furthermore, many of the sources of toxic risk identified within communities, such as lead pollution in the soil, dry cleaners, old gas stations, or ambient air pollution, are not recorded in standard data bases and thus fall outside current TRI-oriented spatial analyses. Finally, there can exist extreme differences even within a neighborhood about what the greatest problems are, reflecting the differently situated experiences of different types of neighborhood residents (cf. Haraway, 1991). Jeff Osleeb (personal communication) reports that in one Bronx neighborhood the Irish residents, living in a largely paved area of the neighborhood, focus on greening public space, whereas Jewish women are concerned about environmental exposure to the risk of breast cancer.

This is a specific example of the concern that standardized geographical techniques tend to reinforce particular ways of thinking about and representing issues that may be very different from those held by others, and that the widespread adoption of such techniques tends to marginalize other ways of thinking about the issue that may be equally legitimate. This possibility has received much attention in recent debates about ‘GIS and society’ which have focused on the types of representations prioritized by current GIS technologies and practices (cf. Rundstrom, 1993; Sheppard, 1995; Harris and Weiner, 1996). In contradistinction to an approach where the researcher defines the problem at hand, Heiman (1997) advocates ‘science by the people.’ In this view, researchers work in a participatory manner with community residents, giving them voice to express their different views about environmental risks, allowing those different views to be altered in debate amongst them, and providing communities with advice about how to collect the information necessary to determine the geography of potential exposure to these risks in the neighborhood.

Table 6. Localized environmental knowledge gathered from four neighborhoods in Minneapolis.

Neighborhood	*Socioeconomic Status*	*Data Needs*	*Environmental Problems*
Hawthorne	Low Income	Parcel-level data	Toxic corridor along
	Racially diverse	Housing data	river
		Crime data	Traffic-related noise
		Little interest census	& pollution
		Plume analysis	Concern w/
			kondirator
Marcy-Holmes	Mid-income	Parcel-level data	Traffic-related noise
	Large student pop	Traffic volume/	& pollution
	Mostly white	pollution	Toxic industrial
		Toxic sites	activity
			Brownfields
			UofM coal plant
Phillips	Low Income	Parcel-level data	Large industrial sites
	Racially diverse	Housing data	Traffic-related noise
	(Af Am & AM Indian)	Crime data	& pollution
		Toxic sites	High lead levels
			Brownfields in SE
Prospect Park	Mid-income	Traffic volume/	Three superfund sites
	Mostly white	pollution	
		Noise levels	Traffic-related noise
		Crime data	& pollution
		Toxic sites	Noxious and
			unpleasant odors
			Degradation of
			watershed

This approach is similar in spirit to the 'geographical expeditions' organized by Bill Bunge in the 1970s into inner city neighborhoods in Detroit and Toronto, to collect information about the exposure of residents to risks missing from standard databases (Horvath, 1971; Bunge and Bordessa, 1975). It provides residents with the power to influence the questions that should be asked. Such research also can contribute to an activist research agenda by raising awareness in communities, and helping them collect information, and learn to use tools, which can make them more effective in struggles to reduce potential exposure by negotiating with local firms and local government (cf. Bryant, 1995).

Such approaches entail certain difficulties. Perceptions of risk may be very different from actual exposure, and concerns have been expressed about the ability of communities to rigorously collect information, and about the politics associated with activist research. Yet it is easy to overlook that fact that similar problems exist with approaches that are labeled as scientific. TRI data are self-reported, and thus at risk of being inaccurate and subjective. The accuracy of census data used for geo-demographic analysis in many inner city areas has been widely questioned, as has the appropriateness of the very categories used for data collection by the census. There is thus a politics to all databases. By contrast, cases exist where activist data collection by communities has been done with care, and has revealed important and unknown environmental risks. A carefully planned participatory research agenda that gives communities a voice over data collection, challenges local perceptions by engaging them with alternative viewpoints, and provides communities with the help necessary to gain the expertise for rigorous data collection and analysis, has the potential to make such community-based studies no more subject to bias than more standardized approaches.

It should be emphasized that this kind of research is far more effective when carried out within local place-based communities. It is at this scale that varieties of local knowledge about toxics and potential exposure can be recorded and confronted with one another, and that it is possible to undertake systematic data collection of missing information. It is at the local scale that differences in soil types, wind flow paths, and groundwater flow can be most accurately determined. Local knowledge and local archives also can be tapped to develop an understanding of the history of the community necessary to interpret both the types of knowledge people hold, why these conflict, and the reasons for the geographical inequities identified. The availability of portable GPS and GIS technologies and environmental recorders and testing kits also make it possible for neighborhood residents to collect and analyze their own data. Important questions remain unanswered about whether technologies developed for public and private can be adapted adequately to meet the purposes of communities; but at least these are now being asked, for example, by those examining the potential for public participatory GIS.

Summary

This study attempts to tackle multiple problems/issues related to environmental risk assessment, in particular as applied to the assessment of environmental justice. Using the Twin Cities region, we have systematically begun to better understand how the scales and resolutions used, the methodologies applied, and a clearer interpretation of place, can provide an improved understanding of this intrinsically geographical problem. Given the examples provided for the Twin Cities region, it is clear that researchers must acquire a deeper knowledge of the underlying spatial and social processes to fully understand and analyze issues of environmental justice. In addition to the obvious effect of scale (and data resolution), researchers must realize the spatial methodology selected--and even the specific GIS applied--will significantly effect the conclusions. In our own work, we feel the application of Monte Carlo simulation has tremendous potential for demonstrating the significance of the relationship between a given toxic site and the specific targeted distribution. Lastly, the actual expertise, knowledge, and concerns of the neighborhood itself must be considered in environmental justice research. As demonstrated in our neighborhood-based research, neighborhoods themselves often maintain an unofficial yet very detailed knowledge of the sources and histories of toxicity within their spaces.

References

Bryant, B., 1995. “Pollution prevention and participatory research as a methodology for environmental justice.” *Virginia Environmental Law Journal*, 14(4), 589-613.

Bunge, W. W. and R. Bordessa, 1975. *The Canadian Alternative: Survival, expeditions and urban change*. Toronto: York University, Atkinson College.

Haraway, D., 1991. *Simians, cyborgs and women: The reinvention of nature*. New York: Routledge.

Harris, T. and D. Weiner, 1996. *GIS and Society: The Social Implications of how People, Space and Environment are Represented in GIS. Report for the Initiative 19 Specialist Meeting, March 2-5, 1996, Koinonia Retreat Center, South Haven MN*. Technical Report 96-7, City: National Center for Geographic Information and Analysis.

Heiman, M., 1990. “From ‘not in my backyard’ to ‘not in anyone’s backyard!’ Grassroots challenge to hazardous waste facility siting.” *American Planning Association Journal*, 56, 359-362.

Inform, Inc, 1995. *Toxics Watch 1995*. New York, Inform, Inc.

Leitner, H. and S. Elwood, 1998. “GIS and Community-Based Planning: Exploring the Diversity of Neighborhood Perspectives and Needs.” *Cartography and Geographic Information Systems*, 25(2), 77-89.

Sheppard, Eric, H. Leitner, R.B. McMaster, and H. Tian, 1999. “GIS-based measures of environmental equity: Exploring their sensitivity and significance,” *Journal of Exposure Analysis and Environmental Epidemiology*, 9(1), 18-28.

McMaster, R. J., H. Leitner and E. Sheppard, 1997. “GIS-based environmental equity and risk assessment: Methodological problems and prospects.” *Cartography and Geographic Information Systems*, 24(3), 172-189.

Mohai, P., 1995. “The determination of dumping revisited: Examining the impact of alternative methodologies in environmental justice research.” *Virginia Environmental Law Journal*, 14(4), 615-653.

National Oceanic and Atmospheric Administration, and the U.S. Environmental Protection Agency, 1996. *ALOHA User’s Manual.* Washington, D.C.: National Safety Council.

Pulido, L., S. Sidawi and R. O. Vos, 1996. “An archeology of environmental racism in Los Angeles.” *Urban Geography*, 18, 419-439.

Rundstrom, R. A., 1993. “The role of ethics, mapping, and the meaning of place in relations between Indians and whites in the United States.” *Cartographica*, 30(1), 21-28.

Sheppard, E., 1995. “GIS and Society: Towards a Research Agenda.” *Cartography and Geographical Information Systems*, 22, 5-16.

Methodological Issues in
GIS-Based Environmental Justice Research

Jeremy L. Mennis
Department of Geography, The Pennsylvania State University
302 Walker Building, University Park, PA 16802

Abstract

Research in environmental justice investigates whether certain unempowered segments of the population, typically minorities and/or the poor, bear a disproportionate burden of environmental risk. Geographic information systems (GIS) have been used to carry out 'conventional' statistical approaches to environmental justice research whereby the socioeconomic character of communities that host environmentally hazardous facilities are compared to non-host communities. However, methodological issues associated with the conventional approach, such as scale of analysis, continue to make GIS-based statistical assessments of evidence of environmental injustice problematic. GIS has the potential to mitigate many of these methodological problems through mapping/visualization, improved modeling of environmental risk, multi-scale analysis, and raster surface-based representations of population. The case study presented in this paper, concerning environmental injustice in the Philadelphia, Pennsylvania region, demonstrates how raster GIS can facilitate the investigation of the relationship between the distribution of demographic character and the location of hazardous facilities across a variety of scales of analysis. This study finds that in the Philadelphia region there is a clear and predictable relationship between socioeconomic status and proximity to environmentally hazardous facilities that can be interpreted as evidence of environmental injustice. By using GIS to create improved representations of population character and environmental risk, environmental justice research can move beyond the simple statistical comparison of groups of areal units to the exploration of demographic patterns and their spatial relationship with the distribution of environmental risk.

Introduction

Research in environmental justice research investigates whether certain unempowered segments of the population, typically minorities and/or the poor, bear a disproportionate burden of environmental risk. The recent attention on environmental justice can be traced to the release of studies by the U.S. General Accounting Office (GAO, 1983) and the United Church of Christ's Commission for Racial Justice (CRJ, 1987) which reported evidence of racially-based discrimination in the locational distribution of environmentally hazardous sites, such as waste treatment, storage, and disposal facilities. Subsequently, the U.S. Environmental Protection Agency (EPA) has recognized the need to mitigate racial and economic discrimination in the siting of environmentally hazardous facilities (EPA, 1992).

Since these two 'early' studies, much statistically-based research has investigated the issue of environmental justice at local, regional, and national scales. Some of these studies have challenged the claims of the landmark CRJ (1987) study as inaccurate and misleading due to the choice of data and methodology, the Anderton et al. (1994) article being perhaps the most prominent. However, all spatial statistical analyses of environmental justice necessarily make some assumptions concerning the data and methodology used in the study. These data and methodology issues broadly concern two representational themes: 1) the definition and measurement of environmental risk and 2) the definition and spatial delineation of 'community.'

Many studies, including those on both 'sides' of the environmental justice debate, have used geographic information systems (GIS) to manage and structure environmental justice analyses. The benefits of using GIS for environmental justice research are relatively straightforward: Environmental justice is an inherently spatial (and temporal) issue (i.e. what is the spatial relationship between the distribution of people and environmental risk) and GIS provides an efficient environment for the management, analysis, and display of spatial environmental justice data. However, GIS software and GIS data also adhere to particular models of the real world that impose representational and methodological constraints and assumptions on the way environmental justice is understood and therefore analyzed. Many of these methodological issues lie at the foundation of the dispute over the interpretation of statistical evidence of environmental injustice.

Unfortunately, the methodological choices made by environmental justice researchers often go unacknowledged in the interpretation of evidence of environmental injustice. The purpose of this paper is to describe the methodological issues associated with using GIS in environmental justice research so that these issues may be brought to the forefront of the environmental justice debate. While there are no methodological 'solutions' that would create a completely accurate and objective assessment of environmental justice, there is certainly value in incorporating the impacts of the methodological assumptions and constraints into the interpretation of study results. The remainder of this paper reviews GIS-based environmental justice research, highlights primary methodological issues, and proposes a novel environmental justice GIS analysis method that is applied to the Philadelphia, Pennsylvania region as a case study.

GIS and the 'conventional' approach to environmental justice research

In nearly all statistically oriented environmental justice studies, justice is defined according to whether the environmentally hazardous facilities in a particular region are spatially distributed in a socioeconomically equitable versus inequitable manner. This *environmental equity* approach to measuring environmental justice generally entails identifying those communities that host environmentally hazardous facilities (however 'community' may be defined), tallying the racial and economic character of those host communities, and comparing that socioeconomic character to those communities in the region that do not host environmentally hazardous facilities (or to the character of the region at large). Evidence of injustice is then defined as when communities that host environmentally hazardous facilities have significantly higher rates of minority and/or poor persons than non-host communities.

This type of analysis is easily implemented in a GIS using U.S. Bureau of the Census demographic and boundary data, hazardous facility data derived from publicly available U.S. Environmental Protection Agency (EPA) databases, and basic statistical functions found in most commercial GIS packages. For example, Glickman et al. (1995) use GIS to examine evidence of environmental injustice in Allegheny county, Pennsylvania, which includes the city of Pittsburgh. These authors investigate the spatial relationship between statistically derived socioeconomic status and proximity to Toxic Release Inventory (TRI) facilities listed in the EPA TRI database as the indicator of environmental injustice. The TRI database is composed of manufacturers that are required by law to report to the EPA certain toxic chemicals that they release to the environment. While the TRI database is certainly not a comprehensive source of information for a region's environmental risk, as Glickman et al. (1995) note, it is easily obtainable and is often used in environmental justice investigations.

Glickman et al. (1995) define community using five different spatial delineations: census block group, census tract, municipality, and half-mile and one-mile distance 'buffers' around each TRI facility. Averages of census socioeconomic variables for each of these community zone schemes were calculated, including percent minority, percent living below the poverty line, percent unemployed, percent over the age of 65, percent under the age of 5, and other census variables that indicate socioeconomic status or at risk populations. Glickman et al. (1995) report mixed, sometimes contrary results concerning evidence of injustice. For instance, when communities are defined by census block groups or tracts, the percentage of minorities in TRI host communities is not significantly different than that in non-host communities. However, when municipalities, a generally larger areal unit than block groups or census tracts, form the basis for defining community, TRI-host communities have significantly higher proportions of minorities than non-TRI-host communities.

These results mirror those found in other studies and indicate one of the primary methodological issues in environmental justice research, the spatial delineation of community and scale of analysis. The CRJ (1987) study was criticized by Anderton et al. (1994) for using zip codes as the areal unit of analysis because these authors felt that zip codes are too large to capture the spatial relationship between socioeconomic status and proximity to hazardous facilities. Instead, Anderton et al. (1994) use census tracts and find that minorities and the poor are not more likely than non-minorities and the non-poor to live in a census tract that hosts a hazardous facility. They therefore conclude that their study finds no evidence of environmental injustice. Significantly, however, these authors did find a positive relationship between disadvantaged socioeconomic status and proximity to hazardous facilities within a 2.5 mile radius of hazardous facilities.

Goldman and Fitton (1994), in a follow-up to the original CRJ (1987) study, note that although the CRJ (1987) and Anderton et al. (1994) studies reach opposite conclusions about the evidence of environmental injustice, their statistical results suggest a similar, and somewhat startling, demographic pattern: ‘bands’ of socioeconomically disadvantaged persons surrounding a ‘core’ of non-socioeconomically disadvantaged persons concentrated around environmentally hazardous facilities. It is open to debate whether this pattern represents a typical demographic scenario, or even if it does, whether it is, in fact, evidence of environmental injustice. However, it is worth noting that the political motivation to ‘objectively’ demonstrate the existence or non-existence of environmental equity often subsumes and sabotages the analysis itself by biasing the interpretation of analytical results (Pulido, 1996) (Anderton et al. (1994) were funded by Waste Management Incorporated, a waste industry organization).

GIS innovations in environmental justice research

Other GIS-based environmental justice studies attempt to expand the conventional approach to the statistical analysis of environmental justice by using the analytical and display capabilities of GIS. Burke (1993) investigates environmental equity in Los Angeles by using various mapping and visualization schemes to expose the subtle relationships between race, class, population density, and the location of TRI facilities. This author finds evidence that “at a given income level, Hispanics and African-Americans are more likely to be living in close proximity to TRI facilities than whites or Asians” (Burke, 1993: 50).

Typically, environmental justice studies do not attempt to explicitly define the spatial distribution of environmental risk as it is an extremely complex task which differs according to type of facility, type of toxic release, and a host of environmental variables that control the dispersion of the toxic material through the environment. Instead, most studies simply consider the people in the ‘community’ (whether defined by census-based areal unit or distance buffer) that hosts the hazardous facility to be at risk. Chakraborty and Armstrong (1997) use GIS to improve on the definition of at risk population by delineating the areas surrounding each toxic facility that are most likely affected by toxic releases based on a numerical model of toxic dispersion. This model generates a ‘plume’ footprint that defines an at risk area within which socioeconomic variables may be tallied.

Chakraborty and Armstrong (1997) also explore the impact of using different representations of population data in environmental justice analyses. Usually, population data are represented by assignment to polygonal areal units. For distance buffer approaches to defining community or at risk population, polygonal population data in certain GIS packages are considered within the distance buffer if any portion of the areal unit overlaps with the buffer. Chakraborty and Armstrong (1997) refer to this method as the *polygon containment* method. This method may lead to misleading calculation of within-buffer population character since the people living within the overlapping areal unit may in actuality be concentrated in a particular portion of the areal unit that is not actually within the distance buffer.

Zimmerman (1994) notes that GIS methods can be developed to partition the population data assigned to an areal unit that is only partially within a distance buffer into inside-the-buffer and outside-the-buffer portions based on the percentage of the of the areal unit that lies within and without the distance buffer, respectively. Chakraborty and Armstrong (1997) refer to this method as the *buffer containment* method. However, this approach assumes an homogeneous distribution of population throughout the areal unit. An alternative is to represent population data as assigned to an areal unit centroid point, called the *centroid containment* method (Chakraborty and Armstrong, 1997). If the centroid falls within the distance buffer, the population data for the entire areal unit represented by that centroid is considered within the buffer. Again, however, error may occur if the centroid falls within the buffer but the actual population is concentrated in a portion of the areal unit outside the buffer.

I performed a brief comparative test of each of these population representation methods in an analysis of the relationship between percent minority and distance to TRI facility in Delaware county, Pennsylvania.. TRI sites in Delaware county are concentrated in industrial and urban waterfront areas, many of which have high concentrations of minority populations. Significantly, I found that the polygon containment and centroid containment methods tended to under-represent the percentage of minorities living in very close proximity to TRI sites as compared to the buffer containment method. These two former methods were less sensitive to variation in demographic character at close proximities to TRI facilities because they tended to incorporate more distantly located, non-minority populations than the buffer containment method using the same distance buffer. In other words, percent minority calculations at close proximities were diluted by the inclusion of a larger area with lower concentration of minority population. Chakraborty and Armstrong (1997) reported similar findings in their comparison of population representation methods.

Another very prominent issue in environmental justice research, related to the issue of defining community, is that of scale of analysis. Scale of analysis concerns both the scope of analysis, the region that the study covers, and the resolution of analysis, which generally refers to the choice of areal unit at which demographic data is represented and tallied. For instance, the CRJ (1987) study was done at the zip code resolution. However, this definition of resolution is problematic because zip codes (and nearly all census-, or other organization-, based zonation schemes) vary widely in their areal extent; they are typically much smaller in urban areas than in rural areas.

This issue of choice of resolution in spatial analysis is associated with what is called the modifiable areal unit problem (MAUP) in the geographic literature (Openshaw, 1983). The MAUP refers to the fact that different aggregation and/or zonation schemes for spatial data may result in vastly different spatial analysis results. The detrimental impact of the MAUP on the analysis of census data is well established (Fotheringham and Wong, 1991; Openshaw, 1984). The difference in results between the CRJ (1987) and Anderton et al. (1994) studies may be attributed in part to the MAUP.

A number of authors (Anderton et al., 1994; Glickman et al., 1995) argue that there exists an ‘appropriate’ areal unit of analysis, or that evidence of environmental equity must not vary with the scale of analysis in order to be regarded as valid. However, simply assuming that there is such a thing as an ‘appropriate’ unit of analysis for environmental justice research immediately violates the principles of the MAUP. Sui (1999) notes that an environmental justice study done at any one scale or based on one particular areal unit cannot, by definition, produce a reliable indication of environmental justice or injustice; there is no such thing as the single ‘best’ or most ‘appropriate’ scale of analysis in environmental justice research.

A number of authors have suggested that GIS be used to support multi-scale environmental justice analysis (McMaster et al., 1997; Sui, 1999). I argue that the purpose of multiscale analysis is not to find the ‘best’ scale of analysis but to investigate how demographic character and its spatial relationship with environmentally hazardous facilities varies *across* scales. This information may indicate the subtle and complex demographic patterns that lie at the root of the environmental justice debate. As Been (1995) notes, environmental justice is infinitely more complex than disproportionate numbers of hazardous facilities being sited in census tracts (or block groups, municipalities, etc.) with a high percentage of minorities. Rather, environmental injustice should be viewed as a complex intertwining of various socioeconomic characteristics distributed in certain spatial patterns. It should be the goal of environmental justice studies to ‘uncover’ these often ‘hidden’ patterns that are embedded in the social and environmental data that is available.

A case study: environmental injustice in the Philadelphia, Pennsylvania region

Nearly all GIS approaches to environmental justice research have been vector- (as opposed to raster-) based because most commercial GIS are vector-based (although there are a growing number of GIS packages offering raster data handling). In addition, most population and hazardous facility data are also vector-based. However, raster modeling of population offers many advantages. Principally, the raster-based approach to representing population allows for data aggregation to nearly any areal unit, facilitates the exploration of how demographic character varies across scales, and provides the means to create more informative visualizations of the distribution of demographic character (Bracken, 1993; Martin and Bracken, 1991).

Here, I describe a combined vector-raster analysis of environmental justice in southeast Pennsylvania which encompasses the city of Philadelphia (which is identical to Philadelphia county) and its four closest counties in Pennsylvania: Bucks, Delaware, Chester, and Montgomery. The goal of this study is to understand the distribution of socioeconomic character and its spatial relationship to environmentally hazardous facilities. I hypothesize that socioeconomic character has a strong relationship with proximity to hazardous facilities; in other words, the socioeconomic character of a location can be predicted as a function of distance to a hazardous facility. I test this hypothesis by modeling population as a raster surface. This allows for demographic variables that indicate population character to be tallied within a series of distance buffers generated from the hazardous facility locations. Regression is then used to test the strength of the relationship between socioeconomic character and distance to hazardous facility.

Three demographic variables that are often used in environmental justice analyses are used to indicate socioeconomic status in this study: number of minorities, number of people living below the poverty line, and number of people over the age of 25 with a bachelors or graduate degree. These population data were acquired from the U.S. Bureau of the Census at the block group level. Data on facilities that store or release toxic materials in the Philadelphia region were acquired from EPA databases including sites listed in the TRI database as well as treatment, storage, and disposal (TSD) facility sites listed in the Biennial Reporting System (BRS) database. Procedures for improving the locational accuracy of these hazardous facility sites and eliminating redundant database listings were followed according to Scott et al. (1997).

A variety of procedures for generating population surfaces from areal unit demographic data have been proposed including areal weighting (Flowerdew et al., 1991), interpolation from areal unit centroids (Bracken and Martin, 1989), and the use of remote sensing imagery and dasymetric mapping (Langford and Unwin, 1994). Dasymetric mapping is a technique that uses ancillary data to redistribute mapped thematic data in a more accurate and logical way. It is used here to improve upon the methods of population data representation that are typically used in environmental justice research. The dasymetric mapping/raster surface generation method described here is a variation on the method described by Langford and Unwin (1994) and uses urban density classification data derived from satellite remote sensing to redistribute population within the original block group data boundaries. This procedure was carried out using ArcView GIS by Environmental Systems Research Institute (ESRI), Inc.

Urban density data for Pennsylvania were acquired from the Environmental Resources Research Institute (ERRI) at the Pennsylvania State University. These data were photointerpreted from Landsat Thematic Mapper (TM) imagery overlaid with a road network to produce a polygon coverage that partitions the state into areas of high density urban, low density urban, and non-urban. Note that 'density' in this case refers to the degree of urbanization (i.e. development), not population density. While degree of urbanization is by no means a perfect proxy for population distribution (Forster, 1985), its utility in modeling population has been demonstrated in a variety of contexts (Langford et al., 1991; Mesev, 1999).

The urban density classification data were converted from vector to raster format with a grid cell size of 100 meters. This resolution was chosen because it meets the analytical requirements and yet is not so fine that it interferes unduly with processing time. Each grid cell was assigned a population value according to three factors: the population of its host block group, the population density of its urban density classification (derived from empirical measurement), and the percentage of the area of the host block group occupied by its urban density classification. This procedure preserves what Tobler (1979) referred to as the pycnophylactic property: summing the population for all the grid cells within any block group produces the same population figure as that originally assigned to that block group. The raster surface generation calculations were carried out primarily in the ArcView GIS Tables module and can be described mathematically as:

$$PGC_{u.c.b} = (PCT_{u.c.b} * PBG_{b}) / GC_{u.b}$$

Where:

$PGC_{u.c.b}$ = Population assigned to one grid cell with urban density classification *u*, in county *c*, and in block group *b*

$PCT_{u.c.b}$ = Percent of population assigned to urban density classification *u*, in county *c*, and in block group *b*

$GC_{u.b}$ = Number of grid cells (area in 10,000 sq. meter units) of urban density classification *u* in block group *b*

PBG_b = Population of block group *b*

Each demographic variable was distributed homogeneously according to the distribution of the total population for each block group. Surfaces of percent minority, percent living below the poverty line, and percent over the age of 25 with a bachelors or graduate degree were created by dividing the 'count' grids for each of these variables by the grid of total population. For a more thorough description of this areal interpolation technique see Mennis (forthcoming).

Distance buffers around each hazardous facility were created that described the area within 500 meters of a hazardous facility, within 1000 meters, and so on up to 10,000 meters, which encompasses 99.9% of the total population. Percent minority, percent living below the poverty line, and percent over the age of 25 with a bachelors or graduate degree were then tallied within each of these distance buffers. *Cumulative* tallies determine these variables within 500 meters of a hazardous facility, within 100 meters, within 1500 meters, etc. while *zone* tallies determine these variables within 500 to 1000 meters of a hazardous facility, within 1000 to 1500 meters, and so on.

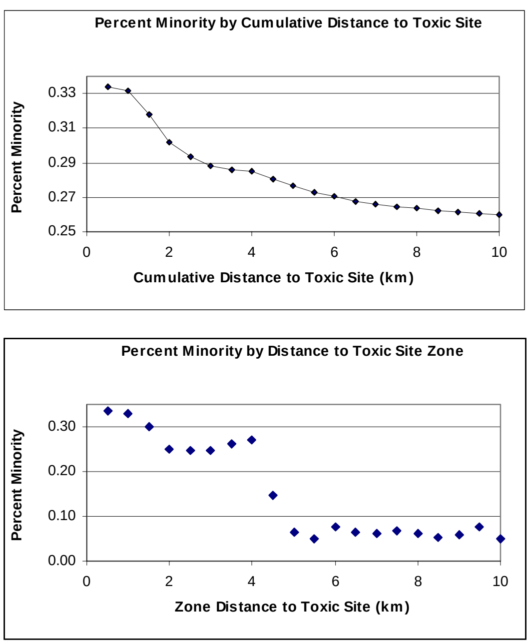


Figure 1. The relationship between percent minority and cumulative distance to hazardous site (top) and zone distance to hazardous site.

The relationship between presence of minorities and distance to hazardous facilities is presented in figure 1. As distance to hazardous facilities increases, percent minority decreases, percent living below the poverty line decreases (not shown), and percent over age 25 with a bachelors or graduate degree increases (not shown). The break in slope at approximately 5000 meters, evident in the graphs of all the variables, is related to the fact that 92.0% of the total population and 98.1% of all minorities live within 5000 meters of a hazardous facility.

Regression tests that predicted percent minority, percent living below the poverty line, and percent over the age of 25 with a bachelors or graduate degree based on cumulative distance to hazardous site up to 5000 meters yielded R^2 values of 0.886, 0.907, and 0.926, respectively. R^2 values for these same variables, but predicted by zone distance to hazardous site up to 5000 meters, yielded values of 0.688, 0.886, and 0.979, respectively. All results were significant at the 0.001 level. Multiple stepwise regression with cumulative distance to hazardous facility up to 5000 meters as the dependent variable and the three demographic variables as independent variables excluded percent minority and percent living below the poverty line and included percent over the age of 25 with a bachelors or graduate degree to yield an R^2 of 0.926 (significant at the 0.001 level). A similar test that predicted zone distance to hazardous site up to 5000 meters excluded percent living below the poverty line and included percent minority and percent over the age of 25 with a bachelors or graduate degree to yield an R^2 of 0.987 (significant at the 0.001 level). Clearly, the poor, minorities, and the lesser educated tend to live in closer proximity to hazardous facilities than the non-poor, non-minorities, and the educated.

These results conjure an image in which each hazardous facility is surrounded by poor, uneducated minorities and that gradually this pattern gives way to wealthier, educated non-minorities as distance to the hazardous facility increases. However, maps that depict the distribution of these variables overlaid with the locations of hazardous facilities demonstrate that this is not at all the case (e.g. figure 2 which shows areas with percent minority greater than the regional mean of 26%). There are, rather, various ‘clusters’ of hazardous facilities that appear to correspond to a variety of interrelated historic, cultural, and infrastructure factors. For instance, many hazardous facilities stretch along the Delaware and Schuylkill Rivers while others are clustered around population centers.

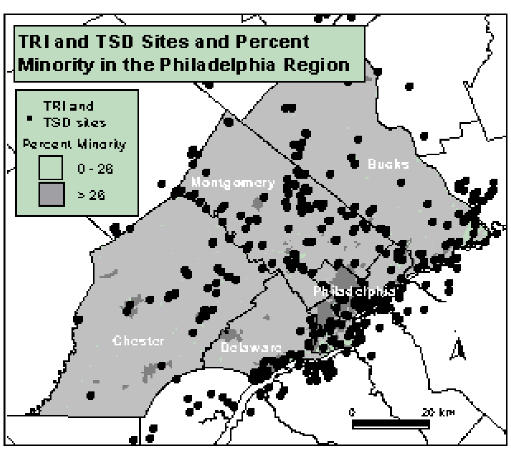
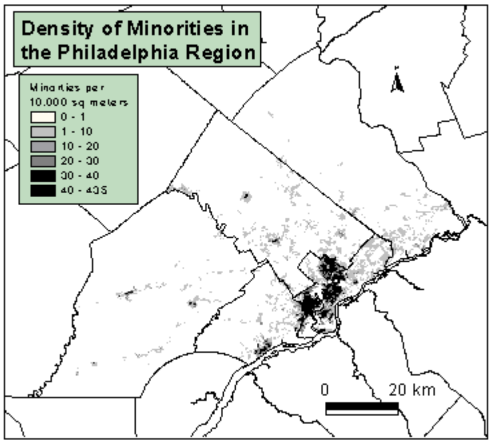


Figure 2. The location of hazardous facilities relative to percent minority.

This apparent, but in fact false, discrepancy between statistical and visual summation can be attributed to the difference between the measurement of percent and density of demographic character. For example, while there are areas outside Philadelphia with high percent minority, nearly all minorities in the Philadelphia region are clustered within certain neighborhoods of Philadelphia (figure 3). While non-minorities are



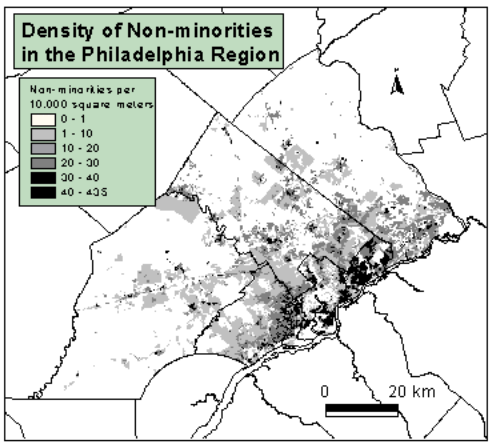


Figure 3. Density of minorities and non-minorities in the Philadelphia region.

also clustered around Philadelphia, they are much less concentrated in specific areas. The same is true with percent and density of people living below the poverty line. Concerning education, it appears that while higher education attainment is concentrated in suburban areas, hazardous facilities are concentrated primarily in urban areas and secondarily in rural areas.

So while hazardous facilities are not necessarily concentrated in poor, uneducated, and minority portions of the greater Philadelphia region, these portions of the population are concentrated in one particular area, the city of Philadelphia. Because the city of Philadelphia is home to one of *many* clusters of hazardous facilities, nearly all those of unempowered socioeconomic status are in relatively close proximity to hazardous facilities compared to other persons of the Philadelphia region. However, there are many non-poor, non-minorities, and educated persons who are also in relatively close proximity to hazardous facilities.

Further statistical analysis and mapping/visualization may reveal other demographic/hazardous facility patterns. For example, spatial autocorrelation measures may indicate the degree of socioeconomic regionalization at a variety of scales. Cluster analysis and point pattern analysis of the hazardous facility data may show a statistical relationship between demographic character and spatial clusters of facilities. Choropleth and bivariate mapping schemes, as well as cartograms, can be used to further visually investigate the demographic patterns embedded in the data.

Conclusion

This paper is intended as both a caution and an encouragement for the use of GIS in environmental justice research. On the caution side, the data representations that are embedded within GIS present potential pitfalls to researchers who do not explicitly acknowledge how GIS data and methods of analysis can control analytical results. While the issue of making explicit an investigation's analytical assumptions exists for nearly any analysis, whether using GIS or not, the ease of use of many GIS often serves to make this issue transparent to the casual user. On the encouragement side, however, GIS provides an environment for creating new and innovative ways of investigating environmental justice. The use of raster representations of population and environmental risk and the use of advanced spatial statistical and visualization techniques hold particular promise in moving environmental justice research forward towards a more exploratory, pattern recognition approach.

References

Anderton, D.L., Anderson, A.B., Oakes, J.M. and Fraser, M.R., 1994. Environmental equity: the demographics of dumping. *Demography*, 31(2): 229-248.

Been, V., 1995. Analyzing evidence of environmental justice. *Journal of Land Use and Environmental Law*, 11(1): 1-35.

Bracken, I., 1993. An extensive surface model database for population-related information: concept and application. *Environment and Planning B: Planning and Design*, 20: 13-27.

Bracken, I. and Martin, D., 1989. The generation of spatial population distributions from census centroid data. *Environment and Planning A*, 21: 537-543.

Burke, L.M., 1993. Race and environmental equity: a geographic analysis in Los Angeles. *Geo Info Systems*, 3(9): 44-50.

Chakraborty, J. and Armstrong, M.P., 1997. Exploring the use of buffer analysis for the identification of impacted areas in environmental equity assessment. *Cartography and Geographic Information Systems*, 24(3): 145-157.

CRJ, 1987. Toxic wastes and race in the United States: a national report on the racial and socioeconomic characteristics of communities with hazardous waste sites, United Church of Christ's Commission for Racial Justice, New York.

EPA, 1992. Environmental equity: reducing risk for all communities, U.S. Environmental Protection Agency, Washington, DC.

Flowerdew, R., Green, M. and Kehris, E., 1991. Using areal interpolation methods in geographic information systems. *Papers in Regional Science: The Journal of the RSAI*, 70(3): 303-315.

Forster, B.C., 1985. An examination of some problems and solutions in monitoring urban areas from satellite platforms. *International Journal of Remote Sensing*, 6(1): 139-151.

Fotheringham, A.S. and Wong, D.W.S., 1991. The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A*, 23: 1025-1044.

GAO, 1983. Siting of hazardous waste landfills and their correlation with racial and economic status of surrounding communities, U.S. General Accounting Office, Washington, CD.

Glickman, T.S., Golding, D. and Hersh, R., 1995. GIS-based environmental equity analysis: a case study of TRI facilities in the Pittsburgh area. In: G.E.G. Beroggi and W.A. Wallace (Editors), *Computer Supported Risk Management*. Kluwer Academic Publishers, Netherlands, pp. 95-114.

Goldman, B.A. and Fitton, L., 1994. Toxic wastes and race revisited, Center for Policy Alternatives, Washington D.C.

Langford, M., Maguire, D.J. and Unwin, D.J., 1991. The areal interpolation problem: estimating population using remote sensing in a GIS framework. In: I. Masser and M. Blakemore (Editors), *Handling Geographical Information: Methodology and Potential Applications*. Longman, London, pp. 55-77.

Langford, M. and Unwin, D.J., 1994. Generating and mapping population density surfaces within a geographical information system. *The Cartographic Journal*, 31: 21-26.

Martin, D. and Bracken, I., 1991. Techniques for modeling population-related raster databases. *Environment and Planning A*, 23: 1069-1075.

McMaster, R.B., Leitner, H. and Sheppard, E., 1997. GIS-based environmental equity and risk assessment: methodological problems and prospects. *Cartography and Geographic Information Systems*, 24(3): 172-189.

Mennis, J.L., forthcoming. Using GIS, spatial statistics, and visualization to investigate environmental justice, Proceedings of the Joint Statistical Meeting of the American Statistical Association, Baltimore, MD, August 8-12, 1999.

Mesev, V., 1999. From measurement to analysis: a GIS/RS approach to monitoring changes in urban density. In: M. Craglia and H. Onsrud (Editors), *Geographic Information Research: Trans-Atlantic Perspectives*. Taylor and Francis, London, pp. 307321.

Openshaw, S., 1983. The Modifiable Areal Unit Problem. *Concepts and Techniques in Modern Geography*, 38. Geo Books, Norwich, UK, 40 pp.

Openshaw, S., 1984. Ecological fallacies and the analysis of areal census data. *Environment and Planning A*, 16: 17-31.

Pulido, L., 1996. A critical review of the methodology of environmental racism research. *Antipode*, 28(2): 142-159.

Scott, M., Cutter, S.L., Menzel, C., Ji, M. and Wagner, D., 1997. Spatial accuracy of the EPA's environmental hazards databases and their use in environmental equity analyses. *Applied Geographic Studies*, 1(1): 45-61.

Sui, D., 1999. GIS, environmental equity analysis, and the modifiable areal unit problem (MAUP). In: M. Craglia and H. Onsrud (Editors*), Geographic Information Research: Trans-Atlantic Perspectives*. Taylor and Francis, London, pp. 41-54.

Tobler, W.R., 1979. Smooth pycnophylactic interpolation for geographic regions. *Journal of the American Statistical Association*, 74: 519-530.

Zimmerman, R., 1994. Issues of classification in environmental equity: how we manage is how we measure. *Fordham Urban Law Journal*, 21(3): 633-670.

Pollution Exposure Index Model
Measures Airborne Pollutants in National Forests

Mike Miller, CH2M HILL
Corvallis, OR

CH2M HILL developed the Pollution Exposure Index (PEI) model for the U.S. Forest Service to measure the exposure of sites located in National Forests to a variety of airborne pollutants. The exposure is calculated for user-specified receptor sites based on pollutant emission rates of sources, seasonal wind frequencies, terrain landforms, and distances between the receptors and pollutant sites.

The source data, which encompasses most of the eastern United States, was pre-processed into a standardized GIS structure for PEI modeling. The data includes; Digital Elevation Model (DEM) and USFS boundary data supplied by the Forest Service, meteorological wind frequencies acquired from the EPA Support Center for Regulatory Air Models (SCRAM) bulletin board system, and pollutant information derived from the EPA Aerometric Information Retrieval System (AIRS) and National Acid Precipitation Assessment Program (NAPAP) inventories.
The PEI model was developed with ESRI's ArcView 3.0 desktop GIS mapping software using its object oriented Avenue programming language. It also accesses the grid-based Spatial Analyst extension package. The model was developed on an NT computer with a Pentium 586/133 processor and 32 mbs of RAM. A large color monitor, at least 14 inches, is recommended.
The PEI model was conceived by and developed under the direction of Bill Jackson of the USFS, Region 8. The model was originally developed as a predictive tool to assess seasonal impacts of controlled forest burns. Recently, it was used for an Ozark Mountain Regional assessment of sulfur, nitrogen and particulate matter. Kent Norville from CH2M HILL's Portland, OR office helped design the model and Mike Miller from CH2M HILL Corvallis, OR programmed the model.

This paper describes the pre-processing procedures and the basic components of the PEI model.

I. Pre-Process Source Data

Source data was acquired for the study area from various sources and pre-processed as GIS data for PEI modeling. The source data are listed as follows:

1. U.S. National Forests in Southern Appalachian Mountains
2. Elevation DEMs
3. Meteorological Stations
4. Regional AIRS Pollutants
5. County Pollutants
6. State Boundaries

The source data were pre-processed as Arc/Info data coverages. The pre-processing tasks include data conversion, modification, and standardization. All the graphic data is standardized to Albers equal area projection, Clark 1866 Spheroid, NAD27 datum measured in single precision meters. The Albers projection parameters are:

1st latitude	= 29 degrees 30 minutes
2nd latitude	= 45 degrees 30 minutes
Central meridian	= -96 degrees
Latitude origin	= 23 degrees

Finally, the coverages were translated as shapefiles because ArcView processes them much faster than coverages.

The remainder of this section contains detailed descriptions of how the source data was pre-processed.

I.1 National Forests and Southern Appalachia Assessment Area

The USFS provided Arc/Info coverages of the Southern Appalachia Assessment Area and seven National Forests contained in the Assessment Area. The National Forests include Cherokee, Jefferson, George Washington, Pisgah, Nantahala, Chattahoochee, and Sumter. Both coverages were translated from latitude/longitude decimal degrees to the Albers projection, then spatially indexed and translated as ArcView shapefiles.

In addition, 250 and 500 kilometer buffer coverages were generated from the Assessment Area coverage and used to define the maximum extent of the pollutant source data and the study area, respectively.

I.2 Elevation DEMs

An elevation DEM file for the study area was purchased from the National Geophysical Data Center. The file contains sample points every 30 seconds or 1 kilometer (1,030 meters) referenced to latitude/longitude decimal degrees and elevations measured in feet. Initially, the DEM file was converted to the Arc/Info grid format (ASCIIGRID), then projected to Albers feet. The 1 km grid was used to interpolate elevations for all the meteorological station and pollutant source points. In addition, it was used to generate a 100 foot contour coverage and shapefile that is accessed by the PEI model to interpolate receptor elevations and derive the highest elevations between the pollutant and the receptor sites.

I.3 Meteorological Stations

Wind frequency data is a critical PEI model input. It is required to identify prevailing winds between pollutant sources and receptors. Wind frequencies are routinely measured and recorded at various National Weather Service (NWS) meteorological stations, generally located at airports. The latitude/longitude coordinates of the stations are also known.

The wind frequency data were retrieved from the Support Center for Regulatory Air Models (SCRAM) section of the US EPA Technology Transfer Network (TTN) bulletin board system and pre-processed. Two ASCII files were developed from the SCRAM data. One file contains information related to the location of the stations and their names. The second lists the seasonal wind frequency data from 16 directions (every 22.5 degrees) for each station. Initially, the first file was converted as an Arc/Info point coverage and projected from latitude/longitude to Albers coordinates. The wind frequencies data in the second file was copied in an Info database table. The station point features and wind frequency table are linked for modeling by station numbers. Appendix I – Meteorological Stations lists the fields in the station locations and wind frequency tables.

Eventually, the elevations of the meteorological stations were interpolated from the 1 kilometer DEM grid.

Meteorological station areas-of-influence (AOIs) were delineated representing homogeneous landforms throughout the PEI study area. The process involved several steps. First, a Thiessen program was run to apportion the station points into regions. The result was a polygon coverage of regions, each containing a single station. Each region has the unique property that any location within a region is closer to the region's station than any other station. Next a plot of the stations, station regions, and topographic relief was produced. The Thiessen generated region boundaries were compared to the relief and the region boundary modifications were drafted on the plot to reflect areas of homogeneous landforms. The boundaries of the region coverage were revised as drafted, resulting in the station AOI coverage which is used for PEI modeling. Finally, the coverages were translated as ArcView shapefiles and spatially indexed for modeling. See Appendix II – Meteorological Station Areas of Influence for description of database.

I.4 Pollutant Data

The PEI model has been developed to measure the emissions of six pollutant types, which consist of carbon monoxides (CO), nitrous oxides (NOx), sulfur dioxides (SO2), volatile organic compounds (VOCs), particulate matter (PM10), and lead (PB).

The pollutant information was derived from the U.S. EPA Aerometic Information Retrieval System (AIRS) database and reformatted by CH2M HILL for model input. Note that although the EPA maintains the database, each individual state agency can change, update, or delete individual source entries. An ASCII file was prepared for each pollutant type from the AIRS data. It contains information related to the locations, names, and emission rates of the pollutants. The latitude/longitude coordinate locations were used to generate Arc/Info point coverages, which were subsequently projected to Albers coordinates. Eventually, the elevations of the pollutant sites were interpolated using the 1 km DEM grid. See Appendix III - EPA AIRS Pollutant Source Points for a description of the AIRS data.

In addition, County Area and Point CO, NOx, SO2, VOC, and PM10 pollutant data were developed from EPA's National Acid Precipitation Assessment Program (NAPAP) inventories. The Area sources refer to roads, landfills, etc, while the Points are usually industrial facilities. A County shapefile, comprised of points located near the centers of their respective Counties, was coded with a unique state/county identifier which is related to the County Area and Point databases containing the pollutant type, category sources, emission rates, and other attributes. The elevations of the County points were also interpolated using the 1 km DEM grid. Appendix IV – National Acid Precipitation Assessment Program describes the County source data and Appendix V – NAPAP County Area and Point Pollutant Tables lists the database structures required for PEI modeling.

Points representing each County were generated, transformed to Albers, spatially indexed, and translated as ArcView shapefiles. The County Area and Point attribute data was also translated as dBASE files. The County points are related to their respective Area and Point attributes by State name and County FIP identifiers.

I.5 State and County Boundaries

The State boundaries, which serve as a display reference, cover most of the eastern United States. The coverage delineates a total of 27 States and the District of Columbia. It was derived from ESRI's ArcUSA data source.

The States coverages were transformed to Albers, then translated as ArcView shapefiles for modeling.

II. PEI Model

The PEI Model is comprised of ArcView display and analysis scripts written in the Avenue object oriented programming language. It is run using the PEI Model pulldown menu, illustrated below. To execute the model, simply select the menu components, from top to bottom, and respond to queries.

PEI Model Pulldown Menu:

Create Receptors
Add Sites
Delete Sites

Interpolate Elev

Identify Source Data

PEI Model
Change Threshold

Print Variables

About PEI Model

The following describes the basic functional components of the model from startup to shutdown. Except for model startup, the descriptions directly relate to the pulldown menu components and include general use guidelines, functionality overview, view, theme and variable lists, and scripts names. There are one or more scripts associated with each menu component.

II.1 PEI Model Startup

The PEI Model is launched when the ArcView PEI3.APR project is activated. Initially, the U.S. Forest Service symbol and a brief overview of the model are displayed. When the user is ready to start modeling, the domain view is opened displaying the State and Southern Appalachian National Forest boundaries for the PEI study area, which includes most of the eastern half of the U.S. The meteorological stations areas-of-influence theme is also added to the domain view. Next, the default values are set for variables required for modeling. Global variables identifying pathnames to the GIS data, scripts, projects, and other file subdirectories in the pei3 directory are defined. Then, default data variables including; the season, mixing heights, pollutant type emission rate thresholds, receptor/pollutant search radius, terrain evaluation mode, and the PEI threshold percentage, are set. Mixing heights, directly related to the seasons, are automatically set for the specified season.

The global variable names and their default values are listed as follows:

Variable Description	Variable Name	Default Value
Directory Pathnames:		
PEI home directory	_peiDir	d:\data\pei3
Shape file directory	_peiShps	_peidir\shapes
ArcView projects directory	_peiProjs	_peidir\projs
Receptors directory	_peiRcps	_peidir\receptor
ArcView scripts directory	_peiScrp	_peidir\scripts
Text directory	_peiText	_peidir\text
Database tables directory	_peiTab	_peidir\tables
Temporary work directory	_peiTemp	_peidir\temp
Elevation grid file	_peiElevGd	_peidir\grid\ele1km
Data Variables:		
Season	_Season	Annual
Annual mixing height	_mhAnn	2689 feet
Winter mixing height	_mhWin	2497 feet
Spring mixing height	_mhSpr	3300 feet
Summer mixing height	_mhSum	2818 feet
Fall mixing height	_mhAut	2136 feet

Carbon monoxide emission threshold	_meCO	100 tons/year
Nitrous oxides emission threshold	_meNOx	40 tons/year
Sulfur dioxide emission threshold	_meSO2	40 tons/year
Volatile organic compound emission threshold	_meVOC	40 tons/year
Particulate matter emission threshold	_mePM10	15 tons/year
Lead emission threshold	_mePB	0.6 tons/year
Receptor/Pollutant search radius	_srchRad (domBufDis)	50 kilometers
Terrain evaluation mode	_ternMod	complex
PEI threshold percentage	_peiThrsh	100
Windrose Directions	_winDirs	16
Windrose degrees/direction	_winDegs	22.5

Script:

- pei.1StartUp

The PEI Model pulldown menu is then accessed to create the receptors, identify pollutant data and execute the PEI model. Descriptions of functions and scripts follow.

II.2 Add and Delete Receptors

The receptors that represent the sites for which PEIs will be calculated must first be sited in the PEI study area. Initially, Create Receptors is selected from the pulldown menu which prompts the user to specify a new or existing receptor file name.

If a new receptor file is entered, the user may add receptors for an area or enter discrete point by selecting the Add Sites option from the PEI pulldown menu. To enter receptors for an area, the cursor is used to 'digitize' points outlining the area in the domain view (eastern U.S.). The area is subsequently filled with receptors at a user-specified increment. Multiple areas may be populated with receptors for a single receptor file. Discrete points may also be added by double-clicking on the site with the cursor. Receptor sites may be deleted using similar techniques after selecting the Delete Sites pulldown menu option. After the receptors are generated or deleted the model refreshes the domain view.

If an existing file name is specified, the receptors will be saved, but all related tables and PEI calculations will be deleted. The user may then access the PEI pulldown menu to identify if new receptors are to be added or existing receptors are to be deleted from the file using the Add Sites or Delete Sites options.

Scripts:

- pei.3InitReceptor
- pei.4AddPolyPts
- pei.4AddPolyPtsMenu
- pei.4DelPolyPts
- pei.4DelPolyPtsMenu

II.3 Interpolate Receptor Elevations

After the receptors are generated, the Code Elevations entry in the PEI pulldown menu must be selected to interpolate elevations for each receptor from the one kilometer grid file. The interpolation is actually calculated by the Spatial Analyst ArcView extension. After the receptors have been assigned elevations, the theme will be enabled for display in the PEI study area view.

Scripts include:

- pei.5ReceptorsElev

II.4 Identify Pollutant Types and Sources

Identifying the pollutant types and sources and other related operations for PEI evaluation are initiated by selecting the Identify Source Data from the PEI pulldown menu. This first prompts the user to specify a distance that will be used as a search radius. Note that 50 kilometers is the default. The radius is measured from the receptors to select the pollutants that will be used for modeling.

The user then specifies the pollutant types for modeling. The types include carbon monoxides (CO), nitrous oxides (NOx), sulfur dioxides (SO2), volatile organic compounds (VOC), particulate matter (PM10), and lead (PB). The pollutant types selected depend on the sources that are subsequently identified. The sources include the EPA AIRS and County Area and Point data. EPA AIRS is contains all six pollutant types, however there is no PB data in the County Area and Point sources. After the pollutant types are identified, the minimum pollutant emission thresholds may be increased, if desired. Next, pollutant source(s) are selected. These include AIRS, County Area and Point, or combinations of these sources. Any combination may be specified except AIRS and County Point data, which are mutually exclusive. At this point, a series of operations are executed to create shapefiles and themes for the specified pollutant types and sources.

AIRS data is relatively simple to develop since the site locations and related attribute codes reside in preprocessed shape files for each pollutant type. AIRS sites that are within the search radius of the receptors and of the specified pollutant type are generated as shapefile(s) for subsequent processing.

The County source data is more complex to process for two reasons. First, the preprocessed County sites are in a shape file and the descriptive attributes are in separate Area and Point dBASE files. Second, not only are the County data identified by pollutant type, but it is further subdivided into pollutant source categories. The following briefly describes how the County data is handled.

Initially, the County sites within the receptor search radius are identified. Next, the Area and Point attribute records within the search radius and of the specified pollutant type(s) are selected. The data may be further refined by selecting pollutant type categories. The Area and/or Point attribute data is appended by pollutant type and summarized by State/County emission rates. That is, if there are several records describing the same County site for a particular pollutant type, then their emission rates are summed in one record for modeling. For example, if the Metals Processing and Other Industrial Processes categories were selected for CO evaluation and information for both were referenced for some Counties, then the emission rates of these will be summed for each County. Finally the summarized attributes are joined to the pollutant type receptor site shapefiles and theme(s) are generated.

The AIRS and County pollutant type files are also referenced with meteorological station identifiers. Pollutants contained in the station areas of influence are assigned that areas station identifier. This is eventually used to develop wind frequencies for each receptor/pollutant pair. Finally, the common pollutant type theme(s) derived from various sources as individual pollutant type theme(s) are merged. For example, SO2 extracted as themes from the AIRS and County Area sources are merged as one SO2 theme for PEI evaluation.
An additional attribute field called 'emissdup' is added to the pollutant database tables. The user may copy the original emission rates into the new field using ArcView's Table document Field Calculator. The user may then modify the emission rates, using ArcView tools, and maintain a copy of the original source values in the 'emissdup' field.

The County Area and Point dbf database templates are attached.

Scripts:

- pei.2SourceData
- pei.2CountySum
- pei.2Category
- pei.2MetAoi
- pei.2MergePolThm

II.5 PEI Model

The PEI model scripts performs several complex functions. First, it creates database tables for each pollutant type containing the receptors and pollutant sites that are located within the specified search radius. The component data comprised of; emission rates, terrain landform factors, wind frequencies, and distances between the receptors are then calculated for each receptor/pollutant pair. The PEIs are then calculated for each receptor/pollutant pair by pollutant type from the component data. Finally, the PEIs of the pairs are summed for each receptor and entered in the receptor database table, by pollutant type.

Before the PEIs are actual modeled global variable defaults representing the season, mixing-height, terrain factor and PEI threshold percentage may be changed.

A detailed description of the PEI module is provided below. The basic components and associated variables used to calculate the PEI are:

Emission rate	= Q
Terrain landform factor	= T
Wind frequency	= F
Receptor/Pollutant distance	= D
Pollution Exposure Index	= PEI

Initially, database tables are developed, by pollutant type, containing receptor/pollutant pairs. The pairings are based on the user-specified search radius. Those pollutants that fall within the search radius of the receptor are paired with the receptor in the receptor/pollutant database table. After the receptor/pollutant pairs are established, the component data are calculated. First, the pollutant emission rates (Q) are copied from the pollutant type database tables to the receptor/pollutant tables. Emission rates less than the rate threshold will not be evaluated. Next, the distances between the pollutants and receptors (D), measured in kilometers, are calculated and also entered in the receptor/pollutant tables. If the distance is less than 100 meters, then it is set to 100.

The seasonal wind frequencies (F) are also derived and entered into the tables for each receptor/pollutant pair. The process involves calculating the azimuthal angles from the receptors to the pollutants, then interpreting the seasonal wind frequencies for the azimuthal angles from the meteorological station wind frequency database tables. Recall that station identifiers are assigned to the pollutants and are used to derive the appropriate wind frequency data. Finally, the wind frequencies are copied to the receptor/pollutant tables.

Terrain landform factors (T) are also interpreted for each receptor/pollutant pair. Two alternative methods are used to derive the terrain factors depending on the evaluation mode. If the evaluation mode is simple, then the terrain factors of all the receptor/pollutant pairs are set to one. If the evaluation mode is complex, the default, then they are calculated from mixing height, pollutant and receptor elevations. The equations used to interpret the complex terrain landforms are described as follows:

1. Ee = max(Er,Em), where E = elevation
   r = receptor site
   m = highest point between receptor and pollutant sites

   – Spatial Analyst to interpolate receptor elevations (Er) from the 5 km elevation grid . Note that pollutant site elevations (Es) are pre-processed.
   – Derive the highest elevation (Em) between each pollutant and receptor (Er) sites. The model draws virtual lines between the locations of the receptor/pollutant pairs that cross 100 foot elevation contours. The highest contour elevation that is intersected is recorded as Em.
   – The receptor elevation (Er) and highest elevation between receptor/pollutant (Em) are compared to identify the higher elevation (Ee) .

2. T = Mh/(Mh + (Ee-Es)), where T = landform factor
   Mh = mixing height

E = elevation

e = highest elevation (above)

s = pollutant site

– Subtract elevations of pollutant site (Es) from highest point between receptor/pollutant (Ee) and add to mixing height. Divide the result into the mixing height to calculate the terrain landform factors (T).

After the basic terrain factor (T), wind frequency (F), emission rate (Q), and distances between receptors and pollutants (D) modeling components are derived for the receptor/pollutants, the PEIs are calculated and recorded for each pair in the receptor/pollutant database table. The equation used to calculate the PEI by pollutant type for each receptor/pollutant pair is:

3. PEIrp = Trs * Frs * Qp / Drs, where PEI = pollutant exposure index measure

R = receptor site

s = pollutant site

p = pollutant type

In addition, the percentage each pollutant contributes to the total PEI for each receptor is calculated and entered into the table. See Appendix VI – Receptor/Pollutant Tables and Appendix VII – Receptor Tables for descriptions of receptor/pollutant and receptor databases. Finally, the PEIs are calculated for each receptor by pollutant type from the data in the receptor/pollutant pair tables. The PEI of each receptor (r) for pollutant (p) is calculated by summing the emission rates Q, distance between receptor/pollutants D, wind frequencies F, and terrain factor T for each receptor/pollutant pair using the equation:

PEIrp = [SUM (Trs * Frs * Qsp / Drs)]

The receptor PEIs may be re-calculated by changing the PEI threshold percentage. Only the PEIs of the pollutants contributing less than the PEI threshold percentage are used to sum the PEI for each receptor. Those above the threshold are ignored. Therefore, the default of 100 percent will result in the summation of all receptor/pollutant pair PEIs to determine the receptor PEIs. This option will change only the receptor PEI summations, not the individual receptor/pollutant table entries. Ultimately, the receptor identifiers, their elevations, and adjusted PEI for the modeled pollutant types are recorded in the receptor database table.

Scripts:

- pei.6Model
- pei.6ModThresh

II.6 PEI Contours

The Spatial Analyst ArcView extension is used to interpret isoline contours from the PEI values calculated for each receptor for a particular pollutant. For best results, it is recommended that the inverse distance weight (IDW) method of interpolation be applied using a grid cell size of approximately 75 meters.

II.7 Shutdown

The shutdown module saves the receptor/pollutant pairs database table, the receptor file, and the contour files. The variable settings and pollutant sources used to produce the receptor file are recorded in the pei/text/variable.txt text file. The script also deletes extraneous polcoX, modX, and CatAre dBASE files, where X is the receptor file number. Finally, the module re-initializes the PEI project model.

Scripts:

- pei.12Shutdown
-

II.8 Print Variables

A text file may be generated by the user in the pei/text/Variable.txt file that describes the values of the global variables when the user selects the Print Variable item from the PEI Model pulldown menu and at the time the model is shutdown.

II.9 About PEI Model

This opens a file with a descriptive overview of the PEI model and each of its scripts.

Appendix I

Meteorological Stations DBF Database

There are 108 meteorological stations derived from the EPA SCRAM bulletin board, which are used to interpret wind frequencies for the PEI model. The attributes are described in two dbf database files. The first, MET.DBF, contains information related to the location of the stations and their names. The second, MET.LUT, lists the seasonal wind frequency data from 16 directions (every 22.5 degrees) for each station. The LUT contains 1728 records.
The elevations for each site were interpolated during pre-processing from the 1 kilometer elevation grid file.

File Name	Field Name	Field Description
MET.DBF	STATION_	Station number
	LONGITUDE	Decimal degree longitude
	LATITUDE	Decimal degree latitude
	REGION	Region identifier
	STATE	State abbreviation
	NAME	Station name
	ELEVATION	Station elevation
MET.LUT	STATION_	Station number
	DIRECTION	Wind direction degrees
	WINWNDHRS	Winter hours of wind
	SPRWNDHRS	Spring hours of wind
	SUMWNDHRS	Summer hours of wind
	FALWNDHRS	Fall hours of wind
	ANNWNDHRS	Annual hours of wind

Appendix II

Meteorological Station Areas of Influence DBF Database

There are 133 meteorological station areas of influence.

File Name	Field Name	Field Description
METAOI.SHP	STATION_	Station number
	ELEVATION	Station elevation
	NAME	Station Name

Appendix III

EPA AIRS Pollutant Source Points DBF Database

Two sets of database files will be generated for each of the six pollutant type shape files, comprised of CO, NOx, SO2, VOC, PM10, and PB. The first file, polltype.SHP, identifies the source plant identifier, name, and emission rates, while the second, polltype.lut, contains the site locations and miscellaneous information. There are a total of 8938 pollutant sources, including 931 carbon monoxide (CO), 2180 nitrous oxides (NOx), 2164 sulfuric oxide (SO2), 1459 volatile organic compound (VOC), 2180 particulate matter (PM10), and 24 lead (PB) sites.
Note that there are no emission years identified in database for NO2. The station number field is populated during the execution of the model.

The database file structure for the AIRS pollutant types follows:

File Name	Field Name	Field Description
polltype.SHP	SOURCE_ID	Pollutant source identifier
	PLANT	Plant name
	EMISSION_RATE	Emission rate (tons/yr)
	STATION_	Met station number
	ELEVATION	Source elevation
polltype.LUT	SOURCE-ID	Pollutant Source identifier
	LONGITUDE	Decimal degree longitude
	LATITUDE	Decimal degree latitude
	REGION_	Region number
	STATE	State abbreviation
	COUNTY_	County code
	ZIP	Zip code
	SIC	SIC code (plant type)
	YEAR	Emission year
	PROCESS_FLAG	Data reliability

Appendix IV

National Acid Precipitation Assessment Program Source Data Description

The Countywide area and point source data came from EPA's 1985 National Acid Precipitation Assessment Program (NAPAP) inventory. These data include Countywide PM, NOx, SO2, VOC, and CO emission estimates for point and area sources. For each county, there are 14 main categories, which are further broken down into 84 different sub categories. For use in the PEI model, these sub-categories were grouped into 17 new categories. The 17 new categories are shown in Table 1.

Table 1. Generalized County-wide Point and Area Source Categories Used in PEI

New Cat Number and Name	
1	FUEL COMB. ELEC. UTIL.
2	FUEL COMB. INDUSTRIAL
3	FUEL COMB. OTHER
4	CHEMICAL & ALLIED PRODUCT MFG
5	METALS PROCESSING
6	PETROLEUM & RELATED INDUSTRIES
7	OTHER INDUSTRIAL PROCESSES
8	SOLVENT UTILIZATION
9	RESIDENTIAL WOOD & OTHER
10	WASTE DISPOSAL & RECYCLING
11	HIGHWAY VEHICLES
12	OFF-HIGHWAY VEHICLES
13	NATURAL SOURCES
14	MISCELLANEOUS
15	OFF-HIGHWAY OTHER
16	AGRICULTURE & FORESTRY
17	FUGITIVE DUST

The distribution of NAPAP original 84 sub categories to the new categories is shown in Table 2. These data were provide by EPA in either an ASCII text or Foxpro database format on a state-by-state level. The data were put into a standard format in an Access database, merged, and formatted for use in PEI.

Table 2. EPA Point and Area Source Category Distribution

Main Category Number and Name / Sub Category Number	Sub Category Name	New Category Number
1 FUEL COMB. ELEC. UTIL.		
1	Coal	1
2	Oil	1
3	Gas	1
4	Other	1
5	Internal Combustion	1
2 FUEL COMB. INDUSTRIAL		
1	Coal	2
2	Oil	2
3	Gas	2
4	Other	2
5	Internal Combustion	2
3 FUEL COMB. OTHER		
1	Commercial/Institutional Coal	3
2	Commercial/Institutional Oil	3
3	Commercial/Institutional Gas	3
4	Misc. Fuel Comb.(ExceptResid	3
5	Residential Wood	9
6	Residential Other	9
4 CHEMICAL & ALLIED PRODUCT MFG		
1	Organic Chemical Mfg	4
2	Inorganic Chemical Mfg	4
3	Polymer & Resin Mfg	4
4	Agricultural Chemical Mfg	4
5	Paint, Varnish, Lacquer, Ename	4
6	Pharmaceutical Mfg	4
7	Other Chemical Mfg	4
5 METALS PROCESSING		
1	Non-Ferrous Metals Processing	5
2	Ferrous Metals Processing	5
3	Metals Processing NEC	5
6 PETROLEUM & RELATED INDUSTRIES		
1	Oil & Gas Production	6
2	Petroleum Refineries &Related	6
3	Asphalt Manufacturing	6
7 OTHER INDUSTRIAL PROCESSES		
1	Agriculture, Food, & Kindred	7
2	Textiles, Leather, & Apparel	7
3	Wood, Pulp & Paper, &Publish	7
4	Rubber & Miscellaneous Plastic	7
5	Mineral Products	7
6	Machinery Products	7
7	Electronic Equipment	7
8	Transportation Equipment	7
9	Construction	7
10	Miscellaneous Industrial Proce	7
8 SOLVENT UTILIZATION		
1	Degreasing	8
2	Graphic Arts	8
3	Dry Cleaning	8
4	Surface Coating	8
5	Other Industrial	8
6	Nonindustrial	8
7	Solvent Utilization NEC	8
9 STORAGE & TRANSPORT		
1	Bulk Terminals & Plants	6
2	Petroleum & Petroleum Product	6
3	Petroleum & Petroleum Product	6
4	Service Stations: Stage I	6
5	Service Stations: Stage II	6
6	Service Stations: Breathing...	6
7	Organic Chemical Storage	4
8	Organic Chemical Transport	4
9	Inorganic Chemical Storage	4
10	Inorganic Chemical Transport	4
11	Bulk Materials Storage	4
12	Bulk Materials Transport	4
10 WASTE DISPOSAL & RECYCLING		
1	Incineration	10
2	Open Burning	10
3	POTW	10
4	Industrial Waste Water	10
5	TSDF	10
6	Landfills	10
7	Other	10
11 HIGHWAY VEHICLES		
1	Light-Duty Gas Vehicles	11
2	Light-Duty Gas Trucks	11
3	Heavy-Duty Gas Vehicles	11
4	Diesels	11
12 OFF-HIGHWAY		
1	Non-Road Gasoline	12
2	Non-Road Diesel	12
3	Aircraft	15
4	Marine Vessels	15
5	Railroads	15
13 NATURAL SOURCES		
1	Biogenic	13
2	Geogenic	13
3	Miscellaneous	13
14 MISCELLANEOUS		
1	Agriculture & Forestry	16
2	Other Combustion	14
3	Catastrophic/Accidental Release	14
4	Repair Shops	14
5	Health Services	14
6	Cooling Towers	14
7	Fugitive Dust	17

Appendix V

NAPAP County Area and Point Pollutant Tables DBF Database

The County Area and Point pollutant information for all pollutant types are stored in individual tables that are related to the County point shapefiles. The ACATEGORY and PCATEGORY fields contain about 37 categories. The values in the 'Total' field are used to derive the cumulative Area or Point emission rates for the Counties by pollutant type.

The attribute fields for the point shapefiles and the COUNTY.ARE and .PNT tables are listed below. The tables are sorted by the stco redefined item. Neither contains PB pollutant source type records. Note that there are 2626 County points referenced with 89,672 Area records and 18,186 Point records. There are 7986 Area records that identify the total emission rates by pollutant types for the Counties and 4466 describing the cumulative rates for the Points.

File Name	Field Name	Field Description
COPOLL.SHP	STATE	State abbreviation
	COUNTY_NUM	County fips code
	LATITUDE	Decimal degree latitude
	LONGITUDE	Decimal degree longitude
	REGION_	Region number
	STATION_	Met station number
	ELEVATION	Source elevation
	STCO	State and County
COPOLL.ARE	SOURCE_ID	Pollutant source identifier
	POLL_TYPE	Pollutant type identifier
	STATE	State abbreviation
	COUNTY_NUM	County fips code
	CATEGORY	Category type
	EMISSION_R	Area emission rate
	POLSTCO	Pollutant, state and county
	STCO	State and County
COPOLL.PNT	SOURCE_ID	Pollutant source identifier
	POLL_TYPE	Pollutant type identifier
	STATE	State abbreviation
	COUNTY_NUM	County fips code
	CATEGORY	Category type
	EMISSION_R	Point emission rate
	POLSTCO	Pollutant, state and county
	STCO	State and Count

Appendix VI

Receptor/Pollutant Tables DBF Database

One pollutant source/receptor database file is generate for each pollutant source type containing the following information.

File Name	Field Name	Maximum Char Width	Type I,C,N	Decimals (inc.)	Field Description
RC#type.DBF	RCPID	5	I	0	Receptor identifier
	SRCID	5	C	0	Pollutant point identifier *
	RCPELEV				Receptor elevation
	SRCELEV				Pollutant elevation
	DISTANCE			2	Receptor/Pollutant Distance
	AZIMUTH			2	Angle - receptor to pollutant
	EMISSION			3	Emission rate
	WINDFRQ				Wind Frequency
	EM				Highest elev between recp/poll
	T			2	Terrain landform factor
	PEI				recp/poll pair PEI
	PEI_PCENT				Percent PEI contribution of recp

* For the AIRs data the srcid is a unique source-id, however for the County Area and Point data the srcid is the appended state abbreviation and the County number (e.g. AL and 102 makes the srcid equal to AL102).

Appendix VII

Receptor Tables Info Database

File Name	Field Name	Field Description
RECnum.SHPRcpid	5	Receptor identifier
	ELEVATION	Receptor elevation
	PEI_TYPE	PEI by pollutant type

GIS Uncertainty and Policy: Where Do We Draw the 25-Inch Line?

James E. Mitchell
Institute for Environmental Studies, Louisiana State University, Baton Rouge, Louisiana

Abstract

The growing availability of improved hardware and software for geographic information systems (GIS) has outstripped most users' ability to identify and represent uncertainty in the available data. In practice, the proliferation and compounding of errors and uncertainty increase as information becomes more easily handled and combined from different sources.

Various stages of GIS database development and analysis generate different forms and amounts of error and uncertainty. In most cases, inherent uncertainty within source data is simply ignored and its nature eventually lost through subsequent processing. Both the location of features and their attributes can include error and uncertainty. By the time decision-makers receive mapped information, it is typically represented as correctly located and attributed.

The use of weather and climate information provided by the National Climatic Data Center (NCDC) is a common example of this scenario. Weather station locations provided by NCDC are reported to the nearest truncated degree-minute. A minute is one-sixtieth of a degree of arc. In the center of the continental United States, 1 minute of latitude averages approximately 6,000 feet and 1 minute of longitude averages approximately 4,800 feet. Thus, the station location is only known to lie within a box of approximately 1 square mile. Map representations of these data should reflect this uncertainty.

Under the Municipal Solid Waste Landfill (MSWLF) Criteria, the U.S. Environmental Protection Agency has dictated that the 25-inch precipitation contour line be used as a regulatory boundary for the level of protection required at municipal landfill sites. The way in which these lines are created and interpreted has important policy implications. Indeed, the cost and practicality of a given location must take this into account. If the 25-inch precipitation figure is critical, characterizing its uncertainty is also important.

In this work, uncertainty is considered a property of the data (1). A Monte Carlo procedure is used to represent the stochastic character of contour lines generated from point data with known locational uncertainty. The 30-year normal precipitation data for Kansas are used as an example. The results of this study are compared with the 25-inch contour used for regulatory purposes in Kansas. This study demonstrates that the method of interpolation greatly influences the resulting contours. In addition, locational uncertainty changes the results unpredictably using four different contouring methods. Finally, the differences have potentially significant policy implications. The nature and origin of these factors are discussed.

Problem Statement

The increasing power of geographic information systems (GIS) and the availability of digital data have enabled users and decision-makers to perform complex spatial analyses for a great variety of environmental applications (2). The rapid expansion of GIS has resulted in a parallel growing concern about the quality of data (3).

An understanding of error and uncertainty is critical for proper use of spatial information. For the purposes of this discussion, error is defined as a deviation between the GIS representation of a feature and its true value (4). For a location, this might arise from rounding or truncating digits. Attribute error can involve misclassification of a feature or some other form of incorrectly accounting for its nature. Error is a measurable value quantifying these differences.

Furthermore, uncertainty shall refer to a characteristic for which the exact location and/or quantity cannot be calculated (5) or an attribute whose value represents a distribution or some other ensemble (composite) measure. Locational uncertainty often arises when inappropriate measurement systems are used. An example of this is the use of a Public Land Survey designation (often referred to as a legal location) to specify a point location. This system is designed to represent a tract of land (an area). It is not accurate for locating points (1).

The uncertainty associated with an attribute is an important characteristic of that feature. It quantifies the precision of a stochastic quantity; that is, one that is not accurately represented by a single value. Annual precipitation is represented by a single number, typically the 30-year mean annual precipitation (30-year normal). This number varies each year, however, and that uncertainty can be quantified by the variance or other statistical measures. In this sense, uncertainty is a known or calculable value that can be used in spatial analyses.

Unreliable GIS data and products may lead to adverse environmental and legal consequences. The National Center for Geographic Information Systems and Analysis (NCGISA) chose data quality as the first initiative on its GIS research agenda (6). Many efforts have been made prior to this, and since, to understand and manage error and uncertainty in GIS applications.

GIS analyses are inherently subject to propagation of error and uncertainty (4, 7). No data set can represent every spatial reality of a geographically dispersed phenomenon. Monmonier (8) points out that as long as the three-dimensional earth's surface is transformed to a two-dimensional plane, error and uncertainty of various forms will be produced. Goodchild and Min-Hua (9) point out two issues that are important when dealing with error and uncertainty:

- Minimization of error in the creation of GIS products.
- Measurement and presentation of error and uncertainty in a useful fashion.

GIS technology introduces error and uncertainty through two major sources: (1) inherent error and (2) operational error. Inherent error is the error present in source data. It is generated when the data are collected. Operational error is generated during data entry and manipulation (7, 10-13). Examples include locational shifts due to projection or combining information from different source scales.

Most error and uncertainty contained in GIS data cannot be eliminated. Instead, they are actually created, accentuated, and propagated through GIS manipulation procedures (14-16). Most operational errors are difficult to estimate.

The selection by the U.S. Environmental Protection Agency (EPA) of a 25-inch per year local precipitation limit as one of the criteria to determine whether small municipal solid waste landfills (MSWLF) are subject to the provisions of Subtitle D provides an excellent example of how uncertainty and errors enter into a GIS analysis and its subsequent products. It demonstrates all of the major forms and purveyors of error and uncertainty:

- Spatial (locational) error
- Statistical (sampling) uncertainty
- Temporal (time domain) error
- Error proliferation (processing error)
- Analytical (choice of methodology) error
- Cartographic representation error

Many of these are avoidable; some are known and understood, yet they remain largely ignored by users of GIS technology. This work presents each of these factors, discusses their origins, and shows how GIS could have been used to better serve the policy and regulatory processes. The Kansas example demonstrates that ignoring the factors influencing error and uncertainty can result in incorrect conclusions and inappropriate policy decisions.

Data Requirements and Sources

To perform an analysis of precipitation, data are typically obtained from the National Climatic Data Center (NCDC), located in Asheville, North Carolina. This is the national repository for such data. These data are also available through state or regional climate centers. The Kansas Weather Library at Kansas State University provided data for this study. The 1990 "normal precipitation" data (17) and locations were obtained and generated into an ARC/INFO point coverage. Figure 1 displays the locations of the precipitation stations used in this study.

Normal precipitation is defined as the average annual precipitation for a three-decade (30 years) period at each station for which reliable data are available. To avoid "edge effects" (processing anomalies due to a lack of data along edges of an area), all stations in Kansas and some from neighboring states were used. A total of 380 stations compose this data set. In addition, precipitation contours from the "Availability of Ground Water in Kansas Map" (18) were digitized from a [paper] source map. The Geohydrology Section of the Kansas Geological Survey provided base map coverages of cartographic features.

All data represent the best available information from the source institutions noted above. Those organizations use the data in their analytical and cartographic research and production operations.

Methodology

To examine the influence of locational uncertainty on the representation of three-dimensional, natural phenomena, a Monte Carlo approach was adopted (1). Using this technique, random realizations of point locations are generated for each rain gauge, in each of 50 separate simulations. From this, 50 possible representations of the unknown locations of each gauge are used to create

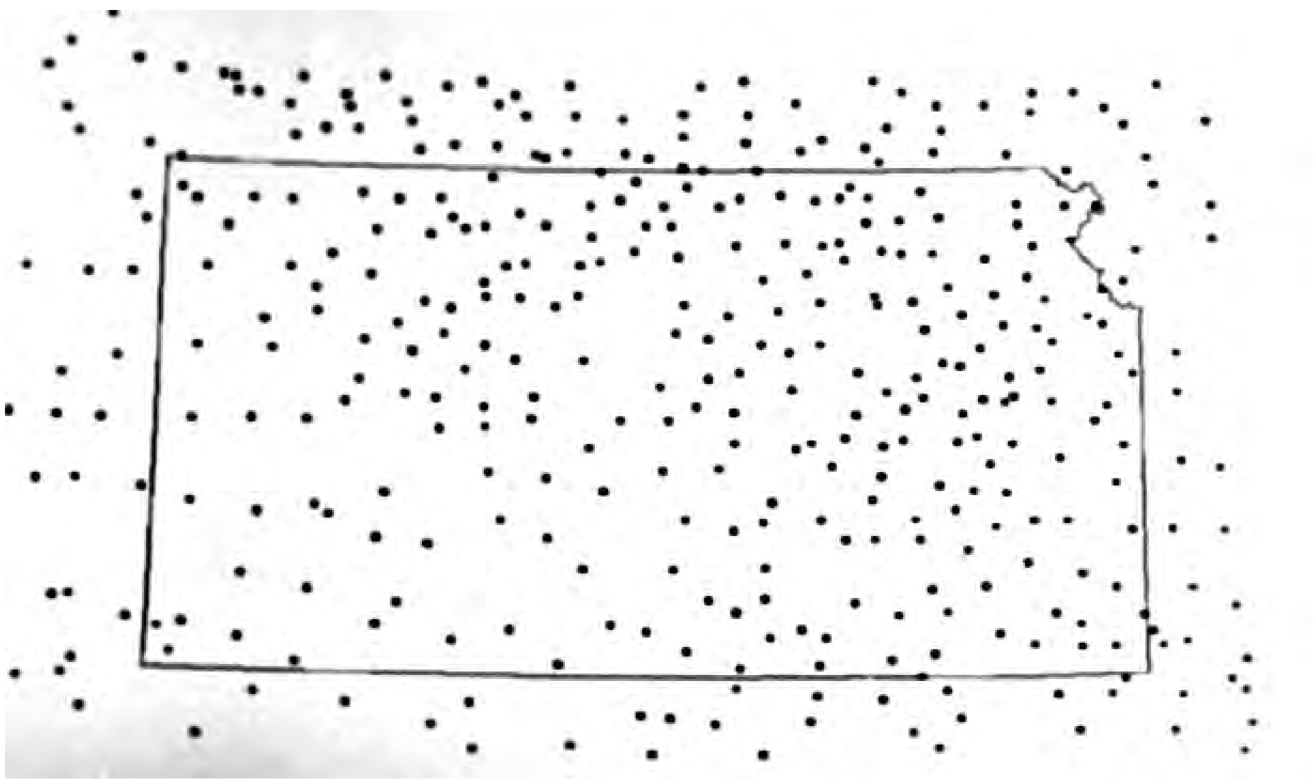


Figure 1. Location of rain gauges used in this study.

50 different sets of contours. All Monte Carlo calculations and data generation were performed using Statistical Analysis System (SAS) (19, 20).

These 50 simulations were sequentially processed using the four different contouring methods available within ARC/INFO. This provided a means to examine analytical error propagation. The first of the four methods is kriging (21, 22). This is referred to in the paper as the UK method, for its use of linear universal kriging (23). The other three are manipulations of the triangular-irregular network (TIN) contouring algorithm available in ARC/INFO. These differ by the number of interpolation points used along the edges of the elements in the TIN data structure (24). The first used the default 1, the second used 5, and the third used 10 (the largest value available). These are labeled D1, D5, and D10, respectively.

GIS operations used in this work include overlay analysis, areal calculation, and arc intersection. ARC/INFO was used for all GIS and cartographic production in this work.

Identifying the Sources of Uncertainty

Spatial Error

Data obtained from NCDC is provided with the knowledge that weather station locations are reported using truncated degrees and minutes of longitude and latitude. NCDC cannot provide any better locational accuracy at this time. Because each location is reported with error, this clearly has the potential to affect any contours or other three-dimensional features interpolated from the data. The magnitude and nature of this influence are unknown and unpredictable (1).

In addition to the poorly defined station locations, examination of the data revealed other anomalies. The locations in the publication reporting normal precipitation (17) were not identical to those identified by the Kansas State climatologist and NCDC. Some of the discrepancies were quite large. These anomalies were brought to the attention of all parties involved. No resolution was provided to this investigator's satisfaction, however.

The contours digitized from the "Availability of Ground Water in Kansas Map" are stated to originate from the 1960 normals (18). No documentation exists, however, concerning the way the lines were derived or the number of rain gauges used. Presumably, they were contoured by hand.

Statistical Uncertainty

This is a sampling consideration based on the size of the data, the nature of the process being sampled, and its variability. Unfortunately, precipitation is a particularly "patchy" phenomenon. That is, rain falls in a discontinuous fashion, and adjacent gauges can depict very different patterns. This is confounded by the fact that most contouring algorithms and other approaches to represent three-dimensional surfaces assume a relatively smooth (locally) and continuous process.

Areal processes are almost always sampled as point information. Most contouring algorithms require a regular grid from which to interpolate surface features. In practice, rain gauges, as well as other environmental sampling programs, are irregularly distributed. Placement often depends on factors other than grid sampling (e.g., convenience, access to communications, finances, Congressional districts). This creates a "nonexperimental sampling" design (25). Nonexperimental sampling can contribute to uncertainty (26, 27).

Temporal Error

The normals are recalculated each decade and can change drastically in local areas. These changes arise for various reasons. First, some stations enter and drop from the database. Stations are deleted due to changes in location or extended periods of data collection problems. On occasion, new stations are added. Thus, the size and areal coverage of the data set changes with time.

In addition, weather patterns change with time. Extended periods of drought or excess rain or snow alter measured precipitation. In turn, three-dimensional representations change unevenly.

Error Proliferation

Once an error enters into the database and is included in GIS operations, spatial analysis, or spatial interpolation, its effect passes into the next stage of processing. In the 1990 normal precipitation data for Kansas, two stations are reported in Garden City. Despite the fact that the two are only a few miles apart, their annual total precipitation differs by 2 inches! In consultation with the state climatologist (Mary Knapp, Kansas State University), one was eliminated from the analysis. This process was repeated for an additional six stations where reported values appeared to be anomalous compared with nearby stations or the previous normal precipitation (1951 to 1980).

Errors can also proliferate through the normal handling of data. With geographic data, this often occurs while converting data from raster to vector and vector to raster forms (28). Some GIS operations are best accomplished in one form or another. As a result, transformations are often "hidden" from the user. Commonly, features "move" slightly after each step in an analysis.

Analytical Error

Different techniques have been developed for performing spatial interpolation, and an abundance of software is available for this purpose. All these methods have strengths and weaknesses. Each is based on a specific set of assumptions about the form and nature of the data. Some are more robust (less sensitive to data anomalies) than others. Most importantly, some provide additional information useful in data analysis.

Unfortunately, users often "take the defaults" when using sophisticated techniques and ignore the assumptions behind the method. Parameters can be varied and their effect evaluated, as in a sensitivity analysis (29). Often, the best approach is to try several methods and evaluate their joint performance (30, 31).

Another difficulty is the need to assign values to areas. By definition, polygons in a GIS are considered to be homogeneous. In reality, they bound areas that are a gradation from one characteristic to another. On the other hand, contours are commonly used to depict surface gradients but are useless (within a GIS) for analytical or modeling purposes. Ultimately, data sampling is accomplished as a point process (except, perhaps, in remote sensing), while many forms of data analysis and processing require areal information.

Cartographic Representation Error

Communicating the uncertainty of map features is not a trivial endeavor. Maps can be produced in two basic forms: as a raster (e.g., orthophotoquads or satellite images) or a composition of vectors (e.g., contour maps). The printing process, however, often reduces all of this to a raster representation at a very fine pixel size. Each method poses its own problems in depicting uncertainty.

Rasters can be used effectively in conjunction with color information theory to produce a continuum of shading within a thematic map layer (32). The choice of colors, however, can influence the interpretation of the data, and no universal scheme exists for depicting thematic variability. For example, blue shades often represent water or cold, while yellow and/or red often represent temperature or heat.

Vectors present a different suite of problems. Contouring is the primary technique for using vectors to depict areal variation. By definition, however, contour lines represent an exact isoline or single value along its length. Uncertainty cannot be represented in a line. Rather, a composite of lines can be displayed that represents a set of possible interpretations of the data. This is not a practical solution for mapping, however, as it can create a jumble of intersecting lines that makes interpretation difficult and is not an aesthetic means of presentation.

Challenges Encountered in This Study

This work attempts to discover and account for sources of error and uncertainty in GIS analysis. Given this information, the challenges are to find the best way to incorporate it into the analysis and to represent it in a useful manner. Another challenge is finding ways to use

GIS uncertainty to support policy and management decisions. Addressing these manifold problems starts with identifying the sources of error and uncertainty, the way they enter the analysis, and the manner in which they are propagated through the use of GIS.

This study includes a number of known sources of uncertainty. In practice, this is not always the case. Users of GIS data and technology should always assume that the sources of uncertainty discussed in this paper are present and attempt to determine their nature. Uncertainty should be considered a property of the data and appropriately represented (1). This is the approach taken in this work.

After examining these factors, a Monte Carlo simulation was deemed the most appropriate approach to capture the nature of the locational uncertainty. Four different methods of contouring were used to examine analytical uncertainty (uncertainty due to the choice of a contouring algorithm). In addition, the contribution of statistical (sampling) uncertainty could have been addressed through incorporating information about the standard error of the point precipitation measurements (normals) used as the base data. Time limitations precluded examining this dimension of the question. Comparing the contours resulting from the 1960 and 1990 precipitation normals demonstrates the effect of temporal variation.

The greatest challenge is communicating the uncertainty in a manner useful to decision-makers. This paper presents a series of maps, figures, and tables aimed at addressing this problem. Some of the maps (see Figures 2 through 5) show the uncertainty resulting from each of the contouring methods. Figure 6 depicts the union (overlay) of the four approaches and displays their correspondence. The pie chart in Figure 7 is a nonspatial representation of this correspondence and the relative area represented within the different combinations of overlapping regions of uncertainty. Tables 1 through 3 further compare these quantities. Figure 8 is the map that the Kansas Department of Health and Environment (KDHE) chose to define the regulatory boundary (the 25-inch contour). The contours resulting from this study can be seen in Figure 9. Finally, the map in Figure 10 is a cartographic comparison of the differences between the contours used by KDHE (based on the 1960 normals) and those generated by the currently available data (the 1990 normals).

There is no single best approach for meeting these challenges, and there may never be one. The real challenge to address is how to educate technical GIS professionals and the users of their work to look for uncertainty and consider its influence on their decision-making process.

Results

The zones of uncertainty defined by the results from the four contouring methods used in this study are displayed in Figures 2 through 5. For each method, these zones represent the areal extent of the overlain contour lines produced in the 50 simulations. Each region is bounded by the furthest west or east contour generated along any length of the region. Table 1 shows the relative area falling within each of these zones as they traverse the state of Kansas. Clear differences exist between the total areas of uncertainty. It is their placement and relative location, however, that have policy and management implications. GIS is required to examine these questions.

Table 1. Comparison of Absolute and Relative Area of Uncertainty Arising From Four Methods of Determining the 25-Inch Precipitation Contour

Method[a]	Total Area (square miles)	Difference Between This and UK Method (square miles)	Percentage of UK Method	Percentage of Combined Area[b] of Uncertainty
UK	289.33	—	—	19.65
D1	494.38	205.05	170.87	33.58
D5	602.13	312.80	208.11	40.90
D10	631.11	341.78	218.13	42.87

[a]UK = universal kriging with linear drift, D1 = TIN interpolation with 1 subdivision, D5 = TIN interpolation with 5 subdivisions, D10 = TIN interpolation with 10 subdivisions (23, 24).

[b]A union (overlay) of all four sets of regions of uncertainty creates a combined area of 1,472.07 square miles. This includes the zones of uncertainty for each method of contouring and areas not included within any of the four regions of uncertainty (gaps between them).

Figures 2 through 5 clearly show differences in both the extent of uncertainty in the 25-inch contour line and its positional interpretation. Each method has a slightly different bend or twist. Islands (isolated regions where the 25-inch line appears as a closed loop) are manifested differently depending on the interpolation scheme. It is interesting to note the relative correspondence between the general shape of the D1 and UK methods. In the south-central border region, D1 and UK represent the local uncertainty as a bulge, while D5 and D10 depict it as an island of lower precipitation.

Areal correspondence and difference are depicted in Figure 6 and Table 2. Figure 7 is a pie chart visualizing the information in Table 2. These results are somewhat surprising in that areas where none of the four methods located the 25-inch line ("None Present") represent the second largest composite area. Because of the large number of "sliver polygons," the graphic representation of the overlay is somewhat difficult to interpret. Table 2 clarifies these interrelationships by breaking down the various categories. The average area per polygon value

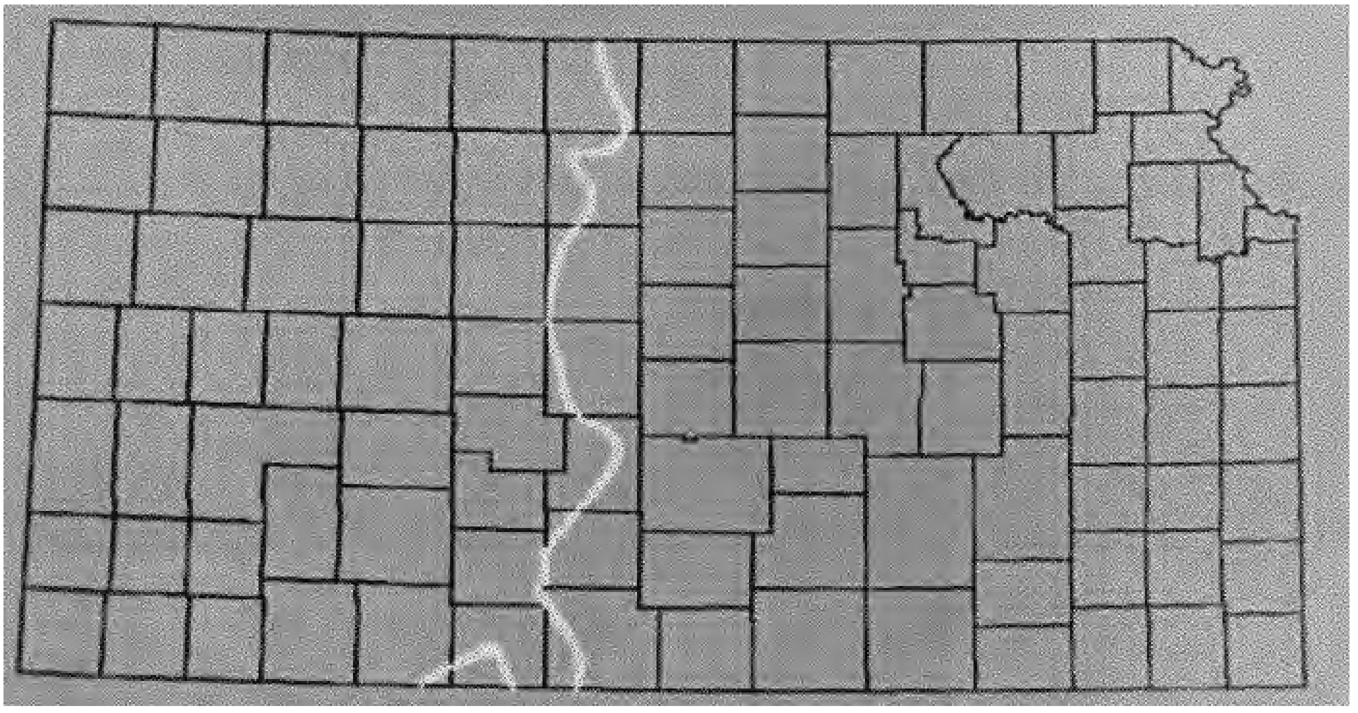


Figure 2. Regions of uncertainty produced by the UK method of contouring.

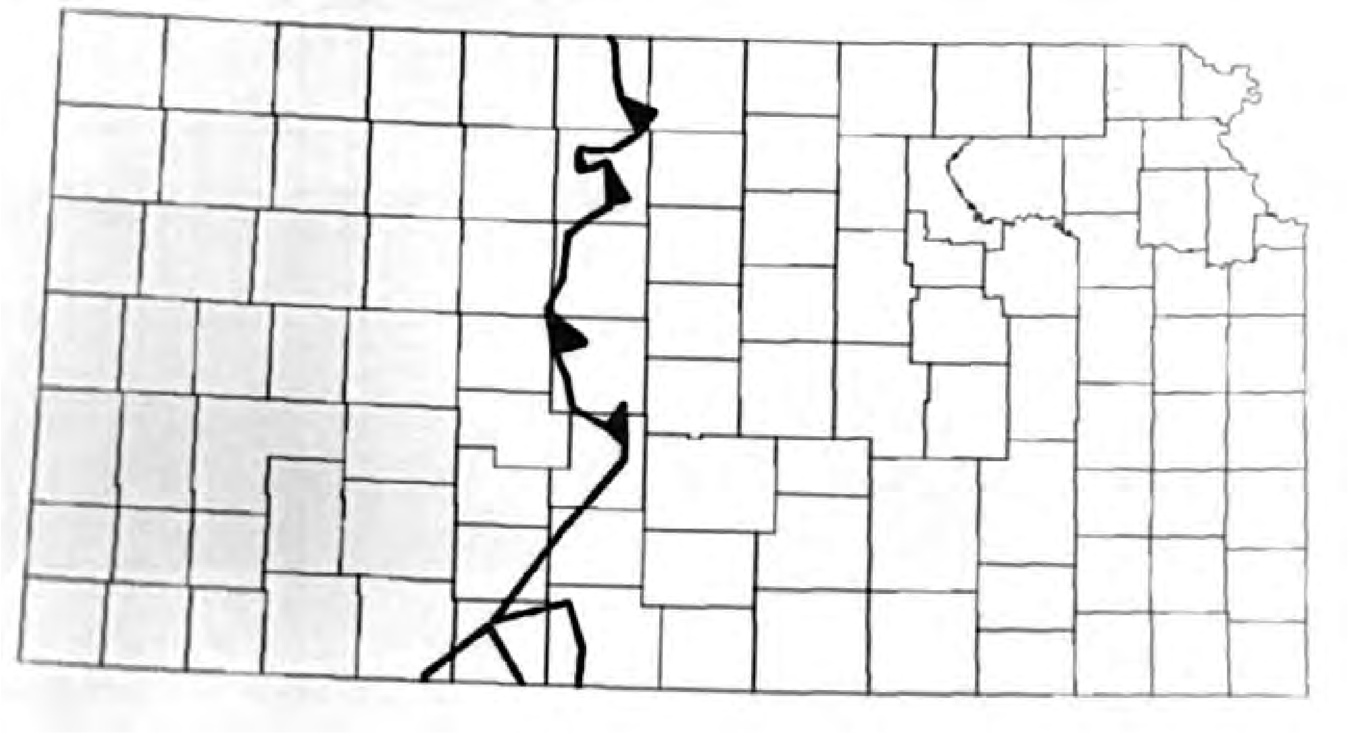


Figure 3. Regions of uncertainty produced by the D1 method of contouring.

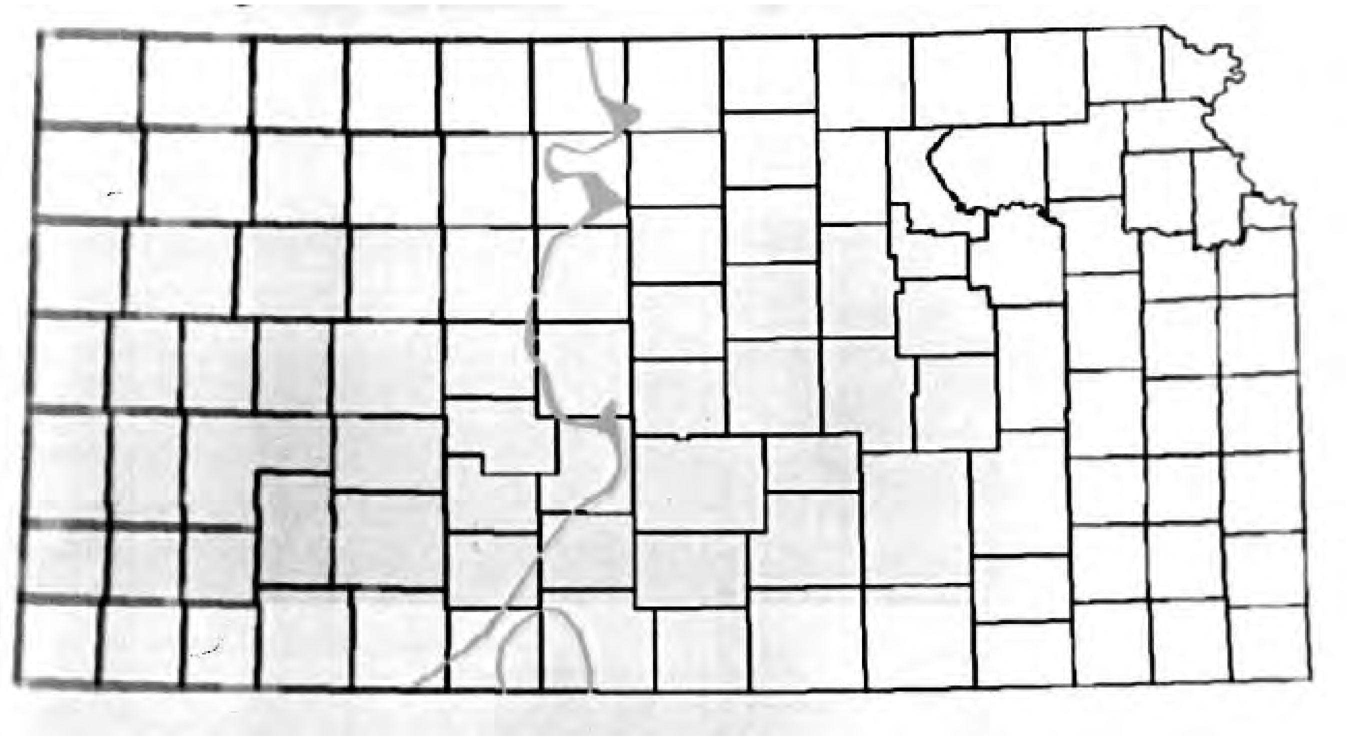


Figure 4. Regions of uncertainty produced by the D5 method of contouring.

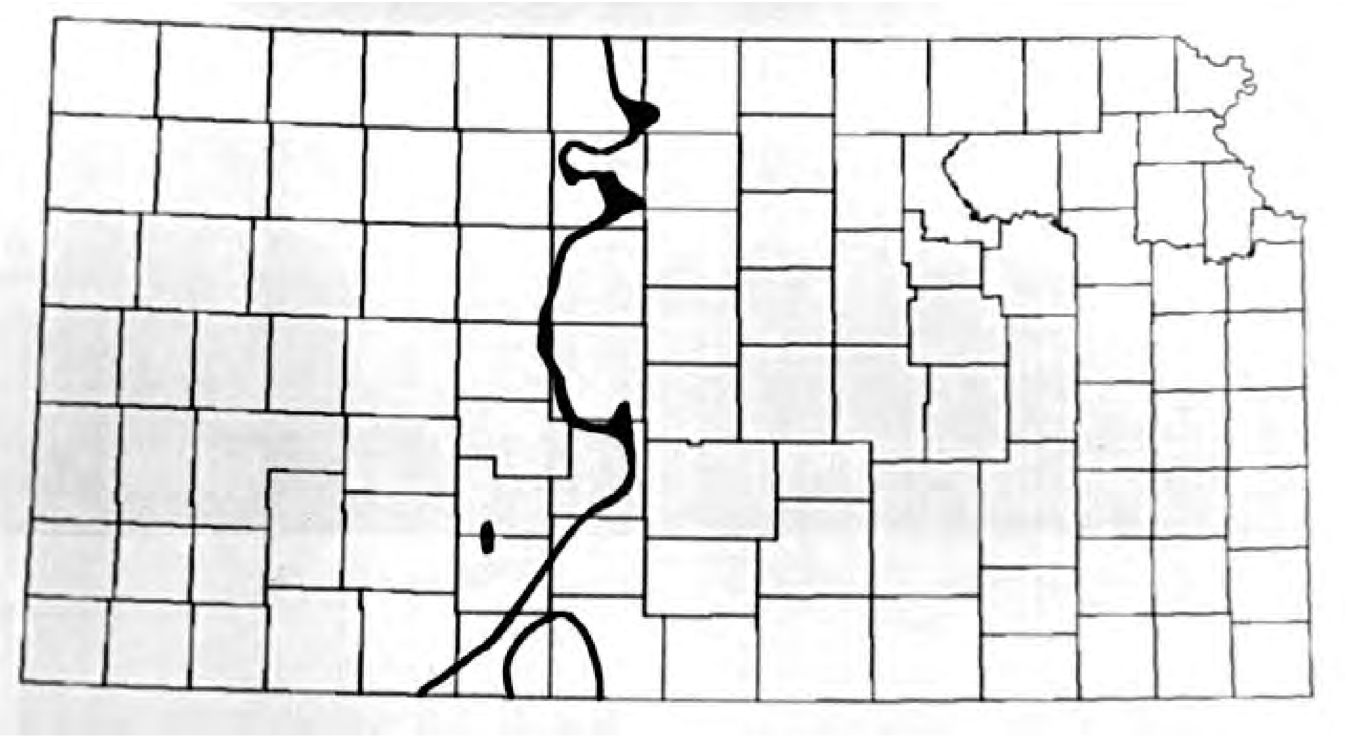


Figure 5. Regions of uncertainty produced by the D10 method of contouring.

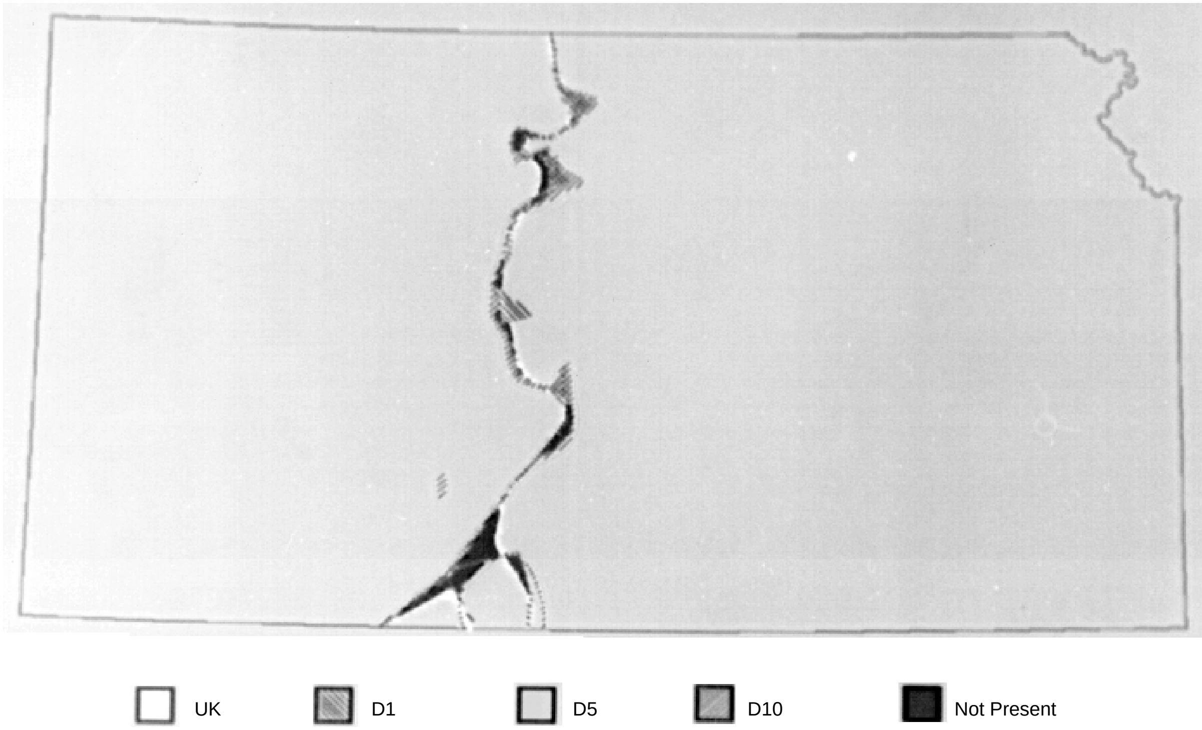


Figure 6. **Union of the regions of uncertainty from all four methods of contouring. The numerous "sliver" polygons make this a difficult presentation to interpret at this scale. The black areas appear prominently, however. These represent areas where no method placed contour lines.**

adds to the interpretation of the relative areas by incorporating the number of polygons in each category. An inspection of this column makes those categories with a multitude of very small polygons stand out. It also displays a number of large jumps in magnitude. As this number increases, the significance of the correspondence increases.

Table 2 highlights the correspondence between the D5 and D10 approaches and the D1 and UK methods of interpolation. This relationship is interesting because the algorithms used by the UK and D1 methods both are forms of linear interpolation. The D5 and D10 algorithms are designed to provide more "smoothing" and appear to create increasingly more "bull's eyes." Only D5 and D10 generate these features. Of the 50 simulations, a particular bull's eye appears west of the 25-inch contour (see Figures 3 and 4) four times using D5 and 39 times using D10. The size and location of these anomalies also vary with the input data. Polygons containing contours from all four methods rank ninth in total area and seventh in average area (out of 17 categories). This supports the conclusion that the four chosen methods have a relatively low spatial correspondence.

Table 3 breaks down the area of uncertainty by county. Although the zones of uncertainty appear to be relatively small when displayed on a statewide basis, they have important impacts in local areas. In particular, combining this information with soils, topography, ground water, and other information can clearly indicate whether a specific location is suitable for a landfill. Often, information developed at one scale is used in another. In this case, statewide information is being used for a site-specific application.

Figure 8 is a copy of the map the KDHE used to delineate the 25-inch precipitation contour. The results from

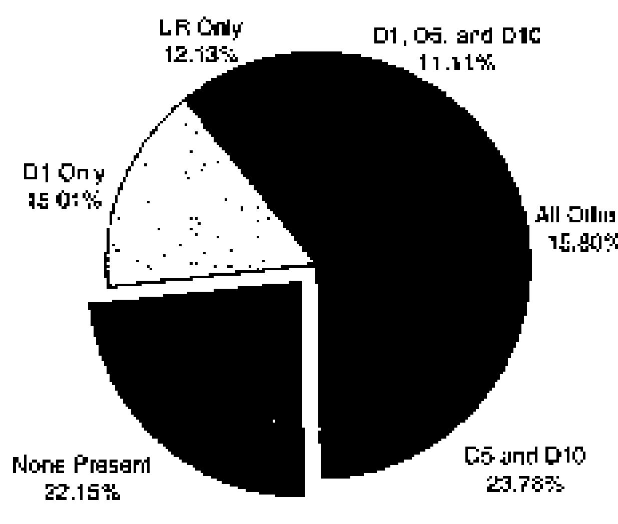


Figure 7. **Breakdown of the total area in each category resulting from the union (overlay) of the regions of uncertainty from all four contouring methods. The "None Present" category represents a surprisingly large proportion among the 17 possible combinations.**

Table 2. Comparison of Absolute and Relative Area of Uncertainty Arising From Four Methods of Determining the 25-Inch Precipitation Contour

Methods of Contouring Found Within Area	N	Total Area (square miles)	Percentage of Total Combined Area	Average Area per Polygon[a] (square miles)
D5 and D10	28	350.13	23.78	12.50
None present	28	326.06	22.15	11.64
D1 only	37	221.02	15.01	5.97
UK only	27	178.60	12.13	6.61
D1, D5, and D10	24	163.61	11.11	6.82
D10 only	69	70.07	4.76	1.02
UK and D1	19	65.82	4.47	3.46
D5 only	73	34.68	2.36	0.48
Common to all	10	23.45	1.59	2.35
UK, D5, and D10	14	15.47	1.05	1.11
D1 and D5	38	10.68	0.73	0.28
D1 and D10	38	6.50	0.44	0.17
UK and D5	19	2.21	0.15	0.12
UK, D1, and D5	12	1.89	0.13	0.15
UK, D1, and D10	12	1.41	0.10	0.12
UK and D10	13	0.47	0.03	0.04
Combined total	461	1472.07	100.00	

[a]Average Area per Polygon = (Total Area) / N. This a useful measure to compare the relative size of each polygon in each classification.

Table 3. Area of Uncertainty for Each County Arising From Four Methods of Determining the 25-Inch Precipitation Contour

County	UK Area (square miles)	D1 Area (square miles)	D5 Area (square miles)	D10 Area (square miles)
Barber	36.38	24.29	28.05	28.31
Barton	28.47	96.84	56.16	59.56
Clark	12.64	11.49	12.97	13.02
Comanche	44.30	52.94	51.43	52.23
Edwards	—[a]	—	—	9.44
Ellis	—	6.14	9.27	8.90
Jewell	—	6.65	23.05	24.26
Kiowa	15.15	14.32	17.81	31.12
Osborne	43.87	99.82	168.2	172.6
Pawnee	—	—	0.11	0.30
Pratt	18.84	11.30	11.56	11.43
Rush	0.04	8.51	31.32	29.51
Russell	27.17	22.71	28.30	28.00
Smith	27.59	68.47	67.38	66.34
Stafford	34.88	70.90	96.49	96.05
Total	289.3	494.4	602.1	631.1

[a]A dash indicates that no contours appeared in that county for the method specified.

the UK method were selected as the best available representation of normal precipitation across Kansas (see Figure 9). The figure displays the unclipped contour lines generated from the data. This is done to point out the importance of "edge effect." Note the incoherent behavior of the contour lines at their termini. If a smaller window of data points were used, interpolation problems would have lain across the region of interest. When present, these features require more handling and time for analysis. They often introduce additional error and uncertainty.

The policy implications of this example are demonstrated in Figure 10. Here, the map shows the combination of the "official" KDHE map and the data interpreted in this study. The pattern of noncorrespondence is noteworthy. The lightest areas are regions that currently experience higher annual precipitation than forecast by the 1960 normals (from the KDHE map). Black areas are expected to have lower precipitation under current climatic conditions. Therefore, large areas of Kansas that should be under regulation according to the MSWLF regulations are not.

In summary, the figures and tables clearly show that locational uncertainty of data measured as points is propagated into contour lines. The nature and magnitude of that uncertainty varies with location and method of interpolation and shows no regular (predictable) pattern. Perhaps most importantly, uncertainty that appears small at one scale can be relatively more significant at another. In addition, seemingly small geographic feature and uncertainty can be an important factor in decision-making.

Discussion and Conclusions

GIS is an established and accepted technology, especially in applications related to natural resource and environmental management. Despite the widespread proliferation of GIS into these areas, the available data are not always appropriate for the intended application. Furthermore, adequate documentation is not always available to determine whether the data are adequate for a given use. The development of metadata standards will play an important role in addressing this problem. Errors and uncertainty will always be present in GIS data. Recognizing their presence, incorporating them into the analysis, and representing them in GIS products will remain a constant challenge.

This study demonstrates the influence that various sources of GIS uncertainty can leverage on the results of an analysis. The example of the 25-inch precipitation

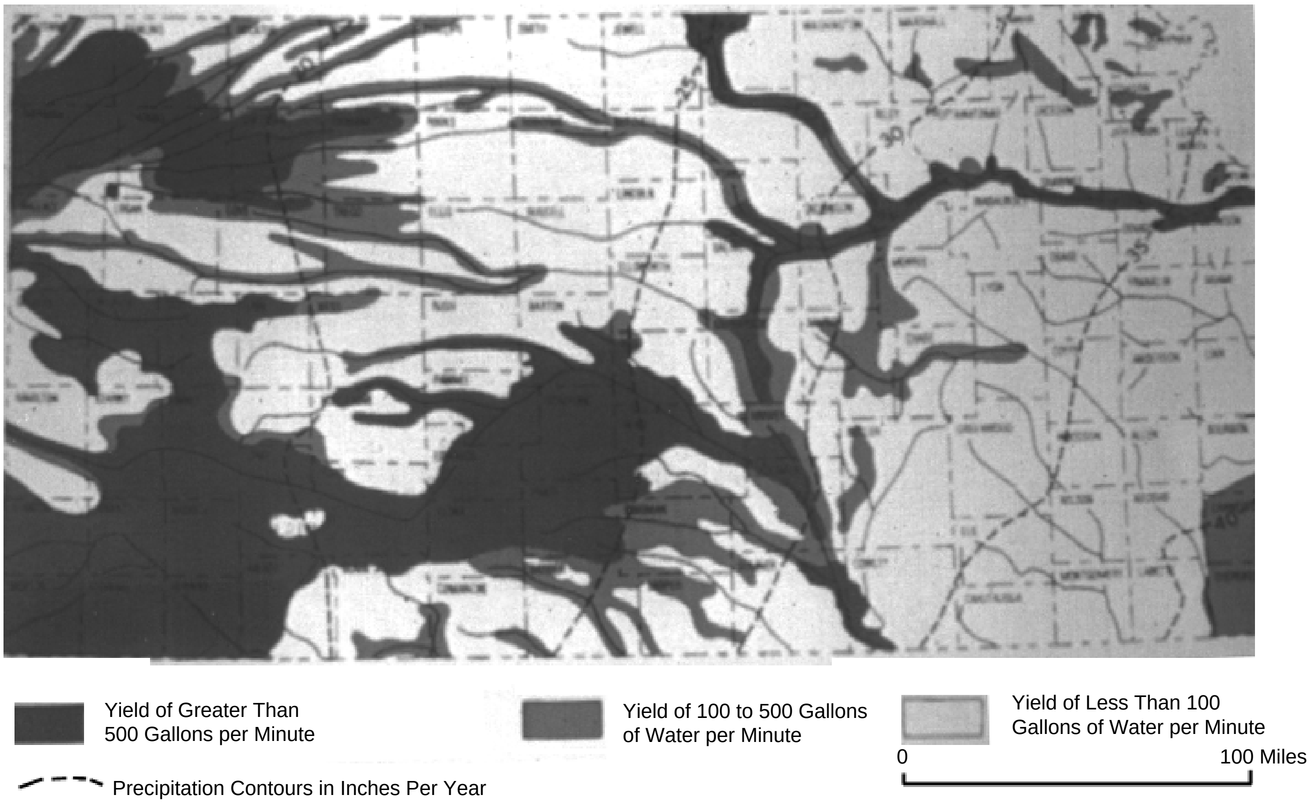


Figure 8. The map that the KDHE selected as the definitive source for the location of the 25-inch precipitation contour (18).

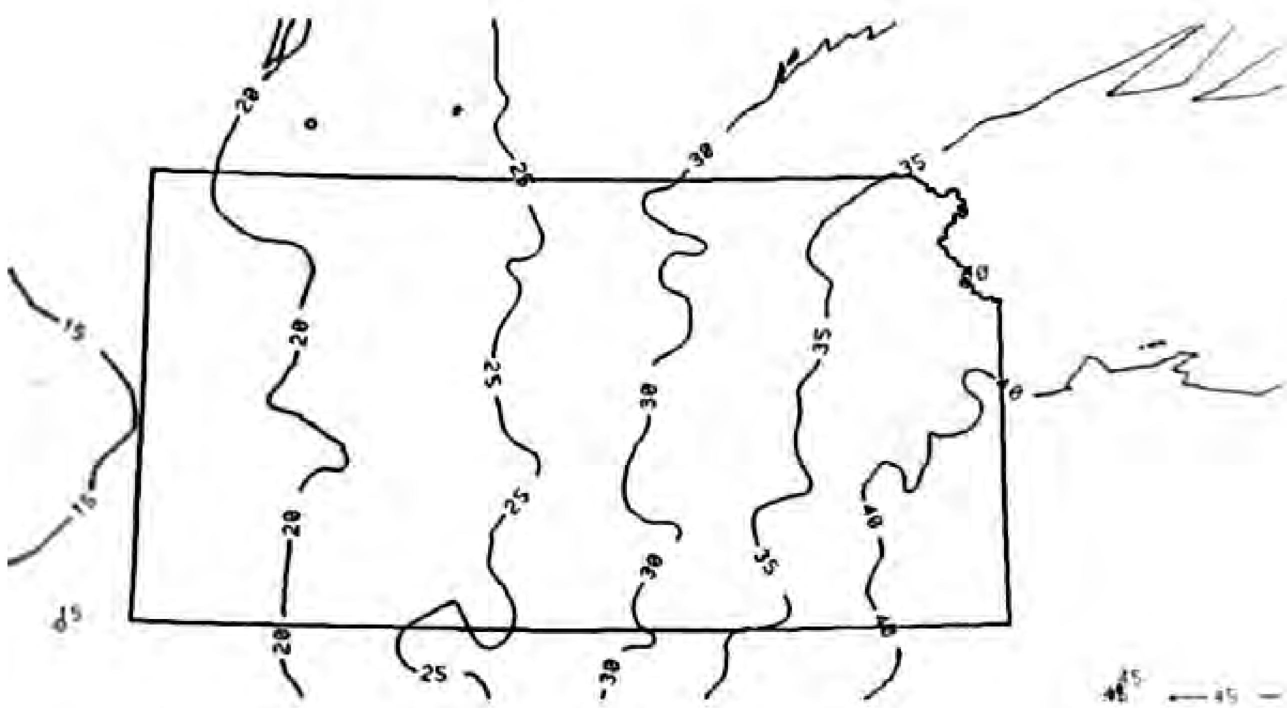


Figure 9. Map of the precipitation contours resulting from applying the UK method with linear drift to the 1990 normals (17). Here, the contour interval is 5 inches. Note the incoherent behavior of the contours around the margins of the map. This is referred to as "edge effect."

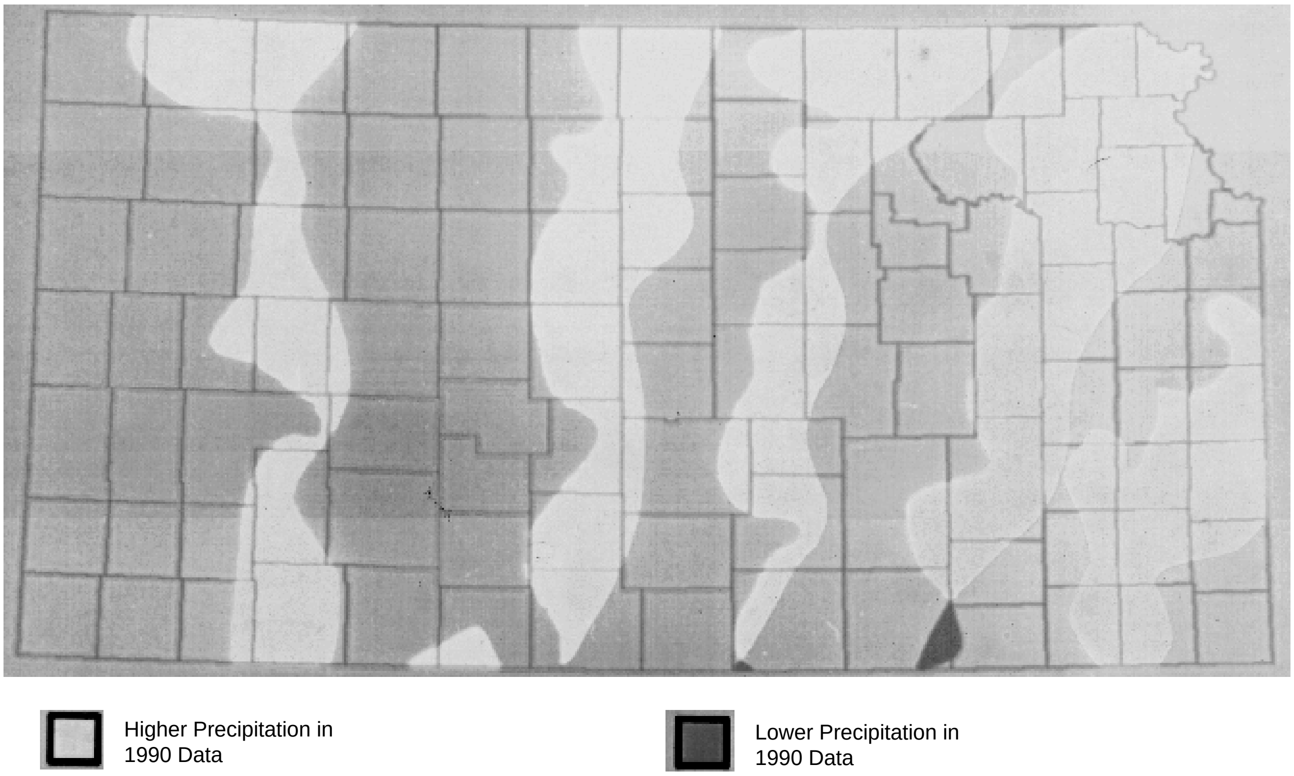


Figure 10. This map represents the union (overlay) of the information in Figures 8 and 9. The lighter regions represent areas exhibiting higher precipitation in the 1990 normals (1961 to 1990) than was apparent in the 1960 normals (1931 to 1960). The darker areas show the opposite relationship.

line in Kansas is a clear example of how the use of inappropriate data can have far-reaching effects on policy and management. The regulatory agency, KDHE, chose the wrong map upon which to base its regulatory authority. As a result, numerous potential sites for small municipal solid waste landfills will be considered that are in violation of the letter and intent of the law.

Ultimately, the responsibility for proper use of GIS technology lies in the hands of practitioners. Technical staff performing GIS analysis must be knowledgeable about sources of error and uncertainty and ensure that users of their work are aware of their influence on GIS output.

The problems demonstrated in the Kansas example could have been avoided simply by investigating the appropriateness of the data. Instead, a convenient source was chosen without seeking any other sources of “better” information. Indeed, familiarity with the nature of the data (30-year normals) should have led the policy analyst to select the most current data and not data that are 30 years out of date! An understanding that contour lines represent a generalization of the point precipitation measurements should also have led to the conclusion that locations near the boundary line ought to be monitored for compliance. Both the temporal and spatial characteristics of climate can change, as exemplified by the difference in the 1960 and 1990 normals. The “Dust Bowl” periods of the 1930s and 1950s significantly influenced the 1960 normals (33). As a result, they are not appropriate for this application.

Although a powerful tool, GIS does not hold all the answers. The technical community and policy-makers must work together to ensure its proper use. In reality, no 25-inch precipitation line floats over Kansas. It is merely the interpretation of scientists and policy-makers who select its location. The only way to arrive at a reasonable answer is to gather the best available information and allow all parties to scrutinize it. GIS can be a wonderful tool to do this.

Acknowledgments

The author wishes to express his gratitude to the Kansas Geological Survey for its computing support before, during, and after his move to Louisiana. Also, the author wishes to thank Mr. M. Schouest, who made the move both possible and tolerable. Mr. C. Johnson of Johnson Controls in Baton Rouge, Louisiana, produced the excellent presentation graphics. Finally, this work could not have been performed without the Internet.

References

1. Mitchell, J.E. 1993. The influence of (x,y) uncertainty on prediction error and contour lines from a three-dimensional surface. Water Resources Bull. 29(5):863-870.

2. Northeast Waste Management Officials Association. 1992. Catalogue and directory of New England states and regional GIS organizations and activities, and an assessment of their future GIS needs. EPA/G-D001456/90-2.

3. Campbell, W.G., and D.C. Mortenson. 1989. Ensuring the quality of geographic information system data: A practical application of quality control. EPA/600/J-89/306.

4. Chrisman, N.R. 1991. The error component in spatial data. In: Maguire, Goodchild, and Rhine, eds. Principles of geographic information systems, Vol. 1. pp. 164-174.

5. Isaaks, E.H., and R.M. Srivastava. 1989. An introduction to applied geostatistics. New York, NY: Oxford University Press.

6. National Center for Geographic Information and Analysis. 1989. The research plan of the National Center for Geographic Information and Analysis. Int. J. Geo. Info. Sys. 3(2):117-136.

7. Vitek, J.D., S.J. Walsh, and M.S. Gregory. 1984. Accuracy in geographic information systems: An assessment of inherent and operational errors. Proceedings of the PECORA IX Symposium. pp. 296-302.

8. Monmonier, M.S. 1977. Maps, distortion, and meaning. Association of American Geographers Resource Paper Number 75-4.

9. Goodchild, M.F., and W. Min-Hua. 1988. Modeling error in raster-based spatial data. Proceedings of the Third International Symposium on Spatial Data Handling, Sydney, Australia (August 17-19). pp. 97-106.

10. Walsh, S.J. 1989. Inherent and operational error within a GIS: Evaluation from a user perspective. In: Goodchild, M.F., ed. Accuracy of spatial databases. p. 290.

11. Burrough, P.A. 1986. Principles of geographic information systems for land resources assessment. Oxford, England: Clarendon Press.

12. Marble, D.F., and D.J. Peuquet. 1983. Geographic information systems and remote sensing: Manual of remote sensing, 2nd ed. Falls Church, VA: American Society of Photogrammetry and Remote Sensing. pp. 923-958.

13. Mead, D.A. 1982. Assessing data quality in geographic information systems. In: Johannsen, C.J., and J.L. Sanders, eds. Remote sensing for resource management. Ankeny, OH: Soil Conservation Society of America. pp. 51-59.

14. Heuvelink, G.B.M., and P.A. Burrough. 1989. Propagation of errors in spatial modeling with GIS. Int. J. Geo. Info. Sys. 3(4):303-322.

15. Dunn, R., A.R. Harrison, and J.C. White. 1990. Positional accuracy and measurement error in digital databases of land use: An empirical study. Int. J. Geo. Info. Sys. 4(4):385-398.

16. Walsh, S.J., D.R. Lightfoot, and D.R. Butler. 1987. Recognition and assessment of error in geographic information systems. Photogrammetric Eng. and Remote Sensing 53(10):1,423-1,430.

17. Owenby, J.R., and D.S. Ezell. 1992. Climatography of the United States: Monthly station normals of temperature, precipitation, and heating and cooling; degree days 1961-1990, Kansas. No. 81. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Climatic Data Center, Asheville, NC (January).

18. Bayne, C.K. 1975. General availability of ground water and normal annual precipitation in Kansas (1:500,000 map). Kansas Geological Survey Map Series M4-A.

19. SAS Institute, Inc. SAS/GRAPH Software: Reference, Version 6, Vol. 1, 1st ed. Cary, NC: Statistical Analysis System Institute, Inc. 1990.

20. SAS Institute, Inc. SAS/GRAPH Software: Reference, Version 6, Vol. 2, 1st ed. Cary, NC: Statistical Analysis System Institute, Inc. 1990.

21. Olea, R.A. 1992. Understanding allays intimidation. Geobyte 7(5):12-17.

22. Bastin, G., B. Lorent, C. Duqué, and M. Gevers. 1984. Optimal estimation of the average areal rainfall and optimal selection of rain gauge locations. Water Resour. Res. 20(4):463-470.

23. ESRI. 1991. Cell-based modeling with GRID. Redlands, CA: Environmental Systems Research Institute, Inc.

24. ESRI. 1991. Surface modeling with TIN, surface analysis, and display. Redlands, CA: Environmental Systems Research Institute, Inc.

25. Davis, J.C. 1982. Statistics and data analysis in geology, 2nd ed. New York, NY: John Wiley and Sons.

26. Mitchell, J.E. 1991. A regionalization approach based on empirical Bayes techniques with applications to hydrologic forecasting, Ph.D. dissertation. Duke University, Durham, NC.

27. Mitchell, J.E., H. Sun, and S. Chakrabarti. 1992. Predicting the form of a surface from random, nonexperimental samples using neural networks. Proceedings of the First Midwest Electro-Technology Conference, Iowa State University, Ames, IA (April 10-11).

28. Congalton, R.G., and D.M. Schallert. 1992. Exploring the effects of vector to raster and raster to vector conversion. EPA/600/R-92/166.

29. Mitchell, J.E., and H. Sun. 1992. Generalization in higher order neural networks when predicting polynomial surfaces, implications for the free-form surface problem. In: Intelligent engineering systems through neural networks, Vol. 2. Proceedings of the Artificial Neural Networks in Engineering (ANNIE '92) Conference, St. Louis, MO (November 15-18).

30. Bates, J.M., and C.W.J. Granger. 1969. The combination of forecasts. Operational Res. Quarterly 20:451-468.

31. Granger, C.J., and R. Ramanathan. 1984. Improved methods of combining forecasts. J. Forecasting 3:197-204.

32. Carr, D.B., A.R. Olsen, and D. White. 1994. Hexagon mosaic maps for display of univariate and bivariate geographical data. EPA/68-C8/0006.

33. Rule, G.K. 1960. Dust bowl. In: The world book encyclopedia, Vol. 4. pp. 315-316.

Design of GIS Analysis To Compare Wetland Impacts on Runoff in Upstream Basins of the Mississippi and Volga Rivers

Tatiana B. Nawrocki
Natural Resources Research Institute, University of Minnesota, Duluth, Minnesota

Introduction

The attention given in hydrologic studies to wetlands differs significantly between the United States and Russia at the present time. Fundamental theories and mathematical models are developed in both countries to describe hydrologic processes and impacts of watershed conditions on surface runoff. In the United States, however, theoretical investigations are directly pointed at wetlands and are supported by large-scale field studies and advanced technological capabilities to manage spatially distributed information. Unlike in Russia, in the United States, special scientific symposia are devoted to wetland hydrology, where major tasks for hydraulic and hydrologic research needs are formulated. Among these tasks are the understanding and assessment of relationships between various hydrologic modifications and wetland functions, especially wetland flood conveyance and water quality protection functions (1). Watershed-scale comprehensive field studies of wetland functions are underway, for example, at constructed experimental wetlands in the Des Plaines River basin in Illinois (2). A new long-term goal—strategic restoration of wetlands and associated natural systems—has been formulated (3).

The intensive efforts of many U.S. scientists yielded numerous results and attracted more attention to the complicated nature of wetlands processes. Wetlands were evaluated as runoff retention basins, and it was found that, in northwestern states, up to 12 inches of water could be accumulated per wetland acre (4, 5). Over time, piecemeal loss and degradation of wetlands in many areas of the United States have seriously depleted wetland resources. Researchers also discovered that adverse impacts from wetland degradation could appear indirectly with little obvious spatial or temporal connections to sources. As described by Johnston (6):

> Cumulative impacts, the incremental effect of an impact added to other past, present, and reasonably foreseeable future impacts, has been an area of increasing concern. . . . Impacts can accumulate over time or over space and be direct or indirect. An indirect impact occurs at a location remote from the wetland it affects, such as the discharge of pollutants into a river at a point upstream of a wetland system.

The process of solving environmental problems related to wetlands is increasingly complex. Analyzing diversified data over increasingly broad areas becomes essential for making competent decisions.

Comparing wetland hydrologic functions in headwaters of the Mississippi River (United States) and the Volga River (Russia) could provide additional information about how alternative management strategies affect runoff, peak flow, and water quality under changing climates. A macro-scale "field experiment" in both of these naturally similar areas is already under way. Wetland conservation as opposed to drainage is now the prevailing policy in the upper Mississippi basin. In Russia, however, economic problems have prevented this type of policy from becoming a priority. Instead, peat mining, reservoir construction on lowlands, and drainage for farming and private gardening are common.

This project, which is being implemented at the Natural Resources Research Institute (NRRI), University of Minnesota, Duluth, has the following goals:

- Developing a multilayered hierarchical base of geographic information system (GIS) data for headwater watersheds of the Mississippi and Volga Rivers.
- Developing a comparative analysis of wetland impacts on the hydrology of the rivers.
- Studying the relationships between natural and human-induced factors on wetland functions under climate change and variable strategies of wetland conservation.

- Defining criteria and thresholds for wetland system stability with regard to flood risk and water quality.
- Outlining recommendations for wetland management in the headwaters.

The methodology for comparative assessments involves statistical analysis, hydrologic models, GIS, and remote sensing. Representative watersheds will be studied in more detail, and procedures for scaling information from the local to the regional level will be developed.

Input Data

In recent years, U.S. governmental and state agencies, as well as a number of private companies, have expended considerable efforts to compile the available data in GIS format for multidisciplinary analysis of watershed problems. Among the major sources of information essential for studies of wetland hydrologic functions are:

- The National Wetlands Inventory, conducted by the U.S. Fish and Wildlife Service.
- Digital elevation maps (DEMs) developed by the U.S. Geological Survey (USGS).
- Major and minor watershed boundaries, outlined for Minnesota by the Minnesota Pollution Control Agency (MPCA).
- The water quality sampling network from the U.S. Environmental Protection Agency (EPA).
- The digital chart of the world (DCW), issued by the Environmental Science Research Institute (ESRI) in scale 1:1,000,000.

These and other sources, listed in Table 1, were used to compile the map illustrations for this paper.

Almost no similar data in GIS form could be found for the territory of the former USSR, however. Any specific data (e.g., detailed maps, hydrology records, water quality sampling data) are generally in paper files dispersed among many agencies and are hard to obtain. The forms of information storage and means of its analysis are out of date, and most maps exist in single or few copies in paper files. In Russia, the time lag grows between the dynamic changes in the environment and the traditional pattern and inertia of management structures.

The GIS situation in Russia developed some interesting paradoxes. During the first few decades of space programs, certain state agencies accumulated an outstanding bank of world image data. When economic hardships hurt the previously privileged space industry, numerous joint ventures with foreign companies were created to distribute images on the world market for hard currency. These data are hardly available for domestic uses, however. The current domestic price for image data is 20,000 to 25,000 rubles for a black-and-white picture of an area 60 by 60 kilometers, or 60,000 to 70,000 rubles for the same image on a computer disk. With the present level of funding for scientific research, the price is too high. Security regulations still restrict access to later data, showing land use changes.

Another paradox is scientists' attitude toward their data. Abandoned by the state, agencies and institutes are reluctant to share their specific data in multidisciplinary projects. Data files are now a commodity. Accomplishing an overlay and integration of special data coverages, which is essential to any watershed GIS study, is almost impossible.

The third paradox is the attitudes of local, regional, and central authorities toward GIS. Many authorities are still ignorant about the potential of this technology. Those who are knowledgeable prefer not to promote GIS for watershed-related tasks because it involves land use analysis. With the onset of land privatization, the best pieces of property (e.g., the waterfront lots adjacent to drinking water reservoirs in the Moscow region) are rapidly allocated to the most powerful landowners. Thus, limiting access to this kind of information is deemed safer.

Experts in Russia have not yet applied GIS to wetland hydrology studies because GIS is still a very new and mostly unfamiliar technology. This makes the current study unique both for its results and for its application of GIS methodology.

Closer review of data sources indicates that most of the input data for the project is available, though dispersed among many agencies. Table 1 is a preliminary list of data and data sources.

Project Design

The project addresses the following questions:

- How do the extent and positioning of wetlands in the headwaters of large rivers affect runoff and peak flow?
- What are the spatial relationships between wetland and other land uses regarding flood risk and water quality under variable climate conditions?
- What is the role of wetlands for diffuse pollution prevention and sediment deposition control under alternative management?
- What determines major criteria for wetland conservation in headwaters, ensuring environmentally sustainable development under multiobjective land and water resource uses?

Table 1. Data Sources for a GIS Study of Wetlands in the Basins of the Mississippi (United States) and the Volga (Russia)

Data	United States	Russia
Level 1		
Base maps	DCW[a]	DCW[a]
Stream network	DCW[a]	DCW[a]
Urban and rural areas	DCW[a]	DCW[a]
Wetlands, unclassified	NWI[a]	DCW[a]
Forests	DCW[a]	MGU[b]
Agricultural lands	LMD[b,c]	MGU[b]
Level 2		
Watershed boundaries	MPCA[a]	RWRC[c]
Digital elevation maps	USGS[a]	NA
Digital orthophotos	USGS, LMIC[a]	CD, RPI[b,c]
Soils	SCS[b,c]	RPI[c]
Hydrologic records	USGS[b]	CHM[b,c]
Water quality records	EPA/MPCA[a,b]	CHM, RCP, RPI[c]
Land uses	LMIC, LSAT[a]	LSAT, CD, RPI[a-c]
Wetlands, classified	NWI[a]	RPI[c]

[a]Data available in GIS ARC/INFO format.
[b]Databases; needed conversion to ARC/INFO.
[c]Data available in paper files; needed digitizing.

Key:

CD = Commercial distributors
CHM = Russian Committee on Hydrometeorology
DCW = ESRI digital chart of the world
EPA = U.S. Environmental Protection Agency
LSAT = Satellite image data
LMD = Published literature and map data
LMIC = Minnesota Land Management Information Center
MGU = Moscow State University
MPCA = Minnesota Pollution Control Agency
NA = Data not available
NWI = U.S. National Wetland Inventory
RCP = Russian State Committee on Natural Resources and Conservation
RPI = Miscellaneous planning agencies and research institutes
RWRC = Russian Water Resources Committee
SCS = U.S. Soil Conservation Service
USGS = U.S. Geological Survey

Tasks established to address these questions include:

- Developing a multilayered hierarchical base of GIS data for headwater watersheds of the Mississippi and Volga basins.
- Performing a comparative analysis and simulation of wetland impacts on hydrology and water quality at representative watersheds.
- Deriving the relationships between natural and human-induced factors and wetland functions under climate change with regard to variable strategies of wetland conservation.
- Defining the criteria and thresholds for wetland system stability with regard to flood risk and water quality parameters.
- Outlining recommendations for land and water resources management and wetland positioning in the headwaters.

GIS is the essential tool for manipulating and integrating the many types of spatial data on water resources, soils, vegetation, land use, economics, and the environment. GIS compiles many sources (e.g., maps, field notes, remote sensing, statistical data) into a consistent, interpretable database used for specific scientific goals and development decisions. The user can run GIS ARC/INFO software on workstation and PC platforms and apply the hierarchical approach to GIS data management, developed earlier (7). At the task level, data resolution and corresponding modeling tools vary.

Level 1 contains the basic reference information for large regions (e.g., the Upper Volga and the Minnesota portion of the Upper Mississippi River basins). It covers an area of several hundred thousand square kilometers, with a map scale approaching 1:1,000,000. Landsat thematic map data and the DCW (8) are used as sources of data at this level. Vogelmann et al. (9) demonstrated the methodology for detection of freshwater wetlands using remote sensing data based on maximum likelihood supervised classification. A graphic data file on GIS focuses on basic physical characteristics such as stream network, geology, soils, wetland classification, and other major land uses. A complementary tabular database or attribute file contains information on stream flow, water quality, pollution sources, and wetland impacts on material fluxes. At this level, the general physiographic and statistical information is accumulated and analyzed, territories are classified, and major problems and typical case study watersheds are defined. This information is compiled from literature, cartographic data in paper and digitized form, statistics, and space image data.

In Level 2, the more detailed GIS analysis and scenario-based modeling is implemented at the watershed scale with a map scale of approximately 1:25,000. The watershed demonstration focuses on alternative approaches to priority-setting in wetland management, climate impact analysis, and resulting interactions with landform, soils, biosphere, and runoff. The sources of data are special, topographic maps and air photo interpretation.

Simulation studies of water balance and fluxes among the various reservoirs are implemented at Level 2. Developing procedures for scaling information from the local to regional level is the important task at this level. GIS assists in handling the input parameter library and analyzing the output. GIS studies, involving area measurements and distribution analysis, evaluate cumulative impacts on runoff and its quality from the loss of wetland area, caused by drainage or filling, under stationary and changing climate.

Wetland functions are considered under two sets of scenarios. Management scenarios compare different wetland and farming allocations, conservation practices, and agricultural chemical use. Climate scenarios assume rainfall and temperature changes under global warming. Scenario-based simulation is applied in the analysis of watershed runoff, wetland moisture regime, soil erosion, and water quality processes.

Methodology

GIS database structure is related to the selected methodology. GIS serves as a linking tool for input-output data analysis and transfer between models, used at different levels and stages of studies.

Scientists in both the United States and Russia developed statistical methods to obtain quantitative relationships between stream flow and wetland area in the river basins. Johnston (6) summarized the U.S. findings:

> Empirical equations for predicting streamflow, developed by U.S. Geological Survey in Wisconsin and Minnesota, indicate that flood flow is proportional to the negative exponent of wetlands and lakes ratio on a watershed (10, 11). This means that relative flood flow is decreased greatly by having some wetlands in a watershed, but a watershed with a large proportion of wetlands does not reduce flood flow much more than a watershed with an intermediate proportion of wetlands. For example, predicted flood flow was 50 percent lower in Wisconsin watersheds with 5 percent lakes or wetlands than it was in watersheds with no lakes or wetlands, but increasing the proportion of lakes and wetlands to 40 percent decreased relative flood flow by only an additional 30 percent (12).

Other estimates agree that wetland encroachment on a watershed of less than 25 percent generally has a minimum influence on peak flow (5, 13, 14).

Johnston and colleagues (15) applied these equations to watersheds in central Minnesota. They found that a watershed with 1.6 percent lakes and wetlands had a flow per unit watershed area that was 10 times the flow predicted for a watershed with 10 percent lakes and wetlands, while watersheds with 10 to 50 percent lakes and wetlands had about the same flood flow per unit area.

Statistical analysis indicates that peak discharge increases with decreasing wetland area within the drainage basin. The regression equation defines the approximation for northwestern Minnesota (16):

$$Q_{AM} = 58.4\, A_W^{0.677}\, (L_S)^{-0.506}$$

$$L_S = 100(A_L + A_M)/A_W + 1$$

where:

Q_{AM} = arithmetic mean of the annual series, cubic feet per second.

A_W = watershed area, square miles.

A_L = lake area within the watershed, square miles.

A_M = marsh area within the watershed, square miles.

A similar statistical approach was developed for peak flow determination in Russia. Maximum flow discharge from snow melt is calculated for the central European zone as (17):

$$Q_m = K_0 * h_p * S_1 * S_2 * S_3 / (A + 1)^n$$

where:

Q_m = flow discharge, cubic meters per second.

K_0 = coefficient of flood concurrence, K0 = 0.006 for plain river basins.

h_p = calculated flood runoff for given probability, millimeters.

A = drainage area, square kilometers.

n = coefficient, n = 0.17.

S_1 = lake storage coefficient, if lake area is less than 1 percent of A, then S1 = 1.

S_2 = pond and reservoir storage coefficient, S_2 = 0.9 with ponds and S_2 = 1 without ponds.

S_3 = combined wetland and forest storage coefficient.

S_3 = 1 - 0.8 lg (0.05 Sf + 0.1 S_w + 1), where S_f and S_w are forest and wetland area, percentage to total drainage area.

Calculated flood runoff for a given probability, h_p, is determined based on average flood runoff h, millimeters, coefficient of variation C_v, and tabulated parameter F, as follows:

$$h_p = (1 + F * C_v)\, h$$

$$h = K_t\, h_k$$

h_k = 100 millimeters for Moscow region.

K_t = land surface coefficient, Kt = 0.9 for plains and sandy soils, K_t = 1.1 for hills and clay soils.

The studies mentioned above generally agree with an assumption that the incremental loss of wetland area would have a small effect on flood flow from watersheds with 10 percent up to 40 to 50 percent wetlands, but a large effect on flood flow from watersheds with less than 10 percent wetlands.

The existence of similar thresholds was found in relation to wetlands abilities to intercept pollutants. As Johnston stated (6),

> The same 10 percent threshold was identified by Oberts (18) for suspended solids, a measure of water quality function. Stream-water draining watersheds having 10 to 20 percent wetlands had about the same loading of suspended solids, so the contribution of suspended solids was relatively constant per unit area of watershed. However, the watersheds with less than 10 percent wetlands had loading rates per unit area that were as much as 100 times greater than the loading rates from the watersheds with more than 10 percent wetlands.

GIS could be especially helpful in determining the impacts on downstream water quality of the spatial positioning of wetlands within watersheds. Studies prove that the location of wetlands can affect their cumulative function with regard to water quality (6). In an earlier work (15), Johnston developed an index of wetland location and applied it to a landscape-level GIS study of urban and rural stream watersheds in central Minnesota. The index is formulated as:

$$PWP = \sum_{i=1}^{j} A_i / \sum_{i=1}^{j} A_i$$

where:

PWP = relative wetland position.

j = stream order (19) of water quality sampling station.

i = stream order of wetland.

Ai = area of *i*th order wetlands.

Calculated values for the index ranged from 0 (i.e., all wetlands were on streams of the same order as that of the sampling station) to 2.6 (i.e., average wetland position was 2.6 stream orders upstream of the sampling station). Watersheds with wetlands located close to sampling stations had significantly better water quality (i.e., lower concentrations of inorganic suspended solids, fecal coliform, and nitrate; lower flow weighted concentrations of ammonium and total phosphorus) than watersheds with wetlands located far from sampling stations (6).

The review of methodological approaches, as shown above, indicates that parameters describing wetland extent, positioning, and land surface characteristics are of universal significance for any comprehensive watershed-scale wetland study.

In the current study, GIS is used at Level 1 to evaluate wetland area per watershed and to develop input parameters for relative wetland position assessment. The comparison and selection procedures for case study watersheds in the Volga and Mississippi basins are based on these values. The parameters, derived from GIS, are as follows:

1. Total watershed area.
2. Lake, pond, and reservoir area.
3. Forest area.
4. Wetland area.
5. Ratio of wetland area to total watershed area.
6. Wetland area by subwatersheds of different order.
7. Relative wetland extent by subwatersheds of different order.
8. Land surface coefficients.

Parameters listed in groups 1 through 4 are obtained directly from GIS attribute tables as values of “area” items for the respective land cover polygons. Parameters 5 through 7 require calculations relating values of area items for different polygon coverages. Land surface coefficients (group 8) could be determined indirectly based on basic soil, land cover, and topography data. Most U.S. methodologies use hydrologic soil groups, based on soil permeability, rates of infiltration, and Soil Conservation Service (SCS) runoff curve numbers (20). Some Russian methodologies have adopted similar empirical land surface coefficients. For example, for central European Russia, this value varies from $K_t = 0.9$ for plains and sandy soils, to $K_t = 1.1$ for hills and clay soils (17).

Level 2 of analysis applies two hydrologic simulation models:

- The Agricultural Watershed Runoff and Water Quality Model (Agricultural Nonpoint Source Pollution Model [AGNPS]), developed by the Agricultural Research Service of the U.S. Department of Agriculture, contains explicit procedures to evaluate the impacts of management practices and landscape feature positioning on watershed runoff. AGNPS is a cell-based runoff model that estimates water volume, peak flow, eroded and delivered sediment, chemical oxygen demand, and nutrient export from watersheds (20-22).
- The Forest Runoff Watershed Model (FRWM) combines analytical and numerical methods for solving hydro- and thermodynamics equations (23). This model considers snow melt constituent in runoff in more detail than does AGNPS. Hydrologic simulation is based on physical process descriptions for snow cover dynamics, freezing and thawing of soil, soil moisture dynamics in frozen and thawed soils, interception of liquid and solid precipitations by vegetation, surface runoff, ground-water aquifers, and channeled streams.

Both models use a similar set of watershed input data, derived from GIS (e.g., elevations, slopes, channel slopes, stream network configuration, soil texture, land cover). The methodology, linking GIS with hydrologic models, was already tested in the wetland study project at the Voyageurs National Park in Minnesota. The ARC/INFO GRID module was used to derive watershed variables for input to AGNPS. GIS then presented and interpreted the scenario-based results of the simulation (24). The typical stages of such an analysis and interpretation for a watershed-scale area are presented in Figures 1 through 5.

Case Study Watersheds

The areas where wetland impacts on runoff are evaluated are located in Minnesota (United States) and Moscow and adjacent regions (Russia) (see Figures 6 through 15). They have mixed urban, rural, recreational, and forest land uses. Both regions have a variety of development pressures. The relative effects of different alterations in watershed management are distinguished and quantified. GIS provides metrics for comparative assessments and analysis of related variables for both areas.

Table 2 and Figures 7 through 14 present a general overview of wetland extent in both areas. The case study subwatersheds used for more detailed analysis will include tributaries of the second and third order. At this stage, several watersheds are considered for more detailed analysis. The limitations imposed by data availability as well as by project resources could affect the final selection. Table 2 serves, therefore, as a preliminary overview of several areas that could potentially be adopted for more detailed studies.

GIS analysis shows that in the Upper Volga, wetlands extent very much depends on allocation of populated areas. The heavily urbanized Moscow metropolitan area affects a large territory of many thousands of square kilometers. The ratio of wetlands as a percentage of total land area is one-tenth of that in the neighboring Tver area, which has the same size but a smaller population (see Figure 11). In areas of intensive agriculture (e.g., the Pronya basin located southeast of Moscow), almost all wetlands were drained and have not existed for several decades.

In Minnesota, despite the growing urbanization (e.g., the Twin Cities area [7,330 square kilometers]), about half of the presettlement wetlands still remain (25); wetlands occupy 442 square kilometers, or over 6 percent of the land area; and shallow lakes constitute an additional 114 square kilometers (1.56 percent). Some watersheds within the Twin Cities metropolitan area have a high wetland percentage, such as 18.9 percent in the Lamberts Creek watershed. Intensive studies with GIS application of landscape feature functioning and wetland impacts on stream flow and water quality demonstrated an innovative approach and made detailed databases available for this area (15).

Preliminary comparative analysis indicates that two pairs of case study watersheds could be initially selected for further studies in the Mississippi and Volga basins:

- Upstream watersheds with wetlands area of 15 to 20 percent (Tver region in Russia and Cass and adjacent counties in Minnesota).
- Tributary watersheds downstream with wetlands area of 1 to 2 percent (the Istra basin in Russia and subwatersheds of the Minnesota River basin, located in Sibley, Scott, and adjacent counties in Minnesota).

Case study watersheds in the Mississippi and Volga basins are situated on gently rolling plains in mixed forest zones with southern portions extending into the forest/steppe and prairies. The Quaternary sediments are of glacial, glaciofluvial, lacustrine, and alluvial origin. Wetlands have hydric soils with various degrees of gley process development and/or peat accumulation, varied by wetland type and soil moisture regimen (28). The annual precipitation is 500 to 600 millimeters with similar

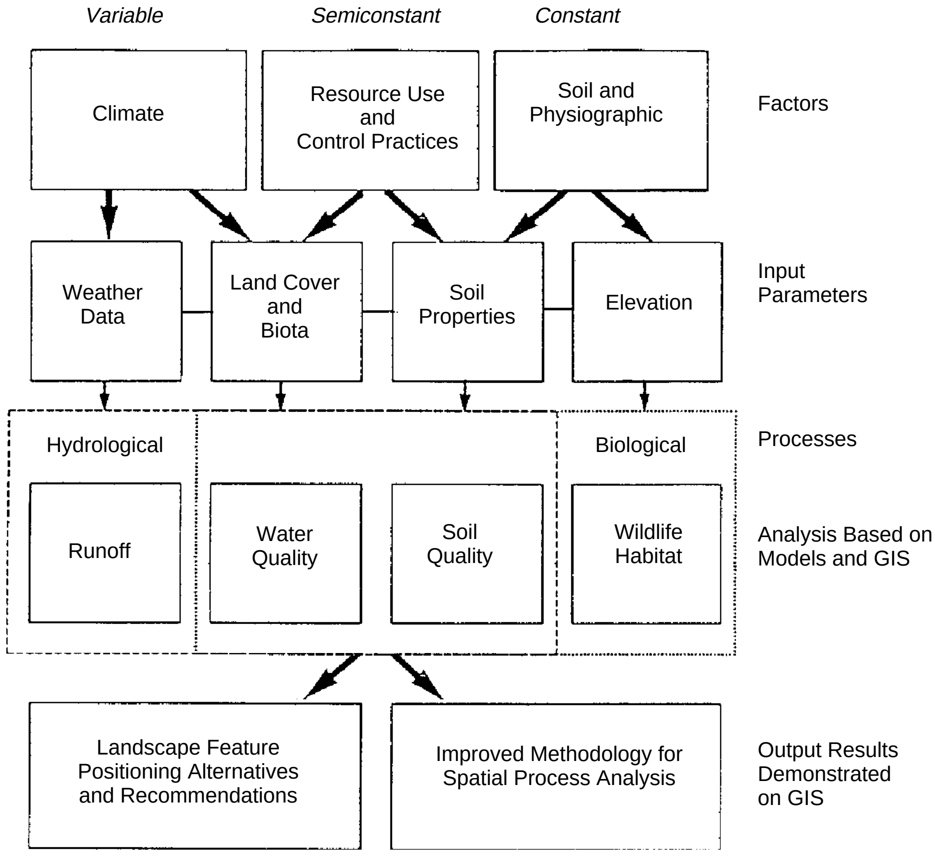


Figure 1. Conceptual framework of linking GIS and models for environmental management.

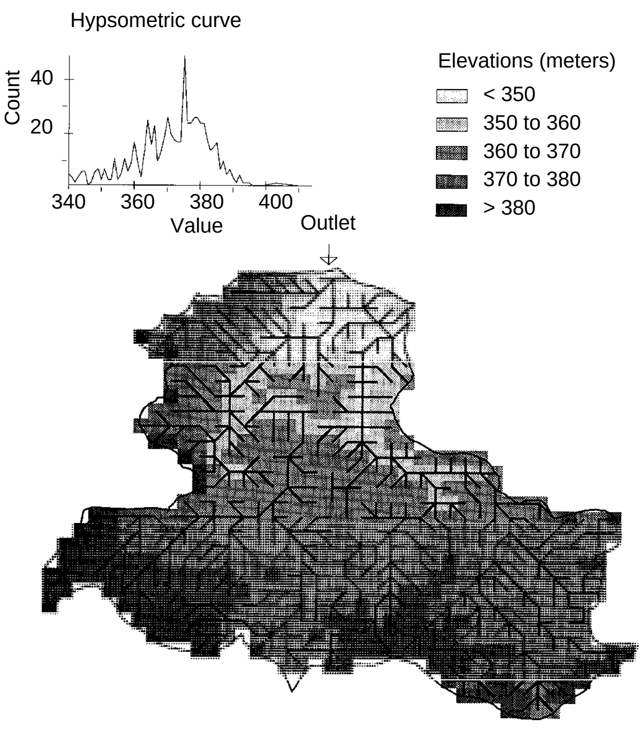


Figure 2. Stream network configuration derived from GIS elevation map.

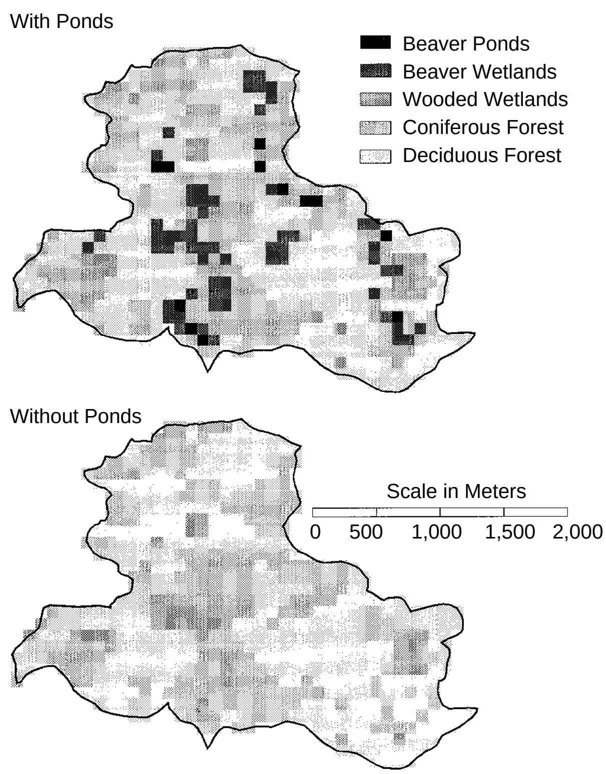


Figure 3. Scenarios of land use.

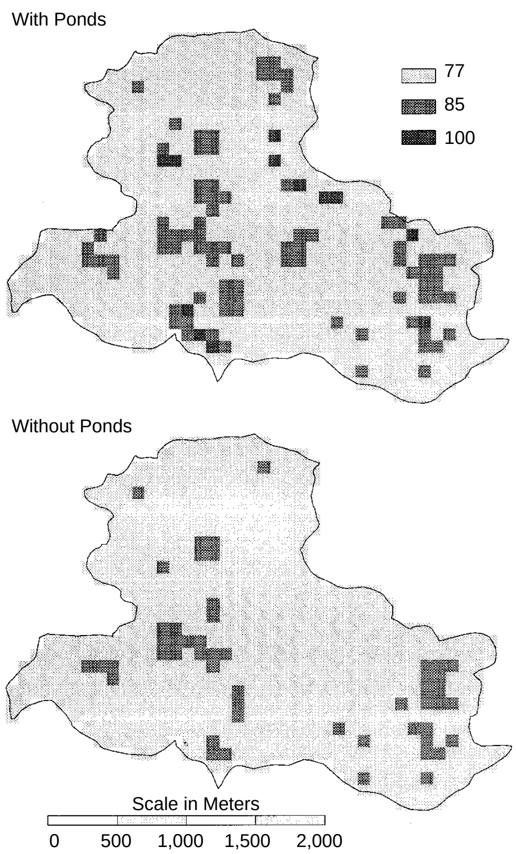


Figure 4. Scenario-related hydrologic curve numbers.

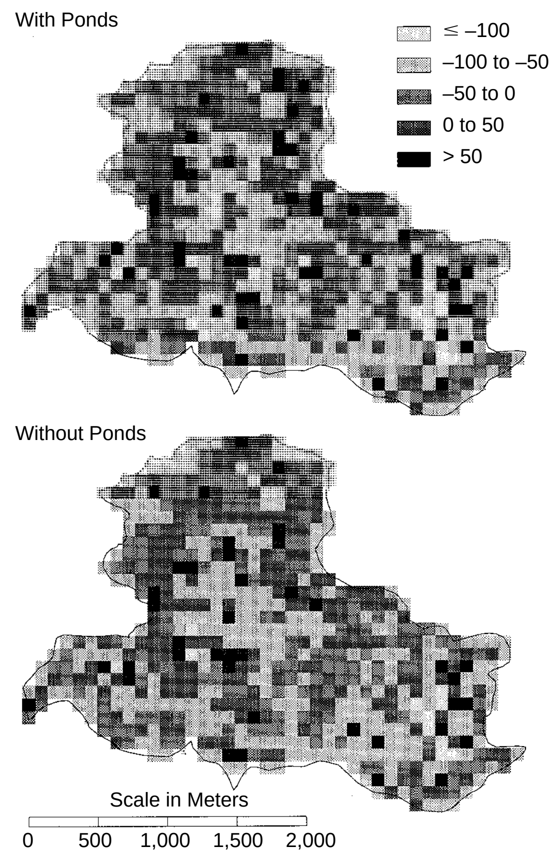


Figure 5. Patterns of sediment transfer between cells (%), + deposition, – erosion.

Table 2. Comparative Data on Wetland Extent in Minnesota and in the Upper Volga Basin (8, 25-27)

Region	Total Area (Square Kilometers)	Wetland Area (Square Kilometers)	Percentage
United States			
Minnesota	205,940.30	30,500.00	14.80
Beltrami Co.	7,923.04	3,909.23	49.34
Cass Co.	6,256.38	1,505.29	24.06
Hubbard Co.	2,624.81	283.22	10.79
Le Sueur Co.	1,204.82	28.31	2.35
Hennepin Co.	1,588.07	36.37	2.29
Sibley Co.	1,555.38	24.26	1.56
Wright Co.	1,852.97	24.27	1.31
Scott Co.	982.56	8.06	0.82
Lambert Creek	9.51	3.69	18.90
Russia			
Tver region	10,000.00	1,169.08	16.90
Moscow region	10,000.00	165.39	1.70
Istra basin	1,827.38	24.07	1.32
Pronya basin	10,200.00	[a]	[a]

[a]Wetland area is insignificant and not identified by available maps.

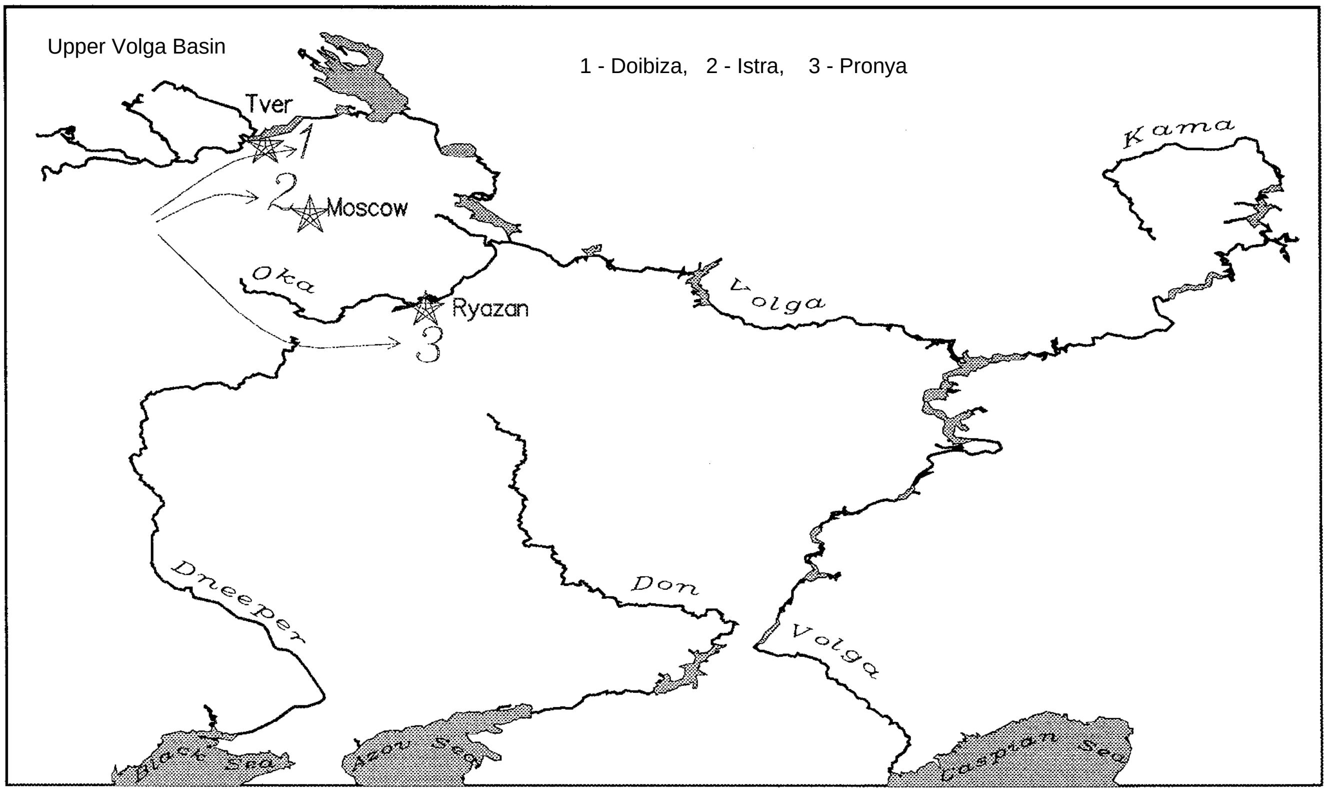


Figure 6. Location of study areas in the Volga basin.

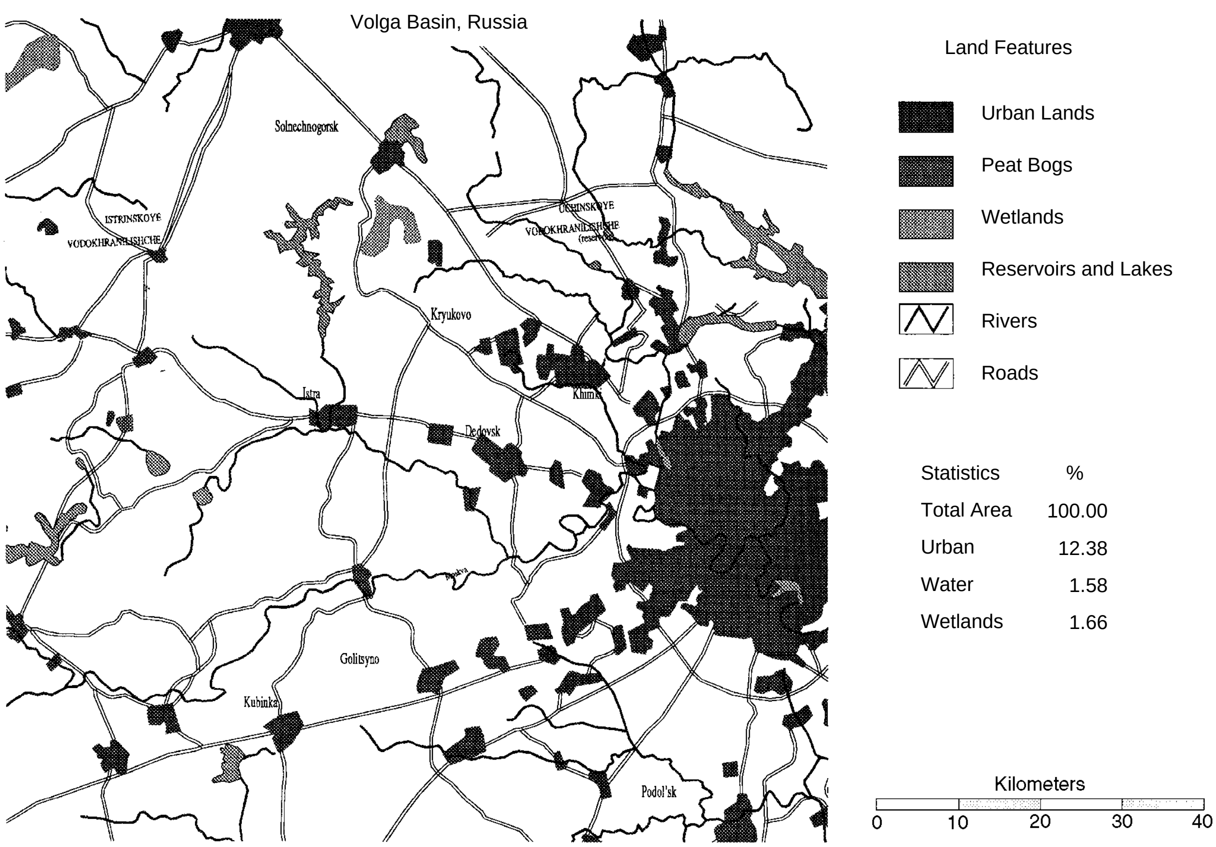


Figure 7. Wetlands and urban lands in the Moscow region.

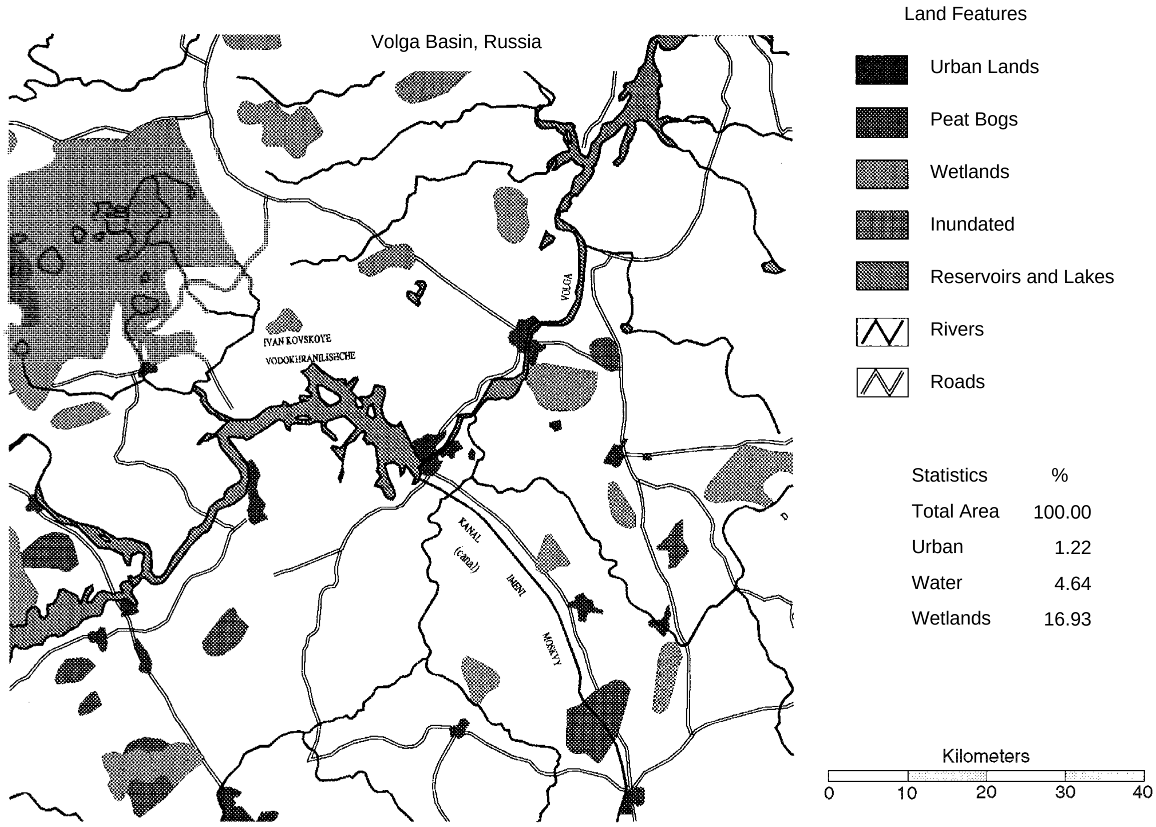


Figure 8. Wetlands and urban lands in the Tver region.

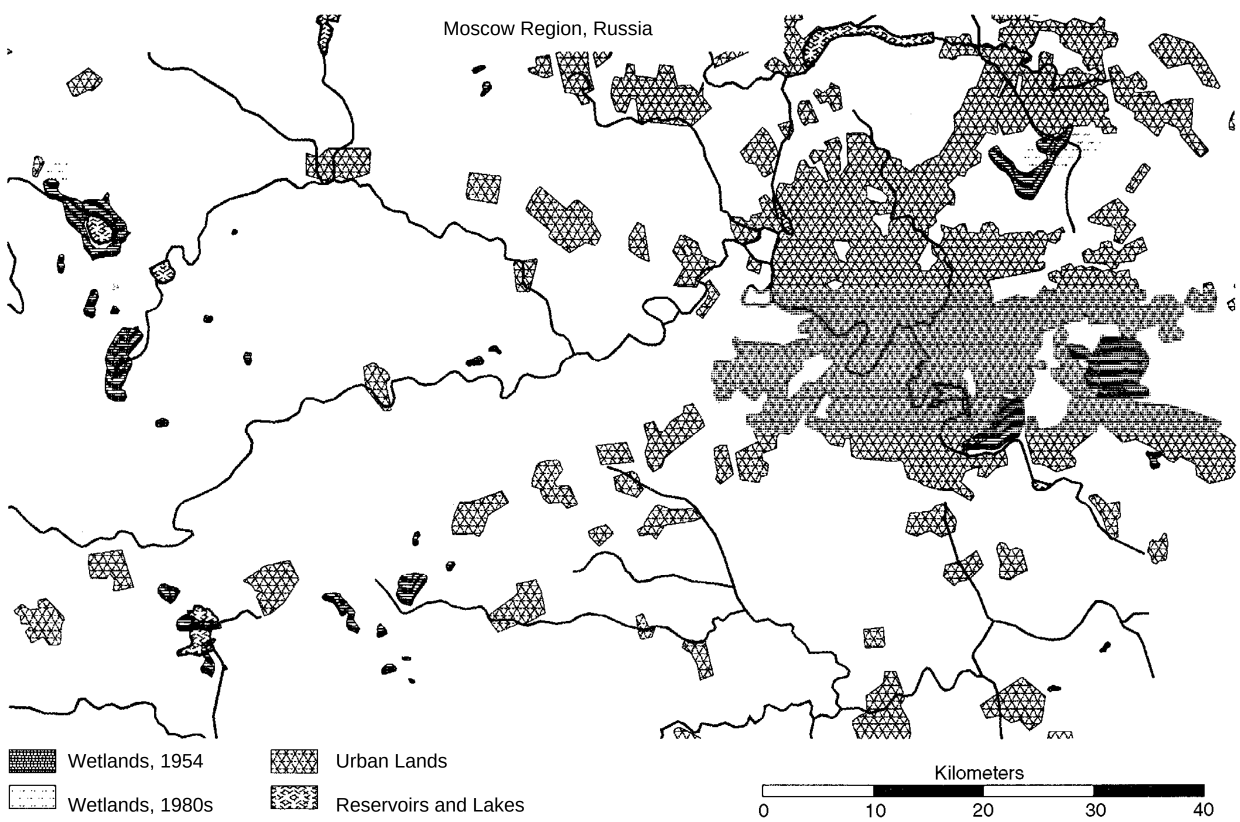
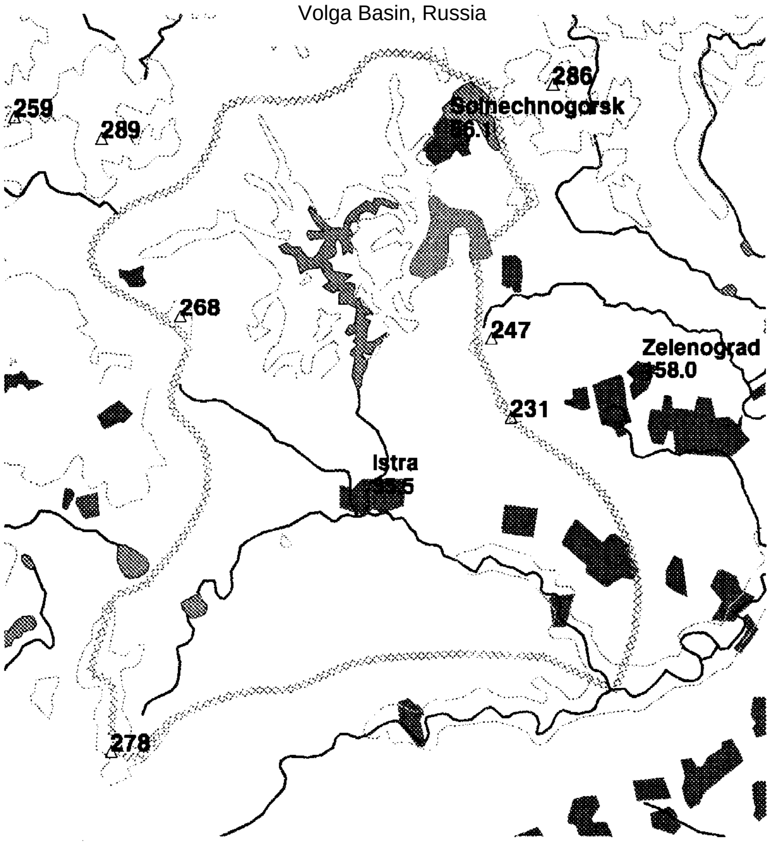
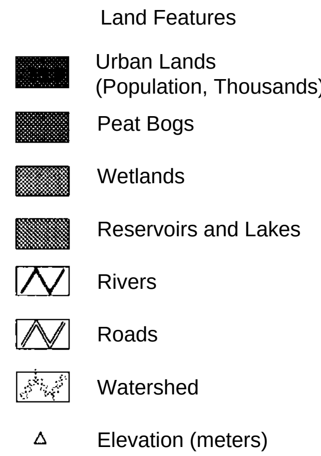


Figure 9. Wetland decline since 1954, Moscow region.





	Square Kilometers	%
Total Area	1,827.38	100.00
Urban	61.31	3.36
Water	40.43	2.21
Wetlands	24.07	1.32

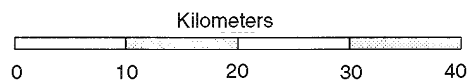


Figure 10. Istra watershed in the Moscow region.

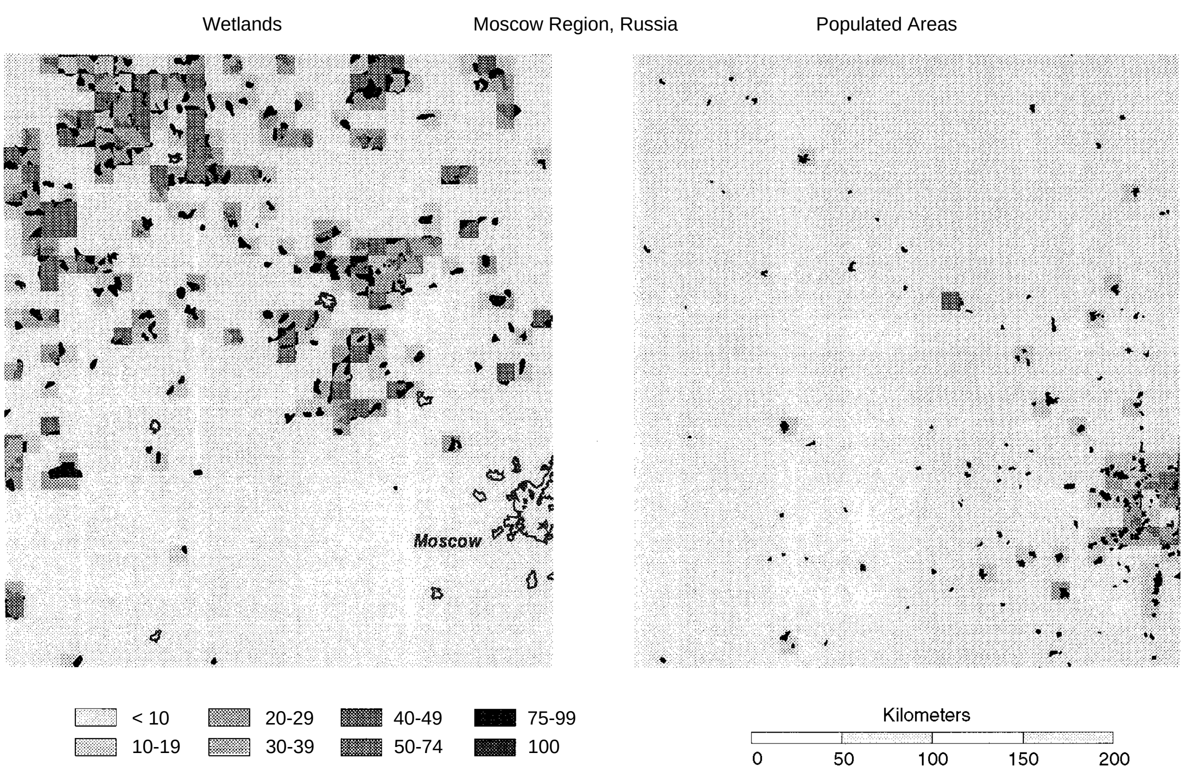


Figure 11. Land cover, percentage of total in Moscow and Tver regions.

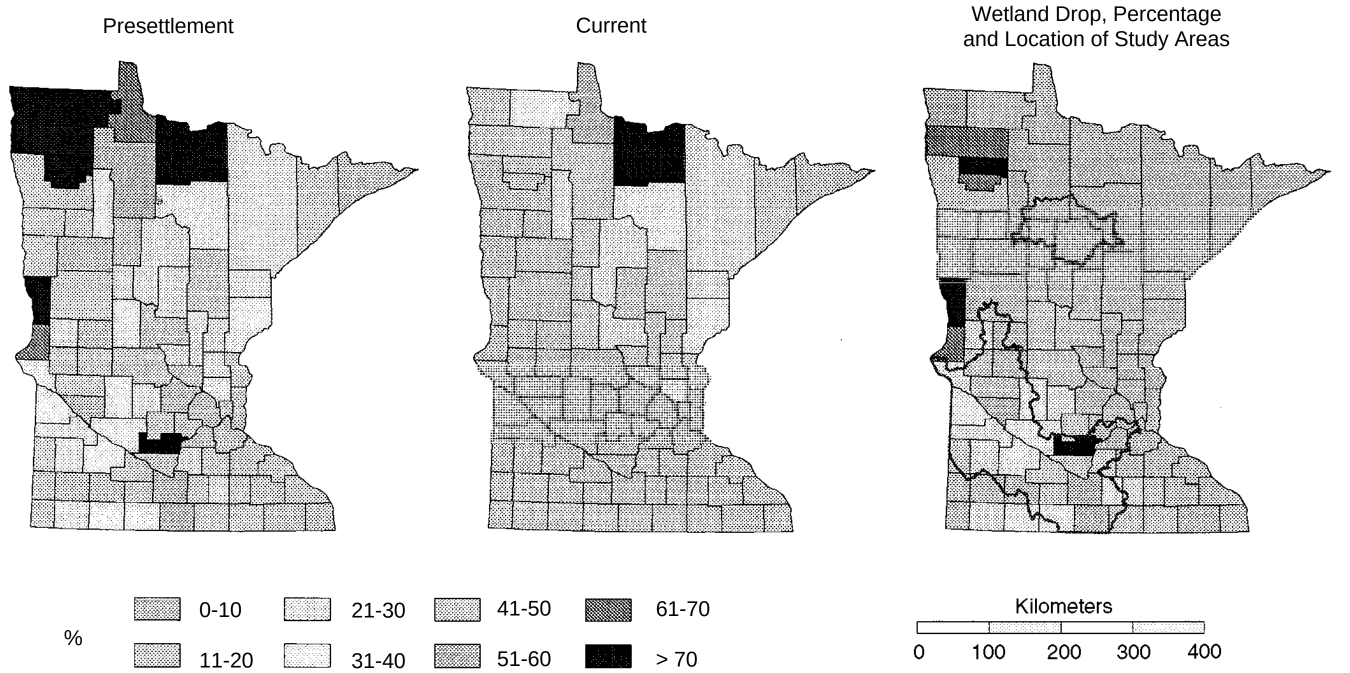


Figure 12. Wetlands in Minnesota, percentage of total area.

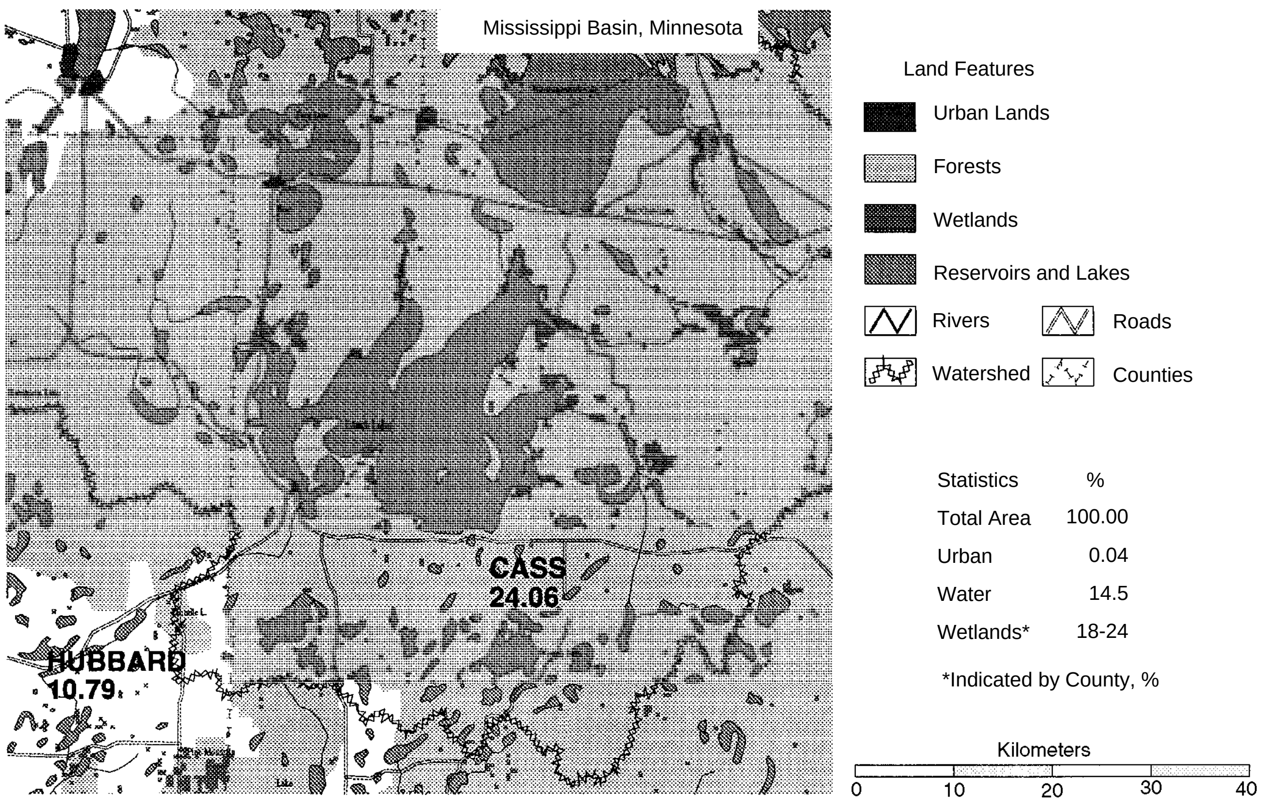


Figure 13. Wetlands and urban lands in Cass County area.

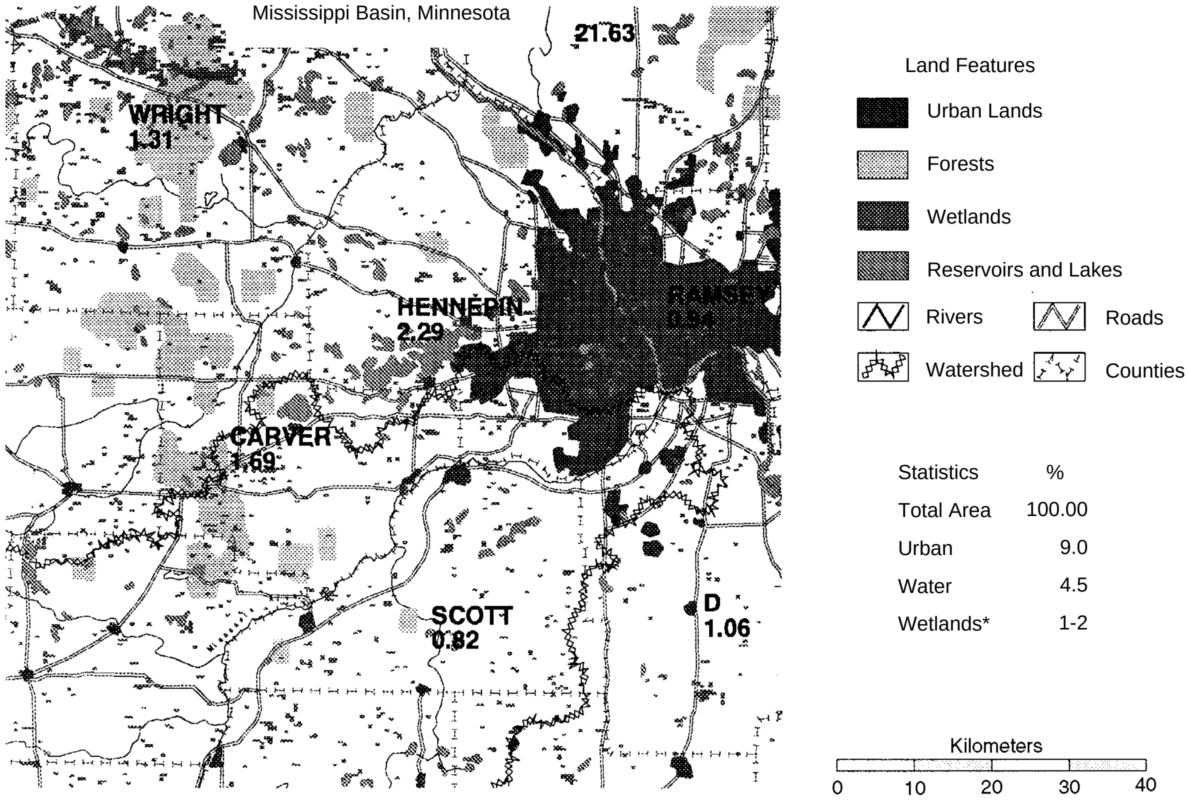


Figure 14. Wetlands and urban lands in Twin Cities area.

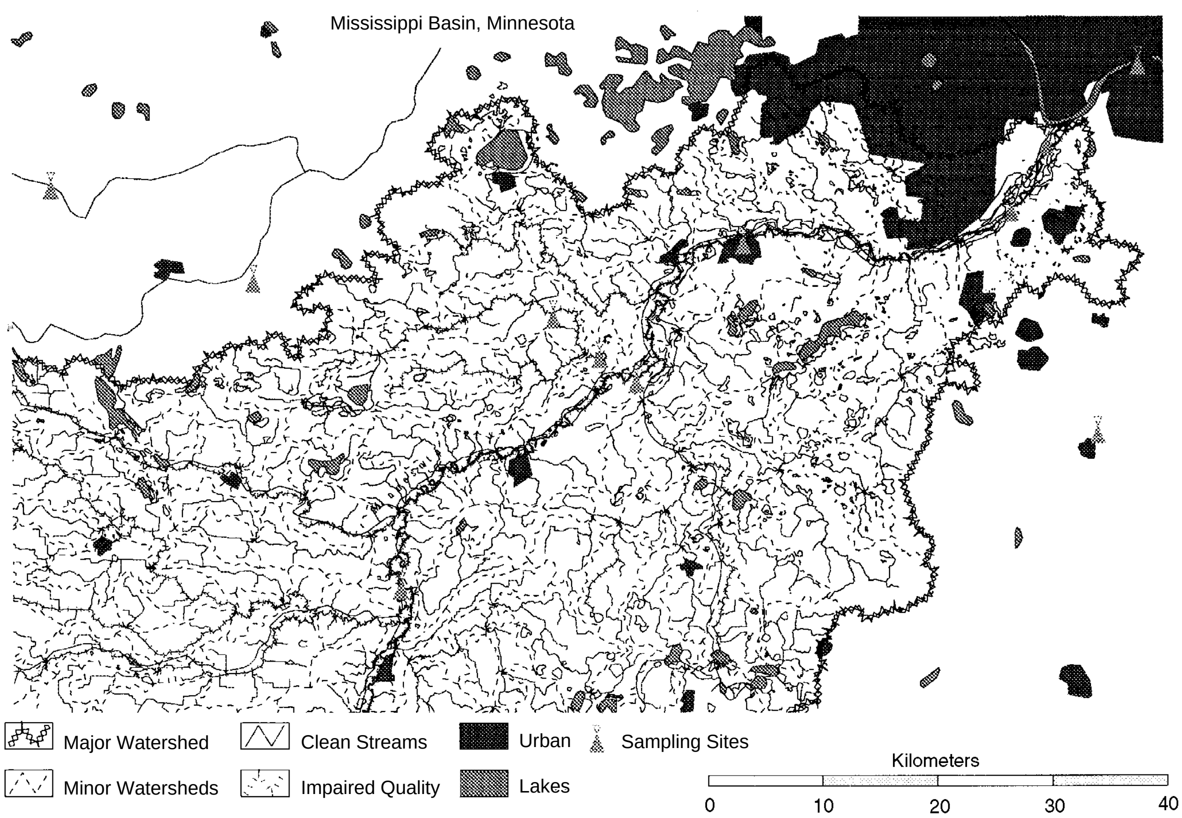


Figure 15. Minnesota River watershed in Twin Cities area.

seasonal distribution in both areas. Average runoff ranges from 150 to 250 millimeters (7, 29).

Conclusion

The current status of the project indicates that most of the input data is available, though dispersed among many agencies. In both the United States and Russia, research methodologies have been developed and applied to study landscape feature impacts on runoff quantity and quality based on simulation and statistical analysis. The comparative analysis of hydrologic and diffuse pollution processes on watersheds in the Upper Mississippi and Upper Volga basins will allow derivation of metrics of wetland loss relative to impacts on runoff and water quality.

The applications of GIS to watershed hydrology are currently much more advanced in the United States than in Russia. Initiatives emerging in the United States, however, could considerably promote GIS use in Russia. Such promotion is beneficial for several reasons. First, this kind of cooperative political activity is in full compliance with the 1992 Freedom Support Act, approved by the U.S. Congress. Second, support of GIS as a new information technology will create a favorable infrastructure in many bilateral economic fields and businesses. Third, a better meshing of the GIS systems in the two countries will lead to further international cooperation in responding to global changes.

Project implementation also helps meet the goal of providing a basis for sound environmental, technical, and economic decision-making on the use of natural resources. This knowledge is essential in developing practical guidelines for sustainable economic development through applied research and technologies.

Acknowledgments

The Water Quality Division of MPCA provided valuable assistance with GIS data for Minnesota. The National Science Foundation (NSF/EAR-9404701) contributed research support. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

1. Wetland management: Hydraulic and hydrologic research needs. 1987. In: Wetland hydrology: Proceedings of the National Wetland Symposium, Chicago, IL (September 16-18). Association of State Wetland Managers, Inc. pp. 331-333.
2. Sather, J.H. 1992. Intensive studies of wetland functions: 1990-1991 research summary of the Des Plaines River Wetland demonstration project. Technical Paper No. 2. Wetland Research, Inc.
3. Hey, D.L. 1985. Wetlands: A strategic national resource. National Wetlands Newsletter 7(1):1-2.
4. Kloet, L. 1971. Effects of drainage on runoff and flooding within the Pembina River basins, North Dakota and Manitoba. Unpublished report by the Bureau of Sport, Fisheries, and Wildlife (December).
5. Simon, B.D., L.J. Stoerzer, and R.W. Watson. 1987. Evaluating wetlands for flood storage. In: Wetland hydrology: Proceedings of the National Wetland Symposium, Chicago, IL (September 16-18). Association of State Wetland Managers, Inc. pp. 104-112.
6. Johnston, C.A. 1994. Cumulative impacts to wetlands. Wetlands 14(1):49-55.
7. Higgins J.M., T.B. Nawrocki, and N.A. Nazarov. 1993. Hierarchical approach to integrated watershed management: Joint TVA/Russian demonstration project. Proceedings of AWWA CONSERV'93 Conference, Sessions 4B-1 through 7C-3, Las Vegas, NV. pp. 1,177-1,197.
8. Environmental Science Research Institute. 1993. Digital chart of the world.
9. Vogelmann, J.E., F.R. Rubin, and D.G. Justice. 1991. Use of Landsat thematic mapper data for fresh water wetlands detection in the Merrimack River watershed, New Hampshire (unpublished).
10. Conger, D.H. 1971. Estimating magnitude and frequency of floods in Wisconsin. Open File Report. Madison, WI: U.S. Geological Survey.
11. Jacques, J.E., and D.L. Lorenz. 1988. Techniques for estimating the magnitude and frequency of floods in Minnesota. Water Resources Investigation Report 87-4170. St. Paul, MN: U.S. Geological Survey.
12. Novitzki, R.P. 1979. Hydrologic characteristics of Wisconsin's wetlands and their influence on floods, stream flow, and sediment. In: Greeson, P.E., J.R. Clark, and J.E. Clark, eds. Wetland functions and values: The state of our understanding. Minneapolis, MN: American Water Resources Association. pp. 377-388.
13. Larson, L.A. 1985. Wetlands and flooding: Assessing hydrologic functions. Proceedings of the National Wetland Assessment Symposium, Portland, ME (June 17-20). pp. 43-45.
14. Ogawa, H., and J.W. Male. 1986. Simulating the flood mitigation role of wetlands. J. Water Res. Planning and Mgmt. 112(1).
15. Johnston, C.A., N.E. Detenbeck, and G.J. Niemi. 1990. The cumulative effect of wetlands on stream water quality and quantity: A landscape approach. Biogeochemistry 10:105-141.
16. Moore, I.D., and C.L. Larson. 1979. Effects of drainage projects on surface runoff from small depressional watersheds in the North Central Region. WRRC Bulletin 99. Water Resources Research Center, University of Minnesota.
17. Maslov, B.S., I.V. Minaev, and K.V. Guber. 1989. Reclamation manual. Rosagropromizdat (in Russian).
18. Oberts, G.L. 1981. Impacts of wetlands on watershed water quality. In: Richardson, B., ed. Selected proceedings of the midwest conference on wetland values and management. Navarre, MN: Freshwater Society. pp. 213-226.
19. Horton, R.E. 1945. Erosion development of streams. Geol. Soc. Amer. Bull. 56:281-283.
20. Chow, V. 1964. Handbook of applied hydrology. New York, NY: McGraw-Hill.
21. Wishmeier W.H., and D.D. Smith. 1978. Predicting rainfall erosion losses: A guide to conservation planning. Agriculture Handbook No. 537. U.S. Department of Agriculture.

22. Young, R.A., C.A. Onstad, D.D. Bosch, and W.P. Anderson. 1987. AGNPS: Agricultural nonpoint-source pollution model. U.S. Department of Agriculture Conservation Research Report 35.

23. Nazarov, N.A. 1988. Model formation of the flood hydrograph of Northern Plain Rivers. Water Resour. 15(4):305-315.

24. Nawrocki, T., C. Johnston, and J. Salés. 1994. GIS and modeling in ecological studies: Analysis of Beaver Pond impacts on runoff and its quality. Voyageurs National Park, Minnesota, case study. NRRI Technical Report 94/01 (February).

25. Anderson, J.P., and W.J. Craig. 1984. Growing energy crops on Minnesota wetlands: The land use perspective. Minneapolis, MN: Center for Urban and Regional Affairs, University of Minnesota. pp. 1-88.

26. Mitsch, W.J., and J.G. Gosselink. 1986. Wetlands. New York, NY: Van Nostrand Reinhold.

27. Brown, R.G. 1987. Effects of wetland channelization on storm water runoff in Lamberts Creek, Ramsey County, Minnesota. In: Proceedings of the National Wetland Symposium on Wetland Hydrology, Chicago, IL (September 16-18). pp. 130-136.

28. SCS. 1975. Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys. Agriculture Handbook No. 436. Soil Conservation Service, U.S. Department of Agriculture.

29. USGS. 1988. Water resources data: Minnesota, water year 1988. Vol. 2, Upper Mississippi and Missouri River basins. U.S. Geological Survey Data MN-88-1.

Mapping Vulnerability of Soils to Nitrate Leaching at Different Scales, Using Different Models

Tamás Németh, László Pásztor, József Szabó, Zsófia Bakacsi
Research Institute for Soil Science and Agricultural Chemistry of the Hungarian Academy of Sciences

ABSTRACT

Various scale environmental information systems together with statistically supported GIS techniques were used for mapping the vulnerability of soils to a specific degradation process. In our work we present approaches for the evaluation of the vulnerability of soils for nitrate leaching. In two pilot areas – with different physiographical conditions – a methodological approach was initiated. A deterministic model family was introduced for the evaluation of the land vulnerability for nitrate hazard at a scale of about 1: 25,000. Compilation of vulnerability maps is a result of an iteration where data characteristics (availability, scale, informativity) and model parameters are in interaction. A stochastic model is also presented for the evaluation of the land vulnerability for nitrate leaching. This latter method was applied to mapping N-leaching hazard in Hungary at a scale of 1: 1M

INTRODUCTION

The increase of soil degradation at an alarming rate all over the world requires modeling and quantification of the regional extent and severity of these processes (FAO 1983, Várallyay 1991). The various scale soil information systems and the inherent techniques of GIS provide unique basis for studies of environmental degradation in modeling of changes in soil characteristics e.g. mapping the vulnerability of soils to degradation or pollution (Batjes and Bridges 1997, Pásztor et al. 1998). The produced maps may increase awareness on the potential nature, severity and extent of soil degradation at regional scales, and also permit identification of environmental hot spots that is sensitive, vulnerable or conflict areas (Pásztor et al. 1999).

In the agricultural practice application of nitrogen to enhance crop yields is necessary in most of the countries. The unproper use of the fertilizer-N (added to the soil N-pool) might play a significant role in nitrate contamination of subsurface waters, which are the main drinking water supplies (Addiscott et al. 1991). The agricultural practice in a certain area exists together with other activities, which also take part in the contamination processes, i.e. animal husbandry, leakage in urban and industrial canalization (Boumans et al. 1999). There are also some natural contamination sources like wet and dry deposit, nitrogen transformation within the soil, nitrogen originating from the subsoil (and rocks), etc.

The proportion of nitrogen present in mineral forms is greatly affected by climatic and soil conditions, by soil microbial population, and by land use. The ratio between the organic and inorganic forms of nitrogen in soil can be modified only slightly over the years. However, under certain environmental conditions (dry climate, negative water balance, overfertilization) a great part of the surplus nitrogen can accumulate in the soil profile, leaving the rooting zone of various crops, in the form of nitrate after the growing season (Németh, 1993a,b; Kovács et al. 1995), even when land is cropped annually. Integration of knowledge related to environmental conditions of a certain area with the soil, water, and crop management practices helps to prevent the simultaneity of the unfavorable processes leading to nitrate leaching, thus water resources may be protected from nitrate pollution of agricultural origin. It is of increasing importance that such an approach be applied in crop production.

It has been estimated that 80 percent of data have some geographic component using economic, environmental, political, social etc. information. A GIS is a tool that uses the power of the computer to pose and answer geographic questions by arranging and displaying data about places in a variety of ways, such as maps, charts, and tables (http://www.esri.com, 1999). People have used maps for thousands of years to clearly present information about places, and GIS is a modern extension of that ancient tradition (Longley et al., 1999). Geographical Information Systems are designed for storing, querying, analyzing as well as for displaying data and/or information with spatial characteristics. GIS represents both the techniques to realize spatially manageable storage of spatial data and extended opportunities of spatial analysis on the stored data (Maguire 1991, Scholten 1995). Recently, this latter is getting more and more attention.

GIS often lacks any means to deal with stochastic processes (Burrough 1999), ignoring the parallel development in spatial and/or multivariate statistics. In our work we are using stochastic models (Linhart and Zucchini, 1986) and apply various methods of multivariate statistical analysis (Lebart et al. 1984, Murtagh and Heck 1987). There are evidences in recent works (e.g. Pásztor and Csillag 1995, Pásztor et al. 1998, Tóth et al. 1998) that automatizable multivariate descriptive data processing methods can efficiently be applied to analyze and model various environmental processes, occasionally supplied with the tools of information theory.

MATERIALS AND METHODS

Hierarchical structure of modeling

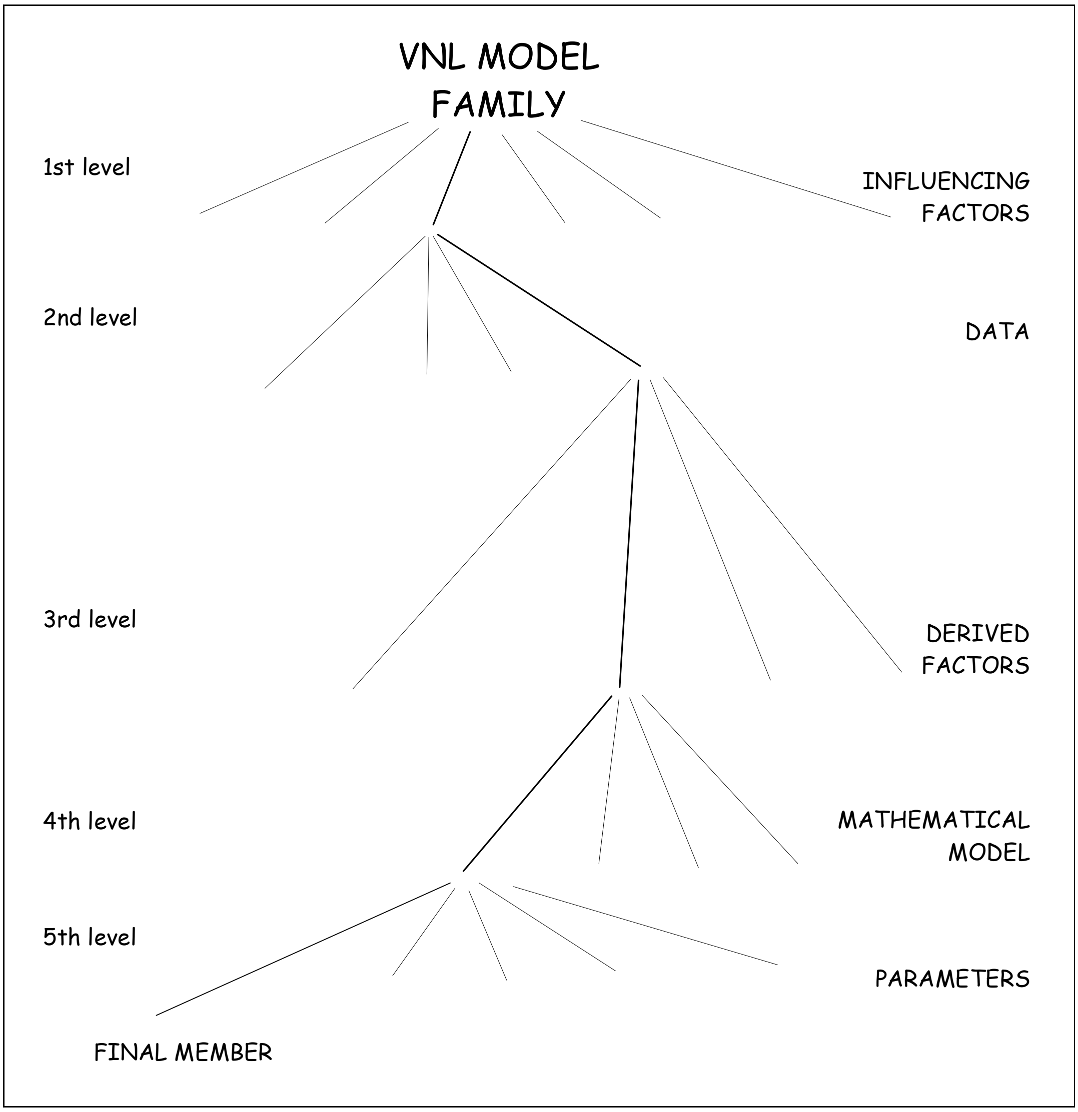


Figure 1. The sketch of hierarchical structure of modeling

GIS based mapping of soil vulnerability is composed of a hierarchical structure of models and the final map is a result of subsequent model selections (Fig. 1). On the first level the influencing factors are identified: 'Which environmental parameters determine the given vulnerability feature?' On the second level the available data on the formerly selected parameters are identified: 'What kind of datum is measured and/or available at all on the given influencing parameter?' On the third level some derived factors may be calculated: 'May the input data be converted/transformed into more appropriate format?' On the forth level the mathematical model is set up: 'What is the functional relation between the vulnerability feature and the determining factors?' On the further levels the refinement of the mathematical model takes place: ’Which value should be assigned to parameters to get good/better/the best result?’

The above-described procedure induces a model family. In this context compilation of the vulnerability map is the result of iteration where data characteristics and model parameters are in interaction.

Vulnerability of soils to N-leaching

Beside soil characteristics the effect of precipitation and the accessibility of groundwater determine the possibility of nitrate contamination of groundwater (Boumans et al., 1999, Németh et al., 1998). Physico-chemical characteristics of soils represent their buffer capacity/resistance feature according to the transfer of pollution. Precipitation surplus induces nitrate pollution to move downwards. Finally, location of groundwater table determines the distance to be done by nitrate to reach the water body.

Large-scale approach

The objective of the large-scale nitrate vulnerability mapping project was to establish an approach which is able to provide nitrate vulnerability maps at a scale of 1: 25,000 for different geographical conditions. Two pilot areas were selected for the implementation. The two regions differ in their physiographical conditions, land use, exposure to pollution and even the two sets of data available for them are distinct. The two pilot areas are well-defined physiographical units each. Csepel Isle is enclosed by River Danube, Watershed of Tetves Creek is a subcatchment of Lake Balaton. Area of the former is 248 km^2, that of the latter is 120 km^2. A simple comparison of the two pilot areas is given in Table 1.

Table 1. Some characteristics of the pilot areas

	Csepel Isle	Tetves Creek
physiography:	plain, island	hilly, catchment
dominant land use:	arable land	forest, pasture
exposure to pollution:	base of drinking water	Lake Balaton
soil information:	PemeTIR database	Kreybig GIS
groundwater information:	detailed, available	incomplete, derived
precipitation information:	poor	poor

The following data representation of the influencing parameters was available. For the description of precipitation annual average precipitation measurements were used. In first approximation it characterizes properly the degree of induction for leaching. For better results seasonal variability in precipitation as well as effects of non-natural water input (irrigation) should be considered. For the description of groundwater table measures on its average depth were used. To achieve more precise results, seasonal changes also in this parameter should be accounted for. For the representation of resistance of soils against the transfer of nitrate pollution their physico-chemical characteristics were used, namely texture and organic matter content together with depth of rootable depth of soils. For better results vertical variability in these parameters should taken into account by description of soils horizon by horizon (e.g. layer by layer).

Two factors were derived from the raw soil data. Organic matter content of soil is a density type parameter, consequently its product with rootable depth provides a quantity featuring column density of organic which is a much more proper parameter in our context. On the other hand, texture information on soil is generally (and as it was in the present case) given in categories. Thus according to their transferability, soils were described by numerical values on ordinal scale based on the knowledge of their texture properties.

The next step was the set up of the mathematical model. As a first approximation, general linear model was used as it is commonly used in multivariate methods.

$$V=\sum_{i=1}^{4} w_i * F_i \quad \text{, where} \qquad (1)$$

V is the measure of vulnerability, F_i is the ith factor and w_i is its corresponding weight.

Small-scale approach

In our small-scale approach a stochastic model is put forward for the evaluation of the land vulnerability for N-leaching. For the determination of nitrate leaching hazard in national scale three influencing factors proved to be relevant and available.

- Hungarian soils were classified into nine main soil water management categories according to their hydrophysical properties (Várallyay et al., 1980). The categories characterize infiltration rate, permeability, and hydraulic conductivity, field capacity and water retention features of Hungarian soils.

- Input map of ’Annual average precipitation’ is based on data collected by the National Meteorological Institute and registered by meteorological stations for the period of 1951-1980.

- Map of ’Groundwater depth’ is based on data collected in the frame of groundwater observation well network of Scientific Research Center for Water Resources and registered for the period of 1961-1980. The scale of this map is also a 1: 1,000,000 and covers non-mountainous region of the country.

Information on precipitation and soil has been available within AGROTOPO digital database compiled in Research Institute for Soil Science and Agricultural Chemistry of the Hungarian Academy of Sciences (RISSAC HAS) (Várallyay et al., 1985; Szabó and Pásztor, 1994). AGROTOPO is composed of territorial information on soils at a scale of 1: 100,000 and meteorological conditions at a scale of 1: 1,000,000 for the whole country.

The input data had to be harmonized because of the following reasons.

- Groundwater and precipitation information was originally collected in points, then interpolated, which resulted in isolines, however our procedure requires polygon features as input. Thus transforming contour information into polygon information restructured these maps.

- Scale of these two maps and that of water management categories map was highly dissimilar. Due to this fact information content of the latter must have been reduced, that is the map was generalized also to a scale of 1: 1,000,000.

- Finally, since the map of groundwater depth is constrained to non-mountainous regions, the other two maps must have been restricted to this extent, too.

Once having the accurate database as a set of topologically constructed, associated coverages, geographic analysis could be performed. To complete our objectives, the maps of various factors, as different layers, were overlaid (intersected). Feature attributes from all coverages were joined, that is a polygon of the resulted map is characterized by three attributes as opposed to the single attributes of the original maps.

The inherent errors of both analogue and digitized maps result in shifts of arcs. Thus even the correlated (e.g. common geographical) boundaries deviate from each other in the different maps. As a consequence, in the course of overlay of maps displaying correlated information sliver (that is small and/or elongated) polygons emerge. The number of this kind of polygons can be significant. Since they apriori distort the results, as well as cause overcomputation, they must have been eliminated. The criterion of elimination was set up based on the distribution of area of resulted polygons. Defining a threshold value under which polygons were eliminated cut the sharp peak of the distribution.

From mathematical/statistical point of view units of the resulted map are elements of a multidimensional factor space. Statistical behavior of the almost 500 units in this three-dimensional feature space was then studied. Since the number of units does not necessarily reflect their extent, in the computations their areas were used as weights.

Applying pure clustering techniques where there is no a priori rule to define the optimum number of groups, one needs some measure of the reality of the groups found by the partitioning algorithm. Finding the extreme value of an information theoretic criterion can provide the best fitting model and the best partition of the sample. Many model selection procedures may be found in the literature. Most of them take the form of a penalized likelihood, where a penalty term is added to the log-likelihood in order to compromise between the goodness-of-fit and the number of parameters. The ancestor of these models was developed by Akaike (1972) and we also turned to it, since it provides a versatile procedure for statistical model identification. The definition of Akaike's Information Criterion (AIC) is:

$$AIC=-2\ln(\text{maximum likelihood})+2(\#\text{ of parameters}). \qquad (2)$$

One of the most desirable properties of AIC is that (as it penalizes for large degrees of freedom) it tends to adopt simpler models and achieves a principle of parsimony. As AIC is basically an estimator of the risk of a model selection, it should be minimized to select among the alternative possibilities, that is the smaller is AIC the better is the classification of the points into groups. Its estimate can be computed by,

$$AIC(\text{estimated})=n\ln(R)+2p, \qquad (3)$$

where n is the number of objects to be grouped, p is the number of estimated parameters and R is the residual sum of squares of deviation from the fitted model (Akaike, 1974).

For the determination of optimal classification a sequence of non-hierarchical clustering was carried out and AIC(estimated) was calculated for each partition.

RESULTS

To get final result, w_l should be determined/defined in formula *(1)*. In the case of vulnerability term of definition is more appropriate. In the simplest case role of the four (original or derived) factors can be considered uniform. In this case weights merely help to standardize factors that is to make their scale comparable. The resulted vulnerability maps of the two pilot areas are displayed in Fig 2 and Fig3.

Spatial categorization (zoning) of land according to its sensitivity to a given pollution is generally required by decision-makers. This involves regionalization of the resulted mapping units, which are characterized by vulnerability values on nominal scale as a result of our large-scale approach. Our small-scale approach, solved this problem together with the 'pure' mapping procedure.

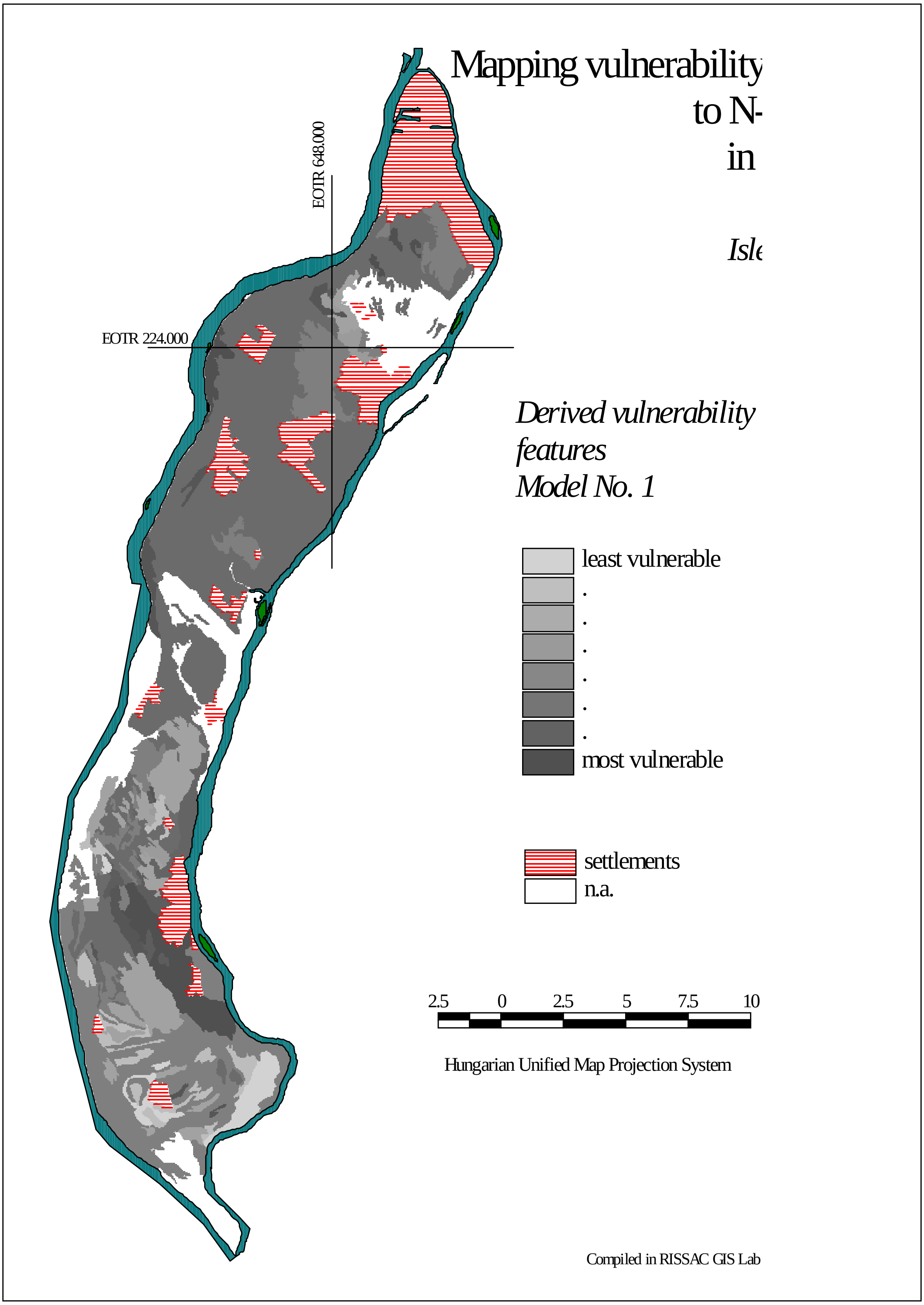


Figure 2. Result of large-scale approach on the Isle of Csepel

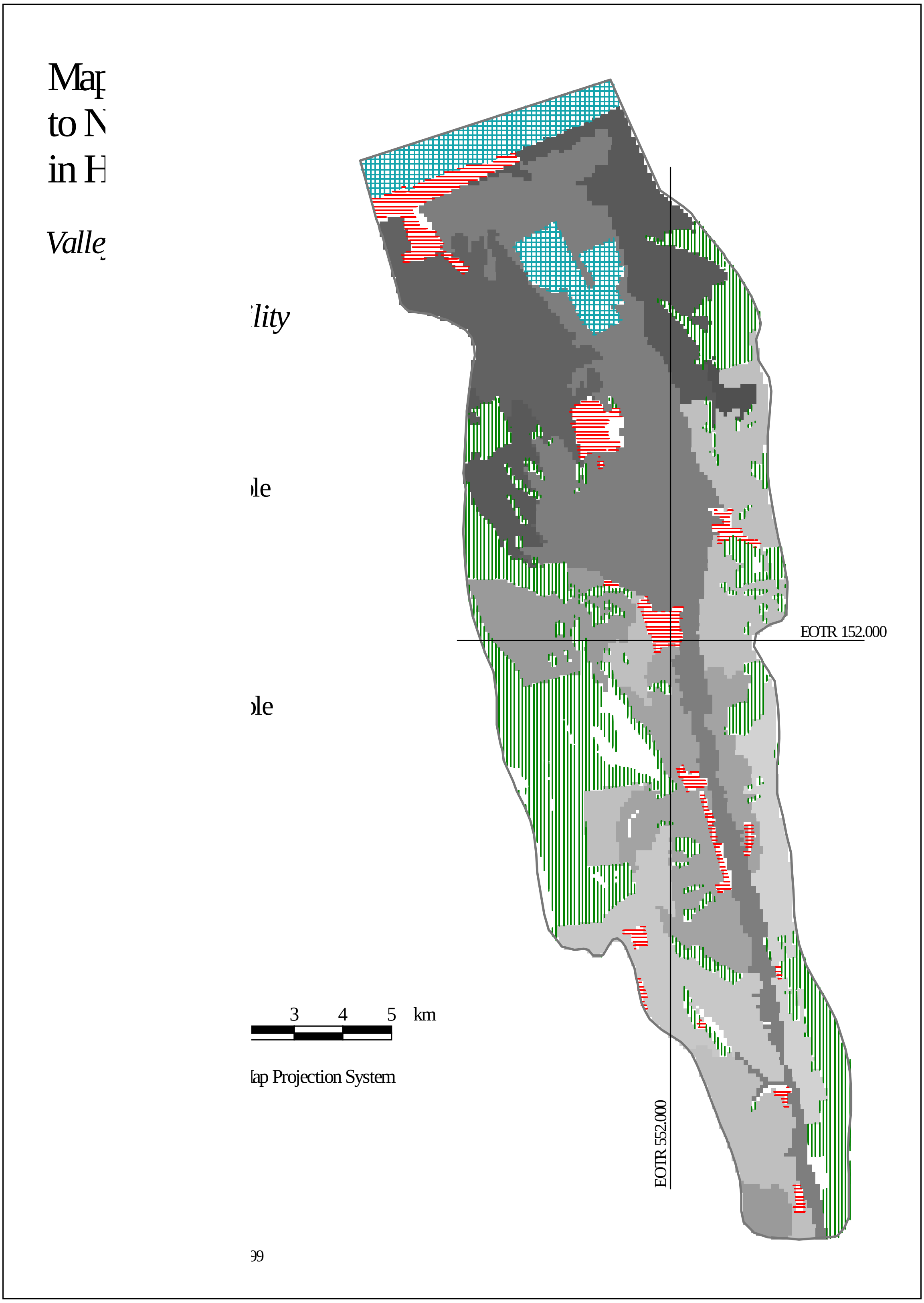


Figure 3. Result of large-scale approach in the watershed of Tetves Creek

Estimated AIC function for our dataset showed two local minima at 5 and 12 categories respectively. The 12-class solution in our scale provides too detailed thematic resolution, however a possible spatial zooming-in in the future may require also this thematic detailedness. Consequently, merely the 5-class solution was further studied. The identification of vulnerability categories was facilitated by displaying of the resulted categories on the units of the intersected map. The different categories showed well recognizable patterns. Analyzing their geographical distribution and extent, it was also possible to rank the resulted categories into a one-parameter sequence from severe hazard to the case of no hazard (Fig. 4).

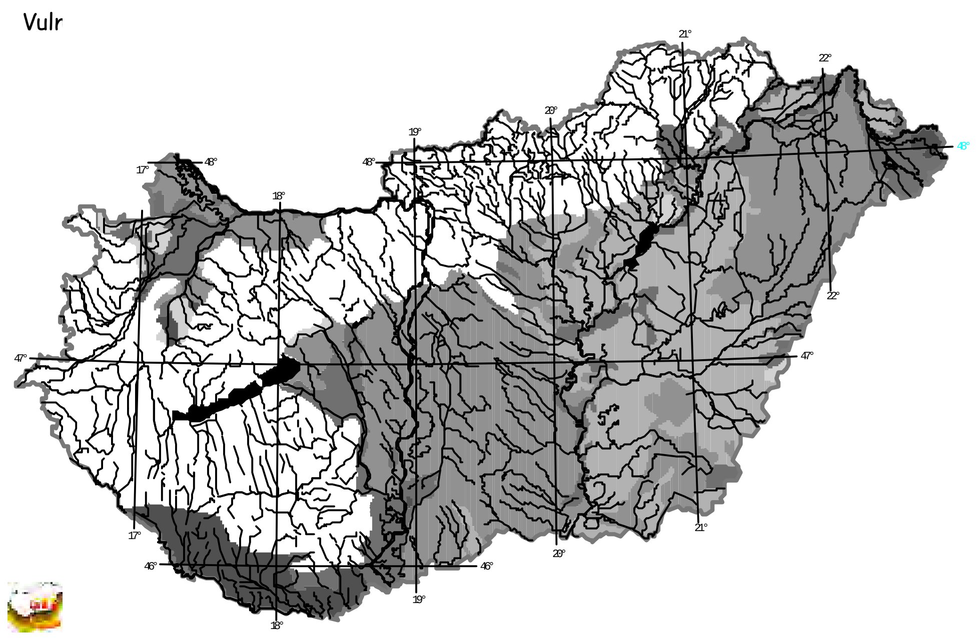


Figure 4. Result of small-scale approach

Severely susceptible soils are located on calcareous alluvial deposits where annual precipitation is relatively and groundwater level is actually high. They are characterized by shallow humus layer but relatively deep and coarse textured pedon, pH is mostly neutral, they have carbonate from the top, their texture class is mostly sand and loam. *Highly susceptible soils* are located on alluvial plains, where annual precipitation is relatively low (under 600 mm/year) and groundwater level is high. The pH, carbonate content and particle size distribution of these soils depend on the flooding material. *Moderately susceptible soils* are located on the coarse texture covered plains and chernozem areas. In spite of good drainage conditions these pour and humus sandy soils are moderately susceptibility because of low annual precipitation and deep groundwater level. *Slightly susceptible soils* are located on poor drainaged hydromorphic soils. Either flat areas are occupied by this category or they are located in depressions, where level of ground water is relatively high, soil texture tends to be heavier, amount of annual precipitation is the lowest in the country. *Relatively unsusceptible soils* are located on salt-affected landscapes; on heavy textured meadows of floodplains and peats; in poorly drained boggy and swampy depressions. In spite of different properties of these soils, they are commonly featured by poor drain conditions. 12-class solution of the procedure, by all means, might require more detailed, e.g. multi-level or multi-furcated explanation.

CONLUSIONS

The presented procedures provide application-sensitive approaches to the general problem of evaluating the vulnerability of various environmental elements for different degradation processes. The methods rely upon the up-to-date tools of GIS, multivariate methods and information theory.

Result of our large-scale approach represent a unique member of a model family where further family members may provide similarly good or even better results. However this member was reached in the course of subsequent model selections and each model selection means a compromise between optimum and executibility.

In our small-scale case study, the resulted map was easily interpretable providing straightforward characterization of vulnerability of Hungarian soils for leaching of nitrate accumulated after the growing season.

Our procedures are proposed for application in other fields with similar conditions.

ACKNOWLEDGEMENTS

The present work was partly founded by the Hungarian National Research Foundation (OTKA, Grant No-s T021275 and F026089).

REFERENCES

Addiscott T.M., Whitmore A.P., Powlson D.S., 1991, ‘Farming, Fertilizers and the Nitrate Problem’, CAB International.

Akaike H., 1972, ’Information theory and an extension of maximum likelihood principle’. In: Petrov B. N. and Csáki F. (eds.) Proc. 2nd International Symp. on Information Theory, , Akadémiai, Budapest, p: 267-281.

Akaike H., 1974, ’A new look at the statistical model identification’, IEEE Transactions on automatic control AC-19, p: 716-723.

Batjes N.H., E.M. Bridges (eds.), 1997, ‘Implementation of a Soil Degradation and Vulnerability Database for Central and Eastern Europe’, ISRIC, Wageningen.

Boumans L., G. van Drecht, D. Fraters, 1999, ’Nitrate in shallow groundwater of the sandy regions of the Netherlands’, In: Proc. ’Conference Spatial Statistics for Production Ecology’ Vol. 1. International Statistical Institute.

Burrough P.A., 1999, ’GIS and Geostatistics: essential partners in spatial analysis’, In: Proc. ’Conference Spatial Statistics for Production Ecology’ Vol. 1. International Statistical Institute.

FAO, 1983, ’Guidelines for the Control of Soil Degradation’, UNEP-FAO, Rome.

Kovács G. J., J.T. Ritchie, T. Németh, 1995, ’Testing simulation models for assessment of crop production and nitrate leaching in Hungary’, Agricultural Systems, 49(4), p: 385-397.

Lebart L., A. Morineau, K.M. Warwick , 1984, Multivariate descriptive statistical analysis, Wiley, New York.

Linhart H., W. Zucchini, 1986, ’Model selection’, Wiley, New York.

Longley P.A., M.F. Goodchild, D.J. Maguire, D.W. Rhind (eds.), 1999, ’Geographical Information Systems, Principles and Technical Issues’, Wiley.

Maguire D.J., 1991, ’Geographical Information Systems: principles and application’, Longman, London.

Murtagh F., A. Heck, 1987, ’Multivariate data analysis’, Reidel, Dordrecht.

Németh T., 1993a, ’Effect of N fertilization on the nitrate-N content of soil profiles in long-term experiment’. Agrokémia és Talajtan, 42, p: 115-120.

Németh T., 1993b, ’Fertilizer recommendations - Environmental aspects.’ Zeszyty Prob. Post. Nauk Roln, 400, p: 95-104.

Németh T., L. Pásztor, J. Szabó, 1998, ‘Stochastic modeling of N-leaching using GIS and multivariate statistical methods’, Water Science and Technology, 38, No. 10, p: 191-197.

Pásztor L., F. Csillag, 1995, 'Reduction of high resolution spectra; Application to characterization of salinity status of soils.' In: Sensors and Environmental Applications of Remote Sensing; Proc. of the 14th EARSeL Symposium, Balkema, Rotterdam, pp. 393-397.

Pásztor L., Zs. Suba, J. Szabó, Gy. Várallyay, 1998, 'Land degradation mapping in Hungary', In: J.F.Dallemand, V. Perdigao (eds.) 'EUR 18050 – PHARE Multi-Country Environment Programme MERA Project Proceedings', European Commission, p: 43-54.

Pásztor L., J. Szabó, T. Németh, 1998, 'GIS-based stochastic approach for mapping soil vulnerability', Agrokémia és Talajtan, Vol. 47, No. 1-4, p: 87-96.

Pásztor L., J. Szabó, Zs. Bakacsi, S.T.D. Turner, T. Tullner, 1999, 'Applicability of GIS tools in environmental conflict mapping: A case study in Hungary'. (in this volume).

Scholten H., 1995, 'Geographical Information Systems, Spatial Data Analysis and Spatial Modeling, A Perspective in European Context', In: Proc. of 10th European ARC/INFO User Conference, Prague, p: 2-14.

Szabó J., L. Pásztor, 1994, 'Magyarország agroökológiai adatbázisa és annak környezetvédelmi felhasználási lehet ségei' (in Hungarian), In: 'Országos Környezetvédelmi Konferencia', Siófok, pp. 156-163.

Tóth T., M. Kertész, L. Pásztor, 1998, 'New approaches in salinity/sodicity mapping in Hungary', Agrokémia és Talajtan, Vol. 47, No. 1-4, p: 76-86.

Várallyay Gy, L. Sz cs, K. Rajkai, P. Zilahy, A. Murányi, 1980, 'Soil water management categories of Hungarian soils and the map of soil water properties' (in Hungarian - English summary) Agrokémia és Talajtan, 29, p: 77-108.

Várallyay Gy., L. Szücs, A. Murányi, K. Rajkai, P. Zilahy, 1985, 'Soil factors determining the agro-ecological potential of Hungary', Agrokémia és Talajtan, 34(Suppl), p: 90-94.

Várallyay Gy., 1991. 'Soil vulnerability mapping in Hungary'. In: Proc. Int. Workshop on 'Mapping of soil and terrain vulnerability to specified chemical compounds in Europe at a scale of 1: 5 M', p: 83-89.

Habitat Filters, GIS, and Riverine Fish Assemblages: Sifting Through the Relationships

Douglas A. Nieman[1], Gregg S. Sermarini, and William S. Ettinger
Respectively: Senior Scientist, Information Systems Manager, and Principal Ecologist
Normandeau Associates, Inc., 3450 Schuylkill Road, Spring City, Pennsylvania 19520

Thomas Proch
Regional Biologist, Pennsylvania Department of Environmental Protection
400 Waterfront Drive, Pittsburgh, Pennsylvania 15222-4745

John Arway
Chief, Environmental Services Division, Pennsylvania Fish and Boat Commission
450 Robinson Lane, Bellefonte, Pennsylvania 16823-9616

Jerry Schulte
Senior Biologist, Ohio River Valley Water Sanitation Commission
5735 Kellogg Avenue, Cincinnati, Ohio 45228-1112

ABSTRACT

Hierarchy theory suggests that natural systems are organized as an endogenous series of levels that reflect differences in process rates acting over a range of spatiotemporal scales. Several hierarchical classifications of spatial scales have been proposed for lotic systems (e.g., basin, network, reach, channel unit, patch). We used a classification similar to one by Lubinski for navigation rivers (basin, network, reach, navigation pool, aquatic area, patch) to develop a hierarchical filter model of habitat quality and assemblage composition in mainstem navigation pools of the upper Ohio River basin. A macroscale filter first considers the geographic distribution of species in the regional assemblage across 10 pools in the study area. This filter reflects the influence of historical events and response of individual species to long-range environmental gradients at the basin, network, and reach scales. Aquatic areas defined by channel geometry comprise a mesoscale filter. Aquatic areas divide a navigation pool into lateral and longitudinal components defining major within-pool gradients of depth, velocity, substrate, and cover. Literature-based information on habitat requirements was used, via canonical correspondence analysis, to group species and life-stages into ten “species-habitat associations” that segregate species by anticipated habitat use along major lateral (margin to midchannel, or primary to secondary/tertiary channel) and longitudinal (tailwater-riverine to lower pool-lacustrine) habitat gradients within navigation pools. After accounting for macroscale and mesoscale influences, filter models can be further refined by considering microscale habitat preferences in relation to the distribution of cover and substrate patches mapped at finer spatial

[1] Corresponding author, Dnieman@normandeau.com

scales. Theoretical underpinnings of the filter concept are discussed with respect to habitat features that appear to shape fish assemblage patterns across multiple spatial scales.

INTRODUCTION

Successful natural resources management requires knowledge of the past, present, and potential future status of ecosystem components that sustain or comprise management units. The elements of required knowledge can be divided broadly into realms of status (inventory) and process (mechanisms). Status recognition involves the questions “what” and “how much”, which are addressed with techniques of classification, measurement, and statistical estimation. Process estimation asks “how” and involves mechanistic understanding, which confers predictive ability. The questions “where” and “when” bridge these realms by challenging us to account for the mechanisms responsible for spatiotemporal variation. The advent of geographic information systems (GIS) has provided tools to effectively manage large datasets, enabling construction of spatially explicit ecological inventories. With GIS, spatiotemporal variation can be measured and analyzed within a multi-scale context, facilitating construction of ecological models that incorporate the kinds of spatial complexity observed in nature. However, effective use of such tools can be hindered by lack of a conceptual framework for guiding applications.

Hierarchy theory (Allen and Starr 1982, O’Niell 1989, Levin 1995) provides a unifying conceptual approach for understanding patterns of natural variation. A main tenet of this theory is that complex systems have an endogenous organization that is structured in part by differences in process rates acting over a range of spatiotemporal scales. For a given scale, activities or patterns at finer, briefer scales appear as unpredictable “noise”, while those at the coarser, slower scales of higher levels appear as near-constant features of the environment that constrain activities or the expression of variability at the chosen level of perception.

There are consequences arising from this conceptual framework that can provide clues to mechanisms that generate ecological pattern. For example, higher-level features may constrain expression of lower-level properties, causing variation in the lower level to appear non-random and spatially organized (Legendre and Fortin 1989, Borcard et al. 1992). Repetition of disturbance at lower levels through space and time eventually alters higher-level properties (Lubinski 1993), leading to synergism, positive feedback, emergence, and other forms of surprise (Casti 1994). Ecologists have learned that large-scale processes drive many patterns at fine scales (Rabeni and Sowa 1996, Richards et al. 1996), and that the problem is not to select a correct scale for description, but to recognize that change occurs at many scales all at once; interactions among scales should therefore be afforded great scrutiny (Levin 1995).

A common goal in applied ecology is to predict status of biological populations from knowledge of physical habitat. Numerous methods, based on various conceptual approaches, exist for making such predictions. A model well-suited for GIS application is the “hierarchical filter”, which contends that in order to “belong” to a local assemblage, a species must possess traits that allow it to “pass” through a nested series of ecological filters operating at successively finer spatiotemporal scales. Furthermore, if a given patch of habitat is to be suitable, the filter model dictates that higher-level features of the environment, within which the patch is embedded, must be suitable as well. Scales may range through several orders of spatial and temporal magnitude, from broad-evolutionary (global, continental) through intermediate-ecological (ecoregion, basin), to fine-stochastic (mosaic, patch). This model has appeared several times in the ecological literature, especially in aquatic contexts (Smith and Powell 1971, Poff and Ward 1990, Tonn 1991, Imhoff et al.1996, Poff 1997).

Our objectives in this paper are first, to briefly review the literature on hierarchical filters, second, to identify potential habitat filters that may shape fish assemblage patterns in the Ohio River and other large navigation rivers; and third, to summarize GIS-based models of habitat quality and fish assemblage composition we developed for navigation pools of the upper Ohio River basin in Pennsylvania. We conclude with some thoughts on potential uses of this model for fisheries research and management. Our goal is to foster a more holistic understanding of fish-habitat relationships in navigation rivers than is provided by approaches that focus on only a single level of ecological description, e.g., on local habitat features.

REVIEW OF HIERARCHICAL FILTER MODELS

Smith and Powell (1971) offer an early description of hierarchical filters in their discussion of “screens” that helped shape fish assemblages in Brier Creek, Oklahoma, a small southwestern warmwater stream. In his recent book on freshwater fish ecology, Matthews (1998) draws heavily on hierarchical perspectives as a framework for organizing factors that shape patterns of fish distribution and abundance. Tonn (1991) filtered the world’s fish fauna through four levels to show how global warming might influence local assemblages via different pathways. Imhoff et al. (1996) adopted a hierarchical system linked to specific mechanisms regulating salmonid abundance. Poff (1997) explored links between macroinvertebrate traits and habitat filters within levels of a nested landscape hierarchy. Other researchers concerned with description of physical habitat have proposed hierarchical systems that reflect “endogenous organization” (Frissell et al. 1986, Lubinski 1993, Wilcox 1993). These examples share the common attribute of hierarchical structure, and are similar in terms of the number and characteristics of levels (Table 1).

An important feature of natural hierarchies is a positive association between spatial extent and temporal constancy. Larger spatial scales are generally associated with processes that occur slowly over long periods of time, while local processes linked to finer spatial scales operate at higher frequency (Poff and Ward 1990). Often, the major processes that drive ecosystem form and function are grouped as distinct levels in a nested hierarchy (Kolasa 1989).

Although hierarchical classifications can organize patterns of variation in physical habitat characteristics, understanding their role as “filters” depends on identification of the environmental characteristics that selectively cull potential assemblage members at successively finer scales. Thus, it is necessary to understand how species’ traits confer ability to resist selective environmental constraints that would act at some level to remove a species from a local assemblage (Poff and Ward 1990, Poff 1997). Such traits encompass morphological, physiological, behavioral, and other life-history characteristics of organisms (Imhoff et al. 1996, Poff 1997, Schlosser 1990, 1991, Winemiller and Rose 1992). In Hutchinsonian terms, the filter model contends that a species’ niche requirements will only be met where “windows of opportunity” overlap across the range of spatiotemporal scales that ultimately influence local assemblages.

Table 1. Examples of hierarchical filter models[a] and habitat classifications[b] developed for aquatic systems.

Type of system (Authors)	Levels in hierarchy	Description of filtering effects or levels (simplified; see original papers for details)
Fishes living in freshwater lakes and streams[a] (Smith and Powell 1971)	Gross physiology	-Filters freshwater fishes from world fauna
	Gross geography	-Determines continental taxonomic composition
	Fine geography	-Determines regional species pool
	Climate	-Affects local distribution patterns
	Fine physiology	-Local suitability affects ultimate composition of assemblages
Fishes living in north temperate lakes[a] (Tonn 1990)	Continental filter	-Determines continental fauna
	Regional filter	-Determines regional species pool
	Lake-type filter	-Separates regional fauna into lake-type species pools
	Local filter	-Local fish assemblages selected from lake-type pools
Benthic macro-invertebrates in streams across a landscape[a] (Poff 1997)	Watershed/basin filter	-Effects of history, climate, geology, and disturbance
	Valley/reach filter	-Effects of stream confinement, slope, lithology, channel morphology
	Channel unit filter	-Effects of stream morphology, hydraulics, substrate distribution
	Microhabitat filter	-Influence of depth, velocity, particle size, food availability, biotic interactions, etc., on habitat selection
Large navigation rivers[b] (Lubinski 1993)	Basin	-Fundamental landscape units defined by watershed boundaries
	Stream network	-Flowing channels within basins; linkage between aquatic and terrestrial ecosystem components
	Floodplain reach	-Length of river defined by degree of interaction between river and valley/floodplain in conjunction with disturbance regimes
	Navigation pool	-Length of river between adjacent dams; artificial element of the landscape that fundamentally alters finer scale habitat features
	Habitat	-Fine-grained structural components most intimately associated with biological assemblages
Controls on stream habitat[b] (Frissel et al. 1986)	Stream system	-Uplift and subsidence, volcanism, and climate change variables that influence planation, denudation, and drainage development
	Segment system	-Minor glaciation, volcanism, earthquakes, landslides, and valley development that control evolution of drainage pattern
	Reach system	-Debris torrents, landslides, meandering, human alterations that influence stream bed elevations through modification of sediment storage, bank erosion, and riparian vegetation.
	Pool/riffle system	-Delivery/transport dynamics of wood and sediment affected by local bank failure, flood scour and deposition, and human activities
	Microhabitat system	-Sediment load, organic matter transport, localized scour, and seasonal plant growth affecting depth, velocity, accumulation of fines, community metabolism, structural habitat heterogeneity
Scales for watershed systems analysis[b] (Imhoff et al. 1996)	Watershed	-Drainage divides between tertiary basins
	Subwatershed	-Boundaries of stream basins nested within watershed
	Reach	-Minimally, two meander wavelengths within a stream segment of a specific type (i.e., Rosgen 1994), bounded laterally by 1:20-year (active profile) and 1:100-year (passive profile) flood elevations
	Site	-Channel segment bounding a single riffle or pool lengthwise, bounded laterally by bankfull elevation
	Habitat element	-Area of relative homogeneity in depth, velocity, and substrate within a site

HABITAT FILTERS FOR NAVIGATION RIVERS

We define navigation rivers generally as those that support commercial barge traffic. We restrict our discussion (but not our definition) to systems where navigability is maintained by “low-head” locks and dams, excluding systems where navigability is maintained by channel training structures (e.g., lower Mississippi River, Baker et al. 1991) or by large dams with massive reservoirs (e.g., many Tennessee River impoundments). We invoke a conceptual model of the upper Mississippi River system (UMRS) by Lubinski (1993) as a framework for developing a hierarchical approach to organizing patterns of habitat and assemblage variation in the Ohio, Allegheny, and Monongahela River mainstems, where navigability is maintained by a chain of 38 locks and dams. These rivers share many similarities but exhibit important differences compared to the UMRS, which we will note during the following discussion.

The model identifies five important spatial scales: basin, stream network, floodplain-reach, navigation pool, and habitat. Table 2 lists important features of this model which are useful in developing a “filtering framework” for estimating habitat suitability and fish assemblage composition from knowledge of environmental conditions. In this section, we discuss each scale within the context of the Ohio River, and identify possible filtering mechanisms, hypothesized here or by others, and how they might influence fish assemblages across the range of identified spatiotemporal scales. In examining potential habitat filters that shape fish assemblages, it is important to remember that assemblages are composed of individual organisms that respond to environmental selective forces operating within the context of local habitats (Poff 1997). Higher level filters operate by constraining local conditions, making certain combinations of local factors possible while ruling out others. Coarse filters may also preclude a species from otherwise acceptable localities because they reflect the influence of history, or because they account for long-range effects of local disturbance, changes in patterns of spatial connectedness, and other phenomena that are determined by the way in which local habitats are distributed across a landscape.

Table 2. Spatial scales of the Ohio River, with major structural, functional, and scale characteristics at each level. Adapted from Lubinski (199

Spatial Scale:	Basin	Stream network	Floodplain-reach	Navigation pool	Habitat
Components:	-Basins and sub-basins	-Headwaters, tributaries, mainstems	-Reaches with valley, floodplain, channels, cutoffs	-Locks and dams and associated navigation pools	-Aquatic areas and habitat conditions
Major features:	-Fundamental units of landscape division; -Provide structural and functional constraints on lower levels of organization	-Interactions between aquatic and terrestrial ecosystem components; -Linear processes emphasized, modified by network shape and position	-Interactions between channel and valley-floodplain	-Artificial component of the landscape -Modifies structural components and functional processes at habitat scale	-Scale most intimately with the abundance and distribution of biota
Abiotic factors:	-Size and location -Climate and geology	-Drainage pattern -Hydrology -Channel morphology -Water quality	-Annual floods -Valley gradient and floodplain morphology -Thermal regime -Sediments and nutrients	-Size and location -Channel geometry -Lateral, longitudinal, and vertical gradients	-Velocity and turbulence -Substrate and cover -Depth and clarity -Water chemistry -Diel variation
Biotic factors	-Vegetation -Dispersal	-Vegetation -Dispersal and migration -POM dynamics	-Vegetation production and diversity	-Vegetative cover	-Vegetation (riparian, submergent, emergent) -Plankton -"Engineer" species -Bioturbators
Natural disturbance	-Glaciation -Flood and drought -Catastrophe (volcanism, wildfire)	$\geq$ Bankfull flow -Intermittency -Stream capture	-Absence of flood pulse -Meandering (local distrubance providing long-range equilibrium)	-Flood and drought	-Windthrow, debris -Ice -Wind-generated waves
Human disturbance	-Agriculture -Urbanization -Industrialization -Resource extraction	-Channelization -Dams and diversions -Contaminants -Exotic species	-Levees and revettments -Harvests	-Water level management	-Snag removal -Dredging and disposal -River traffic, fleeting -Berming
Approximate range in scale					
-space	10^{3}-10^{7} km^{2}	10^{-1}-10^{3} km^{2}	10^{1}-10^{3} km^{2}	10^{1}-10^{3} km^{2}	10^{-3}-10^{4} m^{2}
-time	10^{1}-10^{4} yr	10^{0}-10^{3} yr	10^{-1} - 10^{3} yr	10^{-1} - 10^{1} yr	10^{-2} - 10^{0} yr

Basin Scale

Basins are fundamental landscape units delimited by watershed boundaries (Petts 1989, Lubinski 1993). Basins can be decomposed hierarchically as nested subbasins down to the level of individual watersheds (Frissell et al. 1986). Characteristics of geology, landform, climate, and vegetation are basin attributes that shape and constrain features at finer levels. Lubinski (1993) defines such characteristics as “major factors” that control the range of ecological heterogeneity at finer scales. These attributes are relatively static when viewed from lower levels in a habitat hierarchy, but do change over long periods of time. Changes in such factors are usually associated with high-magnitude, low-frequency events such as glacial advance/retreat, earthquakes, and volcanism that qualify as “disturbance”. Disturbance at coarse scales fundamentally alters finer scale habitat, doing so over relatively brief periods of evolutionary time. Frissell et al. (1986) note that such disturbance is “extrinsic” in that it changes the system’s potential by creating a “new” system with a different range of possible future states from the destruction of the “old” system that had a different range of possible future states. Between disturbances that either create or destroy them, systems are modified by “intrinsic” developmental processes.

Application of the habitat filter concept at the basin scale should incorporate historical legacies that determined biogeographic patterns of species’ distributions and thus the pool of species potentially available to local assemblages at finer spatial scales (i.e., within and between sub-basins). Much of the modern broad-scale distribution patterns in the Ohio River basin were shaped by events associated with glaciation and glacial retreat (Trautman 1981, Strange 1999). Efforts to relate fish assemblages to local habitat variation should therefore be placed in a context that recognizes regional history. Trautman (1981, pages 1-12) provides such a regional analysis by summarizing how fish assemblages of present-day Ohio were shaped in part by broad-scale, regional variation in climate, topography, and the effects of glaciation.

The importance of broad-scale phenomena can also be seen in the swift alterations to native fish fauna in the Ohio River that followed European settlement, largely resulting from the conversion of natural landscapes into agricultural, urban, and industrial systems (Trautman 1981, Pearson and Krumholz 1984). Such changes reflect the cumulative nature of human disturbance, which, when repeated often at many places, begins to impact larger scales (Lubinski 1993). Even small, seemingly isolated disturbances, can “scale-up” by inducing positive feedback. Construction of a dam, or destabilization of a local stream bank, can alter stream channel geometry, hydraulics, and sediment transport over distances that far exceed the boundaries of the local disturbance (Leopold 1994). Such magnification of individual disturbances serves to decrease the number of such events required to modify basin-level controls on aquatic ecosystems.

Stream Network Scale

As described by Lubinski (1993), water-carrying channels above a defined location within a basin comprise the stream network scale, which is similar to the stream system scale of Frissell et al. (1986). Stream networks resolve linkages between the aquatic and terrestrial components of a watershed ecosystem. This scale focuses on interactions between streams and riparian zones (Gregory et al. 1991, Schlosser 1991), longitudinal processes (Vannote et al. 1980), and the influence of tributary spatial position on biotic assemblages (Gorman 1986, Osborne and Wiley 1992, Fairchild et al. 1998). Natural disturbances at this scale include large floods, drought, and events that alter drainage patterns (Lubinski 1993, Strange 1999). Human disturbances include dams, altered hydrologic regimes, and fragmentation and isolation of the network by numerous mechanisms (Sheldon 1987, Lubinski 1993, Pringle 1997).

Since our focus is on the larger mainstems of the drainage network, it is important to identify how certain filters might operate at the stream network scale to influence “main river” fish assemblages. As accumulators of flow from large areas, large rivers reflect their position in a drainage network in terms of channel morphology, energy, sediment, water quality, flow regime, and instream habitat (Vannote et al. 1980, Ryder and Pesendorfer 1989, Sedell et al. 1989, Ward and Stanford 1989, Junk et al. 1989). Thus, the filtering influences on a regional species pool at the stream network scale should reflect the segregation of species according to preferred ranges of stream size, coupled with other habitat differences related to position within the drainage network (Osborne and Wiley 1992). However, such influences are not simply those that reflect differences in preferred habitat use related to stream size, which include inherently local factors, i.e., local filters will remove species incapable of maintaining populations in a local setting. Rather, network filters reflect geomorphologic and fluvial controls on local habitat and patterns of connectivity that helped determine how different fish species came to be where they are today.

Understanding the filtering role of habitat variation at the stream network scale on mainstem assemblages should start by considering pristine mainstems. Accounts of habitat conditions in the pre-settlement Ohio River (Trautman 1981, Pearson and Krumholz 1984) paint a picture of the Ohio River that is much different than the one we see today. The clearer water, lower sediment load, shallower and faster riffles and rapids, and extensive floodplain marshes supported a fish fauna that was markedly different than the present one. The pristine Ohio River and major tributaries likely served as habitat for a greater array of obligate riverine species, some of which specialized in “big river” habitat. Larger species such as paddlefish (*Polyodon spathula*), shovelnose sturgeon (*Scaphirhynchus platorynchus*), gars (*Lepisosteus* spp.), blue sucker (*Cycleptus elongatus*), flathead catfish (*Pylodictis olivaris*), blue catfish (*Ictalurus furcatus*), and some smaller species (e.g., mimic shiner, *Notropis volucellus*; emerald shiner, *N. atherinoides*) probably achieved maximum population densities in the main river and lower reaches of major tributaries. Some species (e.g., many catostomids) maintained large adult populations in mainstems, but relied more or less on tributaries for nursery habitat. Many other species ranged throughout much of the drainage network, with populations found in both mainstems and tributaries. Still others generally occurred in mainstems as strays from populations centered in tributaries. For such species (e.g., many *Etheostoma* and *Percina*), mainstems still served as important migration corridors that maintained connections between separate local populations. Mainstems probably served as refugia for small stream species during episodic disturbance (e.g., drought), conversely, tributaries could shelter fish from poor main river conditions (e.g., floods or low dissolved oxygen, Gammon and Reidy 1981).

Mainstems were altered progressively after European settlement, beginning with large-scale changes in physicochemical conditions linked to deforestation and conversion of the landscape to agriculture (Trautman 1981). Somewhat later, industrial development marked the beginning of an era of severe pollution that abated only recently. Such disturbances directly impacted fishes using mainstems, leading to extirpations throughout many kilometers of river (Trautman 1981, Cooper 1983). Alterations to both tributaries and mainstems undoubtedly changed metapopulation structures by interfering with the normal patterns of migration that maintained gene flow and allowed local populations to recover from natural disturbance. Before the period of severe pollution ended, construction of the navigation system fundamentally altered mainstem habitat, likely influencing patterns of connectivity even further. As a repetitive feature, the chain of navigation dams represents a “scaling up” of the local effects associated with any one dam, possibly influencing a spatial scale higher up in the ecosystem hierarchy by altering patterns of fish movement within the drainage network.

Filtering influences at the stream network scale may interact with heterogeneity at lower levels in the Lubinski (1993) hierarchy. For example, Emery et al. (1999) document changes in the species composition of catostomids within and across navigation pools in the Ohio River. They cited possible causes for shifts in abundance between round and deep-bodied suckers that can be related to heterogeneity across four scales. Round-bodied suckers depend more on lotic conditions (e.g., fast currents, coarse substrates), while deep bodied suckers thrive in lentic conditions. Thus, local assemblages at fine scales differ in sucker representation depending on the relative mix of these habitat types. In navigation pools, lotic and lentic habitats are divided between the upper and lower ends of each pool, respectively (Wilcox 1993). Chaining dams together results in an alternating sequence consisting of a lotic tailwater below an upstream dam that grades into a lentic environment above the next dam downstream. This pattern is then modified over a series of dams by heterogeneity at the reach scale. The lowering of slope downstream along the river results in longer pools, decreased sediment particle size, changes in trophic dynamics, and increased distance between mainstem habitat favorable to round bodied suckers. Off-channel habitat favorable to reproduction of deep bodied suckers is more developed in lower gradient reaches where channel cutoffs, embayments, and flooded tributary mouths are more extensive. An influence at the stream network scale is apparent because of a decrease in the number of tributaries along the lower Ohio River compared to upstream (Emery et al. 1999). As favorable areas for reproduction in the mainstem dwindle, maintenance of round bodied sucker populations would depend on availability of tributary spawning habitat. Fewer tributaries along a reach would therefore reduce round bodied sucker populations. In summary, the shift in sucker composition appears to involve multiscale interactions among several factors (*sensu* Levin 1995, Poff 1997).

Floodplain-reach Scale

This scale resolves floodplain reaches as structurally distinct river segments in which finer-scaled ecological characteristics are controlled in part by the degree of interaction between the river and its valley and floodplain (Lubinski 1993, Sedell et al. 1989, Ward and Stanford 1989). Reaches can be distinguished based on geology, physiography, and large increases in drainage area at major tributary junctions (Ward and Stanford 1989). Morphological aspects that help distinguish reaches include valley width, slope, and floodplain development. Unlike mainstems of the UMRS, those of the upper Ohio River basin are valley constrained and lack broad, dynamic floodplains. As a result, upper basin mainstems have little “off-channel” habitat (*sensu* Wilcox 1993), and flooded tributaries are rarely backwatered by more than a kilometer. Valley slope decreases and width increases in the middle and lower Ohio River, where there is a general increase in off-channel habitat. None of the Ohio River approaches the complexity of off-channel habitat associated with floodplains of the UMRS, however.

To our knowledge, formal reaches for the Ohio River system have not been defined, although the “Three Rivers” union defines a clear reach boundary because of the difference in character between the confluent rivers that create the Ohio River at Pittsburgh, PA. Various researchers have referred to upper, middle, and lower thirds of the Ohio River as distinctive (Pearson and Krumholz 1984, Emery et al. 1999), with the “Falls of the Ohio” at Louisville, KY appearing to be a logical site for another reach boundary.

Filtering influences at the floodplain-reach scale on potential fish assemblages likely reflect differences in the amount and character of off-channel habitat, possibly by filtering species that require off-channel nursery sites from areas where those sites are lacking, or have been disconnected by floodplain modification (Lubinski 1993, Gutreuter 1992, 1993). As the river flows along major climatologic gradients, species may come and go at reach scales in relation to thermal tolerance (Gutreuter 1992). Within and between reaches, assemblages fluctuate in composition as a result of annual variation in flow regime, often associated with the presence or absence of seasonal flood pulses (Junk et al. 1989, Lubinski 1993).

Navigation Pool Scale

The navigation pool scale is unique among spatial scales in representing an artificial element and major modification of the riverine ecosystem (Lubinski 1993). Serial repetition of these structures scales up to influence ecological processes at the broader reach and stream network scales. Dams also constrain habitat variation at finer scales through both passive and active mechanisms. Passively, dams set up longitudinal gradients in fine-scaled habitat features by imposing a step-pool structure along the river continuum (Wilcox 1993). Tailwaters below dams retain some of their original riverine character as sites of greater velocity, turbulence, and substrate particle size, which then grade into slower, deeper impounded areas above dams where finer sediments accumulate. Actively, short-term variation in fine-scale habitat is tied to water-level management, which in the Ohio basin attempts to maintain a constant pool elevation at each dam. As a result, depth varies more over time in upstream areas of the pool, thus influencing inundation frequency and duration of aquatic-terrestrial transition zones along the river margin (Junk et al. 1989). Less variation in depth above the downstream dam limits change in cross-sectional area, resulting in higher average velocity when passing elevated discharge. These factors interact along the spatial axes of a pool to influence temporal variability in fine-scale habitat, which fundamentally affects selective forces acting on biota (Poff and Ward 1990). Similar interactions occur in the UMRS, although spatiotemporal dynamics differ there because water levels are manipulated to approach stability at a mid-pool "hinge point" (Lubinski 1993).

Habitat Scale

Within navigation pools, fine-grained features of the environment are functionally linked to the success of individual fish at the habitat scale. Because each species perceives and responds to environmental heterogeneity uniquely, no single description of fine-grained habitat will suit all cases. Just how species perceive and respond to local heterogeneity, and how such responses help shape fish assemblage pattern, is an area of active inquiry (Poff and Ward 1990, Taylor et al. 1993). A hierarchical perspective on how "patches" of fine-scaled habitat are spatially arranged leads to the general prediction that habitat "specialists" should outnumber "generalist" species, but generalists will on average be more abundant and more broadly distributed in ecological range and habitat use (Kolasa 1989).

Descriptions of habitat use by fish typically refer to hydraulic variables (e.g., depth and velocity), substratum (e.g., particle size, embeddedness, organic matter content), cover (type and amount), and physicochemical limits of tolerance. A common theme in lotic ecology is that habitat conditions are arranged along a continuum at broad scales, while heterogeneity and habitat use are patchy along fine-scaled gradients controlled in part by stream channel geometry at local scales (Vannote et al. 1980, Bain and Boltz 1989, Lobb and Orth 1991).

Predictable environmental gradients occur at the habitat scale along the spatial axes of a navigation pool, which are reflected in habitat classifications developed for navigation rivers (Pearson and Krumholz 1984, Wilcox 1993, Arway et al. 1995). As Table 3 indicates, a longitudinal riverine-to-lacustrine gradient interacts with cross-sectional geometry and meander pattern to influence near-shore to mid-channel gradients in depth, velocity, substrate, and cover. Where islands are present, the primary channel used for navigation usually differs distinctly from smaller secondary or tertiary channels opposite the island. Meander pattern influences near-shore erosion and deposition, resulting in steeply sloping banks with deep water near rocky shores alternating with shallow flats where gravel, sand, and silt predominate.

Natural patterns of regular habitat variation are interrupted by human modifications (e.g., bermed, rip-rapped, and bulkheaded shores and dredged channels) or by other natural, but irregular features (e.g., delta formations at tributary junctions), creating a complex mosaic with both predictable and unpredictable components. Temporally, a mosaic will change in a spatially dependent manner, because some areas will be inherently more dynamic than others. Terrestrial mosaics subjected to unpredictable patchy disturbance may still achieve an equilibrium distribution in the relative proportion of patches within different successional states (O'Niell 1989). Similarly, fluvial processes that continually alter the state of an alluvial channel at a fine scale are also responsible for maintaining an equilibrium pattern in channel dimension at a broader scale (Leopold 1994). Human disturbances at the broader reach and navigation pool scales exert control over local assemblages in part by the way they influence spatiotemporal variation at the habitat scale.

The filtering influence of fine-scale habitat heterogeneity ultimately determines which species in the regional pool will succeed in a given environmental setting. Conditions suitable for growth and reproduction must be maintained over long periods of time for a species to persist in a given environmental setting. The degree to which a species is able to withstand unfavorable conditions on occasion reflects traits that confer resistance to local habitat filters.

Generally, highly variable environments are occupied by assemblages composed of eurytopic species in which abiotic controls on temporal variation predominate, while stenotopic habitat specialists and biotic interactions become more important in stable environments (Matthews 1987, Poff and Ward 1989, 1990).

Temporal variation in lotic habitat is driven strongly by flow regime, which in large rivers has a strong contingent component reflecting predictable seasonal flow patterns (Poff and Ward 1989). Assemblages facing regular cycles of variation typically include species with adaptations for taking advantage of (or avoiding) favorable (or unfavorable) periods within a cycle of variability. Many large river fishes, for example, are strongly dependent on seasonally inundated floodplains and backwaters for reproduction. Such species can be reduced or eliminated from local assemblages by regulated flows that deviate too far from natural pattern (Poff et al. 1997), or by activities that disconnect spawning and nursery areas from the river.

Table 3. Gradients of habitat conditions found within navigation pools of large navigation rivers. Entries in “low” and “high” columns are habitat classification elements devised to compartmentalize variation at the habitat scale.

Habitat	**Relative value**		**Gradient**	**Gradient**
condition	Low	High	type	strength
Depth	near shore zone	midchannel zone	lateral	strong
	tertiary channel	primary channel	lateral	intermediate
	tailwater	lower pool	longitudinal	weak
Velocity and	lower pool	tailwater	longitudinal	intermediate
turbulence	near shore zone	midchannel zone	lateral	intermediate
	tertiary channel	primary channel	lateral	intermediate
Substrate	lower pool	tailwater	longitudinal	strong
coarseness	tertiary channel	primary channel	lateral	intermediate
	very wide border	very narrow border	lateral	intermediate
	island shore	valley shore	lateral	weak
Cover density	midchannel zone	near shore zone	lateral	strong
(e.g., vegetation,	primary channel	tertiary channel	lateral	intermediate
large woody debris)	valley shore	island shore	lateral	weak
	very narrow border	very wide border	lateral	weak

Trautman (1981) frequently comments on the negative or positive influence that the locks and dams had on many fishes of the Ohio River. By and large, the physical changes resulting from impoundment have negatively impacted most obligate riverine species that require moderate to strong current and coarse, or at least silt-free substrates. Impoundment favored “generalist” species, including many that are found naturally in both rivers and lakes.

HABITAT FILTERS AND GIS

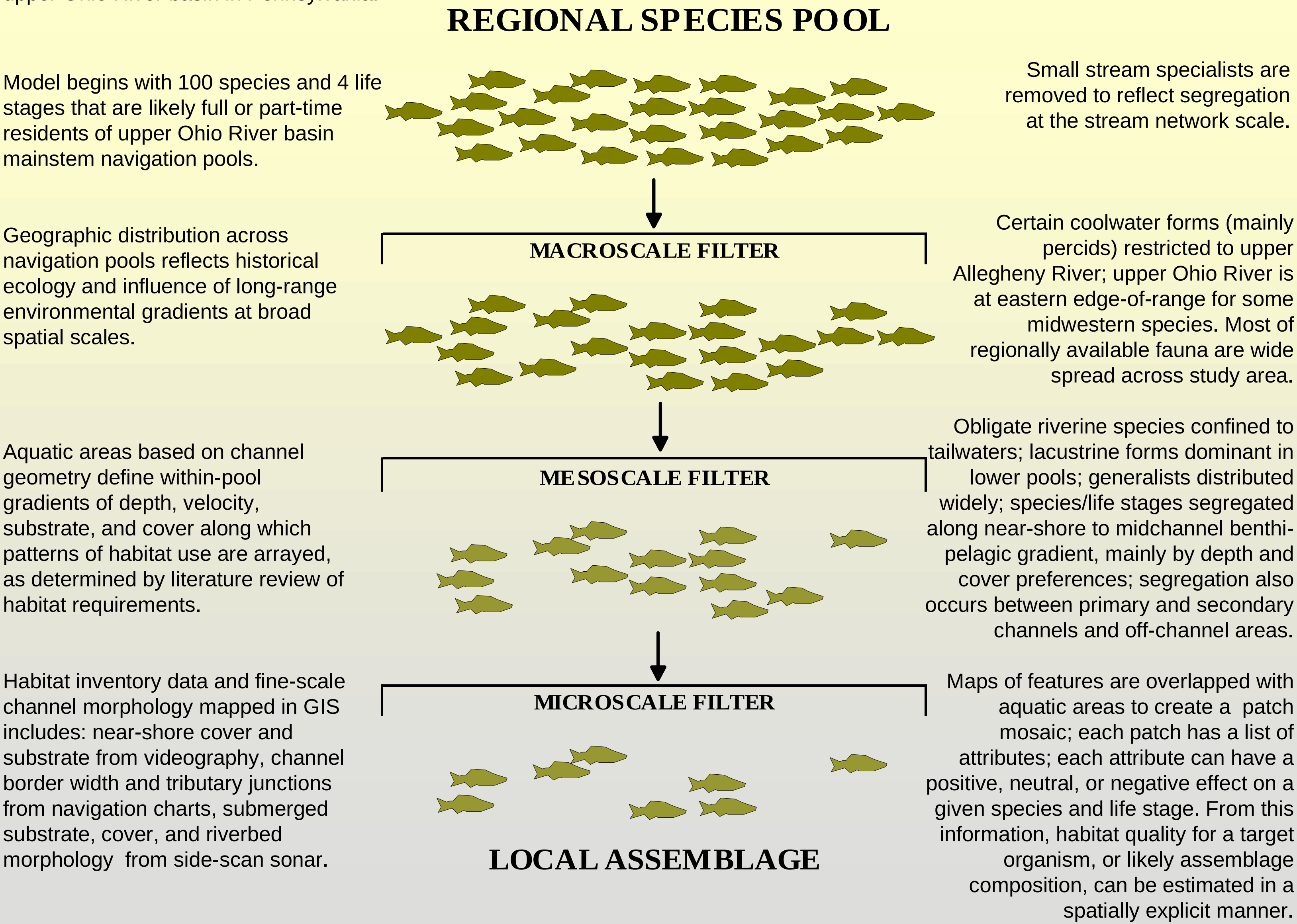
An Aquatic Resource Characterization Study (ARCS) for the Allegheny, Monongahela, and Ohio River in Pennsylvania (Arway et al. 1995) provides an example of a GIS application using the habitat filter concept. The objective of ARCS was to develop a GIS for natural resources management of the “Three Rivers” area (ten navigation pools and 176 km of mainstem) near Pittsburgh, PA. Base maps generated from 1:24,000-scale topography included the drainage network out to the limit of counties containing a defined study area, which incompletely captured the landscape at basin, network, and reach scales. Since our focus was on mainstems, main channels at 1:24,000 scale were revised from aerial photography at 1:8,400 scale to provide details for habitat mapping. Much of ARCS was devoted to habitat inventory and classification (Arway et al. 1995), including videography of near-shore cover (e.g., submerged and emergent vegetation, woody debris) and substrate (Nieman et al. 1996), and side-scan sonar of riverbed substrate and morphology (Nieman et al. 1999). Channel border width was obtained from navigation charts. Inventory data were ascribed to polygons and overlaid with maps of aquatic areas defined by channel geometry. For our purposes, the habitat components of the GIS can be envisioned as a mosaic resulting from the overlay of component coverages. Associated with each polygon are attributes that describe the mix of habitat conditions and aquatic area type found within it.

In a parallel effort, inventory data were linked to literature-based analyses of habitat requirements so that GIS applications could be devised for estimating relative habitat suitability for a selected fish species and life stage, or likely assemblage composition, for any user-defined subset of the study area. Because this process was complex, a conceptual framework was needed to organize the research effort. A habitat filter framework, broadly summarized in Figure 1, eventually arose as the conceptual model around which GIS applications were built. This framework grew out of an effort to devise an ecologically based habitat classification, which led to the realization that habitat influences on riverine biota must be organized hierarchically. The result was a model that focuses on how channel geometry constrains microhabitat variation along the spatial axes of a navigation pool, and that acknowledges further constraints on fish assemblages stemming from broader reach, network, and basin scales.

The left column of text in Figure 1 broadly summarizes model components associated with three levels of a simple hierarchy that correspond to one or more of the five spatial scales discussed previously. To the right are listed the chief ecological consequences of each component. At the "macro scale" level, some species in the regional species pool were removed *a priori* as small-stream specialists, while others were assigned membership in only a few of the ten pools of the ARCS study area, reflecting the influence of broad-scale constraints associated with existing and historical influences at basin, network, and reach scales.

At a "mesoscale", the influence of channel geometry on microhabitat distribution among aquatic areas was compared to literature on fish habitat requirements. Scores were assigned to reflect suitability for each species and life stage based on comparison of literature to "typical" microhabitat within aquatic areas. Since aquatic areas define environmental gradients, scores were analyzed along with descriptions of "typical" microhabitat (=depth, velocity, substrate, cover) variation among aquatic areas using canonical correspondence analysis (ter Braak 1986). The resultant ordination placed life stages of each species into "species-habitat associations" that reflect orientation along a riverine-to-lacustrine gradient and lateral gradients

Figure 1. Summary of a GIS-based filter model for estimating habitat suitability and fish assemblage composition in navigation pools of the upper Ohio River basin in Pennsylvania.



in the use of near-shore versus transitional off-shore and benthi-pelagic areas in mid-channel, along with differential use of primary, secondary-tertiary channels, and off-channel areas.

A “micro scale” filter operating at the finest level of resolution in the environmental data further modifies model output by accounting for local features that have positive, neutral, or negative influence on habitat quality for a target organism. Customized analyses in GIS and spreadsheet programs calculate and summarize weighted area statistics for individual cases, display choropleth maps, and rank all species and life stages by a value that reflects relative likelihood of occurrence in any user-defined subset of the study area.

There are many details of model development that are beyond the scope of this paper. Our main intent was to show how the concept of habitat filters could be used to organize thinking about relationships between fishes and riverine habitat. In summary, fish distribution and abundance in navigation rivers like the Ohio River can be related to historical legacies and environmental constraints at broad spatial scales, to the influence that channel geometry has on the distribution of microhabitat in navigation pools, and to local factors that proximally influence habitat use and the ability of individual species to persist in a given environment.

To date, the ARCS GIS has not been rigorously evaluated against field data. It does portray patterns of habitat use variation found in several empirical studies (Gutreuter 1992, PFBC 1992, Emery et al. 1999). Consideration of how hierarchical systems operate suggests that the ARCS GIS filter model should be better able to predict statistical patterns in large data sets, rather than details of individual sampling events, because of the influence of random noise at fine spatiotemporal scales. For example, comparisons of model outputs from Allegheny River pools 5 and 6 with results of a PFBC (1992) field study compare favorably when aggregated for each pool, despite unpredictable variation within individual field samples. Users of the ARCS GIS and similar models should be aware of scale dependencies in ecological patterns, so that analyses are applied and results interpreted properly. Field studies designed with these concepts in mind are likely to yield new insights into fish-habitat relationships in large rivers.

LITERATURE CITED

Allen, T. F. H., and T. B. Starr. 1982. Hierarchy: perspectives for ecological complexity. University of Chicago Press, Chicago, Illinois.

Arway, J., D. Nieman, T. Proch, and J. Schulte. 1995. Aquatic resource characterization of the upper Ohio River basin using a geographic information system. Pages 381-386 *in* Proceedings of the 1995 International Oil Spill Conference, Washington, D.C.

Bain, M. B., and J. M. Boltz. 1989. Regulated streamflow and warmwater stream fish: a general hypothesis and research agenda. United States Department of Interior, Fish and Wildlife Service Biological Report 89(18). Alabama Cooperative Fish and Wildlife Research Unit, Auburn, Alabama.

Baker, J. A., K. J. Killgore, and R. L. Kasul. 1991. Aquatic habitats and fish communities in the lower Mississippi River. Reviews in Aquatic Sciences 3:313-356.

Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. Ecology 73: 1045-1055

Casti, J. L. 1994. Complexification—explaining a paradoxical world through the science of surprise. Harper-Collins, New York, New York.

Cooper, E. L. 1983. The fishes of Pennsylvania and the northeastern United States. Pennsylvania State University Press, University Park, Pennsylvania.

Dodge, D. P. (editor). 1989. Proceedings of the international large river symposium. Canadian Special Publication of Fisheries and Aquatic Sciences 106.

Emery, E. B., T. P. Simon, and R. Ovies. 1999. Influence of the family Catostomidae on the metrics developed for a Great Rivers Index of Biotic Integrity. Pages 203-224 *in* Simon (1999).

Fairchild, G. W., R. J. Horwitz, D. A. Nieman, M. R. Boyer, and D. F. Knorr. 1998. Spatial variation and historical change in fish communities of the Schuylkill River drainage, southeast Pennsylvania. American Midland Naturalist 139:282-295.

Frissell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10: 199-214.

Gammon, J. R., and J. M. Reidy. 1981. The role of tributaries during an episode of low dissolved oxygen in the Wabash River, Indiana. Pages 396-407 *in* C. F. Bryan, G. E. Hall, and G. B. Pardue, editors. The warmwater streams symposium. Southern Division, American Fisheries Society, Bethesda, Maryland.

Gorman, O. T. 1986. Assemblage organization of stream fishes: the effects of rivers on adventitious streams. American Naturalist 120:423-454.

Gregory, S. V., F. J. Swanson, and W. A. McKee. 1991. An ecosystem perspective of riparian zones. BioScience 40:540-551.

Gutreuter, S. 1992. Systemic features of fisheries of the upper Mississippi River system 1990 fisheries component annual report. Long Term Resource Monitoring Program Technical Report 92-T001. U. S. Fish and Wildlife Service Environmental Management Technical Center, Onalaska, Wisconsin.

Gutreuter, S. 1993. A statistical review of sampling of fishes in the Long Term Resource Monitoring Program. Long Term Resource Monitoring Program Technical Report 93-T004. National Biological Survey Environmental Management Technical Center, Onalaska, Wisconsin.

Imhoff, J. G., J. Fitzgibbon, and W. K. Annable. 1996. A hierarchical evaluation system for characterizing watershed ecosystems for fish habitat. Canadian Journal of Fisheries and Aquatic Science 53 (Supplement 1): 312-326.

Junk, W. J., P. B. Bayley, and R. E. Sparks. 1989. The flood pulse concept in river-floodplain systems. Pages 110-127 *in* Dodge (1989).

Kolasa, J. 1989. Ecological systems in hierarchical perspective: breaks in community structure and other consequences. Ecology 70: 36-47.

Legendre, P., and M.-J. Fortin. 1989. Spatial pattern and ecological analysis. Vegetatio 80:107-138.

Leopold, L. B. 1994. A view of the river. Harvard University Press, Cambridge, Massachusetts.

Levin, S. A. 1995. The problem of pattern and scale in ecology. Pages 275-326 *in* T. M. Powell and J. H. Steele, editors. Ecological time series. Chapman and Hall, New York, New York.

Lobb, M. D. III, and D. J. Orth. 1991. Habitat use by an assemblage of fish in a large warmwater stream. Transactions of the American Fisheries Society 120: 65-78.

Lubinski, K. 1993.A conceptual model of the upper Mississippi River system ecosystem. Long Term Resource Monitoring Report Technical Report 93-T001. U.S. Fish and Wildlife Service Environmental Management Technical Center, Onalaska, Wisconsin.

Matthews, W. J. 1998. Patterns in freshwater fish ecology. Chapman and Hall, New York, New York.

Matthews, W. J. 1987. Physicochemical tolerance and selectivity of stream fishes as related to their geographic ranges and local distribution. Pages 111-120 *in* Matthews and Heins (1987).

Matthews, W. J. and D. C. Heins (editors). 1987. Community and evolutionary ecology of North American stream fishes. University of Oklahoma Press, Norman, Oklahoma.

Nieman, D. A., J. Arway, T. Proch, J. Schulte, G. Sermarini, and R. Shema. 1996. Construction of a geographic information system for natural resources management in a large navigation river system. Pages 924-938 *in* Practical environmental directions: a changing agenda. National Association of Environmental Professionals 21st annual conference proceedings. Washington, D.C.

Nieman, D. A., G. S. Sermarini, S. Sherman, and W. S. Ettinger. 1999. Use of side-scan sonar and GIS for mapping substrate and dredging disturbance in large rivers. Pages 129-148 *in* R. A. Randall, editor. Proceedings of the Western Dredging Association Nineteenth Technical Conference and Thirty-first Texas A&M Dredging Seminar. Center for Dredging Studies, Texas A&M University, College Station, Texas.

O'Neill, R. V. 1989. Perspectives in hierarchy and scale. Pages 140-156 *in* J. Roughgarden, R. M. May, and S. A. Levin, editors. Perspectives in ecological theory. Princeton University Press, Princeton, New Jersey.

Osborne, L. L., and M. J. Wiley 1992. Influence of tributary spatial position on the structure of warmwater fish communities. Canadian Journal of Fisheries and Aquatic Sciences 49:671-681.

Pearson, W. D., and L. A. Krumholz. 1984. Distribution and status of Ohio River fishes. Oak Ridge National Laboratory ORNL/Sub/79-7831/1. Oak Ridge, Tennessee.

Petts, G. E. 1989. Perspectives for ecological management of regulated rivers. Pages 3-24 *in* J. A. Gore and G. E. Petts, editors. Alternatives in regulated river management. CRC Press, Inc., Boca Raton, Florida.

PFBC (Pennsylvania Fish and Boat Commission). 1992. Dredging impacts study, pools 5 and 6, Armstrong County, Pennsylvania. Unpublished internal agency report, Harrisburg, Pennsylvania.

Pringle, C. M. 1997. Exploring how disturbance in transmitted upstream: going against the flow. Journal of the North American Benthological Society 16:425-438.

Poff, N. L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16: 391-409.

Poff, N. L., and J. V. Ward. 1990. Physical habitat template of lotic systems: recovery in the context of historical pattern of spatiotemporal heterogeneity. Environmental Management 14: 629-645.

Poff, N. L., and seven coauthors. 1997. The natural flow regime. BioScience 47:769-784.

Rabeni, C. F., and S. P. Sowa. 1996. Integrating biological realism into habitat restoration and conservation strategies for small streams. Canadian Journal of Fisheries and Aquatic Sciences 53 (Supplement 1):252-259.

Richards, C., L. B. Johnson, and G. E. Horst. 1996. Landscape-scale influences on stream habitats and biota. Canadian Journal of Fisheries and Aquatic Sciences 53 (Supplement 1): 295-311.

Rosgen, D. L. 1994. A classification of natural rivers. Catena 22: 169-199.

Ryder, R. A., and J. Pesendorfer. 1989. Large rivers are more than flowing lakes: a comparative review. Pages 65-85 *in* Dodge (1989).

Schlosser, I. J. 1990. Environmental variation, life history attributes, and community structure in stream fishes: implications for environmental management and assessment. Environmental Management 14: 621-628.

Schlosser, I. J. 1991. Stream fish ecology: a landscape perspective. BioScience 41: 704-712.

Sedell, J. R., J. E. Richey, and F. J. Swanson. 1989. The river continuum concept: a basis for the expected ecosystem behavior of very large rivers? Pages 49-55 *in* Dodge (1989).

Sheldon, A. L. 1987. Rarity: patterns and consequences for stream fishes. Pages 203-209 *in* Matthews and Heins (1987).

Simon, T. P. (editor). 1999. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Smith, C. L., and C. R. Powell. 1971. The summer fish communities of Brier Creek, Marshall County, Oklahoma. American Museum Novitates 2458: 1-30.

Strange, R. M. 1999. Historical biogeography, ecology, and fish distributions: conceptual issues for establishing IBI criteria. Pages 65-78 *in* Simon (1999).

Taylor, C. M., M. R. Winston, and W. J. Matthews. 1993. Fish species-environment and abundance relationships in a Great Plains river system. Ecography 16:16-23.

ter Braak, C. J. F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67:1167-1179.

Tonn, W. M. 1991. Climate change and fish communities: a conceptual framework. Transactions of the American Fisheries Society 119: 337-352.

Trautman, M. B. 1981. The fishes of Ohio. Revised edition. The Ohio State University Press, Columbus, Ohio.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37: 130-137.

Ward, J. V., and J. A. Stanford. 1989. Riverine ecosystems: the influence of man on catchment dynamics and fish ecology. Pages 56-64 *in* Dodge (1989).

Wilcox, D. B. 1993. An aquatic habitat classification system for the upper Mississippi River system. Long Term Resource Monitoring Program, U. S. Fish and Wildlife Service Technical Report 93-T0023. Onalaska, Wisconsin.

Winemiller, K. O., and K. A. Rose. 1992. Patterns of life-history diversification in North American fishes: implications for population regulation. Canadian Journal of Fisheries and Aquatic Sciences 49: 2196-2218.

GIS as a Tool for Predicting Urban Growth Patterns and Risks From Accidental Release of Industrial Toxins

Samuel V. Noe
University of Cincinnati, Cincinnati, Ohio

Introduction

The catastrophic Bhopal incident demonstrated to the world what could happen when industry and population are geographically incompatible. Many believe that the large urban population "should not have been there." A recent publication, "New York Under a Cloud," presents a frightening map of New York State that indicates potential areas of serious population exposure to accidental releases of chemicals stored by area industries and municipalities.

Conventional urban planning and administrative practices at the local level do not adequately provide for the minimization of these risks. Local jurisdictions on the fringes of metropolitan areas may be particularly ill-equipped to respond and plan effectively. Their elected officials, supported by minimal professional staffs and unaware of specific potential risks, may be more interested in soliciting new industrial development along with the tax base it brings. They therefore create industrial zones without restricting facilities that may generate hazardous substances and without recognizing the possibility that underground aquifers, which are current or potential sources of drinking water, may underlie these zones. Jurisdictions often permit facilities that could routinely or accidentally release toxic substances into the air without due regard for prevailing wind patterns or existing or projected urbanized areas that may be affected.

Although the available data and methodology have some gaps, much of the knowledge required to provide adequate protection from these risks exists. We know how to identify the hazardous substances that these sites may produce or store and how to calculate the types and levels of risks associated with them. We can accurately map the locations of streams, underground aquifers, and their catchment areas. Although with less precision, we also can indicate the areas more likely to receive the outfall of airborne and waterborne pollutants.

At the other end of the equation, we can construct models for projecting patterns of urban growth. Those responsible for planning, however, have not made the connections between these techniques. Hazardous facilities sites are thus still permitted in areas that place existing urban residents and their drinking water supplies at risk, and new urban development grows in areas polluted by existing hazardous substance sites.

Clearly, this situation displays a need for the coordinated application of scientific risk assessment techniques and new approaches to regulating urban development. Equally critical, however, is the need to give greater attention to formulating appropriate public policy measures at the local and state levels for dealing with the complex disputes that surround these issues.

Project Background

The project this paper describes addressed these needs. It was undertaken by a team of faculty from the University of Cincinnati's School of Planning, Department of Environmental Health, and the College of Law. The study team focused on the accidental release of hazardous materials both into the air and into the ground-water supply. The team's purpose was to develop an integrated approach to scientific risk assessment, environmental analysis, urban planning, and policy analysis to address conflicts between:

- Expected patterns of suburban residential growth.
- The need to safeguard existing and new residential areas, and their water supplies, from toxic chemical pollution.
- The promotion of industrial development on the peripheries of urban regions, which often leads to the proliferation of hazardous substances sites.
- The need for effective regulation of these sites in complicated multijurisdictional environments.

This project examined these issues within a 100-square-mile area on the northern edge of metropolitan Cincinnati. The study area is not yet completely urbanized but lies in the path of urbanization. It contains a significant number of industrial or storage facilities that house supplies of hazardous materials. A major aquifer serving as a public water supply source passes under the area. Approximately 17 local jurisdictions fall within the study area: two counties, six townships, and nine municipalities. The area encompasses an intricate mix of agricultural, residential, and commercial land uses. In addition, several major industrial concentrations, as well as a number of jurisdictions, are aggressively soliciting new industrial employment. Because of its proximity to most major employment sites in southwest Ohio and to a variety of large retail complexes, the area is experiencing rapid residential development.

Projecting Areas of Future Development

The study team used PC ARC/INFO geographic information systems (GIS) to project the locations of future residential and industrial growth in the study area, to show the locations of areas at various degrees of risk from either airborne or waterborne industrial toxins, and to reveal the potential areas of population exposure resulting from the overlap of these areas.

Although this paper does not describe the models used to project residential and industrial growth in the study area, it does include the criteria used to make projections. The criteria we used to project residential growth were:

- Travel times to major employment concentrations in the region.
- Proximity to interstate highways, interchanges, and main trunk sewers.
- Avoidance of areas composed of steep slopes and flood plains.
- Land currently zoned for agriculture or housing.

The criteria for projecting industrial areas were:

- Relatively flat, not in a flood plain, and zoned industrial.
- Proximity to existing industrial development, main trunk sewers, and interstate highways.
- Relative aggressiveness of local jurisdictions in attracting industrial development.

Identifying Areas at Risk From Airborne Releases

The study team determined the model project areas at risk from airborne releases by using information available from the Ohio Environmental Protection Agency. This information, which included the location, identity, and quantity of hazardous materials recently stored in the study area, was available as a result of reporting requirements mandated by several federal statutes. The study team assumed that a similar ratio reflecting sites containing hazardous materials to the area of industrially developed land would continue into the future. Based on this assumption, the study team could randomly project new potential release sources.

The algorithm and associated software used for calculating the plume size of aerial dispersion of hazardous chemicals was Aerial Locations of Hazardous Atmospheres (ALOHA). The National Oceanic and Atmospheric Administration (NOAA) developed this system, which is in wide use by government and industry for the preparation of emergency contingency plans. This software is available for use on any Macintosh computer, or any IBM-compatible with an Intel 80286 (or better) CPU. This software employs three classes of variables in calculating the plume dispersion for a specific chemical:

- Chemical variables
- Meteorological options
- Source strength options

Chemical variables include both physical properties of the chemical and parameters that define the human health effects of the chemical. In the latter case, the two variables are the threshold limit value (TLV) and the immediately dangerous to life and health (IDLH) value. The TLV is a measure of chronic toxicity of the chemical in humans. It represents the maximum concentration of the chemical in air to which a human can safely be exposed for 8 hours per day on a daily basis. The IDLH is a measure of acute toxicity of the chemical in humans. The IDLH represents the maximum concentration to which a human can be exposed for a short time and not experience death or some other severe endpoint.

Meteorological options describe the ambient atmospheric conditions into which the chemical disperses. ALOHA has the capability of downloading real-time data from NOAA satellites. This case, however, employed average meteorological conditions for the study area over the course of a year. The variables this study used were atmospheric inversion height (or no inversion), wind velocity, air temperature, ground roughness (rural or urban), and stability class (a combined variable describing cloud cover and incoming solar radiation).

The source options quantify the amount of chemical being released and describe how the chemical is released (instantaneous or continuous).

The ALOHA model provides a procedure for showing the IDLH and TLV risk zones from a single accidental release of a single chemical, given specified conditions of atmospheric stability, wind direction, and air temperature. Obviously, climatic conditions change daily, so areas surrounding a single industrial site experience different degrees of risk depending on the variability of these conditions. Moreover, a single site that can potentially release more than one chemical poses a higher risk to surrounding areas. Finally, when release plumes from two or more sites that are located relatively close to each other overlap, risks also increase. To account for these factors of climatic variation and overlapping chemical release plumes, the study team constructed the model described below.

In discussing these procedures, bear in mind that the risk factors are relative. Sufficient data are not available to estimate the absolute probability of an accidental release of industrial toxins into the air. Consequently, no absolute risk levels can be estimated.

1. We acquired NOAA climatic statistics for a full year for the weather station nearest the study. NOAA tabulates eight wind directions (N, NE, E, SE, S, SW, W, and NW). We sorted the average daily wind speeds and average daily temperatures according to the daily prevailing wind direction. Thus, for each of the eight wind directions, it was possible to determine the number of days in the year that the prevailing wind comes from that direction, as well as the average daily wind speed and temperature.

2. We then input the temperature and wind speed data, derived as explained above, into the ALOHA model to prepare plots of the IDLH and TLV zones, or plumes, for each of the eight wind directions. Individually, the plumes emanated downwind from the source of the release. In our study, IDLH plumes varied from 0.17 miles to greater than 10 miles. TLV plumes varied from 0.52 miles to greater than 10 miles.

   When the plumes for the eight different wind directions were combined, the results were translated into a pattern of wedges representing various plume lengths in each of the eight different wind directions. These risk levels vary according to the number of days per year the wind blows in each direction from the source of the release. Plumes blowing in different directions vary in length according to average temperature and wind velocities for the days the wind blew in each direction. The numbers in the wedges represent risk factors assigned as indicated in Table 1.

Table 1. Assigned Risk Factors

Frequency of Wind	Risk Factor
0-25 days/year	1
26-50 days/year	2
51-75 days/year	3
76-100 days/year	4

3. When a single industrial location employed more than one chemical, the IDLH and TLV risk patterns for each were overlaid on one other. Where the overlap occurred, we added the risk factors together.

4. Finally, when the risk patterns from two or more sites overlapped, we added together the risk factors assigned to overlapping areas. The overlap capabilities of GIS allowed us to easily draw and combine the risk patterns, superimposed on a map of the industrial sites under consideration.

When we compared the eight existing combined IDLH and TLV risk patterns with the risk patterns that appear when six projected new sources of releases and projected areas of residential growth are added, the changes are rather dramatic. Of course, we must note that the projected sources of potential releases are hypothetical and their locations selected at random. In reality, precise locations of areas at greater risk cannot be predicted with any degree of certainty. We can reasonably deduce from the maps, however, that any substantial increase in industries storing or using toxic chemicals that might be accidentally released into the air can compound risk levels much more than might be expected. The maps we produced indicated five risk levels.

The study results also showed that an industry capable of generating accidental releases of airborne toxins will very likely place at risk residents not only of the same community but those in neighboring jurisdictions as well. This is particularly true on the fringes of metropolitan areas where highly fragmented political boundaries exist. This fact complicates tremendously the ability of each jurisdiction to protect its citizens and suggests a need for a more comprehensive approach to regulation than conventional land use zoning measures that each locality administers.

Identifying Areas at Risk From Accidental Releases Into Ground-Water Supplies

As the introduction of this paper indicates, local government usually manages conventional land use planning, while air quality is largely a state or federal responsibility. Thus, decisions regarding the regulation of industrial location may not account for the types of industrial operations proposed, the possible use of hazardous materials, and possible risks to local residents from accidental release of toxins into the air. The same problem applies to the location of industries that have the potential for accidental release of hazardous materials into ground-water supplies. In our study, we outlined a technique for predicting where local residents may be placed at risk by drinking water from sources vulnerable to contamination by industrial toxins.

As stated earlier, the study area includes a major aquifer. We noted industrial and related facilities that use and store hazardous materials, and that are located over or immediately adjacent to the area aquifers. We distinguished between sites where a previous spill had been reported, those where a waste well is located, and all other locations where a hazardous material is used or stored. The results clearly indicated a significant potential for contamination.

To predict the number and locations of additional industrial toxin sources that might appear over a 10-year period following 1990 (the base year of the study), we employed a procedure similar to that used in the air pollution section of the study. We assumed that during this period, the number of such sites would increase by 35 percent. The 35-percent increase represented an arbitrarily selected figure approximately midway between estimates of 25- and 50-percent industry growth in the study area. We used this increment to project additional waste well sites and sites that would experience a spill sometime during the decade, as well as the total number of new sites.

We projected the locations of the additional sites by overlaying a 5,000-foot by 5,000-foot grid on the study area and assigning the new sites to grid cells using a random number generator. We considered only cells lying completely or largely over the aquifers and also falling within the area of projected industrial land use. Each cell was assigned a relative contamination risk factor based on the number of projected sites it contained, with a multiplier of 3 applied to sites with a waste well and a multiplier of 4 applied to sites assumed to have had spills. We also assigned existing sites to grid cells and scored them in the same manner.

To obtain more information, we used a simplified version of DRASTIC maps prepared by the Ohio Department of Natural Resources. DRASTIC is an acronym for **D**epth to water, net **R**echarge, **A**quifer media, **S**oil media, **T**opography, **I**mpact of the vadose zone, and hydraulic **C**onductivity of the aquifer. These factors contribute to an index of the relative vulnerability of the aquifer to pollution. Different shades on the map represented the relative vulnerability of sections of the aquifers to ground-water pollution. We assigned two points to aquifers with a DRASTIC pollution potential index less than 180 and four points to aquifers with a higher index. This allowed us to use the GIS to combine the DRASTIC vulnerability map with the maps that showed risks of pollution from the hazardous materials sites. The resultant maps showed the existing and projected potential risk of pollution in different areas of the aquifers.

The next step in the study related the above information to public water companies that extract water from different parts of the aquifers. Thus, we were able to associate risk levels with well locations as well as with the areas served by water companies whose supplies may be at risk. In this case, the projected risks of polluted water supplies were identical to existing risk levels because no public water wells happened to be located in areas where additional potential sources of pollution were projected.

Of course, wide areas in the study area had no public water supply. Residents in these areas may be at risk depending on where they dig private wells. The maps we generated showing areas at risk nonetheless provide useful guides to potentially hazardous areas.

As in the construction of all such models, we needed to make a number of assumptions, simplifications, and value judgments. In this case, these included the projected number of new toxic sites and point scores assigned to hazardous materials sites and the various areas in the DRASTIC maps. Also, for the sake of simplicity, we projected no new well sites in preparing one of the maps. This additional element should probably be included, assuming local water companies could provide projected well locations. Use of a GIS, however, makes it possible to explore the implications of adjustment of any of these factors.

We must note two more significant omissions from the model that the GIS cannot factor in. One was our inability to identify from the data specific chemicals that each sight might release and the relative effects of each. We would have required more complicated techniques for dealing with these variables and for projecting the travel and dilution of plumes of contaminants in an aquifer. In the interest of providing a simple, if relatively crude, model capable of replication by a local planning agency with a simple GIS, we elected not to propose use of more sophisticated techniques.

Another obvious omission was the consideration of water pumping and treatment measures that might mitigate risks of water contaminated by accidental release of industrial toxins. Perhaps, with knowledge of the specific contaminants found in the water at any given time, such mitigation might be possible. We elected not to consider this factor for reasons of simplification but also because this study aimed to provide planning agencies the means to identify potential risks to local residents and to prevent or minimize them through better land use planning and regulatory measures.

Some Final Notes

As stated earlier, the purpose of the study was to propose a technique that planning agencies could use to identify:

- The locations of existing and projected patterns of residential and industrial development in a multijurisdictional suburban area.

- The locations of existing and projected industrial, storage, and disposal sites of hazardous materials.
- Residential areas that might be placed at risk by the accidental release of these hazardous materials into the air or into ground-water supplies.
- The relative levels of risks resulting from potential exposure to more than one hazardous material at a single site, or from multiple sites in the vicinity.

The technique we used in the study permitted projection of *relative* risk levels. Projecting absolute risk levels is impossible without data on the actual incidence of accidental releases of toxins over time in this or in similar areas. A related question is whether it is possible to meaningfully indicate the combined risk from airborne releases of industrial toxins and drinking water contamination to a particular residential area. The issue of weighting relative risk factors is central here. Could we, for example, weight the risk levels of exposure to airborne releases three—or possibly four—times higher than water contamination risks? Or can we say that the risks may be the same, but the danger from airborne releases is three or four times greater?

Obviously, these would be futile exercises, especially because the point scores in the separate mapping studies were arbitrarily assigned. The public should know, however, which present or projected residential areas carry some level of risk from both types of exposure. Thus, after calculating point scores to derive relative risk levels from waterborne pollutants, we multiplied the scores by 4 to bring their maximum ranges into the same order of magnitude. Otherwise, the effect of waterborne pollutants would not be apparent. Consequently, we created two maps to show the combined existing and projected risks; therefore, we highlighted the combined risks that residents face from both airborne and waterborne hazardous materials.

The maps of airborne releases used in these combinations showed TLV. Continuing exposure within the TLV areas over an extended period can also have adverse health effects. This study focused only on accidental releases, however, and a single release is unlikely to sustain continuous exposure. Of course, residential areas at risk from several sites might approach conditions of sustained exposure. This situation is more analogous to prolonged exposure to contaminated drinking water supplies. We did not combine the maps of IDLH airborne releases with the maps of areas at risk from groundwater contamination because they are not analogous conditions. Nonetheless, the IDLH risk maps in themselves reveal the conditions that local planning officials should most seriously consider.

Replicating the procedures outlined in this study should be technically and financially feasible for local planning agencies. Armed with the results of such an investigation, their next step should be to establish the planning and regulatory measures that would minimize both existing and projected levels of risk to area residents. The attorney on our team has outlined a range of possible measures, but detailing them would be the subject of another paper.

Use of GIS in Modeling Ground-Water Flow in the Memphis, Tennessee, Area

James Outlaw and Michael Clay Brown
University of Memphis, Memphis, Tennessee

Abstract

Memphis, Tennessee relies solely on ground water for its municipal and industrial water supply. Memphis Light, Gas, and Water (MLGW) Division owns and operates over 160 water wells in 10 production fields throughout Shelby County. MLGW produces an average of approximately 200 million gallons per day, excluding much of the industrial demand. The city obtains its water from a thick, prolific aquifer known as the Memphis Sand, which was thought to be separated from a surficial aquifer by a thick confining layer. In recent years, evidence of leakage from the surficial aquifer to the Memphis Sand has been found.

The University of Memphis Ground Water Institute (GWI) is developing a hydrogeologic database of the Memphis area to study the aquifer. The database serves as the basis for several ground-water flow models that have been created as well as part of the wellhead protection programs currently being developed for Memphis and other municipalities in Shelby County. A geologic database was developed and is constantly being updated from borehole geophysical logs made in the area. Well locations are being field verified using a global positioning system (GPS).

Use of the database has allowed the development of a three-dimensional model of the Memphis area subsurface. The database also contains locations of and information on both private and public production and monitoring wells, Superfund sites, underground storage tanks, city and county zoning, land use, and other pertinent information. Procedures for linking the database to ground-water flow and solute transport models have been developed. The data visualization capabilities and the ability to link information to geographic features make geographic information systems (GIS) an ideal medium for solving ground-water problems.

An example of GIS use in ground-water flow modeling is the study of the Justin J. Davis Wellfield. The water quality parameters of alkalinity, hardness, sulfate, and barium have significantly increased over the past 10 years at this facility. To understand why these changes are occurring, MLGW, the GWI, and the U.S. Geological Survey (USGS) participated in a joint investigation of the wellfield.

In the spring of 1992, a series of 12 monitoring wells was drilled into the surficial aquifer near the production wells. Geophysical logging and split-spoon sampling revealed an absence of the confining layer, referred to as a window, at one of the monitoring wells. All other wells penetrated various thicknesses of clay. This window in the confining layer suggests that the water quality changes could be due to leakage from the surficial aquifer to the Memphis Sand.

The GIS database was used to construct a flow model of the Davis area. Also, using the surface modeling capabilities of GIS, the extent of the confining layer window was estimated and used to calculate leakage between the two aquifers. The results of these analyses also indicate that further subsurface exploration is needed to more accurately define the extent of the confining layer window.

Introduction

Memphis, Tennessee, relies solely on ground water for its municipal and industrial water supply. The Memphis Light, Gas, and Water (MLGW) Division owns and operates over 160 water wells in 10 production fields throughout Shelby County, as shown in Figure 1. MLGW produces an average of approximately 200 million gallons per day, excluding much of the industrial demand.

The city obtains its water from a thick, prolific aquifer known as the Memphis Sand, which was thought to be separated from the surficial aquifer by a thick confining layer. In recent years, evidence of leakage from the surficial aquifer to the Memphis Sand has been found. The University of Memphis Ground Water Institute (GWI) is developing a hydrogeologic database of the Memphis area to study the aquifer. Several ground-water

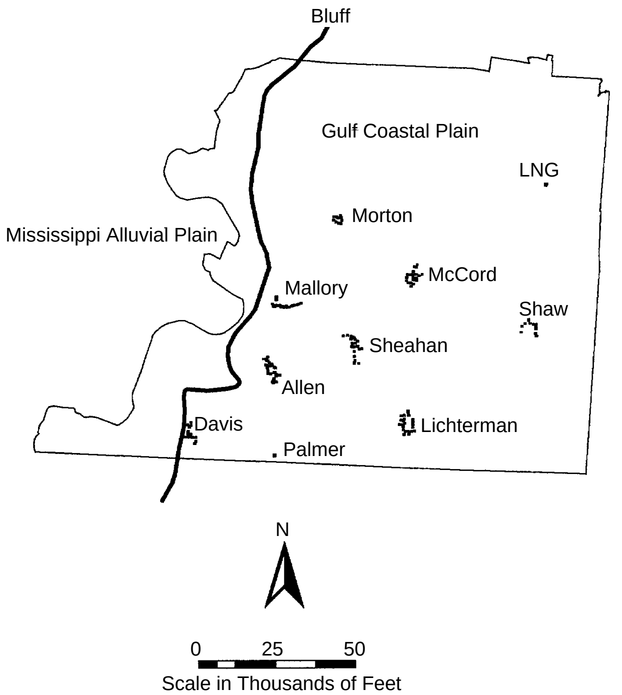


Figure 1. Physiographic description and MLGW Wellfields in Shelby County.

flow models have been developed using the database. Also, the database is an integral part of wellhead protection programs being developed for Memphis and other municipalities in Shelby County. A geologic database was developed and is constantly being updated from borehole geophysical logs made in the area. Well locations are being field verified using a global positioning system (GPS).

Use of the database has allowed the development of a three-dimensional model of the subsurface of the Memphis area. The database also contains locations of and information on both private and public production and monitoring wells; Superfund sites; underground storage tanks; city and county zoning; land use; and other pertinent information. Water quality measurements for every MLGW production well have been obtained, and a history of water quality for the Memphis Sand is being developed. Procedures have been developed for linking the database to ground-water flow and solute transport models (1). The data visualization capabilities and the ability to link information to geographic features make geographic information systems (GIS) an ideal medium for solving ground-water problems.

GIS Database

The GWI has developed and is continuing to update a hydrogeologic database for the Memphis area. ARC/INFO (marketed by Environmental Systems Research Institute, Redlands, California) is the GIS program that the GWI is using. The program runs on a network of 10 SUN SPARC stations. The capabilities of ARC/INFO and the computational speed of the SPARC stations allow very sophisticated ground-water analyses to be performed and have allowed the development of an extensive electronic database.

The basic unit of data storage in ARC/INFO is a coverage. A coverage is a digital representation of a single type of geographic feature (e.g., points may represent wells, lines may represent streets or equipotential lines, and polygons may represent political boundaries or zoning classifications). Information may be associated with an individual geographic feature in a feature attribute table. This information may then be queried and used in analyses. ARC/INFO also has its own macro language that allows the customization and automation of many ARC/INFO procedures.

A relatively new feature of ARC/INFO is address matching. This procedure compares a file containing the street address of a particular feature with an address coverage. This coverage is basically a library of addresses that are linked to a geographic coordinate. As the addresses from the input file are compared with the address coverage, the matching points are written to a second coverage. Any addresses in the input file that do not match an address in the address coverage are written to a "rejects" file. These can be matched by hand on a one-by-one basis.

Address matching serves as an alternative to digitizing, as long as a good address coverage for a specific area exists. The GWI has used this capability extensively and has developed a coverage of underground storage tank (UST) locations inside Shelby County. A database of private and monitoring wells is also being developed and updated using address matching. The raw information was obtained in an ASCII format from the appropriate regulating agencies (i.e., the Tennessee Department of Environment and Conservation Division of Underground Storage Tanks and the Memphis/Shelby County Health Department). The ASCII information was imported into ARC/INFO and address matched. The creation of a suitable address coverage and completion of the address matching of the UST file has taken almost 2 years. The private well coverage is currently being updated from historical information provided by regulating agencies and local well drilling companies.

An important part of the database is the geologic information obtained from geophysical logs in the area. Gamma logs, resistivity logs, and spontaneous potential (SP) logs are three major types of electric geophysical logs. Gamma logs measure naturally occurring radiation emitted from soil in the borehole. Clays and shales emit gamma rays. A high gamma count indicates the presence of clay or shale, and a low gamma count implies that little or no clay is present. Sand layers that contain fresh water are located using resistivity logs. Maximum values of resistivity indicate the possibility of a sand layer. Clays and sands that contain salt water may exhibit similar resistivities. SP logs are used to differentiate between the two (2).

Corroborating data such as formation logs, geologic studies, and available material samples should be consulted when reading and interpreting geophysical logs. The accuracy and reliability of an application based on well logs is completely dependent on a realistic interpretation of the geophysical data. A sample interpretation of a set of geophysical logs is shown in Figure 2. The results of interpretations like this are entered into the point attribute file of a well coverage.

The Triangulated Irregular Network (TIN) module of ARC/INFO is used to create a three-dimensional surface from information stored in a coverage. TIN creates a surface from a set of nonoverlapping triangles defined by a set of irregularly or regularly spaced points. In this study, the points defining the triangular TIN surfaces are the locations of the wells in the model area. TIN uses various interpolation routines to estimate surface values. Once the surfaces have been developed, two-dimensional profiles can be made that show the relative thicknesses of the various soil strata, as shown in Figure 3. These profiles aid in selecting boundary conditions and defining layers in ground-water flow models (3, 4).

In addition to the creation of profiles, a process that extracts surface values for use in a ground-water flow model has been developed. The GWI uses the United States Geological Survey (USGS) flow model, MODFLOW (5). Being a cell-based model, MODFLOW requires a value for each hydrogeologic parameter for each cell in the model grid. A series of FORTRAN programs and arc macro language (AML) programs were coupled to extract the required hydrogeologic data from surface models. For example, piezometric surface values are required to set initial conditions for the model. A coverage of the piezometric surface of the Memphis Sand was created, converted to a TIN surface, and the required values for each cell in the model were extracted using the procedure described above.

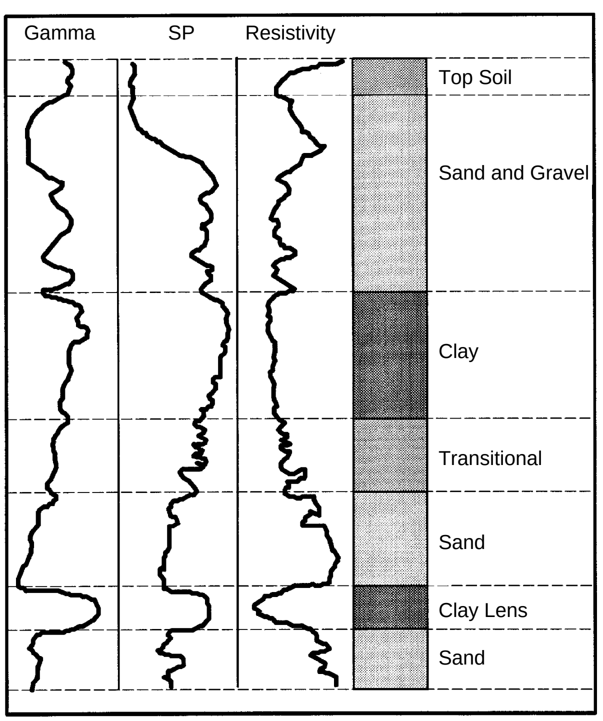


Figure 2. Example interpretation of geophysical logs.

The results and hydrogeologic data from the calibrated model can be read back into the database and converted into coverages. This allows piezometric contours to be developed and displayed with other information in the database to aid in decision-making. Also, capture zones for the wells can be brought into the database and compared with surface features like industries, landfills, Superfund sites, UST locations, or other sites that may have an impact. This has proved especially helpful in developing wellhead protection programs where a complete contaminant source inventory must be performed for the capture zone of each well and within a fixed radius around the well. The procedure used to develop model data from the GIS database is summarized in Figure 4.

McCord Wellfield Wellhead Protection Program

MLGW and the GWI performed a demonstration project funded by the U.S. Environmental Protection Agency (EPA) for the C.M. McCord Wellfield Wellhead Protection Program (6). This wellfield was selected because of multijurisdictional problems that will be encountered during plan implementation. The City of Memphis owns all the wells, but many of them, and all future well lots, are located within the city limits of Bartlett. A wellhead protection plan will have to involve the cooperation of both municipalities. The existing wellfield is shown in Figure 5.

Tennessee wellhead protection regulations require the delineation of two zones of protection for a city the size of Memphis: a 750-foot radius around the wellhead and a 10-year capture zone for the well. The 10-year capture zone (called the Zone 2 area) was delineated using two flow models and information obtained from the GWI database. Results were imported into the GIS database and compared with existing information. The Zone 1 area was delineated by buffering each well point in the coverage by the appropriate radius. Each well location was verified using a Trimble GPS unit and is accurate to within 2 meters. A contaminant source inventory was performed using the coverages developed by address matching. The primary potential sources of contamination in this area are USTs. A windshield survey located other potential sources, such as dry cleaners. These locations were entered into the database also by using

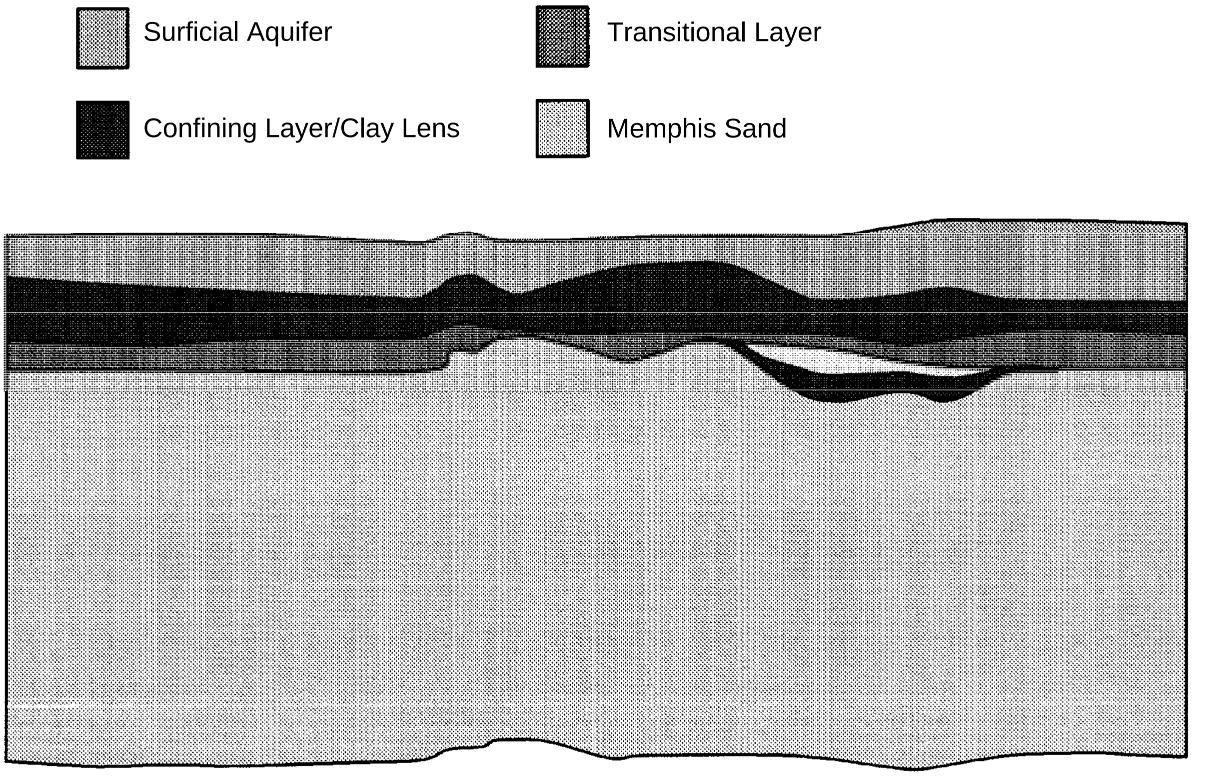


Figure 3. Two-dimensional profile constructed from surface models.

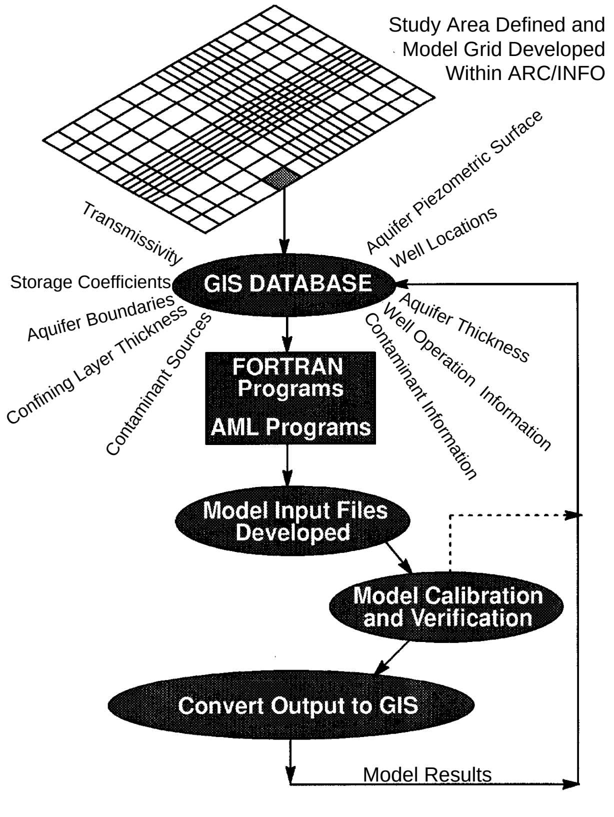


Figure 4. Procedure for integrating GIS and flow model.

the address-matching capabilities of ARC/INFO. The Zone 2 areas for present wells and future wells, along with the potential sources of contamination, are shown in Figure 6.

Davis Wellfield Study

An example of GIS use in ground-water flow modeling is the study of the Justin J. Davis Wellfield. The Davis Wellfield is one of 10 producing fields operated by MLGW. It is located in the southwestern corner of Shelby County and consists of 14 wells, as shown in Figure 1. Production at the Davis Wellfield began in 1971, and an estimated 13 million gallons per day are currently withdrawn from the Memphis Sand aquifer. Since 1972, MLGW has collected water quality data from the wells at the Davis Wellfield, including values for alkalinity, hardness, chloride, sulfate, iron, and barium. Water quality parameters of alkalinity, hardness, sulfate, and barium have significantly increased in the past 10 years.[1] A possible explanation for the change in water quality is water leakage from the upper aquifer through the confining unit to the Memphis Sand aquifer.[1] The water chemistry from the two aquifers is noticeably different. The surficial aquifers generally have a higher total dissolved solids concentration, hardness, and alkalinity than water from the Memphis Sand.[1]

[1] Webb, J. 1992. Memphis Light, Gas, and Water Division. Personal interview.

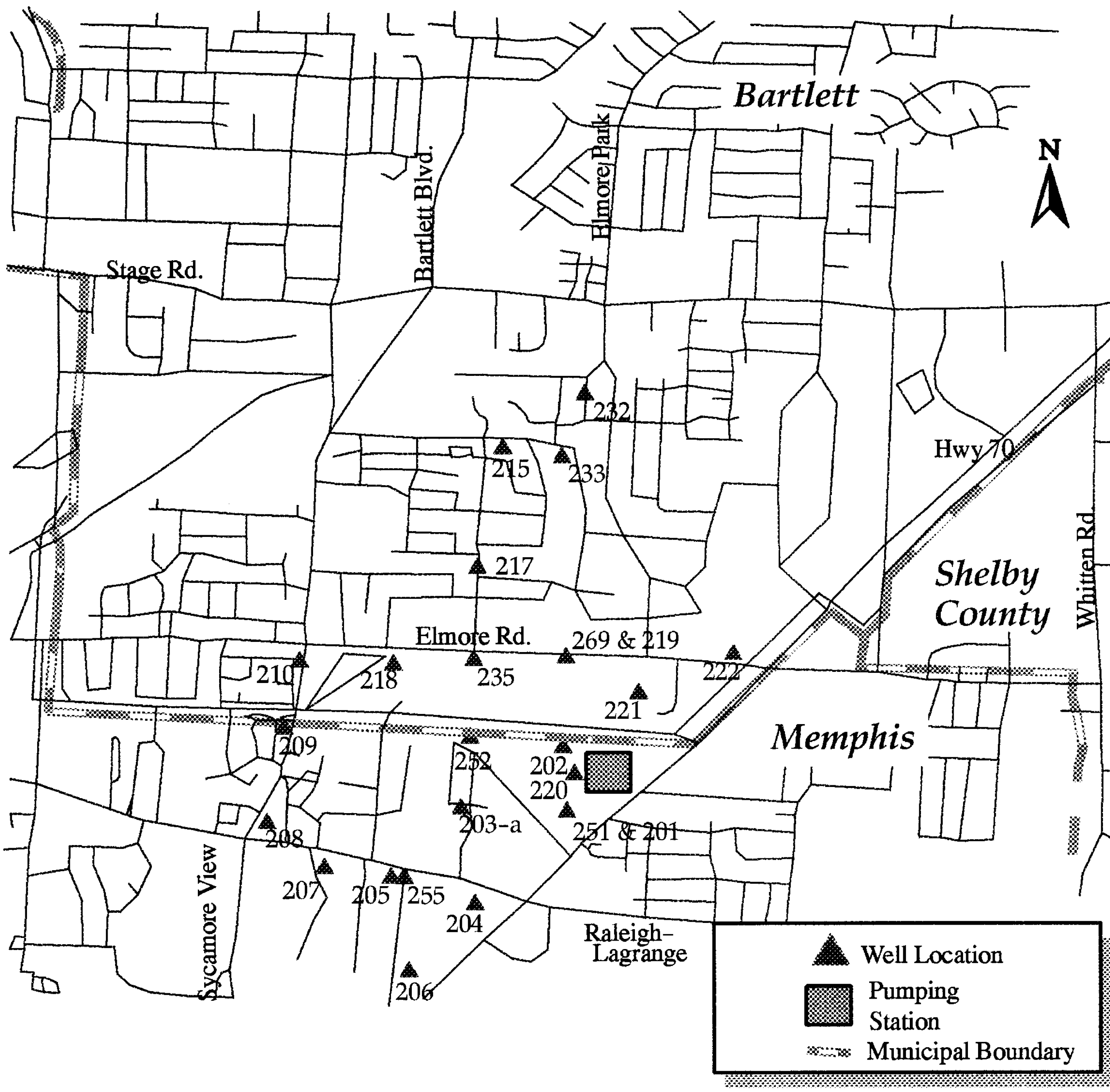


Figure 5. Existing McCord Wellfield.

MLGW, the GWI, and the USGS participated in a joint investigation of the wellfield to determine why the water quality changes are occurring. In the spring of 1992, 12 monitoring wells were drilled into the surficial aquifer near the production wells, as shown in Figure 7.

Geophysical logging and split-spoon sampling revealed an absence of the confining layer at one monitoring well, GWI-3. All other wells penetrated various thicknesses of clay. This “window” in the confining layer suggests that the water quality changes could be due to leakage from the surficial aquifer to the Memphis Sand. The logs for these monitoring wells were combined with an existing geophysical log coverage. The extent of the confining layer window was estimated using GIS surface modeling capabilities, as shown in Figure 8.

Two-dimensional profiles were created to further show the extent of the confining layer window. The locations of the profiles in relation to various surface features are shown in Figure 9. Profiles 1 and 2 were taken across the river bluff, and Profiles 2 and 3 were taken across the window. The profiles are shown in Figure 10.

Many important features of this area’s geology can be inferred by looking at the profiles. A connection of the alluvial and fluvial aquifers is shown in Profile 2. Elsewhere along the bluff, the connection of the two aquifers is less prominent, as shown in Profile 1. The connection of the two aquifers in Profile 2 may be the cause of a peculiar mounding effect in the water table of the alluvial aquifer in that area. The thinning of the top soil in Profiles 1 and 3 may indicate a local recharge area for the

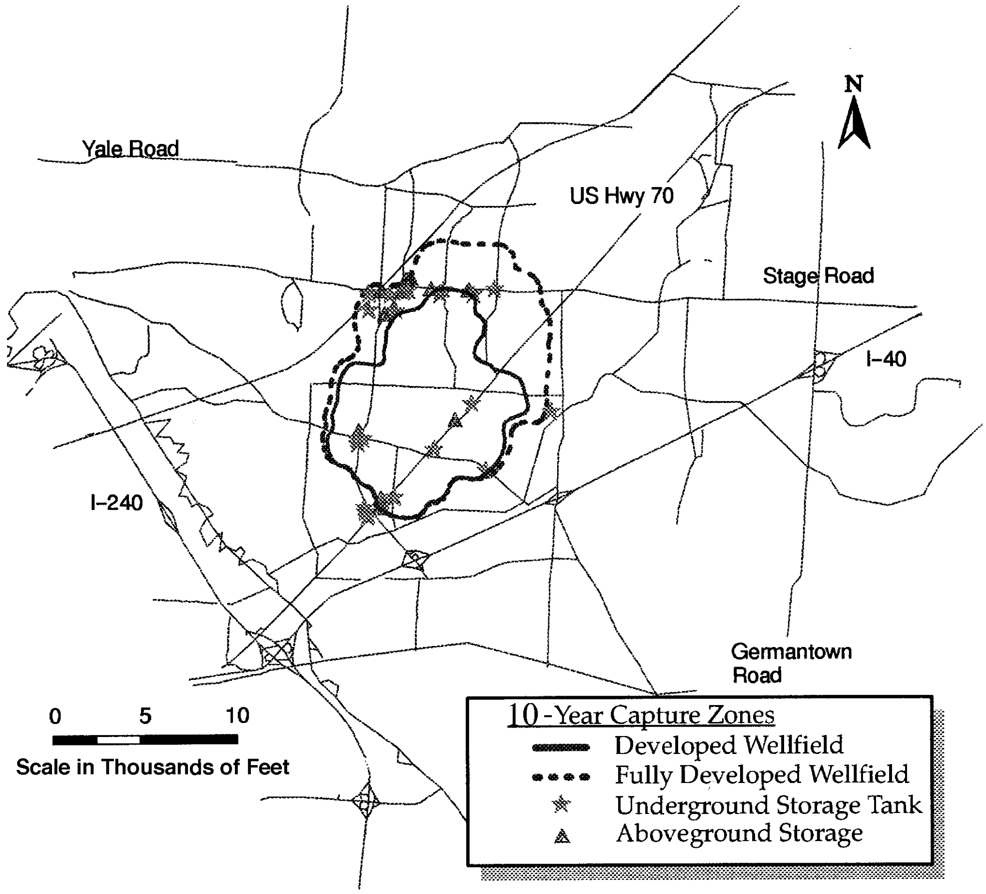


Figure 6. Underground storage tanks and aboveground storage locations.

alluvial aquifer. Profiles 2 and 3 show the confining layer window that suggests a connection between the surficial aquifer and the Memphis Sand.

Following the convention of the USGS, a “window” is defined as any area where the aggregate clay thickness is less than 10 feet (8). A surface of the thickness of the confining layer was generated from the geophysical log coverage. The surface model was converted to a contour line coverage on a 5-foot interval. Using the ARC-EDIT module of ARC/INFO, the contour line coverage was converted to a polygon coverage. The area bounded by the 10-foot contour of the surface model was calculated to be 840,000 square feet (about 19 acres). The area was calculated by adding the areas between the 10- and 5-foot contours and the area within the 5-foot contour.

A flow model of the area was developed based on the hydrogeologic data contained in the database. A steady state model was calibrated to hydraulic conditions recorded during fall 1992 by the USGS and the GWI; the root mean square (RMS) error for this model was 1.76 feet. A second steady state model was developed to simulate conditions recorded during spring 1993 (a period of high water levels in area lakes). The RMS error for this simulation was 5.19 feet. This higher error may indicate that the high water levels in surface water bodies are not realistic for steady state boundary conditions. Realistically, monitoring wells that are relatively far from a surface water body are affected more by average water levels over time rather than relatively short periods of highs and lows.

Using average values of h_1, h_2 (head in upper and lower aquifers), *l* (vertical flow distance), and *VCONT* (a parameter used in MODFLOW to allow for vertical conductance), the estimated flow rate through the window for fall 1992 may be computed as:

$$k = \text{VCONT } 1 = 1.76e^{-3} \times 199.2 = 0.351 \frac{\text{ft}}{\text{day}}$$

$$i = \frac{h_1 - h_2}{l} = \frac{186.5 - 156.1}{199.2} = 0.153$$

$$Q = kAi = 0.351 \times 840{,}000 \times 0.153 = 45{,}111 \frac{\text{ft}^3}{\text{day}}$$

$$45{,}111 \frac{\text{ft}^3}{\text{day}} \times 7.48 \frac{\text{gallons}}{\text{ft}^3} = 337{,}430 \frac{\text{gallons}}{\text{day}} = 0.34 \text{ MGD}$$

The flow rate calculated from average spring 1993 (a period of high water levels in the surficial aquifers) model results was computed as:

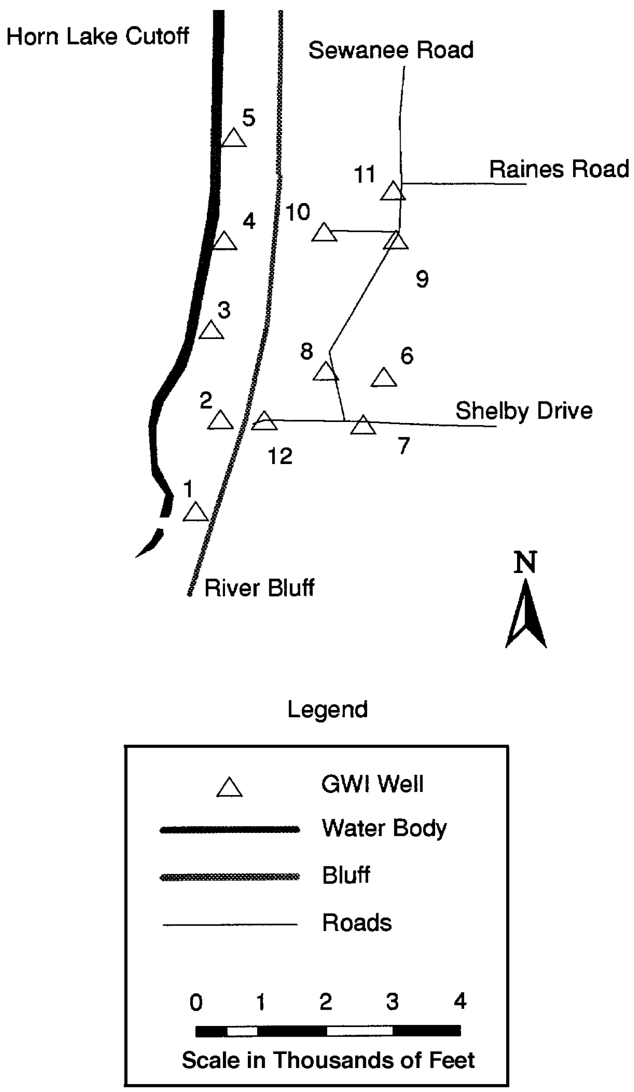


Figure 7. Location of GWI monitor wells.

$$k = \text{VCONT } I = 1.76e^{-3} \times 199.2 = 0.351 \frac{\text{ft}}{\text{day}}$$

$$i = \frac{h_1 - h_2}{l} = \frac{196.4 - 156.8}{199.2} = 0.199$$

$$Q = kAi = 0.351 \times 840{,}000 \times 0.199 = 58{,}673 \frac{\text{ft}^3}{\text{day}}$$

$$58{,}673 \frac{\text{ft}^3}{\text{day}} \times 7.48 \frac{\text{gallons}}{\text{ft}^3} = 438{,}874 \frac{\text{gallons}}{\text{day}} = 0.44 \text{ MGD}$$

Using the GIS-delineated window, an estimated 0.34 to 0.44 million gallons per day flow from the alluvial aquifer to the Memphis Sand aquifer. This variation in the flow rate was due to seasonal variations of water level in the alluvial aquifer.

Since the wellfield pumps approximately 13 million gallons per day, the effect of the window on the entire Davis Wellfield may not be significant. The window lies within the 30-year capture zone of two wells in the field, however.

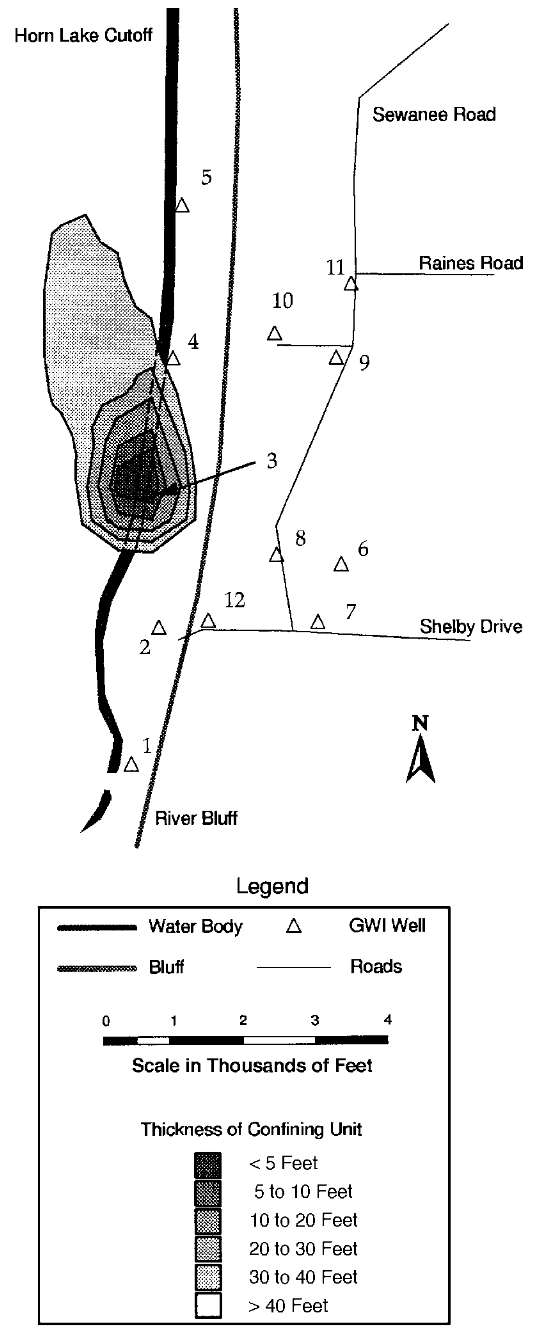


Figure 8. Location and extent of window in confining unit.

The flow rate through the window is approximately 20 percent of the total production of these two wells. Additionally, since the wells would probably not operate simultaneously, the flow rate through the window may account for approximately 40 percent of the flow at either well.

Particle tracking in the Memphis Sand was developed using MODPATH (9) from MODFLOW results. Particles were placed at model well screens and tracked backward for 30 years. The output from MODPATH was read into the GIS database for comparison with other data, as shown in Figure 11. The hole in the confining unit lies

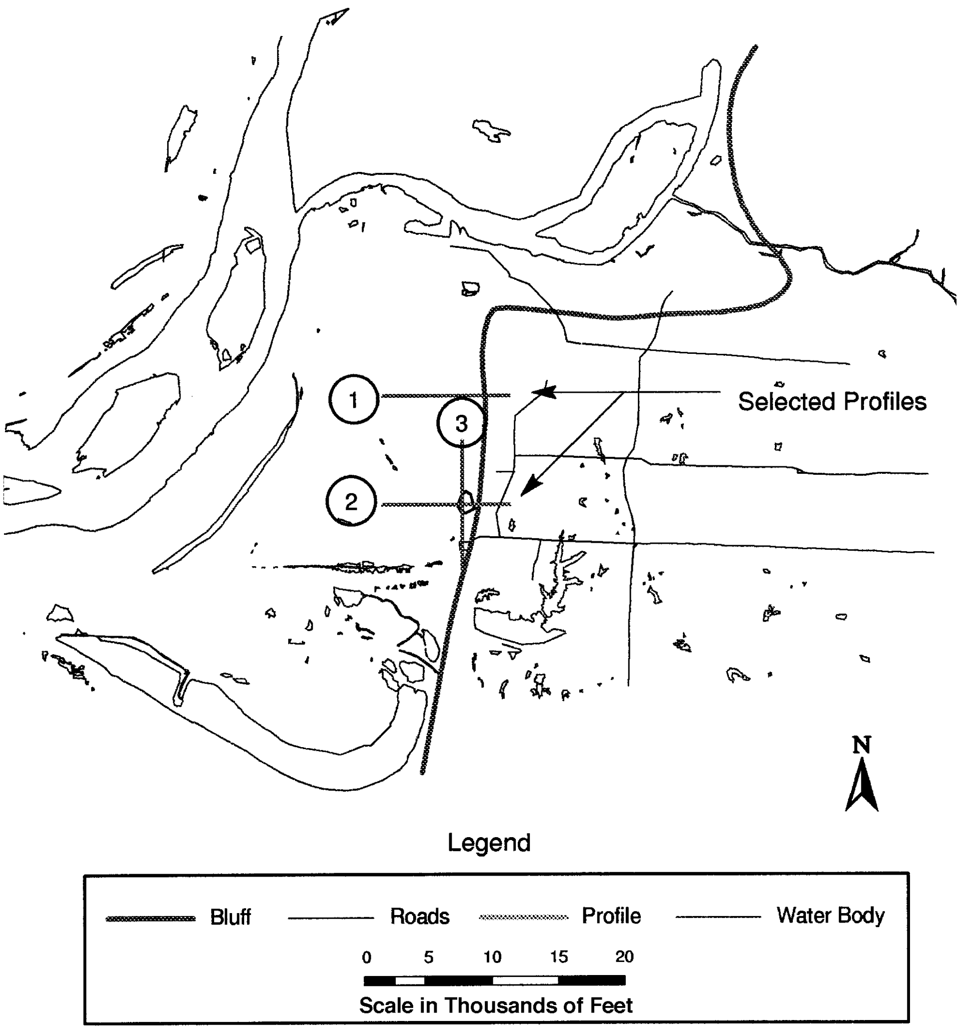


Figure 9. Location of selected profiles.

within the 30-year capture zone of MLGW wells 418 and 419. Historically, 419 was the first well that experienced water quality changes. The water quality is becoming similar to water found in the alluvial aquifer.

A change in the water quality in well 418 is not as immediately noticeable as in well 419. This inconsistency in data may indicate that the window does not extend northward from GWI-3, as the TIN model predicted. To determine which capture zone (418 or 419) encompasses GWI-3, particles were tracked backward for 40 years, as shown in Figure 12. GWI-3 lies on the edge of the capture zone for 419. The flow lines from 418 and 419 move toward the northwest and southwest in the Memphis Sand, pass up through the window, and emerge in the upper aquifer.

Conclusions

The explanation of the Davis Wellfield investigation addressed some limitations of the database. The utility that this hydrogeologic database provides greatly outweighs the disadvantages, however. Without the ability to map and define hydrogeologic features, this project may not have been completed in the allotted time frame or may not have been completed in the same level of detail. GIS greatly enhances the development and evaluation of ground-water flow models.

Specific conclusions that can be drawn from the analysis performed in this project are:

- The delineation of a window in the confining layer using a GIS database is possible.
- Based on the GIS-generated window, an estimated 0.34 to 0.44 million gallons per day flow from the upper aquifer to the Memphis Sand, which may account for as much as 40 percent of the flow at either well 418 or 419.
- The drilling of more monitoring wells north, south, east, and west of GWI-3 may provide for a more accurate delineation of the window.

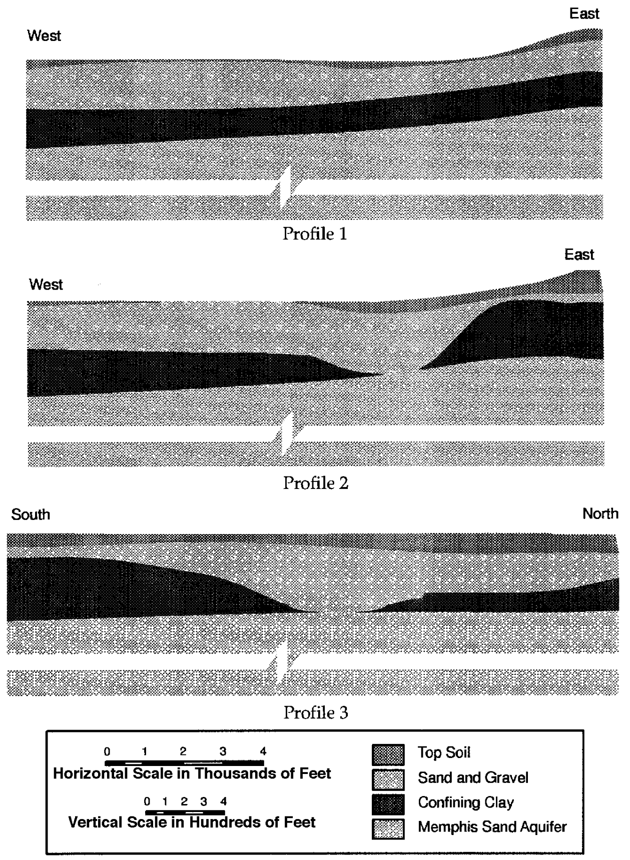


Figure 10. Selected subsurface cross sections in the Davis Wellfield.

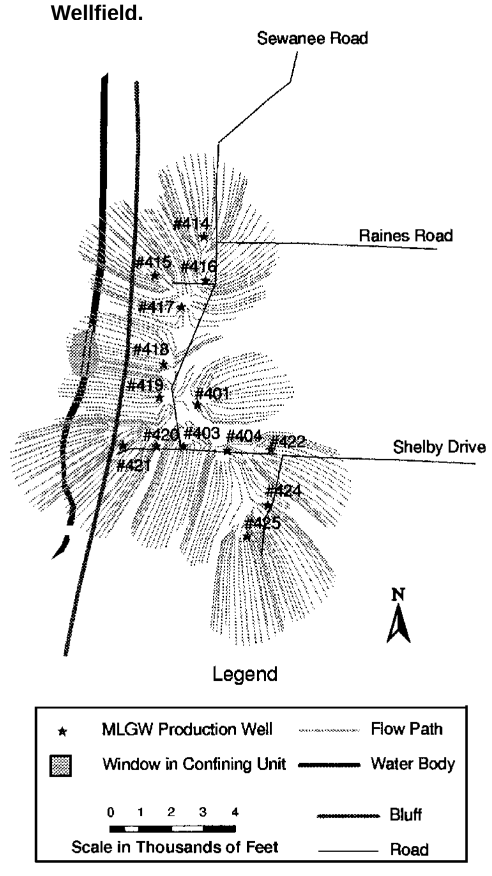


Figure 11. Backward tracking for 30 years.

General conclusions that may be drawn from this discussion are:

- GIS provides a convenient method of viewing flow model input and output.
- A flow model may be developed and evaluated in a relatively short time using GIS.
- GIS provides a convenient means of compiling and managing the information required to develop a wellhead protection program.

Some GIS disadvantages that have been noted are:

- The time required to develop a database and learn to apply the GIS program in a particular situation may be prohibitive.
- GIS-generated results from a limited database may be misleading and should be corroborated with other analysis methods.

Acknowledgments

The authors would like to acknowledge the efforts of the professors and students of the GWI, both present and past, for their contributions to the database and this

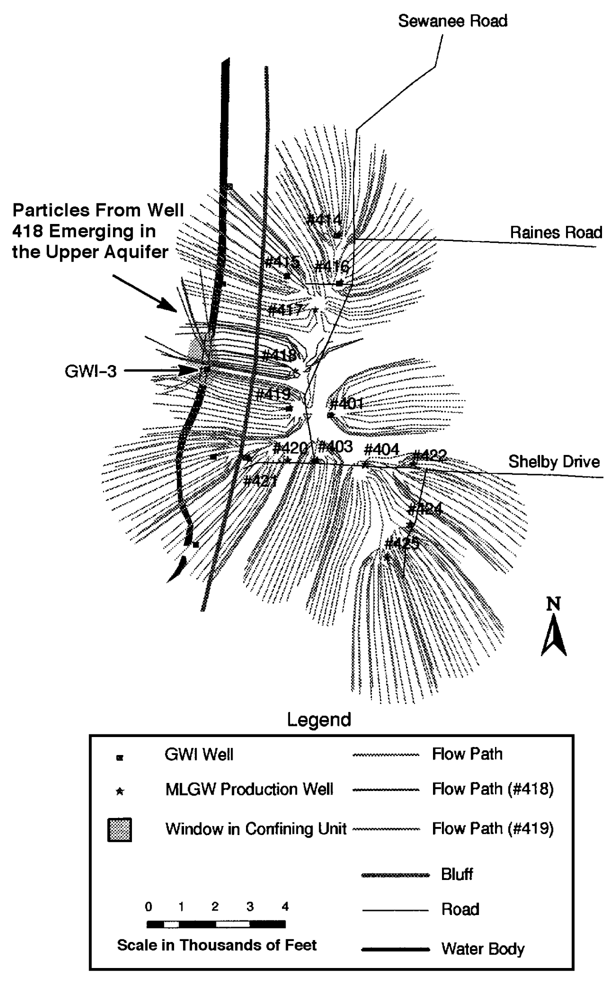


Figure 12. Backward tracking for 40 years.

project, especially: Dr. John W. Smith, Director, and Dr. Charles V. Camp for their patience and guidance in interpreting the model results and hydrogeologic conditions of the area; and David W. Kenley, Brian A. Waldron, and Robert B. Braun for their help in creating the well log database.

References

1. Camp, C.V., J.E. Outlaw, and M.C. Brown. 1994. GIS-based ground-water modeling. Microcomputers in Civil Eng. 9:281-293.
2. Driscoll, F.G. 1986. Groundwater and wells. St. Paul, MN: Johnson Division.
3. Camp, C.V., and M.C. Brown. 1993. A GIS procedure for developing a three-dimensional subsurface profile. J. Computing in Civil Eng. (July).
4. Camp, C.V., and J.E. Outlaw. 1993. Constructing subsurface profiles using GIS. Adv. in Eng. Software. 18:211-218.
5. McDonald, M., and A. Harbaugh. 1988. A modular three-dimensional finite-difference ground-water flow model. Open File Report 83-875. U.S. Geological Survey.
6. Palazolo, P.J., J.W. Smith, J.L. Anderson, and C.V. Camp. 1994. C.M. McCord wellhead protection demonstration grant. Final report. Herff College of Engineering Ground Water Institute, University of Memphis, Memphis, TN.
7. Richardson, G. 1989. A study of potential sources of leakage into the Memphis sand aquifer beneath the Davis well field in Memphis, Tennessee. M.S. thesis. Memphis State University, Memphis, TN.
8. Parks, W.S. 1990. Hydrology and preliminary assessment of the potential for contamination of the Memphis aquifer in the Memphis area, Tennessee. Water Resources Investigation Report 90-4092. U.S. Geological Survey.
9. Pollock, D. 1989. Documentation of computer programs to compute and display pathlines using results from the U.S. Geological Survey modular three-dimensional finite difference ground-water flow model. Open File Report 89-381. U.S. Geological Survey.

Public Participation Geographic Information Systems Applications for Environmental Justice Research and Community Sustainability (Working Paper)

David A. Padgett
Tennessee State University
Nashville, TN

Lynn Pinder, Youth Warriors, Inc.
Baltimore, MD

INTRODUCTION

The primary objective of this study is the development and assessment of a community-based geographic information systems (GIS) method for researching childhood lead-soil exposure and subsequent poisoning. Means to fund and conduct research on childhood lead exposure preventative measures have been addressed under the Residential Lead-Based Paint Hazard Reduction Act (U.S. Congress, 1992) and the Lead Exposure Reduction Act (U.S. Congress, 1993). Among the major action elements listed in the U.S. EPA's 1991 "Strategy for Reducing Lead Exposures" are recommendations for the development of GIS methods to identify and delineate "geographical hot spots" (i.e. highly susceptible areas) (U.S. Environmental Protection Agency, 1991). The U.S. EPA has recommended that GIS be applied in efforts to learn more about the nature and distribution of environmental lead:

> "Continued reanalysis of the data by independent investigators will reveal even more information about the movement of lead in urban environments. These analyses should include…GIS analysis." (U.S. Environmental Protection Agency, 1993)

Several studies utilizing GIS to link lead poisoning sources with specific sensitive population characteristics have been conducted within the past decade (Mielke and Adams, 1989 Huxhold,1991; Guthe et al, 1992; Wartenburg, 1992; North Carolina Department of Environment, Health, and Natural Resources, 1993; Bailey et al, 1994; Brinkmann, 1994; Dakin's 1994; Bocco and Sanchez, 1997; Margai et al, 1997; Bailey, Sargent, and Blake, 1998; Griffith et al, 1998; Lutz et al, 1998). The issue of residential lead contamination has recently become the subject of a major class-action lawsuit (Koff, 1999). While prior research has provided substantial information regarding the spatial nature of environmental lead, there is little to no indication that members of potentially impacted communities have been actively involved in implementing the projects. The application of the public participation GIS method developed in this study brings forth several issues associated with community participation in scientific research. Interest in community-based research is being more frequently raised and discussed at environmental justice-oriented conferences. The general trend among "at-risk" neighborhoods is that they are "fed up" with having their human environment studied by outsiders from academia or the government without being able to fully or even partially participate themselves. The term "guinea pig" has been commonly tossed around during the discourse.

ARGUMENTS FOR A COMMUNITY-BASED PUBLIC PARTICIPATION GEOGRAPHIC INFORMATION SYSTEMS APPROACH

Four key areas have been identified supporting the need for full community involvement in neighborhood level lead-soil GIS mapping research. First, the non-spatial organization of blood-lead screening data may result in inaccurate mapping. Second, mapping the dynamics of long and short-term human migration within the study may prove problematic with the static nature of GIS. Third, the issue of "neighborhood" definition almost certainly requires input by study area residents. Fourth, those same residents may have privacy issues surrounding the release of children's health data, and residential lead contamination levels. Upon learning that they are being used as subjects in research conducted by non-resident investigators, people within study site communities may refuse to cooperate. In a worst case scenario, a community may act to block the release or use of any data collected. Each key issue is discussed in greater detail below.

Non-Spatial Nature of Public Health Data

Norman et al's (1994) research on the differences between rural and urban blood-lead concentrations in North Carolina, focusing primarily upon ethnic, gender and age differences among 20,000 participants, is a recent example of public health data being analyzed spatially. However, the level of analysis is at the county scale without including street addresses, typical public health data format. Earlier research conducted at Leeds, Alabama (U.S. Department of Health and Human Services, 1991) is a classic example of the common omission of critical spatial data from public health investigations. At Leeds, while lead-soil concentrations exceeded 16,000 ppm in some sections of the town, no children participating had blood-lead levels above 25 micrograms per deciliter (ug/dl) (U.S. Department of Health and Human Services, 1991).

However, because the participant's addresses were unavailable, there was no way of knowing whether any of the participants resided in the most contaminated areas. Children's blood-lead levels were derived using parental reporting forms. Missing from the forms were requests for the children's residential addresses. Without street addresses it is virtually impossible to determine the existence of any spatial correlation between lead poisoned children and lead releases. Further, compiling children's residential locations for GIS databases may be problematic as records kept by health departments may list incorrect, inaccurate, or defunct addresses, or no addresses at all. In some cases, the address given by a child's parent or guardian may not be where the child actually resides. Further, even if a correct address is provided, that site may not necessarily be the point of exposure. The spatial inconsistencies associated with blood-lead and lead exposure data may render GIS address geocoding efforts very difficult, if not impossible to do by outside investigators. The primary challenge is capturing the daily movements among residences and public facilities by children under the age of six.

Dynamics of Human Movement Versus Static Data

The static nature of GIS may be limited with respect to the daily migratory patterns of human beings, especially small children. Pin-pointing an exact lead exposure site for a poisoned child may prove difficult considering that a child may move from his/her home to school, back home, and then to a relative or friend's home within a given 24 hour period. Temporal GIS problems and issues are discussed by Langram (1992), Monmonier (1991), and Peuquet (1994). Adams (1995) discusses the difficulties inherent in dealing with people as "point entities." Daily, short distance human migration cannot be captured on a non-real-time GIS image. In a static GIS application, lead-poisoned children would have to be treated essentially as a "dynamic point file." Low-income children, who are most susceptible to lead poisoning, tend to move from one home to another more frequently than those from middle and high income homes, which adds to address matching accuracy problems. With humans being non-stationary objects, tracking them as they move through space during a particular time period may complicate any efforts to draw sound conclusions about conditions within that space which may be impacting the humans present. While recent research indicates that methods for accurate parental estimation of children's soil ingestion rates have yet to be perfected (Calabrese, Stanek, and Barnes, 1997), the community knowledge base remains the best source of information for potential childhood lead-soil exposure incidences.

Neighborhood Delineation

Selecting the spatial limits of the target population in lead-soil exposure studies comes with a variety of challenges. Martin (1991) and Hunter (1979) discuss the difficulty of classifying residential neighborhoods. Coombes et al (1993) discuss the pertinent ingredients which make up a "community," and how the importance of those ingredients may be overlooked during GIS database construction. In most cases, census tracts, school districts, and other readily-available data polygons are not sufficient for targeting specific groups. Neighborhood and/or community lines do not necessarily coincide with municipal boundaries. Tosta (1993) writes "If the bits and bytes on our screens have become our knowledge about a place, and if we are using that knowledge to make decisions . . . and we've never been in that space, then something is terribly wrong." On-the-ground analyses of study areas are necessary for the creation of culturally, socioeconomically, and racially homogenous polygons. In order to delineate the spatial extents of a target population, it is necessary to peruse the study area. During inter-census periods, many communities may undergo significant changes due to gentrification, filtering, and other migratory phenomena. Groundtruthing by study area residents is most effective in establishing the daily movement patterns of children. This step is best conducted during the summer months when children have maximum opportunity for daily contact with playground soil. The observers, practicing "street geography," are able to identify specific exposure points and then spatially match those exposure points with locations of lead poisoned children.

Privacy Issues

Because public health data contains information specific to private individuals, obtaining such data in useful form may prove difficult. Parents of children with elevated blood lead levels will most likely wish to remain anonymous. Privacy issues may interfere with sampling efforts as parents of children and landlords may refuse access to private property for the collection of soil samples. The omission of children's addresses in the aforementioned Leeds, Alabama study may have been an effort to protect the children's privacy. The potential for lead poisoning to result in learning disabilities and lowered I.Q. scores may be reasons for parents wishing for anonymity. Childhood lead poisoning has also been linked to delinquent behavior (Needleman et al, 1996). Such revelations will most likely further influence families' insistence to keep such medical information private. Children with blood-lead levels above 10 ug/dl are now considered at risk (National Research Council, 1993). With indications that the consequences and prevalence of childhood lead poisoning may be more serious than earlier perceived, many municipalities are now making efforts to compile more concise databases. New databases include addresses of children having elevated blood-lead levels; however, due to their sensitive nature, they are relatively difficult to obtain by outside investigators. Potentially impacted stakeholders have greater rights, if not access, to community children's health data. Therefore, it is in their best interest to have some form of proprietorship over any data being used to develop GIS-supported lead lead-soil mapping.

CHILDHOOD LEAD POISONING HEALTH RISKS ASSOCIATED WITH LEAD-SOIL EXPOSURE

Research conducted by the U.S. Agency for Toxic Substances and Disease Registry (ATSDR) and others have indicated that childhood lead poisoning, thought to have been eradicated through the legislated ban on leaded gasoline and lead-based paint during the late 1970s, remains to be a pervasive public health problem, especially for poor urban children (Chadzynski, 1980; Lin-Fu, 1980; Mushak and Crocetti, 1988; Breen and Stroup, 1992; Brewer et al, 1992; Fernandez et al, 1992; National Research Council, 1993; Norman et al, 1994; U.S. General Accounting Office, 1998; Mielke, 1999). The potential neurological, behavioral, and renal damage associated with lead poisoning is well documented in public health related research (Needleman et al, 1990; U.S. Department of Health & Human Services, 1991a; Needleman, 1992; Florini and Silbergeld, 1993; Mushak, 1993; Bernard et al, 1995; Needleman et al, 1996). Youths under the age of seven are identified as being most susceptible to the harmful effects of lead ingestion including nerve damage, brain swelling, and lowered IQ scores (Wriggins, 1997).

Lead contaminated soil is becoming increasingly recognized as a significant exposure pathway for children. Among the most infamous childhood lead-soil poisoning cases are those of West Dallas, Texas (Bullard, 1994) and Leadville, Colorado (Colorado Department of Health, 1990). At Leadville, a study of residential surface soils found lead-soil levels higher than 1000 ppm with 80 percent of homes with concentrations over 500 ppm. Consequently, it was discovered that 41 percent of the children in Leadville had blood-lead concentrations higher than 10 micrograms lead/decilter (ug/dl) blood. A multi-city study conducted by Mielke (1991) concluded that because soil is a major reservoir for lead, there is a strong association between lead-soil levels and childhood blood-lead elevation. White (1992), in an analysis of high lead-soil levels associated with an abandoned landfill in Orleans Parish, Louisiana, found that nearly 70 percent of the children in the area had elevated blood lead concentrations. Other research conducted in Australia (Young, Bryant, and Winchester, 1992), Michigan (Francek, 1992), Minnesota (Mielke et al, 1984), and Louisiana (Mielke, 1993) support the above findings linking lead-soil contamination with childhood lead exposure hazard. Research conducted in Venezuela indicates that proximity of residential backyards to roadways may establish a link between lead-soil levels and traffic volumes, adding yet another parameter to lead-soil pathway analysis (Newsome, Aranguren, and Brinkmann, 1997).

In 1994, the U.S. EPA released guidelines for lead-soil hazard assessment with lead-soil concentrations exceeding 400 ppm being considered potentially hazardous to human health (U.S. Environmental Protection Agency, 1994a). However, as Page and Chang (1993) point out, "There is no universally accepted safe level for lead in soil." The Centers for Disease Control (CDC) recommends that cleanup should be considered when residential lead-soil is between 500 and 1000 ppm (Clickner, Albright, and Weitz, 1995). The CDC recommendation may be complicated by the fact that lead may exist naturally in the upper horizons of all soils at concentrations as high as 500 ppm (Zimdahl and Hassett, 1977). A study conducted during the 1980s at Cincinnati, Ohio estimated that childhood blood-lead levels increase at a rate of 6.2 ug/dl blood for each 1000 ppm increase in lead-soil (Burgoon, Rust, and Hogan, 1995). A U.S. EPA (1993a) study indicated that for each 100 ppm lead in soil above 500 ppm, there is a one to two ug/dl increase in children's blood-lead levels. The above indicates that a standard quantitative relationship between lead-soil and blood-lead levels has yet to be established. Other research indicates that outdoor lead-soil contamination may contribute significantly to indoor lead-dust contamination (Sayre, 1981; Charney, 1982;Calabrese and Stanek, 1992; Clickner, Albright and Weitz, 1995; Lanphear et al, 1995; Rust and Hogan, 1995). Duggan and Inskip (1985) compiled an extensive list of childhood lead-dust exposure studies and estimated the mean lead-dust to blood-lead relationship at about five ug/dl blood for every 1000 ppm lead in indoor dust.

In order to measure the effectiveness of lead-soil abatement efforts, the U.S. EPA (1993a) completed a major three-city study focused primarily upon the relationship between lead contaminated soil and childhood blood-lead levels. Contaminated soil was removed from urban communities having above average childhood blood-lead levels. Following abatement local children's blood-lead concentration was analyzed in order to determine whether removal of the soil would result in any reduction. The study concluded that a decrease of 1,000 ppm lead in soil results in a reduction of approximately one ug/dl blood. With soil being the apparent starting point of the lead exposure pathway, it is necessary that mitigation and/or abatement of contaminated soil be of primary concern in efforts to lessen childhood lead poisoning hazards both inside and outside of the home. The U.S. EPA study includes recommendations for mitigating hazards from contaminated indoor dust among its 1994 guidelines (1994a).

ENVIRONMENTAL RISK ASSESSMENT AND ENVIRONMENTAL JUSTICE USING GEOGRAPHIC INFORMATION SYSTEMS

Geographic information systems is applied in this project primarily as a tool for environmental risk assessment, which is generally considered to be a pro-active approach to solving environmental problems. Environmental risk assessment involves the analysis of physical and population data in order to estimate the potential magnitude of hazard associated with various environmental problems. The U.S. Environmental Protection Agency (EPA) has established that for its risk assessment data needs "Geographic Information System technology could be used to identify high-risk populations…the most exposed and highly susceptible populations in each region would be targeted for enforcement actions" (U.S. Environmental Protection Agency, 1990). Currently, GIS-based models, graphics, and statistical packages are being used and developed in environmental risk assessments by environmental scientists at the U.S. EPA (U.S. Environmental Protection Agency, 1990 and 1992; Stockwell et al, 1993).

In 1993 the "Environmental Justice Act" was introduced to Congress by Representative John Lewis (U.S. Congress, 1993). Among the provisions of the Act are requirements for the U.S. EPA and U.S. Department of Health and Human Services (HHS) to establish the geographical units used for determining Environmental High Impact Areas (EHIAs) which are the 100 geographical areas found to have the highest volumes of toxic chemical releases. Padgett and Robinson (1999) have developed a GIS method for delineating EHIAs. The Agency for Toxic Substances and Disease Registry (ATSDR) has determined GIS to be the "best methodology for identifying potentially impacted minority populations" (Harris and Williams, 1992). The spatial, socioeconomic, and demographic nature of neighborhood level lead-soil contamination analysis offers significant opportunities for academic researchers and people of color and/or low-income communities to cooperatively employ GIS.

A SIX-STEP COMMUNITY-BASED PUBLIC PARTICIPATION GEOGRAPHIC INFORMATION SYSTEMS APPROACH

For this study, a six-step public participation GIS method has been developed specifically for lead-soil research. The steps are to be considered a guide for future research and need not be followed to the letter. Issues and concerns central to community action agendas should take precedence in shaping the exact procedures followed.

Step One: Preliminary Groundtruthing

The assessment should begin by determining the extents of the community site, which must be conducted on-the-ground. The goal is to delineate, to the best extent possible, contiguous demographically, socioeconomically, and culturally homogenous communities. At this point it is prudent to turn to community residents for greater insight into community characteristics. Another critical task to be conducted here is the observation of children at play. Children's daily movement patterns are most likely best assessed via first-hand community sources. An "outsider's" simple observations may not be sufficient; for projects covering large areas with diverse socioeconomic dynamics, an "insider's" viewpoint may be required to ensure an acceptable level of accuracy. Knowledge indigenous to the community population most likely provides the best indication of where children six months to six years old may frequently be exposed to lead in play area soils.

Step Two: Community Soil Sampling and Analysis

Following the delineation of the community site(s) and the play areas therein, samples must be collected from play area soils. The sampling and analysis methods applied will vary from case to case. The amount of resources available to the researcher will determine the numbers of samples he/she will be able to analyze. Costs for having soil samples analyzed may range from $45.00 to over $80.00 per sample. Field assessments requiring several hundred samples could obviously and easily become very expensive. For this study, recently available field test kits for lead in soil, sensitive to 400 ppm, will be used (Hybrivet Systems, 1996). Soil samples will be collected by neighborhood residents on a voluntary basis. Perusal of the aforementioned research on GIS applications in lead-soil research generally indicates that, for community-level studies, one to two samples per census tract is suitable for accurately establishing the nature of neighborhood lead-soil concentration levels.

For this project, 400 ppm lead in soil will be the target concentration in agreement with U.S. EPA sampling protocol (U.S. EPA, 1994); however, it is not a permanent value and should not be treated as a constant. Whatever lead-soil level is most appropriate for an area under investigation is the one which should be utilized. For example, according to the most recent EPA assertions, residential lead-soil concentrations exceeding 400 ppm are considered potentially hazardous to human health (1994); however, in some cases, 1,000 ppm may be considered an appropriate benchmark. It is in the primary interest of potentially impacted communities to locate areas where concentrations exceed this standard. Therefore, in this study with qualitative testing, it is not necessary to know exact lead concentrations above 400 ppm.

Step Three: Location of Sample Sites Exceeding Target Levels

The locations of play areas exceeding target levels should be noted either by street address or through some alternative grid reference system. In this study mapping contaminated sites identified by street addresses will be done voluntarily. Names of participants will not be included on survey forms in order to protect the privacy of community residents. A sufficient number of sampling points will be noted to produce a point attribute file of potential problem sites within the selected community.

Step Four: Analysis of Community Childhood Blood-Lead Data

Blood-lead concentrations for children residing within the community site(s) should be obtained and analyzed. Due to privacy concerns, this step may require a lengthy negotiation period. If this procedure is followed by a public agency or otherwise with free access to blood-lead data, attaining the information should be able to be done in relatively little time. The currently accepted benchmark for blood-lead potentially being a health hazard in children is 10 ug/dl. The numbers of children exhibiting elevated blood-lead levels residing in close proximity to potential hot spots should be noted, as should their respective blood-lead concentrations in ug/dl. Methods used to analyze blood-lead data may vary; however, the primary goal should be to determine whether above-average blood-lead levels are exhibited by relatively high percentages of children within the community site(s). Populations with more than five to ten percent of children having elevated blood-lead should be considered above the norm (Society for Environmental Geochemistry and Health, 1993). For this study, the lead community organization will be charged with determining local childhood blood-lead levels from existing public health data. No blood samples will be collected; however, parents may be advised to have their children tested at residences where lead-soil levels above 400ppm are found.

Step Five: Development of Geographic Information Systems Point Attribute Files

In order to establish a spatial relationship between children and sites of potential lead-soil exposure, home addresses where samples are collected will be submitted confidentially and voluntarily. Residents who have received training in GIS point file geocoding processes and will collect the data from the volunteers and then build a point layer of sampling locations. Samples found to have at least 400ppm lead in soil will be noted as potential hot spots. The final "hot spot" maps will be maintained within the community, most likely by the lead community environmental justice organization. The maps will not be released to the public outside of the community without some form of consent from the voluntary participants.

Step Six: Community Education and Dissemination of Results

If the project results warrant it, potentially impacted volunteers and their children will be notified of the existence and location of lead contaminated soils. Ideally, community residents with proper access will be able to view the maps on-line via the internet at local public libraries or other institutions. Hardcopy maps showing hot spot zones will also be made available to qualified persons. If outside consultation is deemed necessary, and there is a general agreement among the project participants and volunteers, data and graphics will be made available to disciplinary experts, government decision-makers and others who may be better able to analyze and mitigate significant incidences of lead-soil contamination.

CASE STUDY ASSESSMENT METHOD FOR THE SIX-STEP COMMUNITY-BASED RESEARCH APPROACH

An instrument for measuring the effectiveness of the public participation GIS project has been derived from Bullard's environmental conflict resolution assessment framework (Bullard, 1994), and Barndt's method for evaluating public participation GIS programs (Barndt, 1998) (Table 1). The qualitative criteria included at Table 1 are intended for the purpose of determining whether the goals of the project have been attained. However, it should be understood that these criteria are neither those of the community participants nor the lead environmental justice group. The standards included in the instrument are parameters set forth for academic evaluation only. The involved parties themselves are to independently determine their gains and losses from the experience.

IMPLEMENTATION OF THE COMMUNITY-BASED APPROACH: BALTIMORE, MARYLAND: JULY 1998-JUNE 1999

The lead community organization at the Baltimore, Maryland site is Youth Warriors, Inc., an environmental justice organization founded in 1996, the primary purpose of which is to work with local communities in Baltimore's inner-city to educate and mobilize young African Americans around urban environmental education, training, leadership development, and organizing skills.

The group works to "encourage our young people to be stakeholders in their communities -- loving, respecting, and valuing the worth of themselves and their neighborhoods." Youth Warriors has two main campaigns: 1) the Urban Gardens Initiative, a collaborative with local neighborhood associations to identify vacant lots and convert them into flower gardens, and 2) PROJECT LEAD,

Table 1. Case Study Assessment Criteria for Community-Based Public Participation Geographic Information Systems

Assessment Dimensions
Issue Crystallization
Environment
Public Health
Equity and Social Justice
Economic Trade-Off
Type of Leadership Group
Mainstream Environmental
Grassroot Environmental
Social Action
Emergent Coalition
GIS Development and Management Process
Sustainable
Replicable
Efficient
Timely
Immediate
Sophisticated
Results/Outcomes
Appropriate
Actionable
Fits Schedule of Activities and Priorities
Accurate
Insightful
Offers Perspective
Synergistic
Combines Qualitative and Quantitative
Results in Changed Outcomes
Contribution to Broader Community Development Agenda
Integrates the Components of a Working Community Information System
Encourages a More Open Dialog Using the Information Organized
Increases Access to Information and Ensures the Right of Access
Keeps Priority Setting in the Hands of Community
Addresses Process Objectives
Recognizes the Value of Co-Production
Increases the Capacity of Local Community Systems to Use the Technology
Integrates into a Broader Community Development Process

a lead poisoning prevention education program. In April 1998, Youth Warriors was awarded a Conservation Technology Support Program (CTSP) Grant (www.ctsp.org) for GIS support which includes the components listed at Table 2.

Table 2. Youth Warriors' Conservation Technology Support Program Grant

==
Award Components
==

Hardware: (Received - July 1998)
Processor (HP VL 266 64 MB RAM 4.3. GB HDD 24xCD Win95)
Monitor (HP Ultra VGA 17" Display)
Printer (HP Deskjet 1120Cse Printer)
CD-Writer (HP SureStore 7200I Internal IDE with software)

Software:
ArcView 3.0 Windows 95
ArcPress Extension

Training:
Five days ArcView GIS training on a space-available basis. (Training completed by On-Site Coordinator , November 1998)
Two days ArcView 3.1 GIS Basic Conservation Training. (Training completed by GIS Team Leader, July 1998)

The GIS hardware and software is being used in support of PROJECT LEAD, a lead poisoning prevention education campaign through which teenagers from the local community are trained in U.S. Environmental Protection Agency (EPA) approved sampling methods for indoor lead dust content at residents' homes. Selected sampling areas are considered to have high risk for lead exposure due to their socioeconomic and demographic characteristics. There are currently approximately 60 participating households. If samples show high levels of lead dust, residents are (1) informed immediately and told the implications of test results (2) sent a Help Packet listing available resources in Baltimore City, and (3) referred by staff to appropriate Baltimore City agencies for further assistance.

With outdoor soils being the primary source of indoor dust, the scope of PROJECT LEAD is being expanded to include sampling of residential outdoor soils. Spatial analysis of sample test results are being conducted in order to determine the extent and nature of lead soil/dust contamination throughout the community site. ArcView 3.1 GIS is being used to map inner-city residential indoor dust/outdoor soil lead concentrations. Using GIS to support the on-going PROJECT LEAD project is providing an inner-city community with an assessment of local lead exposure risks in an easily understandable graphic format. The community-driven method developed here is applicable at other urban areas. Student participants and organizers in Youth Warriors have gained invaluable skills and training in GIS technology, environmental assessment, database management, and scientific field methods. Student trainees will use Microsoft Excel to build a database of the participants in PROJECT LEAD. Pertinent attribute information will include: residential street addresses, census tracts, parts per million (ppm) lead in soil, household lead dust concentration (ppm), and ages of children present (voluntary). Data is being obtained through surveys and field sampling.

Geographic information systems training adds a community sustainability component to the campaign. Training in ArcView 3.1 GIS and its extensions is enabling Youth Warriors participants to develop skills very attractive to employers, or for entrepreneurial pursuits. The students and young people (ages 13-30) are increasing their awareness about their urban environment, and also gaining practical skills which will enable them access to resources that they may use for long-term community restoration. In May 1999, several participants completed an eight-hour Lead Poisoning Prevention Training course held at the National Safety Council Environmental Health Center.

PROJECT LEAD's implementation plan is built around three components: public education, community service, and grassroots organizing. Interwoven throughout each of these components is the main purpose of providing opportunities for predominantly African Americans from low-income urban communities to participate in high quality environmental training projects that foster environmental education, cultural enrichment, leadership development, and civic responsibility.
It is building community capacity to identify local environmental justice problems and enhance stakeholder understanding of environmental and public health information systems. Pertinent goals include the education of 150 young people in lead poisoning’s dangers and lead sampling techniques, as well as training community-based organizations in GIS technology.

SIX-STEP APPROACH AS APPLIED TO PROJECT LEAD: JULY 1998-JUNE 1999

1. July-September 1998: Training on GIS applications and lead sampling for GIS Team Leader, On-Site Coordinator Assistant, and Youth Warriors participants
2. August-September 1998: Delineation of study area with assemblage of PROJECT LEAD residential sampling sites -- approximately 60 -- and demographic/socioeconomic attribute database construction.
3. October 1998-January 1999: Lead-soil and indoor dust sampling at PROJECT LEAD residential sites and input of results into point file and attribute database format.
4. January-April 1999: Import of lead sampling results point file into ArcView 3.1 GIS for sample site address geocoding for GIS point file map development.
5. February-April 1999: Development of initial community lead exposure hazard maps using ArcView 3.1 GIS.
6. May-June 1999: Development of initial community lead hazard statistical graphics using Arcview 3.1 GIS.
7. May-June 1999: Development of community lead hazard maps and other graphics using ArcPress for local distribution and target group education.

It should be noted that in keeping with Youth Warrior's agenda, the "six-step approach" has been adjusted to seven steps. At the time of this working paper, the project had reached step three above, approximately seven months behind schedule due to a variety of unforeseen and unavoidable circumstances. Work continues at the study area shown at Figure 1. The neighborhood generally encompasses Baltimore City census block group1605-5. The population of 1,068 in the one-tenth mile area is 99.4 percent African American with approximately 10 percent being under six years of age (U.S. Department of Commerce, 1997). With a median household income of $23,750 and only 7.7 percent of the population in poverty, the community is not economically typical of those with high risk for lead exposure. However, the median year built for homes is 1941, and 221 of the 471 housing units were built prior to 1940. Soil sampling, GIS mapping, and assessment should be completed by October, 1999.

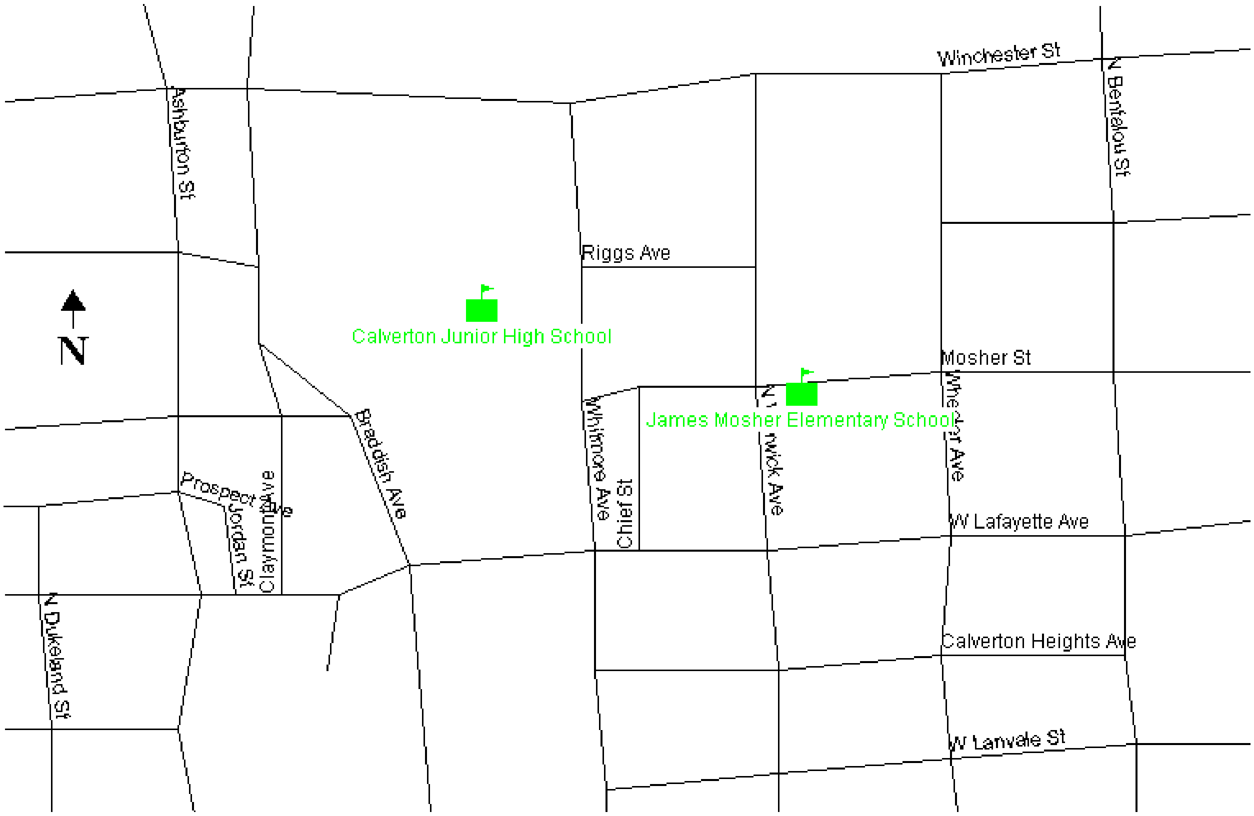


Figure 1. PROJECT LEAD study area: Baltimore, Maryland (U.S. Dept. of Commerce, 1997)

REFERENCES

Adams, P.C. 1995. "A Reconsideration of Personal Boundaries in Space-Time," Annals of the Association of American Geographers, 85(2):267-285.

Bailey, A.J., Sargent, J.D., and Blake, M.K. 1998. A Tale of Two Counties: Childhood Lead Poisoning, Industrialization, and Abatement in New England. In, Economic Geography. Special Issue for the 1998 Annual Meeting of the Association of American Geographers. Edited by Susan Hanson and Richard Peet. Worcester, MA: Clark University, pp. 96-111.

Bailey, A.J., Sargent, J.D., Goodman, D.C., Freeman, J., and Brown, M.J. 1994. "Poisoned Landscapes: The Epidemiology of Environmental Lead Exposure in Massachusetts Children 1990-1991," Social Science and Medicine, 39(6):757-766.

Barndt, M. 1998. "A Model for Evaluating Public Participation GIS Programs," Varenius Specialist Meeting - Empowerment, Marginalization and Public Participation GIS. National Center for Geographic Information and Analysis. Santa Barbara, California,15-17 October, 1998 <http://www.ncgia.ucsb.edu/varenius/ppgis/papers/barndt.html> (accessed April 4, 1999).

Bernard, A.M., Vyskocil, A., Roels, H., Kriz, J., Kodl, M., and Lauwerys, R. 1995. "Renal Effects of Children Living in the Vicinity of a Lead Smelter," Environmental Research, 68:91-95.

Bocco, G., and Sanchez, R. 1997. "Identifying Potential Impact of Lead Contamination Using a Geographic Information System," Environmental Management, 21(1):133-138.

Boulding, J.R. 1994. Description and Sampling of Contaminated Soils. Second Edition. CRC Press: Boca Raton, Florida.

Breen, J.J., and Stroup, C.R. 1992. "Lead Exposure Reduction in Children: The Need for Quality Environmental Laboratories," American Environmental Laboratories, 4:14.

Brewer, R.D., Miller, C.P., Norman, E.H., Freeman, J.I., and Wester, T. 1992. "The Prevention of Childhood Lead Poisoning in North Carolina," North Carolina Medical Journal, 53(4): 149-152.

Brinkmann, R. 1989. Spatial Distribution of Soil Lead Pollution in Milwaukee County, Wisconsin. Doctoral Dissertation. Milwaukee: Department of Geography, University of Wisconsin-Milwaukee.

_. 1994. Lead Pollution in Soils Adjacent to Homes in Tampa, Florida. Unpublished Paper. Tampa, Florida: Department of Geography, University of South Florida.

Bullard, R.D. 1994. Dumping in Dixie: Race, Class, and Environmental Quality. Boulder, CO: Westview Press.

Burgoon, D.A., Rust, S.W., and Hogan, K.A. 1995."Relationships Among Lead Levels in Blood, Dust, and Soil." In, Lead Poisoning: Exposure, Abatement, Regulation. Edited by Joseph J. Breen and Cindy R. Stroup. Boca Raton, Florida: Lewis Publishers, pp. 255-264.

Calabrese, E.J., and Stanek, E.J. 1992. "What Proportion of Household Dust is Derived from Outdoor Soil?" Journal of Soil Contamination, 1:253-263.

Calabrese, E.J., Stanek, E.J., and Barnes, R. 1997. "Soil Ingestion Rates in Children Identified by Parental Observation as Likely High Soil Ingesters," Journal of Soil Contamination, 6:271-279.

Chadzynski, L. 1980. "Finding the Source of Lead." In, Low Level Lead Exposure: The Clinical Implications of Current Research, Edited by Herbert L. Needleman, New York: Raven Press, pp. 239-246.

Charney, E. 1982. "Lead Poisoning in Children: The Case Against Household Lead Dust." In, Lead Absorption in Children. Edited by J. Julian Chisolm, Jr., and David M. O'Hara. Baltimore, Maryland: Urban & Scharzenberg, pp. 79-88.

Clinckner, R.P., Albright, V.A., and Weitz, S. 1995. "The Prevalence of Lead-Based Paint in Housing: Findings From the National Survey." In, Lead Poisoning: Exposure, Abatement, Regulation. Edited by Joseph J. Breen and Cindy R. Stroup. Boca Raton, FL: Lewis Publishers, pp. 3-12.

Colorado Department of Health. 1990. Leadville Metals Exposure Study. Unpublished Report. State of Colorado.

Coombes, M., Openshaw, S., Wong, C., and Raybould, S. 1993. "Community Boundary Definition: A GIS Design Specification," Regional Studies, 3:280-86.

Cox, D.C., Dewalt, G.D., Haugen, M.M., Koyak, R.A., Schmehl, R.L., Balsinger, J., Constant, P.C., Friederich, N.J., Nichols, D.R., Scalera, J.V., and Schwemberger, J.G. 1997. AA Field Test of Lead-Based Paint Testing Technologies,@ American Environmental Laboratory, 9(6):17-19.

Croner, C.M., Pickle, L.W., Wolf, D.R., and White, A.A. 1992. "A GIS Approach to Hypothesis Generation in Epidemiology." In, American Society for Photogrammetry and Remote Sensing/American Congress on Surveying and Mapping/RT 92: Technical Papers: Volume 3 - GIS and Cartography, Washington, DC, August 3-8, 1992, pp. 275-283.

Cutter, S.L. 1993. Living With Risk: The Geography of Environmental Hazards. New York, NY: Routledge, Chapman, and Hall, Inc.

Cutter, S.L., Renwick, H.L., and Renwick, W.H. 1991. Exploitation, Conservation, Preservation: A Geographical Perspective on Natural Resource Use. Second Edition. New York, NY: John Wiley & Sons, Inc.

Dakins, M. 1994. Spatial Patterns and Statistical Models of Childhood Lead Poisoning in Philadelphia, Pennsylvania. NNMEMS Project Report 93-3006. Syracuse, NY: SUNY-College of Environmental Science and Forestry.

Duggan, M.J., and Inskip, M.J. 1985. "Childhood Exposure to Lead in Surface Dust and Soil: A Community Health Problem," Public Health Reviews, 13(1-2):1-54.

Eastman, R.J. 1992. IDRISI Technical Reference. Clark University Graduate School of Geography: Worcester, MA.

Environmental Systems Research Institute. 1995. Understanding GIS: The ARC/INFO Method.

Redlands, CA.

Fernandez, A.M., McElvaine, M.D., Orbach, H.G., and Pulido, A.M. 1992. "An Assessment of Childhood Lead Poisoning: A Demographic Profile of Ten Community Areas in Chicago." In, The National Minority Health Conference: Focus on Environmental Contamination. Edited by Barry L. Johnson, Robert C. Williams, and Cynthia M. Harris. Princeton, New Jersey: Princetion Scientific Publishing Company, pp. 69-75.

Florini, K.L., and Silbergeld, E.K. 1993. "Getting the Lead Out," Issues in Science and Technology, 9(4):33-39.

Francek, M.A. 1992. "Soil Lead Levels in a Small Town Environment: A Case Study from Mt. Pleasant Michigan," Environmental Pollution, 76:251-257.

Golden Software, Inc. 1992. SURFER Reference Manual. Golden, CO.

Griffith, D.A., Doyle, P.G., Wheeler, D.C., and Johnson, D.L. 1998. "A Tale of Two Swaths: Urban Childhood Blood-Lead Levels across Syracuse, New York," Annals of the Association of American Geographers, 88(4):640-665.

Guthe, W.G., Tucker, R.K, Murphy, E.A., England, R., Stevenson, E., and Luckhardt, J.C. 1992. "Reassessment of Lead Exposure in New Jersey Using GIS Technology," Environmental Research, 59:318-325.

Harris, C.H., and Williams, R.C. 1992. "Research Directions: The Public Health Service Looks at Hazards to Minorities," EPA Journal, 18(1): 40-41.

Hunter, A. 1979. "The Urban Neighborhood: Its Analytical and Social Contexts," Urban Affairs Quarterly, 14(3):267-288.

Huxhold, W.E. 1991. An Introduction to Urban Geographic Information Systems. New York, NY: Oxford University Press.

HybriVet Systems, Inc. 1996. LeadCheck Soil: Professional Lead Test Kit. Framingham, Massachusetts.

Koff, S. 1999. "Broad-brush Approach to Suit Over Lead Paint." Cleveland Plain Dealer 13 Jun.: A1+

_________. 1994. Lead Check Instruction Manual. Framingham, Massachusetts.

Langram, G. 1992. Time in Geographic Information Systems. Taylor & Francis: Washington, DC.

Lanphear, B.P., Edmond, M., Jacobs, D.E., Weitzman, M., Tanner, M., Winter, N.L., Yakir, B., and Eberly, S. 1995. "A Side-by-Side Comparison of Dust Collection Methods for Sampling Lead-Contaminated House Dust," Environmental Research, 68:114-123.

Lee, F.G., and Jones-Lee, A. 1992. "Importance of Considering Soil-Lead in Property Site Assessments." Unpublished paper presented at the National Groundwater Association Conference, Orlando, Florida, August 1992.

Lewis, M.W. 1994. "Environmental History Challenges the Myth of a Primordial Eden," Chronicle of Higher Education, 40(35): A56.

Lin-Fu, J.S. 1980. "Lead Poisoning and Undue Lead Exposure in Children: History and Current Status." In, Low Level Lead Exposure: The Clinical Implications of Current Research, Edited by Herbert L. Needleman, New York: Raven Press, pp. 5-17.

Lutz, J., Jorgensen, D., Hall, S., and Julian, J. 1998. "Get the Lead Out: A Regional Approach to Healthcare and Beyond," Geo Info Systems, 8(7):26-30.

Maniates, M.F. 1993. "Geography and Environmental Literacy: Let's Look Before We Leap," Professional Geographer, 45(3):351-354.

Margai, F.L., Walter, S.G., Frazier, J.W., and Brink, R. 1997. "Exploring the Potential Environmental Sources and Associations of Childhood Lead Poisoning," Applied Geographic Studies, 1(4):253-270.

Martin, D. 1991. Geographic Information Systems and Their Socioeconomic Applications. New York, NY: Routledge.

Mielke, H.W., and Adams, J.L. 1989. Environmental Lead Risk in the Twin Cities. Publication No. CURA 89-4. Minneapolis: Center for Urban and Regional Affairs.

Mielke, H.W., Blake, B., Burroughs, S., and Hasinger, N. 1984. "Urban Lead Levels in Minneapolis: The Case of the Hmong Children," Environmental Research, 34, pp. 64-76.

Mielke, H.W. 1991. "Lead in Residential Soils: Background and Preliminary Results," Water, Air, and Soil Pollution, 57-58:111-119.

_____. 1999. "Lead in the Inner Cities," American Scientist, 87(1):<http://www.sigmaxi.org/amsci/articles/99articles/mielke.html>(accessed February 9, 1999).

_. 1993. "Lead Dust Contaminated U.S.A.
Communities: Comparison of Louisiana and Minnesota," Applied Geochemistry, 2:257-261.

Milar, C.R., and Mushak, P. 1982. "Lead Contaminated Household Dust: Hazard, Measurement and Decontamination." In, Lead Absorption in Children: Management, Clinical, and Environmental Aspects, Edited by J.J. Chisholm and D.M. O'Hara. Urban and Schwarzenbert: Baltimore, Maryland, pp. 143-152.

Monmonier, M. 1991. How to Lie With Maps. University of Chicago Press: Chicago, IL.

Mushak, P. 1992. "Defining Lead as the Premiere Environmental Health Issue for Children in America: Criteria and Their Quantitative Application," Environmental Research, 59: 281-309.

Mushak, P., and Crocetti, A.F. 1988. The Nature and Extent of Lead Poisoning in Children in the United States: A Report to Congress. Atlanta, GA: U.S. Department of Health and Human Services.

National Research Council. 1993. Measuring Lead Exposure in Infants, Children, and Other Sensitive Populations. Washington, DC: National Academy Press.

Needleman, H.L., Riess, J.A., Tobin, M.J., Biesecker, G.E., Greenhouse, J.B. 1996. "Bone Lead Levels and Delinquent Behavior," 275(5):363-369.

Needleman, H.L. 1992. "Childhood Lead Poisoning: An Eradicable Disease." In, The National Minority Health Conference: Focus on Environmental Contamination. Edited by Barry L. Johnson, Robert C. Williams, and Cynthia M. Harris. Princeton, New Jersey: Princetion Scientific Publishing Company, pp. 113-116.

Needleman, H.L., Riess, J.A., Tobin, M.J., Biesecker, G.E., Greenhouse, J.B. 1996. "Bone Lead Levels and Delinquent Behavior," Journal of the American Medical Association, 275(5):363-369.

Needleman, H.L., Schell, A., Bellinger, D., Leviton, A., and Allred, E. 1990. "The Long-Term Effects of Exposure to Low Doses of Lead in Childhood: An 11-Year Follow-up Report," New England Journal of Medicine, 322(2):83-88.

Newsome, T., Aranguren, F., and Brinkmann, R. 1997. "Lead Contamination Adjacent to Roadways in Trujillo, Venezuela," The Professional Geographer, 49(3):331-341.

Norman, E.H., Bordley, W.C., Hertz-Picciotto, I., and Newton, D.A. 1994. "Rural-Urban Blood Lead Differences in North Carolina Children," Pediatrics, 94(1):59-64.

North Carolina Department of Environment, Health, and Natural Resources, State Center for Health and Environmental Statistics. 1993. Lead Database Evaluation and GIS Modelling Project. PR-242592. Raleigh, NC.

Padgett, D.A., and Robinson, P.L. 1999. A Geographic Information Systems Method for Environmental High Impact Areas (EHIAs) Delineation for Environmental Justice Research. In, Papers and Proceedings of the Applied Geography Conferences, Volume 22, Edited by F. Andrew Schoolmaster, Denton, Texas: Applied Geography Conferences, Inc. (forthcoming).

Padgett, D.A. 1994. Polygon Design Problems and Improvement Techniques at the Human/Physical Interface: The Case of Lead (Pb) Contamination. In, Papers and Proceedings of the Applied Geography Conferences. Volume 17. Edited by J.W. Frazier, B.J. Epstein, F.A. Schoolmaster III, and K.G. Jones. Kent, Ohio: John W. Frazier, pp. 182-190.

Page, A.L., and Chang A.C. 1993. "Lead Contaminated Soils: Priorities for Remediation?" Hazardous Waste & Hazardous Materials, 10, pp. 1-2.

Peuquet, D.J. 1994. "It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems," Annals of the Association of American Geographers, 84(3): 441-461.

Phoenix, J. 1993. "Getting the Lead Out of the Community." In, Confronting Environmental Racism. Edited by Robert D. Bullard. South End Press: Boston, Massachusetts, pp. 77-92.

Sayre, J. 1981. "Dust Lead Contribution to Lead in Children." In, Environmental Lead. Edited by Donald R. Lynam, Lillian G. Piantanida, and Jerome F. Cole. Academic Press: New York, pp. 23-40.

Society for Environmental Geochemistry and Health. 1993. Lead in Soil: Recommended Guidelines. Science Reviews: Northwood, United Kingdom.

Stockwell, J.R., Sorensen, J.W., Eckert, J.W., Jr., and Carreras, E.M. 1993. "The U.S. EPA Geographic Information System for Mapping Environmental Releases of Toxic Chemical Release Inventory (TRI) Chemicals," Risk Analysis, 13(2):155-164.

Sutton, P.M., Athanasoulis, M., Flessel, P., Guirguis, G., Hann, M., Schlag, R., and Goldman, L.R. 1995. "Lead Levels in the Household Environment of children in Three High-Risk Communities in California," Environmental Research, 68:45-57.

Taylor, P.J., and Johnson, R.J. 1995. "Geographic Information Systems and Geography," In, Ground Truth. Edited by John Pickles, The Guilford Press: New York, NY, pp. 51-67.

Tobias, R.A., Roy, R., Alo, C.J., and Howe, H.L. 1996. "Traking Human Health Statistics in 'Radium City,'" Geo Info Systems, 6(7):50-53.

Tosta, N. 1993. "Sensing Space," Geo Info Systems, 3(9):26-32.

Trimble Navigation Limited. 1996. Pro XL System Training Manual. Sunnyvale, CA.

U.S. Congress. House. 1993. Environmental Justice Act of 1992. 103rd Cong., 1st sess., H.R. 2105, Congressional Record. Vol. 139. Daily ed. (12 May), H2462.

U.S. Congress. House. 1992. Residential Lead-Based Paint Hazard Reduction Act of 1992. Public Law 102-550. House Report 102-1017. Washington, DC: U.S. Government Printing Office.

U.S. Congress. Senate. 1993. The Lead Exposure Reduction Act of 1993. Senate Hearing 103-181. Hearing Before the Subcommittee on Toxic Substances, Research and

Development of the Committee on Environment and Public Works. Washington, DC: U.S. Government Printing Office.

U.S. Department Department of Commerce. 1997. LandView III: Environmental Mapping Software. CD-ROM. Disc 1 of 10. U.S. Bureau of the Census. Dec. 1997.

U.S. Department of Health and Human Services. Agency for Toxic Substances and Disease Registry. 1991. Child Lead Exposure Study, Leeds, Alabama - Final Report. ATSDR/HS-92/13. Washington, DC.

U.S. Department of Health and Human Services. Public Health Service. 1991a. Healthy Children 2000. DHHS Publication No. HRSA-M-CH 91-2. Washington, DC.

U.S. Environmental Protection Agency. 1990. Toxics in the Community: National and Local Perspectives - The 1988 Toxics Release Inventory Report. Washington, DC: U.S. Government Printing Office.

_. 1994. Agency Guidance on Residential Lead-Based Paint, Lead Contaminated Dust, and Lead Contaminated Soil. Unpublished agency memorandum. Washington, DC: U.S. Government Printing Office.

_. 1994. Guidance Manual for the Integrated Exposure Uptake Biokinetic Model for Lead in Children. EPA/540/R-93/081. U.S. Government Printing Office: Washington, DC.

_. 1991. U.S. Environmental Protection Agency Strategy for Reducing Lead Exposures. Unpublished draft report. Washington, DC: U.S. Government Printing Office.

_. 1993. Investigation of Test Kits for Detection of Lead in Paint, Soils and Dusts. EPA/600/R-93/085. Research Triangle Park, NC.

_. 1993a. Urban Soil Lead Abatement Demonstration Project - Volume 1: Integrated Report. EPA/600/AP-93/001A. Washington, DC: U.S. Government Printing Office.

_. 1992. Environmental Equity: Reducing Risk For All Communities. 230 DR-92-002. Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office. 1998. Medicaid: Elevated Blood Lead Levels in Children. GAO/HEHS-98-78. Washington, DC: U.S. Government Printing Office.

Wartenberg, D. 1992. "Screening for Lead Exposure Using a Geographic Information System," Environmental Research, 59:310-317.

White, L.E. 1992. Lead Exposure from the Agriculture Street Landfill. In, National Minority Health Conference: Focus on Environmental Contamination. Edited by Barry L. Johnson, Robert C. Williams, and Cynthia M. Harris. Princeton, NJ: Princeton Scientific Publishing Co., pp. 149-157.

Wixson, B.G., and Davies, B.E. 1994. "Guidelines for Lead in Soil: Proposal of the Society for Environmental Geochemistry and Health," Environmental Science & Technology, 28(1): 26A-31A.

Wriggins, J. 1997. "Genetics, IQ, Determinism, and Torts: The Example of Discovery in Lead Exposure Litigation," Boston University Law Review, 77(5):1025-1088.

Young, A.R.M., Bryant, E.A., and Winchester, H.P.M. 1992. "The Wollongong Lead Study: An Investigation of the Blood Lead Levels of Pre-school Children and Their Relationship to Soil Lead Levels," Australian Geographer, 23(2):121-133.

Zeigler, D.J., Johnson, J.H., Jr., and Brunn, S.D. 1983. Technological Hazards. Washington, DC: Association of American Geographers.

Zimdahl, R.L., and Hassett, J.J. 1977. "Lead in Soil." In, Lead in the Environment. NSF/RA-770214. Edited by William R. Boggess and Bobby G. Wixson. National Science Foundation: Washington, DC, pp. 93-98.

Polygon Development Improvement Techniques for Hazardous Waste Environmental Impact Analysis

David A. Padgett
Austin Peay State University, Clarksville, Tennessee

Introduction

Recently, concern has arisen regarding the effect Superfund sites have on surrounding communities and, specifically, the distribution of those impacts on target populations. In designing geographic information systems (GIS) applications for analyzing potential impacts of hazardous wastes or waste sites on adjacent neighborhoods, many challenges may be encountered. GIS database design requires addressing questions of time, space, and scale.

The U.S. Environmental Protection Agency (EPA) and other federal agencies have conducted studies that indicate that certain sectors of the population may be more vulnerable to exposure to toxics than others. To date, federal departments have enlisted in several GIS-based research projects that attempt to delineate "geographic hot spots" of toxic contamination. Such GIS applications at hazardous waste sites have typically used polygons to represent data from census tracts and/or municipal boundaries. In most cases, however, census tract and other boundaries do not necessarily jibe with community and neighborhood boundaries; therefore, the polygons representing characteristic data for target populations may not be consistent with the actual status of those populations.

The objective of this paper is to demonstrate GIS methods for producing, to the greatest degree possible, socioeconomically and culturally homogenous polygons for impact analysis of specific sensitive populations and/or communities. The paper presents case studies of community/neighborhood characterization problems encountered in developing polygons during previous field investigations involving lead (Pb) contamination, toxic release inventory (TRI) sites, and solid/hazardous waste sites. The paper attempts to demonstrate effective solutions and suggestions for improving polygon development, including GIS data manipulations and software applications. In addition, the paper provides geographic and groundtruthing field methods to support and enhance the accuracy of remotely obtained information. Finally, the discussion includes community and geographic hot spot analyses for potential public health impacts.

Background

In 1992, EPA established the Environmental Equity Workgroup. Its members included personnel from the Offices of Toxic Substances and Civil Rights, as well as Policy, Planning and Evaluation. The workgroup conducted an extensive study on environmental equity issues. Their report offered several recommendations for improving federal agency efforts in protecting minority and low-income populations and recognized a need for more spatial and demographic data. The final report, titled *Environmental Equity: Reducing Risk for All Communities* (1), was released in February 1992 and concluded that "there is (sic) limited data on environmental health effects by race; there are differences by race and income in potential and actual exposures to some pollutants." In response to the above findings, the workgroup offered the following recommendations (1):

> EPA should establish and maintain information which provides an objective basis for assessment of risks by income and race, commencing with developing a research and data collection plan.
>
> It (EPA) should revise its risk assessment procedures to ensure . . . better characterization of risk across population, communities or geographic areas. In some cases it may be important to know whether there are any population groups at disproportionately high risk.

The Agency for Toxic Substances and Disease Registry (ATSDR) formed a community health branch to specifically examine the potential health impact of hazardous waste sites upon people living in surrounding communities. The new branch's personnel direct ATSDR's minority health initiative, which focuses upon health threats to minority populations, including those from environmental contaminants.

In addition, EPA established the Office of Environmental Equity. The office's mission includes analyzing environmental impacts upon minority populations, providing technical assistance to disadvantaged communities, and establishing environmental initiatives at minority academic institutions (MAIs). The office serves as a clearinghouse of environmental data and information for groups and individuals involved in environmental equity activities.

In 1993, Representative John Lewis (Democrat-Georgia) introduced the Environmental Justice Act to Congress. The act requires EPA and the Department of Health and Human Services (DHHS) to establish the geographic units for determining environmental high-impact areas (EHIAs), which are the 100 geographic areas found to have the highest volumes of toxic chemical releases.

GIS Applications

GIS could potentially help address the above data and information needs of EPA. The Agency specifically acknowledges this in other recommendations (1):

> EPA could further develop its enforcement prioritization policy to target high-risk populations. Under this scheme, the most exposed and highly susceptible populations in each region would be targeted for enforcement actions. Geographic Information System technology could be used to identify high-risk populations.

Several recent environmental studies have employed computer applications and spatial data. Goldman (2) used GIS in a major study that graphically displayed counties having high percentages of African-Americans, hazardous wastes, and diseases. Mohai and Bryant (3) applied a linear regression model to show a positive correlation between increasing proportions of minority populations and decreasing distances from hazardous waste sites in Detroit. Lavalle and Coyle (4) conducted an extensive analysis of computer databases that hold hazardous waste law enforcement information for the past 10 years. They found inequity in enforcement and remedial actions in white communities versus nonwhite communities.

EPA has enlisted GIS for community environmental impact projects at Regions II and III. EPA's Office of Health Research (OHR) is investigating methods for linking demographic data with TRI information to evaluate the relationship between levels of hazardous waste releases and exposure risks in minority communities. EPA has also developed the TRI "risk screening" process, which employs TRI, U.S. Census data, and GIS to identify TRI releases that may pose significant risk to human health or the environment (5).

Both EPA (6) and the North Carolina Department of Environment, Health, and Natural Resources (7) have recently completed GIS-based environmental investigations. The EPA study involved GIS analyses of TRI chemical releases in the southeastern United States. The report included numerous GIS-produced maps that show locations where TRI releases may be affecting densely populated areas and sensitive ecosystems. The North Carolina study applied GIS in searching for sources of lead-poisoning in children. Findings indicated a positive spatial correlation between high lead-contamination risk communities and those having certain socioeconomic characteristics, such as low income, above-average African-American population percentages, and above-average percentages of residents receiving public assistance.

The ATSDR recently implemented a study using GIS to evaluate and analyze the demographic characteristics of populations near National Priorities List (NPL) sites. According to the ATSDR, "As a result of our pilot tests, we have determined GIS to be the best methodology for identifying potentially impacted minority populations" (8).

Limitations of GIS

While GIS may be a viable tool for investigating environmental inequity, it is not an absolute solution. Issues involving hazardous waste impact assessments tend to be very complex without the added dimension of racism or discrimination. Efforts to determine a causal relationship between the presence of minority communities and environmental hazards must consider the questions of time, scale, and place. Unfortunately, in many instances, GIS applications may be unable to adequately illustrate these three pertinent issues resulting in skewed or altogether incorrect conclusions.

With respect to scale, among the immediate concerns when applying GIS is selecting appropriate sizes for polygons. As indicated above, EPA is in the process of determining the scale for EHIAs. A polygon may be a county or a census tract. Figure 1 illustrates problems

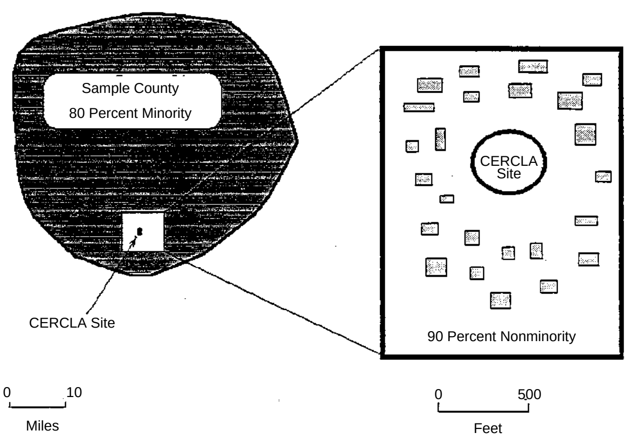


Figure 1. Problems of scale in GIS polygon design.

of scale associated with polygon size selection. At the county-size scale, a case of environmental inequity apparently exists with the presence of a Superfund site in the sample county because 80 percent of the county's population belongs to a minority ethnic group. A closer look, however, reveals that the population residing in the immediate vicinity of the waste site is predominantly nonminority.

Useful as GIS may be, its output in some cases may display static conditions without considering human movements over time. Figure 2 displays a common situation associated with the "filtering" phenomenon, in which a nonminority population moves out of an area while increasing numbers of an ethnic minority group moves into it. The figure shows that in 1950, a nonminority community surrounded a TRI site (i.e., an active industrial site releasing toxic substances). By 1990, the demographics of the neighborhood had changed along with the status of the TRI site, which is now an abandoned Superfund site. A GIS database probably would contain only information on the community from 1990. Such an instance could suggest that some form of environmental injustice exists given the presence of the Superfund site within the minority community. Accounting for the dynamics of time and human movement, however, would show that the waste site preceded the minority population and that, in actuality, the minority community moved toward the site. This conflicts with the prior notion that unsavory forces placed the site in the minority community.

Figure 3, a schematic of polygons used in an investigation into sources of lead-poisoning in children, also displays the limitations of GIS with respect to time and human dynamics, but at a lesser time interval. The study area is divided into low- and high-risk areas for lead contamination. The locations at which children with lead poisoning were found, however, do not correspond with the areal risk factors. In this case, the GIS is limited in its ability to follow human movements on a daily basis. For instance, a parent with a child who exhibits unhealthy blood-lead concentrations may report the child's home address as someplace within the low-risk polygon. The child may attend school in the high-risk area, however. The children's points of contact with lead may not necessarily correspond with their home addresses, resulting in an inaccurate graphic display.

With respect to place, GIS may be limited in its ability to determine the specific borders of a socioeconomically and demographically homogenous human population. Tosta (9) and Coombes et al. (10) discuss the dilemmas associated with neighborhood boundary delineation in GIS applications. Figure 4 displays a schematic of a census tract. The GIS database may list the tract's per

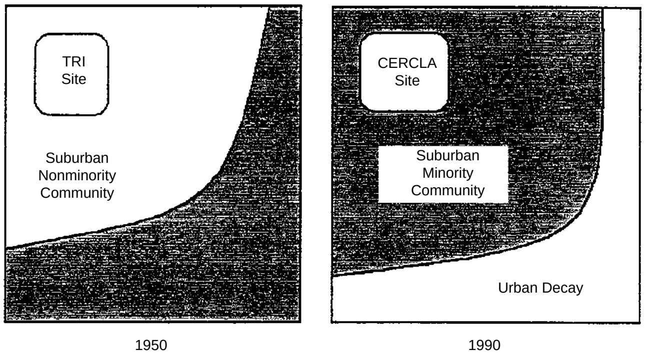


Figure 2. Example of changing community demographics with time near a hazardous waste facility.

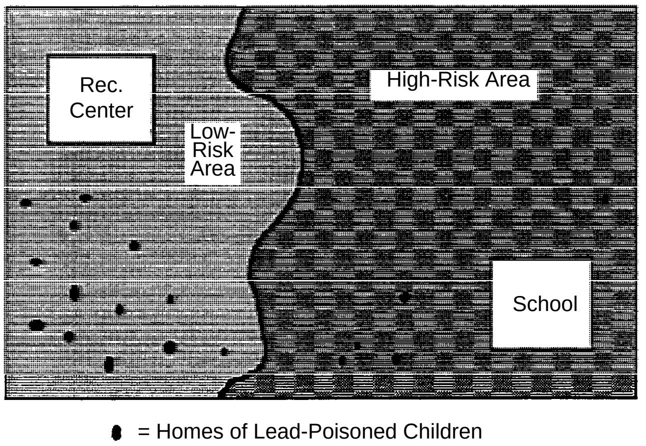


Figure 3. Lack of correspondence between locations of lead-poisoned children and high-contamination risk areas due to daily dynamics of human movements.

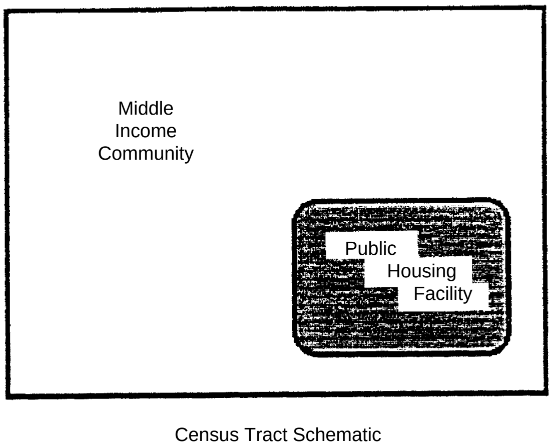


Figure 4. Example of significant neighborhood-type variation within a single polygon, possibly resulting in skewed socioeconomic data.

capita income as relatively low and may list the tract as a low-income neighborhood. Further investigation may find, however, that two very different socioeconomic communities exist within the tract, one middle-class and the other a public housing facility. Frequently, middle-income, African-American communities have low-income housing projects built adjacent to them. With respect to the polygon in Figure 4, if an investigator wanted to research health impacts of toxic wastes for low-income households, using this polygon and others like it would inaccurately depict communities within them.

An additional problem in community definition is determining exactly what defines a minority community. The most common indicator for discerning a minority community would be the existence of a clear majority of some minority group as in Polygon A of Figure 5, or where the minority group makes up more than 50 percent of the population as in Polygon B. In some instances, however, communities have received minority status without the presence of the conditions in Polygons A and B.

Previous investigations show a number of measures used to identify minority communities and census tracts. Greenpeace conducted a 1990 environmental justice study that determined a community's status as minority based upon the relationship between a community's percentage of minority population and the selected minority group's national percentage (11).

Polygon C in Figure 5 depicts Greenpeace's minority community definition. Taking the target ethnic group in Polygon C as African-Americans and the hypothetical extent of Polygon C as the United States (African-Americans make up approximately 12 percent of the total U.S. population), one may determine that Subpolygon C is a minority polygon or community because its minority percentage is over twice that of the national percentage or extent of the large population in Polygon C. The condition that Polygon C illustrates is also evident in a study

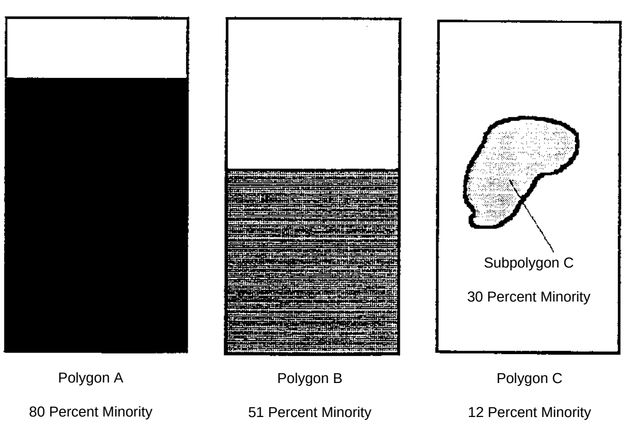


Figure 5. Problems in defining minority polygons.

by Mohai and Bryant (3). The authors claim that environmental inequity exists in Detroit where they found that on average, within a 1-mile radius of the city's hazardous waste facilities, 48 percent of the population is nonwhite.

Solutions With GIS and Supporting Technological Methods

From the above discussion, GIS may appear very limited for use in environmental community impact investigations, but GIS can actually be an extremely effective tool if employed with appropriate supporting technology.

Preliminary Groundtruthing

To design GIS databases that reflect the true nature of target groups and human dynamics, preliminary groundtruthing may be necessary. In some cases, investigators make gross interpretations of suspected environmental inequity without actually visiting the study area. Without groundtruthing prior to final database development, questions of time, scale, and human dynamics may be left unanswered. The integrity of databases produced this way and the antecedent conclusions based upon them may fall into question. Thus, because the nature of environmental and human health impact studies is complex, some on-the-ground work should precede or at least accompany database construction efforts.

Cause-Effect Analyses

Epidemiological studies are increasingly employing GIS. This use is important with respect to environmental investigations because in many cases, proof of a correlation between a waste site and community health problems may be necessary. Croner et al. (12) describe statistics-supported GIS applications for linking "cancer hot spots" with pollution sources. Without conclusive evidence that waste sites and other environmental hazards negatively affect health in socioeconomically disadvantaged populations, claims of environmental injustice may be difficult if not impossible to prove.

In historical analyses of waste facility sitings, GIS may be useful, along with the support of gravity models, in investigating whether the sitings followed the prescribed logic for such siting decisions. Noble (13) wrote that the costliest aspect of waste facilities management is transportation; therefore, siting decisions should favor locations in closest proximity to a selected community's centroid of waste production. Using a GIS gravity model with data for household wastes produced within a given locality, the center of gravity of the volumes of wastes produced could be located. Historical analyses of past siting decisions may find that past siting decisions defied logic. Instead of finding waste sites placed in environmentally safe locations as close as possible to areas of greatest refuse generation, analyses may find instead that sites have been placed farther away in disadvantaged communities. Both the taxpaying and potentially affected residents would pay for such unsavory siting practices.

Conclusion

The potential for technological applications, including GIS, in this arena is great. Increased involvement by technologically trained environmental professionals is imminent. Their future involvement, however, must focus on the scientific soundness of investigative methods, data integrity, and the equitable participation of potentially affected citizens in any subsequent decision-making processes.

References

1. U.S. EPA. 1992. Environmental equity: Reducing the risk for all communities. EPA/230/DR-92/002. Washington, DC.
2. Goldman, B. 1991. The truth about where you live. New York, NY: Times Books.
3. Mohai, P., and B. Bryant. 1992. Environmental racism: Reviewing the evidence. In: Bryant, B., and P. Mohai, eds. Race and the incidence of environmental hazards: A time for discourse. Boulder, CO: Westview Press.
4. Lavalle, M., and M. Coyle. 1992. Unequal protection: The racial divide in environmental law. The Environmental Professional 153:S1-S12.
5. U.S. EPA. 1990. Toxics in the community: National and local perspectives. EPA/560/4-90/017. Washington, DC.
6. Stockwell, J.R., J.W. Sorensen, J.W. Eckert, Jr., and E.M. Carreras. 1993. The U.S. EPA geographic information system for mapping environmental releases of toxic chemical release inventory TRI chemicals. Risk Analysis 132:155-164.
7. North Carolina Department of Environment, Health, and Natural Resources. 1993. Lead database evaluation and GIS modeling project. PR-242592. Raleigh, NC: National Institute for Environmental Health Sciences.
8. Harris, C.H., and R.C. Williams. 1992. Research directions: The Public Health Service looks at hazards to minorities. EPA J. 181:40-41.
9. Tosta, N. 1993. Sensing space. Geo. Info. Sys. 39:26-32.
10. Coombes, M., S. Openshaw, C. Wong, and S. Raybould. 1993. Community boundary definition: A GIS design specification. Regional Studies 3:280-86.
11. Bullard, R.D. 1993. Environmental equity: Examining the evidence of environmental racism. Land Use Forum (Winter). pp. 6-11.
12. Croner, C.M., L.W. Pickle, D.R. Wolf, and A.A. White. 1992. A GIS approach to hypothesis generation in epidemiology. In: American Society for Photogrammetry and Remote Sensing/American Congress on Surveying and Mapping/RT 92, Washington, DC (August). Technical papers, Vol. 3: GIS and cartography. pp. 275-283.
13. Noble, G. 1992. Siting landfills and LULUs. Lancaster, PA: Technomic Publishing Company, Inc.

Applicability of GIS Tools in Environmental Conflict Mapping: A Case Study in Hungary

László Pásztor[1*], József Szabó[1], Zsófia Bakacsi[1],
Simon T.D. Turner[2], Tibor Tullner[3]

[1]*Research Institute for Soil Science and Agricultural Chemistry of the Hungarian Academy of Sciences, GIS Lab*

[2]*ADAS International, UK*

[3]*Geological Institute of Hungary*

ABSTRACT

Present paper demonstrates technical problems and their solution emerged during the compilation of first draft Environmental Conflict Maps (ECM) in Hungary. The development of a methodology elaborated for a pilot area, and the produced ECMs themselves are presented. In developing this ECMs, local knowledge, that captures and identifies existing environmental problems, was integrated with agri-environmental spatial databases. Various input information sources, their spatial and/or thematic compatibility/incompatibility, data availability, spatial analysis are presented and discussed. Definition of and examples for direct and indirect conflicts are demonstrated.

INTRODUCTION

The Integrated GIS of the Ministry of Environment and Regional Policy of Hungary (MERP) laid the foundation of a standard environment and user interface for environmental data management first of all in Thematic Information Centres (TICs) and other regional authorities of the Ministry. The completion of this project brought up the necessity to provide this kernel system with extensions furnishing ministerial authorities decision-making tools for environmental management and state of the environmental assessment. The ECM methodology is regarded as one of the cornerstones of these extensions facilitating the understanding, assessment, as well as the mitigation and/or elimination of environmental conflicts occurring as a result of the interaction of socio-economic and natural factors.

The purpose of compiling ECMs upon standard methodology within the framework of the related subproject can be summarised as follows:

---

* H-1022 Budapest, Herman Ottó út 15. , Hungary, Fax: (36-1) 212-1891, E-mail: lacus@rissac.hu

- identifying the users of the related data;
- identifying conflicting environmental factors;
- raising public awareness on the gravity of the issue;
- assessing the need, feasibility and format of communication links between TICs and municipalities;
- setting up an evaluation scheme supporting complex assessment of different factors;
- full compatibility of the MERP with the Integrated GIS; and
- harmonising with EU standards and priorities.

Definition of environmental conflicts and their management is a sophisticated issue considering both the complexity of interacting factors and local and regional organisations involved in their evaluation. The possibility of working up a standard method is further impeded by the multitude of local factors specific for the regarded geographic area. It was thus necessary to involve a broad range of regional and local experts possessing professional knowledge in handling these problems.

DEFINITIONS

The issue of ECM is extremely complex, the successful outcome of which will largely depend on the issues involved, the stakeholders and scope for negotiation or mediation. However, the application of maps, and notably environmental conflict maps can play a useful role in the process of ECM. The use of a map to communicate complex environmental information is well recognised. An environmental conflict will almost certainly be unique to a particular set of circumstances and therefore to prescribe a rigid approach to producing an ECM would probably give rise to a product that will not be useful to the stakeholders involved.
Despite the broad range of factors acting in environmental conflicts it is necessary to provide its explicit definition as well as to distinguish between direct conflicts occurring invariably as a result of a well-defined pollution load on a target surface and indirect conflicts always related to actual or foreseen modification of the land use pattern. Environmental problems can be defined as damage or the threat of causing damage to the quality of the environment.

Environmental Conflict: within and/or between spatial and/or temporal coincidence of environmental pressure (pollution, /threshold of land and susceptibility of target surface) and/or general environmental elements which is recently and/or might prove to be in the future disturbing or harmful for the receptor (actual- and future state of landuse).

Direct conflict: between pollution load and susceptibility of target surface.

Indirect conflict: between suitability for a human induced environmental management and actual/future state land use/management.

PILOT PROJECT

In order to present a practical case study of ECM methodology, a pilot area was selected along the NE shore of Lake Balaton. It represents an approximately 5 km. wide belt amounting to approximately 100 km^2. Selection of this area was based on the following aspects:

- it is considered as the most important recreation area and tourist target in Hungary;
- it is affected by a number of factors providing sources to environmental conflicts; and the
- availability of a number of thematic data necessary for complex evaluation.

ELEMENTS OF ECM GIS

Huge work was invested in the collection of all relevant data for the pilot area. Compromising between data requirements and availability proved to be the greatest challenge. Finally the following dataset was set up.

➢ Topography is represented by the following:
  - DTM with a cell size of 50 meters,
  - Road, railway, water networks,
  - Lake Balaton is represented by a further theme, and
  - Watersheds of Balaton are also represented by a polygon theme.

➢ In addition to the topography of the pilot area, the
  - Country border of Hungary and the
  - Main rivers of the country are added as a background.

➢ A satellite image theme covers the majority of the pilot area. Involvement of further remotely sensed data was hindered due to their high cost. Actually, land cover information of (CORINE; see later) is also based on satellite imagery.

➢ Information on soil is represented by the polygon theme AGROTOPO. AGROTOPO is a soil information system on Hungarian soils in a scale of 1:100.000. Its mapping units (agroecological units) are characterized with nine different attributes.

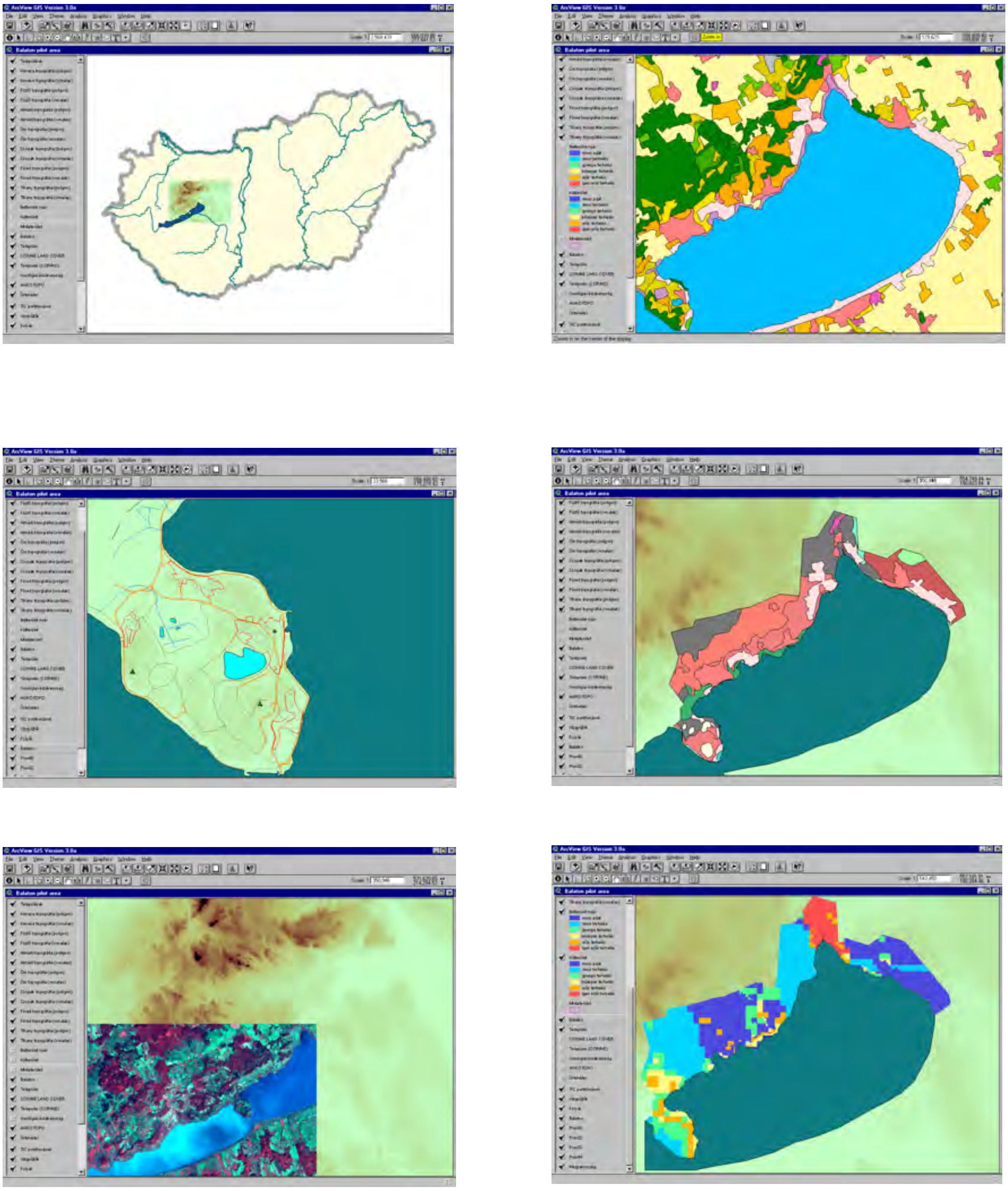


Figure 1

Some quick views on ECM ArcView project file

- Information on geology is also represented by one theme. An expert system based value characterising the vulnerability of the geological environment is involved in the ECM methodology.

- Information on land cover is also represented by one theme (CORINE). Land cover categories are given with a code of three digits where the first two digits refer to two further, higher level categorization of actual land cover.

- A raster was filled up by local municipal experts. Environmental impact categories were determined upon the experience of EU and Hungarian studies but considering the Sofia priorities and the 5th Environmental Action Program as well. Moreover, they indicated conflicting land use schemes occurring in their area and made a summary of the major problems occupying them. The specified criteria (namely 15 different factors) were to be classified and indicated by the municipalities for each cell of the raster on a scale ranging between 0-4. The spatial resolution of the raster is 500x500 meters inside settlements and 1000x1000 meters around settlements. This local knowledge based information is shared in four individual themes. Data were provided for settlements and outer regions for low and high (touristic) season.

- TIC data involves two kind of information:
  - sub-catchments of Lake Balaton as a polygon theme without any important attribute furthermore only three of them are touched by the pilot area.
  - potential and/or actual polluter sites within pilot area as a point theme with a huge number of descriptive data, which includes detailed information on industrial and communal waste depositories, 75 are lying inside the pilot area.

ECM ArcView project collects all relevant and available data for Balaton pilot area in one session. In one hand this project is highly suitable for computer demonstrations of the data collection activities, on the other hand it also represents the starting point of any spatial analysis carried out based on the ECM database in order to achieve ECMs.

SAMPLE ECMs

Sample ECMs are stored in various ArcView project files. These files originate from ECM project file. They are supported with ArcView extension: Spatial Analyst for the execution of spatial analysis. Unrelevant data layers are deleted, new themes created either as auxiliary themes or ECM results. A final softcopy map as a layout is always compiled.

1. Groundwater vulnerability under TIC censused sites

The vulnerability of the geological environment is of outstanding importance in accelerating or preventing the progress of pollution in groundwater and deep subsurface water horizons. Various geological factors determine the capacity of preventing the migration of pollution into subsurface water tables. TICs are responsible for the management and operation of regional environmental data as well as to collect information on pollutants/potential pollutants: location, ownership, cadaster emission data, amount of processed and produced hazardous waste, site and state of communal waste depositories. Potential and/or actual polluters situated over locations with higher hazard of pollution vulnerability should be treated/checked more seriously. Consequently, a co-evaluation, spatial analysis of the conflicting environmental factors should be carried out to get a complex picture on groundwater vulnerability endangered by TIC censused sites.

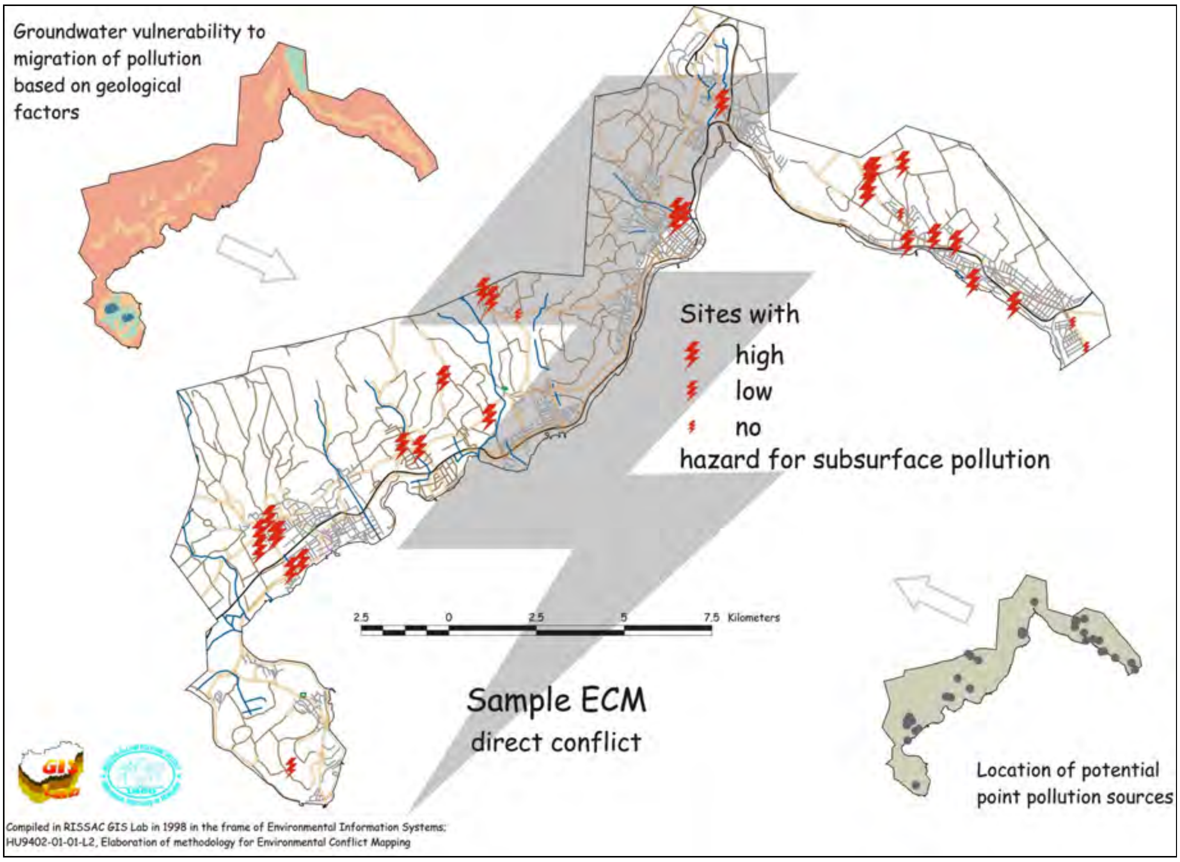


Figure 2

Sample ECM: Groundwater vulnerability under TIC censused sites

An expert system based methodology of pollution vulnerability assessment subdivides the considered area into 4 classes according to the combination of three geological factors (namely permeability of the superficial geological formations, thickness of the eventual impermeable overburden and the highest groundwater level). High, medium and low vulnerability classes are defined considering the capacity of preventing the migration of pollution. TIC censused sites were categorized according to their profiles into two groups:

sites producing or not producing hazardous waste with risk to subsurface water tables. Sites belonging to the former class then were further classified according to their geographical location within the various geology based vulnerability categories. Sites are ranked as ones with high, low or no hazard of subsurface pollution. The output ECM entitled ‘Groundwater vulnerability under TIC censused sites’ is displayed in Fig. 2.

2. Impact of traffic on Natural Conservation

Territory of Balaton National Park covers major part of northern Balaton region, consequently overlaps the present ECM pilot area. Balaton region is also characterized as tourist attraction which requires infrustructural background: road and railway network. Noise and pollution of the traffic network conflicts with the functions of National Park, parts of National Park close to elements of traffic network are less valuable and strongly exposed to damages than farther areas. The co-evaluation, spatial analysis of the conflicting environmental factors gives a complex picture on status of Balaton National Park endangered by traffic caused impacts.

Impact of environmental elements represented as line features was modeled using buffering techniques. Buffering generally means creating a polygon around spatial objects with a given radius. Spatial Analyst module of ArcView however provided a more sophisticated opportunity. According to the distance of the elements of a grid from a spatial unit grid cells were classified; thus the whole territory was characterized with one operation. Co-evaluation of buffer zones resulting from railway and road system and further merging with spatial extension of National Park then provided vulnerability classification of Balaton National Park. The output ECM entitled ‘Impact of traffic on Natural Conservation’ is displayed in Fig. 3.

3. Overall perception on state of local environment

Simultaneous environmental impacts sum up in overall perception on state of environment in human mind. Too much load results in bad perception, which is a conflict in itself. This situation can be treated as a ‘polarized’ conflict where is no other component, but any, even the slightest, potential negative change in state of environment results in real conflicts. Highly impacted sites are ‘preconflict’ areas.

Values for specific environmental elements provided by municipalities for pilot area raster were summed up. Since high value for a given factor means bad feeling on state of environment, higher the summed up value the more the environmental impact within a given raster cell. Actually, summarizing category values representing as different factors involved

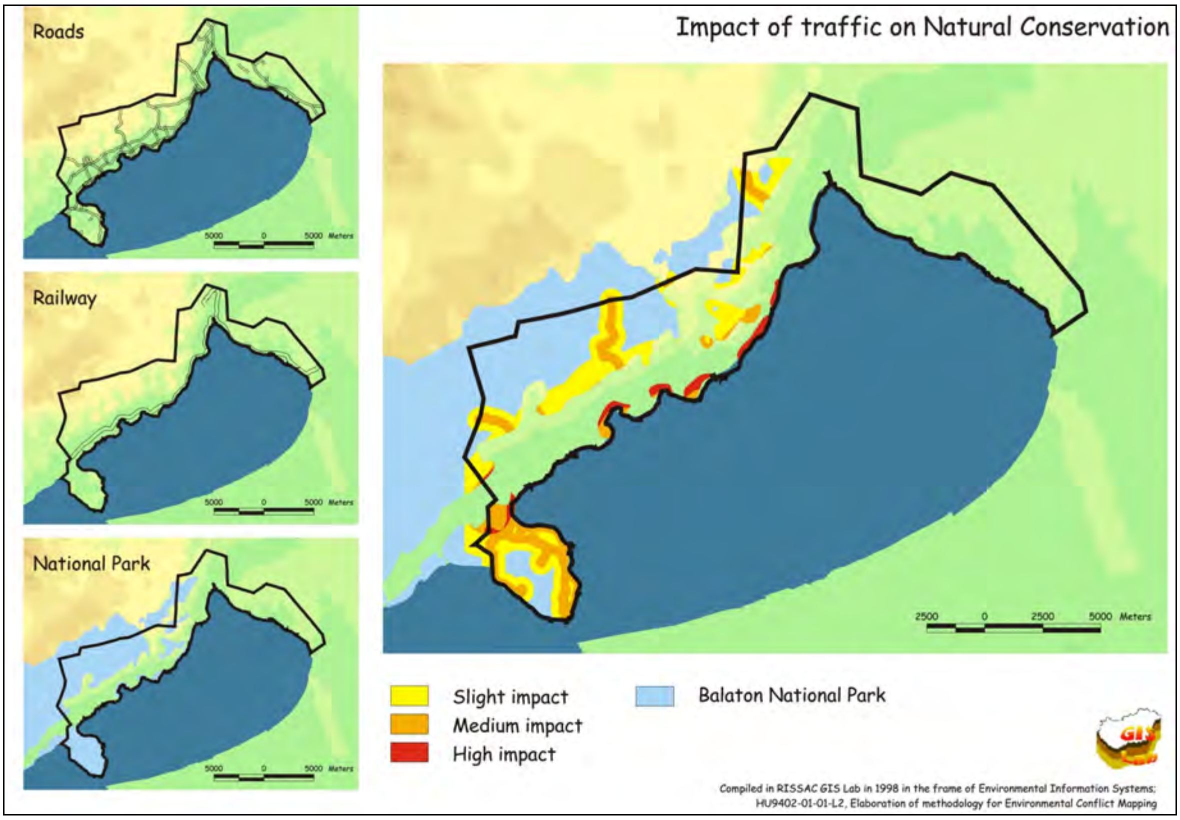


Figure 3

Sample ECM: Impact of traffic on Natural Conservation

in data collection can be judged at the first sight. Nevertheless overthinking the procedure it can be concluded that the result does give an indication about the state of local environment. The output ECM entitled ‘Overall perception on state of local environment’ is displayed in Fig. 4.

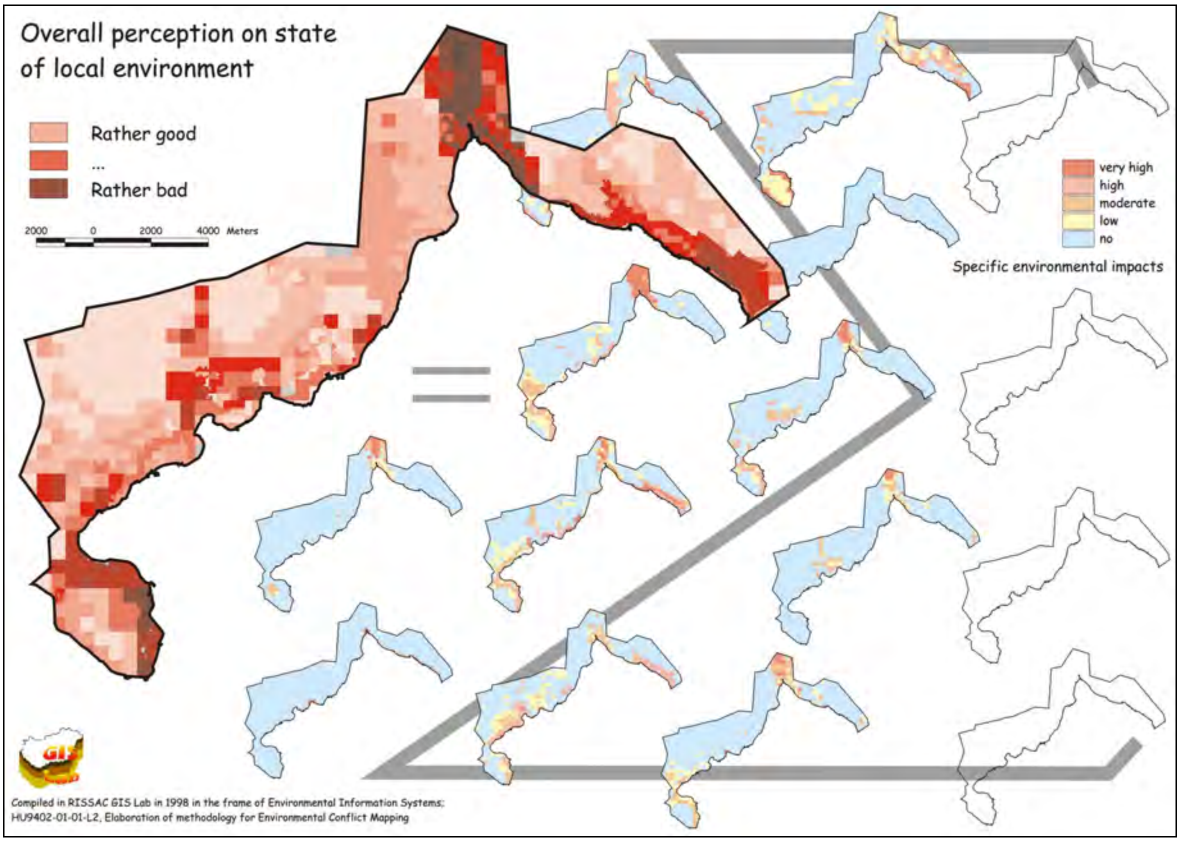


Figure 4

Sample ECM: Overall perception on state of local environment

4. Water quality evaluation

The greatest attraction of Lake Balaton, as a tourist target, is its water. Consequently, water quality is of prime significance. Quality of water is affected by multiple factors and shows temporal/seasonal changes. About forty water quality parameters area measured in the course of the year at sample sites along the shoreline. Cartographic treatment of this problem also produces ECMs. The visualisation of the point, but multitemporal and/or multivariate information was a challenging task for us. Our proposed solutions area displayed in Fig 5 and Fig 6.

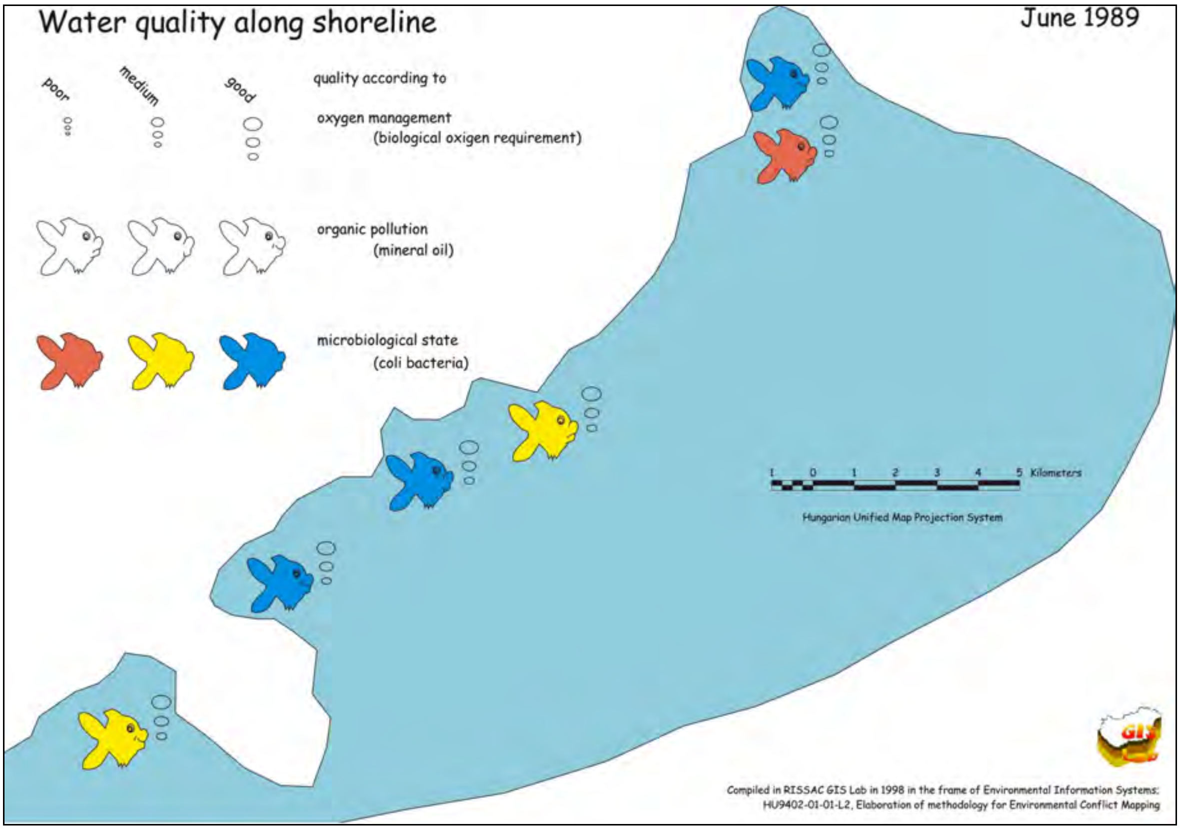


Figure 5

Sample ECM: Water quality along shoreline: multivariate evaluation

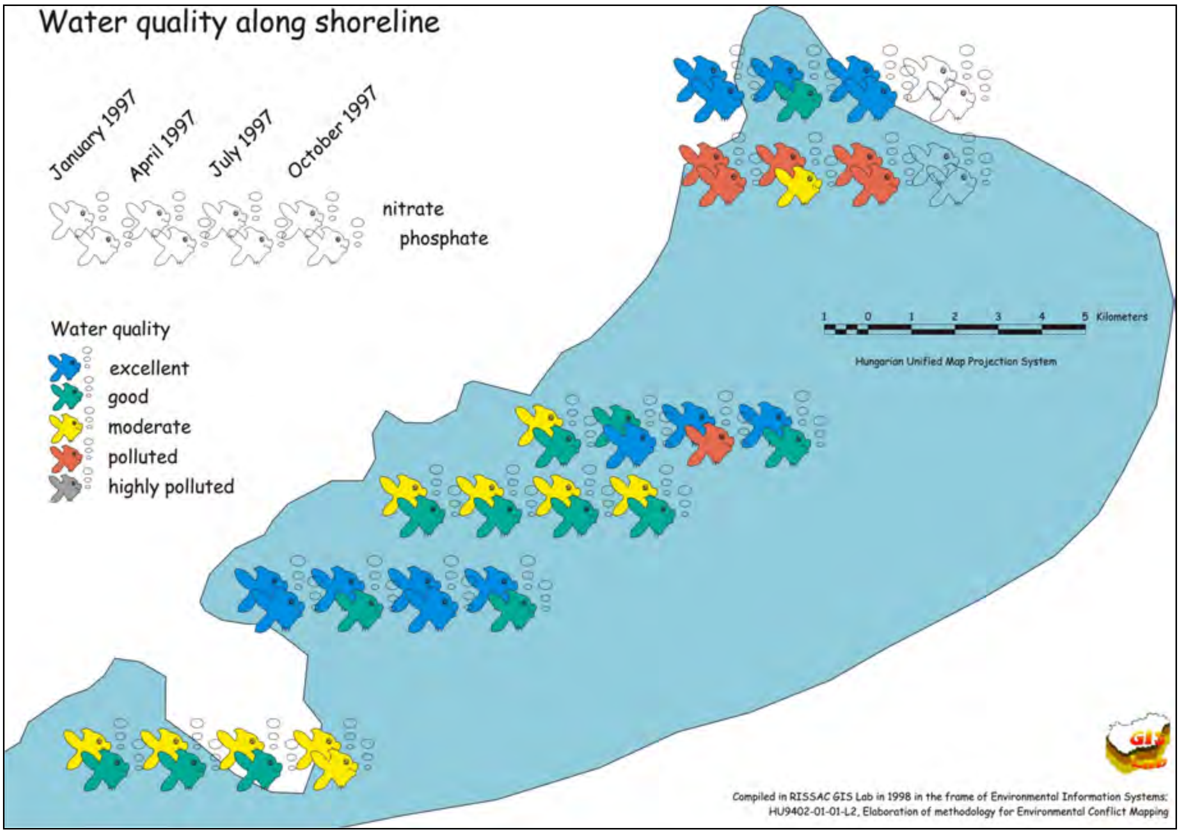


Figure 6

Sample ECM: Water quality along shoreline: multivariate/multitemporal evaluation

ACKNOWLEDGEMENT

Present work was carried out in the frame of PHARE project No. HU 9402-01-01-L2. The ECM Project was very effectively aided by Chief Information Office, Departement for Environmental Information Systems, Directorate for Environmental Strategy, Ministry of Environment and Regional Policy, under chairmanship of *Pál Bozó*. The Ministry of Environment and Regional Policy has made significant progress in the establishment of an Environmental Information System. This has included, through close co-operation with the UNEP-GRID programme, the establishment of GRID Budapest as a node on the GRID network; the establishment of two Thematic Information Centres (TIC's) at Hortobágy and Székesfehérvár covering the themes Nature Conservation and Environmental Protection respectively. Our special thanks go to the following persons who were very active in the project: *Kálmán Rajkai, Balázs Zágoni,* (Research Institute for Soil Science and Agricultural Chemistry of the Hungarian Academy of Sciences), *Endre Farkas and Imre Szepesi,* (Environmental Protection Thematic Information Centres at Székesfehérvár), *Tibor Cserny, Gábor Turczy,* (Geological Institute of Hungary, Division of Information Management) and *Zsuzsanna Flachner.*

Watershed Stressors and Environmental Monitoring and Assessment Program Estuarine Indicators for South Shore Rhode Island

John F. Paul and George E. Morrison
Environmental Research Laboratory, U.S. Environmental Protection Agency, Narragansett, Rhode Island

Abstract

The U.S. Environmental Protection Agency has initiated the Environmental Monitoring and Assessment Program (EMAP), a nationwide ecological research, monitoring, and assessment program whose goal is to report on the condition of the nation's ecological resources. During the summers of 1990 through 1993, data were collected from approximately 450 sampling locations in estuarine waters of the Virginian Biogeographic Province (mouth of the Chesapeake Bay to Cape Cod). During this period, sampling stations were located in the coastal ponds and coastal area of south shore Rhode Island.

One objective of EMAP is to explore associations between indicators of estuarine condition and stressors in the watersheds of the sampled systems. Extensive watershed information for south shore Rhode Island is available in geographic information system (GIS) format. Watershed stressors along south shore Rhode Island were compared with EMAP indicators of estuarine conditions using GIS analysis tools. The indicator values for coastal EMAP stations (those offshore from coastal ponds) were associated with all of the aggregated south shore watershed stressors. The coastal pond indicator values were associated with stressors in the individual coastal pond watersheds. For the total south shore watershed, the major land use categories are residential and forest/brush land, followed by agriculture. Closer to the coast, residential land use is more prevalent, while further from the coast, forests/brush lands dominate. All coastal EMAP stations, with one exception, exhibited unimpacted benthic conditions, indicating no widespread problems. For the individual watersheds, the major land use categories are residential and forest/brush land. The population density (persons per square mile) shows an increasing trend from west to east. Impacted benthic conditions were observed at EMAP sampling sites in two coastal ponds. These two impacted benthic sites appear to be organically enriched.

Introduction

Since its inception in 1970, the U.S. Environmental Protection Agency (EPA) has had the responsibility for regulating, on a national scale, the use of individual and complex mixtures of pollutants entering our air, land, and water. The Agency's focus during this period centered primarily on environmental problems attributable to the use of individual toxic chemicals. Regulatory policy, while continuing to control new and historical sources of individual chemicals (i.e., "end of the pipe") and remediate existing pollution problems, will have to address the cumulative impacts from multiple stresses over large spatial and temporal scales.

In this decade, the focus of environmental problems, or "scale of concern," has shifted from point-source and local scales to regional and global scales. Concurrently, the focus has shifted from chemical to nonchemical stressors. The threat posed by nonchemical stresses (e.g., land use, habitat alteration and fragmentation, species loss and introduction) presents a substantial risk to the integrity of both specific populations and ecosystems, and entire watersheds and landscapes.

The shift in the scale of concern for environmental problems presents a unique challenge for environmental decision-making. Traditionally, environmental information has been collected over local spatial and short temporal scales, focused on addressing specific problems, limited in the number of parameters measured, and collected with a variety of sampling designs that were neither systematic nor probabilistic. It is not surprising, then, that several scientific reviews concluded that the information needed to assess, protect, and manage marine and estuarine resources was either insufficient or unavailable and recommended a national network of regional monitoring programs (1, 2). Two key recommendations resulted from these reviews: (1) the need for a national monitoring program designed to determine the status and trends of ecological resources, and (2) the need for an assessment framework for synthesizing and interpreting the information being produced in a timely manner and in a form that the public can understand and decision-makers can use. EPA's response to these recommendations was to institute a long-term monitoring program, the Environmental Monitoring and Assessment Program (EMAP), and to adopt a risk-based strategy for decision-making.

EMAP is a nationwide ecological research, monitoring, and assessment program whose goal is to report on the condition of the nation's ecological resources. During the summers of 1990 through 1993, data were collected from approximately 450 sampling locations in estuarine waters of the Virginian Biogeographic Province (mouth of the Chesapeake Bay to Cape Cod) (3-5). During this period, some of the sampling stations were located in the coastal ponds and coastal area of south shore Rhode Island. One objective of EMAP is to explore associations between indicators of estuarine condition and stressors in the watersheds of the sampled systems. Extensive watershed information for south shore Rhode Island is available in geographic information systems (GIS) format.

The intent of this paper is to compare watershed stressors with EMAP indicators of estuarine condition along south shore Rhode Island using GIS analysis tools. The indicator values for coastal EMAP stations (those offshore from coastal ponds) are associated with all of the aggregated south shore watershed stressors. The coastal ponds indicator values are associated with stressors in the individual coastal pond watersheds. The project reported on in this paper served as a pilot for integrating watershed information with wide-scale ecological data collected to assess condition of estuarine waters.

Ecological Risk Assessment Context

Robert Huggett, EPA's Assistant Administrator for Research and Development, is using the risk assessment-risk management paradigm as a framework to reorganize the EPA research laboratories (6). Huggett is also reorienting the research that EPA conducts to be risk based (both human and ecological). The major thrust of the research to be conducted in the EPA laboratories will be directed toward reducing the uncertainties in the risk assessment process. In this way, the risk assessment context provides the "why" for the research conducted.

Ecological risk assessment is defined as a process for evaluating the likelihood that adverse ecological effects have occurred, are occurring, or will occur as a result of exposure to one or more stressors (7). The value of the risk assessment framework lies in its utility as a process for ordering and analyzing exposure and effects information, and in its flexibility for describing past, present, and future risks.

One way of depicting the ecological risk assessment process is shown in Figure 1 (8). The key points are that the process is continuous; the process can be oriented in either direction, dependent upon the form of the question or issue being addressed; and monitoring is at the hub, providing information to all activities. The end result of the effort is to provide better information for making environmental management decisions.

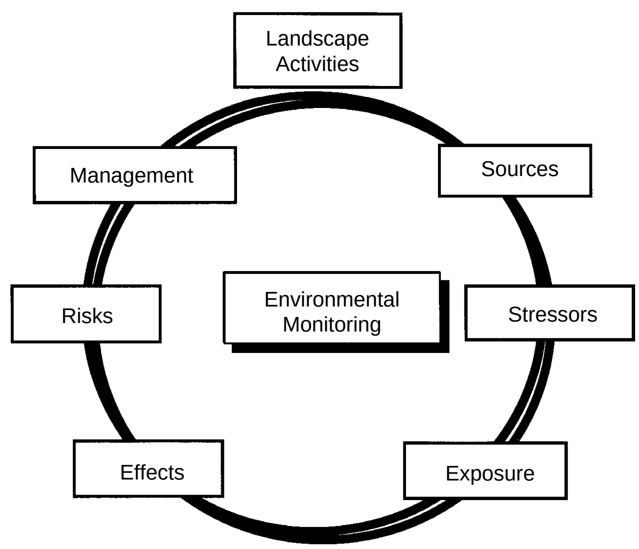


Figure 1. Ecological risk assessment framework (8).

Overview of EMAP and Estuarine Results

EMAP has been described as an approach to ecological research, monitoring, and assessment (9). It is not the only approach but is an approach that is driven by its goal to monitor and assess the condition of the nation's ecological resources. The objectives of the program to address this goal are to:

- Estimate the current status, trends, and changes in selected indicators of the condition of the nation's ecological resources on a regional basis with known confidence.
- Estimate the geographic coverage and extent of the nation's ecological resources with known confidence.
- Seek associations among selected indicators of natural and anthropogenic stress and indicators of ecological condition.
- Provide annual statistical summaries and periodic assessments of the nation's ecological resources.

The approach used by the program to meet its objectives and address its goal includes:

- Use of a large, regional scope that encompasses the entire county but provides information on the scale that federal and regional environmental managers require.
- Emphasis on ecological indicators to provide the information to assess condition (i.e., collect information on the ecological systems themselves to determine their condition or "health").
- A probability-based sampling design to produce statistically unbiased estimates on condition and to provide uncertainty bounds for these estimates.
- A vision of the program as long-term, continuing into the next century, which is consistent with the large, regional spatial scale being addressed.
- Development through partnerships with other agencies that have natural resource stewardship responsibility.

The estuarine component of EMAP was initiated in 1990, with monitoring in the estuarine waters of the Virginian Biogeographic Province (mouth of Chesapeake Bay northward to Cape Cod) (10). Figure 2 depicts the biogeographic provinces of estuarine resources of the country. These provinces have been delineated based upon major climatic zones and the prevailing offshore currents (11). This is comparable with the ecoregion approach used to describe the distribution of terrestrial ecosystems (12). The biogeographic province is the comparable approach for coastal ecosystems. Monitoring in the Virginian Province continued through 1993; monitoring was conducted in the Louisianian Province from 1991 to 1994; monitoring was initiated in the Carolinian Province in 1994; and monitoring will be initiated in the West Indian Province in 1995.

A suite of measurements was collected at each of the EMAP-Estuaries sampling sites that were selected with a probability-based sampling design (13, 14). As indicated above, the measurements emphasized ecological conditions indicators, which included biotic indicators such as benthic and fish abundance, biomass, diversity, and composition, and also included abiotic indicators such as dissolved oxygen, sediment contaminant concentration, and sediment toxicity (15).

In the Virginian Province, approximately 450 probability-based sampling sites were visited during the summer periods in 1990 through 1993 using consistent indicators and collection and analysis procedures. An example of the results is shown in Figure 3, which presents the condition of benthic resources (16). The benthic condition is reported using a benthic index, which is an aggregate of individual benthic measurements that were combined using discriminant analysis to differentiate impacted from unimpacted sites (3, 17). The figure presents results for values of the benthic index that were determined to be impacted. The bar chart is the standard EMAP format for results: province-scale results with 95-percent confidence intervals about estimates. The large, small, and tidal categories refer to the strata used in the probability-based sampling design: large systems are the broad expanses of water such as in Chesapeake Bay, Delaware Bay, and Long Island Sound; small

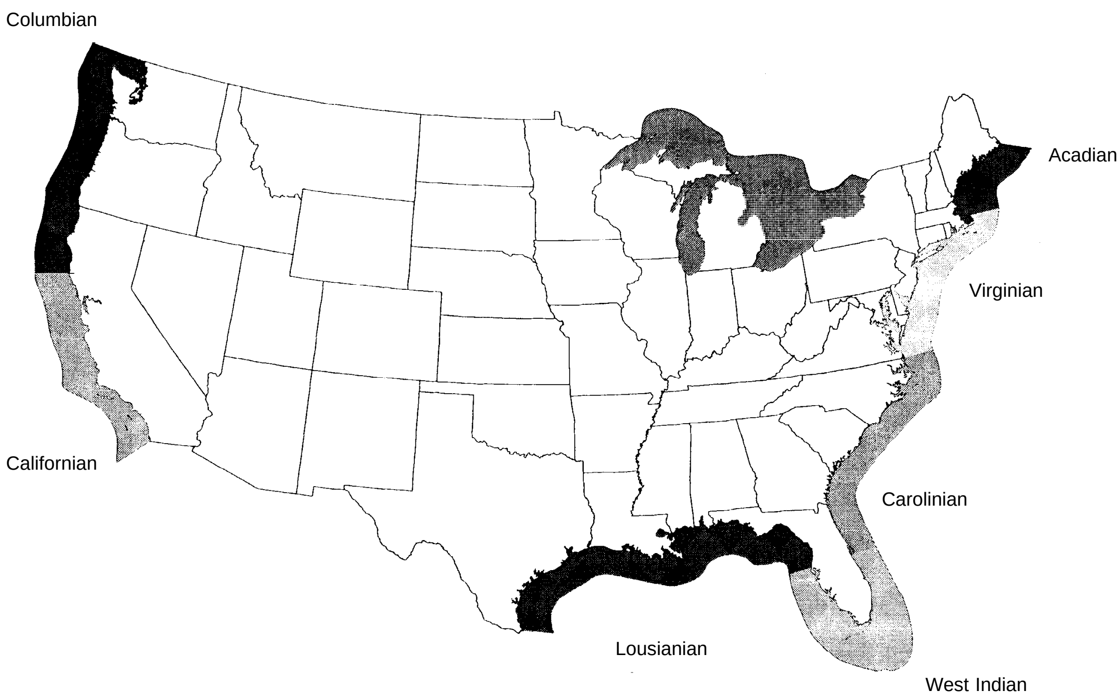


Figure 2. Biogeographic provinces used by EMAP-Estuaries (13).

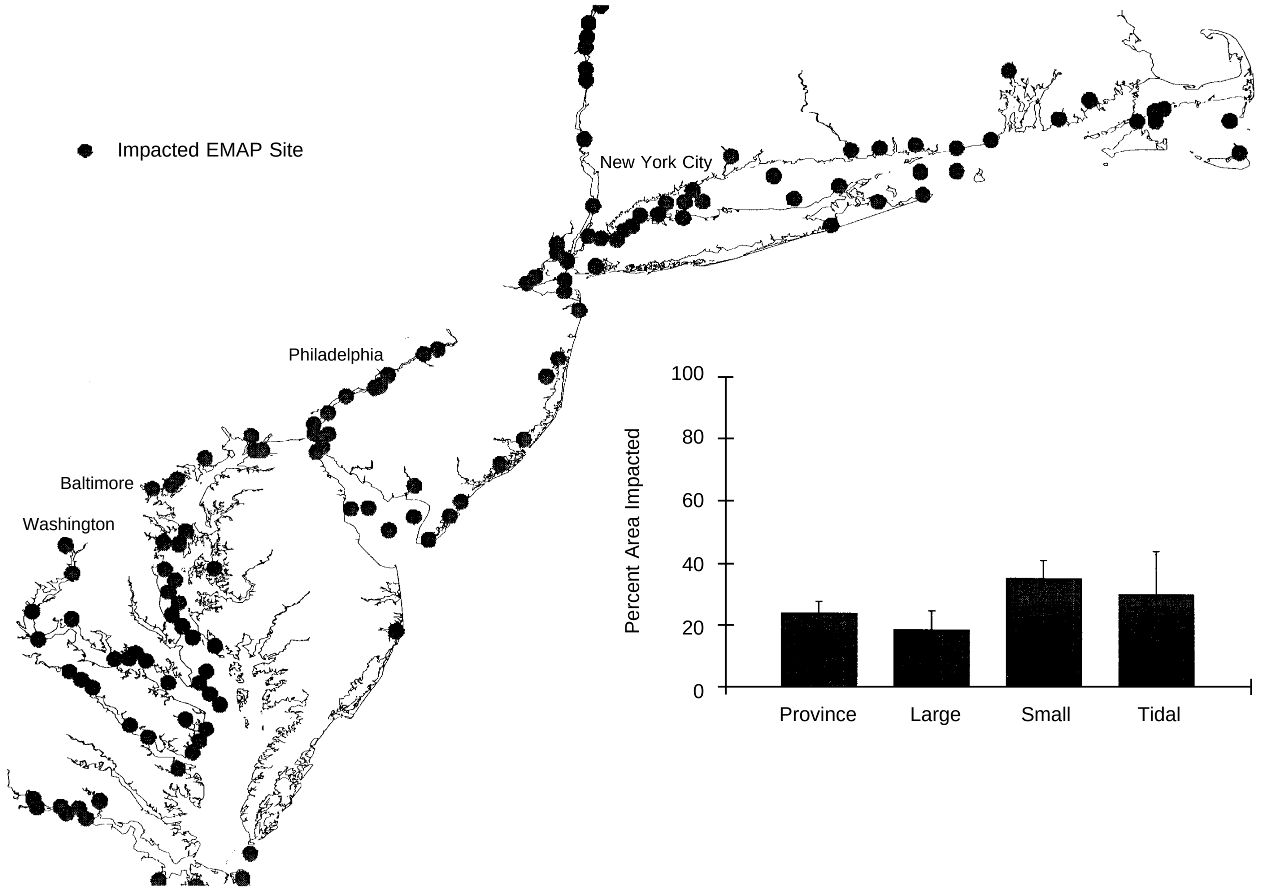


Figure 3. Condition of benthic communities in Virginian Province.

systems include the bays and harbors along the edges of the major systems and embayments along the coast; and large tidal rivers include the Potomac, James, Rappahannock, Delaware, and Hudson Rivers.

The results indicate that 24 percent ± 4 percent of the estuarine waters of the Virginian Province have impacted benthic communities. The small and tidal river systems have proportionately more impacted area than the large systems.

All the EMAP data are geographically referenced; therefore, the data can be spatially displayed to explore patterns. The spatial display of the impacted benthic community information is a simple spatial analysis of the EMAP data. This analysis shows that the impacted benthic resources are distributed across the entire province, with more impacted sites in the vicinity of the major metropolitan areas.

In addition to analyzing the EMAP results at the regional scale, analyses have been conducted at the watershed scale (see Figure 4). The probability-based sampling design permits the data to be aggregated (poststratified) in ways other than the way the original design was stratified. The only restriction to the aggregation is the number of available sample sites for the aggregation; a small number of sites leads to large uncertainties in the results. Figure 4 shows the aggregation for four major watersheds: Chesapeake Bay, Delaware Bay, Hudson-Raritan system, and Long Island Sound. This watershed scale is close to the practical scale at which environmental management decisions are implemented. The data need to be analyzed at smaller scales, however, to focus on environmental management of the smaller watersheds (e.g., contaminated sediments). This leads into the need to conduct the pilot project addressing watershed information.

South Shore Rhode Island Pilot Project

EMAP's third objective relates to exploring associations between indicators of estuarine condition and watershed stressors. Note that the word "watershed" was added. One way to address environmental management remediation strategies is to look at the watershed activities that could possibly be modified or changed to improve estuarine conditions.

Watershed stressors for the estuarine environment include land-based sources of pollution, such as point sources of pollution, and land use activities (i.e., how the land is actually used, including landscape patterns).

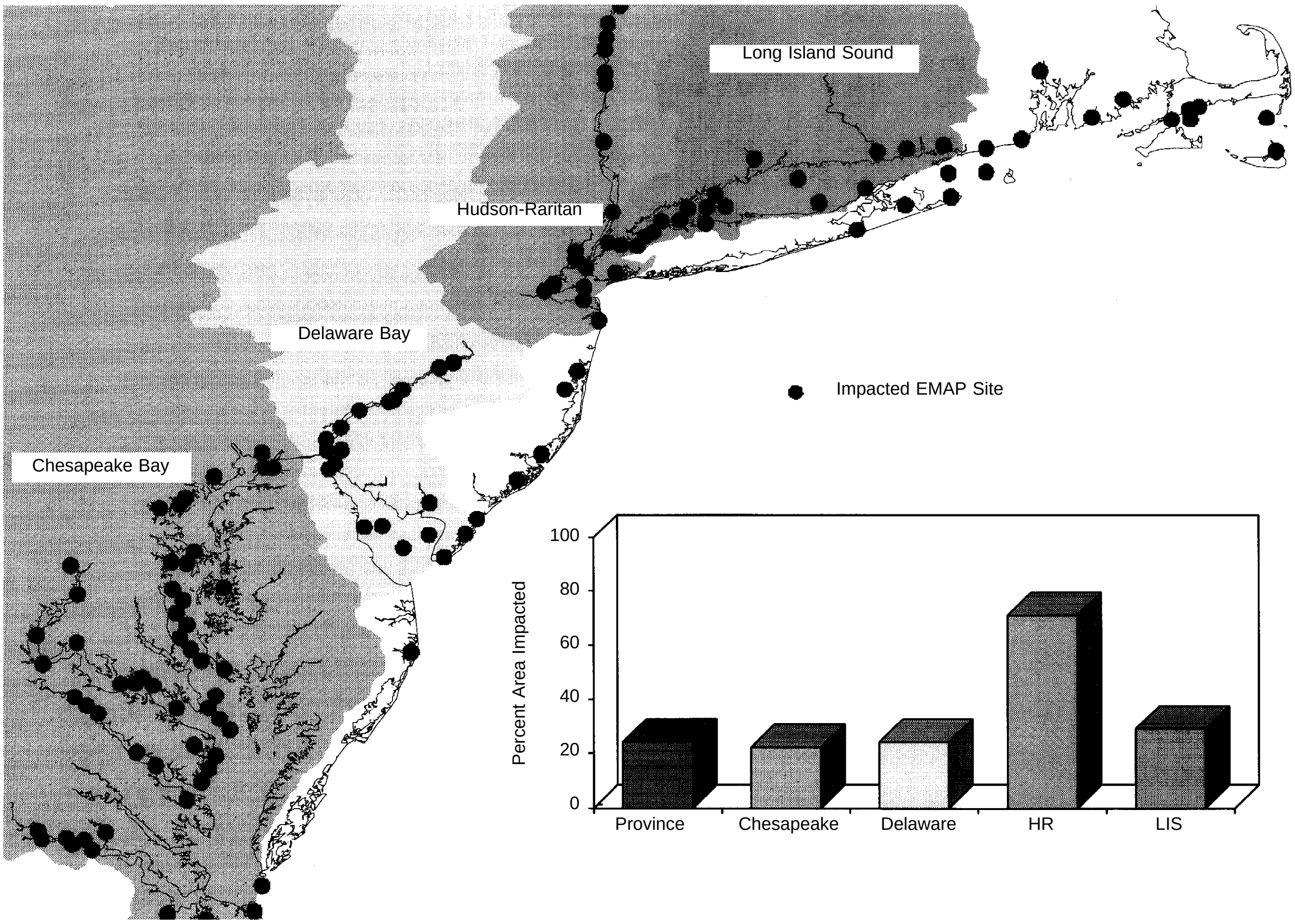


Figure 4. **Condition of benthic communities in major watersheds in Virginian Province (Chesapeake Bay, Delaware Bay, Hudson-Raritan system, and Long Island Sound).**

Which of these stressors is more important for a particular situation depends on the types of estuarine impact (localized or systemwide) and the management question that is being addressed.

The specific objective of the south shore Rhode Island pilot project was to compare watershed stressors with EMAP indicators of estuarine condition using GIS analysis tools. This project was not intended to be a definitive study by itself of south shore Rhode Island but to explore the process necessary to undertake the comparisons, to investigate the feasibility of pulling the necessary information together, and to identify potential problems before undertaking this comparison on a much larger scale.

The south shore Rhode Island study area is depicted in Figure 5. This coastal area drains into the coastal waters of Block Island Sound. The project was intentionally restricted to a limited geographic area to avoid being overwhelmed with the tremendous volumes of data that could have been encountered.

All data sources used in this project were available electronically. Digitized U.S. Geological Survey quad maps were available from the Rhode Island Geographic Information System (RIGIS) at the University of Rhode Island. The National Pollutant Discharge Elimination System (NPDES) was available for major dischargers from the National Oceanic and Atmospheric Administration's National Coastal Pollution Discharge Inventory (18). The 1990 census was also available from RIGIS. The EMAP 1990 through 1993 estuarine data were available from the EMAP-Estuaries Information System at the EPA Environmental Research Laboratory in Narragansett, Rhode Island. The RIGIS data were already available as ARC/INFO coverages. The NPDES and EMAP data had to be converted to ARC/INFO point coverages.

Two approaches were used to conduct spatial analyses. Buffer zones at 1, 3, 5, 10, and 20 kilometers from the south coast of Rhode Island were created and used to clip the south shore area coverages (e.g., land use, population, point sources). The watershed boundaries of three south shore coastal ponds were manually delineated and used to clip the south shore area coverages. The ponds were Quonochontaug, Ninigret, and Point Judith (west to east).

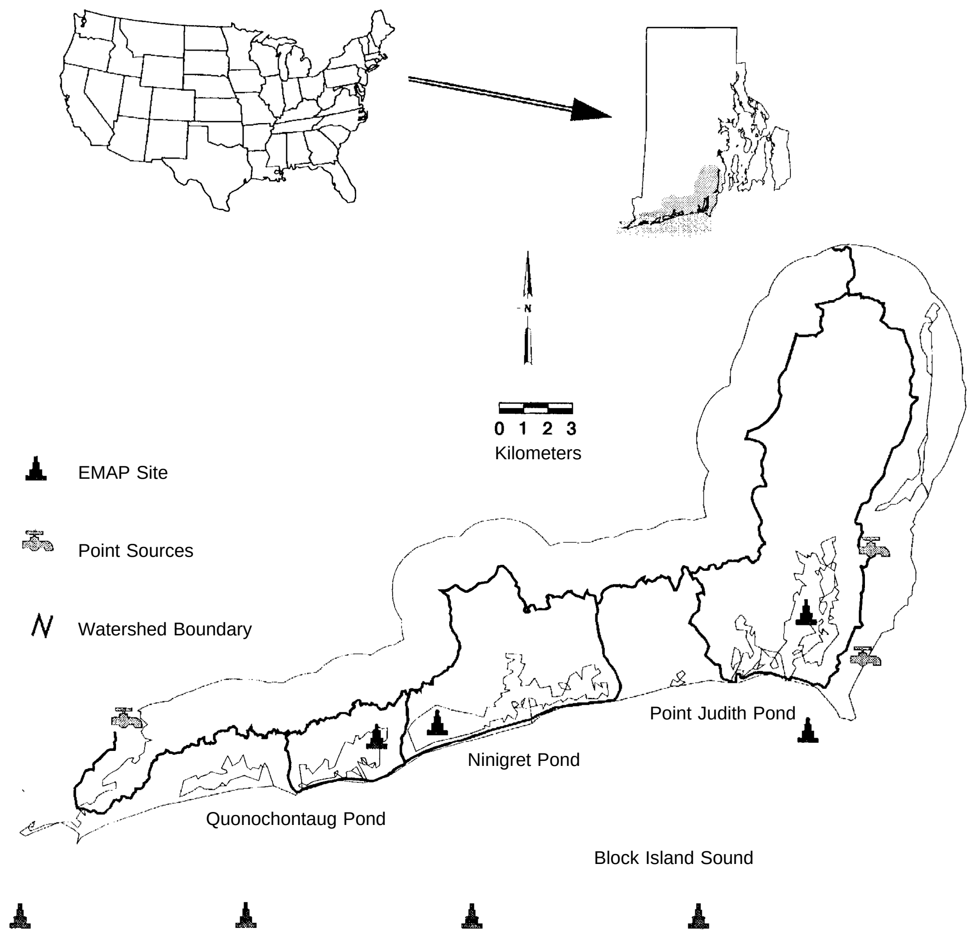


Figure 5. South shore Rhode Island study area.

South Shore Rhode Island Pilot Project Results

The results for land use by distance from the coast are presented in Figure 6. These results give the broad-scale coastal perspective. For the total south coast watershed, the major land use categories are residential and forest/brush land, followed by agriculture. Closer to the coast, residential land use is more prevalent, while farther from the coast, forests/brush lands dominate. Population (see Figure 7) increases with distance from the coast, but population density does not appear to be a function of distance from the coast. Only one out of five coastal EMAP stations exhibited impacted benthic conditions, indicating no widespread benthic problems in the coastal waters. The one station that was classified as impacted was dominated by an extremely high number of individuals of one species.

A smaller scale view can be gained by looking at the results for the individual watersheds. This view also provides an east-west perspective compared with the south-north perspective with the distance from the coast. Again, for the individual watersheds, the major land use categories are residential and forest/brush land (see Figure 8). The population increases from west to east, and population density shows an increasing trend from west to east (see Figure 9). Impacted benthic conditions were observed at the EMAP stations in Quonochontaug and Point Judith Ponds. These stations exhibited organic enrichment (total organic carbon in the sediments exceeding 2 percent), possibly from historically improperly treated sewage. No benthic data were available for the Ninigret Pond station; however, dissolved oxygen was observed to be low at this station. No major NPDES point sources are located in the coastal pond watersheds, although two are located on the eastern edge of the Point Judith Pond watershed boundary.

Discussion

A pilot project was conducted for south shore Rhode Island to compare watershed stressors with EMAP indicators of estuarine condition. The results indicate that such a comparison can be accomplished, with the watershed information providing a qualitative link to the estuarine conditions observed. One potential problem is the need to delineate the watershed boundaries for all

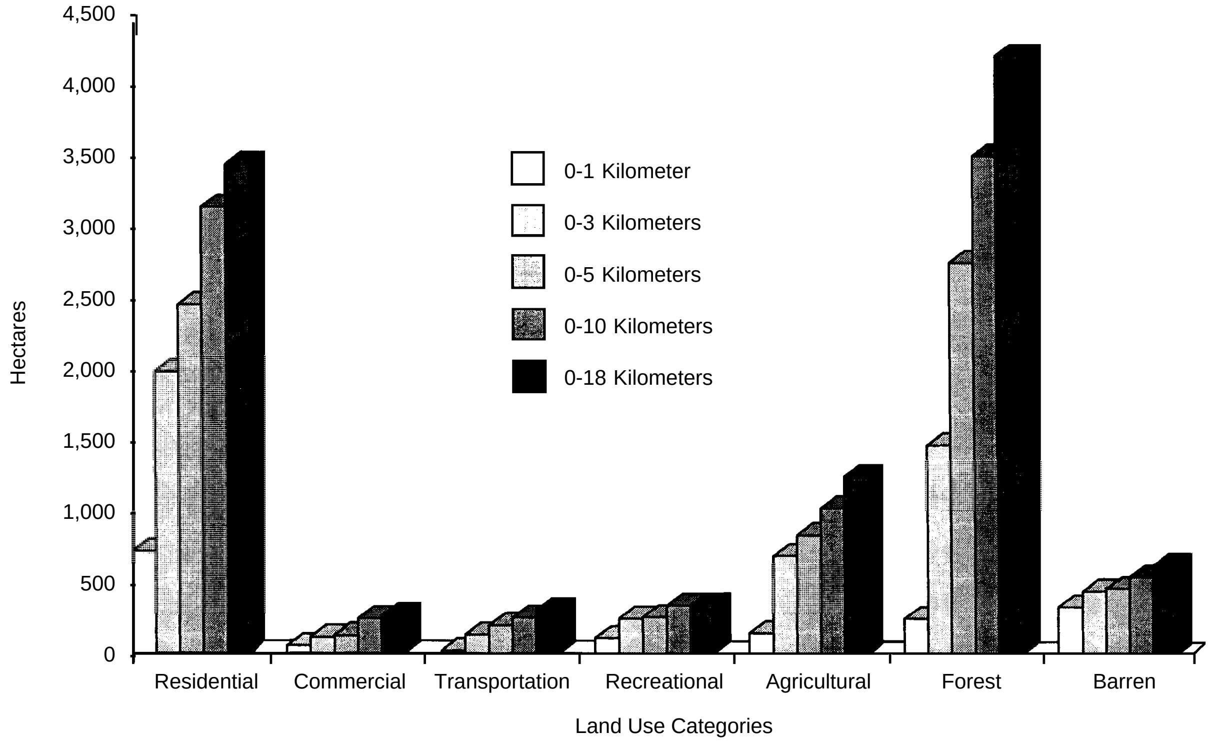


Figure 6. South shore Rhode Island land use by distance from south coast.

watersheds for which EMAP data are available. ARC/INFO provides tools for doing this, but practical application indicated that difficulties are encountered when the topographic relief is relatively flat, and on-screen corrections needed to be applied (19).

A restriction that needs to be understood before applying the procedures used in this project to a much wider geographic area is that the data sets for the watershed stressors need to be available over the wider geographic area. Further, these data sets need to be temporally consistent and constructed with consistent methods and land use classification schemes.

This project was conducted with only a small number of actual EMAP sampling sites, particularly for the individual watersheds. Because of this restriction, no statistical analyses were conducted with the EMAP data for comparison with the watershed information. Only qualitative comparisons were attempted. The next step is to increase the number of individual watersheds so that a rigorous statistical analysis can be conducted. This is being conducted by Comeleo et al. for comparing watershed stressors for subestuary watersheds in the Chesapeake Bay with estuarine sediment contamination (19). The steps after this will be to (1) apply the techniques to the entire EMAP Virginian Province data

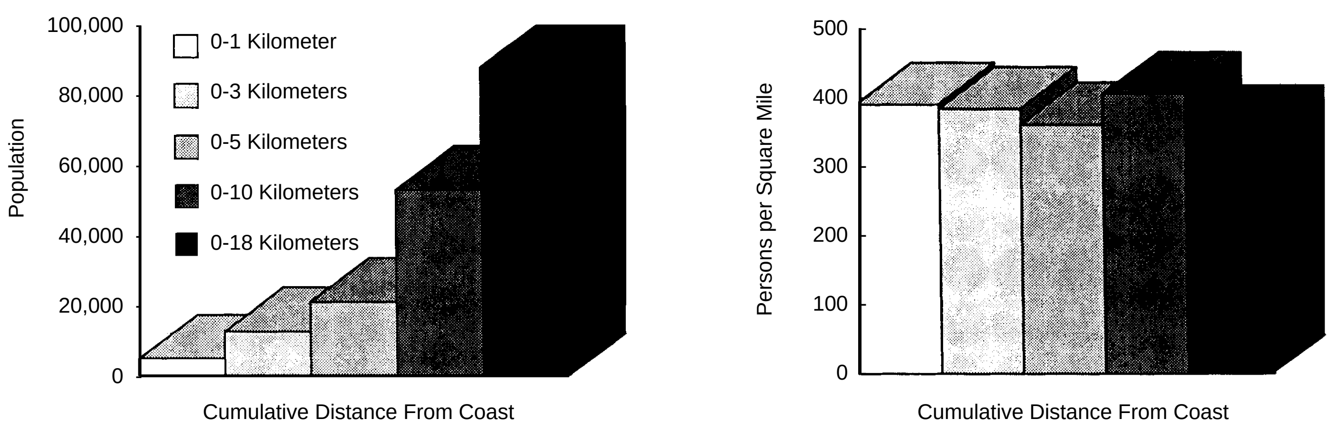


Figure 7. South shore Rhode Island population by distance from south coast.

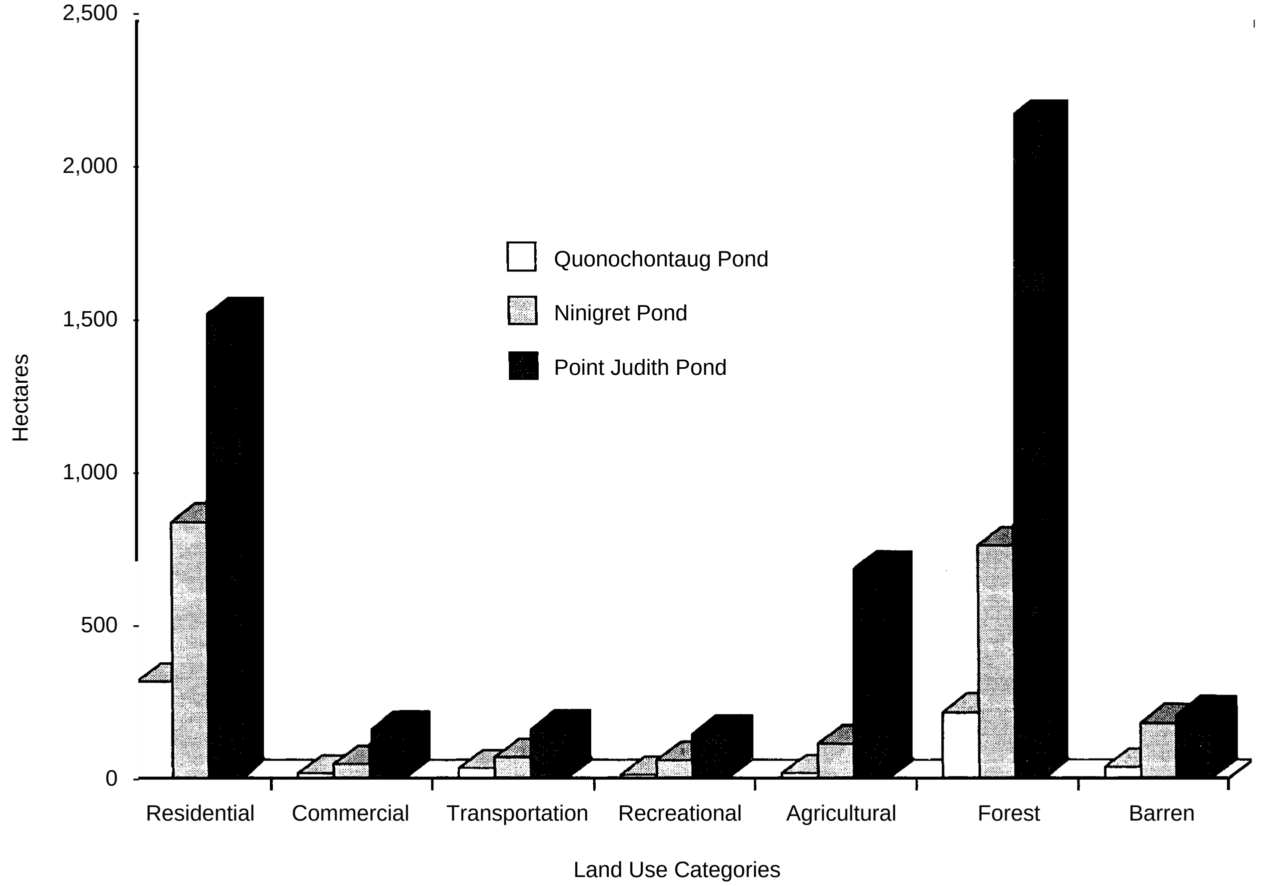


Figure 8. South shore Rhode Island land use by individual watershed.

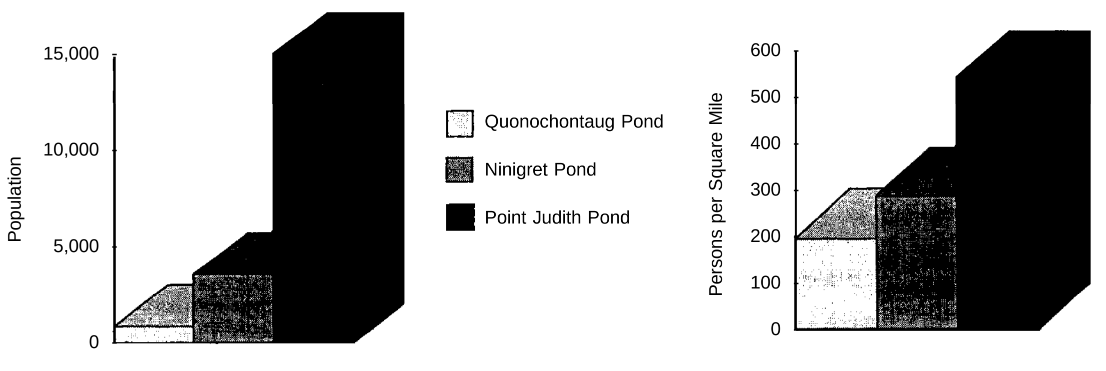


Figure 9. South shore Rhode Island population by individual watershed.

set, and (2) relate the watershed stressor information to estuarine benthic condition.

Acknowledgments

We thank S. Hale, J. Pesch, J. Copeland, and D. Keith for their review of this paper. This is contribution No. 1648 of the Environmental Research Laboratory-Narragansett.

References

1. SAB. 1988. Future risk: Research strategies for the 1990s. Science Advisory Board Research Strategies Committee, U.S. EPA, Washington, DC.

2. NRC. 1990. Managing troubled waters: The role of marine environmental monitoring. Washington, DC: National Academy Press.

3. U.S. EPA. 1993. Virginian Province demonstration project report, EMAP-Estuaries (1990). EPA/620/R-93/006. Washington, DC.

4. U.S. EPA. 1994. Statistical summary: EMAP-Estuaries Virginian Province (1991). EPA/620/R-94/005. Narragansett, RI.

5. U.S. EPA. 1994. Statistical summary: EMAP-Estuaries Virginian Province (1992). EPA/620/R-94/019. Narragansett, RI.

6. Newman, A. 1995. A new direction for EPA's Office of Research and Development. Environ. Sci. and Technol. 29(3):126A-129A.

7. U.S. EPA. 1992. Framework for ecological risk assessment. EPA/630/R-92/001. Washington, DC.

8. Gentile, J.H., J.F. Paul, K.J. Scott, and R.A. Linthurst. 1995. The role of regional monitoring in ecological risk assessment: An EMAP example for watersheds. Manuscript in review.

9. Messer, J.J., R. Linthurst, and W. Overton. 1991. An EPA program for monitoring ecological status and trends. Environ. Monitor. and Assessment 17(1):67-78.

10. Paul, J.F., K.J. Scott, A.F. Holland, S.B. Weisberg, J.K. Summers, and A. Robertson. 1992. The estuarine component of the U.S. EPA's Environmental Monitoring and Assessment Program. Chemistry and Ecology 7:93-116.

11. Bailey, R.G. 1983. Delineation of ecosystem regions. Environ. Mgmt. 7:365-373.

12. Omernik, J.M. 1987. Ecoregions of the conterminous United States. Annals of the Assoc. of Amer. Geographers 77:118-125.

13. U.S. EPA. 1990. Near coastal program plan for 1990: Estuaries. EPA/600/4-90/033. Narragansett, RI.

14. U.S. EPA. 1991. Design report of the Environmental Monitoring and Assessment Program. EPA/600/3-91/053. Chervils, OR.

15. U.S. EPA. 1990. Environmental Monitoring and Assessment Program: Ecological indicators. EPA/600/3-90/060. Washington, DC.

16. U.S. EPA. EMAP-Virginian Province four-year assessment (1990-1993). EPA report in review.

17. Engle, V.D., J.K. Summers, and G.R. Gaston. 1994. A benthic index of environmental condition of Gulf of Mexico estuaries. Estuaries 17(2):372-384.

18. NOAA. 1993. Point-source methods document. NOAA technical report. Silver Spring, MD: Strategic Environmental Assessments Division, National Ocean Service, National Oceanic and Atmospheric Administration.

19. Comeleo, R.L., J.F. Paul, P.V. August, J. Copeland, C. Baker, S. Hale, and R.W. Latimer. 1995. Relationships between landscape stressors and sediment contamination in Chesapeake Bay estuaries. Manuscript in review.

Nonpoint Source Pesticide Pollution of the Pequa Creek Watershed, Lancaster County, Pennsylvania: An Approach Linking Probabilistic Transport Modeling and GIS

Robert T. Paulsen
The Paulsen Group, Bowie, Maryland

Allan Moose
Southampton College, Southampton, New York

Abstract

The U.S. Environmental Protection Agency (EPA) has mandated that each state prepare a state management plan (SMP) to manage pesticide residues in the state's environment. One aspect of an SMP involves identifying specific soils and sites that may be vulnerable to the transport of pesticides into water resources. A recently developed system identifies vulnerable areas by coupling probabilistic modeling that uses the Pesticide Root Zone Model (PRZM) with a desktop geographic information system (GIS-MAPINFO). A limited test of this system succeeded in identifying and mapping individual soil series in a watershed that were shown to have transported atrazine to surface and ground water.

During this project, various digital data sources were evaluated for availability and ease of use, including:

- STATSGO.
- U.S. Geological Survey (USGS) digital line graphs (DLGs).
- National Oceanic and Atmospheric Administration (NOAA) climate data.

This study documents hands-on hints and tricks for importing and using these data.

From 1977 to 1979, the USGS measured the movement of atrazine off fields of application into water resources in the Pequa Creek basin in Lancaster County, Pennsylvania (1). Atrazine in surface water appeared at levels exceeding 20 parts per billion in storm flow and above the 3 parts per billion maximum contaminant level (MCL) during base flow from Big Beaver Creek, a tributary to Pequa Creek. Each soil series in the subbasin was digitized into a GIS. PRZM allowed simulation of runoff, erosion, and leaching of atrazine (applied at 2.24 kilograms per hectare in conventionally tilled corn) for each soil. This process included simulating each soil under different slopes for an 11-year period from 1970 to 1980. Interpreting the results for each soil series determined the probability distribution of atrazine in kilograms per hectare for each mode of transport. GIS used these data to thematically map each soil series for atrazine loss.

The results of this demonstration project suggest that the Manor silt loam, with slopes varying from 6 percent to 20 percent, had a high potential to transport atrazine residues to surface water. This type of analysis could suggest that this soil series be:

- Farmed using conservation tillage.
- Managed to install grass waterways or buffer strips to stop runoff.
- Set aside from production to protect water resources.

Digital databases were available for the study area, but many technical problems were encountered in using the data. Researchers embarking on these types of modeling and GIS projects should prepare themselves for significant expenditures of time and finances.

Introduction

A significant volume of published literature documents pesticide residues in ground water, and the volume of investigations of residues in surface water is expanding. The growing acceptance of immunoassay techniques for the determination of pesticide residues in water has given the field of pesticide monitoring an accurate and economical analytic methodology. This will result in an increase in monitoring capability at the federal, state, local, and university levels. These increases in monitoring capability have documented and will continue to document the occurrence of pesticides in water resources as the result of past transport through the soil profile. The U.S. Environmental Protection Agency (EPA) has mandated that each state prepare a state management plan (SMP) to manage pesticide residues in the state's environment. Lacking, however, is a reliable pesticide screening technique to indicate which soils, on a countywide scale, may be sensitive to the transport of a specific pesticide to deep within the soil profile or to the surface water resources. These assessments would greatly supplement the usability and validity of SMPs.

Electronic databases such as State Soils Geographic (STATSGO), Data Base Analyzer and Parameter Estimator (DBAPE), or the SOIL5 subsets found in Nitrate Leaching and Economic Analysis Package (NLEAP) provide easy access to detailed soil data and model input estimator subroutines, thereby simplifying data entry to numerical models. Two groupings define soils: soil series and soil associations. Soil series are the individual soil taxa found in a field. Soil associations represent groups of soil series, usually three or four soil series occurring together in an area, and are mapped as a single unit on a county scale. Mapping of most soil associations across the United States is complete, with open access to the county scale maps. A digital soils mapping data set called SSURGO contains many of the soil series maps for the United States. Climatologic databases also provide easy access to long-term data from the National Oceanic and Atmospheric Administration (NOAA) weather stations, allowing a user the opportunity to input realistic climate data to pesticide transport models.

Many numerical pesticide transport models, such as the Pesticide Root Zone Model (PRZM), Ground-Water Leaching Effects of Agricultural Management Systems (GLEAMS), and Leaching Estimation and Chemistry Model (LEACHM), can produce transport estimates for specific pesticides in specific soils. Each model has its own strengths and weaknesses, and detailing these characteristics is beyond the scope of this paper. Several authors, however, have described comparisons between models (e.g., Smith et al. [2], Mueller et al. [3], and Pennell et al. [4]). These numerical models all generally require extensive site-specific soil, agronomic, and climatologic databases. The results from these models are extremely detailed. Their pesticide transport estimates, however, are only valid for those locations for which site-specific data are sufficient to allow calibration of the model. Applying such site-calibrated model results to larger scales (county scales) is inappropriate.

In one procedure, a user could identify soils and use transport models that may have a limited ability to retard pesticides from reaching water resources. This type of modeling has recently been called probabilistic modeling (5). The concept behind this procedure is to use an existing transport model, such as PRZM, and vary certain input parameters (e.g., slope, organic content, pesticide K_{oc}) to produce a probability of a given output being equaled or exceeded. For example, a PRZM model could be created for a soil series with an average organic content of 1 percent and a slope of 8 percent in the eastern corn belt. The model would use the 30 years of historical climate data for a nearby station. The model would vary the organic content and surface slope within given ranges for the soil series and run 1,000 simulations. The analysis could then entail plotting the results (i.e., monthly runoff loads, erosion loads, and leaching through the root zone) in a frequency diagram and generating probability curves. This analysis would allow the user to estimate the anticipated pesticide losses, runoff, erosion, and leaching for any given soil in the county. The soils with greater probabilities for pesticide loss could be identified and mapped using GIS.

Recent advances in computing speed and efficiency have reduced the amount of time and expense needed to run numerical pesticide transport models. This makes it possible, in a relatively short amount of time, to quantitatively model not just one soil series, but the hundreds of major soil series that occur in an entire state (e.g., 357 major soil series combined into 464 different soil associations in Wisconsin). This type of model can be very useful to the development of SMPs as well as to a variety of users, including pesticide registrants, bulk pesticide handlers, custom appliers, county agricultural extension agents, and individual growers.

Objective

The objective of this study was to use probabilistic modeling analyses and a geographic information system (GIS) to determine which soil series in a watershed may contribute to nonpoint source pollution through runoff of agricultural chemicals. Specifically, the study aimed to locate a watershed with historical atrazine runoff, map the soils, and perform transport modeling using historical precipitation. The results of this procedure would determine which soil series had a high potential to contribute to the nonpoint source pollution of the watershed. Once the transport modeling was completed, a GIS would help map the distribution of the sensitive soil series. The mapping would act as a base for implementing best management practices (BMPs) to reduce nonpoint source pollution.

Background

Ward (1) described the water quality in the Pequa Creek basin in Lancaster County, Pennsylvania, for the years 1977 through 1979. Flow from Pequa Creek (154-square-

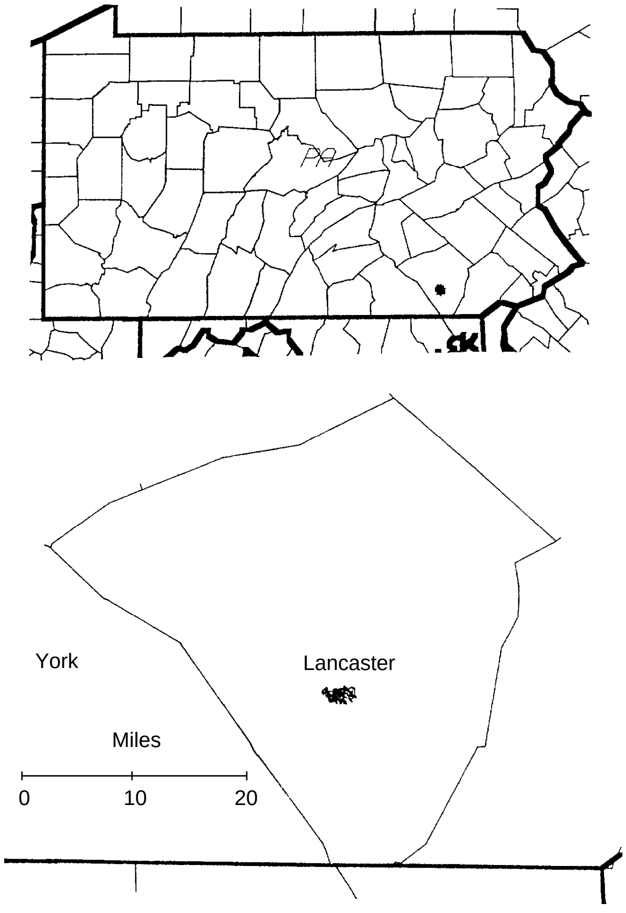


Figure 1. Location of Pequa Creek basin, Lancaster County, Pennsylvania.

mile drainage area) eventually discharges into the Chesapeake Bay (see Figure 1). The data collection efforts (6) documented the occurrence of atrazine (2-chloro-4-ethylamino-6-isopropylamino-s-triazine), a commonly used herbicide for weed control in corn-growing regions, and other agrichemicals in both base-flow and storm-flow conditions of the Pequa Creek. A subbasin of Pequa Creek, Big Beaver Creek, had the greatest reported atrazine concentrations during the sampling period. The maximum reported atrazine concentrations at the Big Beaver Creek sampling station, near Refton, Pennsylvania, were 0.30 parts per billion during base-flow conditions and 24.0 parts per billion during storm-flow conditions. The Big Beaver Creek basin is 20.4 square miles in area, and agriculture constituted about 66 percent of the land use in 1979. Corn was grown on 26.6 percent of the agricultural lands in this subbasin. The average rainfall for this basin is about 37 inches annually (1).

As noted, agriculture represented the major land use in the area. The primary agricultural soils in Lancaster County are silt loams (Typic Hapludults and Hapludalfs) in texture with slopes that range from 0 percent to 8 percent (7). Upon inspection of the air-photo soil series maps found in the county soil survey, however, agricultural crops grew on lands with slopes of up to and exceeding 15 percent, with soils such as the Manor silt loam, Pequa silt loam, and Chester silt loam (7).

Natural soil organic contents in the agricultural soils range from 0.1 percent to 2.0 percent. Water contents of the agricultural soils range from 10 percent to nearly 30 percent. The soil Erosion Factor (K) for the surface layer ranges from 0.17 for the relatively stable Ungers Series to 0.43 for the Pequa Series. The greater the value, the greater the susceptibility to sheet and rill erosion. The soil Erosion Factor (T in tons/acre/year) for the entire soil profile ranges from 2 for the relatively stable Clarksburg Series to 5 for the Elk Series.

Methods

Determining which soil series in the Big Beaver watershed may contribute to nonpoint source pollution through runoff of agricultural chemicals entailed performing a combination of probabilistic modeling analyses and a GIS data manipulation.

Physiographic and Soil Series Boundaries

The orientation of Pequa Creek, Big Beaver Creek, and other surface water bodies was digitized directly from the 1:50,000-scale county topographic map for Lancaster County (8). This map also provided the basis for digitizing the Pequa Creek drainage divide, location of urban areas, and roadways. The MAPINFO GIS allows for the creation of boundary files by tracing the boundary off the topographic map with a digitizing tablet configured to the latitude and longitude coordinates of three points on the map. The latitude and longitude are displayed while the boundary is being traced, allowing the user to verify the accuracy of the boundary against known coordinates on the map. GIS contains a self-checking boundary closure program to ensure that the polygons are closed and that the boundary contains no extraneous line segments. These boundary data are already available from the USGS in a digitized format, digital line graphs (DLGs). Because digitizing is an easy task, however, and to minimize costs, the project used manual digitizing rather than purchase the data.

The roadways and urban areas were digitized to allow use of standard control points, such as road intersections and benchmarks, to configure the U.S. Department of Agriculture (USDA) Soil Conservation Service air-photo-based, 1:15,840-scale soil series maps for the Big Beaver Creek watershed. Known land grid coordinates were noted on the air-photo maps (7). We concluded, however, that using known reference points, such as roadways and towns, allowed for a better configuration of the digitizing tablet to the air photos and eliminated concerns over scale distortions sometimes common in air photos.

After configuring the air photos to the digitizing tablet, a 2-square-mile area around the surface water sampling points was digitized. The next step entailed digitizing all the mapped soil series units within this area. The locations of crop areas, as plowed fields, were noted. In noting forested areas, it became apparent that only minor acreages were not in agricultural production. Those mapped units were generally the Manor and Pequa Series soils with slopes exceeding 25 percent.

Pesticide Transport Modeling

The PRZM pesticide transport model helped to quantify the ability of several soil series to retard the transport of atrazine through the root zone as leachate, dissolved in surface runoff and adsorbed on sediment that moved during erosion. The PRZM model performed in an uncalibrated or screening model mode. The input values for soil properties came from both the EPA DBAPE database and the Lancaster County Soil Survey (7). The modeled soil profile was 150 centimeters thick and divided into 5-centimeter compartments. The soil half-life of atrazine was set at 57 days in accordance with values that the PRZM manual listed (9). The primary soil property that varied in this demonstration project was surface slope. All other parameters, such as soil organic content, moisture content, and bulk density, appeared as midpoint values for the ranges listed in DBAPE.

The agronomic scenario that the model simulated was for corn grown continuously for 10 years using conventional tillage practices and planted on May 7 of each year. Atrazine was surface applied at a rate of 2 pounds per acre (2.24 kilograms per hectare) on May 1 of each year. For climatic input, the model used the historical precipitation regimen from 1970 through 1980, as measured at the Harrisburg, Pennsylvania, station.

PRZM simulations were made for each of the following soil series:

- Chester
- Conestoga
- Elk
- Glenelg
- Glenville
- Hollinger
- Letort
- Manor
- Pequa

Monthly values were calculated for leachate, runoff, and erosion per hectare. Unfortunately, no data for the Pequa Series were available in the DBAPE database; therefore, this portion of the analyses omitted it. In addition, analyses of the Manor Soil Series included more detailed probabilistic modeling where the surface slope held constant (6-percent slope) and the surface soil organic content varied to include the high, average, and low organic contents as listed in DBAPE. PRZM also calculated the volume of water as evapotranspiration, runoff, and recharge through the root zone.

Results

The results of this study should demonstrate the application of transport modeling to the possible protection of water resources. Regulatory decision-makers should not consider these results in their current form because such decisions would require a much more rigorous simulation strategy to increase the level of confidence in the data. As a demonstration study, however, the results do show the usefulness of this approach. Table 1 contains the cumulative frequency data for the simulated atrazine residues in runoff, erosion, and leaching that occurred under 30 years of historical precipitation. The data cover the 12 soils mentioned, with the surface slope held constant at 6 percent.

Atrazine in Runoff and Erosion

The results of this analysis suggest that the Haggerstown Series had the greatest potential for yielding atrazine in runoff; approximately 50 percent of the simulated monthly atrazine in runoff values equaled or exceeded 0.0001 kilogram per hectare. Conversely, the Elk Series yielded the least atrazine to runoff; 50 percent of the runoff data were at residue levels of 1×10^{-6} kilograms per hectare. Within the Big Beaver Creek subbasin, the Manor Series had the greatest potential to yield atrazine in runoff.

As with the runoff data, the Haggerstown Series had the greatest potential to yield atrazine in eroded sediments, and the Elk Series yielded the least atrazine in erosion. Within the Big Beaver Creek subbasin, the Manor Series had the greatest erosion potential regarding atrazine.

GIS Analyses

After entering the results from the transport modeling into a database, GIS could produce maps showing the location of soils with high runoff potentials. Figure 2a shows the orientation of soil series around the surface water sampling points in Big Beaver Creek. Figure 2b represents the same scene but fills in the soils with high runoff potential. Using this type of analysis can help areas that may be sources of nonpoint source runoff contamination. Once identified, these soils can be targeted for alternative management practices that may reduce the amount of runoff and the degree of nonpoint source contamination.

Table 1. Cumulative Frequency of Simulated Atrazine Residues in Runoff for 12 Major Soils in Lancaster County, Pennsylvania (values are percentage of data)

Load (kilograms per hectare)	Soil Series in Lancaster County, Pennsylvania											
	Bucks	Chester	Clymer	Connestoga	Elk	Glenelg	Haggerstown	Hollinger	Lansdale	Letort	Manor	Ungers
IE-10	83.33	84.85	84.09	84.08	78.79	84.85	93.94	83.33	88.64	85.61	86.36	84.09
IE-9	78.79	80.30	80.30	81.82	75.76	80.30	93.94	81.06	85.61	84.85	84.61	84.09
IE-8	74.24	76.52	75.00	78.03	69.70	76.52	93.94	77.27	80.30	84.09	84.85	81.82
IE-7	66.67	68.94	66.67	72.73	59.09	68.18	91.67	68.94	75.00	83.33	75.52	75.76
IE-6	56.06	46.97	56.06	61.36	50.00	56.06	84.85	59.09	65.15	76.52	66.67	67.42
IE-5	47.78	35.61	46.21	50.00	43.18	47.73	74.24	47.73	51.52	67.42	46.97	53.03
IE-4	36.36	21.97	32.58	35.61	31.82	34.85	53.79	34.85	40.15	46.97	31.82	39.39
IE-3	24.24	15.15	21.21	21.21	21.21	21.21	37.12	21.21	24.24	31.82	16.67	24.24
0.00	15.91	6.82	14.39	15.12	15.15	13.64	19.70	15.15	15.91	17.24	4.55	15.15
0.01	6.82	0.00	6.06	6.06	6.82	3.79	7.58	6.06	6.82	4.55	0.00	4.55
0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

This table reads as follows: Given the Elk Series, the first value reads that 78.79 percent of the simulated data were greater than or equal to IE-10 kilograms per hectare. Similarly, within the same column, 6.82 percent of the data were greater than 0.01 kilograms per hectare but less than 0.1 kilograms per hectare.

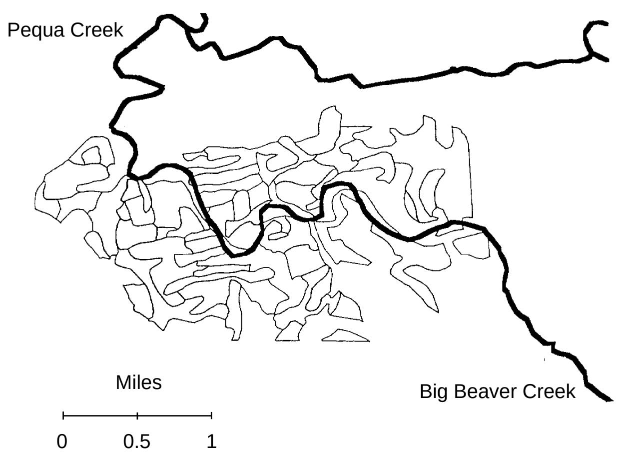


Figure 2a. Soil series in the Big Beaver Creek basin.

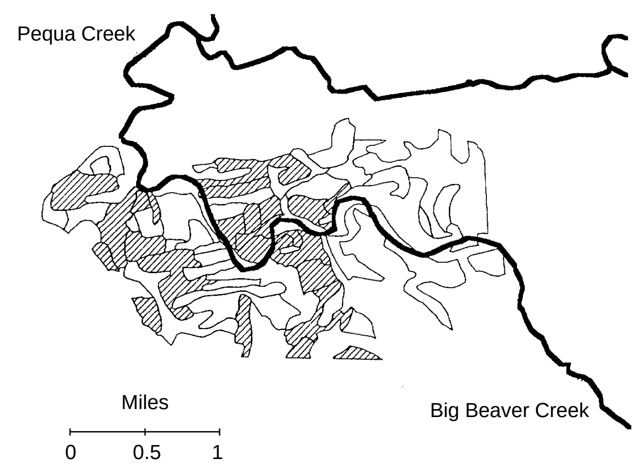


Figure 2b. Soils sensitive to atrazine runoff in the Big Beaver Creek basin.

Detailed Modeling of the Manor Series

Performing an introductory probabilistic modeling exercise allowed further investigation of the potential of the Manor Series to release atrazine into runoff. The existing 30-year climate data and the stated agronomic data were retained from the previous modeling. The organic carbon content of the surface soil layer, however, was allowed to vary between the published low, average, and maximum values found in the DBAPE database. This exercise followed the principles set forth by Laskowski et al. (5) and others who describe probabilistic modeling approaches. In essence, by varying input parameters within known endpoints, the probabilistic approach can generate a distribution of pesticide residue values that statistically reflects the anticipated residues. Parameters to vary may include:

- Organic carbon content
- Surface slope
- Kd (distribution coefficient)

- Moisture content

By allowing input variables to vary according to a normal distribution, this approach thereby eliminates some of the uncertainty associated with pesticide transport modeling. The probabilistic modeling approach requires the creation of a significant database by performing many runs (e.g., 1,000 model runs that generate 12,000 monthly values for each soil).

This study included a limited probabilistic modeling exercise. Table 2 lists the results for the mean atrazine residues in runoff, erosion, and leaching for the Manor Soil Series during the:

- Entire year
- Growing season
- Winter months

The surface slope was held constant at 6 percent, but the soil organic carbon content varied within the published range. The means for all months show limited variation in mean residues. Runoff was by far the major source of atrazine. Erosion and leaching values were on similar scales (trace amounts).

Table 2. Summary Statistics for Detailed Modeling of the Manor Soil Series in the Pequa Creek Watershed (statistics based on 1,080 values)

	Mean Atrazine Residue (kilograms per hectare)		
Percent Organic Carbon[a]	**Runoff**	**Erosion**	**Leaching**
All Months			
Low	0.01381	0.00024	0.00028
Average	0.01229	0.00042	0.00012
High	0.01105	0.00057	< 0.000006
Growing Season			
Low	0.03290	0.00058	0.00017
Average	0.02881	0.00100	< 0.000008
High	0.02540	0.00130	< 0.000004
Winter Months			
Low	0.00014	< 0.000002	0.00036
Average	0.00045	< 0.000001	0.00014
High	0.00075	< 0.000003	< 0.000007

[a]Data taken from DBAPE soils database as low, midpoint, and maximum reported organic contents.

The greatest runoff and erosion values occurred during the growing season. The greatest leaching, however, occurred during the winter months. These results support the general observations that surface residues run off during spring and summer but that as the crops grow and evapotranspiration increases, recharge to ground water decreases, subsequently limiting pesticide transport to ground water. Conversely, during the winter months, the surface soil pesticide residues generally decrease because of exposure to months of photolysis, hydrolysis, and biodegradation. Subsurface residues have been protected from degradation, however, and increased ground-water recharge, due to great reductions in evapotranspiration, transports the residues through the soil column.

This limited exercise provided a valuable learning experience regarding probabilistic modeling. As computing techniques and hardware advance, the cost in time and money for each simulation should decrease dramatically. Although researchers tend not to have great faith in pesticide transport modeling, the advances in this field will reduce uncertainty and instill greater confidence in the modeling process.

GIS Pitfalls

GIS is a powerful tool and has great promise for use in environmental problem-solving. Several points or pitfalls, however, hinder broad acceptance of GIS. As with most new technologies, cost is the overriding concern in using GIS. Although technical staff and project scientists understand the power of GIS and the effort that data preparation requires, management and corporate staff often do not see the benefits for the costs. Many managers assume that current GIS systems resemble those seen on “Star Trek,” and when reality becomes apparent, managers tend to discard GIS as too costly and complex. Several points need consideration when contemplating the use of GIS. Although various products exist, this discussion focuses on ARC/INFO and MAPINFO products.

Hardware

Computer hardware is plentiful if the available budget can support a purchase. Many high- powered GIS packages (e.g., GRASS, ARC/INFO, INTERGRAPH, IDRISI) run best on mainframes or minicomputers. Most technical staff, however, only have access to PC machines. Corporate purchasing departments more readily expend funds for PC technology because they will eventually find use for these machines even if they are not used for GIS. A recent ARC/INFO advertisement (August 1994) lists costs for SUN SPARC minicomputer systems with ARC/INFO software at $12,000 to $15,000 depending on configuration.

Minicomputers and mainframes require specialized staff to configure and maintain the hardware. Today, many staff level personnel can open and augment their PC machines with a minimum of external support. GIS performance reflects the tradeoff in hardware, particularly when a considerable amount of data manipulation is required. For example, if linking discrete depth soil series data to STATSGO soil associations is necessary, then a minicomputer system may be best. The postprocessed data could, however, be exported to a format that will run on PC-based systems. If the user wants to import and manipulate remote sensing imaging (e.g., SPOT or Landsat data), then minicomputers are recommended. If the user wants to display already edited images and preprocessed GIS data, then PC-based computing may be sufficient. The ultimate use of GIS drives the hardware selection.

Software

A great number of GIS software packages are available to meet almost any level of use and expertise. Software runs under both UNIX and DOS/Windows (denoted as DOS for the remainder of this paper) operating systems. The UNIX-based software tends to be more powerful and flexible than the DOS-based software. UNIX-based packages require more specialized staff to optimize GIS, however.

UNIX-based software packages include GRASS, ARC/INFO, INTERGRAPH, and IDRISI. Costs vary from public domain charges for GRASS and IDRISI to vendor supplied ARC/INFO and INTERGRAPH, which can cost several thousand dollars each.

DOS versions of ARC/INFO (e.g., PC ARC/INFO, ARCAD, ARC/VIEW) are also available and provide the user with various levels of data editing and manipulation abilities. Generally, PC ARC/INFO is the same as the UNIX version, varying in speed of processing. ARCAD is a GIS engine that uses AutoCAD for drawing and displaying, giving the user most of the abilities of the UNIX-based version. ARC/VIEW I was primarily a display and simple analysis tool. It allowed the user to view, display, and manipulate existing arc data but did not support image editing. Currently, ARC/VIEW II provides more support for image editing and data manipulation. Costs range from about $3,000 for PC ARC/INFO and ARCAD (AutoCAD also costs about $2,000) to around $500 for the ARC/VIEW products.

MAPINFO is a DOS-based GIS that was designed for marketing and demographic applications. Several researchers, however, have used MAPINFO for environmental applications. The most outstanding feature of MAPINFO is that it easily imports data layers as it reads dBASE type files directly. MAPINFO V3 also reads database files and recreates them as *.TAB files. In contrast to the “coverage and entity” concepts of the ARC/INFO line of programs, MAPINFO reads latitude and longitude coordinates and displays the results. This simplifies data management because many researchers who have already created custom databases can easily import those data as long as latitude and longitudes coordinates are present.

As with the ARC/INFO line of programs, many common data layers can be purchased for use in MAPINFO. These layers can be expensively priced, costing approximately $1,000 per county for roadway, census, and demographic data. One major lapse is the poor library of environmental layers, USGS topography, hydrography, soil boundaries, or climate stations. MAPINFO does sell a module that allows users to convert to and from ARC/INFO coverages so that common data layers can be established. Experience shows, however, that conversion programs do not always work as advertised. For example, large boundary files (STATSGO data for Indiana) do not readily convert from ARC/INFO to MAPINFO. Third-party vendors may be needed to convert data for use in MAPINFO.

One very important factor supporting the use of MAPINFO is that it has a business application slant; therefore, it is slightly easier to convince corporate management to invest in GIS because marketing and sales data (territories) can be relatively easily overlain onto environmental data.

Finally, some packages that are add-ons to spreadsheet programs tend not to be powerful or versatile enough for use in environmental GIS work. These software packages may be valuable as an introduction to GIS techniques, however.

Data Availability and Format

After compiling the hardware and software into GIS, the next step entails accessing data layers such as:

- State and county boundaries
- Land use covers
- Water boundaries

Currently, USGS DLGs for hydrography, land use, transportation, and cultural features are available for minimal costs. Shareware programs can convert the USGS DLG3 formats into DXF (data transfer files) for import to GIS packages. These data require conversion to DXF or ARC coverage type formats for use in either ARC/INFO or MAPINFO.

The USDA Soil Conservation Service produces digital data for soil types (STATSGO and SSURGO) that users can import to ARC/INFO relatively easily. The STATSGO data cost approximately $1,000 per state and are available for most states. The detailed soil series maps, SSURGO, cost approximately $500 per county and are not available for every county in the United States. Many data layers are available for direct use by GRASS. As of yet, however, no convenient conversion utilities exist to move GRASS data to ARC/INFO or MAPINFO. The

U.S. Fish and Wildlife Service now distributes data layers from the National Wetlands Inventory on the Internet (enterprise.nwi.fws.gov).

Other data sources available through private vendors are listed in the MAPINFO and ARC/INFO user guides and in any issue of *GIS World*. The user should be prepared to absorb significant costs if purchasing all the required data layers.

Conclusion

This study shows that the technology and software exist for a water resource manager to couple pesticide transport modeling with GIS to identify areas or individual soils that may contribute to nonpoint source pollution of water resources. The study used PC-based computing system and software. Soils maps and hydrographic maps can be easily digitized for limited cost. A skilled scientist or technician, without being a GIS expert, can run GIS to answer specific questions. The technologies this study demonstrated may be extremely valuable to managers responsible for producing SMPs.

The pesticide transport modeling performed during this study was intended for illustrative purposes. More detailed analyses, and additional simulations, would be necessary to use these data for regulatory actions or land use management. The study did, however, succeed in identifying the Manor Soil Series, with slopes exceeding 6 percent, within the Big Beaver Creek subbasin of the Pequa Creek basin, as a potential source for atrazine in runoff.

References

1. Ward, J.R. 1987. Surface-water quality in Pequa Creek Basin, Pennsylvania, 1977-1979. U.S. Geological Survey. Water Resources Investigation Report 85-4250.
2. Smith, M.C., A.B. Bottcher, K.L. Campbell, and D.L. Thomas. 1991. Field testing and comparison of the PRZM and GLEAMS models. Trans. Amer. Soc. of Agric. Eng. 34(3):838-847.
3. Mueller, T.C., R.E. Jones, P.B. Bush, and P.A. Banks. 1992. Comparison of PRZM and GLEAMS computer model predictions with field data for alachlor, metribuzin, and norflurazon leaching. Environ. Toxicol. and Chem. 11:427-436.
4. Pennell, K.D., A.G. Hornsby, R.E. Jessup, and P.S.C. Rao. 1990. Evaluation of five simulation models for predicting aldicarb and bromide behavior under field conditions. Water Resour. Res. 26(11):2,679-2,693.
5. Laskowski, D.A., P.M. Tillotson, D.D. Fontaine, and E.J. Martin. 1990. Probability modeling. Phil. Trans. Royal Soc. of London 329:383-389.
6. Ward, J.R., and D.A. Eckhardt. 1979. Nonpoint-source discharges in Pequa Creek Basin, Pennsylvania, 1977. U.S. Geological Survey. Water Resources Investigation Report 79-88.
7. Custer, B.H. 1985. Soil survey of Lancaster County, Pennsylvania. Washington, DC: U.S. Department of Agriculture, Soil Conservation Service.
8. U.S. Geological Survey. 1977. Lancaster County, Pennsylvania: County map series (topographic) 1:50,000 scale.
9. Carsel, R.F., C.N. Smith, L.A. Mulkey, D.J. Dean, and P. Jowise. 1984. Users manual for Pesticide Root Zone Model (PRZM), Release I. EPA/600/3-84/109.

Integration of GIS With the Agricultural Nonpoint Source Pollution Model: The Effect of Resolution and Soils Data Sources on Model Input and Output

Suzanne R. Perlitsh
State University of New York, Syracuse, New York

Abstract

The assessment of agricultural nonpoint source pollution has been facilitated by linking data contained in a geographic information system (GIS) with hydrologic models. One such model is the Agricultural Nonpoint Source (AGNPS) Pollution Model, which simulates runoff, nutrients, and sediment from agricultural watersheds. Vector-based (ARC/INFO) and raster-based (IDRISI) GIS systems were used to generate AGNPS input parameters.

The objectives of this project were to generate AGNPS input parameters in GIS format from GIS data at different resolutions and different levels of detail (soil survey soils data versus soils data currently available in digital format from the United States Department of Agriculture). Differences in the AGNPS model sediment output based on the variations in GIS-generated AGNPS model input were evaluated.

The study also evaluated the influence of cell size resolution and soils data on sediment generated within each cell in the watershed (SGW), sediment yield from each cell in the watershed (SY), sediment yield at the watershed outlet, and peak flow. Model output was validated by comparison with measured values at the watershed outlet for a monitored storm event. Results of this study indicate that the use of different resolution GIS data and different soils data sources to assemble AGNPS input parameters affects AGNPS model output. Higher resolution data do not necessarily provide better results. Such comparisons could affect decision-making regarding the level and type of data analysis necessary to generate sufficient information.

Introduction

Agricultural runoff is a major contributor to nonpoint source pollution. Fifty-seven percent of the pollution in impaired lakes and 64 percent of the pollution in impaired rivers of the United States can be attributed to agricultural nonpoint source pollution (1). Sediment is one of the most common agricultural nonpoint source pollutants and is the largest pollutant by volume in the United States (2). More than 3 billion tons of sediment enter surface waters of the United States each year as a result of agricultural practices (1).

Accurate assessment of the effects of agricultural activities on water quality within a watershed is vital for responsible watershed management and depends on our ability to quantify the spatial variability of the watershed and the complex interactions of hydrologic processes (3). Computer models have been developed to simulate these hydrologic processes to provide estimates of nonpoint source pollutant loads. Adequate simulation of a watershed's spatial variability helps provide the best representation of hydrologic processes within the watershed.

Preservation of spatial variability within hydrologic models can be accomplished using a distributed parameter model. The distributed parameter model is more advantageous than lumped parameter models, which generalize watershed characteristics, because distributed parameter models provide more accurate simulations of the systems they model (4). One of these models is the Agricultural Nonpoint Source (AGNPS) Pollution Model. AGNPS is a distributed process model because it produces information regarding hydrologic processes at grid cells within the watershed, thus enabling preservation of the spatial variation within the watershed. Distributed parameter models integrate well with GIS because GIS can replicate the grid used in a distributed parameter model. Manual compilation of AGNPS input parameters required to evaluate small areas at low resolution (large grid cells) is relatively easy. Manually assembling data to evaluate larger areas at finer resolutions becomes tedious, however. The integration of GIS data with the AGNPS model facilitates data assembly and manipulation (5).

Several researchers have integrated AGNPS with GIS (4-11). Smaller cell sizes within distributed parameter models are thought to best represent spatial variability within a watershed (5, 10). Certain AGNPS input parameters show sensitivity to changes in grid cell size, affecting sediment yield output (11). The use of GIS to generate input parameters for the AGNPS model enables analysis of watersheds at higher resolutions than would be practical using manual methods (5).

Research Hypotheses

The project investigated the following research hypotheses:

- AGNPS output at the highest resolution will better approximate sediment yield at the watershed outlet.
- AGNPS output for sediment generated within each cell in the watershed at highest resolution will best reflect the watershed processes.
- AGNPS output generated from the more detailed soil survey data will better estimate watershed processes.

The project also investigated other questions: will certain AGNPS input parameters (cell land slope, soil erodibility [K], the cropping factor [C], and the U.S. Department of Agriculture [USDA] Soil Conservation Service [SCS] curve number [CN] show sensitivity to changes in grid cell size? How does slope affect model output? Does a qualitatively significant difference exist between model input parameters and output calculated from data sets generated at different resolutions with different levels of detail in soils output?

Objectives/Tasks

The research in this project included analyses of:

- Certain AGNPS input parameters generated at different resolutions (10- x 10-, 30- x 30-, 60- x 60-, and 90- x 90-meter resolutions).
- AGNPS output for sediment yield (SY) and sediment generated within each cell (SGW) at different resolutions (center cells of 10- x 10-, 30- x 30-, 60- x 60-, and 90- x 90-meter resolutions).
- AGNPS output generated from different levels of detail in the soils input data (soil survey versus STATSGO data sources).

Significance

Version 4.03 of AGNPS was released in June 1994. Version 4.03 allows for evaluation of 32,767 cells. This version allows for cell sizes from 0.01 to 1,000 acres (approximately 6.36- x 6.36-meter resolution to 2,012- x 2,012-meter resolution). Previous versions of AGNPS limited the number of cells to 3,200 and the cell size resolution to 0.4 hectares (or 63.25 x 63.25 meters) (12). Reviewed literature provides no evidence that AGNPS has been used to evaluate a watershed at 10- x 10-meter resolution.

The soils data in this study were compiled at two different levels of detail. Soils data at the 1:20,000 soil survey level were generated in digital format. This level of detail was compared with soils data at the State Soil Geographic (STATSGO) database level with a scale of 1:250,000. Reviewed literature mentions no previous studies comparing AGNPS output with input generated from these two different levels of detail in soils input.

Technology for collecting and processing geographic data is continuously improving. Currently, the United States Geological Survey (USGS) 1:24,000 digital elevation models (DEMs) are available at 30- x 30-meter resolution. New satellite technology will enable DEM data to be available at 10- x 10-meter resolution, or higher. Certain satellites currently provide land cover data at 10- to 30-meter resolution (13).

An important objective of this project was to determine whether higher resolution data provide different results when routed through AGNPS. Does spatial data at higher resolutions provide better information? This paper describes the results of an analysis of AGNPS output based on different levels of both resolution and soils detail in GIS data input sources.

Materials

The Study Area

An ongoing effort is underway to clean up Onondaga Lake, Onondaga County, New York. To accomplish this effort, areas contributing agricultural nonpoint source pollution to the lake are being evaluated.

The Onondaga Lake watershed is approximately 287.5 square miles, with 40 subwatersheds. The subwatersheds in the agricultural portion of the Onondaga Lake watershed (south of Syracuse, New York) have been isolated for study of their potential nonpoint source contributions to Onondaga Lake (see Figure 1). The study area watershed (1.84 square miles, 1,177.5 acres) is one of these agricultural subwatersheds. GIS data were collected within the Otisco Valley quadrangle (USGS 1:2,400), which includes the southern portion of the Onondaga Lake watershed. Elevations in the study watershed range from 1,820 feet to 1,203 feet, with an average elevation of 1,510 feet. The watershed perimeter is approximately 6.5 miles (34,505 feet). The streams in the watershed flow from south to north to Rattlesnake Gulf, with a stream length of approximately 3.08 miles (16,265.4 feet). The stem fall of the main stream stem is quite steep at 283 feet per mile. The drainage density of the watershed is 1.67 miles of stream per square mile. Land use in the watershed is predominantly agricultural (82.8 percent).

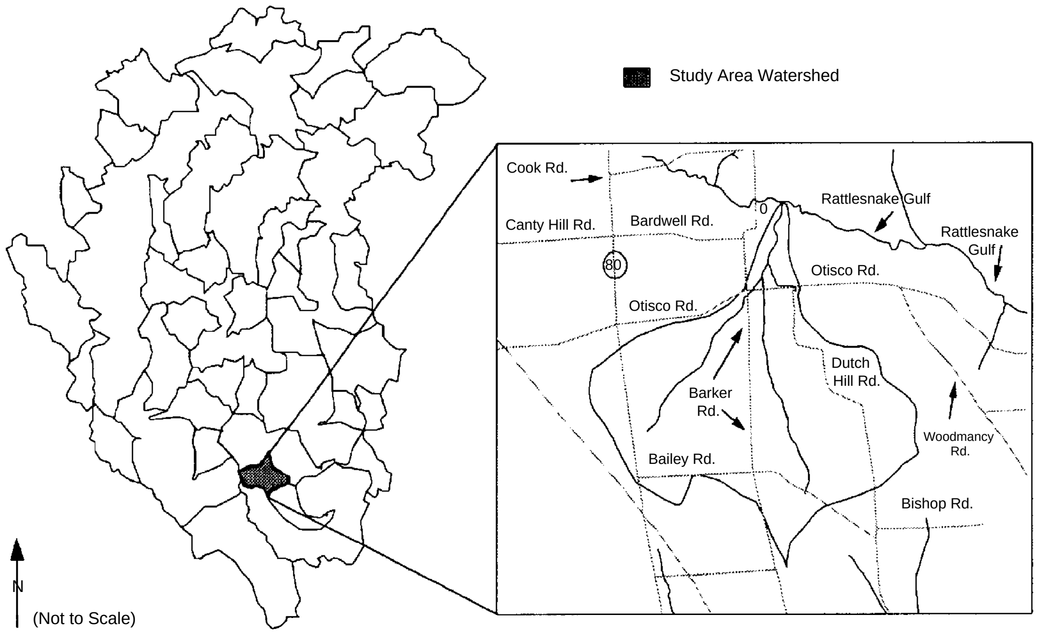


Figure 1. Onondaga Lake watershed and study area (not to scale).

The AGNPS Model

AGNPS was developed to analyze and provide estimates of runoff water quality, specifically to evaluate sediments and nutrients in runoff from agricultural watersheds for a specific storm event (11). To use AGNPS, a watershed is divided into cells of equal area. Calculations for each of the model output values are made within each cell based on the watershed data contained in each cell. Approximately 1,000 people in 46 different countries use the AGNPS model. Users include students, university professors, government agencies, lake associations, and environmental engineers.[1] AGNPS was developed in 1987 by the Agricultural Research Service (ARS) in cooperation with the Minnesota Pollution Control Agency (MPC) and the SCS. The model runs on an IBM-compatible personal computer.

Data Sources

The GIS packages of ARC/INFO Version 3.4D (14) and IDRISI Version 4.1 (15) were used to prepare input parameters for the AGNPS model. AGNPS input parameters were derived from three base maps—land use, a DEM, and soils. Table 1 shows the 22 input parameters that AGNPS required, and the base source for the data.

The land use map was obtained from a classified ERDAS image (resolution of 28 x 28 meters). The image was converted to IDRISI, brought into ARC/INFO, and regridded based on the resolution requirements of each data set. USGS could not provide a DEM for the study area, so the DEMs were interpolated from points digitized in ARC/INFO based on the Clarke method (16). The DEMs were interpolated in IDRISI on 10- x 10-, 30- x 30-, 60- x 60-, and 90- x 90-meter resolution surfaces.

Soil survey data were obtained from Onondaga County Soil Survey air photographs. An orthophoto of the 7.5 minute quadrangle was obtained from the USGS and was used with a zoom transfer scope to ortho-correct the soil survey data. The corrected soil polygons were then digitized in ARC/INFO. The Otisco Valley quadrangle comprises 79 soils mapping units. Thirty-eight different mapping units occur in the study area watershed.

The USDA SCS (now the Natural Resource Conservation Service [NARCS]) provides digital soils data from its STATSGO database. The mapping scale of STATSGO data is 1:250,000, thus it is best suited for broad planning and management uses. The number of soil polygons per quadrangle is between 100 and 400, and the minimum area mapped is 1,544 acres. The STATSGO soil data used in this project were obtained from the Onondaga County Soil Conservation Service. Approximately seven STATSGO soil groups were identified for the Otisco Valley quadrangle. Only one STATSGO soil type occurs in the study area watershed (Honeyoe silt loam).

Methods

AGNPS input parameters that showed sensitivity to changes in grid cell size in previous studies were compared between the resolutions. The AGNPS model was run eight times using precipitation values from the actual

[1] Personal communication from AGNPS Technical Support, September 1994.

Table 1. AGNPS Input Parameters

#	AGNPS Parameter	Root Data Source	General Derivation of Input
1.	Cell number	Watershed map	Program written to determine #
2.	Cell division	Not applicable	No cell division, assumed 0
3.	Receiving cell number	Aspect map from DEM	Program written to determine #
4.	Receiving cell division	Not applicable	No cell division, assumed 0
5.	Flow direction	Aspect map from DEM	Reclassed 1-8 from azimuth map
6.	SCS curve number	Land use and soils coverage	SML written to determine CN
7.	Land slope percentage	Slope map from DEM	Provided in slope percentage from IDRISI
8.	Slope shape factor	Algorithm	Assume uniform slope
9.	Average slope length	Table of values	Obtained from SCS
10.	Manning's n coefficient	Literature values	Attached to land use database
11.	USLE K factor	SCS and soil survey	Attached to soils database
12.	USLE C factor	Literature values	Attached to land use database
13.	USLE P factor	Literature values	Attached to land use database
14.	Surface condition constant	Land use coverage	Attached to land use database
15.	Chemical oxygen demand	Land use coverage	Attached to land use database
16.	Soil texture	Soil survey	Attached to soils database
17.	Fertilizer indicator	Land use coverage	Assumed for agricultural land class
18.	Pesticide indicator	Land use coverage	Assumed for agricultural land class
19.	Point source indicator	USGS 1:24,000 map	Points in ARC/INFO and IDRISI
20.	Additional erosion	Field survey, known gullies	Assume no additional erosion
21.	Impoundment indicator	1:24,000 map, field survey	Assume no impoundments
22.	Channel indicator	Streams coverage	Assume no significant channel

storm that was monitored. Each time, the model was run using an input file created with the different input data sources as follows:

1. 30- x 30-meter resolution-soil survey data.
2. 30- x 30-meter resolution-STATSGO data.
3. 60- x 60-meter resolution-soil survey data.
4. 60- x 60-meter resolution-STATSGO data.
5. 90- x 90-meter resolution-soil survey data.
6. 90- x 90-meter resolution-STATSGO data.
7. Center cells of the 10- x 10-meter resolution—soil survey data.
8. Center cells of the 10- x 10-meter resolution—STASGO data.

As grid cell size increases, the time required to assemble data as well as the space required to store the data files increase. If a cell size resolution is cut in half, the number of cells in that coverage quadruples. In the study area watershed, increasing grid cell size from 90- x 90-meter resolution to 60- x 60-meter resolution created 778 more cells within the watershed. Moving from 60- x 60-meter resolution to 30- x 30-meter resolution added 3,646 cells to the watershed, and moving from 30- x 30-meter resolution to 10- x 10-meter resolution added 40,115 cells to the watershed data set (see Figure 2). Due to the 32,767-cell limitation of AGNPS Version 4.03, AGNPS output for SY and SGW at the 10- x 10-meter resolution (which contains 45,104 cells) could not be obtained. Input parameters at 10- x 10-meter resolution, however, could be compared with input parameters at 30- x 30-meter resolution.

Methodology of Data Analysis

The input parameter "maps" were converted to IDRISI files and combined in a format that could be routed through the AGNPS model. AGNPS model output for soil generated within each cell and for sediment yield was assembled. The 30- x 30-meter resolution maps were compared with the 60- x 60-meter resolution maps; the 60- x 60-meter resolution maps were compared with the 90- x 90-meter resolution maps; and the 30- x 30-meter resolution maps were compared with the 10- x 10-meter resolution maps.

A method for comparing maps with different grid sizes was developed so that maps of different resolutions

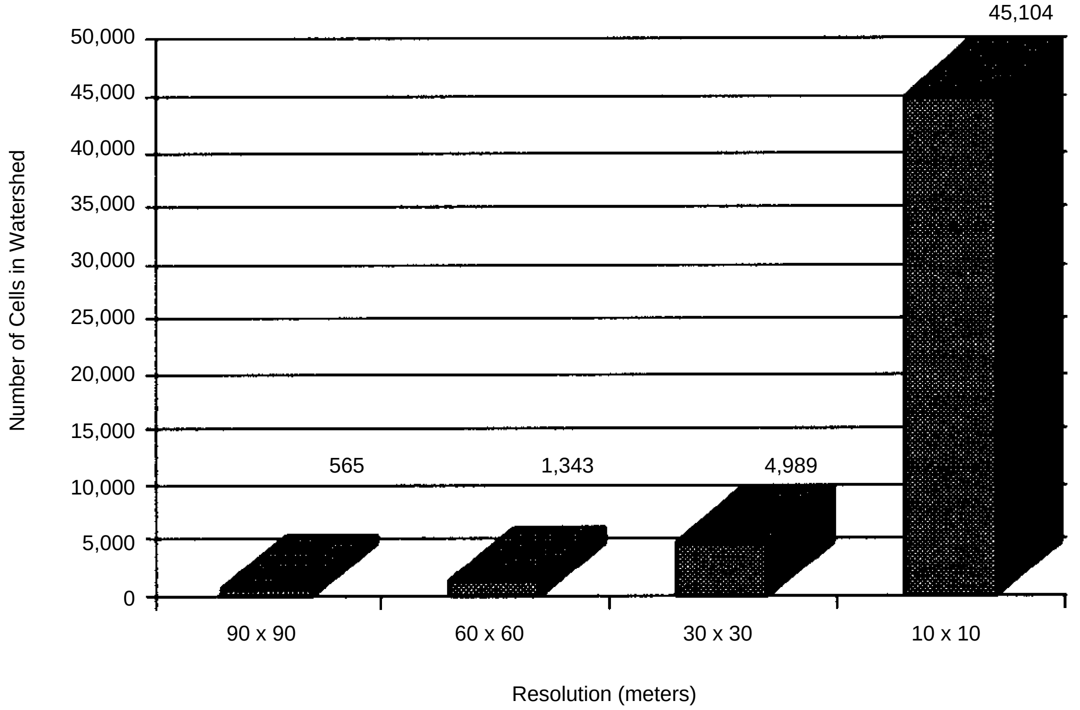


Figure 2. Effect of resolution on study area data set.

could be compared using the root mean square (RMS) statistic. This method was selected to provide a means for statistical analysis that could accommodate the spatial variability of this data set. The 60- x 60-meter resolution maps were "expanded" in IDRISI. This duplicated each grid cell so that the original 60- x 60-meter resolution map was equivalent to the 30- x 30-meter resolution map. The IDRISI command "RESAMPLE" was used to bring the 60- x 60-meter resolution map data onto the same grid size as the 90- x 90-meter resolution map. The center cells of the 10- x 10-meter resolution map were selected for comparison with the 30- x 30-meter grid cells. This comparison is based on the fact that the center cell of the 10-meter resolution is the cell that best corresponds to the entire cell of 30-meter resolution (see Figure 3).

Once maps were registered on comparable grids, the RMS difference between the maps of differing resolution was used to compare the difference between the maps

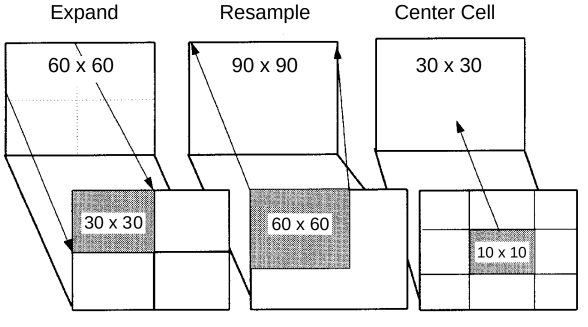


Figure 3. Method for comparing cells of differing resolution.

and the effect of cell size resolution. The RMS statistic is a measure of the variability of measurements about their true values. The RMS is estimated by comparing values in one grid system with the values in the comparison grid system. The difference between corresponding values in each grid system is squared and summed. The sum is then divided by the number of measurements in the sample to obtain a mean square deviation. Finally, the square root of the mean square deviation is calculated. The RMS difference quantifies the discrepancy between two data sets.

$$\mathrm{RMS} = \sqrt{\sum_{i=1}^{n} \frac{(\mathrm{grid1} - \mathrm{grid2})^2}{n}}$$

Results

The Storm Event

The storm that was monitored for the purposes of this field validation occurred on August 28, 1994, at approximately 8:45 p.m.; the storm duration was approximately 1.5 hours. It was a high intensity, short duration thunderstorm (see Figure 4). A global flow probe (Model FP101) was used to obtain discharge velocity measurements in feet per second. These values were then multiplied to obtain cubic feet per second. The peak discharge occurred at 11:30 p.m. on August 28, 1994, with a flow discharge of 11.15 cubic feet per second. The average runoff for the period was 4.458 cubic feet per second, or 2.42 cubic feet per second per square mile. The runoff volume per day was 0.09 inches.

Total sediment yield was derived from the analysis of total solids measured in field samples throughout the 24-hour storm period. The samples were processed to evaluate total suspended solids (TSS) using the vacuum filtration procedure (17). A total of 1.204 tons of suspended sediment was predicted at the watershed outlet from field data samples. A LaMotte field nutrient test kit was used to measure nitrate and phosphate concentrations in the stream. Nutrient values in this watershed for this storm event were so small (phosphorous below 0.1 parts per million and nitrogen 0.3 parts per million), they were not selected as parameters to be used in evaluating and validating the AGNPS model. The AGNPS predicted nutrient output for the storm was not measurable (0.00 parts per million). The low levels of nitrogen and phosphorous in the stream channel during the storm event can be attributed to the time of year in which the stream was monitored. At the time of field validation, agricultural activities were not operating.

Results at the Watershed Outlet: Peak Flow

The peak flow values that AGNPS calculated are largest at the highest resolution and decrease as cell size increases. The peak discharge from the watershed during the monitored storm event was 11.15 cubic feet per second. Comparisons of the AGNPS predicted peak flow and the actual field-validated peak flow showed that the 30- x 30-meter resolution cells best approximate the peak flow of the watershed for the sampled storm event. As grid cells increase from 30 x 30 to 60 x 60 and from 60 x 60 to 90 x 90, the peak flow is underestimated. As grid cells decrease from 30 x 30 to 10 x 10, the peak flow is grossly overestimated (see Table 2).

Results at the Watershed Outlet: Sediment Yield

Sediment yield at the watershed outlet was determined to be 1.204 tons. In all of the resolutions, the amount of sediment deposited at the watershed outlet cell increased as cell resolution increased (see Table 2). For this particular watershed in this particular storm, the AGNPS model overestimated the sediment yield predicted at the watershed outlet at the 10- x 10-, 30- x 30-, and 60- x 60-meter resolutions and underestimated sediment yield at the 90- x 90-meter resolution. Table 2 includes the information that AGNPS predicted for the cell designated as the watershed outlet within each resolution. (The results reported include output from the center cells of the 10- x 10-meter resolution data set, routed through the AGNPS model. Although these values are reported, the results from this data set cannot be assumed to approximate the sediment output that would result had the entire 10- x 10-meter resolution data set been simulated.)

Soil Survey Versus STATSGO Data

The Kappa statistic (14,18) was used as an indicator of similarity to describe the differences between the AGNPS output for SY and SGW generated from STATSGO and soil survey data. Results (see Table 3) indicate that no significant difference exists between the output derived from the STATSGO and soil survey data inputs within

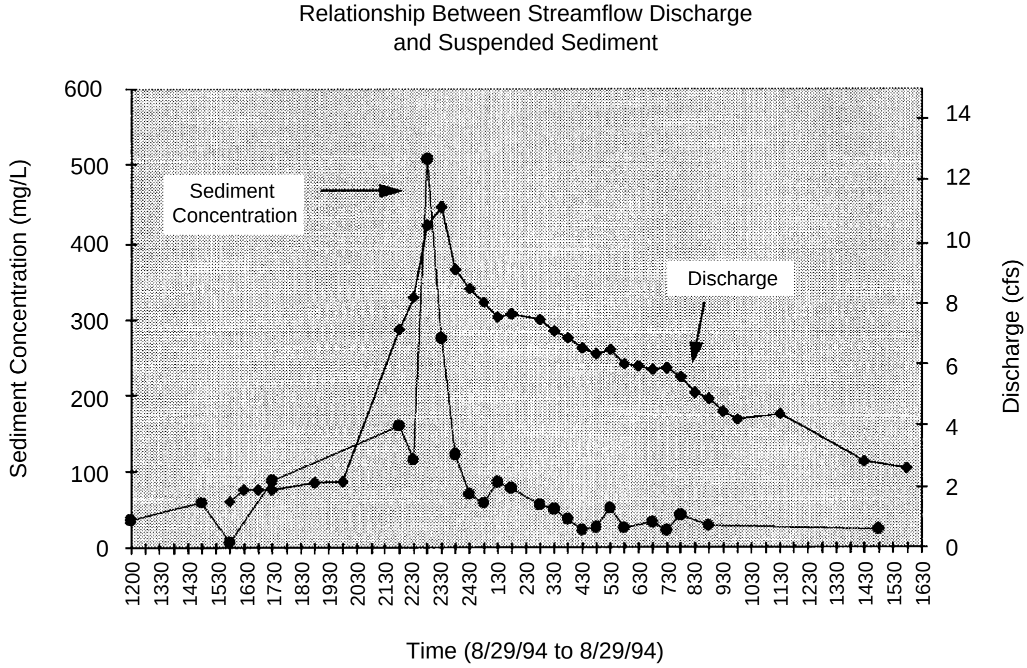


Figure 4. Storm event hydrograph and pollutograph.

Table 2. AGNPS Results at the Watershed Outlet Versus Actual Field Values

Resolution (meters)	Predicted Soil Survey Data	Difference From Actual	Predicted STATSGO Data	Difference From Actual
10 x 10 (center cells)				
Peak runoff rate (cfs)[a]	29.88	+18.73	29.88	+18.73
Total sediment yield (tons)	4.48	+3.49	4.69	+3.28
30 x 30				
Peak runoff rate (cfs)	11.27	+0.12	11.27	+0.12
Total sediment yield (tons)	2.84	+1.64	3.11	+1.91
60 x 60				
Peak runoff rate (cfs)	9.96	-1.19	9.96	-1.19
Total sediment yield (tons)	2.11	+0.91	2.12	+0.92
90 x 90				
Peak runoff rate (cfs)	8.87	-2.28	8.87	-2.28
Total sediment yield (tons)	0.86	-0.34	0.87	-0.33

[a]cfs: cubic feet per second.

Table 3. Kappa Coefficient of Similarity

Resolution (meters)	Kappa
10 x 10 (center)	
Soil survey versus STATSGO SY	0.9866
Soil survey versus STATSGO SGW	0.9703
30 x 30	
Soil survey versus STATSGO SY	0.9859
Soil survey versus STATSGO SGW	0.9785
60 x 60	
Soil survey versus STATSGO SY	0.8960
Soil survey versus STATSGO SGW	0.7542
90 x 90	
Soil survey versus STATSGO SY	0.8743
Soil survey versus STATSGO SGW	0.6574

the same resolutions. This may be due to the homogeneity of the soil textures in both soils data sets (both dominated by silty soils).

Effects of Resolution on SGW and SY

AGNPS output for SGW and SY was evaluated for every cell within the watershed. The RMS difference was applied to determine the relative effect of input data resolution on SY and SGW output (see Table 4 and Figure 5).

Table 4. RMS for AGNPS Sediment Loss Output

Description	SGW Pounds	SY Pounds
Soil Survey Data Constant		
10 centers to 30-meter resolution	168.38	344.22
30- to 60-meter resolution	125.43	739.75
60- to 90-meter resolution	93.50	498.01
30- to 90-meter resolution	29.91	711.49
STATSGO Data Constant		
10 centers to 30-meter resolution	164.45	312.48
30- to 60-meter resolution	125.37	782.07
60- to 90-meter resolution	97.07	501.81
30- to 90-meter resolution	122.04	747.51

Moving to higher cell resolutions increasingly affects sediment generated within each cell; the largest difference in sediment generated within each cell occurs as cell resolution increases from 30 x 30 to 10 x 10 meters. Sediment generated within each cell is least affected by moving from 60 x 60 to 90 x 90 meters. Sediment yield per cell is most affected when cell resolution increases from 60 x 60 to 30 x 30 meters and least affected by increasing resolution from 30 x 30 to 10 x 10 meters.

These results prompted an assessment of the methods used to compare resolutions, to determine whether the effect on sediment yield between the 30- x 30- and 60- x 60-meter resolutions could result from the method used in comparing the resolutions (expansion of the 60- x 60-meter resolution). The procedure of comparison between the 30- x 30- and the 60- x 60-meter resolutions was repeated; however, rather than expanding the 60 x 60 data file, the 60 x 60 data file was “resampled” onto the 30- x 30-meter resolution grid, then the files were compared. The RMS results (see Table 5) show that both methods for comparing data between the resolutions provide essentially the same results. The effect of resolution on sediment yield per cell is, in fact, greatest as resolution increases from 60 x 60 to 30 x 30 meters.

Results: AGNPS Input Parameters of Concern

Previous AGNPS analyses have shown sediment yield (and sediment-associated nutrient yields) to be most affected by AGNPS inputs for cell land slope, the soil erodibility factor (K), the Universal Soil Loss Equation’s

Table 5. Comparison of RMS for 30 x 30 to 60 x 60

Method	RMS SGW Pounds	RMS SY Pounds
Expansion of 60 x 60 to 30 x 30	125.43	739.75
Resampling 60 x 60 to 30 x 30	125.37	739.30
Difference	0.06	0.45

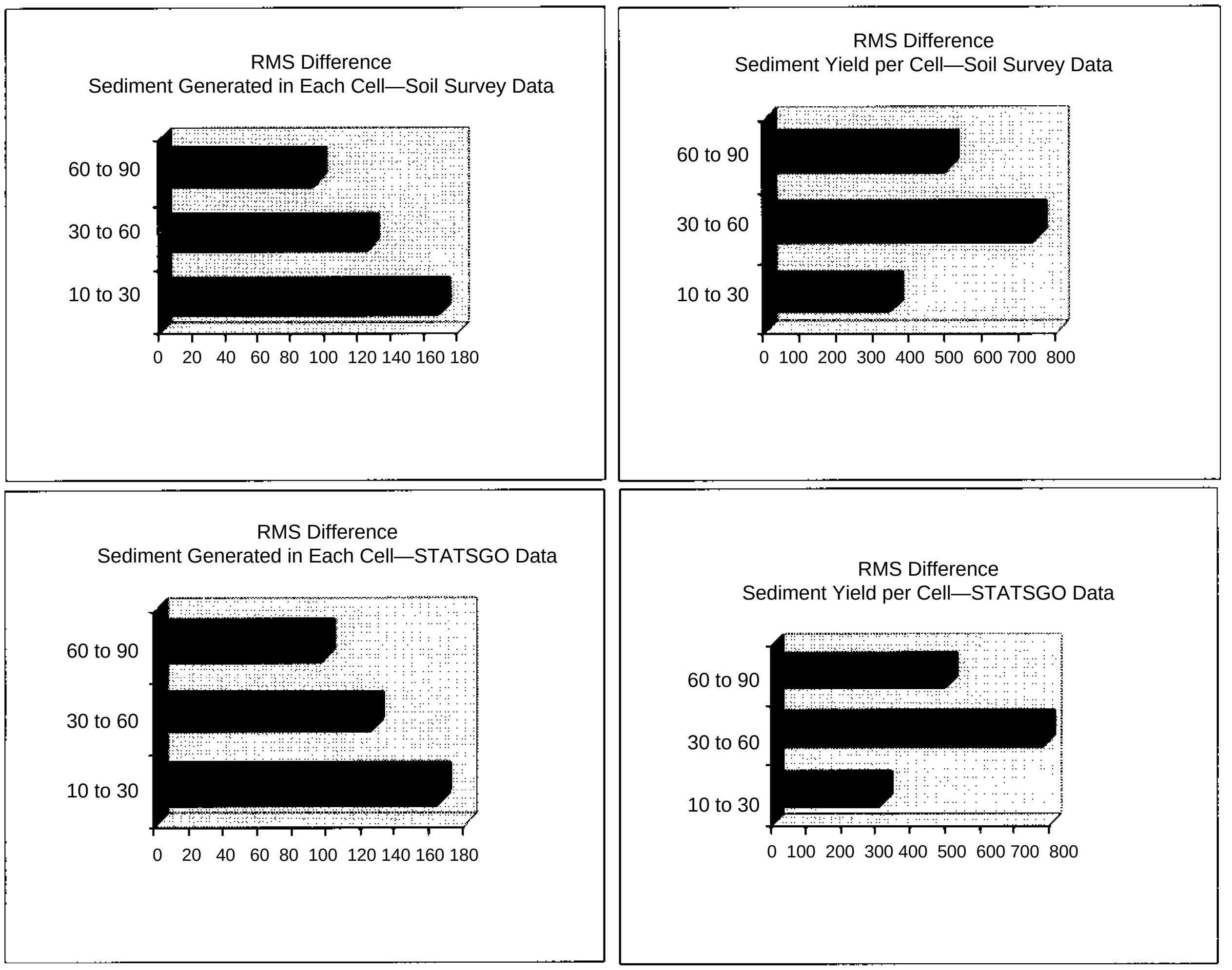


Figure 5. RMS for AGNPS sediment output.

(USLE) cropping management factor (C) and the SCS curve number (CN). To address the concerns regarding the influence of these parameters on sediment yield, the RMS differences (see Table 6 and Figure 6) and general statistics (see Table 7) for these parameters were computed.

Discussion

When evaluating the RMS as an indicator of the effect of resolution on input parameters and output sediment values, looking at the overall trend between resolutions, rather than focusing on specific values, is important. The RMS statistics for the soil erodibility factor (K), the cropping management factor (C), and the SCS curve number (CN) are least affected by a decrease in cell size resolution from 10 x 10 meters to 30 x 30 meters. These parameters are most affected when cell size resolution decreases from 30- x 30-meter to 60- x 60-meter resolution. As resolution decreases further from 60 x 60 meters to 90 x 90 meters, the effect on RMS decreases. The small-large-smaller trend in the RMS for these parameters is the same trend seen in the RMS for sediment yield throughout the watershed. The sediment yield within each cell therefore seems to be most affected by these input parameters. The general statistics for each of these parameters of concern show that very little difference exists in the values within each resolution, with the exception of slope. Slope values are higher at the higher resolutions and decrease as resolution increases. This is related to the method in which the GIS calculates slope.

Table 6. RMS Difference: AGNPS Input Parameters of Concern

Parameters of Concern	10 to 30	30 to 60	60 to 90
K value	0.0058	0.054	0.051
Cropping factor	0.03	0.173	0.015
SCS curve number	2.11	11.40	8.40
Slope	8.17	5.59	3.68

The RMS statistics comparing resolutions for sediment generated within each cell follow the same trend as the RMS statistics for slope percentage. As resolution increases, so do the discrepancies between the compared

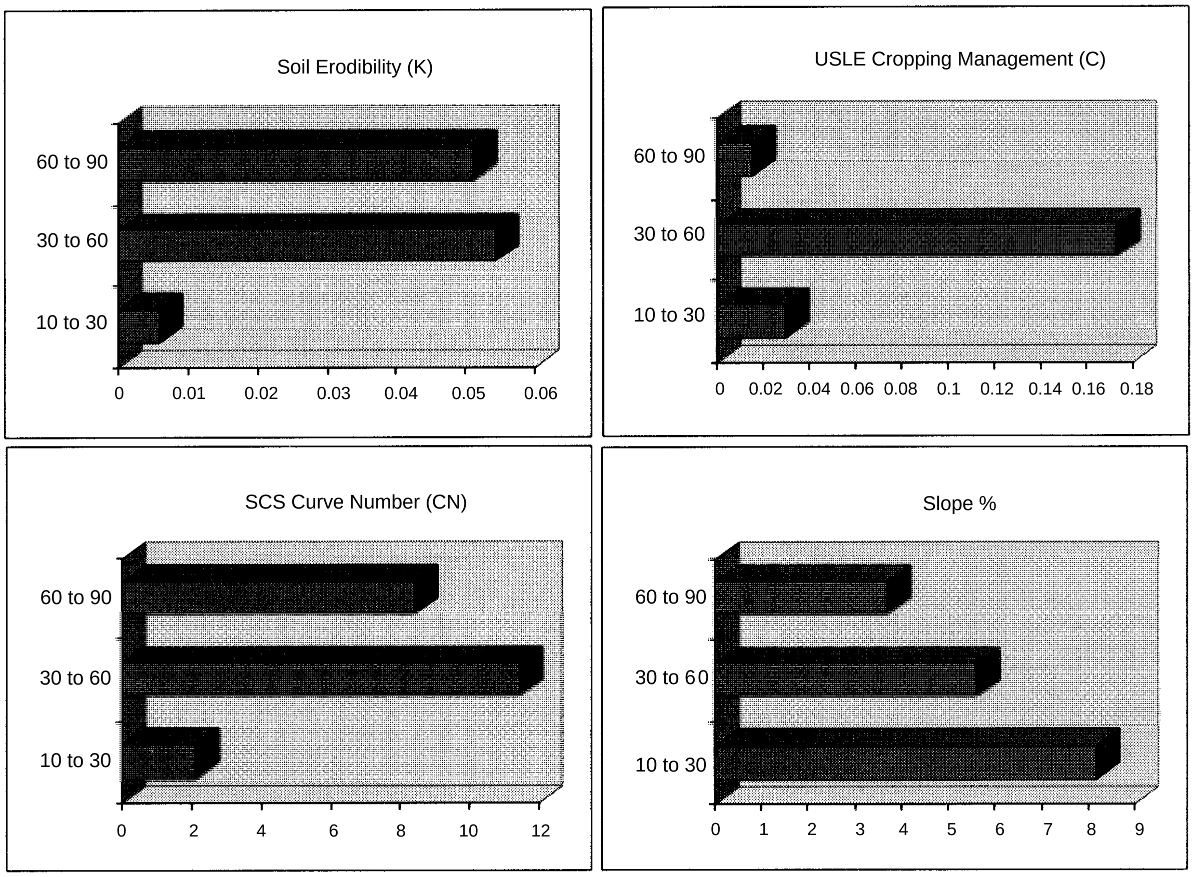


Figure 6. RMS difference for AGNPS input parameter of concern.

data sets. This trend between resolutions indicates that slope values influence sediment generated within each cell in the watershed.

The results from the 10- x 10-meter resolution data set were obtained by selecting the center cells of the 10- x 10-meter resolution data set and routing the data from this set through AGNPS using the flow pathways developed for the 30- x 30-meter resolution. The results do not provide the same information as would be provided had the entire 10- x 10-meter data set been routed through AGNPS. The RMS values obtained from comparisons of the 10- x 10-meter resolution input parameters with the 30- x 30-meter resolution input data reveal that little difference exists between the data in these resolutions. Comparison of the center cell 10- x 10-meter data set output with field monitored data shows that the 10 x 10 center cell data overestimates both peak flow and sediment yielded at the watershed outlet. This can be attributed to the larger slope values in this resolution.

Conclusion

This study used GIS to generate data files for application to the AGNPS model. The objectives of this project were to evaluate the effect of different levels of detail used in generating the input files on selected input and output parameters. The results show that, for a watershed with characteristics equivalent to those of the study area watershed, differences exist in model output based on the cell size resolution of the watershed.

The selected cell size resolution directly affects slope values. The influence of the slope parameter dominates AGNPS predictions for sediment generated within each cell and sediment yield at the watershed outlet in the study area watershed. The indicated parameters of concern have the most influence on sediment yield for each cell in the watershed. The greatest variation in the indicated parameters of concern and thus the sediment yield output occurs between the 30- x 30-meter and 60- x 60-meter resolutions. AGNPS estimates for sediment yield in files generated from STATSGO data were larger than sediment yields from files generated with soil survey soils data in the 30- x 30-, 60- x 60-, and 90- x 90-meter resolutions. For this watershed, however, no significant difference existed between data generated from soil survey and STATSGO data sources as indicated by the kappa coefficient of similarity.

Results predicted by the AGNPS model at the watershed outlet were compared with results from an actual storm monitored at the watershed outlet. The 30- x 30-meter resolution data set provided the most accurate

Table 7. Statistics for AGNPS Input Parameters of Concern

Value	10 x 10	30 x 30	60 x 60	90 x 90
K Value (units of K)				
Average	0.2989	0.2989	0.2882	0.2889
Maximum	0.49	0.49	0.37	0.49
Minimum	0.17	0.17	0.17	0.17
Standard deviation	0.0453	0.0453	0.0513	0.0456
C Factor (units of C)				
Average	0.0306	0.0306	0.0295	0.0295
Maximum	0.076	0.076	0.076	0.076
Minimum	0	0	0	0
Standard deviation	0.0212	0.0211	0.0213	0.0208
CN (units of CN)				
Average	71.004	71.003	70.75	70.68
Maximum	100	100	100	100
Minimum	55	55	55	55
Standard deviation	9.13	9.08	8.97	8.84
Slope (%)				
Average	34.13	30.153	27.67	26.38
Maximum	567	224	152	99
Minimum	0	0	0	0
Standard deviation	33.75	23.45	18.97	16.44

prediction for peak flow at the watershed outlet. AGNPS output in the 10- x 10-meter center, 30- x 30-meter, and 60- x 60-meter resolutions overestimated the actual sediment yield recorded at the watershed outlet for the validated storm event.

For the study area watershed, cell size resolution of 30 x 30 meters seems appropriate based on the accurate AGNPS model prediction for peak flow when validated with the field-monitored storm. The steep slopes created in the 10- x 10-meter resolution data set may lead to an overestimation of sediment output, rendering data at this resolution unreliable. At this time, the 10- x 10-meter resolution is both impractical and infeasible for use with the AGNPS model.

AGNPS output at the highest resolution does not provide a better approximation of sediment yield at the watershed outlet. AGNPS output for sediment generated within each cell in the watershed at the highest resolution does not accurately simulate the watershed processes. AGNPS output generated from the more detailed soil survey data is not significantly different from data generated by the STATSGO digital soils database.

This study raises the following questions:

- What level of detail is both practical and acceptable for policy-making and decision-making?
- What constitutes a cost-effective analysis?

Ultimately, these questions are best answered on a case-by-case basis and should be determined based on the size of the study area and on how the results of the analysis will be used (i.e., to make a direct land use decision or for broader planning). For broad planning analyses on large watersheds, the benefit of digitizing the soil survey data is outweighed by the cost in time and effort to generate this detailed database. STATSGO data may be sufficient. If a direct land management decision is being made for a small area such as a farm within a watershed, however, the analysis should use the most detailed soils data.

Recommendations for Future Work

The original intent of this study was to use the capabilities of AGNPS Version 4.03 to evaluate a watershed using data generated at a high cell size resolution—10 x 10 meters. AGNPS h

ad never been used to evaluate data at such a high resolution. As discovered during this project, the newest version of AGNPS is not, at this time, capable of handling a data set that has more than 32,767 cells (19). Once this limitation with the AGNPS model is remedied, the entire 10- x 10-meter resolution data set should be routed through the model so that definite conclusions regarding the applicability of such a detailed data set can be made.

References

1. U.S. Department of Agriculture (USDA). 1991. Riparian forest buffers: Function and design for protection and enhancement of water resources. N1-PR-07-91.
2. Ashraf, M.S., and D.K. Borah. 1992. Modeling pollutant transport in runoff and sediment. Trans. Amer. Soc. Agric. Eng. 35:1,789-1,797.
3. Barten, P., and K. Stave. 1993. Characterization of streamflow and sediment source areas for the Little Beaver Kill Watershed, Ulster County, NY (July).
4. Vieux, B., and S. Needham. 1993. Nonpoint-pollution model sensitivity to grid cell size. J. Water Resour. Planning and Mgmt. 119(2).
5. Kumar, V. 1993. Geographic information system application for nonpoint source pollution management. Logan, UT: Utah State University.
6. Yoon, J., L. Padmanabhan, and L. Woodbury. 1993. Linking Agricultural Nonpoint Source Pollution Model (AGNPS) to a geographic information system. Proceedings of the AWRA Conference on Geographic Information Systems and Water Resources (March).
7. Haddock, G., and P. Jankowski. 1993. Integrating nonpoint source pollution modeling with a geographic information system. Department of Geography, University of Idaho.

8. Mitchell, J., B. Engel, R. Srinivasan, and S. Wang. 1993. Validation of AGNPS for small watersheds using an integrated AGNPS/GIS system. Proceedings of the AWRA Conference on Geographic Information Systems and Water Resources (March).

9. He, C., J. Riggs, and Y. Kang. 1993. Integration of geographic information systems and a computer model to evaluate impacts of agricultural runoff on water quality. Proceedings of the AWRA Conference on Geographic Information Systems and Water Resources (March).

10. Freezor, D., M. Hirschi, and B. Lesikar. 1989. Effect of cell size on AGNPS prediction. Paper No. #89-2184. ASAE/CSAE Meeting, St. Joseph, MI.

11. Young, R., C. Onstad, D. Bosch, and W. Anderson. 1989. AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. J. Soil and Water Conserv. 44(2).

12. U.S. Department of Agriculture (USDA). 1987. Agricultural Research Service, AGNPS: Agricultural nonpoint-source pollution model, a watershed analysis tool. Conservation Research Report No. 35.

13. Star, J., and J. Estes. 1990. Geographic information systems: An introduction. Englewood Cliffs, NJ: Prentice Hall.

14. ESRI. 1987. PC ARC/INFO user's guides. Redlands, CA: Environmental Systems Research Institute, Inc.

15. Eastman, R. 1993. IDRISI technical reference. Worcester, MA: Clark University.

16. Clarke, K.C. 1990. Analytical and computer cartography. Englewood Cliffs, NJ: Prentice Hall.

17. Clesceri, L., A. Greenberg, and R. Trussell. 1989. Standard methods for examination of water and wastewater, 17th ed. pp. 2-75 to 2-76.

18. Rosenfield, G.H., and K. Fitzpatrick-Lins. 1986. A coefficient of agreement as a measure of thematic classification accuracy. Photogrammetric Engin. and Remote Sensing 52(2).

19. U.S. Department of Agriculture, Soil Conservation Service. 1995. AGNPS Newsletter.

Comparing Experiences in the British and U.S Virgin Islands in Implementing GIS for Environmental Problem-Solving

Louis Potter
Office of the Chief Minister of the Government, British Virgin Islands

Bruce Potter
Island Resources Foundation, St. Thomas, U.S. Virgin Islands

The British and U.S. Virgin Islands: Comparisons and Contrasts

British Virgin Islands

Three miles to the north and east of the U.S. Virgin Islands (USVI) lie the British Virgin Islands (BVI), a group of 36 islands and cays with a total area of 60 square miles. The four largest are Tortola (24 square miles), Anegada (14 square miles), Virgin Gorda (8.5 square miles), and Jost Van Dyke (4.5 square miles).

Geologically, the BVI belong to the Greater Antilles, and like the USVI and Puerto Rico, rise from the Virgin Banks (or Puerto Rican Shelf). Rocks of the BVI, except Anegada, consist of thick, steeply inclined, metamorphosed volcanic and sedimentary stratified series of Cretaceous age, with dioritic and pegamitic intrusions. Anegada is a 30-foot-high emergent coral limestone platform, apparently from the Pleistocene age.

The BVI are a crown colony of the United Kingdom (UK) with a total population in 1991 of 16,108. Most of the population resides on Tortola (13,225 inhabitants). Between 1980 and 1991, population increased 46.6 percent. The BVI have internal self-government with an elected council headed by a chief minister. The UK appoints a governor to represent the queen and to manage defense, internal security, external affairs, civil service, and court administration.

The BVI economy is based mainly on tourism and servicing international business. Sailing and diving are pivotal features in BVI tourism. The average number of tourists per capita in the BVI is 221, compared with 119 for the USVI (1). The BVI 1990 per capita income was recorded at $10,125. Prior to the 1970s, the BVI economy was based on subsistence agriculture and remittances from British Virgin Islanders who worked in the USVI.

U.S. Virgin Islands

The USVI are an unincorporated territory of the United States, purchased from Denmark in 1917. The total population in 1990 was about 101,000, divided among St. Thomas (48,000), St. Croix (50,000), and St. John (3,500). The tourism-dominated economy of the Virgin Islands generated a per capita Gross Territorial Product in 1990 of approximately $11,000—the highest in the Caribbean (2).

Geographically, geologically, and topographically, St. Thomas (28 square miles) and St. John (20 square miles) are similar; they are both largely volcanic, have deeply indented coastlines, and lie on the Puerto Rican Shelf. St. Thomas and St. John are close to the BVI. St. Croix is a relatively large (84 square miles) and mostly limestone island that lies on its own submarine ridge, which rises more than 4,000 feet from the bottom of the Caribbean Sea. St. Thomas and St. John are about 5 miles apart, and St. Croix is 48 miles south of them.

During the height of tourism development (from the late 1950s through the mid-1970s), the USVI experienced average annual compound population growth rates of over 6 percent, as well as a doubling in real incomes. This unprecedented paroxysm of growth is still being assimilated by a population that differs greatly from the 30,000 people who lived in a predominantly agricultural USVI in 1950. In 1990, the USVI received a daily average of 37 visitors per square kilometer. This compares with a visitor load of 23 visitors per day per square kilometer in the BVI, which is also a high-density tourist destination (1).

Background to GIS Implementation Activities

British Virgin Islands

The idea of geographic information systems (GIS) applications in the BVI first arose with a presentation about a proposed project for St. Lucia, made by Dr. Jan Vermeiren of the Organization of American States at the Caribbean Conference of Planners in Kingston, Jamaica, in 1984 (3). The Town and Country Planning Department recognized that it could use GISs analytical and display properties to make presentations to the chief minister and the BVI Executive Council. This proposal fell on fertile ground, given a relatively long-held tradition of support for the Town and Country Planning Department by the United Nations Development Programme (UNDP) and the British Development Division, dating back to the early 1970s.

Subsequent to this inspiration, the Town and Country Planning Department requested budget authority to develop a land use database. This database would include buildings, property boundaries, and constructed and natural features of importance. The Finance Department hoped this project would help combat growing competition to the postal services by independent package delivery services. They renamed the project the National Addressing System, and the legislature provided $200,000 to provide a physical address for each property in the territory.

The Town and Country Planning Department conducted a pilot project, focusing on Road Town, the capital. The pilot project demonstrated that the hard copy land ownership or cadastral maps that the Survey Department was then using were inadequate for accurately accounting for properties, even in the BVI's most developed urban areas. Therefore, the Town and Country Planning Department expanded the scope of the National Addressing System project to identify options for increasing the accuracy of property ownership records, including maps.

Because the Town and Country Planning Department had little experience with computerized land information systems, departmental managers sought support from the UNDP office in Barbados. The UNDP had previously assisted with the department's development control applications database. This mode of operation, in which the BVI purchase services provided through the UNDP, which acts as a "vetting agent" for consultants and other technical assistance, continues to this day. An expert from the United Nations Community and Housing Services (UNCHS) Nairobi office provided the first such consultation by exploring how the BVI might implement a GIS. Over the next few years, three other experts provided input, which the Town and Country Planning Department gradually integrated into a picture of how to use a GIS within the technical and financial limitations of a small island government.

In the meantime, external conditions were improving the chances for the success of the BVI program. British Virgin Islanders were receiving formal and informal training in computer applications in general, and specifically in AutoCAD drafting systems, which increasing numbers of local architects and engineers are using. In addition, the power of rugged microcomputer systems that could withstand the harsh operating conditions of the Virgin Islands was also improving, and local dealers were increasing their skills in support of such systems.

U.S. Virgin Islands

In 1988, a proposal to develop a locally supported GIS was being discussed in detail in the USVI (4). This project resulted in a formal application from the government of the USVI for financial assistance from technical assistance funds provided by the Office of Territorial and Insular Affairs of the U.S. Department of the Interior. The grant was awarded in March 1991. The proposed project combined existing information from several USVI government agencies to produce GIS overlays, as shown in Table 1 (5).

In addition, according to the grant application, (which was written by the Virgin Islands government and may not have represented U.S. Geological Survey's [USGS's] intentions) the National Mapping Division (NMD) of the USGS agreed to digitize the eight USGS 1:24,000-scale quad sheets ("quadrangles") that cover the USVI, including the following categories:

- Roads and trails
  - Power transmission lines
- Hydrography
  - Stream networks
  - Shorelines
  - Wetlands
  - Mangroves
  - Reefs

Table 1. GIS Overlays

Agency	GIS Overlays
DPNR	Zoning Flood plain Subdivision
WAPA	Water distribution Aquifer profiles Electrical distribution
VITEMA and emergency services	Critical routes Critical facilities
DPW	Sewer line network Transportation Flood plain

DPNR = Department of Planning and Natural Resources
WAPA = [Virgin Islands] Water and Power Authority
VITEMA = Virgin Islands Emergency Management Agency
DPW = Department of Public Works

- Topographic contours
- Political boundaries

(These categories were subsequently adjusted based on discussion between the USGS, WAPA, and government agencies. Documentation for these new coverages is not available, but the general idea holds, although no VITEMA data were used, and the USGS did not digitize the WAPA power system.)

The underlying notion was that the whole would be bigger than the parts—each organization would bring its own corporate data and maps so all could share in the digitized product.

Five objectives were identified for the project:

- Provide the USGS in St. Thomas with a complete microcomputer (GIS) workstation.
- Develop a digitized database from USGS topographic maps.
- Contract for the digitizing.
- Acquire digital data to load into GIS.
- Enter data not in digital format.

The proposal required a 50-percent local cost contribution to the project, including a matching suite of hardware for data maintenance, backup, and analysis. The application stated that itemized costs of $305,000 "will be provided by the Virgin Islands Water and Power Authority (WAPA) and the Office of Territorial and International Affairs."

The project application is ambiguous about the nature of the hardware and operating systems that the GIS requires. A list of hardware for USGS use refers to a "microcomputer workstation." Later references to "ESRI ARC/INFO workstation software," however, indicate to sophisticated users that these applications require UNIX workstations and UNIX language operating systems. In general, UNIX is not used or supported in the Virgin Islands. Some Virgin Island officials associated with the GIS project feel they were not fully informed about the GIS operating environment and the long-term support costs to which they were committing.

A frustration for USGS in the project stems from the somewhat limited role the agency has had in providing high-quality data conversion services (digitizing). One USGS staff member explained that unfortunately the agency's mandate only extends to providing data; the agency "can't get involved in applications."

The initial project proposal also was unclear about ownership of the digitized data. The proposal spoke in general terms: "All available digital data and attributes will be complete, accurate, and up-to-date at the end of the project, and will be available for use and transfer to the Department of Planning and Natural Resources (DPNR) (5)."

Given the conditions and environmental constraints discussed, a number of specific differences developed between the GIS implementation processes of the BVI and the USVI.

Initial System Planning Activities

British Virgin Islands

Although the BVI had no formal systems plan, consultants worked with the Town and Country Planning Department on four different occasions to provide insight into some aspect of GIS applications. Sometimes, the benefits from the consultations were neither the type nor the quality the department originally expected. In general, however, each provided some additional perspective on the possible benefits and perils of implementing a GIS.

U.S. Virgin Islands

The USVI apparently made few initial system planning efforts, although the U.S. Department of the Interior and USGS have wide experience in the territory. (An informed source claims a grant was made, possible by EPA, for a preceding $50,000 GIS project, but no mention of this has been found in the materials available for this article, either as a proposal, or in terms of specific products.) Possibly, this very familiarity led to a series of unexamined assumptions and diminished communications about the exact terms of the assistance and services that the USGS exchanged with several agencies of the Virgin Island government.

One indicator of the lack of system planning activities in the Virgin Islands is a proposal that the Virgin Islands Emergency Management Agency (VITEMA) circulated for a "Geographical Information Systems: Technical Operators Meeting." This proposal, from an agency that has always been one of the most important participants in GIS activities, called for a "technical working group" to examine existing database management systems in the territory to develop a planning strategy for implementing GIS.[1] This proposal was dated September 28, 1992, 10 days before the USGS announced a demonstration of the completed products of Phase I of the "comprehensive geographic information system being developed for the USVI (6)."

[1] Ward, R.G. 1992. Geographical information systems: Technical operators meeting. Memorandum to Cyrille Singleton. VITEMA, St. Thomas, U.S. Virgin Islands (September).

Software Selected and Rationale

British Virgin Islands

In part because of the extended timeframe for the BVI planning and initiation of the GIS system, the Town and Country Planning Department never committed to a specific system configuration until the last stages of the planning process.

This process of “creative procrastination” had three synergistic results:

- PC power increased (and prices decreased) to the point where reasonably priced systems could perform many of the compute- and data-intensive operations demanded for graphics software mapping.
- ARC/INFO released the ARC/CAD version of its GIS software, which worked on PCs within the well-known AutoCAD drafting software. Architect and engineering offices in the BVI already used computer-aided design (CAD) software, so upgrading to include GIS functionality was relatively easy.
- The fourth GIS consultant to work in the BVI was experienced in implementing systems in the Caribbean and had special knowledge and access to early versions of both ARC/CAD and Version 1.0 of ARC/VIEW. These two tools, based on a last minute proposal by International Development Advisory Services (IDAS) of Miami, Florida, a private GIS support contractor, became the basis for the BVI GIS.

U.S. Virgin Islands

Workstation ARC/INFO 6.1 was selected as the basic software for the USVI GIS project because it is the USGS standard software. In the environment surrounding the USVI project, this seemed to be a sufficient explanation, although there may have been other reasons. Because of this decision, however, WAPA and the DPNR lack any means of updating map or attribute data files. Outside providers, such as the USGS’s NMD, must perform that service. The USGS has noted that the only reason WAPA and DPNR lack these capabilities is because they (WAPA and DPNR) failed to provide the matching suite of hardware and software specified in the grant application.

The digitized water supply system offers an example of the extra costs that such a condition creates. The USGS built an ARC/INFO coverage by converting mapped data from AutoCAD source files, which they then linked to detailed attribute information about each element (e.g., pipe, valve, elbow) in the system. The USGS then used ARC/INFO network software, purchased with project funds, to build a network model that analyzes and displays the operation of the entire water distribution system—but no software in the Virgin Islands can run the network model.

Hardware Platforms

British Virgin Islands

The BVI GIS was originally installed on a Compaq 486-50, with a 21-inch screen. The Town and Country Planning Department soon learned, however, that data input would be more efficient if two or three smaller machines split the work, with the Compaq available for analysis and data quality checking. The department upgraded its existing office computers to handle the data entry. Users already feel the need for networked applications to share data more quickly. Plans for GIS expansion to other offices, such as the Electricity Corporation, increase the pressure for an extended local area network.

U.S. Virgin Islands

The system that the USGS used to build the USVI GIS database was a Data General UNIX workstation with one large digitizing tablet and one pen plotter. No matching or comparable hardware are installed anywhere in the USVI, as the original project proposal had foreseen. Observers tend to agree, however, that the failure to provide a specific hardware configuration is less significant than the lack of committed, senior, full-time technical staff. This staff is required to operate the level of GIS facility that the USGS envisioned.

Base Map Priorities and Layers Constructed

British Virgin Islands

Building a map database is proving to be a long process for the Town and Country Planning Department. This is complicated by the failure of a key digitizing contractor in Texas to provide property lines in a format conducive to constructing accurate property polygons. Operators in the Town and Country Planning Department have increased their data entry efficiencies, however, and most properties on the most densely inhabited islands have now been digitized.

Producing demonstration data displays accounts for a significant part of the cost of developing databases for the early phases of the BVI GIS implementation. These demonstrations aim to illustrate possible new application areas for other agencies and departments of the BVI that are interested in cooperating and sharing costs of additional system development. For example, the Electricity Corporation and the National Disaster Preparedness Agency need to map emergency services.

Converting the data (i.e., digitizing) in house in the BVI has produced costs and benefits. The costs revolve around the steep learning curve for data entry procedures and the constant distractions of responding quickly to "outsiders" who may be important long-term supporters of the GIS. The benefits include increasing staff skills and the ability to build constituencies for the program by promptly responding to real needs.

Coverage priorities for the BVI GIS include a national addressing system, completion of the territorial land use mapping, and accurate cadastral mapping (which has major environmental planning and management implications).

U.S. Virgin Islands

The USGS produced 44 coverages for the Virgin Islands from a variety of sources. Table 2 shows the major coverages and scales, by island.

Table 2. Major Coverages and Scales for the USVI

St. Croix	St. John	St. Thomas
STC water distribution 1:2,400	STJ water distribution 1:2,400	STT water distribution 1:2,400
STC roads 1:2,400	STJ roads 1:2,400	STT roads 1:2,400
STC building footprints 1:2,400	STJ building footprints 1:2,400	STT building footprints 1:2,400
STC shorelines 1:2,400	STJ shorelines 1:2,400	STT shorelines 1:2,400
STC DLG boundaries 1:24,000	STJ DLG boundaries 1:24,000	STT DLG boundaries 1:24,000
STC DLG roads 1:24,000	STC DLG roads 1:24,000	STT DLG roads 1:24,000
STC DLG hydrography 1:24,000	STJ DLG hydrography 1:24,000	STT DLG hydrography 1:24,000
STC DLG hypsography 1:24,000	STJ DLG hypsography 1:24,000	STT DLG hypsography 1:24,000

In addition, the following National Park Service data were converted and added to the data set but were not produced by the USGS:

- STJ NPS boundaries, 1:24,000
- STJ NPS roads, 1:24,000
- STJ NPS hydrography, 1:24,000
- STJ NPS hypsography, 1:24,000
- STJ NPS benthic communities, 1:24,000
- STJ NPS historical sites, 1:24,000
- STJ NPS vegetative cover, 1:24,000

Mapping for St. Thomas, St. John, and St. Croix identified a total of 10 coverages for each island. They are based on information from the USGS ("quad sheets" specifically demarking political boundaries, shorelines and streams, topography, and roads) and higher precision WAPA mapping, which derives from 1986 aerial photogrammetry, including left and right road boundaries, building footprints, shorelines, and water supply system data. The WAPA data are at 1:2,400 scale, an order of magnitude more precise than the USGS base map. St. John mapping consists of 10 added layers based on data that the Virgin Islands National Park (VINP) provided.

The original USVI project proposal referred to a two-phase process of database development, shown below (5):

Phase	Tasks
I. Base system development	Water distribution network Power distribution network Flood plain maps
II. Agency extensions (i.e., by USVI agencies)	Land use maps Transportation networks Emergency facility networks Tax parcel/land value

The USGS announced that Phase I was completed in October 1992 (6). Supposedly, the contents of these two phases were subsequently adjusted to reflect a different range of coverages, but the notion of a "Phase II" in which local agencies would assume more operating responsibilities was retained.

Ownership of, access to, and terms that govern the use of this digital data are confused. The USGS says it is unable to provide an authoritative catalog of the coverages because "one has not been produced." WAPA says it has several diskettes of data in the safe but no equipment to manipulate them. The VINP has learned that it can use its own data as well as WAPA data converted and attributed by the USGS, but the park does not possess or use USGS digital line graph (DLG) data. To personnel in USVI agencies, USGS statements have clouded the question of access to the GIS information. For example, one such statement announced that the USGS Water Management Division cannot make the digitized data available to Virgin Island government agencies.

The DPNR apparently has no means of making direct use of the digital data. First, DPNR has no hardware or software that can use the data. Secondly, it has no operators who can build the GIS systems to actually apply the data to decision-making needs. The department is said to be preparing a new GIS proposal for training, hardware, and software for a new GIS system. According to unconfirmed rumors, this system will be based on MapGrafix, a Macintosh mapping system.

In operational terms, the data seem to belong to WAPA, which contributed major financial support and map resources to the project. WAPA has been helpful in providing copies of the digital data to other groups and agencies.

Environmental Problem-Solving in Local Decision-Making

British Virgin Islands

In the BVI, the first priority of the GIS facilities is to extend the National Addressing System and to improve the property ownership system. This will both improve postal services, as originally proposed, and provide better information for important revenue and financial analyses. Land use mapping and environmental impact assessments are important second priorities for GIS applications. Other features already developed for interim studies and analyses include mapping of significant coastal and wildlife features and environmentally sensitive areas from the Anegada Development Plan and mapping of important submarine habitats adjacent to Virgin Gorda.

The BVI have concentrated on developing GIS applications to address strategically important issues in the territory. Marine and coastal resources are vitally important to the BVI economy. They embody historical and cultural values, as well as maintain a high-quality environment to support charter yacht-based tourism, which is integral to the BVI economy. The Conservation and Fisheries Department is working with the Town and Country Planning Department to convert the country's coastal atlas to digital form (see section on Principal Users), as has been done on a demonstration basis for the Anegada and Virgin Gorda mapping.

U.S. Virgin Islands

In the USVI, environmental decision-making generally follows an adversarial, rather than a problem-solving format. A combination of historical and cultural factors have created the general assumption that the development process creates winners and losers. In this environment, information becomes an important tactical weapon, making it difficult to gather support for activities or programs that aim to make information more widely accessible. Technology is more acceptable, and more likely to receive leadership support, if it is justified on technical, less "political" terms.

The USGS team made a presentation on the Virgin Islands GIS in October 1992, after just completing the digital coverages for Phase I of the GIS project. The presentation emphasized that the GIS is intended to provide decision-makers with easily accessible spatial information (6).

Originally, the GIS was expected to benefit primarily the territory's three coastal zone commissions in their assessment of environmental effects of major development proposals. The ground-water protection program of the Division of Environmental Protection of the DPNR is using GIS analyses produced by the Water Resources Division of the USGS, employing the USVI GIS coverages with added data (e.g., wells) that the Water Resources Division is digitizing.

GIS Support Factors

According to the GIS support contractor for the BVI, a successful GIS requires three key support elements:

- GIS policy leadership
- GIS technical leadership
- Competent outside expert assistance

The following summarizes the comparative experience of the two programs for these three key implementation support factors:

Support Factors	BVI	USVI
GIS policy	Town and Country Planning Department director led project from chief minister's office	No GIS manager in government
GIS technician	BVI technician trained in the United States	No GIS specialist in government (draftsman at WAPA)
Outside support	UNDP and IDAS	USGS technical support and U.S. Department of the Interior financial support

Principal Users: Planned and Actual

British Virgin Islands

The BVI Department of Finance is the first user of data products from the GIS, based on the initial funding for the addressing system. This system is based on detailed parcel maps of the BVI so that the effective base map resolution of the BVI system is 1:2,500. This is considerably finer than the 1:24,000 scale of the USVI maps.

The Town and Country Planning Department, however, is working to recruit other users to the system, including:

- Public Works.
- Water and Sewerage Department for a systemwide map (which may eventually spin off as a separate system, given this department's long-term interest in engineering-quality facilities management information).
- Conservation and Fisheries.
- British Overseas Development Administration, which funded a coastal atlas for the BVI (7). At the suggestion of the Town and Country Planning Department, this mapping was developed in ARC/INFO. A proposal has been made to convert the coastal atlas to digital form for natural resource management applications, with a demonstration already developed showing the distribution of sensitive marine communities around Virgin Gorda.

In addition to these uses, the GIS group is starting to experiment with the use of remote sensing products in GIS production, which would encourage the use of GIS for natural resource change detection.

The Town and Country Planning Department's GIS specialist, Mikey Farara, is being reassigned to provide networking support (including GIS distribution over the network) for several government agencies. Meanwhile, the GIS operation is adding a cartographer to assist in tailoring GIS products to users' needs.

As enthusiasm for the GIS has blossomed in the BVI, managing for realistic expectations and stressing the investment costs that participating agencies can expect have become problems for the Town and Country Planning Department.

U.S. Virgin Islands

Complex evaluation issues face the USVI's three coastal zone commissions (one on each island). Therefore, land use planning in general and coastal zone permitting specifically were assumed to be important first users of the GIS. The DNPR, however, had no process to prepare the Division of Comprehensive and Coastal Zone Planning to implement this system. In addition, the scale of permitting decisions may be too fine for the GIS base map. (See discussion of scale below.)

WAPA is not using the GIS data. One senior manager characterized their experience with the GIS project as "paying a lot of money for a diskette of data that we keep locked in the safe."

The VINP (part of the U.S. National Park Service) and Biosphere Reserve have purchased a PC-based GIS system and employed an analyst to implement it for the park and adjacent areas on St. John and the surrounding seas. With this system, they plan to enter the USGS-developed data into the database. In addition, the Virgin Islands Resource Management Cooperative (a collaboration of research and resource management organizations) makes the VINP GIS data and analytical capabilities available to members, including government members.

At this time, the only major Virgin Islands government user of the GIS is the ground-water protection group of the DPNR's Division of Environmental Protection. Because they lack equipment or software to manipulate the GIS data already available, they use the Water Resources Division of USGS as a GIS contractor. This arrangement has two problems:

- *High costs:* Although the USGS "owns" the existing digital data, and processing private contracts would be complex, DPNR believes it could get similar services at cheaper prices from other vendors.
- *Inappropriate scale:* Environmental management processes in the Virgin Islands (and in most other small island states) require knowledge of property ownership, implying maximum map scales of 1:5,000 to 1:10,000. The USGS quad sheet scale of 1:24,000 is too coarse for many management purposes. Costs of remapping areas of concern at the higher resolution are high, and the problems of maintaining multiple map resolutions and sources are not trivial.

What GIS Can Do

Joseph Berry has proposed seven basic categories of "What GIS Can Do for You" (8). These applications can be related to the GIS products and proposals for the BVI and USVI, with special attention given to natural resource and environmental issues. Table 3 shows what coverages that have been or are being developed for the two systems can do.

Table 3 illustrates two contrasting issues separating the two jurisdictions. The USVI have the data available to perform a number of relatively complex analytical processes, especially in St. John. They have no capability to actually execute any such studies, however. The BVI, on the other hand, have proposed and often developed pilot or demonstration applications for several GIS uses but still need to develop the data resources to support these on a territorywide basis.

Lessons Learned

The comparative experience of these two very distinct GIS programs reinforces three basic lessons of information system design and implementation:

- *Plan, don't assume:* The prolonged, sometimes repetitious, planning process that evolved in the BVI

Table 3. Coverage Capabilities

Questions:	Analytical Function	USVI Application	USVI Status	BVI Application	BVI Status
Can you map it?	Mapping	USGS and WAPA-based coverages	Done	Land use cadastral	Done and proposed
Where is what?	Natural resource management	DLG hydrography	Done	Coastal atlas Sensitive areas	Proposed and partial
		DLG hypsography	Done	Significant features Population data	Major islands done
		Well inventory	Done	Land use	
		STJ national park coverages	Done	Sensitive areas	
Where has it changed?	Temporal	DLG and WAPA	Done	Land use updates population	Proposed
		Boundaries and roads 1982 to 1989			Proposed
What relationships exist?	Spatial			Land use Population data Coastal atlas	Partial Proposed Partial
Where is it best?	Suitability	STJ national park coverages (limited application)	Done	Land use Coastal atlas Sensitive areas Significant features	Partial
What affects what?	System	STJ national park coverages (limited application)	Done	Land use Coastal atlas Sensitive areas Significant features Population data	Partial and proposed
What if...?	Simulation	None discussed		Speculation, but no plans to implement yet	

involved multiple consultants providing often conflicting advice. This process served to educate policy-makers and managers in a much broader range of possibilities and avoidable problems than were available to the USVI. A corollary to the need for careful planning is the need to avoid making decisions or commitments to specific systems before such decisions are absolutely necessary. Especially in systems involving high technology, premature decisions often mean early obsolescence.

- *Implement in phases with early demonstration products:* Some issues, such as cadastral mapping and scale, are so subtle to inexperienced users that they need practice in real-life situations. If the USGS had spotted the scale problems at an early stage in the data conversion process, the USGS may have been able to provide a better solution. Some USVI critics claim the "1:24,000—one size fits all" attitude characterizes the federal approach.

- *Identify critical success factors for each situation:* In some environments (e.g., USVI), GIS is most attractive for its ability to provide enhanced powers of analysis. In others, such as the BVI, it is seen as a data integration tool and as a way to better inform political leadership and the public. To ensure success, major GIS implementations also need to meet the three major support requirements:
  - A political/senior management "chief"
  - A technical "chief"
  - Competent outside technical assistance

Finally, implementers should recognize that they have a stake in open information sharing. They should seek ways to redefine the decision-making process as a nonzero sum game: more information should benefit all parties. Of course, such changed attitudes require fundamental value shifts that take a long time to achieve and may have high short-term costs.

References

1. McElroy, J.L. 1991. The stages of tourist development in small Caribbean and Pacific islands. In: Proceedings of the International Symposium on the Island Economies: Policy Models for Planning Development, Lesbos, Greece (November).
2. Bureau of Economic Research, Virgin Island Department of Commerce. 1994. Economic indicators. St. Thomas, U.S. Virgin Islands (July).

3. Potter, L. 1984. Program notes from the Caribbean Conference of Planners, Kingston, Jamaica.
4. Potter, B., K. Green, and M. Goodwin. 1988. Management of natural resource information for the Virgin Islands National Park and Biosphere Reserve: Special biosphere reserve report. St. Thomas, U.S. Virgin Islands: Island Resources Foundation.
5. Government of the U.S. Virgin Islands. 1991. Application for technical assistance funds: Virgin Islands GIS system. Office of Territorial and International Affairs, U.S. Department of the Interior (April).
6. Parks, J.E. 1992. Invitation to a demonstration of the Virgin Islands GIS. U.S. Geological Survey, St. Thomas, U.S. Virgin Islands (October).
7. University of Manchester, British Overseas Development Administration. 1993. British Virgin Islands coastal atlas. Project was collaboration between graduate students at Manchester University, the British government, and the government of the British Virgin Islands.
8. Berry, J.K. 1994. What GIS can do for you. GIS World. Boulder, CO (May).

The Application of Geographic Information Systems in the Development of Regional Restoration Goals for Wetland Resources in the Greater Los Angeles Drainage Area

Charles Rairdan and Richard F. Ambrose
U.S. Army Corps of Engineers - Los Angeles District

ABSTRACT

The 2,320 square-mile Greater Los Angeles Drainage Area historically supported a broad diversity and extent of wetland resources that have been almost entirely lost or transformed by a system of upstream impoundments, channelization of lowland watercourses, depletion of ground water supplies, and widespread development.

Restoration planning efforts are now mostly site-specific, concentrated in the coastal zone, and do not adequately consider the functional linkages with adjoining ecosystems or the effects of urbanization in the contributing watersheds. Because it is unlikely that a significant likeness of the region's wetland heritage can be recovered, it will be important to instead optimize the functions performed by the remaining and restored systems.

Geographic Information Systems (GIS) was used in conjunction with a modified Hydrogeomorphic (HGM) classification method to reconstruct recent historic (ca. 1870) wetland resource conditions, and to perform a landscape-level comparison with current conditions to determine the nature and extent of the associated functional losses.

Results indicated the widespread loss of individual wetland resource functions within the categories of: hydrologic, water quality, nutrient cycling/food chain support, and habitat; and severe disruption of the broadscale functions of landscape maintenance and biodiversity, which relate the successful performance and continuity of the individual functions and the biological integrity of supporting ecosystems.

In response to these findings, a set of fifteen, function-based regional restoration goals was developed, the objectives of which are to rehabilitate key elements of the current landscape (hydrology, sediment and nutrient budgets, water quality, and habitat) with restoration measures that will enhance the long-term success and regional sustainability of remaining and restored

wetlands, and that will lead to the creation of new restoration opportunities in a highly constrained urban environment.

This study effectively demonstrates the capabilities of GIS in conducting landscape-scale analyses of multiattribute systems in support of regional and watershed planning activities.

INTRODUCTION

Background

Wetland losses in southern California, among the highest in the U.S., are on the order of 90 percent (Dahl 1990), and are frequently higher in coastal areas where development is concentrated. As momentum builds to restore the region's wetland resources, the range and intensity of stakeholder interests, the extent to which the landscape has been altered, and ongoing development pressures all act to hinder the coordinated progress of these efforts (Sorensen and Gates 1984).

Restoration planning and research in the region have typically focused on individual tracts of wetlands. Moreover, the functional linkages of coastal and inland systems have not been sufficiently recognized and incorporated into the planning process (Rairdan 1998).In order to promote a more cohesive approach to wetlands restoration planning in the metropolitan Los Angeles area, a landscape-level comparison of recent historic (ca. 1870) and current wetland resource conditions was performed with Geographic Information Systems (GIS) to serve as a basis for the development of a comprehensive set of function-based restoration goals.

This paper outlines the role of GIS in that research and provides examples of the methodologies and results generated. A full description of the study and its data are available in Rairdan (1998).

Study Area

The Greater Los Angeles Drainage Area (GLADA) comprises a 2,320 square-mile area and five constituent watersheds (Figure 1). The majority of the study area is covered by Los Angeles County, but also contains portions of Ventura, San Bernardino, Riverside, and Orange Counties. The current population of the Los Angeles metropolitan area is about 14 million people.

Figure 1.The greater Los Angeles drainage area and constituent watersheds.

The major landforms in the region include the high-ranging (3,000 to 10,000 feet elevation) interior San Gabriel mountains, the interior (San Fernando and San Gabriel) valleys, and the ring of coastal mountains (1,500 to 2,000 feet elevation) that encloses the coastal plain and geologically separates it from the interior valleys. At the transition between these two landforms are the mountain gaps (Glendale and Whittier Narrows) through which the Los Angeles and San Gabriel Rivers pass, and where ground water flows from the interior valleys historically re-emerged as surface flows during the dry season. In a similar manner, the Lower Santa Ana River emerges from a gap in the coastal range in the southeastern portion of the study area and formerly debauched its alluvial contents onto the coastal plain.

The orographic effect of the interior mountains and steep elevational gradient (approximately 5,000 feet over a 45-mile distance) of this coastal drainage result in periodic, high-volume, high-velocity flood flows (flashy hydrology) that accompany the large winter storms of the region.

The overall climate in GLADA is Mediterranean, marked by warm, sunny winters having widely variable precipitation, and hot, rainless summers moderated by coastal breezes. The diversity and arrangement of major landforms, however, coupled with the basin's proximity to the ocean, result in a set of distinct microclimates--ranging from marine conditions in the coastal plain to alpine weather in the interior mountains. The complex interactions of climate and major landforms, in turn, once resulted in a broad diversity of wetland ecosystems across the semi-arid landscape (Rairdan 1998).

Research Questions

Given the apparent extensive loss of wetland resources in the study area due to widespread development over the last 120 years, GIS was utilized to facilitate the resolution of the following research questions:

1. What were the recent historic wetland resource conditions in GLADA?

2. What are the current wetland resource conditions?

3. What has been the nature (in terms of ecosystem structure and function) and extent of wetland resource losses in the region?

4. How might some of these losses be reversed or ameliorated at the regional scale of planning?

METHODS

GIS Database

Applying the principles of landscape ecology and computer modeling techniques (Forman and Godron 1986, Haines-Young et al. 1993, Stow 1993), ArcInfo® GIS was employed to reconstruct recent historic and current wetland resource conditions. However, only through the availability of complete sets of early U.S. Coast and Geodetic Surveys (1859-1893; 1:10,000) and USDA Soil Surveys (ca. 1910), augmented by a variety of other manuscript maps, historic photos, and narrative accounts in university and regional archives, was the proposed research considered feasible. The period of circa 1870 for recent historic conditions was decided by the timing of *en masse* American settlement of the region, and the fact that most of the early landscape-level impacts to wetland resources occurred in the coastal zone, for which the original wetland configurations were known.

The standardization of the modeling scale (approximately 1:100,000) and wetland resource feature resolution (about a 5-acre minimum patch size) were determined by the soil survey map specifications, which included a scale appropriate for the level of detail needed for the study and which could be consistently reconciled among the various historic data sources and between recent historic and current conditions.

To reduce the amount of digitizing needed to construct the GIS model, USGS 1:100,000 hydrography data was imported from the Internet (http://nhd.usgs.gov/) and adapted as a template layer for both recent historic and current conditions. Because the interior mountain areas of GLADA had not been extensively altered by development, manual digitization of recent historic conditions was limited primarily to the lowland areas and, for current conditions, the revision of existing wetland resource features.

In addition to these principal data layers, supplemental layers of topography (DEM), land use, former artesian areas, and microclimatic zones were adapted from existing data sources to facilitate the landscape-scale analysis of the region's wetland resources.

To ensure topologic consistency between recent historic and current conditions, a system of cross-referencing the available source maps, and field verification of the mapped features and vestigial landforms (for recent historic conditions) was employed. Thomas Brothers® street maps were indispensable to accomplishing this task in the large, ubanized study area. Access to a full set of high-resolution aerial photographs, however, would have greatly facilitated the data and model validation process.

Wetland resource features were categorized in the GIS database as either point, linear, or areal feature types depending on how much was known about them geographically and how they were depicted on source maps. For instance, because little is known about the former extent of the region's vernal pool complexes other than their approximate location, these systems were identified as point features and, along with the other feature types, associated with a characteristic set of structural and functional attributes. Riverine wetland resources were mostly described as linear features, whereas marsh-like wetlands or systems having a discernable area were delimited as polygon features.

Once the wetland resource features were established within the GIS database and delineated with the below-described classification system, their relative quantities and distributions within GLADA, the fate (converted, extirpated, or remaining) and extent of recent historic wetland resources, and the extent of newly created wetland resources could all be determined by query and overlay analyses.

<u>Wetland Resource Classification System</u>

Wetland resource types were classified with a system adapted from Brinson's (1993) Hydrogeomorphic (HGM) method that grouped structurally-similar wetland resources into the regional classes of riverine, depressional, slope, and estuarine fringe. The wetland resource types within each class (Table 1) were further distinguished according to their functional differences along the landscape gradients of geomorphic setting, substrate and vegetation, water supply, and hydrodynamics. Table 2 presents an example of this classification scheme, which served, in effect, as the primary attribute table for the digital wetland resource features.

Table 1. Recent historic and current (in italics) wetland resource types by regional class occurring within the study area.

RIVERINE: Upper Riverine Dry Wash Lower Riverine Riverine Marsh Braided Lower Riverine *Soft-Bottomed Channel* *Concrete-Lined Channel* SLOPE: Slope Marsh Peat Marsh	DEPRESSIONAL: Vernal Pool Ephemeral Lake/Pond Depressional Marsh Non-Tidal Salt Marsh *Flood Control Basin* *Reservoir/Recreational Lake* *Spreading Ground* *Constructed Lake/Pond* ESTUARINE FRINGE: Tidal Marsh *Marina/Harbor*

Table 2. Example of the wetland resource classification system.

Wetland Resource	Geomorphic Setting	Substrate and Vegetation	Water Supply	Hydrodynamics
				RIVERINE:
Braided Lower Riverine	Lower reaches of mainstem channels characterized by a wide flood plain and meandering streams; wetlands occur in linear arrangements immediately adjacent to streams and in patchy distributions within the broader flood plain.	Medium- to fine-grained sediments; vegetation charac-terized by riparian shrub and tree communities on point bars and river banks.	Fresh; intermit-tent to seasonal; meandering and overbank flows, shallow ground water table.	High volume, low velocity flows result in wider flood plain and longer residence times; some ponding may occur.

Assessment of Wetland Resource Functions and Values

From the characterization of the physical attributes of the region’s wetland resources, and in conjunction with an extensive literature review, functions within the categories of hydrologic, water quality, nutrient cycling/food chain support, and habitat (Table 3) were related to the individual wetland resource types and evaluated with a qualitative rating scale as being either:“not supported”, “weakly supported”, “moderately supported”, or “strongly supported”.

Table 3. Wetland resource functions described in the study.

HYDROLOGIC: Flood Attenuation Ground Water Recharge Sediment and Bank Stabilization	WATER QUALITY: Sediment and Toxic Substance Removal Nutrient Removal
NUTRIENT CYCLING /FOOD CHAIN SUPPORT: Primary Production Decomposition Nutrient Export	HABITAT: Macroinvertebrates Herpetofauna Fish and Shellfish Mammals Perching Birds Waterfowl and Shorebirds

The broadscale functions of landscape maintenance (relating to the successful performance and continuity of the individual wetland resource functions along the landscape gradients of hydrology, sediment and nutrient budgets, water quality, and habitat) and biodiversity (landscape-level support of the richness and diversity of wetland-associated species) were also considered, along with the ecosystem values of aesthetics and recreation, education and research, and wetlands heritage with regard to their potential to further the realization of the regional restoration goals ultimately defined by the study.

In combination, these evaluations were performed to: (1) characterize the functional attributes (and ecosystem values) of GLADA's wetland resources; (2) assess the functional linkages of the different wetland resource types within the landscape; and, (3) to estimate the functional losses from recent historic times to the present. A complete description of these assessments is contained in Rairdan (1998).

RESULTS

GIS Maps and Data

The graphic results of the completed GIS database were displayed on 44" x 36" color maps to adequately capture the scale, resolution, and connectivity of the wetland resource features within the context of the broader landscape, while quantitative data were summarized in tabular form.

By comparing recent historic and current conditions, the former diversity and magnitude of the region's wetland resources were clearly revealed. Figure 2 provides an example of the dramatic differences in wetland environments that have occurred between the two resource periods, and illustrates the extent to which the landscape has been fundamentally altered from recent historic times.

An example of the fate and current status (converted, extirpated, remaining, and created) accounting of the region's wetland resources, generated by an overlay analysis of recent historic and current conditions, is provided in Table 4. In addition to these quantities, the types of conversion (from one wetland resource type to another) were also determined to ascertain the nature of individual wetland resource transformations (Rairdan 1998).

Table 4. Example of the fate and current status accounting of wetland resources in GLADA.

	Extent (in miles)					
	Recent Historic (Total)	Converted	Extirpated	Remaining	Created	Current (Total)
Lower Riverine	705	368	318	21	13	34

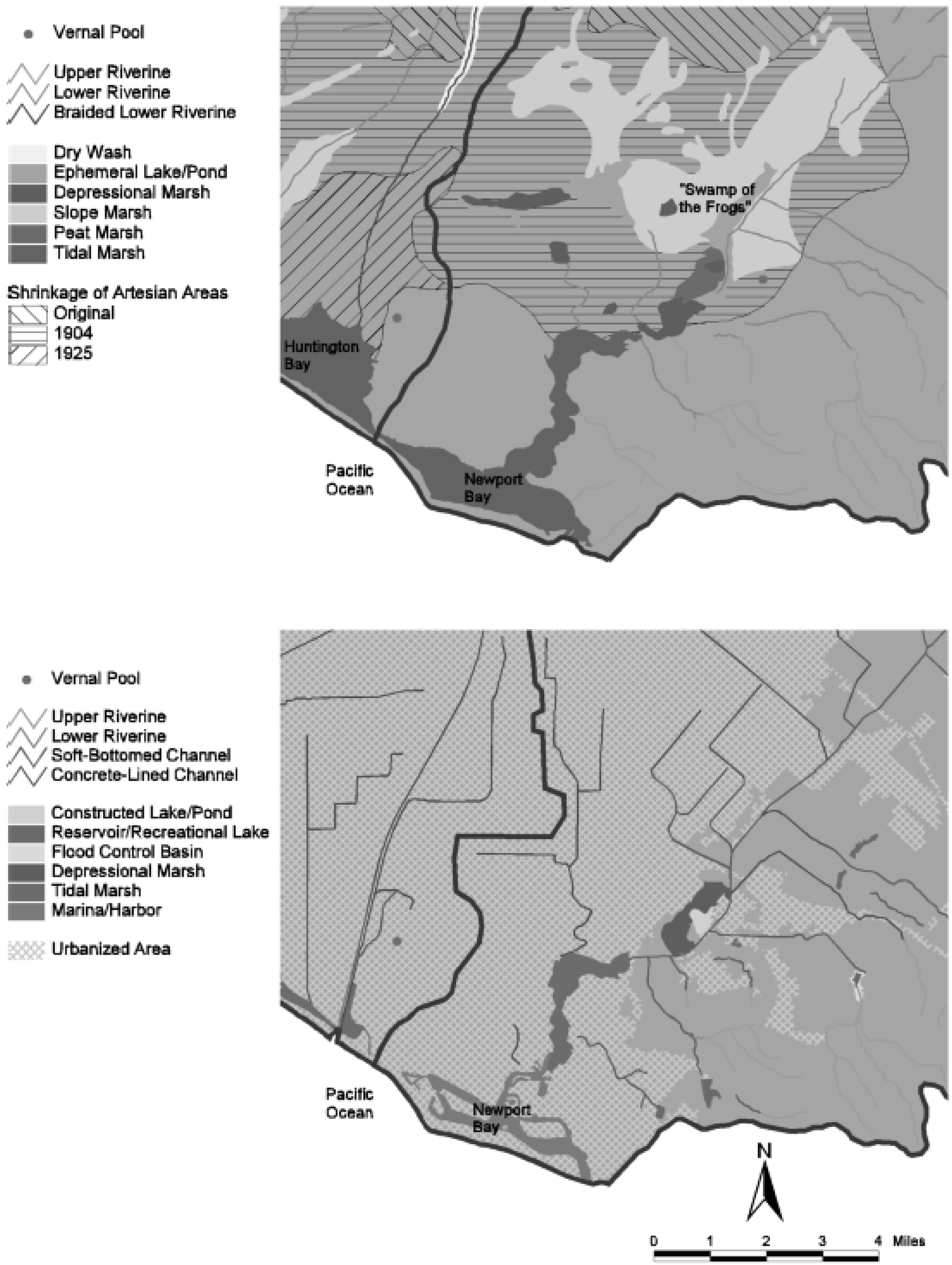


Figure 2. Gray-scale images of recent historic and current wetland resource conditions in the coastal reaches of San Diego Creek watershed.

From the comprehensive analysis of the GIS-based model, the estimated functional losses of wetland resources in GLADA can best be described as widespread and extensive for the individual functions, and as resulting in the severe disruption of the broadscale functions of landscape maintenance and biodiversity. This examination of the data also validated a main study premise that, at a regional scale of planning, the ecological integrity of the broadscale functions is paramount to the successful maintenance and performance of the individual wetland resource functions. This finding provided perhaps the greatest impetus for the development of the regional restoration goals, and for reinforcing the ideas that a landscape approach to restoration planning and implementation is necessary to achieve the optimum restoration benefits, and will more likely result in the restoration or creation of regionally sustainable wetland ecosystems.

Regional Restoration Goals

From the foregoing analyses and results, a set of 15 function-based regional restoration goals was derived. Table 5 presents an abbreviated description of these goals, which were developed in conjunction with a literature review of case studies in which similar landscape-scale goals and restoration measures had been either identified or implemented.

The order in which the goals are presented approximates the manner in which the individual wetland resource functions propagate within the natural environment and also considers the practical constraints imposed by the existing urban environment. In Rairdan (1998), the identification of these goals was followed by a discussion and examples of some of the ways in which the goals might be applied throughout the region. Finally, it should be recognized that, in order to be successful, these goals would have to be implemented in a coordinated fashion, and that their time-frame for implementation (20 - 50 years) is commensurate with the spatial scale at which they were developed.

DISCUSSION

The application of GIS in this study allowed for the development of an integrated, landscape perspective of the region's wetland resources and associated functional capacities. And although it was apparent at the outset of the study that the historic losses were substantial, the revelation of the former diversity and extent of wetland resources by the GIS database was nonetheless striking. The degree to which the current landscape has been altered had in effect obscured the perception of a past landscape that was capable of supporting a wide range of highly-productive wetland ecosystems.

Table 5. **Abbreviated description of the regional restoration goals.**

Goal No.	Function Type(s)[1]	Restoration Goal
1	Hyd	Increase wetlands with high flood attenuation capacity at strategic locations within the drainage.
2	Hyd, LM	Restore historic flow regimes to wetland ecosystems in conjunction with regional flood control measures.
3	Hyd, NC/FCS, LM	Restore the continuity of fluvial processes in each watershed by, for example, modifying upstream reservoir operations, restoring natural flood plain systems, and by active management of sediment and nutrient budgets along the watershed continuum.
4	Hyd, LM, B	Restore or locate wetland resources with high ground water recharge capacity or injection wells in areas that, in addition to replenishing ground water supplies, would benefit downgradient wetland habitats.
5	WQ	Restore wetland resources that are conducive to the improvement of water quality near storm water outfalls.
6	WQ	Implement effective storm water management practices in upland areas to reduce the levels of non-point source pollutants entering wetland environments.
7	WQ, B	Restore historic temperature regimes to the region's wetland resources (e.g., by modified dam operations, revegetation of riparian corridors, etc.).
8	All	Reduce the extent of concrete lined channel.
9	All	Establish riparian corridors either within or parallel to existing flood control channels.
10	Hab, B	Eradicate or control the propagation of invasive non-native species in existing and restored habitats.
11	Hab, B	Enhance the habitat values of flood control basins and water conservation facilities.
12	Hab, B	Restore/enhance lost and degraded tracts of tidal and freshwater marsh; increase the size and connectivity of viable habitats.
13	B	Create a network of freshwater inland habitats that function as stopover sites and species pools to larger, more sustainable wetland sites.
14	B	For species having poor dispersal capabilities, initiate restocking programs in conjunction with the creation/restoration of wetland habitats.
15	B	Increase the connectivity of existing and restored wetlands with adjacent and upstream ecosystems.

[1] B = Biodiversity; Hab = Habitat; Hyd = Hydrologic; LM = Landscape Maintenance; NC/FCS = Nutrient Cycling/Food Chain Support; WQ = Water Quality.

By compiling the available wetland resource data in the GIS database according to a consistent set of topologic and classification criteria, and by combining the primary data layers with other supporting themes, a landscape-scale model of the region's wetland resources was successfully generated. The completed model was then able to provide reasonable approximations of both recent historic and current wetland resource distributions, the fate and current status of these resources, and a tangible depiction of the complex interactions of climate, major landforms, hydrology, and land use in the creation and maintenance of the associated functional capacities.

Fundamental data gaps (e.g., only knowing the approximate location of historic vernal pool complexes or not knowing the lateral extent of most riverine wetlands) were successfully overcome by attributing the basic physical characteristics of each wetland resource type (i.e., its geomorphic setting, the typical distribution of wetlands within that setting, substrate/vegetative qualities, water supply, and hydrodynamics) to the corresponding point, line, or polygon feature type. The physical attributes were then coupled with the functional qualitites (Rairdan 1998) of each wetland resource type to evaluate their relative support levels and to assess the functional linkages of the region's wetland resources along the landscape gradients of hydrology, sediment and nutrient budgets, water quality, and habitat.

However, even when the areal extents of certain wetland resource types (e.g., estuarine and inland freshwater marsh) were sufficiently known, the qualitative nature and quantitative uncertainty of the model analysis required only an approximation of the magnitude at which the individual and broadscale functions were performed. In other words, for the purposes of defining a comprehensive and functionally-integrated set of regional restoration goals, we were more interested in assessing the processes by which the wetland resource functions perform within the landscape than in just producing a numerical accounting of estimated wetland resource extents.

Other key advantages of the GIS database were the ease with which it could be modified and refined during its development and subsequently navigated for analysis. The numerous methodological questions that arose during the concurrent tasks of data entry, development of the wetland resource classification system, wetland resource feature delineation, and attribute coding all required iterative and sometimes substantial changes to the GIS model as it was being built. Without the automated GIS capabilities, the heuristic qualities of the model and the manual implementation of its many changes would have likely proved too time-consuming and costly to warrant the construction of a non-digital version. Once the database was complete, the overlays and quantitative analyses and the production of large color maps for visual assessments were generally straightforward and justified the initial database efforts.

Another time-saving aspect of the GIS resulted from the application of the principles of landscape ecology to the digital technology. For example, by definition, a landscape consists of a heterogenous mosaic of interacting ecosystem, or wetland resource, units that are characterized by spatially distinct landforms and a number of other interrelated physical and biological factors (Rairdan 1998).This definition was instrumental in determining where the boundaries between the different wetland resource types occurred (e.g., the transition between upper and lower riverine systems at the outlets of mountain canyons) within their characteristic landscape or geomorphic setting. By verifying in the field what the typical landscape configuration was at the boundary between adjoining wetland resource types and adjacent upland ecosystems, these regionally consistent relationships were employed to delineate the wetland resource boundaries without having to physically inspect every resource juncture across the 2,300 square-mile study area. This technique also ensured greater data consistency between recent historic and current conditions than could otherwise be achieved by cross-referencing the available manuscript maps, narrative accounts, and historic photographs.

In conclusion, the GIS model facilitated an in-depth understanding of both past and present wetland resource conditions in the greater Los Angeles drainage area and ultimately led to the development of 15 function-based restoration goals that, if properly implemented over time, would result in the increased functional capacities of remaining wetland habitats, the creation of substantial new restoration opportunities in a highly constrained urban environment, and the long-term sustainability of the region’s wetland resources.

REFERENCES

Brinson, M. M.1993.A hydrogeomorphic classification for wetlands. U.S. Army Corps of Engineers, Washington DC. Technical Report WRP-DE-4.

Dahl, T. E.1990.Wetland losses in the United States, 1780's to 1980's. United States Department of the Interior, Fish and Wildlife Service, Washington, DC.

Forman, R. T. T., and M. Godron.1986. Landscape ecology. John Wiley & Sons, Inc., New York, NY.

Haines-Young, R., D. R. Green, and S. H. Cousins.1993.Landscape ecology and spatial information systems. pp. 3-8.*In* R. Haines-Young, D. R. Green, S. H. Cousins (eds.).Landscape ecology and geographic information systems. Taylor & Francis, London, England.

Rairdan, C.1998.Regional restoration goals for wetland resources in the greater Los Angeles drainage area: a landscape-level comparison of recent historic and current conditions using geographic information systems. Dissertation, Environmental Science and Engineering Program, University of California, Los Angeles.

Sorensen, J., and S. Gates.1984.New directions in restoration of coastal wetlands. pp. 1427-1443.*In* O. T. Magoon, H. Converse, and L. T. Tobin (eds.).Proceedings from the third symposium on coastal and ocean management .American Shore & Beach Preservation Association/California State Lands Commission, Sacramento, CA.

Stow, D. A.1993.The role of geographic information systems for landscape ecological studies. pp. 9-21.*In* R. Haines-Young, D. R. Green, S. H. Cousins (eds.).Landscape ecology and geographic information systems. Taylor & Francis, London, England.

Use of GIS for the Investigation and Classification of Land Redeveloped Under the Ohio Voluntary Action Program (VAP)

Andrew Rawnsley, Ravensfield Geographic Resources, Ltd.

Craig A. Kasper, P.E., and W. Lance Turley, P.G., Hull & Associates, Inc.

Ohio Voluntary Action Program, Background

In an effort to entice developers, business owners, and investors to redevelop brownfields, the federal government and a number of state governments instituted various brownfield redevelopment programs. They allow communities, property owners, and developers of brownfield properties to voluntarily assess and remediate environmentally contaminated sites. Many states also provide that sites which successfully participate in these programs are released form future state environmental liability, eliminating one of the chief barriers to private investment and site redevelopment.

Ohio's Voluntary Action Program (VAP) (Ohio Revised Code 3746) establishes a new approach to traditional environmental cleanup programs. The new program offers a number of incentives and new standards to prompt investigation and cleanup of contaminated properties or properties suspected of contamination. The establishment of such a law should move Ohio several steps forward by effectively joining environmental management with business development. The VAP offers a mechanism to address contaminated or potentially contaminated sites in Ohio, sites that might not otherwise be addressed. The program provides practical standards that consider the future use of the property and cost/benefit in developing the standards, and provides a mechanism that ends the owner's responsibility for further cleanup of the property.

A provision of the VAP is the Urban Setting Designation (USD), which designates areas in cities or urban townships where groundwater is not used for drinking water purposes, and is not required to meet drinking water quality standards. This reduces or eliminates the need to remediate groundwater at a brownfield site. Groundwater within the USD may need to meet other water quality standards.

Given the breadth of the program, applications for GIS in the redevelopment of properties under this sort of regulatory umbrella are quite extensive, although, as is the case with many environmental projects, cost considerations can limit its use short of its full potential. This discussion centers on two specific case studies, a complex investigation, construction, and development effort involving a major industrial facility, and an equally complex scenario involving the designation of a large portion of the City of Cleveland as a USD area. Both projects were very public and politically charged, adding several very pervasive concerns to already technically challenging projects.

An important issue with both of these case studies is that GIS and other mainstream data management techniques were integrated from the beginning. Effective data management of any kind is extremely difficult to shoe-horn or retrofit into an existing project. This is not unique to regulatory projects; it is the classic “collection/conversion of the data is 75 percent of the cost” scenario endemic to all GIS projects. With a long-term management system such as one developed by utility or municipal government, this is an accepted portion of the cost of a project. Many environmental projects, aside from ones that involve massive cleanup, are often on the fringe of affordability for GIS application. Advanced GIS will simply not hold up in a cost/benefit if it has to overcome the initial burden of extensive data conversion. The projects are simply more finite in scope, and the use of GIS must be made be as efficient as possible. To do so, it is of paramount importance that the data management issues be resolved and planned for as part of the original scope of work, in the same way that sampling and analysis or QA/QC plan is part of the original scope. In addition, while remediation activities are typically long-term, investigations, particularly ones involving redevelopment, are typically accelerated, even more so if a construction or resale schedule must be met. Under these conditions, data management issues must be planned for in advance, or their full benefit will not be realized.

The specific goal of either construction or resale is one of the primary things that differentiates many VAP projects from a standard site investigation. While redevelopment of a site may not be the specific goal of all VAP projects, most do have that goal, and economic pressures are added to regulatory ones. While under investigation or remediation, property is essentially idle, which if not costing money is at least not making any. As will be seen in the Jeep case study, construction schedules are often immutable. In these situations GIS is a valuable tool, if planned for in advance. The entire purpose of effective data management is about standardization of collection, form, and retrieval, allowing for increased efficiency and consistency. In situations with compressed schedules, it is extremely useful.

Another of the unique aspects of GIS use in this sort of redevelopment is that it combines several classic GIS applications, from more economic, infrastructure, and property oriented analysis to the modeling and data management required in extensive environmental site investigations. In addition, the GIS professional must prepare for any number of fringe roles or applications, such as construction management and high-profile, public projects where geopolitical relationships must be understood.

Roles of GIS in Environmental Redevelopment

VAP projects vary considerably in scale and complexity. While some may involve fairly standard Phase I style site assessments, where GIS is limited to searching and presenting information from standard databases, some are full-blown Phase II investigations involving several hundred borings, wells, or other sampling points and tens to hundreds of thousands of sampling results. In projects of this scale, effective data management is critical, and use of GIS as a keystone of a data management strategy solves many problems. The reasons for this are fairly obvious – everything that happens in a site investigation occurs at some point in space.

One cannot in all honesty draw neat lines around the complex roles GIS can play in an any situation, particularly in an environmental investigation, although for sake of organization everyone tries. This discussion is no different, and centers on the management, analytical, and presentational roles of GIS in environmental investigations. The discussion does not, however, spend too much time on the role of specific technology, for the basic reason that it is relatively unimportant. There are few tasks that are specific to a specific software package, and no software package functions with complete satisfaction over a wide range of applications. In situations such as environmental redevelopment projects, with compressed budgets and time schedules, one uses the tools one has in the most expedient manner possible.

Management Roles

The management aspect of GIS in these situations cannot be underestimated. Much of the GIS literature slights management in favor of analysis; however, one cannot analyze prior to proper collection and arrangement, and certainly one cannot analyze *repeatedly* without proper collection and arrangement. In addition, GIS tends to be too costly when applied as a surgical tool, particularly in environmental projects where *any* expenditure is likely to displease someone. While unique analysis and visualization may pique someone’s interest, long-term efficient management often convinces clients to use GIS.

Repeatability is stressed in the above paragraph for a reason: redevelopment projects under a regulatory framework such as the Ohio VAP are often either governed by sets of standards, or are risk-based. The choice of what standard or risk level to target for a specific property is mostly, if not wholly, a cost/benefit analysis. Different combinations of remediation and construction scenarios are weighed against various standards or risk levels to produce the optimal redevelopment strategy. One must be able to adjust and repeat analyses in an efficient manner. Consistency and predictability of error are two of the main by-products of proper data management.

Again, this centers on cost. As mentioned, many environmental projects are on the fringe of affordability for GIS application. Redevelopment projects are unique in that there is some component of added value to the investigation, and not viewed only as a cost. They are still not, however, value-added to the degree that a government or utility GIS is, and is considerably more finite in scope. In short, to the environmental professional the utility or benefit of GIS is not always as immediately apparent, or at least not as obvious when costs are added. This becomes a considerable part of the GIS professional's challenge when working in this area.

Virtually everything involved in an environmental investigation involves space and location. This makes GIS the ideal tool for managing at least the spatial component of any site data. These site data include:

- Aerial photographs.
- Infrastructure (roads, rail, utilities, etc.).
- Topography.
- Geologic/hydrogeologic features.
- Sampling points.

In addition, particularly true with regards to construction projects, GIS analysts may be required to be the de facto coordinate transformation authority. In general, GIS analysts and operators abhor local coordinate systems, and many state agencies are requiring coordinates in accepted standard coordinate systems (State Plane, UTM, etc). Contractors and surveyors, however, are strongly attached to them. It is not a duty to be taken lightly; in fact it is one to be avoided, as it may expose the GIS operator to unneeded liability.

While some ancillary attribute data is stored internal to the specific GIS, a complex data such as analytical results are typically stored in an external database system of some type. The ability of most GIS systems to link to external data sources, therefore becoming an effective clearinghouse for any type of information, is one of its strong suits. The performance and utility of individual GIS packages in this regard vary.

Analytical Roles

The analytical roles for GIS in redevelopment typically fall into two general categories, one ‘introverted’ (analysis of the site itself) and the other ‘extroverted’ (analysis of what surrounds the site). Each of the case studies exemplify one or the other of these concepts. The Jeep project is an extensive site investigation for purposes of meeting regulatory risk standards, while the Cleveland USD is a optimization of an areas meeting regulatory criteria.

Of the two general categories, the first is often more complex, as it often involves longer-term projects and greater amounts of information. Typical analytical processes in these situations may be considered a subset of more general procedures done at any environmental site: modeling contaminant distribution and travel, ground-water conditions, site geology, impacted wetlands, etc. Again the goals of a VAP project will color what one does and how one does it. Sites slated for redevelopment or resale have a different potential than one which is merely seeking alternate remediation programs. The analysis often centers on analyzing risk levels to meet certain regulatory goals in a fairly direct cost/benefit sort of fashion.

The second, ‘extroverted’ type of analysis, typified in the VAP by USD projects, resembles a classic facility placement exercise more than an environmental management system. One is attempting to either maximize a placement of a ‘facility’ (in this case a USD area), or prove that a current ‘facility’ meets specific guidelines, based upon regional conditions. Data management is not as critical, as it is more of a one-time, surgical procedure than a long-term need.

Presentational Roles

The presentation role that GIS may play in these sort of projects is critical, and one not always completely acknowledged. In short, the data must be presented. Site layouts, analytical results, standards comparisons, impacted areas: all of it must be shown at some time to *someone* who will need to understand it, whether it be a property owner, a professional, a regulatory representative, or the public.

Lost in all the hype in the “digital revolution” and typical GIS marketing literature is the fact that a paper map is still the most functional product of a GIS, and will continue to be so. Both case studies discussed in this paper produced large numbers of unique large-format maps, into the hundreds for the Jeep project. When discussing design issues, one quickly shifts from the mechanical to the perceptual, and as many, many pages of print have been consumed with the topic, little more will be said here. Effective design is *not* a mechanical process, and that ‘unscientific’ fact does not diminish the importance of effective design and presentation.

Chrysler Corp Stickney Ave. Jeep Plant

Case Studies

Jeep

The Chrysler Corporation operates two plants in the City of Toledo: the Jeep Parkway Plant and the Stickney Avenue Plant in North Toledo. Toledo has been home for the production of the Jeep automobile since 1940 when Willys Overland developed the first prototype and began to mass produce the Jeep military vehicle in Toledo for use in World War II. Chrysler is the city’s largest employer, providing over 35,000 high-paying direct and peripheral jobs, and generates millions in tax revenues. Clearly, the loss of Chrysler would not only devastate the Toledo regional economy, but would also have a direct impact on the state’s economic well-being.

In late 1996, Chrysler announced its plans to construct a new facility to expand production and replace the out-of-date Jeep Parkway Plant. After a year of intensive studies and assessments of the Stickney Avenue property and infrastructure, and a relentless campaign by the City of Toledo, Chrysler announced it would stay in Toledo and invest $1.2 billion into the expansion and production of its Toledo operations. The city is contributing $27.5 million for roadway, environmental, and other infrastructure improvements. The Stickney Avenue property, which possesses much of the necessary infrastructure such as adequate rail lines and access to major highways, contains 400 acres of land and 13 businesses. Land use of the 400-acre property includes railroad operations, aluminum smelting, a brass foundry, auto repair, and various other commercial activities.

Phase I and Phase II Environmental Assessments are being conducted at all the properties, whether acquired or not. Remedial measures will be integrated with plant construction to take advantage of “dual-purpose” activities and engineering/institutional controls. The city is also in the process of acquiring and relocating 83 residential properties, relocating a railroad spur, and addressing a variety of environmental issues, including 43 acres of jurisdictional wetlands, of which 25 acres will be impacted by the plant expansion. Although environmental issues often represent only a small portion of a large-scale industrial redevelopment project such as this one, they are critical in the overall plans to determine the most economical, yet protective solution for a variety of property conditions.

Due primarily to availability of a comprehensive GIS, RGR was able to assist HAI in rapidly assessing how changes in remediation goals would influence the project schedule and cleanup costs. HAI was in turn able to alert the City to new investigative requirements and cleanup strategies that would fit within development time frames. The primary challenge for this project centered on *time*. Once Chrysler made the decision to build, the construction schedule was quite aggressive, and the environmental investigation, while minor in comparison to the massive construction plans, had to happen prior to any construction. Any delays due to the environmental investigation would not have been received well.

To reiterate, it is difficult to draw neat lines around the roles GIS can play in a project of this nature. In fact, it is a measure of the success of a GIS implementation if one cannot. If the use of GIS is essentially transparent through all phases of the project, and one cannot easily tell where it starts and ends, then it is a success. If it is not so pervasive, and one can therefore easily distinguish its presence, then it may be necessary to reevaluate how and why GIS was used, with particular attention paid to project planning stages. Again, for sake of organization, all function and utility of a GIS flows from its management role. Transparent presence use and pervasive use is a direct consequence of a planned, central management role in a project.

Project Details

All spatial data collected in the investigation processes was managed with ARC/INFO. RGR managed the extensive analytical sample information using a custom database system with a web-based front end. AutoCAD Map was used frequently for complex editing, as ARC/INFO is deficient in this area, and Arcview was used for convenience from time to time. Map finishing was done in Corel DRAW. While this appears on the surface to add step and an extra layer of management, and therefore extra time in a time-sensitive project, the trade off was well worth it. Professional publishing software has far better fine-grained page control than either ARC/INFO or Arcview, and the quality of the final maps was very important as they were window from the GIS to professionals, public officials, and the public itself. So while basic layout of the GIS data was done using ARC/INFO's layout manager, final text, color, symbolization, map layout, and printing was done in Corel.

A list of the data collected for this project includes:

- Aerial Photographs
- Transportation (roads, rails)
- Site structures
- Parcel boundaries
- Surface hydrography
- Surface topography
- Sampling locations
- Surveyed wetlands
- Impacted soils areas

HAI completed Phase I Environmental Assessments at all properties identified for potential delivery to Chrysler. The assessments determined past management and disposal practices for hazardous substances, areas of potential contamination, and other environmental concerns, cumulatively making up VAP identified areas. HAI also conducted preliminary evaluations to determine eligibility of the identified areas for the VAP.

HAI and RGR conducted GPS and total station surveys of identified areas, entered their spatial attributes and descriptions into a GIS and presented the areas on an aerial photograph that also included property lines, wetland delineations, initial plant designs and proposed infrastructure

Soils sampling points and related analytical data

improvements. The aerial photograph formed a base map that was used to show proposed sampling locations, findings from various phases of investigation and changes to proposed land use as the project progressed.

HAI developed a scope of work for an initial Phase II investigation that stressed characterization of site hydrogeology and sampling of "hot spots" within the identified areas (i.e., location within identified areas where evidence of contamination was the strongest). Soil analytical samples were preferentially collected from within two feet of grade – the point of compliance for direct contact in the context of industrial land use. Provisions were made for sampling at greater depth based on field observations and screening results.

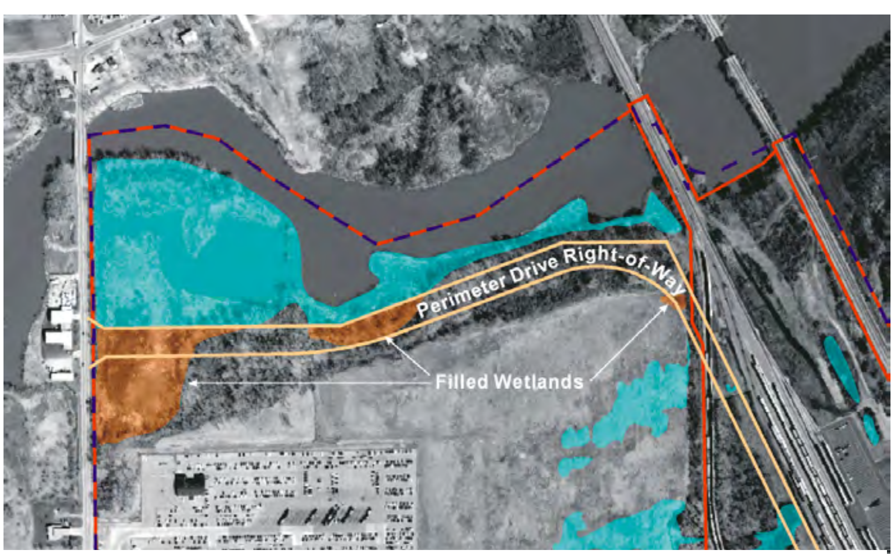
The laboratory delivered Phase II analytical data directly to RGR, who immediately incorporated into the GIS integrated sample management system. In general, the data could be exported by RGR from the GIS and sent to HAI or shared with Chrysler's consultants as spreadsheets or on maps (e.g., isoconcentration contours for arsenic) within hours of RGR receiving an electronic report from the laboratory. Due to restrictions on HAI's ability to access some of the properties, multiple changes in the proposed footprint of the plant (daily at times), tight deadlines for investigation and remediation of the properties, and the need to present findings to a number of interested parties, the flexibility and rapid data management capabilities of the GIS proved invaluable to the ultimate success of the project.

Following an initial Phase II investigation at a given property, HAI evaluated the data and, where necessary, collected additional samples to delineate areas of concern. Site-specific cleanup standards were developed based on multiple chemical adjustment for all chemicals of concern (COCs) detected on all properties combined. Data evaluation was iterative because as additional properties were investigated new COCs were detected; therefore requiring revised cumulative adjustments. The GIS database was an essential element in comparing and manipulating large data sets and mapping areas requiring remediation.

The initial Phase II investigation identified direct contact with soils containing metals (primarily arsenic) and semi-volatile organic compounds (SVOCs) as being the driver for cleanup at most of the properties. At the project's onset, the City of Toledo envisioned that plant foundations and parking lots, which per Chrysler's design covered approximately 95 percent of the proposed plant property, would serve as engineering controls for limiting exposure to soils exceeding direct contact standards. During early phases of data analysis, HAI and RGR worked closely with Chrysler's engineers to identify final construction grades to map and identify volumes of soils requiring excavation during construction activities and additional volumes that would be excavated to preserve a two-foot point of compliance. However, as plant design proceeded, Chrysler decided that leaving contaminated soils beneath engineering controls would be unacceptable due to the following reasons:

- Future changes in facility operations and structures over an approximately 100-year lifetime for the plant could result in contaminated soils being brought to the surface.
- Contaminated soils brought to the surface would require time-consuming and costly administrative tasks.
- Operations and maintenance requirements within the VAP would create an additional tier of administrative responsibilities.

Therefore, Chrysler and the City prepared a development agreement specifying that cleanup would generally conform to VAP standards, but that soils containing COCs at concentrations exceeding standards be remediated regardless of whether they were below the point of compliance.



Wetlands impacted by a construction scenario

In addition to the soils investigation, HAI assessed impacts to approximately 45 acres of wetlands, knowing that a portion of them would be destroyed. Existing wetlands were delineated using GPS, and a coverage created from that. The wetlands layer was then overlaid repeatedly on new plant footprints and road designs, trying to find the optimal mix of minimizing wetland impacts and maximizing plant design.

A by-product of proper data management, in this case unplanned, is the ability of the system to behave as a comprehensive clearinghouse in projects such as this, where a wide variety of consultants worked on the project. This is similar to a large Superfund site with a large PRP group, where each requires a comprehensive view of the site, but cannot achieve this on their own. A properly managed GIS in that case becomes the de facto repository for information, and its ability to transform data from format to format and coordinate system to coordinate system becomes critical.

The Jeep project had a wide variety of professionals working on it, such as the plant process designers, wetland scientists, construction contractors, and geotechnical consultants. Each has its own (for lack of a better term) world view, and sees the data, or needs to see the data, in wildly different ways. In addition, there were three separate coordinate systems in use at the site, two local ones and State Plane. Any GIS management and analysis system required to be a clearinghouse in this situation had to be both responsive and nimble to adjust to the continual requests from different sectors of the project.

Cleveland Urban Setting Designation

Urban Setting Designations are sub-programs within the Ohio VAP that designate areas within cities and urban townships where groundwater is not consumed. While they are frequently sought by individual property owners similar to a VAP No Further Action, they are increasingly being sought by cities a development incentive, providing blanket coverage for large areas within their boundaries to spur redevelopment of brownfield areas.

A portion of the Industrial Valley section of the

On May 10, 1999, the Ohio Environmental Protection Agency approved a USD submitted by the City of Cleveland, covering approximately 11,500 acres, or nearly a quarter of the city. Neighborhood Progress, Inc., and MacLaren-Hart, both of Cleveland, oversaw the project. RGR provided the GIS support for the project. The final submittal consisted of 6 separate areas across the city (West, Northeast, Southeast, Inner East Side, Inner West Side, and Industrial Valley).

For approval, a USD submittal must prove the following (summarized from Ohio Revised Code 3745-300-10 *Ground Water Classification and Response Requirements*):

> (a) The property or properties for which designation is requested is entirely within the boundaries of a township with a population of twenty thousand or more residents in the unincorporated area of the township or entirely within the corporation boundaries of a city;
>
> (b) Not less than ninety percent of the parcels within the city or township where the property or properties for which designation is requested is located is connected to a community water system.
>
> (c) The property or properties for which designation is requested is not located in a wellhead protection area.
>
> (d) Wells installed or used for potable water supply purposes are not located within one-half mile of the property boundary of the property or properties.
>
> (e) When the property or properties is located over a sole source aquifer in a consolidated saturated zone or an unconsolidated saturated zone capable of sustaining a yield greater than one hundred gallons per minute, the certified professional, must demonstrate that there is not a reasonable expectation that there will be any wells installed or used for potable water supply purposes within one-half mile of the property boundary.

All of the above stipulations are spatial phenomenon, and can be addressed in a GIS. (a), (c), and (e) are relatively simple in that they only require a single overlay, and (b) is often solved by questioning the local water authority. (d) involves a greater amount of effort, as it involves collection of potential well locations from a number of sources (the Ohio Department of Natural Resources (ODNR), local health departments, local water authorities, etc.) and a fair amount of address matching.

GIS proves itself in the well location phase of a USD. The USD requestor must verify that any possible well location within the half-mile radius is not being used. The records for the well locations can be very old (pre-1950), and wildly inaccurate in terms of actual location. Without 'pre-location' via GIS, field workers would require an inordinate amount of time sorting and verifying potential well locations.

Data Sources

The City of Cleveland, from the Cuyahoga County Auditor, provided both the base parcel map used for the project, and the coverage containing the ward lines. In addition, Cleveland State University provided a data set based upon the same materials, the difference between the two being that the CSU data has been re-projected into State Plane while the original was in a local county coordinate system. RGR addressed-matched the well locations using the Dynamap 2000 product from Geographic Data Technologies (GDT), of Lebanon, New Hampshire. Well locations were compiled from several sources: parcel ID's from the Cuyahoga County Health Department, and Ohio Department of Natural Resources (ODNR) wells logs located by address, coordinate, or standard well logs.

Procedures

RGR compiled the individual USD areas by aggregating parcels from the base materials. Unlocated potential well locations were address-matched against the GDT Dynamap 2000 database for Cuyahoga County using ARC/INFO. The well screening procedure left field workers with 83 potential well locations to verify. Of these 83 potential locations, 2 were found to still be in operation (original estimates had no wells still in operation). The preliminary USD areas had to be redrawn around the two wells.

Final Products

USD requestors are required to provide an acceptable legal description of the area. There was considerable debate as to what was an 'acceptable' legal description in this instance. On other USD projects RGR had been involved with, either the number of parcels was small enough that a complete legal description was feasible, or the areas were regular enough in shape that a legal description of the outside boundary was possible. Neither was practical in this case. The number of parcels involved (@14,000) made individual descriptions impossible, and the outside boundary of the areas was too complex. The original set of paper maps produced by RGR was deemed unacceptable to the State, and the State's suggested modifications were deemed too expensive by the project coordinators. While on the surface it seems that a parcel list compiled from the GIS would work, in reality a permanent parcel number is not, in fact, permanent, due to frequent splits and parcel combinations. In the end, the parcel list combined with the actual GIS data was accepted as the legal description for the area.

References

Kasper, Craig A., *Chrysler-Jeep Industrial Redevelopment Project, Toledo, Ohio*, The 27th Annual Conference on Environmental Law, March 12-15, 1998.

Using GIS To Identify Linkages Between Landscapes and Stream Ecosystems

Carl Richards, Lucinda Johnson, and George Host
University of Minnesota, Duluth, Minnesota

Introduction

Factors that operate on a variety of temporal and spatial scales influence the structural and functional components of stream ecosystems (1). Quantifying the effects of factors that operate across multiple scales has challenged aquatic scientists over the last several decades. Recently, scientists have recognized that they cannot successfully protect or restore ecosystem integrity without taking into account all appropriate scales; therefore, they are focusing on understanding interactions between terrestrial and aquatic components of entire watersheds (2). Although awareness of the importance of watershed and landscape-scale influences on streams is growing, the tools to examine these influences are still in their infancy.

Most watershed and landscape studies to date have focused on the role watershed-scale parameters play on water chemistry (3-5). These studies usually examined nutrient and sediment inputs from various watershed land covers. Methods for evaluating the patterns in the terrestrial segment of the watershed were awkward and laborious, involving use of planimeters or cutting and weighing maps. More recent watershed studies have attempted to integrate both longitudinal and lateral influences of the terrestrial ecosystems on water quality in streams and wetlands (6-8). This approach takes advantage of newly available tools (geographic information systems and multivariate statistics) for quantifying landscape structure.

Relatively few studies have examined how watershed features influence biological communities. Most studies examining stream biota have concentrated on single land use types (9, 10) or on the relationship between watershed land use and stream physical habitat (11, 12). Typically, study designs have not addressed questions concerning variability of stream communities over relatively large geographic scales.

Our own work centers on identifying linkages between landscape features (watershed scale) and stream reach environments (physical habitat, chemistry), and relating these parameters to major patterns of community variation. In this manner, landscape and reach environment interactions probably control the influence that specific landscape components have on biological communities (13). This is the general premise when using biological communities to assess watershed status. To represent stream biota, we examine benthic macroinvertebrates, which have been used extensively for biomonitoring numerous environmental stresses (14). Macroinvertebrates are sensitive to watershed conditions and exhibit sufficient stability in assemblage structure over time to make them useful as long-term monitors of stream health (15).

This paper presents an overview of our attempts to identify the relative strengths of landscape variables on macroinvertebrate communities. We classify landscape variables into two general categories. The first category, geology and landscape structure (GEOS), considers variables that are fixed on the landscape and are largely uncontrollable by management activities. The second category, land use (LU), includes variables that have anthropogenic origins and may be influenced by land management activities. By understanding the relative strengths these two sets of variables possess in determining community structure, we hope to identify specific species groups that can act as land use and land form indicators. We also hope to identify ways to predict the outcome of specific large-scale land management activities (e.g., silviculture, agriculture) or other large-scale environmental changes (e.g., global warming) on stream ecosystems.

Study Area

This study was conducted in the Saginaw River basin, a 22,562-square-kilometer watershed in east-central Michigan (see Figure 1) that flows into Lake Huron. The Saginaw River watershed was chosen for this study because its component drainages range from heavily affected agricultural to relatively pristine areas.

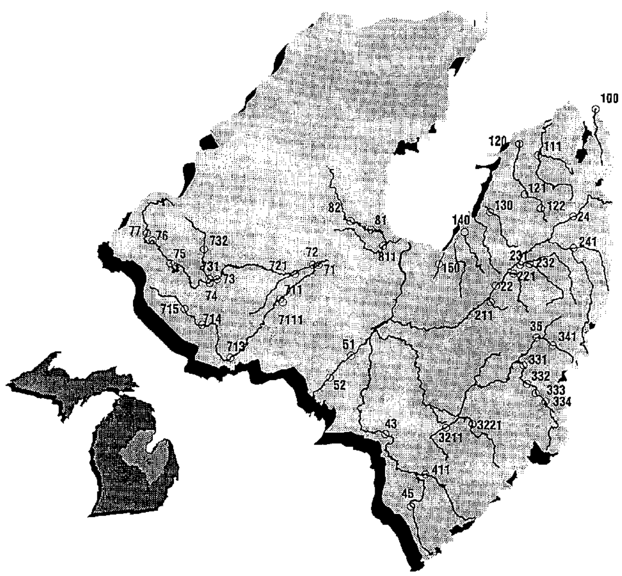


Figure 1. The Saginaw basin study area.

Dominating the soils in the lake plain are medium- and fine-textured loams to clays, with sand found in the outwash plains and channels. Artificial drainage and tile systems extensively drain the clay regions. Glacial features such as ground moraines and outwash plains are common. The western sector is characterized by rolling plains with coarse-textured ground moraines. This region contains a high percentage of the forested land, while agricultural land use dominates the eastern sector.

The Saginaw basin covers 16,317 square kilometers, including four major subbasins: the Tittabawassee (6,734 square kilometers), Shiawassee (3,626 square kilometers), Flint (3,108 square kilometers), and Cass (2,331 square kilometers) Rivers. The Tittabawassee subbasin further divides into three principal watercourses—the Chippewa, Pine, and Tittabawassee Rivers. Watersheds adjacent to Lake Huron (Kawkawlin and East basins) are characterized by low topographic relief and elevations averaging 203 and 206 meters, respectively. The Flint and Chippewa/Pine basins average about 278 meters in elevation. These drainages also exhibit the greatest variation in topography.

Study Design

The analysis covers 45 stream sites within the study area. These sites reflect a gradient of land use and physiographic conditions in the Saginaw River drainage. Researchers obtained biological, chemical, and physical samples at one 200-meter stream segment at each site. In addition, a geographic information system (GIS) database was compiled reflecting a number of landscape parameters for the watershed of each stream segment.

Sampling Methods

Macroinvertebrates

At each sampling site, we deployed Hester-Dendy artificial substrate samplers (16) for macroinvertebrate community characterizations twice, in early summer and during base flow conditions in late summer and fall of 1991 or 1992. We allowed samplers to colonize for 6 to 8 weeks. In the laboratory, macroinvertebrates were counted and identified to genus whenever possible. A series of derived variables from the original species abundance tables was used to describe community characteristics. We chose metrics based on their relative utility for examining macroinvertebrate communities, as suggested by Barbour et al. (17) and Karr and Kearns (18). Because macroinvertebrate assemblages are relatively stable through time (15) and preliminary analysis indicated no significant differences between sampling years at stations for which we had 2 years of data (unpublished data), we combined macroinvertebrate data into one database.

Chemistry

We assessed nutrients and other chemical properties related to water quality at each stream site during several periods in the summer and fall of 1991 or 1992. Stream flow during fall sampling was typically less than median flow rates and was considered to represent base flow levels. We used the maximum values of samples taken in June and July to represent summer conditions and the maximum values from September and October to represent fall base flow conditions. The nutrients measured were ammonium (NH_3), nitrate-nitrogen, total nitrogen (TN), orthophosphate (PO_4), and total phosphorus (TP). In addition, we assessed alkalinity (ALK), conductivity, total dissolved solids, and total suspended solids (TSS). Standard methods were used for all measurements (19).

Physical Habitat

We assessed physical habitat during base flow conditions at each stream site in a stream reach that is at least 8 to 12 times the width of the stream segment. A suite of quantitative habitat structure measurements and observations was made at each site. We derived values for six general habitat attributes:

- Substrate characteristics
- Instream cover
- Channel morphology
- Riparian and bank conditions

- Riffle/Run quality
- Pool quality

Landscape Descriptors

Land use patterns, surficial geology, hydrography, and elevation databases helped to quantify landscape characteristics in the study area (see Table 1). Land use patterns were derived from existing digital data at the Michigan Department of Natural Resources (Michigan Resource Information System [MIRIS] database) (see Table 2). We based classification of land use/cover categories on a modified version of the Anderson (20) scheme, which was constructed specifically for natural resource applications. The result was the following nine land use/cover categories:

- Urban
- Row crop/agriculture
- Other agriculture
- Herbaceous range land
- Shrubby range land
- Nonforested wetlands
- Forested wetlands
- Mixed hardwood forests
- Deciduous forests

In this region, nonrow-crop agriculture is largely represented by pasture, and range lands are predominantly abandoned fields (old fields).

The U.S. Department of Agriculture (USDA) STATSGO soils database enabled the compilation of soil data. The database consists of U.S. Soil Conservation Service (SCS) soil surveys and includes information on dominant texture and drainage in large landscape units. We aggregated soils into simplified categories based on glacial landform.

We delineated watershed boundaries for each sampling station on United States Geological Survey (USGS) topographic maps and digitized them using ARC/INFO (Environmental Systems Research Institute [ESRI], Redlands, California). We identified stream order for each stream segment and coded it as an attribute of the stream reach file. All databases were transformed into a common digital format as necessary and projected into a common coordinate system. We stored data in vector format and analyzed them in ARC/INFO.

Table 2 lists the landscape variables we derived for each watershed. Land use/cover values were reported and analyzed as a percentage of the total watershed area. Patch heterogeneity measured landscape fragmentation and was reported as the number of patches per hectare. We derived slope from elevation data using ARC/INFO. The standard deviation of elevation was used as a surrogate measure of topographic variability.

Statistical Analysis

Using redundancy analysis (RDA), a canonical extension of principal component analysis (PCA), we detected relationships among the individual multivariate data sets. RDA is a form of direct gradient analysis that describes variation in a multivariate data set (e.g., habitat variables or macroinvertebrate metrics) based upon environmental data (21). In RDA, the station scores from a PCA are regressed on a specified set of environmental variables with each iteration, and the fitted values of the regression become new station scores (22). Thus, environmental or predictor variables constrain PCA. Two important outputs from this method are the interset correlations of environmental variables with the RDA axes, which indicate the environmental variables that have the strongest influence in the ordination, and the fraction of

Table 1. Spatial Data Used for Landscape Characterization

Data Layer	Source	Scale	Format Received
Hydrology	U.S. EPA stream reach	1:100,000	ARC/INFO
Elevation	USGS-DEM	1:250,000	Digital elevation model
Land use/cover	USGS	1:100,000	Digital line graph
Land use/cover	MIDNR	1:24,000	Intergraph
Watershed boundary	USGS topographic maps	1:24,000	Manual delineation
Station locality	USGS topographic map	1:24,000	Manual digitizing
Soils	USDA SCS STATSGO	1:250,000	ARC/INFO
Major basin	USGS topographic	1:24,000	Manual digitizing
Quartenary geology	University of Michigan	1:250,000	Manual digitizing

Key: USGS = United States Geological Survey
DEM = Digital elevation model
MIDNR = Michigan Department of Natural Resources
USDA SCS = United States Department of Agriculture Soil Conservation Service

Table 2. Landscape Variables Measured or Derived for Each Watershed

Variable	Units
Land use/cover	Proportional area
Slope	Degrees
Standard deviation elevation	Meters
Patch heterogeneity	Patches/hectare
Soils	Proportional area
Total area	Hectares

total variance of each predicted variable that is explained by the RDA axes (22).

Performing Monte Carlo permutation tests determined the statistical validity of the association between predictor and predicted variables. Tests were conducted by random permutation of the site numbers in the predictor variables. We randomly linked the predictor data to the predicted data. We conducted 99 simulations to approximate a normal distribution with which to compare our data with random combinations.

We first determined which of the reach variables had strong influences on macroinvertebrate distributions by conducting separate RDAs with physical habitat and chemistry variables as environmental descriptors. We then examined the ability of the landscape data to predict the variation in the important reach variables.

To determine the relative influences of LU and GEOS landscape variables on stream chemistry and physical habitat, we used partial RDA, where one landscape variable type was held constant and variation due to the other landscape set was examined independently. Using this approach, total variation in a multivariate data set can be decomposed in a manner analogous to analysis of variance (23, 24). For this analysis, we attributed variation in the reach variables (habitat and chemistry) to four separate components:

- The variation in reach variables that LU variables explained independently of GEOS variables.
- The variation in reach variables that GEOS variables explained independently of LU variables.
- The variation in reach variables that both GEOS and LU variables shared. This shared variation could have been due to both the dependence of one type of variable on the other as well as noncausal relationships (e.g., the types of soil found in a watershed determine in large part the types of agriculture that can be practiced).
- The variation in reach variables that were unexplainable. This may have been attributable to sampling error, stochastic variation, or other variables not sampled.

Results

Regional Characteristics

Land Use

Land use within the study region was dominated by row-crop agriculture (see Table 3). Individual watersheds ranged from 14 to 99 percent in agricultural land uses, with the East basin watersheds exhibiting the greatest proportion of agricultural land use and the Flint having the lowest proportion of agricultural land use. The Chippewa/Pine and Kawkawlin watersheds exhibited the greatest diversity of land use and cover types within the study region.

Wetlands represented a minor land use component with most watersheds having between 0 and 15 percent land area. The Cass and Kawkawlin basins had the greatest proportion of wetlands, with a median of 6.8 percent for individual watersheds.

Macroinvertebrates

Considerable variation existed among the major basins with respect to the 15 macroinvertebrate community metrics during summer (see Table 4). Metric values for the Flint, Shiawassee, and Chippewa watersheds were similar. Sites within the Kawkawlin and East basins differed considerably from the Flint, Shiawassee, and East basins in several of the metrics. The Kawkawlin watershed was notable for high shredder and filterer proportions and a low proportion of detritivores. The East basin also had a high proportion of shredders. Both the East and Kawkawlin basins had lower proportions of strictly erosional taxa and higher proportions of depositional taxa than the other major basins.

Taxa at the East and Kawkawlin basins also exhibited lower oxygen tolerance than at other major basins. In addition, their Hilsenhoff Biotic Index (HBI) scores (which are sensitive to oxygen availability) were higher than other basins, and they had the lowest EPT (Ephemeroptera, Plecoptera, Trichoptera) richness. Total richness at Kawkawlin was relatively high, however. Richness was highest in the Chippewa/Pine watershed and lowest in the East basin.

In general, fall patterns of macroinvertebrate metrics resembled those of summer. The Kawkawlin and East basins had high HBI scores, low EPT scores, low proportions of erosional taxa, and high proportions of depositional taxa. The proportion of predators was exceptionally high in the Kawkawlin basin due to the abundance and trophic classification of one chironomid genus.

Table 3. Summary of Landscape Metrics in Six Major Basins of the Saginaw River Drainage

Landscape Variables	East Basin	Cass	Flint	Shiawassee	Chippewa/Pine	Kawkawlin
n	**8**	**7**	**8**	**5**	**15**	**3**
Row crops	86.4 79.7-98.0	58.4 45.1-73.3	38.0 25.6-65.5	43.5 26.7-71.8	48.2 13.9-91.3	26.4 18.5-74.9
Other agricultural land	0.2 0.0-1.2	0.6 0.4-1.4	3.1 0.6-4.1	0.4 0.1-2.0	4.0 0.3-6.6	1.9 0.2-2.3
Urban	0.8 0.1-2.0	1.9 0.4-3.7	8.56 1.4-23.2	10.3 2.4-15.0	2.1 0.8-4.0	1.2 1.0-6.8
Deciduous forest	6.4 0.5-12.1	16.9 5.8-33.4	17.5 10.2-20.5	19.2 16.4-30.4	23.2 3.2-42.1	56.8 13.2-64.5
Mixed hardwood	0.03 0.0-0.3	0.3 0.01-1.7	2.6 0.0-3.8	0.7 0.4-0.9	5.4 0.1-8.5	0.3 0.1-0.3
Range: Herb	1.0 0.1-3.6	6.7 2.1-12.7	13.0 6.6-14.5	12.6 0.9-17.3	6.9 1.0-9.0	2.5 0.7-3.1
Range: Shrub	1.3 0.0-2.3	4.7 3.2-7.7	6.5 2.6-10.5	7.3 2.2-10.1	5.6 0.7-9.0	3.0 2.7-3.2
Forested wetlands	0.1 0.0-0.5	0.1 0.0-0.3	2.3 0.3-3.2	0.7 0.0-3.7	1.0 0.0-2.5	1.8 0.1-2.2
Non-forested wetlands	1.7 0.0-5.3	6.7 1.9-14.5	3.9 0.4-5.2	3.6 0.7-5.0	4.3 0.1-7.9	5.0 1.3-5.9
Slope (degrees)	0.15 0.07	0.29 0.12	0.41 0.11	0.27 0.10	0.35 0.10	0.14 .03
Elevation (meters)	206.5 19.8	239.8 6.7	277.2 22.5	252.6 47.1	278.7 34.9	203.1 9.8
Patch heterogeneity	241.5 102.0	711.4 250.6	950.4 228.6	762.7 223.6	703.6 185.9	519.6 156.7
Watershed area (hectares)	14,968.9 11,132.0	38,117.7 58,321.9	28,926.1 19,236.7	46,530.0 50,798.0	53,704.2 44,872.8	22,240.2 3,848.0

Land use/cover variables (agricultural land through nonforested wetlands) are reported as median and range; landscape structure variables (slope through watershed area) are reported as mean and standard deviation. Land use/cover represents proportional areas of each watershed.

Identification of Important Reach-Scale Variables

Chemistry

RDA showed that chemical variables explained 26 percent of the variation in macroinvertebrate data in summer and 33 percent in fall. The most important variables in summer were TN and TSS (see Table 5). Fall macroinvertebrate communities were influenced by a greater number of variables, including NH_3, TP, ALK, and TSS.

Physical Habitat

The 13 physical habitat variables explained 37 percent of the macroinvertebrate data in summer and 46 percent of the macroinvertebrate data in fall. In summer, the percentage of deep pools and canopy extent along with channel dimensions, such as bank full width (BFW) and bank full depth (BFD), were the most important variables. In fall, the percentage of fines and deep pools as well as canopy extent were among the most important variables (see Table 6).

Landscape Influences on Surface Water Chemistry

In summer, the landscape data explained 55 percent of the variation in chemical variables. The proportion attributable to LU was larger than that attributable to GEOS (see Figure 2). The two data types shared 12 percent of

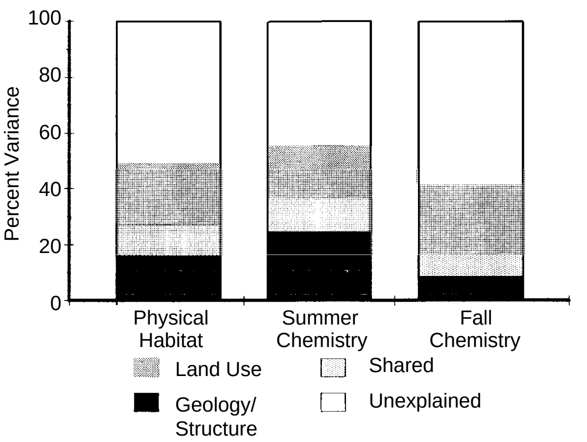


Figure 2. Results of variance decomposition from partial RDA.

Table 4. **Mean and Standard Deviation of Macroinvertebrate Metrics Calculated for Summer Collection Periods for Six Major Basins of the Saginaw River Drainage**

	East Basin	Cass	Flint	Shiawassee	Chippewa/Pine	Kawkawlin
n	**8**	**7**	**8**	**5**	**15**	**3**
Chironomidae	59.1 35.3	57.9 20.3	45.9 27.9	32.2 28.5	45.5 29.6	67.1 18.6
Omnivores	19.4 13.7	19.1 7.1	18.1 13.9	14.4 3.8	21.5 9.8	22.0 16.7
Detritivores	57.1 34.1	69.7 9.7	75.3 16.0	79.9 6.4	70.4 9.8	29.0 26.1
Shredders	30.3 33.9	18.7 5.1	10.6 6.0	7.7 6.8	14.4 15.6	51.0 26.9
Gatherers	59.8 32.9	18.7 5.1	64.3 12.8	65.0 14.1	65.8 17.2	39.8 30.1
Filterers	27.4 38.1	23.4 18.5	22.4 14.1	18.9 12.8	17.6 15.1	39.9 35.7
Grazers	32.2 34.4	13.6 16.2	26.2 21.0	40.1 22.7	25.5 21.1	25.4 22.9
Predators	1.5 2.2	1.2 1.2	1.5 2.0	1.0 1.0	1.4 0.8	1.9 0.8
2 Dominants	64.5 25.5	54.3 6.1	50.3 15.2	54.2 9.5	51.9 11.6	60.0 21.7
Total abundance	2077 4951	650 739	574 622	325 91	497 230	433 297
HBI	7.1 1.4	5.6 ?	5.6 0.8	6.0 0.5	5.1 1.1	8.1 0.8
Erosional taxa	25.9 12.4	36.1 5.5	35.5 9.5	38.9 14.7	36.1 11.0	14.9 5.3
Depositional taxa	35.5 13.2	23.7 9.6	27.5 11.5	27.0 6.6	25.4 10.7	52.4 6.5
Species richness	17.2 4.5	18.3 9.6	22.1 8.2	20.6 4.7	26.6 3.0	23.3 4.9
EPT taxa richness	5.0 2.7	5.7 2.8	7.3 3.0	8.0 2.3	10.0 3.7	3.3 0.5

the variation. The relationship between LU variables and chemistry was significant ($p < 0.05$), and the relationship between GEOS variables and chemistry was not significant ($p > 0.05$).

In fall, variation explained by LU was proportionally less than during summer (see Figure 2). GEOS landscape variables explained 25 percent of the total variation while LU variables accounted for less than 10 percent of the total variation. LU and GEOS variables shared approximately 8 percent of the variation. In contrast with summer, GEOS variables were significant and LU variables were not significant when examined with the Monte Carlo test ($p < 0.05$).

The importance of GEOS and LU variables in explaining variation in the chemistry variables, as well as the total amount of variation explained, differed considerably among the chemistry variables. Figure 3 shows only summer data. For example, the landscape variables explained almost 80 percent of the total variation in TN (see Figure 3). The largest proportion of variance was explained by the shared influences of GEOS and LU. Alkalinity, which was also well predicted, however, was much more influenced by variation attributable to LU variables. In comparison, LU variables explained less than 45 percent of the total variance in TP, and the majority of this variance was attributable to GEOS.

To further examine the influence of specific landscape variables, we compared the various axes in the significant partial ordinations (see Figure 4). In summer, when we observed a significant effect of LU variables on water chemistry, forested land covers and nonrow-crop agriculture had their greatest influence on TSS and ALK. LU heterogeneity and shrub vegetative cover most strongly influenced both TN and NH_3. In fall, when GEOS variables had a significant relationship to water chemistry, ALK and NH_3 were more influenced by peat land soils and watershed size. The proportion of sand and gravel soils, as well as clays, explained much of the variation in TN and TSS.

Table 5. Chemistry Variables That Had a Correlation (r) of at Least 0.30 With One of the RDA Axes With the Summer or Fall Ordinations; a Monte Carlo Analysis Indicated That Both Summer and Fall Ordinations Were Significant ($p < 0.05$)

	RDA 1		RDA 2		RDA 3	
Variable	**Summer**	**Fall**	**Summer**	**Fall**	**Summer**	**Fall**
TN	0.44	0.03	-0.04	-0.07	0.05	-0.14
NH_3	0.05	0.04	0.09	0.52	0.07	-0.09
TP	0.12	0.40	0.07	0.26	0.25	0.14
PO_4	-0.04	0.19	0.01	-0.08	0.26	0.34
TSS	0.01	-0.31	0.43	0.36	0.22	-0.13
ALK	-0.08	-0.36	-0.16	0.09	0.15	0.2

Table 6. Physical Habitat Variables That Had Correlations (r) Over 0.3 With the Ordination Axes; Results of the Monte Carlo Simulation Indicated That the Fall but Not the Summer Ordinations Were Significant ($p < 0.05$)

	RDA 1		RDA 2		RDA 3	
Variable	**Summer**	**Fall**	**Summer**	**Fall**	**Summer**	**Fall**
Percentage of fines	0.27	0.47	0.04	0.28	0.18	-0.09
Percentage of shallows	-0.06	0.17	-0.28	-0.11	0.32	-0.38
Wood	0.09	0.34	-0.1	0.02	0.10	-0.17
Percentage of deep pools	0.50	0.54	0.16	0.09	-0.17	0.26
Erosion	0.16	0.30	-0.3	-0.31	0.01	0.27
Maximum depth	0.08	-0.14	0.08	0.16	-0.39	0.13
Canopy extent	0.75	0.21	-0.47	-0.42	0.21	-0.25
BFW	-0.10	-0.23	0.52	0.36	0.07	0.03
BFD	-0.09	-0.13	0.32	0.02	-0.08	0.22
Flood ratio	0.03	-0.30	0.07	0.23	-0.41	0.02

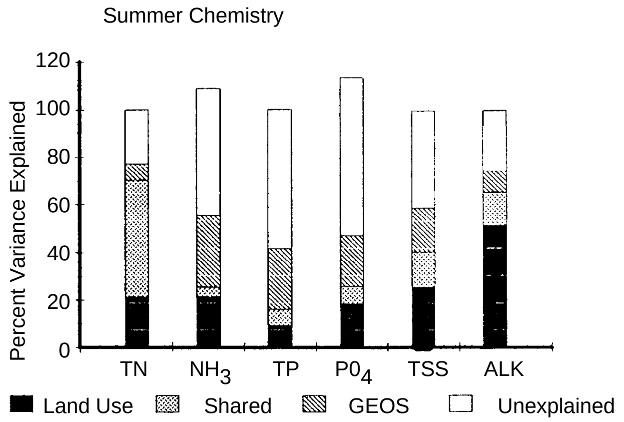
Landscape Influences on Physical Habitat

GEOS landscape variables attributed for the largest portion (22 percent) of the explained variation in physical habitat variables (see Figure 2). LU variables accounted for 16 percent of the explained variance. The partial ordination for GEOS but not LU was significant as the Monte Carlo procedure determined.

As noted with chemistry variables, there were considerable differences in the ability of the landscape variables to predict individual habitat characteristics. Landscape variables were best at predicting BFW and least powerful for predicting the percentage of deep pools (see Figure 5). BFW was influenced predominantly by GEOS and only minimally by LU. The most influential GEOS variables for BFW related to watershed area (see Figure 6). Woody debris was predominantly influenced by LU variables. The most influential LU variables for woody debris related to forested wetlands. Flood ratio was intermediate to these examples. Both sets of landscape variables shared the largest proportion of explained variance for this parameter.

Discussion

Our studies demonstrate the distinct influences landscape features have on stream macroinvertebrate communities through modifying surface water chemistry and stream habitat. Land use most strongly influences stream chemistry during summer months when surface runoff and soil leaching are greatest. In addition, fertilizer application in row-crop agriculture is highest in the first part of the growing season. The strong relationship between some aspects of land use and stream water chemistry were similar to those observed in other studies (4, 6, 7). The specific mechanism by which stream chemistry influences macroinvertebrates is not clear. The addition of nutrients can significantly affect stream productivity (25-27); however, light often limits primary production in agricultural areas (28-30). Nutrients may



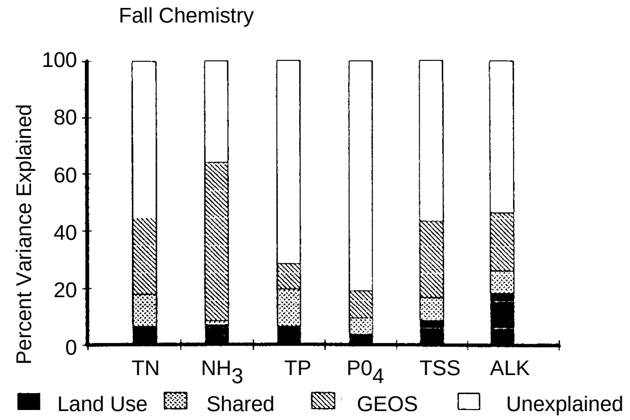


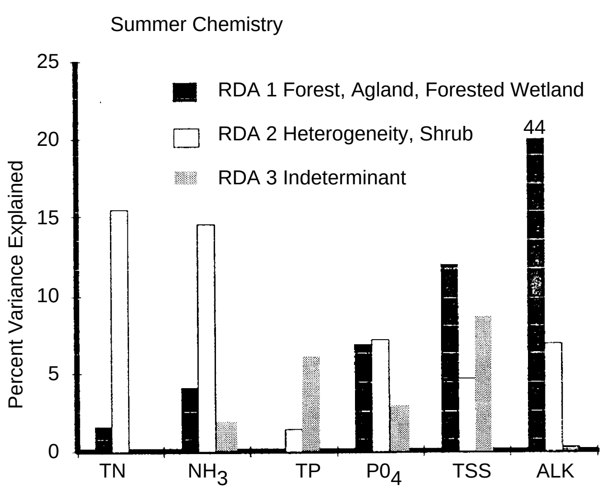
Figure 3. Results of variance decomposition for chemical parameters from partial RDA.

also act as indicators of other agricultural chemicals (e.g., herbicides, pesticides) and carbon sources, which we did not measure in our study, but their inputs are moderated by agricultural land use.

The influence of landscape features on stream habitat may be even more complex than stream chemistry can show. Low-gradient midwestern streams in the United States are under a complex set of controls related not only to subtle topographic gradients but also to interactions among land use, soil type, riparian vegetation, and historic anthropogenic modifications. Ditching, channelization, and hydrologic changes associated with large-scale land-cover changes, such as logging and wetland drainage, have a significant effect on current stream physical conditions.

Our analysis did not quantify many of these effects. Nonetheless, some land use influences seem clear. First, we found that the extent of permanent vegetation in a watershed played a central role in land use impacts on stream habitats. All of the key factors that described variation in the habitat variables included land covers with permanent vegetation. In addition, our initial land-cover analysis found that row-crop agriculture highly negatively correlated with almost all permanently vegetated land covers. Second, although wetlands represented a relatively small proportion of total watershed area (less than 5 percent), they had a relatively high influence on habitat features. Both forested and non-forested wetlands were important to habitat features such as woody debris and some aspects of channel dimensions.

We found that one of the most important landscape influences on stream habitat was watershed area, which explained much of the variation in channel dimensions. These habitat parameters, in turn, were among the most important for explaining variation among macroinvertebrate communities. Stream size, including channel dimension, is an important determinant of several macroinvertebrate community characteristics that stem from longitudinal phenomena associated with stream ecosystems (31, 32). Because these relationships are fixed, they are not amenable to manipulation by management. Consequently, if the monitoring objective is to



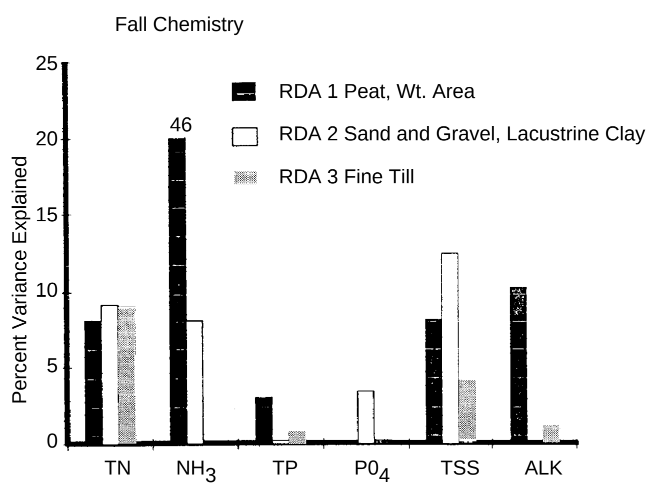


Figure 4. Factors influencing chemical parameters in summer (land use) and fall (geology/structure).

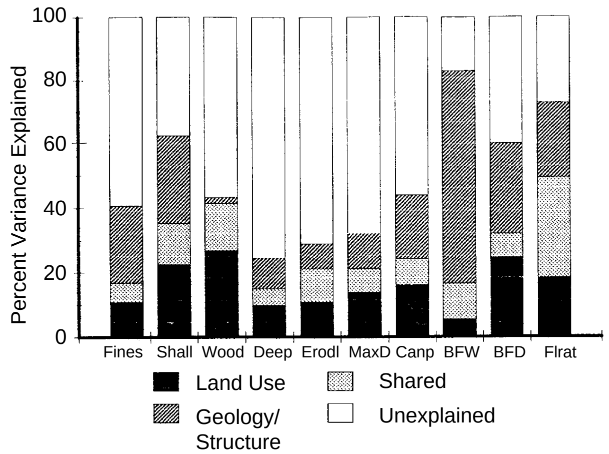


Figure 5. Results of variance decomposition from partial RDA for physical habitat.

detect changes resulting from land use, then specific macroinvertebrate community characteristics that respond to changes in channel dimensions are not good biological indicators.

Our study indicates that land use and geology/structure have similar magnitudes of influence on reach environments and biotic communities. This suggests that although some community and water quality characteristics are controlled by landscape features that are difficult to modify or manage, other aspects may be altered through specific activities designed to increase or decrease land uses for a desired effect.

The methods described in this study show promise for examining landscape-stream interactions. Much of the variation we found in both stream reach environments and macroinvertebrate communities, however, remained unexplained. Undoubtedly, we can attribute much of this unexplained variation to the abbreviated sampling procedures that large-scale field studies such as this dictate. Questions of scale and position, however, may also be responsible for much unexplained variation.

The minimum mapping unit used in our study (1 hectare) may be insufficient to detect local land use practices that influence streams. For example, riparian zones along many streams may only be a few meters in width and yet exert considerable control over nutrient inputs (33). Higher spatial resolution may improve predictive power. In addition, the position of certain landscape elements may determine consequent effects on stream conditions. Johnston et al. (7) found that wetland position within a watershed influenced the ability of the wetland to modify stream chemistry. With respect to biologic communities, the abundance of some taxa may relate to the presence of upstream refugia, which can provide sources of colonists (34) and other biogeographic effects.

The further development of empirical relationships between landscape parameters and streams will provide important insights into understanding the spatial and temporal scales needed for modeling and monitoring large watersheds. Government and commercial geographic databases that can be easily adapted to GIS technology should become more readily available. As they do, the ability to identify relationships across scales, and consequently predict stream community composition from watershed-scale attributes, will facilitate the assessment of ecological risk associated with land use changes.

Acknowledgments

Jim Wessman, Frank Kutka, and Tim Aunan developed much of the data reported in this study. John Arthur and Tom Roush collected most of the chemistry and macroinvertebrate data and provided invaluable logistical support. Cajo ter Braak provided helpful insights to statistical analysis. This research was funded by the U.S. Environmental Protection Agency through CR-818373-01-0.

References

1. Minshall, G.W. 1988. Stream ecosystem theory: A global perspective. J. North Am. Benthological Soc. 7:263-288.
2. Naiman, R.J. 1992. New perspectives for watershed management: Balancing long-term sustainability with cumulative environmental change. In: Naiman, R.J., ed. Watershed management. New York, NY: Springer-Verlag. pp. 3-11.
3. Likens, G.E., F.H. Bormann, N.M. Johnson, D.W. Fisher, and R.S. Pierce. 1970. Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook Watershed ecosystem. Ecol. Monogr. 40:23-47.
4. Correll, D.L., E.D. Friebele, and J. Miklas. 1977. Nutrients in land runoff from Rhode River watersheds in 1975 and 1976. In: Nonpoint source studies on Chesapeake Bay, Vol. II. Publication No. 55. Chesapeake Bay Center for Environmental Studies/Smithsonian Institution, Edgewater, MD.
5. U.S. EPA. 1976. The influence of land use on stream nutrient levels. EPA/600/3-76/014. Washington, DC.
6. Osborne, L.L., and M.J. Wiley. 1988. Empirical relationships between land use/cover patterns and stream water quality in an agricultural watershed. J. Environ. Mgmt. 26:9-27.
7. Johnston, C.A., N.E. Detenbeck, and G.J. Niemi. 1990. The cumulative effect of wetlands on stream water quality and quantity: A landscape approach. Biogeochemistry 10:105-141.
8. Detenbeck, N.E., C.A. Johnston, and G.J. Niemi. 1993. Wetland effects on lake water quality in the Minneapolis/St. Paul metropolitan area. Landscape Ecol. 39-61.
9. Duda, A.M., D.R. Lenat, and D.L. Penrose. 1982. Water quality in urban streams: What can we expect? JWPCF 54:1,139-1,147.

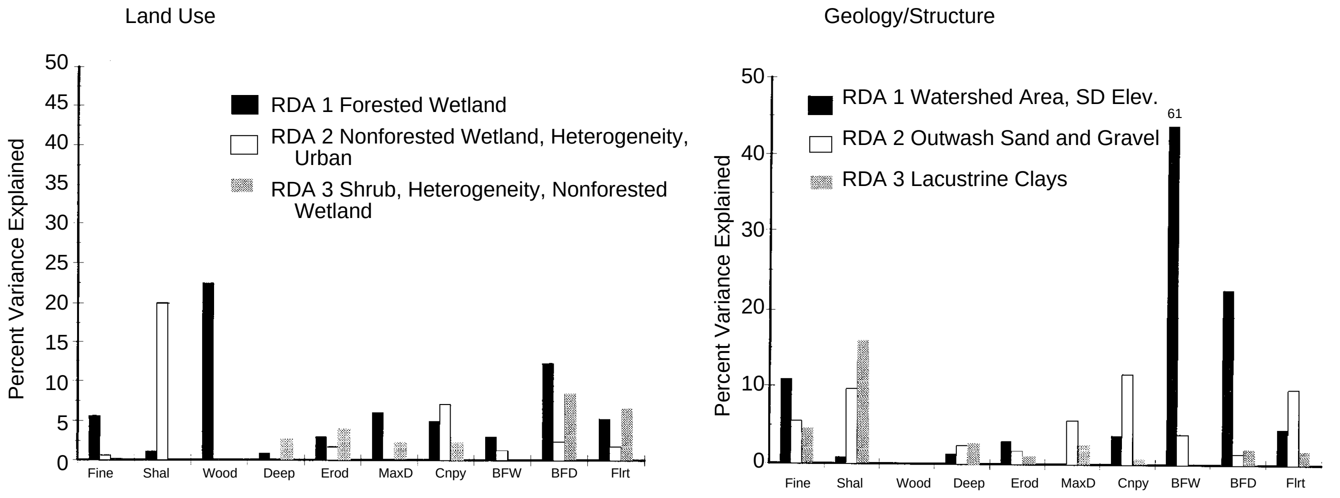


Figure 6. Factors influencing physical habitat parameters.

10. Jones, R.C., and C.C. Clark. 1987. Impact of watershed urbanization on stream insect communities. Water Resources Bull. 23:1,047-1,055.

11. Bisson, P.A., R.E. Bilby, M.D. Bryant, C.A. Dolloff, G.B. Grette, R.A. House, M.L. Murphy, K.V. Koski, and J.R. Sedell. 1987. Large woody debris in forested streams in the Pacific Northwest: Past, present, and future. In: Salo, E.O., and T.W. Cundy, eds. Streamside management: Forestry and fishery interactions. Contribution 57. Institute of Forest Resources, University of Washington, Seattle, WA. pp. 143-190.

12. Everest, F.H., R.L. Beschta, J.C. Scrivener, K.V. Koski, J.R. Sedell, and C.J. Cedeholm. 1987. Fine sediment and salmonid production: A paradox. In: Salo, E.O., and T.W. Cundy, eds. Streamside management: Forestry and fishery interactions. Contribution No. 57. Institute of Forest Resources, University of Washington, Seattle, WA. pp. 98-142.

13. Richards, C., and G.E. Host. 1994. Using GIS to examine the influence of land use on stream habitats and biota. Water Resources Bull. 30:729-738.

14. Rosenberg, D.M., and V.R. Resh. 1993. Freshwater biomonitoring and benthic macroinvertebrates. New York, NY: Chapman and Hall.

15. Richards, C., and G.W. Minshall. 1992. Spatial and temporal trends in stream macroinvertebrate species assemblages: The influence of watershed disturbance. Hydrobiologia 241:173-184.

16. Ohio EPA. 1987. Biological criteria for the protection of aquatic life, Vol. II. User's manual for biological assessment of Ohio surface waters. Columbus, OH.

17. Barbour, M.T., J.L. Plafkin, B.P. Bradley, C.G. Graves, and R.W. Wisseman. 1992. Evaluation of EPA's rapid bioassessment benthic metrics: Metric redundancy and variability among reference stream sites. Environ. Toxicol. Chem. 111:437-449.

18. Karr, J.R., and B.L. Kearns. 1992. Components of biological integrity: Their definition and use in development of an invertebrate IBI. In: Simon, T.P., and W.S. Davis, eds. Proceedings of the 1991 Midwest Pollution Control Biologists Meeting: Environmental indicators: Measurement and assessment endpoints. EPA/905/R-92/003. U.S. EPA Region 5, Environmental Sciences Division, Chicago, IL. pp. 1-16.

19. APHA. 1989. Standard methods for the examination of water and waste water. American Public Health Association.

20. Anderson, J.R., E.E. Hardy, J.T. Roach, and R.E. Witmer. 1976. A land use and land cover classification system for use with remote sensor data. Geological Survey Professional Paper 964. U.S. Government Printing Office, Washington, DC.

21. ter Braak, C.J., and I.C. Prentice. 1988. A theory of gradient analysis. Advances in Ecol. Res. 8:271-317.

22. Jongmann, R.H., C.J. ter Braak, and O.F. van Tongeren. 1987. Data analysis in community and landscape ecology. Wageningen, The Netherlands: Pudoc.

23. Borchard, D., P.L. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. Ecology 73:1,045-1,055.

24. ter Braak, C.J. 1990. Update notes: CANOCO Version 3.0. Agricultural Mathematics Group, Wageningen, The Netherlands.

25. Petersen B.J., J.E. Hobbie, T.J. Corliss, and K. Kriet. 1983. A continuous-flow periphyton bioassay: Tests of nutrient limitation in a tundra stream. Limnology and Oceanography 28:583-591.

26. Grimm, N.B., and S.G. Fisher. 1989. Stability of periphyton and macroinvertebrates to disturbance by flash flooding. Ecol. Monogr. 52:93-110.

27. Pringle, C.M. 1987. Effects of water and substratum nutrient supplies on lotic periphyton growth: An integrated bioassay. Can. J. Fish. Aquat. 44:619-629.

28. Munn, M.D., L.L. Osborne, and M.J. Wiley. 1989. Factors influencing periphyton growth in agricultural streams of central Illinois. Hydrobiologia 174:89-97.

29. Bushong, S.J., and R.W. Bachmann. 1989. In situ enrichment experiments with periphyton in agricultural streams. Hydrobiologia 178:1-10.

30. Wiley, M.J., L.L. Osborne, and R.W. Larimore. 1990. Longitudinal structure of an agricultural prairie river system and its relationship to current stream ecosystem theory. Can. J. Fish. Aquat. 47:373-384.

31. Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing. 1980. The river continuum concept. Can. J. Fish. Aquat. 37:130-137.

32. Statzner, B., J.A. Gore, and V.H. Resh. 1988. Hydraulic stream ecology: Observed patterns and potential applications. J. North Am. Benthological Soc. 7:307-360.

33. Osborne, L.L., and D.A. Kovacic. 1993. Riparian vegetated buffer strips in water-quality restoration and stream management. Freshwater Biol. In press.

34. Sedell, J.R., G.H. Reeves, F.R. Hauer, J.A. Standford, and C.P. Hawkins. 1990. Role of refugia in recovery from disturbances: Modern fragmented and disconnected river systems. Environ. Mgmt. 14:711-724.

A Watershed Approach to Source Water Assessment and Protection Utilizing GIS-Based Inventories: A Case Study in South Carolina

James M. Rine, and Elzbieta R. Covington
Earth Sciences and Resources Institute
University of South Carolina
Columbia, SC 29208

ABSTRACT

The Earth Sciences and Resources Institute at the University of South Carolina (ESRI-USC) developed a watershed -based methodology for conducting contaminant inventories and susceptibility analyses required under the U.S. Environmental Protection Agency (US EPA) *State Source Water Assessment and Protection Program Final Guidance* (*SSWAPP* Final Guidance, August 1997). *SSWAPP* is an offspring of the 1996 Safe Drinking Water Act. The watershed-based methodology was tested within a medium-size public water system which utilizes surface water. The project, initiated at the request of the South Carolina Department of Health and Environmental Control (SCDHEC), helped prepare South Carolina's initial response to the *SSWAPP* Final Guidance (US EPA, 1997). In the initial response, US EPA required all states to submit a source water assessment plan for their state. The three components of that assessment plan are: 1.) Delineation of the source water area(s); 2.) Inventory of contaminants within a source area; and 3.) Determination of source water susceptibilities. SCDHEC tasked ESRI-USC to perform the assessment component #2, An Inventory of contaminants. Following the termination of the SCDHEC project, ESRI-USC also completed the other components of the assessment. The ESRI-USC version of the *SSWAPP* assessment plan, which differs from the plan adopted by SCDHEC, is presented in this paper.

In the plan proposed by ESRI-USC for South Carolina (or any State), the inventory assessment methods follow a watershed-approach which is comprised of three key elements. First, a master database is maintained for each public water supply (PWS) source area by the respective PWS operator/owner or by the State environmental agency (SCDHEC). Each PWS master database lists all inventoried sites within the entire source-area watershed(s) upstream of the PWS intake structure and lists all databases where pertinent environmental data are stored for those inventoried sites. Second, each watershed database is linked to original databases maintained and updated by their originating agency (e.g. SCDHEC, US EPA, U.S. Department of Agriculture, other SC agencies, etc.). This second element will allow routine updating of

inventories without requiring additional programs. Assuming the inventories are updated routinely, watershed susceptibility analyses can be revised with current and timely information. Third, the proposed plan has each PWS database linked to a statewide master database under the supervision of the State environmental agency.

The first step in field testing of the ESRI-USC plan was the development of a contaminant inventory for the pilot study area, noting all the known and potential release sites of contaminants listed in US EPA’s National Primary Drinking Water Standards. Initial work concentrated on synthesizing numerous Federal and State electronic databases of known or reported contaminant releases. These databases proved to be the most complete source of information regarding the general type of known releases or discharges. However, inventory development of potential contamination sites utilizing existing databases was difficult due to the lack of accurate location data, unpopulated databases, and incomplete listings of the types and amounts of contaminants present at each permitted site. Consequently, field surveys were needed to complete the contaminant inventories. The field test demonstrated the need for consistency and coordination in database development and management, both within and among Federal and State agencies.

Regarding the determination of source water susceptibilities, ESRI-USC proposes that inventories of the entire topographically-bound watershed are necessary to develop meaningful relative susceptibilities. With this approach, the contaminant inventory can be joined with other digital data sets of the watershed, such as hypsography, land use, soils, and infrastructure to more accurately determine the relative susceptibilities of specific contaminants at specific sites anywhere in the watershed.

INTRODUCTION

According to the US Environmental Protection Agency's *State Source Water Assessment and Protection Program Final Guidance* (US EPA’s *SSWAPP* Final Guidance, August 1997, page A-25) susceptibility analysis should be conducted with, “a clear understanding where the significant potential sources of contamination are located.” The Earth Sciences and Resources Institute at the University of South Carolina (ESRI-USC) interprets this US EPA guidelines to mean that the source water assessment process should fully integrate susceptibility analysis with the inventory of present and future sources of contaminants, wherever they may be located within the source area. To have a "clear understanding" of the "significant potential sources of contaminants" an entire watershed should be inventoried. Or, at the minimum, significant sources should not be ignored just because their locations lie outside of some arbitrary boundary.

In 1998, ESRI-USC conducted a comprehensive contaminant inventory of the upper portion of the watershed for Shaw Creek (HUC3050204) in the Coastal Plain of South Carolina (Fig. 1).

Shaw Creek is the source for the run-of-river withdrawal for the City of Aiken public water supply (PWS) which has over 14,300 connections to homes and businesses. The pilot study portion of the Shaw Creek watershed occupies 181 km^2 (70.5 sq. miles). South Carolina Department of Health and Environmental Control (SCDHEC) initiated this project to help develop a plan for implementing South Carolina's response to US EPA's *State Source Water Assessment and Protection Program Final Guidance* (*SSWAPP*; August 1997). A report on this pilot study can be viewed via the Internet by visiting http://gisweb.esri.sc.edu and selecting *Projects Contaminant Inventory for Source Water Assessment and Protection Project (SWAP)*. *REPORT: Development of a Contamination Inventory Methodology for the Source Water Assessment and Protection Program in South Carolina.* In addition, the contaminant inventory can be queried via the internet at the same location: *Projects: Prototype Model of a Interactive GIS Database for SWAP Contaminant Inventory.*

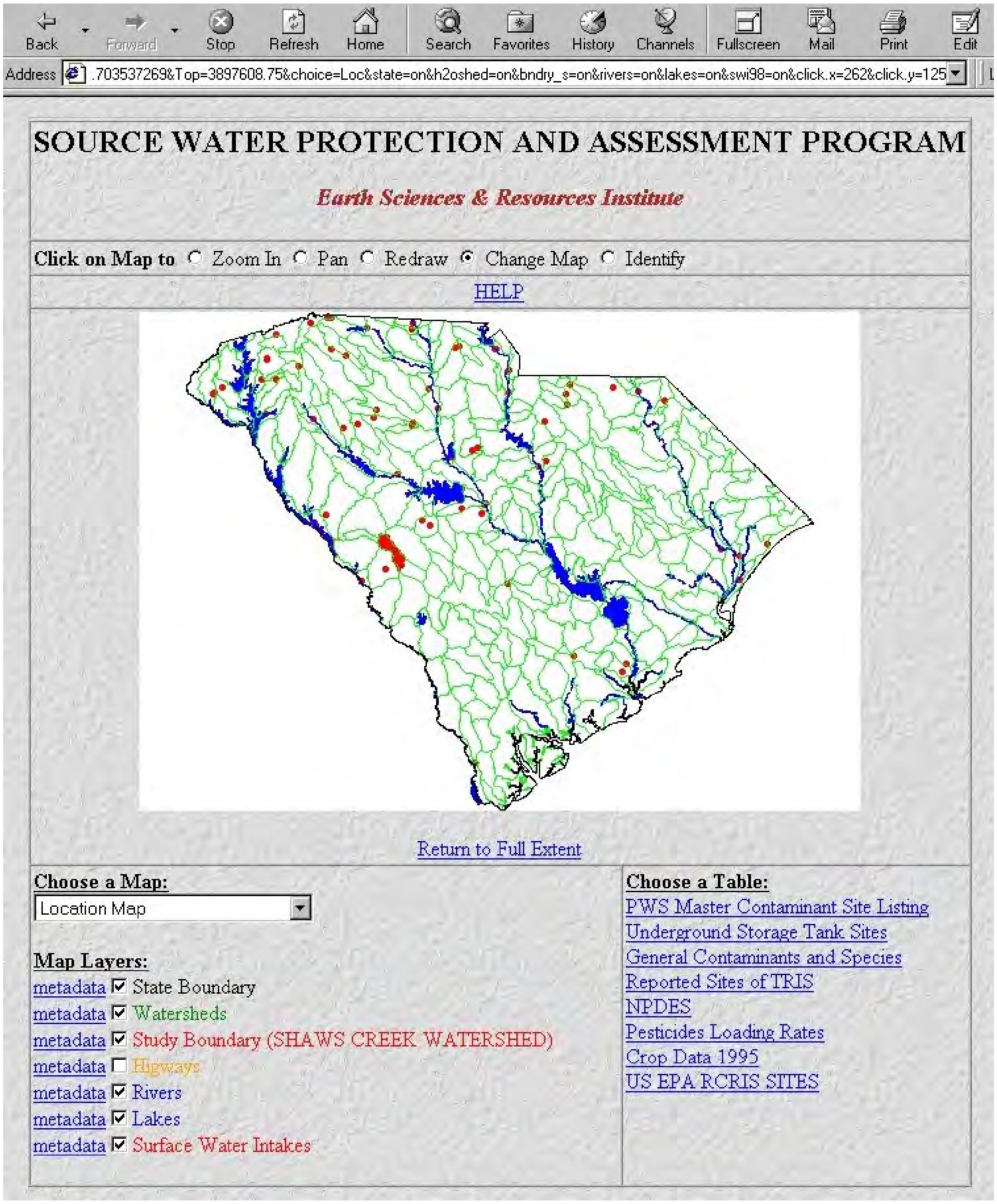
Via a narrative of the pilot study results, this paper will describe how the assessment of the entire watershed (topographically delineated boundaries) is a superior approach to restricting the assessment to some arbitrary segment of the watershed. Secondly, a template for organizing State *SSWAPP* databases is suggested. Third, it is suggested that watershed wide assessments are economically feasible, and in the long term more economical than the piece-meal assessments that require additional surveys and added funding.

DESCRIPTION OF PILOT STUDY

Development of a Contaminant Inventory

The objective of the pilot study was to develop a contaminant inventory for the entire watershed area upstream of the City of Aiken Public Water Supply (PWS) intake. In this inventory are the locations and the characteristics of the sites of all known or reported contaminant releases and the sites of all potential contaminant releases (i.e., sites where hazardous materials are used but no release has been reported to date). The substances inventoried were restricted to only

Figure 1



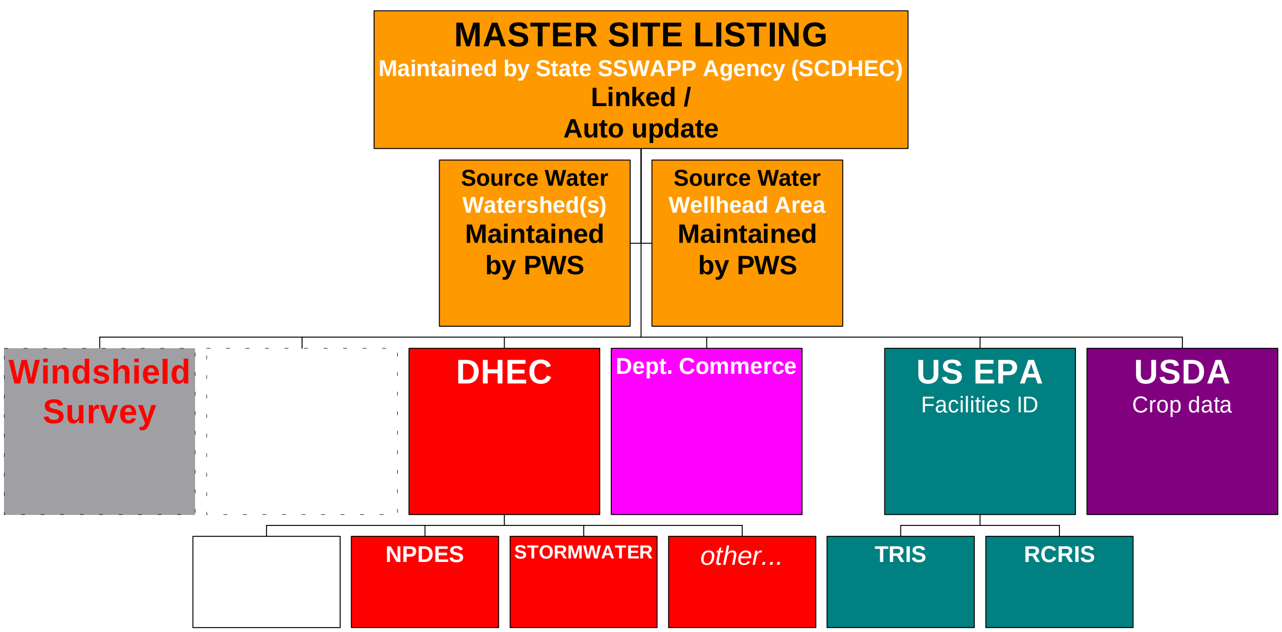
Location map of pilot study (colored in watershed) within the upper coastal plain in west central South Carolina. This upper portion of Shaw Creek is the source area for a run-of-river withdrawal for the City of Aiken Municipal Water Supply. The goal of this pilot study was to develop an assessment methodology for use by all PWS utilizing surface water (PWS intakes shown as dots). Location Map is screen save from interactive, web based contaminant inventory (http://gisweb:esri.sc.edu).

those listed in US EPA's National Primary Drinking Water Standards (NPDWS) with groupings under the categories of: petroleum; chlorinated solvents and other volatile organic compounds (VOC's); pesticides / herbicides; inorganics, mobile inorganics; and pathogens. (For the current National Primary Drinking Water Standards visit www.eps.gov/OGDW/wot/appa.html.)

Survey of Existing Databases for Reported Contaminant Releases

Work concentrated first on synthesizing and cross-referencing numerous Federal and State electronic database-listings of reported contaminant releases. Cross-referencing between USEPA databases is relatively straight forward because each site is given a unique Facility ID number. Within most of the South Carolina (SC) databases examined, however, the absence an unique ID for each site makes it impossible to easily cross-reference between databases. In order to accomplish the objectives of this project, ESRI-USC compiled a prototype master database for the entire pilot study watershed (see Table 1: Master Watershed Site List for Source Area to City of Aiken PWS). This master database is designed to link each inventoried site to every database on which the site is listed (Fig. 2).

Figure 2



Organizational chart showing database structure as proposed for South Carolina by ESRI-USC. MASTER SITE LISTING would include entire state. PWS people would be responsible for updating and correcting data from within their respective watershed source area. The State agency responsible for SSWAPP (SCDHEC) would be responsible for overseeing the entire State MASTER SITE LISTING as well as maintaining liaison with upstream States. Links to routinely updated databases (e.g., US EPA TRIS or the states underground storage tank database [UST]) will allow routine updating of the source water assessments.

Federal and State electronic database-listings of reported contaminant releases generally contain complete information on the general type and location of known releases or permitted discharges. Federal databases, such as the Toxic Release Inventory (TRIS; http://www.epa.gov/enviro/html/tris/state/south_carolina.html) indicate the reporting period of the release and list types and amounts of contaminants released. Within the computerized statewide databases in SC, however, data on amounts of contaminants released are generally incomplete.

Figure 3 shows the distribution of reported contaminant release sites within the study area.

Survey of Existing Databases for Potential Contaminant Releases from Commercial and Public Facilities

Regarding potential contaminant release sites, initial efforts centered on researching the following electronic databases:

1. SCDHEC databases listing owners of various use permits, such as, underground or above ground storage tanks;

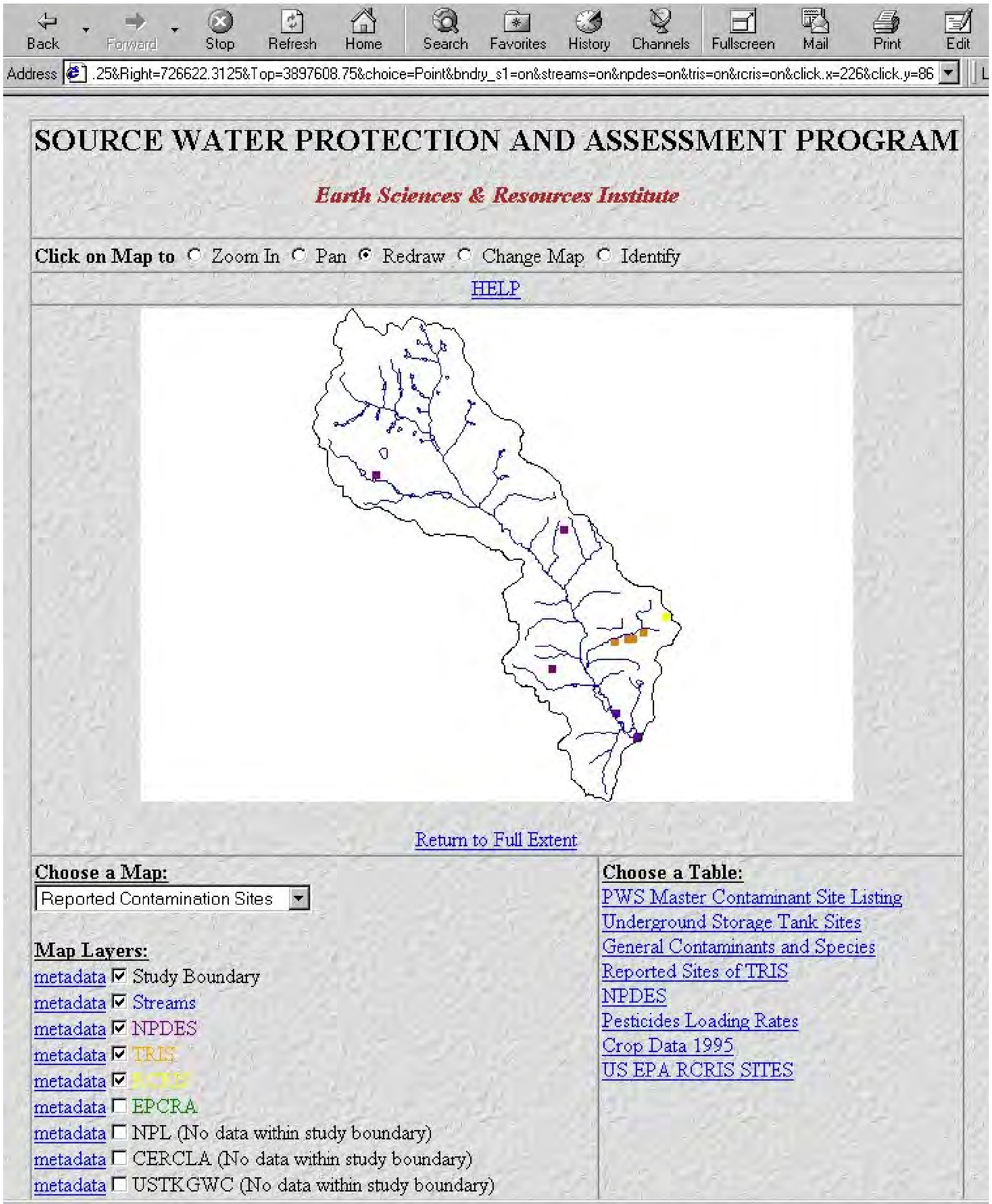
2. The *Emergency Planning and Community Right to Know Act* database (EPCRA/SARA Title III; www.state.sc.us/dhec/eqchome.html) which lists facilities within SC that contain 10,000 lbs. or more of hazardous raw materials;

3. The industrial facilities and land use databases from the South Carolina Department of Commerce.

For those sites that have no listing of potential contaminants present, an interpreted listing was made based on the commercial activities attributed to Standard Industrial Code (SIC) designated to the site (Table 2). For these descriptions refer to OSHA SIC website at www.osha.gov/cgi-bin .

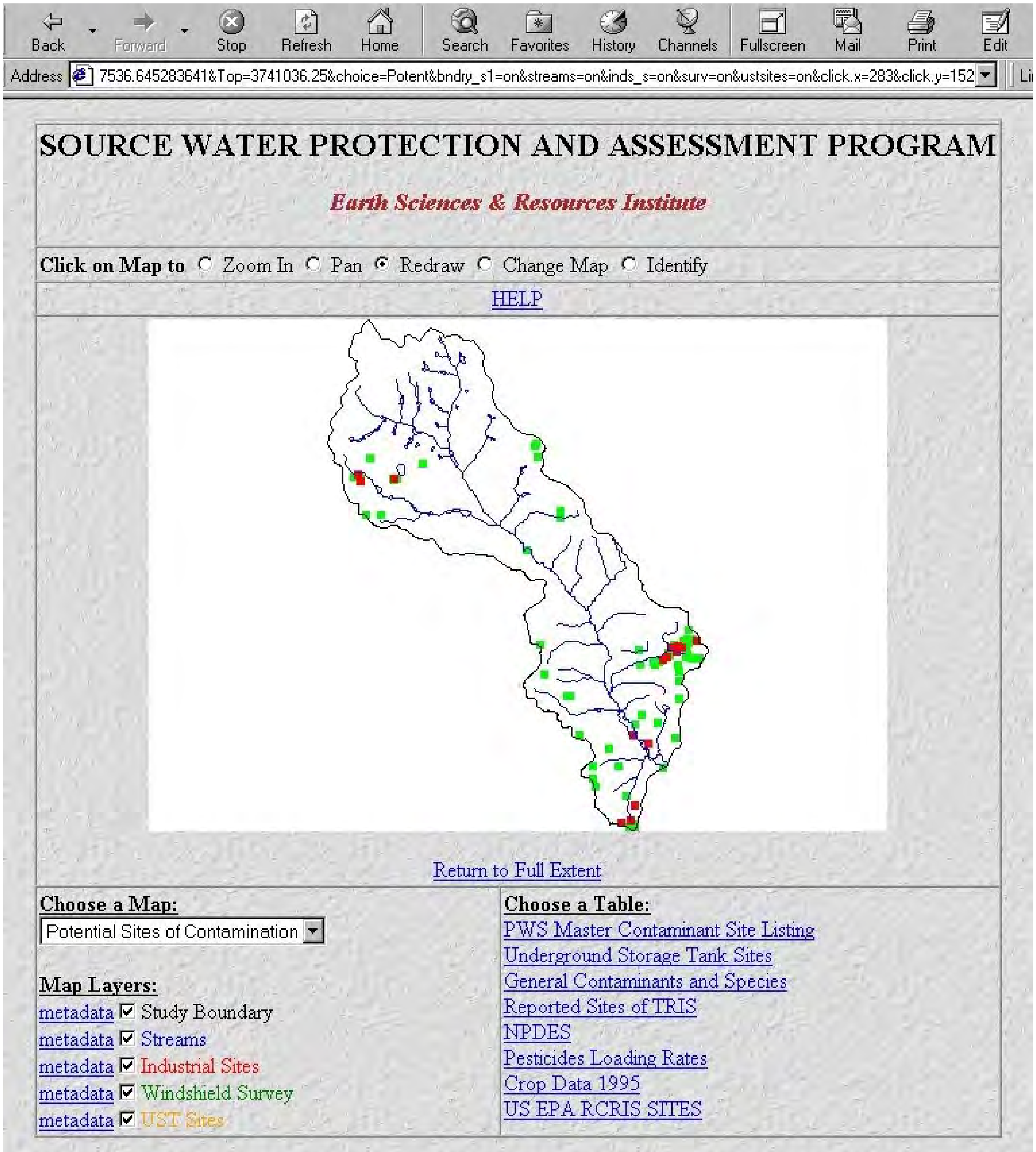
Figure 4 shows the distribution of potential contaminant release sites within the study area.

Figure 3



Sites of reported contaminant releases are displayed on this map. Reported sites are derived from SCDHEC's National Pollutant Discharge Elimination System database (NPDES) and US EPA's Toxic Release Inventory sites (TRIS) and Resource Conservation and Recovery Information System database (RCRIS) which contains comprehensive information on hazardous material handlers in the U.S. and its territories. Map is screen save from interactive, web based contaminant inventory (http://gisweb:esri.sc.edu).

Figure 4



Screen view is of potential sites of contamination derived from databases within SCDHEC. Because of a lack of correct location data, a field survey (windshield survey) was necessary to complete an accurate contaminant inventory of the study area. Out of a total of 62 potential sites identified and mapped in the final inventory of the pilot study area, only 14 of those sites were identified and located by computer searches of pre-existing data bases. No underground storage tanks sites (UST) could be located in the study area via computer searches prior to the windshield survey. After the survey, 15 sites were correlated with SCDHEC USTSITES database (Bureau of Underground Storage Tanks sites). In summary, a windshield survey was needed identify and locate over 50% of the inventoried sites found in the pilot study area.

Field (Windshield) Survey

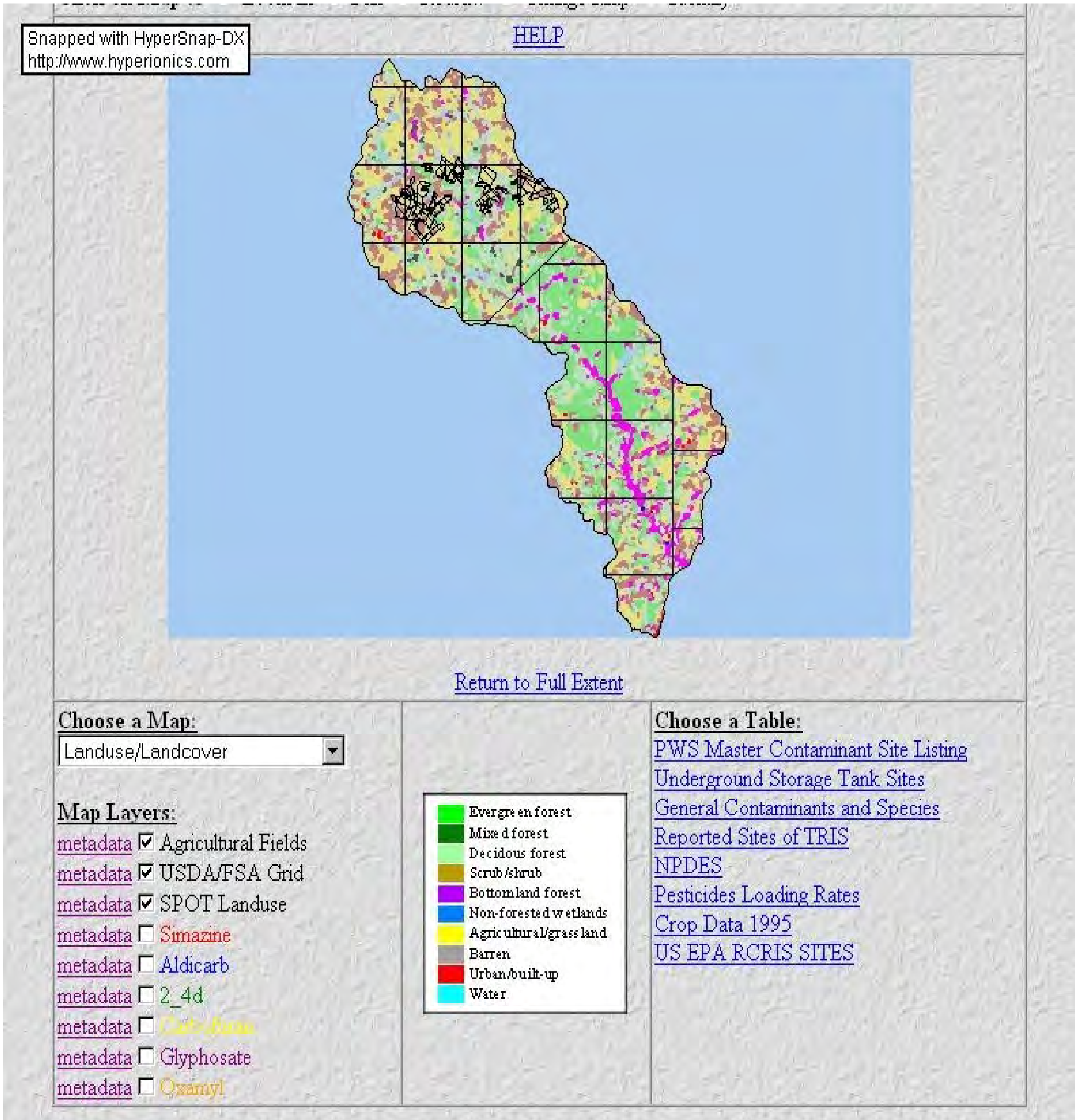
A field survey (i.e., windshield survey) was critical to completing an accurate contaminant inventory of the study area. Out of a total of 62 potential sites identified and mapped in the final inventory of the pilot study area, only 14 of those sites were identified and located by computer searches of pre-existing data bases. No underground storage tank sites (UST) could be located in the study area via computer searches prior to the windshield survey. After the survey, 15 sites were correlated with SCDHEC *USTSITES* database (Bureau of Underground Storage Tanks sites). In summary, a windshield survey was needed to identify and locate over 50 % of the inventoried sites found in the upper Shaw Creek watershed.

Survey of Potential Agriculturally-derived Nonpoint Source Pollutants

"Agriculture accounts for 75% of total pesticide use in the United States and is the primary source of pesticides to surface waters in most areas" (Larson and others, 1997). In South Carolina and in many other states, however, there are no established procedures for determining the extent of the potential problem before it reaches the municipal water intake.

Data available on agricultural use in the study area (and state-wide) are derived from two primary sources, SPOT Land-use / Land-cover data from the SC Department of Natural Resources and annual crop use data from the Farm Services Agency of USDA (FSA). The Land use / Land cover data are based on 1989-1990 multi spectral SPOT images. According to the Land-use/Land-cover data agricultural/grassland use occupies 32% of the upper Shaw Creek study area (Fig. 5). More specific agricultural data (i.e., land use according to individual crops) are available from the FSA with complete crop listings, according to specific fields, up to 1995. As of 1996, revisions to the Farm Act require farmers to report only a few selected crops. FSA crop data originate from individual farm holders and consist of reported use for each cultivated/farmed plot, crop type planted or maintained, and the number of acres used. Each field has an assigned ID number referenced to aerial photographic survey grids. The current FSA record of the individual fields consist of hand-drawn field boundaries on hard copies of aerial photographs. Presently USDA is conducting a nation wide program to convert this hardcopy mode of recording farm locations into a digital format. Visit http://www.nhq.nrcs.usda.gov/ITD/gis.html for details on this USDA program.

Figure 5



Land use / Land cover plot of the pilot study area. Coverage is ERDAS-generated GIS files utilizing 60 meter SPOT data from the Land Resources group within the SC Department of Natural Resources (DNR). The ten categories of land use / land cover are shown in key at bottom of image. Grid pattern is of aerial photographic survey used by FSA to locate farm track and fields, the boundaries of which are hand drawn on the photos and given has an unique identifying code that is referenced to the crop-use database. The irregular-shaped polygons in the NW part of the study area are digitized fields. By transferring the hard copy location data sets into a digital format, individual fields, crop use, and pesticide use can be spatially correlated with other data sets, such as topography, soils, geology to expedite the development of susceptibility analyses. USDA is beginning to implement a nationwide conversion of their data into such a GIS approach. Image is screen save from interactive, web based contaminant inventory (http://gisweb:esri.sc.edu).

The potential value of the crop data is that it presents a reasonable means to estimate the loading of pesticides/herbicides per farm field based on the maximum recommended application rates in the *1997 Pest Management Handbook* from the Clemson Agricultural Extension (see Table 3). Digital crop use data that are available from the FSA on a seasonal basis, can be immediately summarized for each aerial survey grid to arrive at a loading rate for that grid. If actual cultivated field boundaries were available, then the exact location where the pesticide was applied ccould be plotted. USDA is moving towards converting field locations into a GIS format. This USDA initiative should be strongly encouraged because of the benefits to the SWAP program. Precise delineation of cultivated areas will aid in the estimating the amount of specific pesticides introduced annually into the watershed. It will also benefit in susceptibility analyses. Some of these benefits are as follows.

1. Exact routes of any overland flow from a particular application site can be determined by overlaying field locations on topographic coverage (hypsography). This will allow accurate targeting of any necessary use restrictions or remediation efforts.

2. Precise positioning of field locations in a digital format greatly facilitates the correlation of pesticide use with reported soils data to improve reliability of modeling the post-application fate of contaminants (i.e., transport of contaminant into groundwater or surface waters).

Design of Database Structure

In the plan proposed by ESRI-USC for South Carolina (or any State), the inventory assessment methods follow a watershed-approach which is comprised of three key elements. First, a master database is maintained for each public water supply (PWS) source area by the respective PWS operator/owner or by the State environmental agency (SCDHEC). Each PWS master database lists all inventoried sites within the entire source-area watershed(s) upstream of the PWS intake structure and lists all databases where pertinent environmental data are stored for those inventoried sites. Second, each watershed database is linked to original databases maintained and updated by their originating agency (e.g. SCDHEC, US EPA, U.S. Department of Agriculture, other state agencies, etc.). This second element will allow routine updating of inventories without requiring additional large survey projects such as required by *SSWAPP*. Assuming the inventories are updated routinely, watershed susceptibility analyses can be revised with current and timely information. Third, the proposed plan has each PWS database linked to a statewide master database under the supervision of the State environmental agency (Fig. 2).

RELATIONSHIP OF WATERSHED-WIDE CONTAMINANT INVENTORY TO SOURCE WATER ASSESSMENT

The source water assessment process should fully integrate susceptibility analysis with the inventory of present and future sources of contaminants, wherever they may be located within the source area. Following US EPA guidelines, to have a "clear understanding" of the "significant potential sources of contaminants" an entire watershed must be inventoried, or, at the very least, significant sources should not be excluded from the inventory just because their locations lie outside of some arbitrary boundary. This section describes the advantages of a watershed approach to contaminant inventory and the advantages of susceptibility analyses using delineations defined by natural boundaries versus fixed set-back or buffers.

Absolute Distances vs. Buffers / Segmentation

Numerous State *SSWAPP* plans include delineation schemes that consist of offsets or buffers from the surface water body and segmentation of the watershed at some arbitrary time of travel, such as the “24 hour travel distance computed for 10% exceedance flow…” (SCDHEC, 1999). Such a scheme will group potential contamination sites into broad categories. The South Carolina plan delineates three zones of susceptibility (Fig. 6):

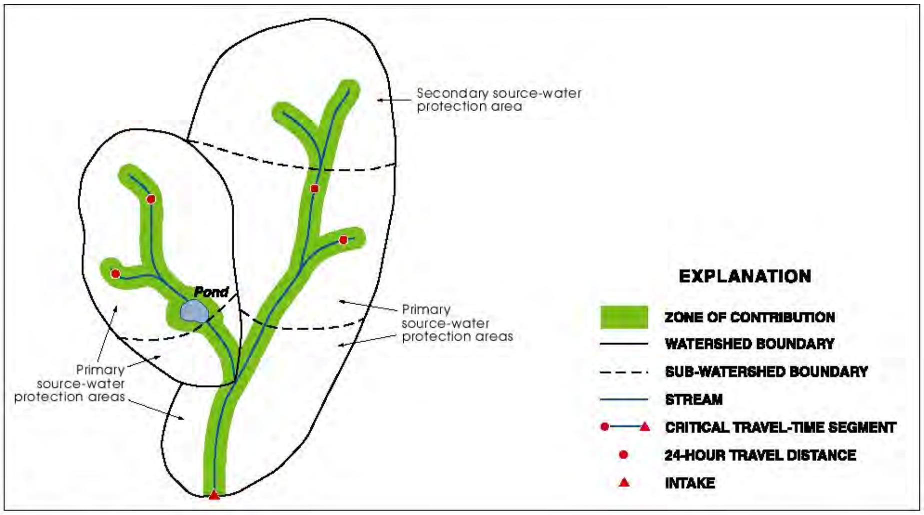
1. the *surface water zone of concern* (61 m offset from edge of stream or geomorphic flood plain);

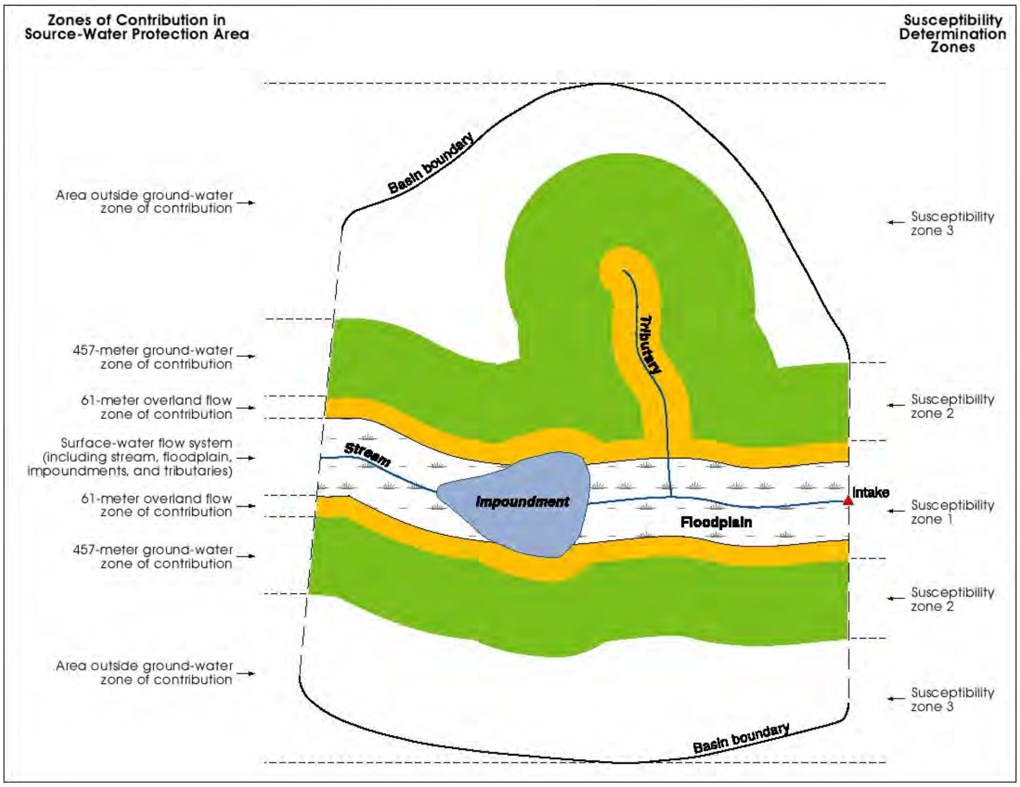
2. the *ground water zone of concern* (457 m offset from edge of stream or geomorphic flood plain);

3. the *area outside the ground water zone of concern* , (The *ground water zone of concern* is defined as that area where shallow ground water may migrate into stream; SCDHEC, 1999).

These three zones are delineated in both the primary and secondary protection areas within a watershed, areas demarcated by a 24 hour time of travel determination. Consequently, the SCDHEC plan delineates a total of six segments within the pilot study watershed (Fig. 7).

Figure 6



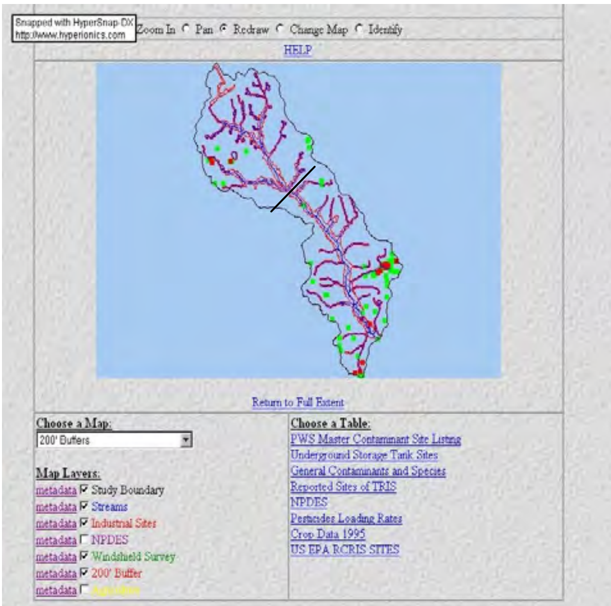


Proposed delineation plan for surface water systems in South Carolina contains *primary* and *secondary source water protection areas* (*SWPA*) and susceptibility offset zones 1,2,and 3. The *primary SWPA* is defined as “all subwatersheds adjoining the 24-hr travel distance upstream from the PWS intake” (SCDHEC, 1999). Off set from the stream trace, the SC plan delineates three zones of susceptibility. Figure is from SCDHEC (1999).

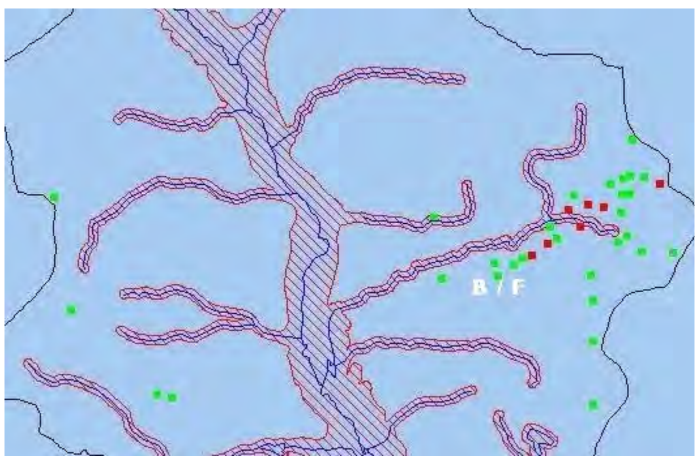
1. The *surface water zone of concern* consists of a 61 m off set from edge of stream or geomorphic flood plain.

2. The *ground water zone of concern* is a 457 m off set from edge of stream or geomorphic flood plain (the *ground water zone of concern* is defined as that area where shallow ground water may migrate into stream).
3. The *Area outside the ground water zone of concern.*

Figure 7



A **B**



C

Application of the proposed delineation plan for surface water systems in South Carolina to the pilot study area segments the watershed into a total of six segments within the pilot study watershed. The black line denote the 24-hr travel distance upstream from the PWS intake" (SCDHEC, 1999). **A.** The *surface water zone of concern,* which consists of a 61 m off set from edge of stream or geomorphic flood plain, only includes 12 of the 62 inventory sites. **B.** The *ground water zone of concern,* which is a 457 m off set from edge of stream or geomorphic flood plain, contains 37 of the 62 inventory sites. **C.** A problem with arbitrary delineations is that "reality" often ignores their presence. For example, in the pilot study area, the

Bridgestone/Firestone (B /F) which lies outside of the SCDHEC designated *Surface water Zone of Concern*. This designation, however, did not prevent water used to fight a fire at the facility from overflowing the site's storm runoff basin and traveling down stream to the PWS intake where it was introduced into the public water supply.

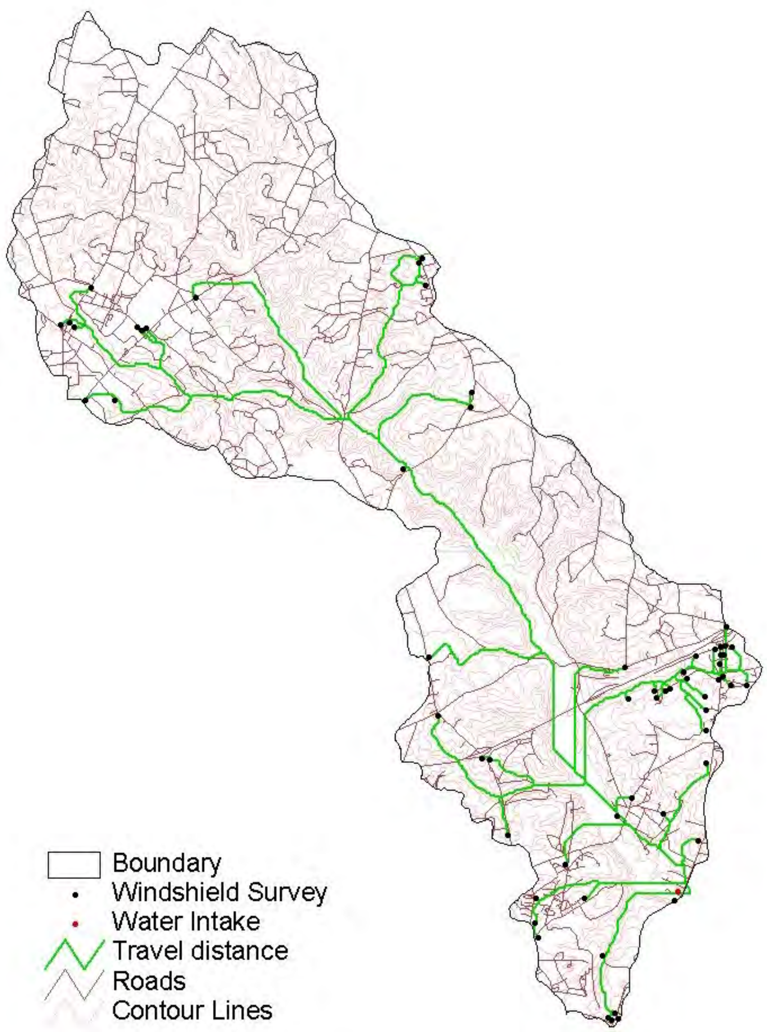
The spatial analytical capabilities inherent within a GIS, however, allow for a much more rigorous analysis, which includes the relative ranking of individual sites.

Using Network Analyst ™ the individual flow-path distances from each site to the PWS intake can be automatically calculated. Figure 8 presents the flow lines, which have flow distances values as attributes from each survey site to the intake. A major advantage to this approach is that not only is there a parameter to individually rank each site, but arbitrary, and often misleading designations are avoided. For example, in the pilot study area, the Bridgestone/Firestone site lies outside of the SCDHEC designated "Surface water Zone of Concern" (Fig. 7 C). This designation, however, did not prevent water used to fight a fire at the facility from overflowing the site's storm runoff basin and traveling the 10.3 km (6.4 miles) down stream to the PWS intake where it was introduced into the public water supply.

Environmental Fate of Contaminant Species vs. Single Susceptibility Determinations

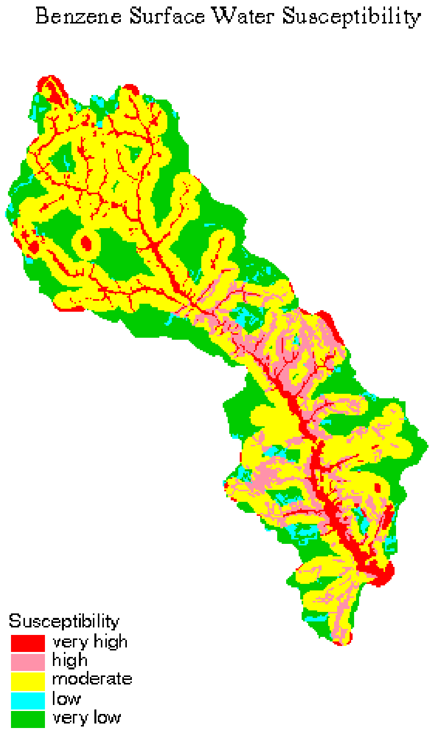
In *SSWAPP*, the US EPA suggests that susceptibility determinations "take into account … inherent characteristic(s) of the contaminants…." One interpretation of this guideline would be separate susceptibility analyses for each species of contaminant, or, at a minimum, maps made for groupings of contaminants that share the same environmental fate. Figure 9 compares hypothetical sensitivity maps for benzene and cadmium. These maps are "stack- unit maps" or "spatially joined" maps, which sum various ranked environmental parameters. How these parameters are ranked depends on the environmental fate of the contaminant in question. For instance, cadmium adsorbs readily onto sediment surfaces, with particles with higher surface areas (e.g., clays) adsorbing higher amounts of cadmium. Consequently, the higher ranked areas on the sensitivity map for cadmium are controlled by spatial distribution of finer grained soils, such as C and D soils which occur as levees and overbank deposits within areas adjoining rivers (Fig. 10). In contrast, for benzene, proximity to the stream is considered the most important parameter as far as benzene's impact on a surface water system. Since benzene volatilizes rapidly, if benzene is released too far from a stream, evaporation combined with infiltration into soils, may prevent benzene from reaching a stream. Consequently, the buffer delineations are the most prominent parameter in the benzene soil sensitivity map (Fig. 7 and 9).

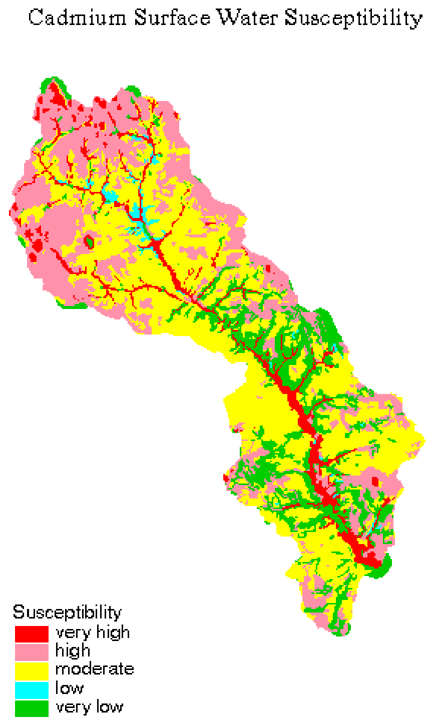
Figure 8



Using Network Analyst ™ (from Environmental Systems Research Institute, Inc.) the individual flow-path distances from each survey site to the PWS intake can be automatically calculated. Using the stream flow calculations for overland as well as in-stream distances, time of travel from each site can be calculated and compared. Time of travel can also be calculated for low flow as well as high flow conditions. Using this tool, a coverage consisting of intersections of streams with roads, rails, and pipelines would give PWS operators quick access to time of travel information in case of an emergency spill due to a transportation accident. This information would be very accessible via the internet.

Figure 9





BENZENE PARAMETERS

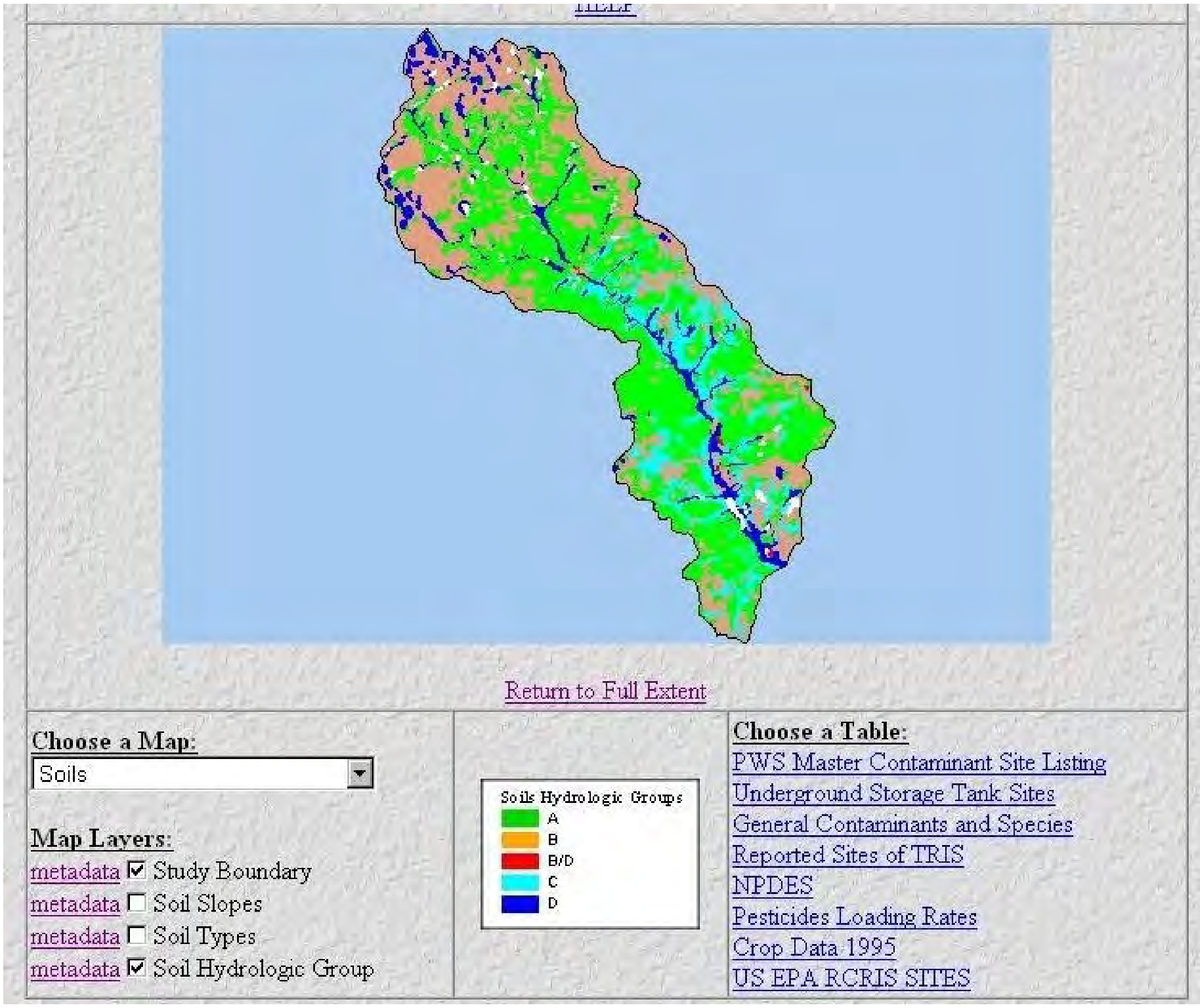
1. SURFACE WATER ZONE OF CONCERN (ZOC)
2. SOIL HYDROLOGIC GROUPS
3. GROUND WATER ZOC
4. SOIL SLOPE

CADMIUM PARAMETERS

1. SOIL HYDROLOGIC GROUPS
2. SOIL EROSION INDEX
3. SURFACE WATER ZOC
4. SOIL SLOPE

Two contaminants on the NPDWS list, benzene, a VOC, and cadmium, a phase II inorganic, behave differently when released into the environment. Consequently, different sensitivity maps where developed for each. These maps are “stack- unit maps” or “spatially joined” maps, which sum various ranked environmental parameters. How these parameters are ranked depends on the environmental fate of the contaminant in question. For instance, cadmium adsorbs readily onto sediment surfaces, with particles with higher surface areas (e.g., clays) adsorbing higher amounts of cadmium. Consequently, the higher ranked areas on the sensitivity map for cadmium are controlled by spatial distribution of finer grained soils, such as C and D soils which occur as levees and overbank deposits adjoining rivers (see Fig. 10). In contrast, for benzene proximity to the stream is considered the most important parameter as far as benzene’s impact on a surface water system. Since benzene volatilizes rapidly, if benzene is released too far from a stream, evaporation combined with infiltration into the soil, may prevent benzene from reaching a stream. Consequently, the buffer delineations are the most prominent parameter in the benzene soil sensitivity map (see Fig. 7).

Figure 10



Screen view is of the soil hydrologic group distribution within the upper Shaw Creek watershed area. Hydrologic soil groups are delineated based on their relative run-off versus infiltration characteristics. Group A soils have a high infiltration rate and low run-off potential, being deep, well-drained, and sandy/gravelly. The other end member, Group D soils, has a very slow infiltration rate, a high run-off potential, and have a high-clay-content layer near the surface or are shallow soils over a low-permeability bedrock or other material. Soil hydrologic group designations are just one of 14 soil attributes from USDA-NRCS Map Unit Interpretation Database.

The variability of the environmental fates for benzene and cadmium is not unique. Given the wide variations in the environmental fate and transport characteristics of hazardous materials, it does not seem scientifically defensible that a single susceptibility analysis should be applied for all the contaminants listed in the NPDW Standards.

General descriptions of the characteristics of hazardous materials, including those listed in the NPDW Standards, can be obtained through the US EPA web site (for an example visit http://www.epa.gov/OGDW/dwh/t-soc.html).

Using Natural Spatial Variability vs. Arbitrary Buffers for Susceptibility Analysis

The diverse data sets describing the geographic, hydrogeologic, hydrologic, and land use characteristics are readily available for watersheds nationwide. These data sets can be easily integrated utilizing the GIS capability of spatial joining (also known as, polygon overlay or stack unit mapping).

The process of integrating varied data sets necessary to develop a susceptibility analysis is relatively straight forward, utilizing a GIS. The next step of developing a logical and defensible scheme for ranking the relative importance of these data is not a simple process. To develop a logical ranking scheme will require the incorporation of ideas and experiences from a variety of sources. To enlist the appropriate expertise, we recommend establishing an advisory board, having discussions at public meetings, and posting of the plan on a web site for review.

When conducting a susceptibility analysis of a watershed or a specific site within a watershed, more than the physical characteristics of the area must be considered. The factors ESRI-USC recommend are considered for a susceptibility analysis of a specific contaminant or groups of contaminants with similar environmental fate characteristics and toxicity levels are as follows:

1. PHYSICAL CHARACTERISTICS
    - soils
    - geology
    - hydrology

2. SITE STRUCTURAL / PROCEDURAL CHARACTERISTICS
    - history of reported releases, complaints, etc.
    - BMP or Emergency Plan in place

3. CONTAMINANT LOADING
    - % of product that is a hazardous material
    - travel distance from release site to PWS intake or well

If a susceptibility analysis lumps contaminants with different environmental fate characteristics and toxicity levels it is suggested that the most conservative environmental fate (environmentally harmful) and most lethal toxicity level of the contaminant group be used in the analysis.

The susceptibility analysis proposed by ESRI-USC is a "relative comparison" approach as described in the USEPA SSWAPP guidelines (1997, p. 2-18). In implementing this approach, the ranking scheme will be clearly documented for critique and later revision. In addition, such an analytical approach can be applied to either surface or ground water systems.

CONCLUSIONS

1. In the susceptibility plan proposed by ESRI-USC for South Carolina (or any State), the inventory assessment methods follow a watershed-approach which is comprised of three key elements. First, a master database is maintained for each public water supply (PWS) source area by the respective PWS operator/owner or by the State environmental agency (SCDHEC). Each PWS master database lists all inventoried sites within the entire source-area watershed(s) upstream of the PWS intake structure and lists all databases where pertinent environmental data are stored for those inventoried sites. Second, each watershed database is linked to original databases maintained and updated by their originating agency (e.g. SCDHEC, US EPA, U.S. Department of Agriculture, other State agencies, etc.). This second element will allow routine updating of inventories without with out requiring additional programs. Assuming the inventories are updated routinely, watershed susceptibility analyses can be revised with current and timely information. Third, the proposed plan has each PWS database linked to a statewide master database under the supervision of the State environmental agency, such as SCDHEC in SC.

2. Given the wide variations in the environmental fate and transport characteristics of hazardous materials, it does not seem scientifically defensible that a single susceptibility analysis should be applied for all the contaminants listed in the NPDW Standards.

3. The spatial analytical capabilities inherent within a GIS allow for a rigorous susceptibility analysis utilizing diverse data sets describing the geographic, hydrogeologic, hydrologic, and land use characteristics which are readily available for watersheds nationwide. These

data sets can be easily integrated utilizing the GIS capability of spatial joining (also known as, polygon overlay or stack unit mapping).

ACKNOWLEDGEMENTS

ESRI-USC wishes to thank SCDHEC for the opportunity to participate in our portion of the pilot studies for *SSWAPP*. ESRI-USC also extends it appreciation to the U. S. Geological Survey (USGS) for their cooperation while we worked coincidentally on our separate portions of this project. ESRI-USC offers a special thanks to the SC office of the Farm Services Agency of USDA for helping us examine and analyze crop use information for the study area.

CITATIONS

Clemson University Cooperative Extension Service, 1997. 1997 Pest Management Handbook, Field Crops, Fruits, and Vegetables. Volume 1: Clemson Agricultural Extension, Clemson, SC, 700 pp.

Larson, S.J., P.D. Capel, and M.S. Majewski, 1997. Pesticides in Surface Waters, Distribution, Trends, and Governing Factors; Volume 3 of the Series, Pesticides in the Hydrologic System: Ann Arbor Press, Inc., Chelsea, MI, 373 pp.

Rine, J.M., R.C. Berg, J.M. Shafer, E.R. Covington, J.K. Reed, C.B. Bennett, J.E. Trudnak, 1998. Development and testing of contamination potential mapping system for a portion of the General Separations Area, Savannah River Site, South Carolina: Environmental Geology, vol. 35, no. 4, p. 263-277.

Rine, J.M., E.R. Covington, J. B. Atkins, 1998. Development of a Contaminant Inventory Methodology for a SWAP Program in South Carolina, Earth Sciences and Resources Institute, University of South Carolina, ESRI-USC Technical Report 98-10-F140, Columbia, SC

South Carolina Department of Health and Environmental Control, Bureau of Water (SCDHEC), 1999. South Carolina Source Water Assessment & Protection Program: Final Draft for Advisory Committee Review January 1999; Columbia, SC

U.S. Environmental Protection Agency, 1997. State Source Water Assessment and Protection Programs Guidance, Final Guidance: U.S. EPA Office of Water, EPA 816-R-97-009, Washington, D.C.

Table 1.
Master Watershed Site List for Source Area to City of Aiken Public Water Supply

USCWID	FAC_NM	AKA FAC NM	ADDRESS	CITY
1	Beaulieu of America	Beaulieu Fibers	136 Frontage Rd.	AIKEN
2	RaceTrac	RACETRAC #427	Columbia Hwy (address#?)	AIKEN
2	RACETRAC #427	RACETRAC #427	2664 N COLUMBIA HWY	AIKEN
2	RACETRAC #427	RACETRAC #427	2664 N COLUMBIA HWY	AIKEN
3	BP Gas & convenience store	DEPOT FOOD STORE 135	2655 Columbia Hwy (US HWY 1 AT I-20)	AIKEN
3	BP Gas & convenience store	DEPOT FOOD STORE 135	US HWY 1 AT I-20	AIKEN
3	BP Gas & convenience store	DEPOT FOOD STORE 135	US HWY 1 AT I-20	AIKEN
4	***Amoco VACANT***	Amoco	Columbia Hwy	AIKEN
5	Smile Gas	SMILE GAS #77	2645 COLUMBIA HWY I-20 & US 1	AIKEN
5	Smile Gas	SMILE GAS #77	2645 COLUMBIA HWY I-20 & US 1	AIKEN
5	Smile Gas	SMILE GAS #77	2645 COLUMBIA HWY I-20 & US 1	AIKEN
5	Smile Gas	SMILE GAS #77	2645 COLUMBIA HWY I-20 & US 1	AIKEN
5	Smile Gas	SMILE GAS #77	2645 COLUMBIA HWY I-20 & US 1	AIKEN
6	S.E. Peters, Inc.		127 W. Frontage Rd.	AIKEN
7	Firestation		Columbia Hwy	AIKEN
8	Pepperidge Farm, Inc.		10 Windham Blvd (Verenes Industrial Park)	AIKEN
8	Pepperidge Farm, Inc.		10 Windham Blvd (Verenes Industrial Park)	AIKEN
9	Beloit Manhatten MillPro Service		25 Beloit St.(Verenes Industrial Park)	AIKEN
10	Smith Kline Beecham	Beecham Products	65 Windham Blvd (Verenes Industrial Park)	AIKEN
11	Carlisle Tire	NI Industries MirreY	25 Windhan Blvd (Verenes Industrial Park)	AIKEN
12	Gorham Bronze	Gorham Bronze Div. Textron	45 Windhan Blvd (Verenes Industrial Park)	AIKEN
13	United Defence LP Ground Systems Div.	FMC Corp.	15 Windhan Blvd (Verenes Industrial Park)	AIKEN
13	United Defence LP Ground Systems Div.	FMC Corp.	15 Windhan Blvd (Verenes Industrial Park)	AIKEN
13	United Defence LP Ground Systems Div.	FMC Corp.	15 Windhan Blvd (Verenes Industrial Park)	AIKEN
14	Aiken Airport BP fuel pump	AIKEN AVIATION ENTERPRISES INC	AIKEN AIRPORT - HWY #1	AIKEN
14	Aiken Airport (fuel pump?)	AIKEN AVIATION ENTERPRISES INC	AIKEN AIRPORT - HWY #1	AIKEN
14	Aiken Airport (fuel pump?)	AIKEN AVIATION ENTERPRISES INC	AIKEN AIRPORT - HWY #1	AIKEN
14	Aiken Airport (fuel pump?)	AIKEN AVIATION ENTERPRISES INC	AIKEN AIRPORT - HWY #1	AIKEN
15	Aiken Airport Sunococ fuel pump	AIKEN AVIATION ENTERPRISES INC	AIKEN AIRPORT - HWY #1	AIKEN
16	Bridgestone/Firestone Special Training		(Verenes Industrial Park)	AIKEN
17	Airport Industrial Park building ***(VACANT)***			AIKEN
18	Aiken Co. Drop off Center #3			AIKEN
19	Southeastern Clay Co.			AIKEN
20	Kentucky-Tennessee Clay Co.		REYNOLDS POND RD N	AIKEN
21	Shell gas & convenient shore	?AIRPORT STOP AND SHOP?	US 1 AT I-20	AIKEN
21	Shell gas & convenient shore	?AIRPORT STOP AND SHOP?	US 1 AT I-20	AIKEN
21	Shell gas & convenient shore	?AIRPORT STOP AND SHOP?	US 1 AT I-20	AIKEN
22	***VACANT (Abandoned Gas station)***	Abandoned Gas station	US 1 (Columbia Hwy)	AIKEN
23	Morris Quick Lube		Columbia Hwy	AIKEN
24	Jordans Auto Service & Sales			AIKEN
25	Shiloh Baptist Church			AIKEN
26	Quick Stop & Gas	QUICK STOP II	2142 Edgefield Hwy	AIKEN
26	QUICK STOP II	QUICK STOP II	HWY 19 & 191 NORTH	AIKEN
26	QUICK STOP II	QUICK STOP II	HWY 19 & 191 NORTH	AIKEN
27	Shell Food Mart & Blimpies	I-20 SHELL	I-20 & HWY 19	AIKEN
27	Shell Food Mart & Blimpies	I-20 SHELL	I-20 & HWY 19	AIKEN
27	Shell Food Mart & Blimpies	I-20 SHELL	I-20 & HWY 19	AIKEN
28	Exxon / Subway (Garvin Oil, Inc.)	KENT'S KORNER #15	1925 Edgefield Rd. (survey notes)	AIKEN
28	Exxon / Subway (Garvin Oil, Inc.)	KENT'S KORNER #15	1925 EDGEFIELD HWY(I-20 @ SC19)	AIKEN
29	Automotive machine Shop		1625 Edgefield Rd.	AIKEN
30	B&B Paint & Body Shop		599 Hazel Dr	AIKEN
31	Curry's Auto Service		Shilo Hts. Rd.	AIKEN
32	Graves Auto Salvage		Edgefield Rd.	AIKEN

Table 1.
Master Watershed Site List for Source Area to City of Aiken Public Water Supply

33	BP Depot		US 1 & Rutland Rd.	AIKEN
34	Williams Garage & Auto Salvage		SC HWY 19	AIKEN
35	2 Fine Fellows US 1		1139 Rutland Rd.	AIKEN
36	Gerry's Automotive		Rutland Rd.	AIKEN
37	Speedway/Starvin' Marvin		US 1 (Columbia Hwy)	AIKEN
38	Grace Div. Kaolin Products			AIKEN
39	City of Aiken Waterworks		US 1 (Columbia Hwy)	AIKEN
40	Chave's Auto		1955 US 1 (Columbia Hwy)	AIKEN
41	Noth Side Golf Club		US 1 (Columbia Hwy)	AIKEN
42	Thomas Used Cars		State Hwy 191	AIKEN
43	Mt. Sinai Baptist Church			AIKEN
44	Satterfield Construction Co., Inc	Asphalt plant #1	SC HWY 191	AIKEN
45	Trenton Quick Stop / Phillip 66	TRENTON QUICK SHOP	HWY 121	TRENTON
45	Trenton Quick Stop / Phillip 66	TRENTON QUICK SHOP	HWY 121	TRENTON
45	Trenton Quick Stop / Phillip 66	TRENTON QUICK SHOP	HWY 121	TRENTON
46	Bill'y Super Store		intersection US1 &SC19 & SC 121	TRENTON
46	Bill'y Super Store		CORNER OF HWY 121 & US 25	TRENTON
46	Bill'y Super Store		CORNER OF HWY 121 & US 25	TRENTON
46	Bill'y Super Store		CORNER OF HWY 121 & US 25	TRENTON
46	Bill'y Super Store		CORNER OF HWY 121 & US 25	TRENTON
46	Bill'y Super Store		CORNER OF HWY 121 & US 25	TRENTON
47	Martin Color Fi		Trenton 121	AIKEN
48	Carlisle Tire & Wheel		intersection US1 &SC19 & SC 121	TRENTON
49	Trenton Correctional Institute	TRENTON CORRECTIONAL CENTER	85 OLD PLANK RD	TRENTON
49	Trenton Correctional Institute	TRENTON CORRECTIONAL CENTER	85 OLD PLANK RD	TRENTON
50	Edgefield Co. Collection & Recycling center			TRENTON
51	Ebenezer Baptist church			TRENTON
52	Carlisle Engineering Products		E. Wise Trenton	TRENTON
53	Moore Craft Cabinet Co.		E. Wise Trenton	TRENTON
54	Yonce Field (grass airfield)		off SC HWY 121 connector	TRENTON
55	191 Country Convenience Store		177 Rt. 191	SC
56	Resale shop and Auto service (abandoned gas pumps)		Box 28 Rt. 2 (SC Hwy 191)	SC
57	Lewis E. Holmes Farms			AIKEN
58	Condrey's Auto and Cycle Shop			AIKEN
59	Aiken Correction Center		ROUTE 4 BOx 494-B US 1 N	AIKEN
60	Ramada Inn		110 Frontage Rd.	AIKEN
61	Edgefield Co. Water & Sewer Authority		???	TRENTON
62	Feldspar Production Inc.		P.O. Box 2455	AIKEN

KEY TO HEADER ABBREVIATIONS:

USCWID	**unique ID# for each inventoried site in watershed**	**DHECSPCCID**
FAC_NM	**facility name**	**DHECNPDES**
AKA FAC_NM	**other / previous facility name**	**DHECSTORMW**
DHECUSTID	**UST permit ID#**	**EPA FACILITY ID**
surv-id	**site ID numbers utilized in windshield survey**	**RCRIS CITE**
		TRIS CITE

Table 2.
Listing of General Types of Potential Contaminants from Industrial/Govermental Facilities as Derived from Facility SIC Descriptions

SIC	SIC DEFINITION	SIC GROUP	SIC MAJOR GROUP DEFINITION	**PETROLEUM**	Benzene	Ethylbenzene	Styrene	Toluene	Xylenes (total)		**CHLORINATED SOLVENTS & OTHER VOCs**	Carbon Tetrachloride	p-Dichlorobezene	1,2-Dichloroethane	1,1-Dichloroethylene	Trichloroethylene	1,1,1 Trichloroethane	Vinyl chloride	Chlorobenzene	o-Dichlorobenzene	cis & trans Dichloroethylene	PCB's	Tetrachloroethlylene	Dichloromethane	Benzo(a)pyrene (PAHs)	Phthalate	1.2 Dichloropropane	1,2,4 - Trichlorobenzene	1,1,2 Trichloethane	*Total trihalomethanes (Interim)*
0191	General farms, primary crop	01	Agricultural production crops	**Y**	Y																									
1455	Kaolin and ball clay, mining	14	Mining & quarrying of nonmetallic minerals, ex+D6cept fuels																											
1629	Heavy construction not elsewhere classified	16	Heavy construction other than building construction contractors	**Y**	Y	Y					**Y**	Y								Y				Y						
2051	Bread and other bakery products, eYcept cookies & crackers, manufacturing	20	Food & kindred products	**Y**	Y																									
2282	Yarn texturizing, throwing, twisting, and winding mills, manufacturing	22	Textile mill products																											
2511	Wood household furniture,except upholstered	25	Furniture and fixtures	**Y**	Y		Y				**Y**	Y		Y	Y		Y							Y						
2821	Plastics materials and resins, manufacturing	28	Chemicals and allied products	**Y**	Y		Y	Y			**Y**	Y				Y		Y		Y		Y							Y	
2824	Organic fibers, noncellulosic, manufacturing	28	Chemicals and allied products	**Y**			Y	Y			**Y**				Y	Y	Y	Y		Y		Y							Y	
2834	Pharmaceutical preparations	28	Chemicals and allied products	**Y**	Y		Y																							
2899	Chemicals and chemical preparations, not elsewhere classifiec, manufacturing	28	Chemicals and allied products	**Y**	Y	Y	Y	Y	Y		**Y**	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y			Y	Y	Y	
2951	Asphalt paving mixtures and blocks	29	Petroleum refining and related products	**Y**	Y	Y	Y	Y	Y																					
3011	Tires and inner tubes manufacturing	30	Rubber and miscellaneous plastic products	**Y**			Y	Y			**Y**																		Y	
3069	Fabricated rubber products, not elsewhere classified, manufacturing	30	Rubber and miscellaneous plastic products	**Y**			Y				**Y**	Y				Y	Y												Y	
3089	Plastic products, not elsewhere classified, manufacturing	30	Rubber and miscellaneous plastic products	**Y**	Y		Y	Y			**Y**	Y			Y							Y				Y				
3295	Minerals and earths, ground or otherwise treated, manufacturing	32	Stone, clay, glass and concrete products																											
3297	Nonclay refractories	32	Stone, clay, glass and concrete products																											
3364	Nonferrous die-castings, eYcept aluminum, manufacturing	33	Primary metal industries								**Y**					Y	Y		Y					Y						
3492	Fluid power valves and hose fittings, manufacturing	34	Fabricated metal products, eYcept machinery and transportation equipment	**Y**				Y			**Y**	Y				Y	Y		Y			Y		Y						
3499	Miscellaneous fabricated metal products	34	Fabricated metal products, eYcept machinery and transportation equipment	**Y**				Y			**Y**	Y				Y	Y		Y			Y		Y						
3599	Industrial and commercial machinery and equipment, not elsewhere, manufacturing	35	Industrial and commercial machinery and computer equipment	**Y**				Y			**Y**	Y				Y	Y		Y			Y								
3714	Motor vehicle parts and accessories	37	Transportation equipment	**Y**	Y			Y			**Y**	Y			Y	Y	Y		Y	Y				Y			Y			
3795	Tanks and tank components	37	Transportation equipment																											
4581	Airports, flying field and airport	45	Transportation by air	**Y**	Y						**Y**	Y							Y	Y				Y						
4941	Water supply	49	Electric, gas, and sanitary services								**Y**																			Y
4952	Sewerage systems	49	Electric, gas, and sanitary services																											
5015	Motor vehicle parts, used (junkyards)	50	Wholesale trade-durable goods	**Y**	Y	Y					**Y**	Y			Y	Y	Y													
5093	Scrap and waste materials	50	Wholesale trade-durable goods																											
5411	Grocery stores	54	Food stores																											
5521	Motor vehicle dealers (used only)	55	Automotive dealers and gasoline service stations	**Y**	Y	Y		Y			**Y**	Y			Y	Y	Y			Y				Y						
5541	Gasoline service station	55	Automotive dealers and gasoline service stations	**Y**	Y	Y		Y			**Y**	Y			Y	Y	Y			Y				Y						
5812	Eating places	58	Eating and drinking places																											
6512	Operators of nonresdential buildings	65	Real estate																											
7011	Hotel and motels	70	Hotels, Rooming Houses, Camps, And Other Lodging Place																											
7538	General automotive repair shops	75	Automotive repair, services, and parking	**Y**	Y	Y		Y			**Y**	Y			Y	Y	Y			Y				Y						
7549	Automotive services, eYcept repair	75	Automotive repair, services, and parking	**Y**	Y	Y		Y			**Y**	Y			Y	Y	Y			Y				Y						
7992	Public golf courses	79	Amusement and recreation services																											
8661	Religuous organizations (church yard with cemetery)	86	Membership organizations																											
9223	Correctional institutions (w/UST?, vehicle maintenance, small manufacturing)	92	Justice, public order, and safety	**Y**	Y	Y		Y			**Y**	Y			Y	Y	Y			Y				Y						
9223	Fire protection ((Fire station)	92	Justice, public order, and safety																											

Table 2.

Listing of General Types of Potential Contaminants from Industrial/Govermental Facilities as Derived from Facility SIC Descriptions

	PESTICIDES/HERBICIDES	Acrylamide	Alachlor	Aldicarb, etc.	Atrazine	Carbofuran	Chlorodane	2,4 -D	Dibromochloropropane	Epicchlorohydrin	Ethylene dibromide	Hepachlor + metabolites	Lindane	Methoxychlor	Pentachlorophenol	Toxaphene	2,4,5-TP	Dinoseb	Diquat	Dioxin	Endothall	Endrin	Glyphosate	Hexachlorobenzene	Hexachlorocyclopentadiene	Oxamyl	Picloram	Simazine	Adipate	Dalapon		**INORGANICS**	Asbestos	Barium	Cadmium	Chromium	Mercury	Lead	Copper	Antimony	Berylium	Cyanide	Nickel	Thallium	Selenium	*Sulfate (proposed)*	*Arsenic (Interim)*		**MOBILE INORGANICS**	Nitrate	Nitrite		*RADIONUCLIDES*	*Radium (proposed)*	*Radon (proposed)*	*Uranium (proposed)*		**PATHOGENS**	Giardia	Total Coliform	Turbidity	Viruses
	Y		Y	Y	Y	Y			Y									Y			Y					Y				Y																			**Y**	Y	Y											
																																																									Y			Y	
	Y														Y																	**Y**		Y																												
	Y									Y																						**Y**										Y																				
	Y									Y																																																				
																																Y												Y																		
	Y								Y	Y										Y												**Y**		Y	Y	Y				Y	Y	Y	Y	Y					**Y**	Y												
	Y																												Y																																	
	Y																												Y																																	
																																Y										Y																				
																																Y		Y				Y				Y																**Y**			Y	
																																Y				Y							Y		Y		Y											**Y**			Y	
																																Y				Y		Y	Y		Y	Y	Y																			
																																Y		Y	Y	Y	Y			Y		Y	Y																			
	Y							Y									Y						Y					Y		Y																																
	Y									Y																																																			
	Y	Y								Y																																							**Y**	Y	Y											
	Y										Y																																																		
	Y							Y									Y						Y																																		**Y**		Y		Y
	Y							Y									Y						Y																																							
																																	Y								Y																						

Table 3.
Estimated loading rates of specific pesticides/herbicides per acre farmed land based on maximum recommended application rates in the 1997 Pest Management Handbook from Clemson University Cooperative Extension Service

CROP ABREV.	CROP NAME	PESTICIDES / HERBICIDES	Acrylamide	Alachlor	Aldicarb	Atrazine	Carbofuran	Chlorodane	2,4 -D	Dibromochloropropane	Epicchlorohydrin	Ethylene dibromide	Hepachlor + metabolites	Lindane	Methoxychlor	Pentachlorophenol	Toxaphene	2,4,5-TP	Dinoseb	Diquat	Dioxin	Endothall	Endrin	Glyphosate	Hexachlorobenzene	Hexachlorocyclopentadiene	Oxamyl	Picoloram	Simazine	Adipate	Dalapon
APPLE	APPLE		0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0	2-4pt/1000 gal	0.0	4.0	0.0	0.0
BEANS	BEANS		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CABAG	CABBAGE		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CORN	CORN		0.00	3.00	0.00	2.50	1.00	0.00	1.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.50	0.00	0.00	0.00	0.00	2.40	0.00	0.00
CRP	Conservation Reserve Plan		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CUCUM	CUCUMBER		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DPEAS	PEAS		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
FALOW	FALLOW		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GRAPE	GRAPE		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.50	0.00	0.00	0.00	0.00	4.00	0.00	0.00
GRASS	PASTURES		0.00	0.00	0.00	0.00	0.00	0.00	3.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4lb/gal	0.00	0.00	0.00	0.00	4.00	0.00	0.00
MELON	MELON		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MILET	MILLET		0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OATS	OATS (SMALL GRAINS)		0.00	0.00	0.00	0.00	0.25	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
OKRA	OKRA		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PEACH	PEACH		0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.00	0.00	0.00	0.00	0.00	4.00	0.00	0.00
PEARS	PEARS		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PEPRS	PEPPERS		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PLUMS	PLUMS		0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.00	0.00	0.00	0.00	0.00	4.00	0.00	0.00
RYE	RYE(SMALL GRAINS)		0.00	0.00	0.00	0.00	0.25	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SORGH	SORGHUM		0.00	2.75	7.00	1.20	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SOYBN	SOYBEANS		0.00	3.00	0.00	0.00	0.00	0.00	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SQASH	SQUASH		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SWCRN	CORN (SWEET)		0.00	2.75	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.00	0.00	0.00
SWTPO	SWEET POTATO		0.00	0.00	3.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.72	0.00	0.00	0.00	0.00
TMATO	TOMATO		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	9.00	0.00	0.00	0.00	0.00
UPCN	UPLAND COTTON		0.00	0.00	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.50	0.00	0.00	8.5-17 oz/ac	0.00	0.00	0.00	0.00
WHEAT	WHEAT (SMALL GRAINS)		0.00	0.00	0.00	0.00	0.25	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Management and Reuse of Contaminated Soil – The SoilTrak™ Method

Edward Rogers, Jr., BEM Systems, Inc.[1]

Introduction

In 1991, BEM Systems, Inc. was contracted to conduct environmental investigations along a 20+ mile corridor for development of a new light rail transit system project. During these investigations, suspicions that widespread and varying degrees of contamination would be found were confirmed.

At about the same time, a new approach to contamination had recently been developed by the State regulatory agency wherein low-level contamination could be left in place based on proposed land-use and the imposition of engineering controls to limit potential exposure to the contaminants. This would allow for the redevelopment of mildly contaminated properties for industrial use where the likelihood of exposure to contamination would be far less than that encountered for residential development, for example. If such an approach were implemented, the developer would have to negotiate site-specific cleanup standards, design and implement engineering controls to limit any potential exposure to remaining contamination and map out the type and extent of contamination allowed to remain on-site. This information would be attached to the property deed. The document would further stipulate that as long as the land use remained as proposed, the contamination could remain in-place. Should the land use change, the issue of contamination and remedial options would have to be revisited. A document stating the allowable land use and identifying the contaminants, their locations and concentrations would be attached to the property deed. This document is referred to as a “Declaration of Environmental Restrictions” (DER).

As part if its investigations, BEM recommended that contaminated soil reuse be implemented as the preferred remedial alternative and that a Memorandum of Agreement (MOA) should be negotiated with the State. The proposal involved reuse of over 400,000 cubic yards of “moderately” contaminated soil using engineering controls to minimize exposure potential to the contaminants that would remain. This would minimize the costs associated with off-site disposal

[1] Edward Rogers, Jr., BEM Systems, Inc., 100 Passaic Avenue, Chatham, NJ 07928;

of soil and would provide much needed engineering fill material. Since the proposed reuse of the property was a rail corridor, the probability of public exposure to contamination would be nearly zero. The state agency agreed to this approach with the following stipulations. Reused soil must not have contaminant levels that are greater than 10 times the State agency's "Non-Residential Direct Contact Cleanup Standards". The contaminated soil would have to be placed at least two feet above the seasonal high water-table level and two feet below grade. Another stipulation of the MOA was that mapping would be required which showed where contaminated soil was placed, which contaminants traveled with it and at what concentrations.

For the typical property, implementation of such a plan is fairly straightforward. Such properties are generally no more than a few acres in size and identifying the degree and extent of contamination is "text book" environmental work. BEM, however, was working with a single "site" that was 20 miles long, involved most urban land uses and passed through six municipalities and two counties. BEM would have to develop a detailed system of tracking and reporting soil movements and identifying their associated contaminants.

As though this task were not complex enough, the client added another twist. It was the transportation agency's intention to recover the costs of remediation from the owners of the properties they planned to acquire. Since the transportation agency could approach the corridor as one site, they would enjoy an economy-of-scale with regard to investigation costs, remediation costs, and more favorable cleanup standards that no individual property owner could achieve alone. Using this information, it was reasoned that property owners would settle on the final sale price of the property sooner if we could demonstrate that the cost of remediation, which would be deducted from the sale price, was lower than the property owner would have paid. Timely property acquisition was key to the success of this rail project. It would be necessary to provide a detailed accounting of the remediation cost, by property owner, to defend our client's cost recovery efforts.

It quickly became clear that BEM would have to develop a method of tracking soil movement activity throughout the corridor. In light of the amount of data that would be generated by the remediation/construction activity, a database of some type would also have to be developed. In order to meet the objectives of cost recovery and the DER, this database would have to integrate tightly with other parts of BEM's environmental database management system. This

---

908-598-2600, x143; erogers@bemsys.com

would be particularly true for sample analytical data and property owner information, as well as existing spatial databases of the corridor that included sample locations and tax lot boundaries.

This paper will present the design and implementation of the SoilTrak Method, including the issues leading to design decisions, the successes and failures of SoilTrak, an evaluation of whether intended objectives were met, and where SoilTrak is headed now.

Developing the SoilTrak™ Strategy

The first step in developing the SoilTrak Method was determining precisely what information would be needed and what would have to be tracked. It seemed clear that we would have to track soil movement itself. Since it was necessary to know not only the contaminants at source locations, but also the specific source property, a grid system based on tax lot boundaries and the width of the corridor seemed ideal. This would allow the contractors to report soil movement based on a source cell and a destination cell. This seemed an ideal solution to everyone except the first of the contractors we would work with. SoilTrak would be developed and tested on a small segment of the corridor that would become the largest of the several “Park & Ride” facilities. “Contractor 1” was hired to do the construction of this site. Contractor 1 stated that the unusual shapes of the tax lot parcels in this area, many of which had sweeping curves in them, were too difficult to map and track during field activities. They proposed developing an irregularly shaped grid system based upon activity and depth of excavation. The cells would either be “cut” (excavation) cells or “fill” cells. Since anything else would be too complex to manage in the field, this approach was adopted. The contractor supplied us with a CAD file mapping the proposed cut, which would later be critical to implementing the system. From this point, development of the SoilTrak database and application began.

Database Design

As with most database projects, reporting needs would primarily dictate what information was stored in the database. It was already known that the primary needs were to support cost recovery by property owner and to report the source, destination and contaminant content of reused soil. BEM’s environmental management information system (BEMIS™) already had a database and associated GIS coverage for tax lot information along the corridor, so SoilTrak would be designed to integrate with these. BEMIS also had tables containing analytical results data and a sampling location table with an associated sample location GIS coverage. It is important to understand that BEMIS was designed to be data-centric, rather than GIS-centric.

All attribute information is stored in a central database of many related tables with GIS coverages and applications like SoilTrak acting as data “peripherals”. They could draw data from, or add data to, the specific tables they supported. This strategy allowed SoilTrak to focus only on the tables needed to track soil movement and cost information. These tables would be SoilTrak’s contribution to BEMIS.

The first, and main, table would track soil movements (SoilMove). A key question on the contents of a record centered on whether a record should record the source *and* destination, or only source *or* destination. That is, would a record report the cut and subsequent fill activity, or the cut and fill activities separately. The first approach, and seemingly the most intuitive, could produce half as many records as the second could. However, it was determined the soil cut on one day could be temporarily stockpiled and filled on a later date. This would create confusion during data entry since SoilTrak would require both a source and destination cell. In many cases, soil would also be cut from a single cell and filled to several cells. This would result in significant redundancy in the data since the single cut would have to be written for every fill. It was decided to record the activities as separate records.

The next decision centered on precisely what it was that should be tracked. Since soil movement was the focus, it seemed logical to track the movement of each discreet amount of soil removed from the ground, which we will refer to as “slugs” of soil. Every time a soil excavation activity occurred, the soil removed would now become an entity in the database and would receive a unique ID, a “Soil Code”. Other information such as excavation date, volume, a cell ID and an activity (was the record of this soil reporting a cut or fill?). Fill records would be virtually identical to cut records. The Soil Code would be the primary key of SoilTrak, an approach that would later prove not to be optimal.

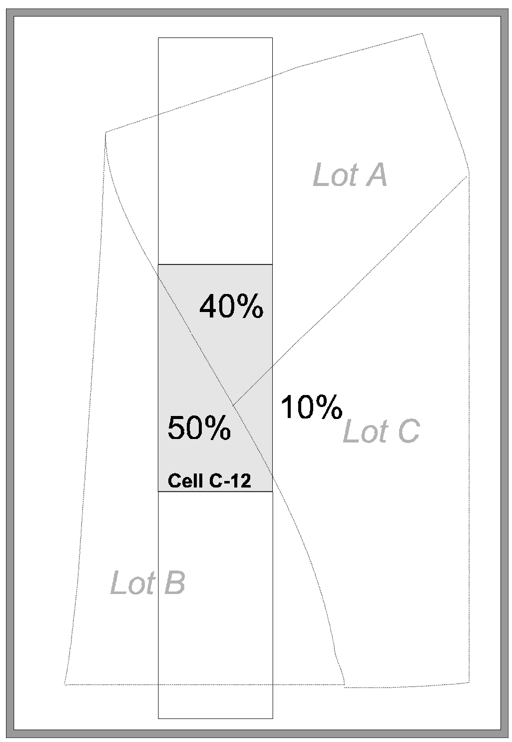


Figure 1

Soil volumes would be assigned as a ratio of the volume to the area a tax lot occupied within the source cell.

This information was adequate for tracking soil movements but could not meet either of SoilTrak’s objectives of cost recovery and contaminant tracking. Another piece of information was necessary, property ID. Since BEM already had a database of tax blocks and lots, It seemed reasonable to use this existing data. Each cut or fill activity would occur within a tax block and lot combination identified within BEMIS by a property code. The inclusion of this field would provide SoilTrak with the needed link into the rest of BEMIS.

Database Design – Cost Recovery

However, the inclusion of prop_code in a soil move record created a dilemma. The cut and fill grid was drawn without regard to property boundaries. If soil were excavated from a cell that was subdivided by multiple property codes, which would you report? Since the smallest geographic unit the contractors could report on was the cut or fill cell, it would be necessary to report all properties present in the cell. Cost recovery objectives dictated knowing how much soil was removed from each property, not cut or fill cell. In the case where a cell were divided by three tax lots, SoilTrak's data approach would require the three cut records be entered for every cut from that cell.

With this in mind, the next question centered on assigning the soil volume. For all intents and purposes, soil excavation and reuse was the cost-recoverable remediation activity. If a rate per cubic yard could be arrived at, cost recovery reporting would be as simple as multiplying that rate by the total volume of soil removed from a specific tax lot (property code). However, soil volumes were being reported by cut cell. This problem was solved using BEM's GIS. At the time, BEM was using ESRI's ArcCAD™ GIS product. This product used AutoDesk's® AutoCAD™ as its graphics engine and created PC ArcInfo coverages for storing GIS related data. BEM had polygon coverages of both the tax lot boundaries and the proposed cut cells. In ESRI's model of the GIS universe, area features are recorded as polygon coverages, with each polygon having an area and perimeter. These two pieces of information are reported in every record of an ESRI ArcInfo polygon attribute table. If one were to assume that any soil removed from a cell was removed from the entire area of the cell and to a constant depth, the area information could be useful (Figure 1). Since ArcCAD is a true GIS application, it would be possible to intersect the tax lot coverage with the cut and fill cell coverage. This would produce a new polygon attribute table containing the property code, cut or fill cell ID and the area of the polygon (Figure 2). This table, a Dbase file, was lifted from the coverage and put into a spreadsheet. From there the percent of the total area of a cell present in a tax lot was computed. This number would be used as a volume multiplier. If an excavation volume were reported as 150 cubic yards for a cell, and the cell were subdivided equally by three different tax lots, three records would be entered into SoilTrak with 50 cubic yards of the soil attributed to each record, and thus, each lot.

Pat

Area	Perimeter	Cnf_prop_	Cnf_prop_i	Cnf_real_i	Prop_code
-7142316.0000	0.9918399E+05	1	0	0	0.00000
0.1858992E+04	0.1975570E+03	2	1	312	526.00000
0.8201172E+01	0.5457860E+02	3	2	312	531.00000
0.7009863E+04	0.4816174E+03	4	3	313	526.00000
0.1054805E+03	0.7981525E+02	5	4	313	0.00000
0.6209745E+05	0.1980884E+04	6	5	314	100.17000
0.9944193E+04	0.5001563E+03	7	6	315	526.00000
0.4609180E+02	0.3903281E+02	8	7	315	0.00000

Figure 2

Intersecting the tax lot coverage (prop_code) with the cell coverage (cnf_real_i) yielded a table with areas for each fraction of a lot within a cell. These were used to find the multipliers needed to assign volumes to owners.

Knowing the volume of soil removed from a cell alone was not enough to determine recoverable costs. SoilTrak was programmed with a default value for cost per cubic yard. However, a module was added that allowed for line item costs to be entered by cell. Thus, while one truck may have been used for cut cell C-1, two may have been used for cut cell C-2. Off-site disposal may have been required for only a few cells, though this could add greatly to the cost of remediation within that cell. The cell cost module allowed for cell-specific rates that could be calculated and applied to cost recovery reports. This approach also provided a more defensible and accurate picture of the actual remediation costs for property owners.

Database Design – Contaminant Tracking

This aspect of SoilTrak was quite different from the property considerations of cost recovery. Contaminated soil was to be reused as fill material. In this case, the only information of interest was what contaminants were present in the reused soil, where the soil was placed and what volume of soil was placed there. Every soil placed as fill would have to have an origin and a means of determining which contaminants were present. Soil would be cut from cell C-12 and placed in cell F-15. Recall, however, that during the data entry phase, if property lines split a cell, multiple records for the cut activity were entered. That is, multiple “soil codes” were created for a single cut cell in a single soil move event. Which one do we list as the source? The answer is all of them. This was facilitated through a utility that allowed mixing of soil. If three soil codes existed for a cut activity, this module would allow you to mix them together to form a single new soil code. This would then be the code used for filling activities. A later re-analysis of the SoilTrak method would reveal that this was far from an optimal approach, though it did work. A separate table called “SoilMix” was used to store this “fake” mixing information.

The issue then centered on determining which contaminants were present. In an ideal model, the distribution of contamination would be known in at least two-dimensions for the entire corridor. However, creating such fine resolution of contaminant data over 20 miles would require far more analytical samples, and thus cost, than could be supported or justified. Another approach would find all sample locations within the cut cells and assume that any contaminants present at those locations were present throughout the cell. An examination of the data showed that too few samples were located directly within the cut cells. Since many of the samples were collected to characterize an entire tax lot, they were spread too thin to use only those that fell within the cut cells. The final, and selected approach, involved using the GIS to determine all samples collected by tax lot. Then, any cell that crossed into the tax lot was assumed to have any contamination found within the lot, regardless of whether the sample was collected within the cell (Figure 3). This provided a conservative, but reproducible and systematic approach. The GIS was used to identify samples by tax lot, and tax lots within cells. These data were tabulated and provided to SoilTrak which then used its connection to the BEMIS database to determine which contaminants were present, at what concentrations and whether or not they exceeded State concentration standards.

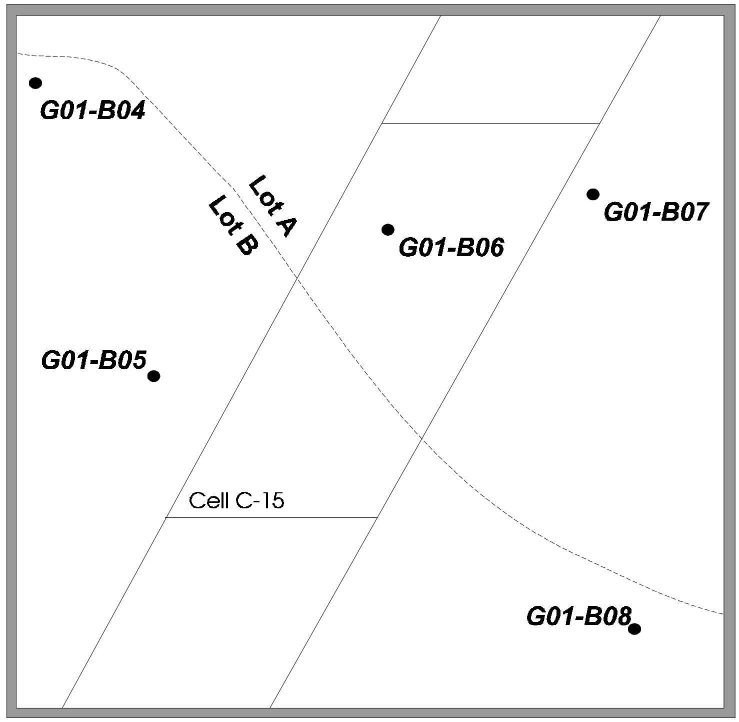


Figure 3

The GIS was used to determine in which lots samples had been collected. This allowed SoilTrak to assign any sample in a lot to its associated cell. Thus, the contaminants found in any of the sample locations above would be said to be present in Cell C-15.

Recall that the MOA stated that any reused contaminated soil would have to be at least two feet below grade and two feet above the mean high water table. This was handled by allowing the starting and finishing elevations for each fill to be reported.

Implementation – Contractor 1

In order to provide the data that would be needed for the SoilTrak method, forms were developed for use by field personnel. These essentially reported the date, source, destination and volumes of soil moved. Cell specific cost sheets were later provided by the contractor and used in the cell-specific cost module described above.

The contractor then provided a CAD map showing the proposed cut cells. Fill cell mapping was not provided, as the cells had not yet been determined. The CAD file was converted into an ArcCAD polygon coverage and later intersected with an existing tax lot coverage to provide information that SoilTrak would need.

The SoilTrak application was developed primarily in the spring of 1996 and was on-line by July of the same year. The application was created using Microsoft FoxPro® version 2.6 with all data tables in FoxPro's native dBase® (*.dbf) format. This choice worked well with our then current GIS platform, ArcCAD, since it too stores its attribute data as dBase files.

Initial implementation of SoilTrak and the SoilTrak method went well and the required objectives were met. During implementation however, one issue did arise. As described above, the contractor provided CAD files of the proposed cut cells, with each cell having a constant depth of excavation. Fill cells however were not mapped at all. Fill areas were decided and recorded in the field. Initially, fill cells were so large in areal extent that making any statements about the locations of specific contaminants within the cells would be meaningless. At BEM's request, the contractor began to subdivide the fill cells. These new cell boundaries were largely dictated by "lifts" of soil that were placed, usually from a particular source area. Unexpectedly, however, these cells began to overlap each other. Recall that filling was occurring in three dimensions. These cells were not truly overlapping since they were being defined at different depths. To a two-dimensional GIS however, these cells did overlap. This apparent overlapping made mapping fill cells in the GIS impossible since the GIS would try to create new polygons every time a cell was overlapped. This would have made assigning attributes to fill cells unmanageable, as there would be multiple many-to-one relationships between cell polygons and cell attributes. At that point in time however, the SoilTrak application was not GIS enabled and thus the issue of unmanageable fill cell mapping had no impact.

Implementation – Contractor 2 – 1998

Several months had passed and SoilTrak had been a success. A new and larger phase of the corridor construction was about to begin and BEM again prepared to use this new tool. Some lessons had been learned from the Park & Ride construction implementation. Among these were:

- The first version did not include a module for entry of cell cost details (line-item costs). Such a module was added by the conclusion of the Park & Ride effort;
- Cells would not be allowed to overlap at all, each cell would be discrete. This would support GIS mapping;
- All cell mapping was to be completed prior to the start of soil movement activities.

The Park & Ride test area was fairly small, measuring only 14 acres. The next activity would involve half of the corridor, covering a distance of approximately 11 miles.

Work began in the spring of 1998 with meetings between BEM and the new contractor. The field reporting form was refined and initial efforts at cell mapping were made. The corridor consisted of several Park & Ride facilities and a large light rail maintenance facility. Work was to begin at the large maintenance facility so mapping was first provided for this facility with Park & Ride facilities close behind.

Cell Mapping

As planning continued one new change became apparent, there would be no “cut” or “fill” cells. Previously, cell mapping had been based on the depth of excavation and cutting or filling activity. This contractor chose a systematically based cell system instead, where cells were 200 feet long and 100 feet wide in the aforementioned areas. The effect of this system was that both cutting and filling could occur within the same cell. Cell IDs were based primarily on station numbers along the corridor. Each cell would be named using the station number it started at going from south to north. The prefix “C” or “F” would identify the cell as a cut or fill cell. Since both activities could occur within a cell, each cell was given both a cut and fill cell ID. This would be no problem for the SoilTrak application but posed a problem for the GIS, as it could not handle one (polygon) to many (Cell IDs) relationships. This would be addressed later since the SoilTrak application did not yet support GIS mapping.

With cell mapping completed for the Park & Ride areas and the large maintenance facility, construction began. Initially, the SoilTrak Method worked smoothly, but soon, problems began to surface.

Incomplete Cell Mapping

As stated previously, cell mapping had been received for only certain areas. Eventually, construction began along the actual corridor portions of the project area for which no cell mapping had been received. Many requests were made for this information, however Contractor 2 was not as “computer savvy” as Contractor 1 and was slow to respond with any additional mapping. Maps that were provided often consisted of small-scale paper maps showing only tick marks every 200 feet along the corridor. However the width of the cells was never provided. Without this information, it was impossible to create the CAD drawings needed to build a GIS coverage. It was later stated that the width of the cells was the corridor “right-of-way”. Obtaining CAD files with complete mapping of the right-of-way lines turned out to be more difficult than one would have anticipated. CAD files existed with complete design information, but right-of-way lines were hard to identify as many were not labeled and easily lost among many other unlabeled lines. Recall that a cut-fill cell GIS coverage was needed to provide the cell-to-property data needed for data entry. While soil movement activities continued, SoilTrak data entry ground to a halt. Many months would pass before BEM was able to build a usable cell coverage while questions continued as to the accuracy of the right-of-way width designations. This delay allowed field data reporting problems to go unnoticed, and thus unaddressed, for months.

Significant Issues

By this point, flaws in the SoilTrak method and approach began to surface; these centered on designation of contaminated areas and the practice of soil mixing via stockpiling. Each is addressed separately below.

Contaminated Cells

Recall that SoilTrak was created to track the reuse of contaminated soil. Inherent to this mission is the determination of what is, and is not, contaminated. Daily soil movement sheets reported all soil movement, contaminated or not. There was a place on the form to designate which soil movements involved contaminated soil. With this information, only entries for contaminated soil movement were entered into the SoilTrak application. It was later discovered that there was a lack of clarity between contractors and BEM staff on how to determine if a cell was contaminated. Different methods were being used, creating data accuracy problems. Some cells had been reported as being contaminated and later were considered “clean”, and vice-versa. This again brought SoilTrak data entry to a halt while a time-consuming purge of bad data ensued. This was made additionally painful by the fact that “fake-mixing” and filling using these “mixes” pervaded the two main database tables. A consistent standard was developed and necessary data was reentered and bad data expunged. Data entry resumed, but a new and more serious problem was being identified.

Stockpiles

When the SoilTrak application was designed, some discussions centered on the likelihood of soil “mixing”. The Soil Reuse Plan that was driving the construction stipulated that soil “blending” (mixing) was not permitted. Clean soil could not be mixed with contaminated to make it “less” contaminated or to change it’s engineering characteristics. It was decided, however, that mixing of contaminated soil from different cells was possible. Thus, a mixing module was added to SoilTrak to address the anticipated cases where soil from varying cells had been temporarily stockpiled together. This type of activity would have no effect on the cost recovery objective of SoilTrak, but would have a profound impact on the tracking of contaminants.

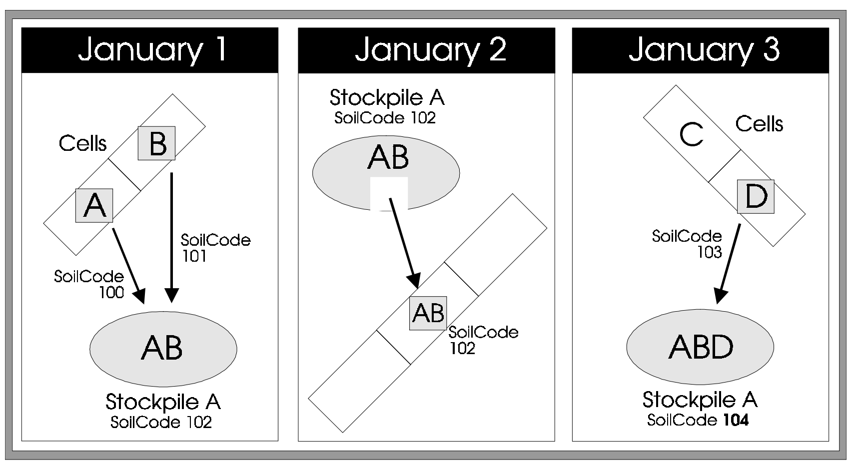


Figure 4

The new Soil Code created for Stockpile A was required to prevent Contaminant D from showing up in the fill activity of January 2nd.

With contaminant tracking in mind, a model had to be developed that could accurately account for the addition of “new” contaminants in a stockpile, while assuring that any new contaminants did not show up in fill records occurring prior to addition of the new contaminants. Recall that SoilTrak tracked “slugs” of soil removed from the ground via “Soil Code” numbers. It was determined that a mix could simply be a new Soil Code to which the comprising soil codes were attached. In this way, SoilTrak could identify a record as a “mix” record, look into the source soil code field of each mix record and identify which contaminants were present in the source area, and thus, the mix. Filling would be done using the mix soil code. However, what if soil containing new contaminants is added to an existing soil mix/stockpile? If SoilTrak used the mix’s soil code to identify contaminants, then early fills would be identified as having contaminants that weren’t part of the stockpile until after the filling occurred. The solution to this was to create a new soil code every time a new soil was added to an existing mix. The new soil code would always represent the current suite of contaminants.

This model proved effective in the first phase where nearly all of the mixing was “fake” mixing required for tracking costs. During the second phase, stockpiling (mixing) became common place. Three serious issues began to appear, the combination of which would prove a fatal flaw to SoilTrak’s DER contaminant tracking objective.

Multiple Additions to Stockpiles

During SoilTrak’s programming, soil mixing was not a major consideration since it was not anticipated to occur. The above model had been anticipated but it was later discovered that the code could not backtrack more than one mix at the reporting end. The data entry module could support entry of these records, however, so data entry continued. This flaw in the code could be addressed at a later time.

Stockpile Nomenclature

Mixing of soil from multiple cells into stockpiles became a common practice of Contractor #2. Since the mixing was occurring with contaminated soils from within the corridor, no rules were being broken. However, the degree to which stockpiling was occurring was unanticipated, so no agreements were reached regarding stockpile naming. Reviews of soil movement reporting sheets coming from the field were found to have stockpile names not known to those doing data entry. Upon further investigation, it was discovered that stockpiles often had more than one “name”, one had at least five synonyms.

Upon discovery of this issue, meetings were arranged between BEM, the Contractor and their field crews to develop a consistent nomenclature system for the stockpiles. Each stockpile would receive a name that would include the construction area in which it was located, followed by a sequential number identifying the order in which the stockpile had been created. This approach worked, but considerable damage had been done to the data requiring a substantial effort to correct. Most of the corrections had to be done directly to the data since the SoilTrak application could not edit Soil Code numbers. Since Soil Codes are program assigned and represent a key field, SoilTrak application users were not permitted to edit these numbers.

Combining Stockpiles

Frequent additions of soil to an existing stockpile created a large number of new soil codes. Tracking all of these was proving to be a data entry nightmare for the SoilTrak operators. Extreme care always had to be taken when selecting the stockpile (Soil Code) to mix new soil into, since each stockpile was represented by many soil codes. While difficult, however, data entry through SoilTrak was still possible. One more twist was discovered which would break the SoilTrak application; Contractor #2 was mixing stockpiles into other stockpiles. By this point the new soil code - source soil code method of tracking contaminants broke down. Stockpiles could now have multiple layers of mixing which would be very difficult to backtrack to original cut soil codes. SoilTrak also had no means of allowing data entry of these kinds of mixes. Data entry was forced to continue directly into the SoilTrak tables. A major code re-write would be required to support the combining of stockpiles. Since nearly all of the soil filling during this second phase of construction originated from stockpiles, it would be some time before the SoilTrak application could meet its soil contaminant-reporting requirement.

Lessons Learned

In the spring of 1999, the SoilTrak process and application were re-evaluated. The first most important question was whether or not SoilTrak met its two objectives, cost recovery support and contaminant tracking. The cost recovery support objective was cleanly met. Accurate reports could be generated for any tax lot, property owner or cut cell. A default value for cost per cubic yard of soil excavation could be used or calculated from line item actual costs for a cell. Many difficulties and issues arose during the initial use of SoilTrak, but none of these affected the cost recovery tracking or reporting. Periodic mapping of soil movements involving both cutting and filling was greatly simplified by SoilTrak's ability to generate the attribute tables needed to link-up with BEM's GIS. SoilTrak successfully tabulated this data for reporting as well.

From the contaminant tracking and reporting perspective, SoilTrak fell significantly short of its goal. The inability to handle complex soil mixing associated with stockpiling crippled both the data entry and reporting capabilities. This wasn't the only shortcoming of SoilTrak however.

During the evaluation, and subsequent re-design of SoilTrak, many issues were addressed. Each is discussed below.

Preparing Data for Data Entry was too Complex

SoilTrak's user interface was not difficult to use; the failing was in the amount of preparation needed to start entering data. In order to enter data, the user first had to know which properties a cell consisted of. Using the GIS, a spreadsheet was supplied that listed the properties in a cell and the area multipliers needed to assign soil excavation volumes to owners. Still, this required the user to take a field entry of "Cut, 500 yards from cell C-15" and divide it into as many entries as there were tax lots within cell C-15. Every cut typically needed a fake mix from which fill records could later be completed. A cell having only three tax lot fragments in it required six record entries for each cut from that cell. While programming could have eliminated much of the data entry problem, the database model still required that this many records be written to the database.

Keeping track of stockpile Soil Codes proved to be nearly impossible for both operators and SoilTrak. The SoilTrak database was re-evaluated and a more fundamental issue was about to be identified.

Tracking Soil Codes Created Needless Complexity

The initial idea was to track each "slug" of soil, each discreet cutting activity, from cradle to grave. However, the smallest geographic unit being reported was the cell. Without the ability to map precisely where within a cell a "slug" of soil came from, no useful information was gained by tracking soil at this level. What then was the level of detail that should be tracked? It became evident that what was really being tracked was cells, not soil. All activity occurred at that level, soil was either removed from, or placed within, a cell. Stockpile tracking was also complicated by soil codes. Every time a soil was added to a stockpile, a new Soil Code was created. This was true even if soil were added from a cell that had been previously added.

It was decided that SoilTrak's database should be redesigned to track soil movement by cells rather than Soil Codes. Soil would either go into or out of a cell. Stockpiles would be treated as cells into which soil could be added or removed at any time. Dates, rather than Soil Codes, would identify which contaminants were present in fills made from stockpiles. By dealing with cells only, no distinctions had to be made as to which lots were in a cell during data entry eliminating the need for multiple record entries. Fake mixes were also eliminated.

Another benefit of this approach was that by treating stockpiles as cells, a separate “soilmix” table was no longer required. This paradigm shift promised to greatly simplify the data entry and programming tasks.

Handle Cell-Property Relationships at the Back-End

Part of SoilTrak’s complexity-of-use centered on relating cells to properties at data entry time. Moving to a cell-based tracking model changed the need for this. These relationships could be derived at report-time when the information was needed, rather than stored record-by-record. This would reduce the number of records in the database by 6:1 for cells that had three lots within them. Also, if real-time embedded GIS were used to determine these relations at run-time, any updates to the cell or property coverages would be immediately applied to any reports. BEM had already had success developing embedded GIS applications using Microsoft Visual Basic and ESRI’s Map Objects LT™. At a minimum, the GIS could be used to re-create a table of the property-cell relationships that could then be stored as a look-up table in the new SoilTrak database.

Another advantage would be the ability to enter data into a new SoilTrak without the need for completed cell or property mapping. This could allow data entry at the outset of the project, assisting in early identification of any field procedural issues.

Identify Contaminated Cells at the Back-End

SoilTrak’s fundamental goal was to track reuse of contaminated soil. In order to achieve this objective, determinations had to be made up-front regarding which cells were, and were not contaminated. The field sheets used by the contractors had a check box to identify which cuts occurred in contaminated cells, allowing them to report all soil movement easily. But it was later discovered that Contractor #2 was using a different means of determining which cells were contaminated than the BEM staff. Then, during the re-evaluation of SoilTrak, questions were raised about whether the method we were using was too conservative and perhaps indefensible. Such questions at the near end of the data entry process were devastating. Since only contaminated soil movements were supposed to be present in SoilTrak, any change in which cells were designated as contaminated could force major editing of the data and entry of new data.

In order to prevent this chaos in the future, it was decided that SoilTrak should store *all* soil movement without regard to classifications of "clean" or contaminated. Determination of which cells were contaminated could then be done at report time. This would also allow for multiple methods of making this determination and the ability to change the method at-will, with no changes in the data. Another advantage of this approach was the elimination of the contractor as a factor in reporting clean versus contaminated soil reuse.

Providing users with this flexibility would require additional data tables and complicate the programming task, but the cost was deemed well worth it.

Handle Depth to Contaminated Soil at the Back-End

One minor failing of the SoilTrak application was its inability to report the starting and ending elevations of contaminated soil filling for a cell. Although the database was designed to handle this information, it was requested from the user for each filling activity. Rarely was this information available from the field. More typically, construction firms will survey the beginning elevation, complete all filling activities and then re-survey. The paradigm shift toward cell-based tracking resulted in a new data table that stored the current net cut-fill volumes of each cell. This created a convenient place to store the starting and final elevations of the filled soils by cell. This would also require only one data entry effort for each cell and could be done when final elevation data was available.

Agree on all Nomenclature Up-Front

BEM and both contractors had agreed at the outset of each phase what name each cell would have. In Phase 2, stockpiles became a factor but were named in a haphazard fashion. Often times multiple names were associated with the same stockpile. The discovery of this problem resulted in substantial data editing. In a new, cell-based model where even stockpiles are considered a cell, consistent nomenclature would be critical. Any naming schemes would be decided by all parties up-front.

Summary

The SoilTrak method and application were designed to track and report on contaminated soil reuse. It used a database and application developed in Microsoft FoxPro and depended heavily upon data provided by GIS coverages managed in ESRI's ArcCAD GIS. SoilTrak had two primary objectives. One was to track and report on costs associated with this cost-effective remedial option for the purposes of recovering these costs from property owners. The second was to report on the contaminants present in contaminated soils used as fill throughout the corridor. This would satisfy the state environmental agency's requirements for the Declaration of Environmental Restrictions that was a foundation of the soil reuse option.

In the initial phase of construction, a large Park & Ride facility, the SoilTrak method and application were able to meet these objectives. During the second phase, the southern half of the rail corridor, the complexity associated with soil mixing via stockpiling proved to be too great for the SoilTrak application to manage, and the contaminant tracking objective could no longer be met. Cost tracking and cost recovery objectives were unaffected. A later addition to the SoilTrak application supported regular mapping and reporting of both cutting and filling activities.

An evaluation of SoilTrak's effectiveness showed that too much data preparation was required prior to data entry. It was also clear that SoilTrak had failed its contaminant-tracking objective. An examination of SoilTrak's data model showed that the tracking of individual "slugs" of soil via Soil Codes was complicating data management and yielding no benefits. The paradigm was changed to track activity by cell instead. This solved many problems by simplifying both data entry and management as well as allowing for back-end determination of cell-property relationships and resulted in many fewer database records.

In order to provide greater flexibility and minimize confusion, the new paradigm also stated that all soil movements should be logged. Determination of which source (cut) cells were contaminated would be done at the back end, allowing for multiple methods of determination. Moving this determination to the back-end also eliminated any need to edit or add soil movement data if the determination method was changed.

With these changes, and better communication between all involved parties, SoilTrak can meet all of its objectives with significantly less labor and reduced chance of error. At this writing (June

1999) SoilTrak is still being examined and development of a more powerful version, including embedded GIS mapping, is scheduled to begin shortly.

Application of a Geographic Information System for Containment System Leak Detection

Randall R. Ross[1], Milovan S. Beljin[2] and Baxter E. Vieux[3]

ABSTRACT

The use of physical and hydraulic containment systems for the isolation of contaminated ground water associated with hazardous waste sites has increased during the last decade. Existing methodologies for monitoring and evaluating leakage from hazardous waste containment systems rely primarily on limited hydraulic head data. The number of hydraulic head monitoring points available at most sites employing physical containment systems may be insufficient to identify significant leakage. A general approach for evaluating the performance of containment systems based on estimations of apparent leakage rates is used to introduce a methodology for determining the number of monitoring points necessary to identify the hydraulic signature of leakage from a containment system. The probabilistic method is based on the principles of geometric probability. A raster-based GIS (IDRISI) was used to determine the critical dimensions of the hydraulic signature of leakage from a containment system, as simulated under a variety of hydrogeologic conditions using a three-dimensional ground-water flow model. MODRISI, a set of computer programs was used to integrate ground-water flow modeling results into the hydraulic signature assessment method.

INTRODUCTION

Subsurface vertical barriers have been used to control ground-water seepage in the construction industry for many years. Recently, the industrial and regulatory communities have applied vertical barrier containment technologies as supplemental or stand-alone remedial alternatives at hazardous waste sites to prevent or reduce the impact of contaminants on ground-water resources (Rumer and Ryan, 1995). While subsurface barriers appear to be useful for isolating long-term sources of ground-water contamination at many sites, the potential exists for leakage

[1]Hydrologist, U.S. EPA Nat. Risk Mgmt. Res. Lab, Subsurface Protection and Remediation Div., P.O. Box 1198, Ada, OK 74820
[2]Consulting Hydrologist, M.S. Beljin & Associates, 9416 Shadyoak Court, Cincinnati, OH 45231
[3]Assoc. Prof., School of Civil Eng. and Env. Sci., Univ. of Oklahoma, 202 West Boyd Street, Norman, OK 73019

of contaminants through relatively high hydraulic conductivity zones ("windows") within the barriers.

This paper describes the application of a Geographic Information System (GIS) as a tool to help identify leakage through discrete zones within a subsurface vertical barrier. The proposed techniques could be useful for evaluating existing containment systems by providing insight as to how many monitoring points are necessary to determine the approximate locations of discrete leaks, given specified confidence and constraints.

Containment Systems

Subsurface containment systems may be active (e.g., ground-water extraction to manage hydraulic gradient), or passive (e.g., physical barriers only) depending on the remedial objectives and complexity of the hydrogeologic setting (Canter and Knox, 1986). Frequently, containment systems employ a combination of active and passive components, which commonly incorporate vertical barriers keyed into underlying low-permeability units. Many containment systems also include a low permeability cover to reduce the rainfall infiltration, extraction and injection wells, and trenches for ground-water management.

Soil-bentonite slurry cutoff walls (slurry walls) are the most common type of subsurface vertical barriers used at hazardous waste sites and are generally installed around suspected source areas (U.S. EPA, 1984). Construction defects or post-construction property changes are potential failure mechanisms of subsurface vertical barriers (Evans, 1991). Construction defects may result in the formation of relatively high hydraulic conductivity windows in a barrier. Some of the mechanisms responsible for the formation of such windows include emplacement of improperly mixed backfill materials, sloughing or spalling of *in situ* soils from trench walls, and failure to excavate all *in situ* material when keying wall to the underlying low permeability unit (U.S.EPA, 1987). Post-construction property changes may result from wet-dry cycles due to water table fluctuations, freeze-thaw degradation, or chemical incompatibility between the slurry wall material and groundwater contaminants.

Monitoring of Containment Systems

The performance of hazardous waste containment systems has generally been evaluated based on construction specifications. Most subsurface vertical barriers are required to maintain a hydraulic conductivity of 1×10^{-7} cm/s, or less. The use of appropriate construction quality assurance (QA) and quality control (QC) testing during installation is essential to ensure that the design performance specifications are achieved. The regulatory community recognized the need to develop procedures to verify post-construction performance and identify unsatisfactory zones in containment systems (U.S.EPA, 1987). While construction dewatering systems are deemed successful if the barriers limit ground-water leakage to reasonably extracted quantities, there are no uniform methods to reliably measure and document the hydrologic performance of existing and proposed hazardous waste containment systems (Grube, 1992).

The minimum number of monitoring points necessary to determine whether a containment system is functioning as designed depends on site-specific conditions. For example, in some cases it may be possible to determine whether leakage has occurred by analyzing the water level trends in monitoring wells (Ross and Beljin, 1998). Subtle variations in the hydraulic head distribution associated with leakage through a subsurface barrier may be identifiable if sufficient hydraulic head data are available for analysis. Such an undertaking would generally be considered prohibitively expensive due to the high cost of installing a piezometer network capable of adequately defining the hydraulic head distribution. However, the recent development of relatively inexpensive installation techniques may make it feasible to install a sufficient number of small diameter piezometers to identify the hydraulic signatures associated with containment system leakage.

A New Monitoring Method

The process of locating a leak in a hazardous waste containment system can be analogous to mineralogical prospecting where a compromise is sought between the cost of exploration and the thoroughness of the search. For mineral exploration applications, the expected benefit of a search is the sum of the value of each target multiplied by the probability of finding it, assuming that the target exists in the search area (Singer, 1972). For containment system leak detection, the expected benefit of a search is the potential reduction in risk to human health and the environment associated with the detection and abatement of significant leaks.

Gilbert (1987) presents a methodology based on the work of Savinskii (1965), Singer and Wickman (1969), and Singer (1972) that can be used to determine the grid spacing required to detect highly contaminated local areas or hot spots at a given level of confidence, or estimate the probability of finding a hot spot of specified dimensions, given a specified grid spacing. Given a specific grid spacing, the probability of detecting a target is determined by the method of geometric probability, which is a function of the ratio of the area of the target to the area of the grid cell. The method assumes that the highly contaminated areas are circular or elliptical in shape, the boundaries of the hot spot are clearly identifiable based on contamination levels, hot spot orientation is random with respect to the sampling grid, and the distance between grid points is much larger than the area sampled. In order to address variations in the distribution of hydraulic head, rather than contaminant concentrations, the assumptions were modified for the methodology presented in the paper.

METHODOLOGY

The hydraulic signature associated with leakage from a containment system is simulated using a numerical model for a variety of hydrogeological settings. The modeling results provide the data on which the hydraulic signature assessment method is demonstrated. A set of computer programs was developed (Ross and Beljin, 1995) to import modeling data into a raster-based GIS, for further processing. The GIS was used to generate the input data for the ground-water model.

Ground-Water Modeling

A model may be defined as a simplified version of a real system that approximates the stimulus-response relationships of that system (Bear and others, 1992). By definition, the use of a model requires the application of simplifying assumptions to describe the pertinent features, conditions, and significant processes that control how the system reacts to stimuli. In this study, one of the primary objectives of the modeling was to predict the hydraulic head distribution associated with leakage through discrete leaks in a vertical barrier under different hydrogeologic conditions.

The conceptual model presented in this paper is based on characteristics of several specific hazardous waste sites that incorporate physical containment as a major component of the remedy. The sites which influenced the development of the model used in this study include the Gilson Road Superfund site (Nashua, New Hampshire), the G.E. Superfund site (Moreau, New York), and the Velsicol/Michigan Chemical Company Superfund site (St. Louis, Michigan). The conceptual model for the containment system consists of a slurry wall fully penetrating an unconsolidated surficial aquifer, keyed in to an underlying low permeability aquitard (Figure 1).

Hydraulic head values are assumed to be higher in the interior of the containment system, simulating a worst-case scenario for potential contaminant losses from the system (Figure 1).

The elevated water levels within the conceptual containment system are assumed to be derived from deficiencies in the system (i.e., leakage under or through the upgradient wall and infiltration through the cap), and water levels are assumed to be relatively stable over time. Ground-water flow is assumed to be horizontal, except in the immediate vicinity of the vertical barrier. Given the long-term nature of most hazardous waste containment systems, the hydraulic heads are averaged over long time periods. Consequently, steady-state flow conditions are assumed for all simulations used in this study.

The hydraulic head distribution associated with a linear segment of a conceptual vertical barrier was simulated using Visual MODFLOW[7] (Guiger and Franz, 1995), a commercial version of the three-dimensional, finite difference ground-water flow model MODFLOW, developed by the U.S. Geological Survey (McDonald and Harbuagh, 1988).

<u>Data Processing with a GIS</u>

The hydraulic head data generated by the numerical simulations are extracted, visualized, sampled, analyzed, and appropriately manipulated using several software packages. Hydraulic head data from a vertical cross-section parallel to, and immediately down gradient from the simulated vertical barrier are used throughout this study. The data are extracted from MODFLOW output files and reformatted as image files for analysis using MODRISI (Ross and Beljin, 1995). The GIS software used in this study is IDRISI (Eastman, 1995), a raster GIS that provides numerous analytical capabilities that are directly applicable to this, and other hydrogeologic studies. The uniform grid spacing facilitates the transfer of data from one software package to another. The raster format allows import and export of uniform grid model data and also provides a robust platform for the analysis, visualization and data manipulation.

<u>Model Setup</u>

The model domain consists of 51 rows, 51 columns, and 25 layers (Figure 2) and is discretized into uniform 1 m^3 blocks. This configuration is sufficiently large to reduce boundary effects and provides sufficient resolution to allow identification of subtle variations in hydraulic heads

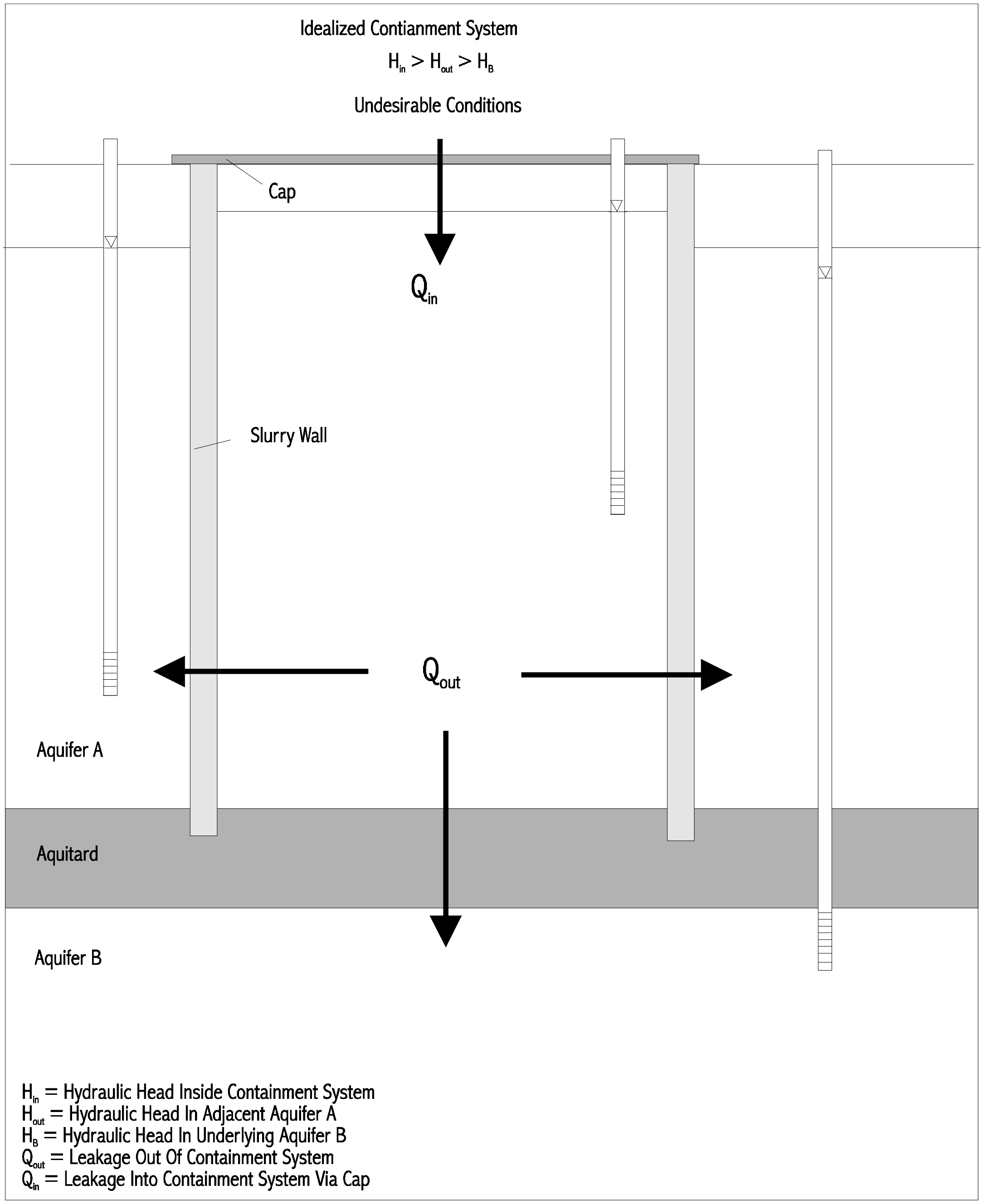


Figure 1. Major components of an idealized hazardous waste containment system exhibiting unfavorable conditions (e.g., outward hydraulic gradient).

associated with leakage through a vertical barrier. The uniform grid size allows consistent precision over the entire model domain and simplifies data management and transfer between software packages.

he slurry wall is simulated as a one-meter thick barrier with uniform properties, except for the window. The hydraulic conductivity values for the aquifer and window are scenario dependent. Leakage through the wall is simulated as a window with dimensions of 2 x 3 cells (6 m^2), located in the approximate center of the vertical barrier (row 25, columns 24-26, layers 12 and 13).

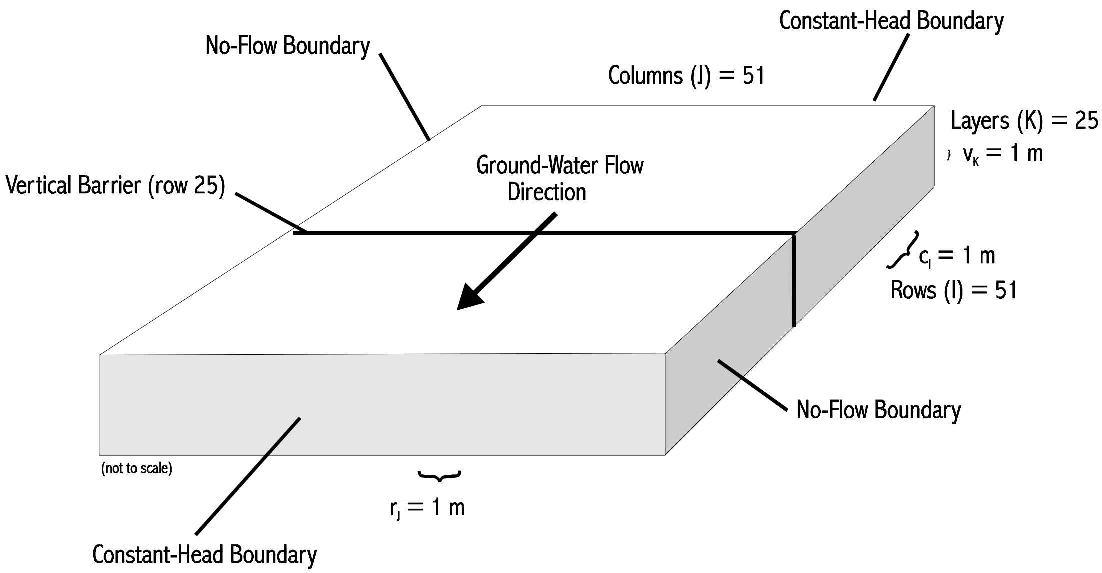


Figure 2. Conceptual model domain and boundary conditions.

Boundary conditions are depicted in Figure 2. The upgradient and downgradient sides of the model are constant-head boundaries, resulting in a horizontal hydraulic gradient across the model domain of 0.0196 m/m. This value falls within the range of hydraulic gradients commonly observed in the field. The sides and lower surface of the model oriented parallel to ground-water flow are simulated as no-flow boundaries.

The applicability of the numerical model for simulating the hydraulic head distribution associated with leakage from a containment system was demonstrated by comparing model results to data generated from a laboratory bench scale model of a cutoff wall (Ling, 1995). Simulation results agreed favorably with the physical model results, indicating that the approach described in this study is appropriate for simulating the hydraulic head distribution associated with leaking vertical barriers.

General Simulation Scenarios

Several hypothetical hydrogeologic conditions are evaluated in this study. Different scenarios are used to better understand the potential variability of the hydraulic signatures associated with different subsurface conditions and to account for potential uncertainties associated with predictive modeling.

A range of homogeneous and isotropic conditions were simulated in an effort to provide a reference case for evaluating the effects of varying average aquifer hydraulic conductivity values on the hydraulic signature of a simulated leak. The scenarios spanned a wide range of hydraulic conductivity values with respect to the aquifer material and zone of leakage. The hydraulic conductivity values for the aquifer range from 1 10^{-2} cm/s to 1 10^{-5} cm/s. The hydraulic conductivity of the vertical barrier is maintained throughout the study at 1 10^{-7} cm/s. The hydraulic conductivity values for the window ranged from 1 10^{-2} cm/s to 1 10^{-5} cm/s. The hydraulic conductivity value for the window is assumed to be less than or equal to that of the adjacent aquifer materials. The scenarios simulate the general effects of layering by varying the horizontal to vertical hydraulic conductivity ratios of aquifer materials.

One of the primary limitations of using ground-water flow models as a predictive tool results from the uncertainty associated with input parameters. This uncertainty is directly related to the spatial variability of hydrogeologic properties of the porous medium (i.e., aquifer material). To account for some of the spatial variability and uncertainties associated with three-dimensional predictive flow modeling, several scenarios utilizing heterogeneous distributions of hydraulic conductivity were assessed. The assumption of lognormally distributed hydraulic conductivity is used for the heterogeneous, isotropic and heterogeneous, anisotropic simulations. Unique lognormal hydraulic conductivity distributions were generated for each of the 25 layers using built-in functions of the GIS software. This approach resulted in the generation of approximately 63,000 hydraulic conductivity values within the model domain.

Hydraulic Signature Assessment Method

The methodology used to address the hydraulic head distribution associated with leakage from a containment system was developed based on the work of Singer and Wickman (1969) and

Gilbert (1987). The proposed method is directly applicable to determining the grid spacing necessary to detect the hydraulic signature associated with a discrete leak in a subsurface vertical barrier. The methodology requires the following assumptions:

- the hydraulic signature of the leak is circular or elliptical;
- hydraulic head data are acquired on a square grid;
- the criteria delineating the hydraulic signature are defined; and
- there are no measurement misclassification errors.

The model results indicate that the hydraulic signatures associated with the simulated leaks range in shape from approximately circular to elliptical when viewed in vertical cross-section. An increase in the anisotropy results in the elongation of the signatures in the horizontal directions. As expected, the greater the anisotropy, the more elliptical the hydraulic signature of the leak.

The criteria for delineating the hydraulic signature of a leak from background noise are based on the average hydraulic head value ($_h$) of the model cross-sectional surface. For this study, hydraulic head values of $_h$+0.05 m and $_h$+0.1 m were identified as critical values (C_v), indicating the presence of a hydraulic anomaly associated with containment system leakage. This follows the assumption that any background noise associated with the hydraulic head measurements is significantly less than 0.05 m. The dimensions of the hydraulic anomalies are determined using GIS software by image reclassification to delineate nodes exceeding the average hydraulic head by the specified critical values. The dimensions of the hydraulic signatures delineated by the two values for C_v are expressed as shape factors (S), defined as the ratio of the length short axis to the length of the long axis of the hydraulic signature. The shape factor for a circular feature is 1. An increase in anisotropy results in the elongation of the feature and a decrease in S, where $0 < S \leq 1$.

The probability tables of Singer and Wickman (1969), were used to determine the probability of not detecting a leak when a leak is present (β) to the ratio of the semi-major axis to grid size (L/G). The semi-major axis is defined as one half the length of the long axis of an elliptical feature. The general procedure for determining monitoring point spacing necessary to detect a hydraulic anomaly of given dimensions and specified confidence is outlined in Table 1, and in the following example.

Table 1. General steps for determining monitoring point grid spacing.

1.	Specify the radius or one half the length of the long semi-major axis (L) of the hydraulic signature (mound) associated with the leak;
2.	Assuming a circular hydraulic signature, let the shape factor (S) equal one; for elliptical features, S may be calculated using equation (9);
3.	Specify the maximum acceptable probability (β) of not detecting the hydraulic feature (=0.1);
4.	Knowing L, S and assuming a value for β, determine L/G from Figure 4, and solve for G (minimum grid spacing required to detect the hydraulic anomaly associated with the leak, given the specified constraints).

In order to determine the minimum grid spacing necessary to identify a hydraulic feature of specified dimensions, an acceptable probability of not detecting the feature must be established. For this example, a value of β= 0.1 is assumed for a leak signature with dimensions of 5 m by 4 m, as delineated by C_v = 0.1 in Figure 3a. From Figure 4, a value of approximately 0.64 is indicated for the ratio of the length of the semi-major axis to grid size (L/G), given = 0.1 and S = 0.8. Therefore, solving for G using L = 2.5, it is determined that a minimum grid spacing of approximately 3.9 m is necessary to identify the specified feature with a 90% probability of success. The resulting grid spacing (G) may be used to determine the minimum number of block-centered monitoring points required to detect the feature for a specified area by dividing the total area by the area of one square grid (G^2).

The probability tables were also used to generate nomographs relating the probability of not detecting a leak (β) of specified dimensions (L), for different grid dimensions (G). Figure 5 illustrates this relationship for circular hydraulic signature (S = 1.0). The nomographs may be used to estimate the dimensions of the smallest hydraulic signature capable of being identified by a monitoring network of known dimensions within an acceptable level of confidence (β). For example, given a monitoring point spacing of 20 m, what is the smallest circular hydraulic anomaly that can be detected with 80% probability of success (β= 0.2). From Figure 5 it is noted that a circular feature with a radius of approximately 10.1 m can be detected with the specified probability and grid spacing. The probability of not detecting the anomaly will increase as the radius of the hydraulic signature decreases.

RESULTS AND DISCUSSION

The dimensions of the hydraulic signatures associated with leakage through a subsurface vertical barrier are a function of the hydrogeologic properties of the aquifer, vertical barrier, and zone of leakage. Assuming all other variables remain constant, the magnitude of the hydraulic signature diminishes significantly as the hydraulic conductivity of the window decreases (Figure 3). The hydraulic signature of leakage through the hydraulic conductivity window becomes less prominent as its value is reduced by one order of magnitude (Figure 3g). As the value is further reduced, the hydraulic signature becomes discernable only immediately adjacent to the window (Figure 3j). The decrease in hydraulic signature corresponds to a decrease in flux through the window, as the window hydraulic conductivity is reduced (Table 2).

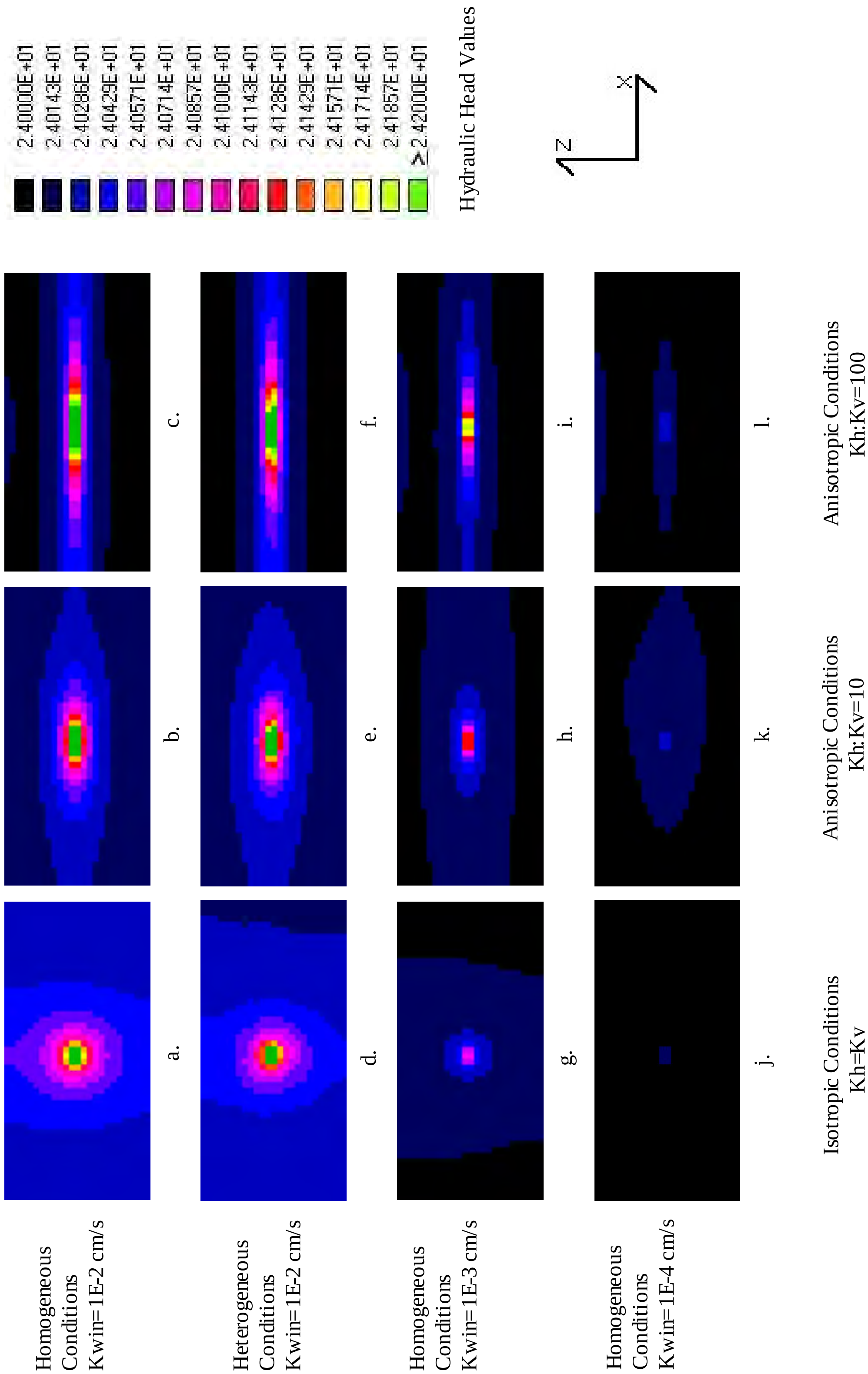


Figure 3. Vertical cross-section of model results illustrating hydraulic signature (head) variations due to changes in conceptual hydrogeological setting.

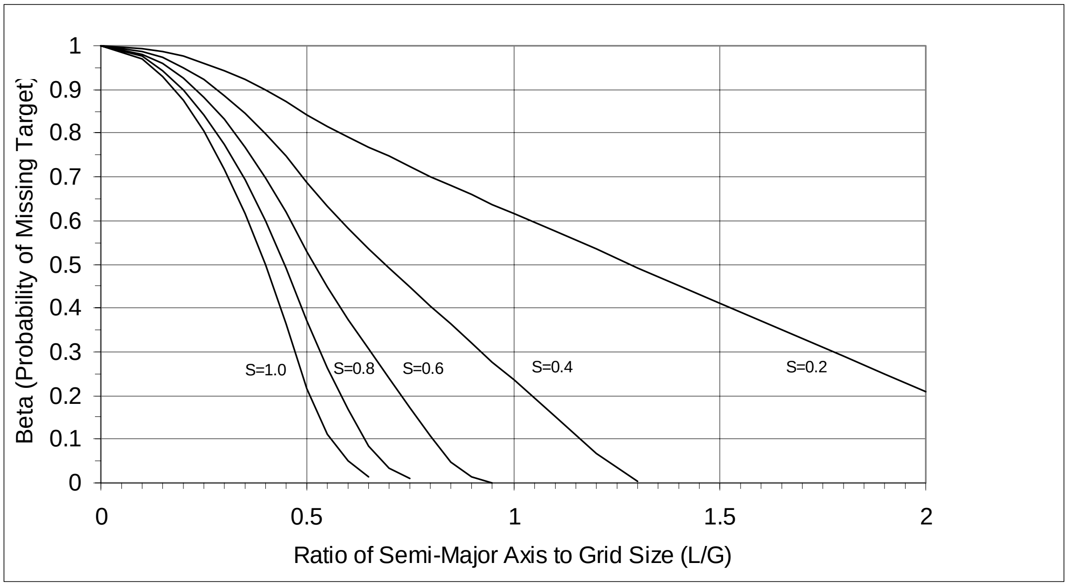


Figure 4. Nomograph relating ratio of semi-major axis of elliptical target and grid size to the probability of missing the target (Beta) for different shape factors using a square grid pattern.

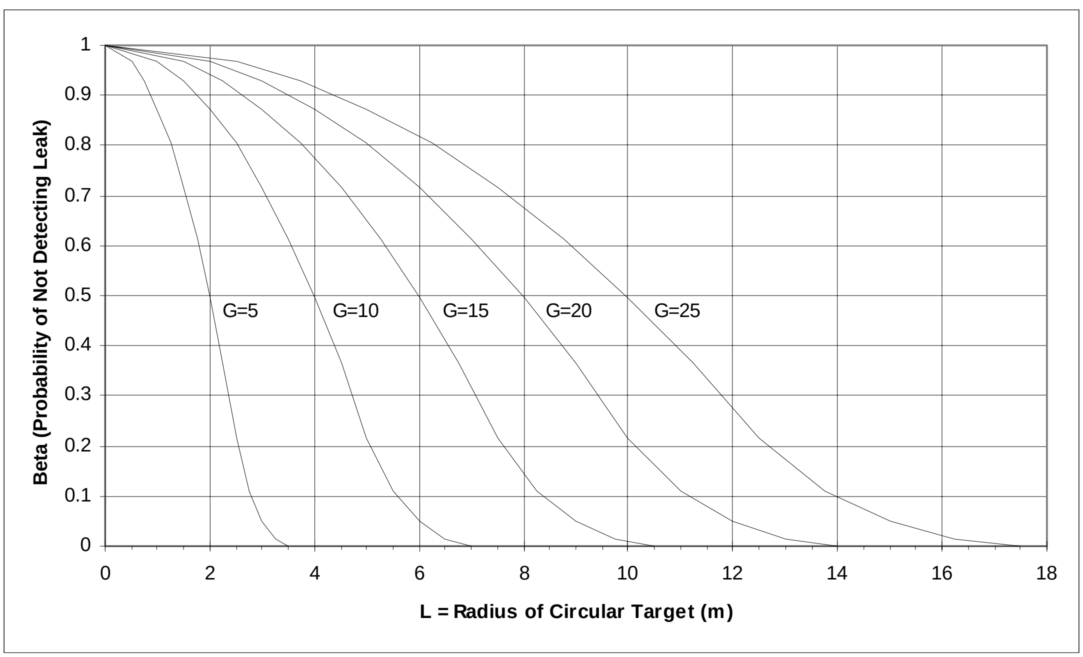


Figure 5. Nomograph relating radius of circular hydraulic signature to probability of not detecting leak (Beta) for different grid spacings.

Table 2. Simulated flux through windows of varying hydraulic conductivity.

Window Hydraulic Conductivity (cm/s)	Minimum Head Value (m)	Maximum Head Value (m)	Range (m)	Flux Through Window (m^3/d)
$1\ \ 10^{-2}$	24.0293	24.2627	0.2334	1.3110^{1}
$1\ \ 10^{-3}$	24.0117	24.0826	0.0709	3.98
$1\ \ 10^{-4}$	24.0071	24.0165	0.0094	4.9610^{-1}
$1\ \ 10^{-5}$	24.0063	24.008	0.0017	5.0910^{-2}

The effect of varying the horizontal to vertical hydraulic conductivity values is illustrated in Figure 3. For example, the hydraulic signature from leakage through a window under homogeneous and isotropic conditions forms an approximately circular feature (Figure 3a). However, as the horizontal to vertical hydraulic conductivity ratio increases, the hydraulic signature of the leak becomes more elliptical (Figure 3b,c). Similar trends are observed with respect to increasing the horizontal to vertical hydraulic conductivity ratio for the heterogeneous simulations (Figure 3d,e,f) and other homogeneous simulations with smaller hydraulic conductivity values for the windows (Figure 3g-l).

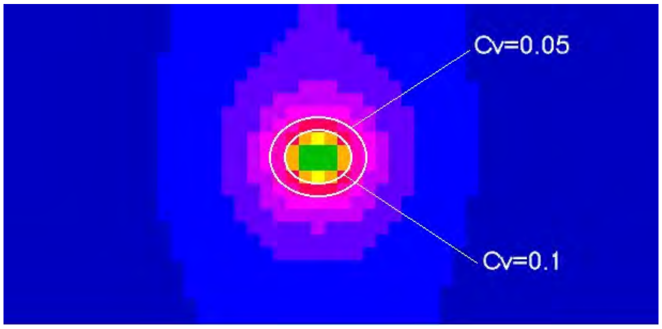
The method was applied to different hydraulic signatures developed from ground-water flow simulations of leakage through a vertical barrier. The criteria used to differentiate the hydraulic signature of leakage from background noise are $C_v = \ _h$+0.05 m and $_h$+0.1 m. Figure 6a depicts the head distribution associated with hydraulic signature of leakage through a window located in the approximate center of a vertical barrier in a homogeneous, isotropic aquifer. The approximate dimensions of the vertical hydraulic mound as defined by $C_v= \ _h$+0.05 and $_h$+0.1 are 7 m by 6 m, and 5 m by 4 m, respectively.

An increase in the anisotropy of the simulated aquifer by one order of magnitude produces a vertically compressed and horizontally elongated hydraulic signature (Figure 6b). Similarly, increasing the anisotropy of the simulated aquifer by two orders of magnitude results in even greater compression and elongation of the hydraulic signature in the vertical and horizontal directions, respectively (Figure 6c).

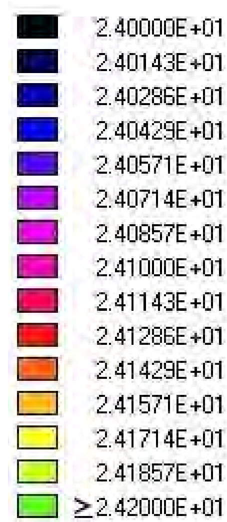
Hydraulic signatures for leakage through a window with a hydraulic conductivity value of 1 10^{-3} cm/s exhibit similar trends in response to increases in anisotropy (Figure 7a,b,c). However, the overall hydraulic signature of the window is decreased significantly relative to that of the base case. This results in a lack of head values greater than the elevation threshold for $C_v = {}_h+0.1$ for the homogeneous, isotropic simulations. The hydraulic head values associated with leakage through windows with hydraulic conductivities < 1 10^{-3} cm/s were all less than $C_v = {}_h+0.05$, and therefore, could not be evaluated as described above.

The grid sizes necessary to identify the hydraulic features described below with a 90% probability of success (β=0.1) were obtained using the nomograph in Figure 4. The number of sampling points (N_s) necessary to identify the hydraulic features within the domain of the model cross-section is determined by dividing the cross-sectional area of the model (1,275 m^2) by the area of one square grid spacing (G^2). The results are listed in Table 3.

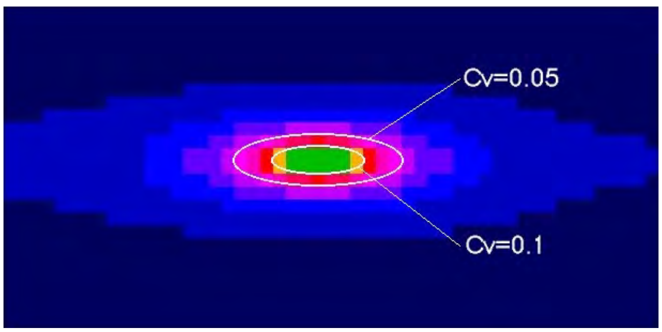
The number of monitoring points required to identify the hydraulic signatures of the simulated leaks using the prescribed constraints and confidence ranges from approximately 40 to over 300. The wide range of values is a function of the variability in the size and shape of the hydraulic features. This variability results from the use of different critical values to define the hydraulic signatures of the leaks and the wide range of shape factors resulting from the three orders of magnitude range of the anisotropy values.



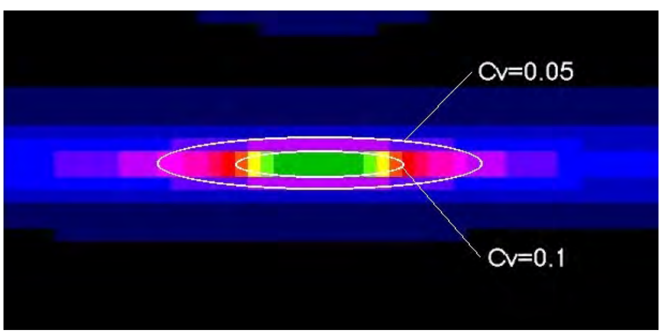
a. Homogeneous and Isotropic Simulation Results

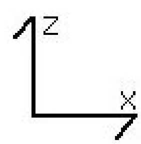


Hydraulic Head Values



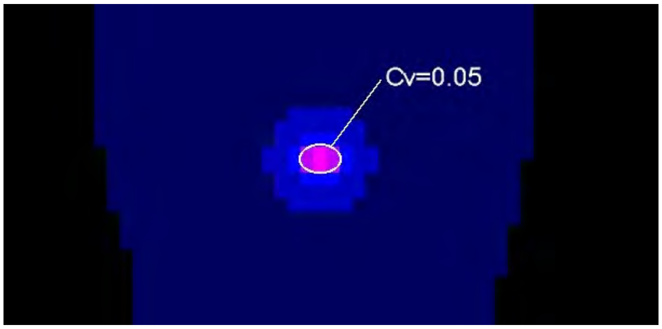
b. Homogeneous and Anisotropic Results (K_h:K_v=10)



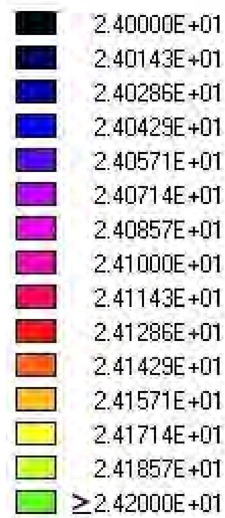


c. Homogeneous and Anisotropic Results (K_h:K_v=100)

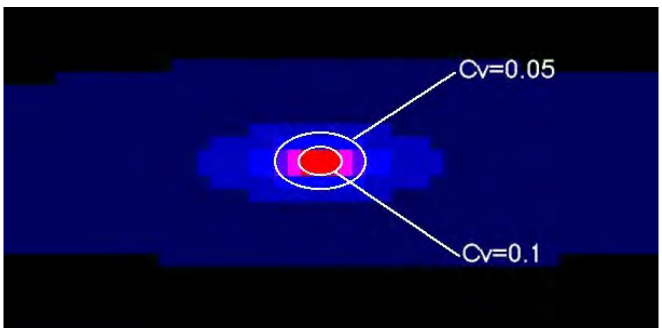
Figure 6. Vertical cross-section of model results illustrating variations in hydraulic head values due to changes in anisotropy (K_{aq}=1 10^{-2} cm/s, K_{win}=1 10^{-2} cm/s). The ellipses define the approximate boundaries of the hydraulic features defined by specified critical values (C_v= +0.1 and +0.05).



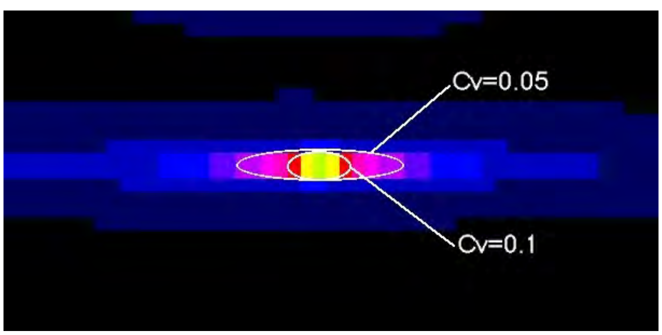
a. Homogeneous and Isotropic Simulation Results

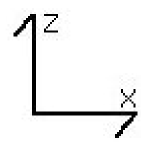


Hydraulic Head Values



b. Homogeneous and Anisotropic Results (K_h:K_v=10)





c. Homogeneous and Anisotropic Results (K_h:K_v=100)

Figure 7. Vertical cross-section of model results illustrating variations in hydraulic head values due to changes in anisotropy (K_{aq}=1 10^{-2} cm/s, K_{win}=1 10^{-3} cm/s). The ellipses define the approximate boundaries of the hydraulic features defined by specified critical values (C_v= +0.1 and +0.05).

Table 3. Parameters and Results Obtained from Hydraulic Assessment Method.

Kwin (cm/s)	Kh:Kv	C_v	S	L	L/G	G	N_s
1 10^{-2}	1	0.1	0.8	2.5	0.64	3.91	84
1 10^{-2}	1	0.05	0.85	3.5	0.62	5.65	40
1 10^{-2}	10	0.1	0.28	3.5	1.64	2.13	280
1 10^{-2}	10	0.05	0.31	6.5	1.51	4.3	69
1 10^{-2}	100	0.1	0.13	7.5	3.5	2.14	278
1 10^{-2}	100	0.05	0.16	12.5	2.9	4.3	69
1 10^{-3}	1	0.1	BCL	-	-	-	-
1 10^{-3}	1	0.05	0.67	1.5	0.74	2.03	311
1 10^{-3}	10	0.1	0.67	1.5	0.74	2.03	311
1 10^{-3}	10	0.05	0.4	2.5	1.17	2.14	280
1 10^{-3}	100	0.1	0.4	2.5	1.17	2.14	280
1 10^{-3}	100	0.05	0.15	6.5	3.05	2.13	281
1 10^{-2}*	1	0.1	0.8	2.5	0.64	3.91	84
1 10^{-2}*	1	0.05	0.85	3.5	0.62	5.65	40
1 10^{-2}*	10	0.1	0.28	3.5	1.64	2.13	280
1 10^{-2}*	10	0.01	0.31	6.5	1.51	4.3	69
1 10^{-2}*	100	0.1	0.13	7.5	3.5	2.14	278
1 10^{-2}*	100	0.05	0.16	12.5	2.9	4.31	69

BCL = All head values below critical value threshold.
*Heterogeneous simulations; all other simulations homogeneous

CONCLUSIONS

Numerical modeling of ground-water flow through high hydraulic conductivity windows in subsurface vertical barriers was conducted to provide data sets for use with a probabilistic method for determining the grid spacing necessary to identify the hydraulic signature associated with the leaks. The proposed method of combined ground-water modeling and GIS represents a potential tool that may be used by the regulatory community and others to evaluate the adequacy of existing and proposed hazardous waste containment systems for identifying containment system leakage. The utility of the proposed method is demonstrated using simulated data. Based on the application of the method presented, the following conclusions were made:

- The number of points necessary to identify the hydraulic signature of a discrete leak within prescribed constraints is a function of the criteria used to delineate the feature;
- The hydraulic signature associated with a minor leak in a vertical barrier may be difficult to detect with a realistic number of monitoring points;
- By using the nomographs described above, the probability of failing to detect the hydraulic signature of a leak can be estimated for a given monitoring well spacing and specified confidence;
- The dimensions of the smallest hydraulic signature detectable with a given monitoring point spacing can be estimated, given the appropriate constraints and specified confidence;
- The monitoring point spacing used at many hazardous waste sites is likely inadequate to detect the hydraulic signatures of all but the largest leaks; and
- The method for delineating the hydraulic signature of a leak using the average hydraulic head plus specified values does not appear to be as sensitive to the heterogeneity of the aquifer as it is to anisotropy.

DISCLAIMER

REFERENCES

Bear, J., Beljin, M.S., and Ross, R., (1992). *Fundamental of Ground-Water Modeling, Ground Water Issue*, U.S. Environmental Protection Agency, Office of Research and Development, R.S. Kerr Environmental Research Laboratory, Ada, OK, EPA/540/S-92/005.

Canter, L.W., and Knox, R.C. (1986). *Ground Water Pollution Control*, Lewis Publishers, Boca Raton, FL.

D'Appolonia, D.J. (1980). A soil-bentonite slurry trench cutoff, Journal of Geotechnical Engineering, ASCE, Vol. 106, no. 4, pp.399-417.

Eastman, J.R. (1995). *IDRISI for Windows: User's Guide*, Version 1.0, Clark University, Worchester, MA.

Evans, J.C. (1991). Geotechnics of Hazardous Waste Control Systems, Chapter 20, Foundation Engineering Handbook, 2nd ed., H.Y. Fang, ed., Van Nostrad-Reinhold Company, New York.

Gilbert, R.O. (1987). *Statistical Methods for Environmental Pollution Monitoring*, Van Nostrand Reinhold, New York, N.Y.

Grube, W.E., Jr. (1992). ASlurry Trench Cut-Off Walls for Environmental Pollution Control, Slurry Walls: Design, Construction and Quality Control. ASTM STP 1129, David B. Paul, Richard R. Davidson, and Nicholas J. Cavalli, Eds., American Society for Testing and Materials, Philadelphia.

Guiger, N., and Franz, T. (1995). AVISUAL MODFLOW, The Integrated Modeling Environment for MODFLOW and MODPATH, Version 1.1, Waterloo Hydrogeologic, Ontario, Canada.

Ling, K., (1995b). *Windows Development and Detection in Soil-Bentonite Cutoff Walls*, Ph.D. Dissertation, University of Cincinnati.

McDonald, M.G., and Harbuagh, A.W. (1988). AA Modular Three-Dimensional Finite-Difference Ground-Water Flow Model (MODFLOW), U.S. Geological Survey Techniques of Water-Resources Investigation, Book 6, Chapter A1.

Ross, R.R., and Beljin, M.S. (1995). AMODRISI: A PC Approach to GIS and Ground-Water Modeling, Proceedings, National Conference on Environmental Problem-Solving with Geographic Information Systems, Cincinnati, OH, September 21-23, 1994. EPA/625/R-95/004.

Ross, R.R., and Beljin, M.S. (1998). An Evaluation of Containment Systems Using Hydraulic Head Data, J. Envir. Eng., 124(6), 575-578.

Rumer, R.R., and Ryan, M.E., eds. (1995). *Barrier Containment Technologies For Environmental Remediation Applications*, John Wiley & Sons, Inc., New York.

Savinskii, J.D. (1965). A Probability Tables for Locating Elliptical Underground Masses with a Rectangular Grid, Consultants Bureau, New York, pp. 110

Singer, D.A., (1972). AELIPGRID, a Fortran IV program for calculating the probability of success in locating elliptical targets with square, rectangular and hexagonal grids, Geocon Programs, Vol. 4, No. 1, p.1-16.

Singer, D.A., and Wickman, F.E. (1969). Probability Tables for Locating Elliptical Targets with Square, Rectangular and Hexagonal Point-Nets, Mineral Sciences Experiment Station Special Publication 1-69, Penn. State University, University Park Pennsylvania, 100 p.

U.S. EPA (1987). A Construction Quality Control and Post-Construction Performance Verification for the Gilson Road Hazardous Waste Site Cutoff Wall, Hazardous Waste Engineering Research Laboratory, Office of Research and Development, Cincinnati, OH, EPA/600/2-87/065

MODRISI: A PC Approach to GIS and Ground-Water Modeling

Randall R. Ross
Robert S. Kerr Environmental Research Laboratory, U.S. Environmental Protection Agency, Ada, Oklahoma

Milovan S. Beljin
University of Cincinnati, Cincinnati, Ohio

Abstract

It is widely accepted that ground-water contamination problems cannot be adequately defined or addressed until the governing physical, chemical, and biological processes affecting the transport and fate of contaminants are adequately characterized. Recent research has led to a better understanding of these complex processes and their effect on the movement of contaminants in the subsurface. The compilation and application of such information has yet to be accomplished at many hazardous waste sites, however. Too often, copious quantities of data are collected, only to be stored, ignored, or misplaced, rather than used for problem-solving. Geographic information systems (GIS) are computer-based tools that are relatively new to many environmental professionals. GIS allows the manipulation, analysis, interpretation, and visualization of spatially related data (e.g., hydraulic head, ground-water velocity, and contaminant concentration). GIS is more than a cartographic utility program, however. The analytical capabilities of GIS allow users to display, overlay, merge, and identify spatial data, thereby providing the basis for effective environmental decision-making.

IDRISI is a widely used PC-based raster GIS system that provides numerous analytical capabilities that are directly applicable to hydrogeologic studies. Raster systems are particularly well suited for analysis of continuous data such as elevation (e.g., water table, land and bedrock surfaces), precipitation, recharge, or contaminant concentrations and may be readily integrated with finite-difference ground-water models. Because the formats for IDRISI and ground-water model input data sets are different, a need exists for a program to integrate these two types of robust tools.

MODRISI is a collection of utility programs that allows easy manipulation and transfer of data files between IDRISI and ground-water models (e.g., MODFLOW, ASM, MOC). In addition, MODRISI integrates other widely used commercial and private domain software packages, such as SURFER, Geopack, GeoEas, AutoCAD, CorelDraw, and various spreadsheet programs. Two-dimensional arrays of models' input data sets can easily be created from IDRISI image files. AutoCAD vector files obtained by digitizing model boundaries, well locations, rivers and streams, or U.S. Geological Survey digital elevation model (DEM) files can also be translated into model input file formats. MODRISI can process model output files and prepare GIS image files that can be displayed and manipulated within IDRISI. Thus, MODRISI is more than a pre- and postprocessor for ground-water models; it is a complete GIS/ground-water modeling interface that is accessible to most ground-water hydrologists.

Introduction

Hydrogeologists collect and analyze large volumes of data during a ground-water modeling process. These data are stored and presented in many different forms such as maps, graphs, tables, computer databases, or spreadsheets. To most hydrogeologists, geographic information systems (GIS) are relatively new tools. They have been developed and applied in other natural and social science fields for over two decades, however, and can also be used in the ground-water modeling process.

GIS represents a new, powerful set of tools that can significantly improve the usefulness of results obtained during the ground-water modeling process. Bridging the disciplines of ground-water modeling, computer graphics, cartography, and data management, GIS represents a computer-based set of tools to display and analyze spatial data (e.g., water level elevations, ground-water quality data, modeling results, ground-water pollution potential). Efficient use of increasingly large volumes of data can be achieved only with powerful systems capable of acquiring information from a variety of sources, scales, and resolutions.

GIS can be defined as a computer-assisted system for the efficient acquisition, storage, retrieval, analysis, and representation of spatial data. Most GIS platforms consist of numerous subsystems that perform the listed tasks. The subsystems have the ability to query spatially related information and incorporate statistical analyses and modeling of relations and their temporal changes within the database. More than just a mapping system, GIS allows the user to analyze spatially related data and visualize results in either paper map form or graphically on screen. The data to be analyzed are a collection of spatial information represented by points, lines, and polygons and their associated attributes (characteristics of the features such as elevation or concentration). The cartographic tools of GIS allow the analyst to display, overlay, measure, merge, and identify the data to support a particular analysis. By allowing spatial data analysis and display, GIS provides the means necessary for effective environmental decision-making and implementation of environmental management plans.

GIS uses two basic map representation techniques: vector and raster. Vector representations describe features with a number of connected points. Raster representations subdivide a study area into a mesh of grid cells, each cell containing either a quantitative attribute value or feature identifier. Raster systems are well suited for analysis of continuous data (e.g., water level elevations, infiltration and recharge rates). This makes raster-based systems ideal for integration with ground-water models that use regularly spaced nodes. The objective of this paper is to illustrate such an integration of MODFLOW (1), a widely used U.S. Geological Survey (USGS) finite-difference ground-water flow model, and IDRISI (2), a raster-based GIS.

Previous Studies

To enhance understanding of a hydrogeologic system, and also to develop a credible ground-water model of the system, hydrogeologic features such as lithological logs, recharge and withdraw rates, estimates of spatial distribution of hydraulic conductivity, or specific storage can be plotted using GIS capabilities of data retrieval and overlay options to interactively define an area of interest (3-6). Two previous studies that combined GIS and ground-water modeling are briefly described.

Torak and McFadden (7) used GIS to facilitate finite-element modeling of ground-water flow. Complex aquifer geometry and irregularly distributed aquifer-system characteristics that influence ground-water flow affect the design of the finite-element mesh. GIS systems represent the complex arrangement of nodes and elements and the distribution of aquifer properties to provide input to the flow model. Point-data coverages of pertinent aquifer characteristics are rated from a relational database and are displayed using GIS.

Contoured surfaces based on point-data coverages produce triangulated irregular networks (TINs) that are superposed on the finite-element mesh to delineate zones of elements having similar aquifer properties. Zone boundaries are identified using the contoured TIN surface and by manually determining where boundaries align with the element sides. The allocation of well pumping rates to nodes in the finite-element mesh is performed efficiently with GIS for model input. Well pumping rates are accumulated by element from the combined coverages of the pumping data and the mesh, and element data are distributed to the node points for input. GIS is also used to prepare data for model input and to assess the adequacy of the data prior to simulation.

Three-dimensional perspectives showing TIN coverages of aquifer-property data are used to analyze and interpret complexities within the flow system before zonation. Additionally, GIS is used to display computed hydraulic heads over the finite-element mesh to produce contour maps of the simulated potentiometric surface. Because the node points in the finite-element mesh are not arranged in an orthogonal fashion, such as a finite-difference grid, a map display of the computed values of hydraulic head at the nodes is prepared for efficient and accurate interpretation of simulation results.

Harris et al. (8) conducted the Remedial Investigation/Feasibility Study (RI/FS) of the San Gabriel basin. Vast amounts of hydrogeologic data have been gathered, and a comprehensive systematized GIS database has been developed. The identified hydrologic boundaries, recharge basins, stream locations, well locations, and contaminant distributions are some of the features considered in developing a base map. The GIS-generated base map has allowed development of a finite-element grid for the basin. For each finite element, the initial estimates of the hydraulic conductivity, specific yield, recharge rates, and other input parameters were provided.

Using simple interfacing programs, the retrieval GIS nodal and elemental data were converted to required formats for the input files of the Couple Fluid, Energy, and Solute Transport (CFEST) code. Simulated groundwater levels were compared with the GIS-generated potentiometric surfaces. In areas of wide variations between simulated and observed data, the zonal distribution of controlling parameters was reevaluated, analyzed, and updated. Data processing, development of input files for computerized analysis of ground-water flow, and analysis of simulation results with different alternative conceptualizations is time consuming and tedious. Efficient use of GIS and CFEST not only eased the burden of conducting multiple simulations but reduced the probability of errors as well as the amount of time and effort required for each simulation.

IDRISI

IDRISI is a grid-based geographic information and image processing system developed by the Graduate School of Geography at Clark University and supported by the United Nations Institute for Training and Research (UNITAR) and the United Nations Environment Programme Global Resource Information Database (UNEP/GRID) (9). IDRISI is a collection of over 100 program modules that are linked through a menu system. These programs are organized into several groups:

- The core modules provide data entry and database management capabilities.
- The geographic analysis modules provide tools for database analysis.
- The statistical analysis modules allow statistical characterization of images.
- The peripheral modules provide a series of utilities.

IDRISI and other raster-based systems divide data sets into map layers; each layer contains data for a single attribute. For the example of a ground-water model, these layers could correspond to the MODFLOW two-dimensional arrays (e.g., initial water levels, transmissivity distribution, IBOUND arrays, computed hydraulic heads). IDRISI provides many analytical tools that are useful in hydrogeologic studies.

Three of the most important categories of these tools are database query, map algebra, and context operator. A semihypothetical case described below illustrates the use of these analytical tools. IDRISI provides an extensive set of tools for image processing, geographic and statistical analysis, spatial decision support, time series analysis, data display, and import/export and conversion. In addition, as a set of independent program modules linked to a broad set of simple data structures, the system is designed such that researchers may readily integrate into the system their own modules, written in any programming language.

IDRISI uses three types of data files: image, vector, and attribute. Image files contain rasterized information relating to a spatial variable. Vector files contain the coordinates of points, lines, and polygonal features. An attribute file lists the identifiers of features and the associated attribute values. Values files can be extracted from the existing image files, or image files can be created from existing values files. The values files can be combined and stored in a dBASE format. Each image, vector, or attribute file has a corresponding documentation file that contains information about the data file (e.g., title, number of rows and columns).

MODRISI: MODFLOW/IDRISI Interface

MODRISI is a set of utility programs that allows the transfer of data files between MODFLOW, IDRISI, Golden Software SURFER, GeoEas, and other software. Preparation of two-dimensional arrays for the MODFLOW input files is generally tedious and time consuming. The arrays can be created easily from the IDRISI image files, however. Thus, MODRISI serves as a preprocessor for MODFLOW. For example, when the values of a variable are available only for irregularly spaced points, interpolation routines in SURFER or GeoEas may be used to estimate the values of the variable on regularly spaced grid-nodes. MODRISI translates SURFER or GeoEas files into IDRISI image files for manipulation, analysis, and display. IDRISI recognizes either latitude and longitude geodetic coordinates or arbitrary Cartesian plane coordinates. IDRISI assigns the lower left grid-block of a raster image as a zero-row, zero-column block.

Vector files, such as model boundaries, well locations, and rivers, may be created within IDRISI and translated into a MODFLOW input file format. For example, the location of a river may be digitized on screen in IDRISI. The vector-to-raster function may be invoked, assigning all blocks through which the river passes as river nodes. Similarly, the positions of wells may be digitized and translated into the row-column positions and saved as a MODFLOW input file for the well package. Once the MODFLOW input files are prepared, MODFLOW simulations may be initiated. The MODFLOW hydraulic head output files may be read by MODRISI and modified to create IDRISI image files. Again, the image files may be displayed and evaluated within IDRISI. Thus, MODRISI is used as a postprocessor for MODFLOW.

Case Study

The utility of MODRISI was demonstrated at a hazardous waste site. Previous investigations provided site-specific information, including water level and bedrock and land surface elevations, which was then analyzed using GeoEas, a public domain geostatistical software program. These data were kriged to produce a grid of regularly spaced data. These data were imported into MODRISI and converted to IDRISI image files. IDRISI was used to visualize the surfaces that GeoEas generated.

Several prominent features are obvious upon inspection of the kriged bedrock topography (see Figure 1). A bedrock ridge trending northwest to southeast is flanked by a minor trough to the east and a major trough to the southwest. The outline of the site is visible on all figures. The kriged water level elevation map (see Figure 2) illustrates the general hydraulic gradient to the west. Land surface elevations (see Figure 3) range from greater than 200 feet in the northeastern portion of the site to a low of 166 feet on the western boundary.

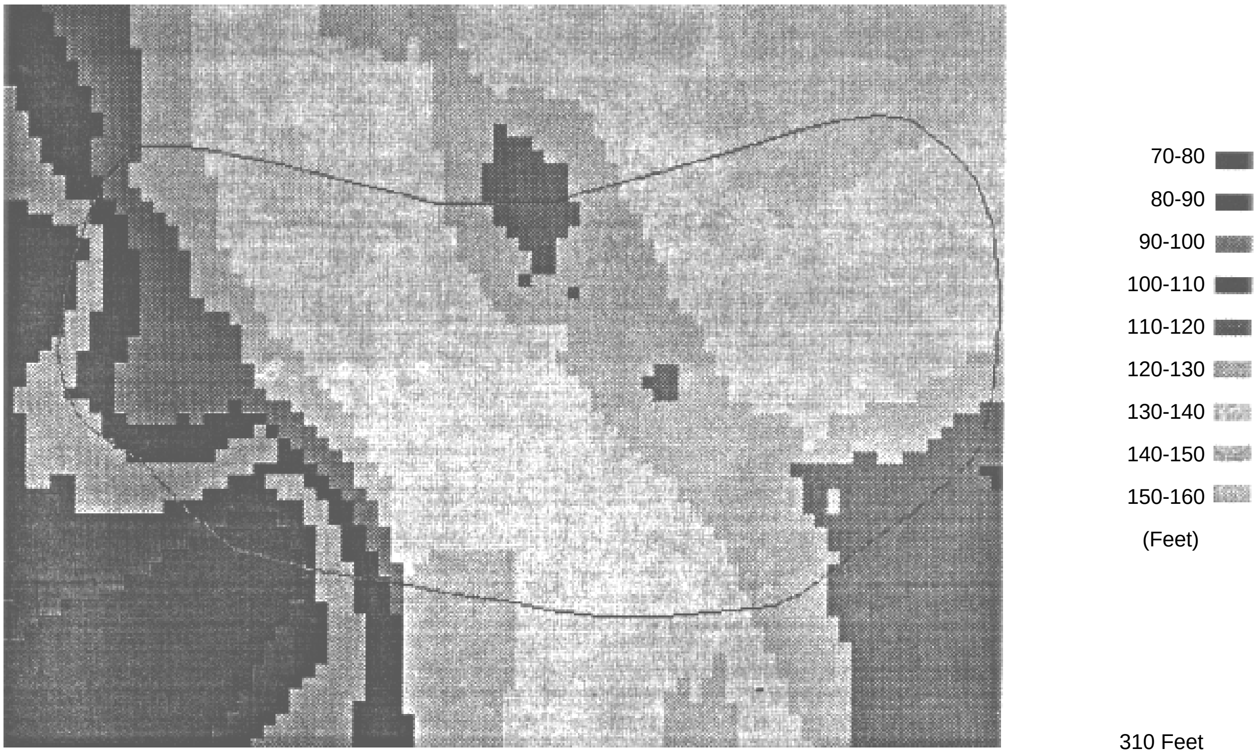


Figure 1. Bedrock elevation contour map derived from kriged data and transformed by MODRISI into IDRISI image file format.

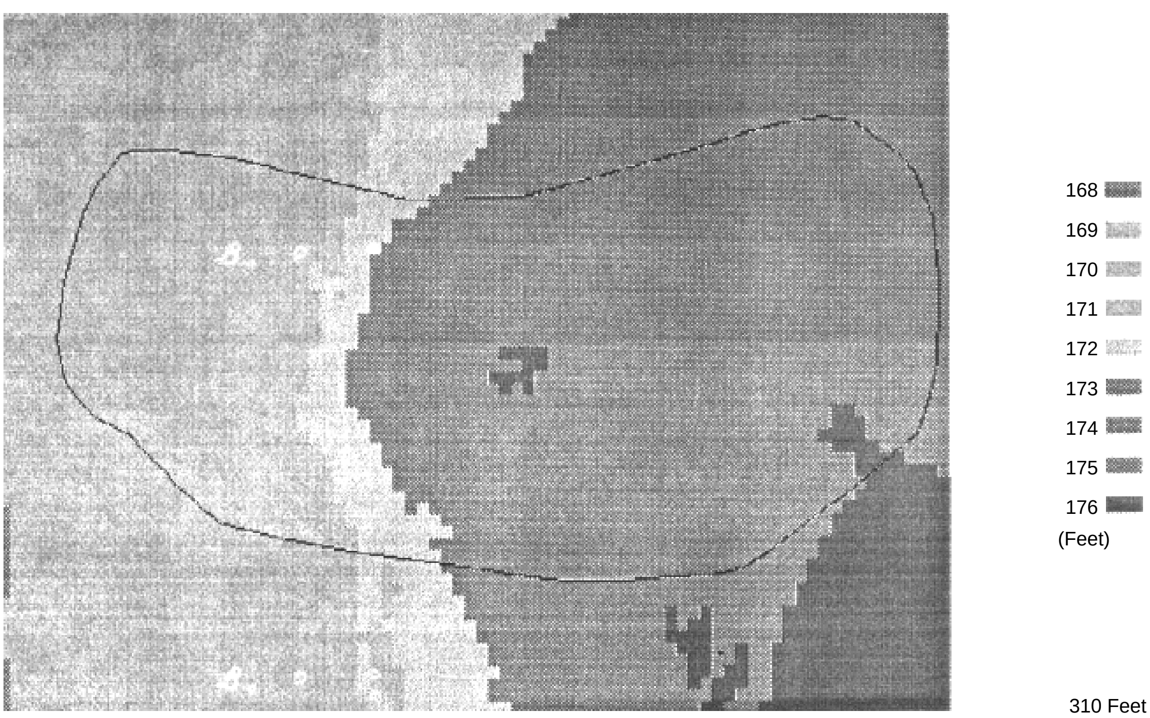


Figure 2. Water level contour map derived from kriged water level data and transformed by MODRISI into IDRISI image file format.

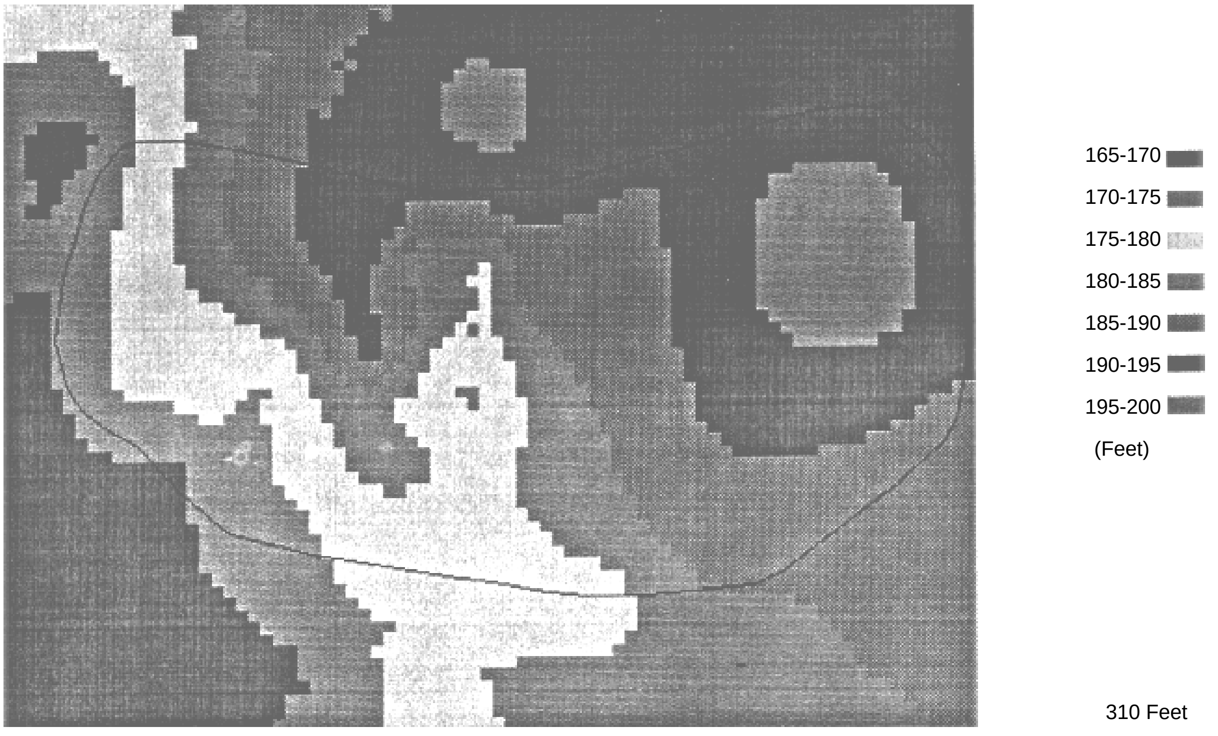


Figure 3. Land elevation contour map generated from kriged data and transformed by MODRISI into IDRISI image file format.

Additionally, the value (e.g., elevation) and (x,y) coordinates may be queried for any point, line, or area of an image file.

The analytical capabilities of IDRISI are illustrated in the following example. The saturated thickness was determined by subtracting the bedrock surface from the water level surface. The results illustrate the spatial variability of saturated thickness of the overburden aquifer (see Figure 4). The results correspond favorably with the general bedrock topography, as would be expected given the relatively low hydraulic gradient across the site. This OVERLAY (subtract) function may also readily be used to evaluate the adequacy of model calibration. For example, predicted values may be compared with observed values graphically, allowing the modeler to quickly visualize and identify areas of the model domain requiring additional consideration and manipulation.

Conclusions

One of the most tedious tasks in a ground-water modeling process is preparing input data and postprocessing model results. GIS allows rapid incorporation and evaluation of new site characterization information. The proposed combination of IDRISI, a raster-based GIS system, and MODRISI, a set of utility programs, could significantly reduce the amount of time necessary for entering data in required array formats. The visualization capabilities of IDRISI in conjunction with MODRISI and MODFLOW allow project managers to better understand the three-dimensional nature of subsurface environmental problems.

Acknowledgments

The authors wish to thank Chet Janowski, U.S. Environmental Protection Agency Region I, for his support on this project.

References

1. McDonald, M., and A. Harbough. 1988. A modular three-dimensional finite-difference ground-water flow model: USGS techniques of water resources, Book 6, Chapter A1.
2. Eastman, J.R. 1992. IDRISI: Technical reference. Clark University, Worcester, MA (March).
3. Van Metre, P.C. 1990. Structure and application of an interface program between a geographic-information system and a ground-water flow model. USGS Open File Report 90-165. Denver, CO: U.S. Geological Survey.
4. Hinaman, K.C. 1993. Use of a geographic information system to assemble input-data sets for a finite-difference model of ground-water flow. In: Proceedings of the AWRA Symposium on Geographic Information Systems and Water Resources, Mobile, AL, (March 14-17). pp. 405-412.
5. Rifai, H.S., L.A. Hendricks, K. Kilborn, and P.B. Bedient. 1993. A geographic information system (GIS) user interface for delineating wellhead protection areas. Ground Water 31(3):480-488.
6. Orzol, L.L., and T.S. McGrath. 1992. Modifications of the USGS modular three-dimensional finite-difference, ground-water flow model. U.S. Geologic Survey Open File Report 92-50.

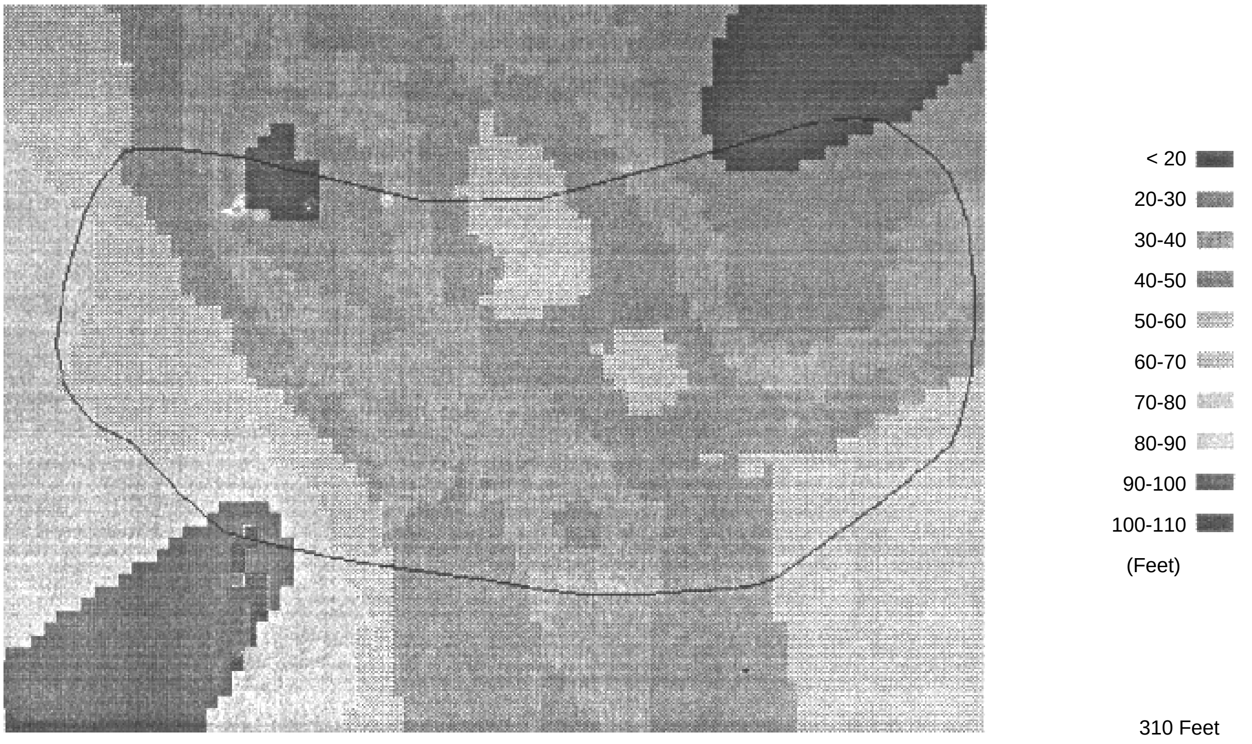


Figure 4. Saturated thickness contour map developed by subtracting bedrock image file from water level image file within IDRISI.

7. Torak, L.J., and K.W. McFadden. 1988. Application of a geographic information system to finite-element modeling of ground-water flow. In: Proceedings of the GIS Symposium Integrating Technology and Geoscience Applications, Denver, CO (September 26-30).

8. Harris, J., S. Gupta, G. Woodside, and N. Ziemba. 1993. Integrated use of GIS and three-dimensional, finite-element model: San Gabriel basin ground-water flow analyses. In: Goodchild, M., B. Parks, and L. Steyaert, eds. Environmental modeling with GIS. New York, NY: Oxford University Press.

9. Eastman, J.R., P.A. Kyem, J. Tolendano, and W. Jin. 1993. GIS and decision making. Explorations in geographic information systems technology, Vol. 4. United Nations Institute for Training and Research, Geneva, Switzerland.

Integration of EPA Mainframe Graphics and GIS in a UNIX Workstation Environment To Solve Environmental Problems

William B. Samuels
Science Applications International Corporation, McLean, Virginia

Phillip Taylor
Tetra Tech, Inc., Fairfax, Virginia

Paul Evenhouse
Martin Marietta, Inc., Research Triangle Park, North Carolina

Robert King
Assessment and Watershed Protection Division, U.S. Environmental Protection Agency, Washington, DC

Abstract

The Assessment and Watershed Protection Division of the Office of Wetlands, Oceans, and Watersheds has developed water quality analysis software on the U.S. Environmental Protection Agency (EPA) mainframe computer. This software integrates national on-line environmental databases and produces maps, tables, graphics, and reports that display information such as water quality trends, discharge monitoring reports, permit limits, and design flow analyses.

In the past, this graphic software was available only to users connected to the mainframe with IBM graphics terminals or PCs with graphics emulation software. Recently, software has been developed that can be used to: 1) access the EPA mainframe from a UNIX workstation via the Internet, 2) execute the Water Quality Analysis System (WQAS) procedures, 3) display WQAS graphics in an X-Window on the workstation, and 4) download data in a geographic information system (GIS) format from the mainframe. At the same time, this workstation can execute ARC/INFO and ARC/VIEW applications in other X-Windows. This capability allows analysts to have the power of GIS, the mainframe databases (e.g., Permits Compliance System [PCS], STORET, Reach File, Industrial Facilities Discharge File, Daily Flow File, Toxic Chemical Release Inventory), and the retrieval/analysis/display software (Environmental Data Display Manager, Mapping and Data Display Manager, Reach Pollutant Assessment [RPA], PCS-STORET Interface, UNIRAS) available to them on one desktop. This capability extends the tool set available to GIS analysts for environmental problem-solving.

This paper discusses application of these tools and databases to several problems, including EPA's watershed-based approach to permitting, and the RPA, an automated method to identify priority pollutants in watersheds.

Introduction

The purpose of this paper is to explore new geographic information systems (GIS) data integration tools that are applicable to a wide range of environmental problems, including the U.S. Environmental Protection Agency's (EPA's) watershed-based approach to permitting and the Reach Pollutant Assessment (RPA), an automated method to identify priority pollutants in watersheds. The ultimate goal is to make these tools and databases accessible to a wide range of users.

Understanding aquatic resource-based water quality management depends on access to and integration of diverse information from many sources. To date, the techniques to perform this integration, and thus yield meaningful analyses supporting environmental decision-making, are neither fully developed nor documented. New tools and information resources are now available, but not used to their full potential, for more valuable water quality and watershed analyses. EPA headquarters is responsible for ensuring that integrated data management tools are available for water quality analyses and data reporting as well as making national data systems more useful. EPA will accomplish this by upgrading and crosslinking systems, developing interactive data retrieval and analysis mechanisms, and providing easy downloading of data to client workstations.

The Assessment and Watershed Protection Division (AWPD) of the Office of Wetlands, Oceans, and Watersheds (OWOW) has developed water quality analysis software on the EPA mainframe computer (1). This software integrates national on-line environmental databases and produces maps, tables, graphics, and reports that display information such as water quality trends, discharge monitoring reports, permit limits, and design flow analyses. In the past, this graphic software was available only to users connected to the mainframe with IBM graphics terminals or PCs with graphics emulation software. Recently, software has been developed that can be used to:

- Access the EPA mainframe from a UNIX workstation via the Internet.
- Execute the Water Quality Analysis System (WQAS) procedures.
- Display WQAS graphics in an X-Window on the workstation.
- Download data in a GIS format from the mainframe.

At the same time, this workstation can execute ARC/INFO and ARC/VIEW applications in other X-Windows. This capability allows analysts to have the power of GIS, the mainframe databases (e.g., Permits Compliance System [PCS], STORET, Reach File, Industrial Facilities Discharge File, Daily Flow File, Toxic Chemical Release Inventory), and the retrieval/analysis/display software (Environmental Data Display Manager, Mapping and Data Display Manager, RPA, PCS-STORET Interface, UNIRAS) available to them on one desktop. This extends the tool set available to GIS analysts for environmental problem-solving. This paper discusses how these tools and databases have been applied to two examples: 1) a watershed-based approach to permitting and 2) the RPA, an automated procedure for identifying watersheds with priority pollutants.

Mainframe Databases and Tools

The EPA IBM ES9000 mainframe computer, located in Research Triangle Park, North Carolina, contains a large volume of digital water quality and environmental data available on-line through a number of data retrieval and display tools (see Figure 1). Other documents describe these databases and tools in detail (2, 3).

This effort focused on showing how these databases and tools can complement GIS activities. In some cases, data can be directly downloaded to a workstation in GIS format. An example of this is accessing EPA's Reach File (Version 1) from GRIDS (4). In other cases, databases are accessed by mainframe tools, the data are processed, and a GIS data set is produced that can be downloaded to a workstation. An example of this case is the RPA (RPA3) tool that integrates data from the Reach File, STORET, and PCS to identify priority pollutants in watersheds (5).

The mainframe can be accessed through several paths: Internet, PC dialup, or dedicated line into a terminal controller (see Figure 2). In the applications presented here, the Internet connectivity is emphasized because this is the mechanism that makes these databases and tools available to GIS analysts at their workstations. Figure 3 shows the hardware and software requirements for Internet access to the water quality data integration tools. The basic components are a UNIX workstation with an X-Window manager, the X3270 software, and Internet connectivity. The X3270 software is required to emulate an IBM 3270 full-screen terminal. This software is publicly available through EPA's National Computer Center User Support. In addition, an account on the mainframe computer is required. Once this account is established, an additional software module, GDDMXD, is required to map IBM host-based graphics to the workstation's X-Window. The GDDMXD software resides on the mainframe and is loaded when the user logs in. Once the hardware and software are set up, a single UNIX workstation can provide access to mainframe and workstation tools and databases on one desktop (see Figure 4).

Applications

To illustrate how these mainframe and workstation tools/databases can work together to solve environmental problems, we present two applications. The first shows a watershed-based approach to permitting; the second describes the RPA.

Watershed-Based Approach to Permitting

The watershed approach is a process to synchronize water quality monitoring, inspections, and permitting to support water quality protection activities on a geographic basis. It is a coordinated and integrated method to link science, permits, and other pollution control and prevention activities to meet state water quality standards. Numerous local, state, and federal agencies have recognized watershed approaches as the best way to manage natural resources effectively and efficiently. Establishing a schedule for data collection, permit issuance, and other elements of this approach affords the opportunity to coordinate and integrate other natural resource management efforts to make better use of

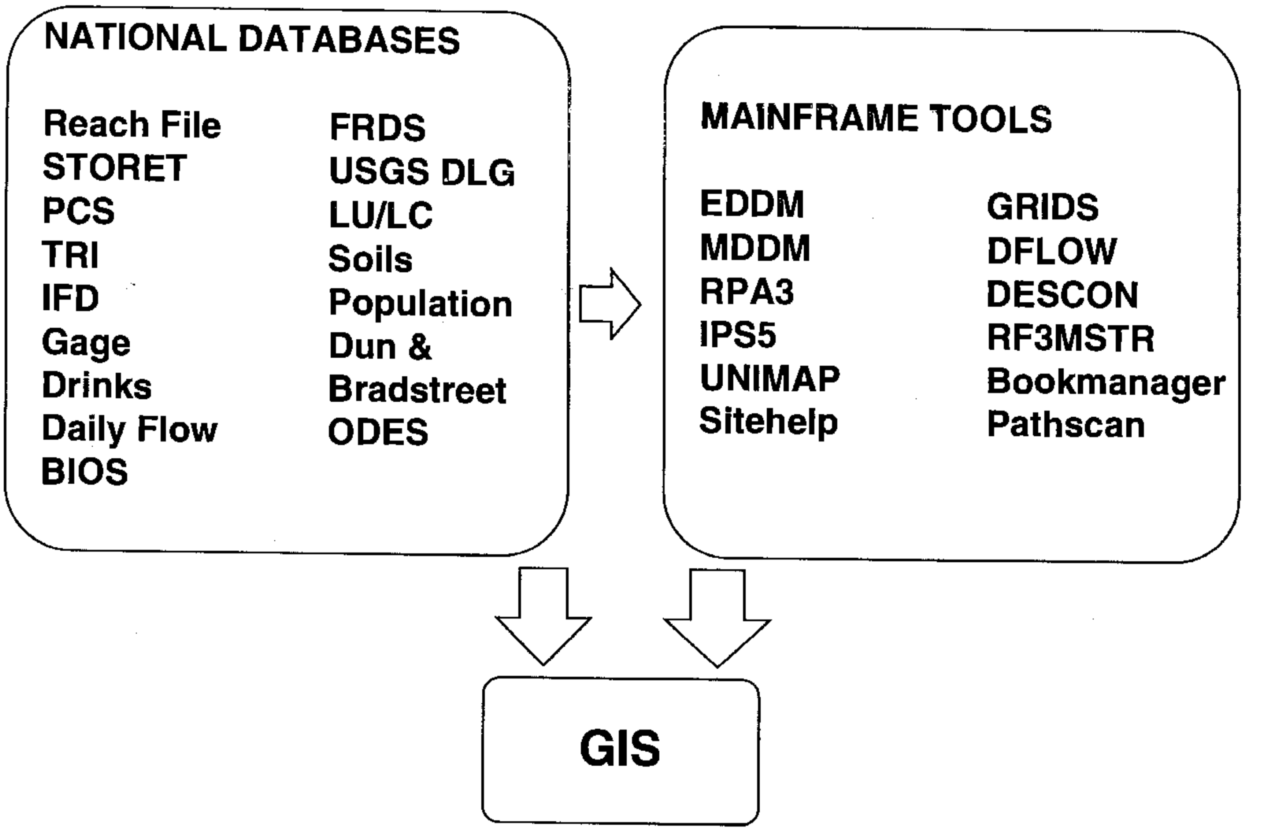


Figure 1. EPA mainframe databases/tools and linkage to GIS.

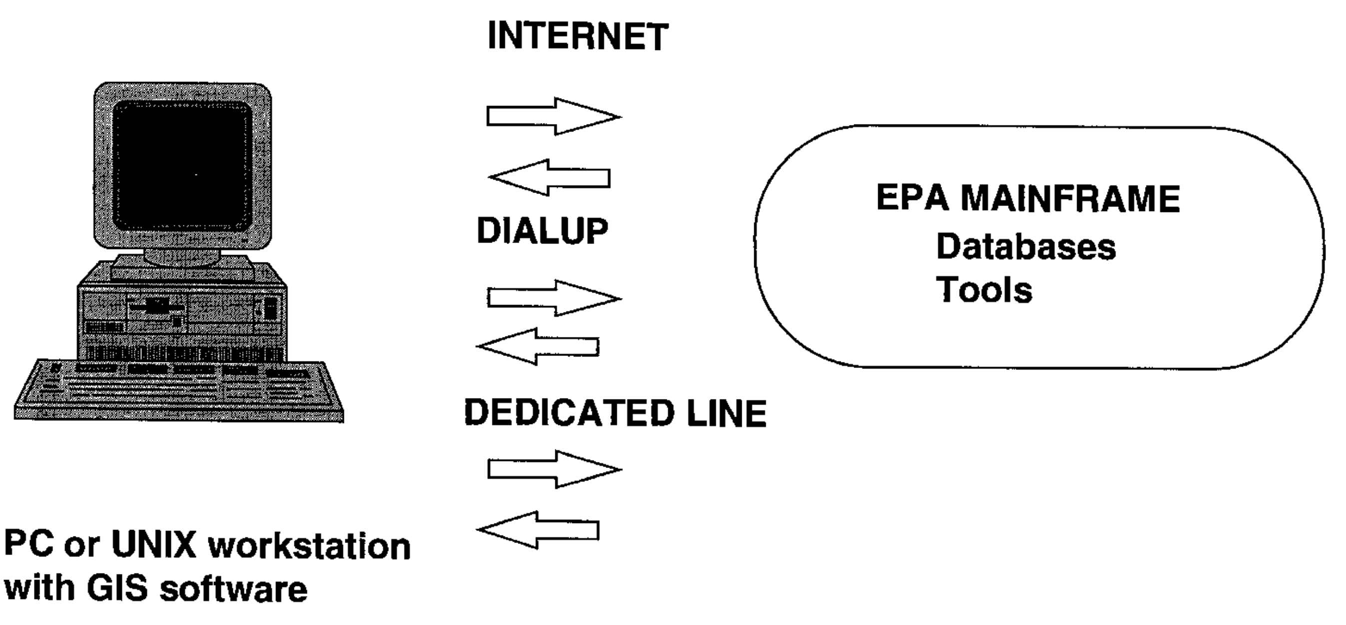


Figure 2. Access to the EPA mainframe databases and tools.

LOCAL WORKSTATION

HARDWARE

UNIX workstation (with X-Windows) on Internet (i.e., DG Aviion)

SOFTWARE

X3270 Software - creates an X-Window, which emulates an IBM 3270 full-screen session (provided by EPA)

EPA MAINFRAME ACCOUNT

GDDMXD software provided on the mainframe to map GDDM graphics into X-Windows

Figure 3. Hardware and software requirements for accessing water quality data integration tools via INTERNET.

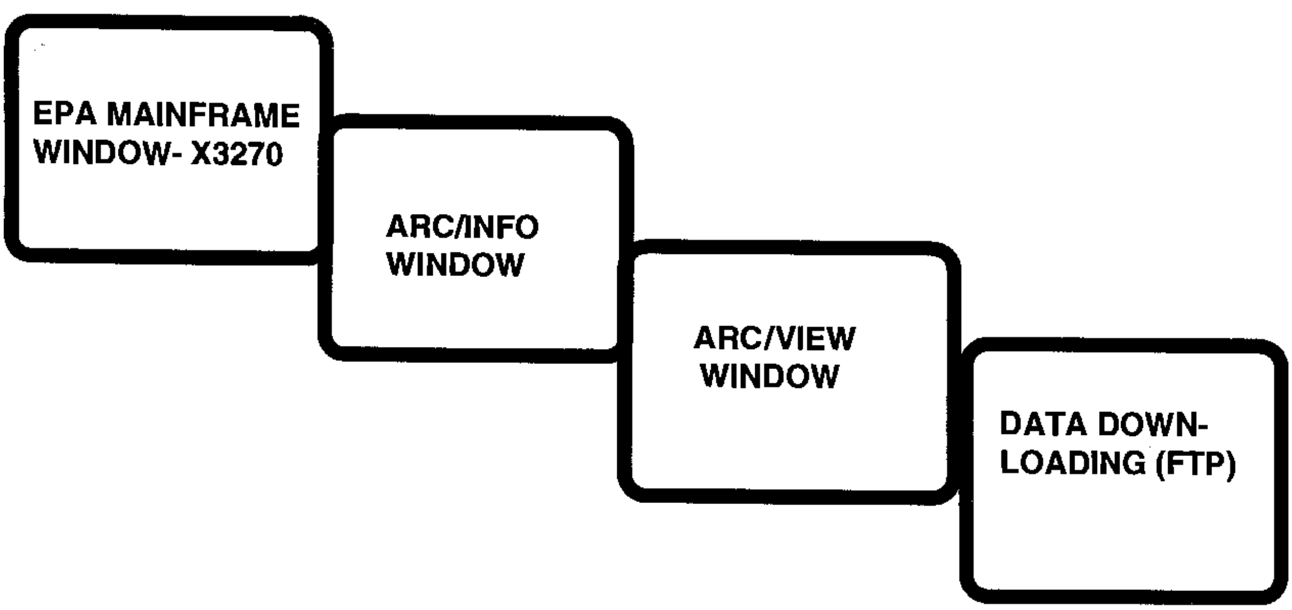


Figure 4. A UNIX workstation environment provides access to mainframe and workstation tools and databases on one desktop.

limited local, state, and federal financial and human resources (6).

This application illustrates how the watershed approach used GIS and EPA mainframe databases and tools. As an example, a four-step approach (see Figure 5) has been developed and applied to an impaired watershed (Saluda River basin) in South Carolina. Steps one and two identified watersheds of concern through their nonattainment of designated uses (see Figure 6) and highlighted the cause of nonattainment, in this case pathogens (see Figure 7). The data sets used were U.S. Geological Survey (USGS) hydrologic unit boundaries, Soil Conservation Service (SCS) watershed boundaries, and data from the EPA waterbody system, which were indexed to the SCS watersheds.[1] These data sets were integrated into an ARC/INFO arc macro language (AML) to allow users to pose queries and prioritize watersheds for further investigation.

Once priorities were set, the third step was to evaluate, in detail, the sources and causes of nonattainment. The Saluda River basin, which had pathogens as its cause of nonattainment, was selected for further analyses. In this step, the mainframe tools supplement the workstation GIS capabilities illustrated so far. A STORET retrieval was performed for ambient water quality stations monitoring for fecal coliforms. The STORET stations were partitioned into three categories (low, medium, and high) according to the state fecal coliform standard (7), which reads:

> not to exceed a geometric mean of 200/100 mL, based on five consecutive samples during any 30 day period; nor shall more than 10% of the total samples during any 30 day period exceed 400/100 mL

The categories in Figure 8 correspond to the standard as follows:

low: < 200/100 milliliters

200/100 milliliters $\leq$ medium < 400/100 milliliters

high: $\geq$ 400/100 milliliters

Figure 8 illustrates the use of ARC/VIEW to visualize the location of fecal coliform “hot spots” in the Saluda River basin. Figure 9 focuses on one SCS watershed (03050109-040) where pathogens cause nonattainment. The locations of industrial and municipal dischargers are plotted, and facilities with fecal coliform limits and their respective permit expiration dates (captured from the PCS) are shown in a table included with the figure. The GIS capabilities used to generate Figure 9

[1] Clifford, J. 1994. Personal communication with Jack Clifford, U.S. EPA, Washington, DC.

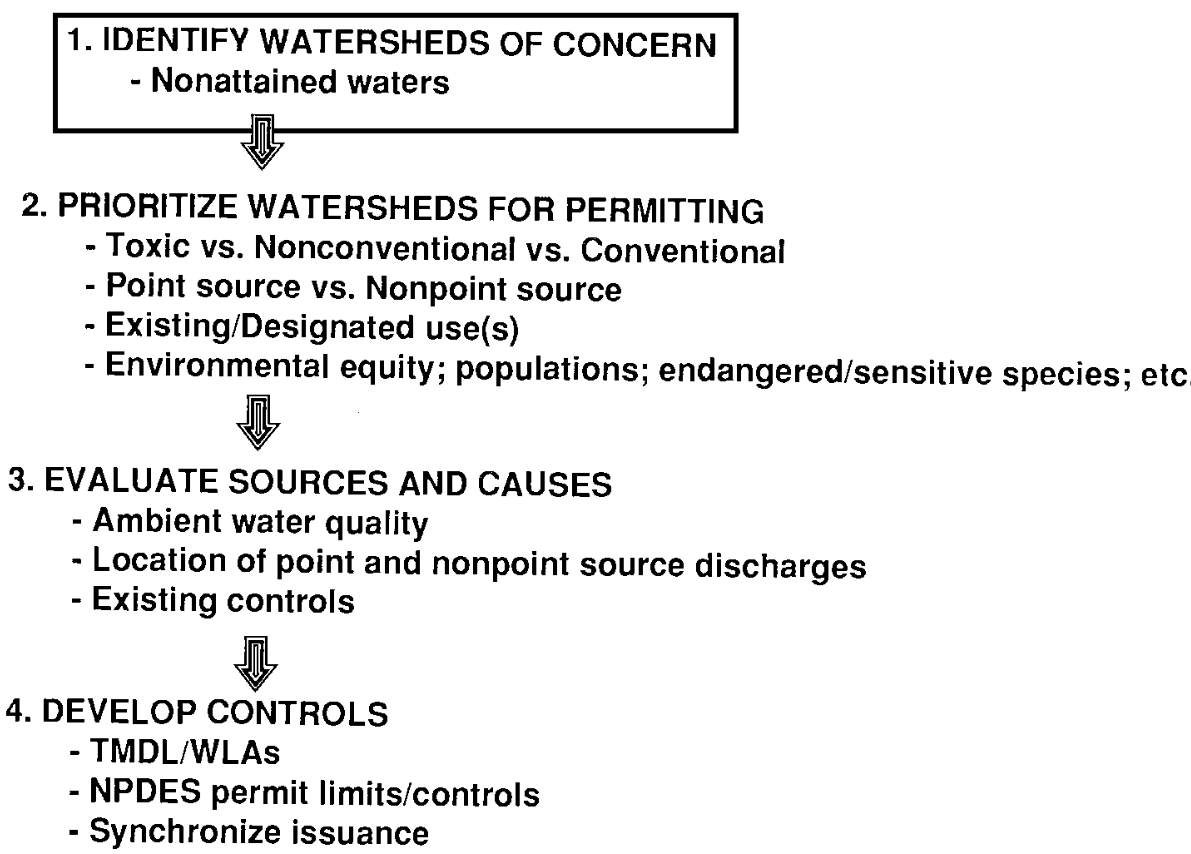


Figure 5. Four steps illustrating an example approach to permitting on a watershed basis.

Step 1: Identify Watersheds of Concern

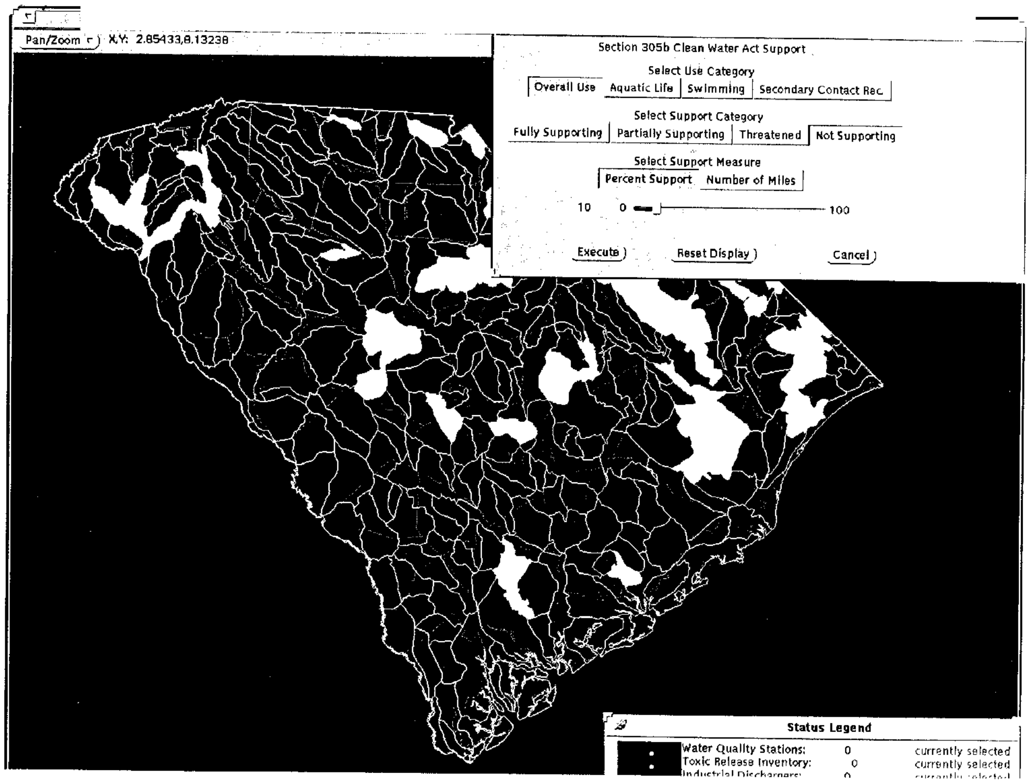


Figure 6. Identification of watersheds of concern—watersheds where at least 10 percent of the reaches are not fully supporting overall designated use (data and ARC/INFO AMLs provided by Jack Clifford).[2]

Step 2: Evaluate Priorities

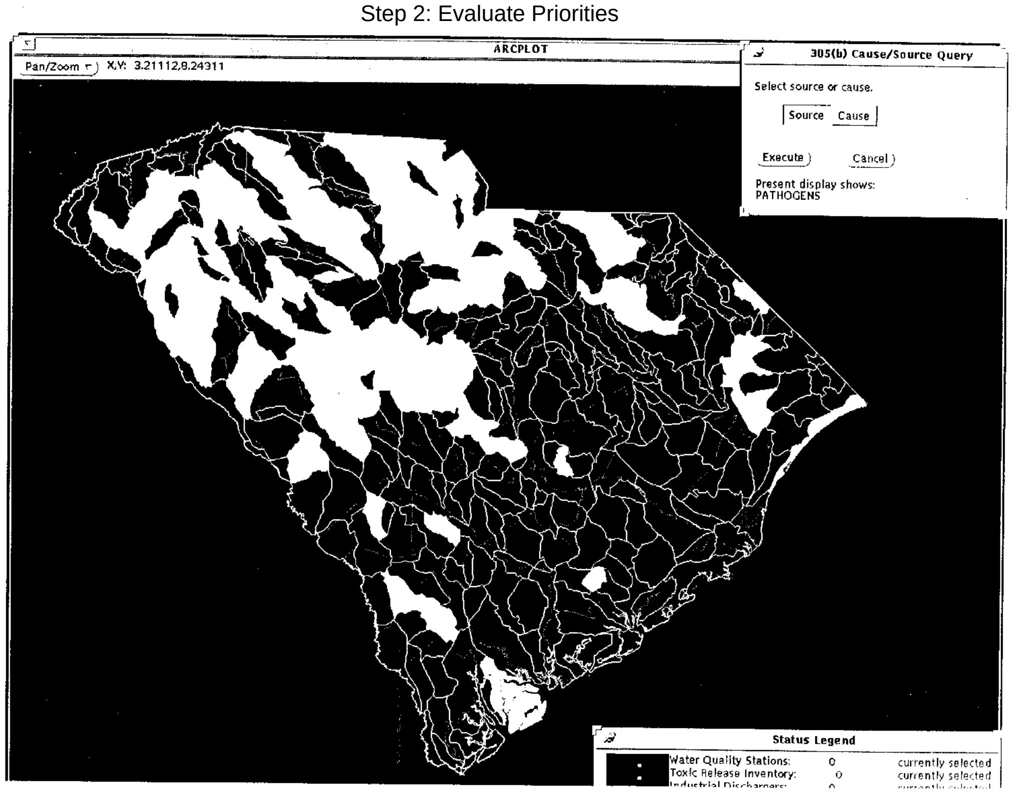


Figure 7. Watersheds where the cause of nonattainment is pathogens (data and ARC/INFO AMLs provided by Jack Clifford).[3]

[2] See note 1.
[3] See note 1.

Step 3: Evaluate Ambient Water Quality—Low, Medium, and High Levels of Fecal Coliforms

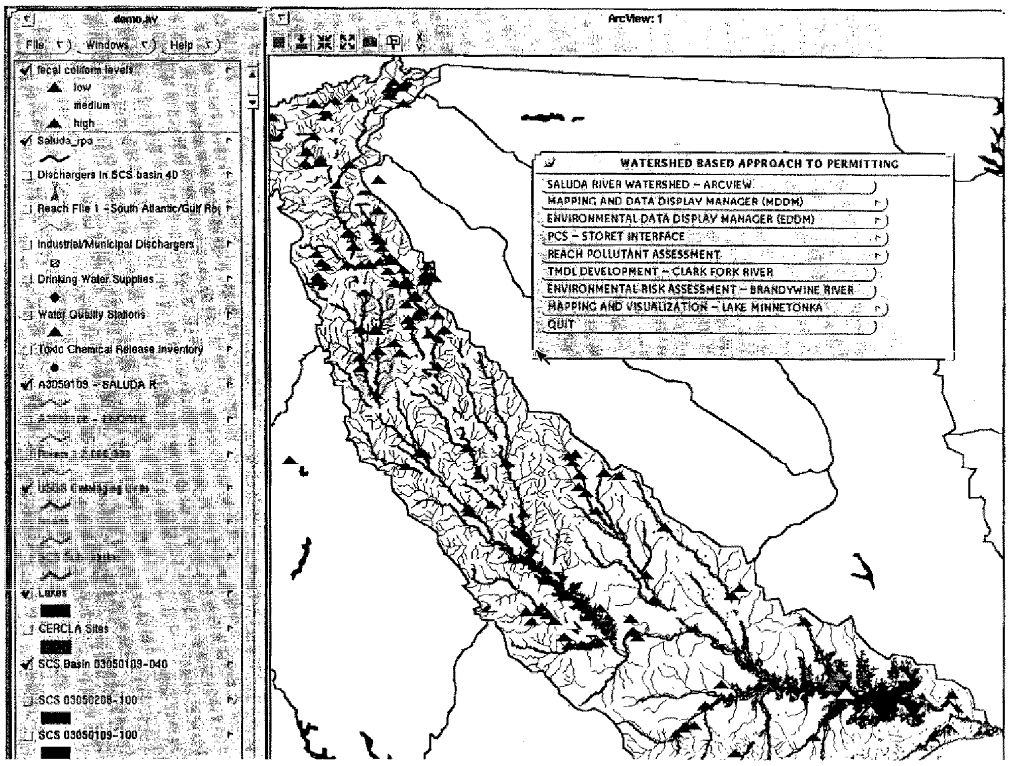


Figure 8. Map of the Saluda River basin showing the location of STORET monitoring stations and fecal coliform levels.

Step 3: Existing Controls—Dischargers That Have Limits for Fecal Coliforms

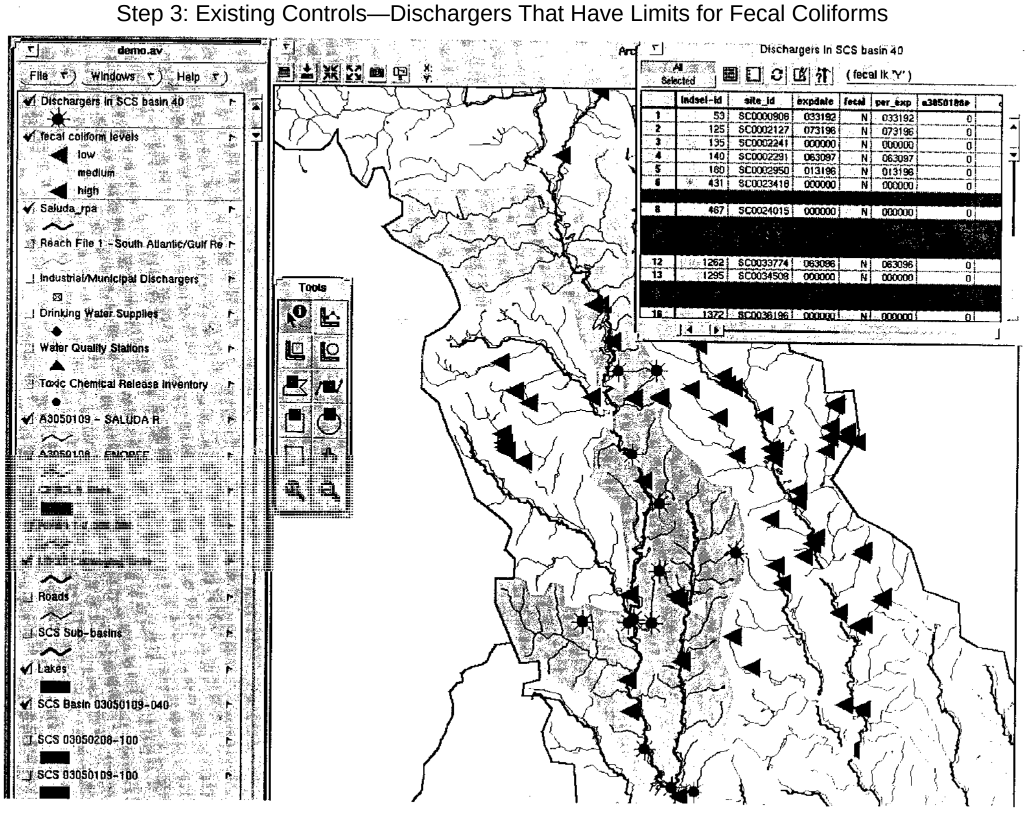


Figure 9. Focus on SCS watershed 03050109-040 (shaded in gray). Locations of STORET stations, municipal and industrial dischargers are also shown. The table in the upper right highlights dischargers with fecal coliform limits and their respective permit expiration dates.

show that permit issuance is not synchronous, which is a key element in the watershed approach.

Within this local workstation GIS environment, the attributes associated with the STORET monitoring stations and the PCS dischargers are limited with respect to the large amount of time series sampling data that exists in these databases. Figures 10 through 12 illustrate how an EPA mainframe procedure, the Environmental Data Display Manager (EDDM), can be accessed from an X-Window on the GIS workstation to query the entire STORET and PCS databases and thus provide additional data analysis and display capabilities to the GIS workstation. In Figures 10 and 11, a water quality inventory was performed for a STORET station, and a time series plot of fecal coliform levels is displayed. In Figure 12, the limits and discharge monitoring report (DMR) data were accessed from PCS for a sewage treatment plant. Excursions beyond the PCS limits for fecal coliforms are easily visualized in the plot.

The fourth and final step in the watershed approach was to develop controls for achieving water quality standards. This might include the development of total maximum daily loads (TMDL), waste load allocations (WLA), and the synchronization of permit issuance. Another mainframe tool, the PCS-STORET INTERFACE (referred to as IPS5 on the mainframe), can be used to access and compute design flows for TMDL development (see Figure 13) and to find all facilities discharging to a particular reach (see Figure 14), an initial step in the synchronization of permit issuance.

RPA

The RPA is a procedure on the EPA mainframe that automates identification of reaches where priority pollutants have been detected. It can be run for a user-selected state or USGS hydrologic unit.

Section 304(l) of the Clean Water Act (CWA) identifies water bodies impaired by the presence of toxic substances,

Step 3: Using EDDM To Evaluate Ambient Upstream Fecal Coliform Levels—Regulation of Upstream Dischargers Is Necessary

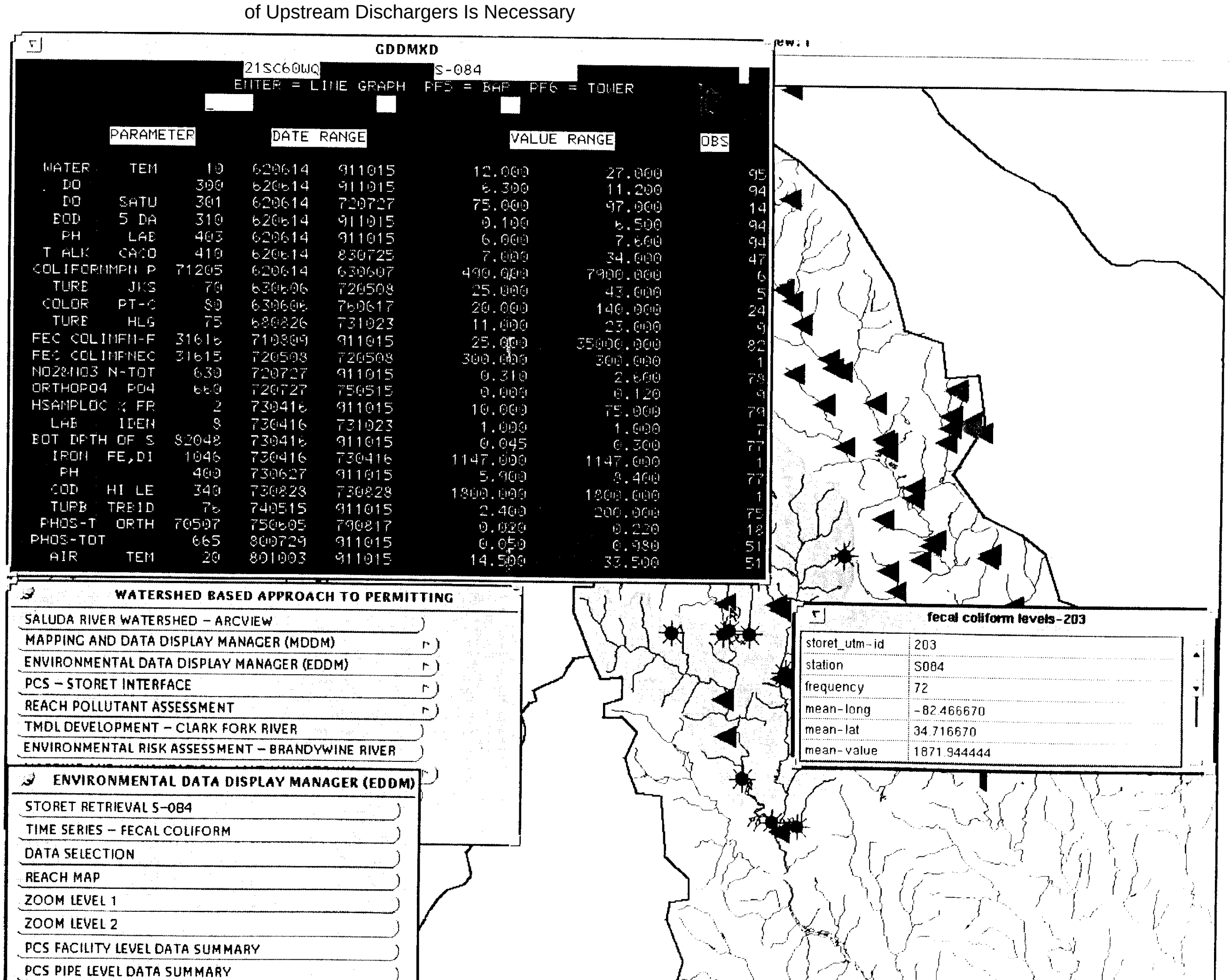


Figure 10. Using EDDM to perform a water quality inventory for STORET station 21SC06WQ S-084 in the Saluda River basin, South Carolina.

Step 3: Using EDDM To Visualize Trends in Ambient Coliform Levels Upstream of a Discharger

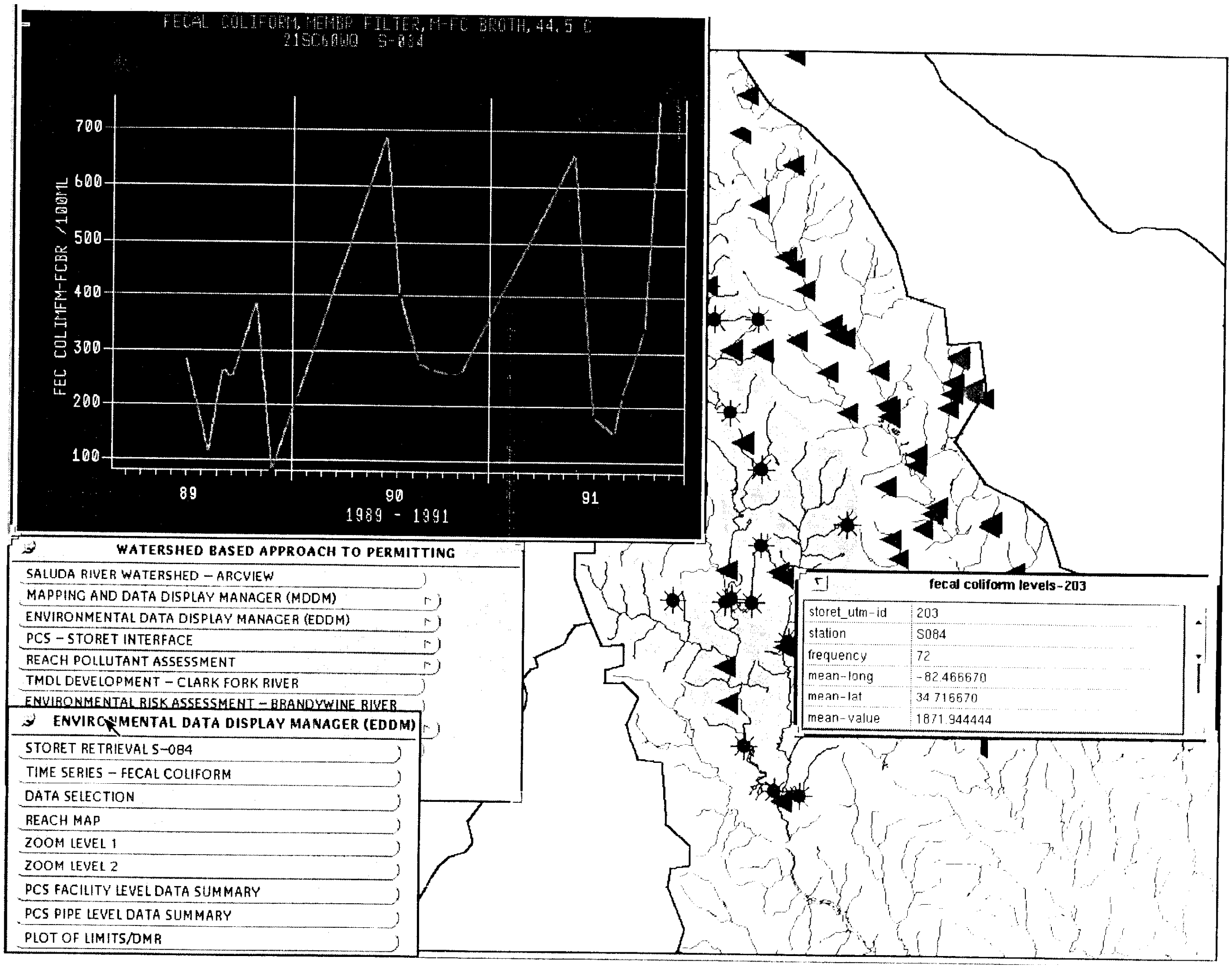


Figure 11. From the EDDM water quality inventory table, the fecal coliform parameter (31616) was selected for a time series plot.

identifying point source dischargers of these substances and developing individual control strategies for these dischargers. To meet these requirements, the EPA Office of Wetlands, Oceans, and Watersheds prepared guidance identifying criteria to be used in reviewing state reports.

The RPA was designed to address the requirements under criterion 7 of Section 304(l): identification of state waters with likely presence of priority toxic pollutants. This assessment was accomplished by identifying and summarizing reaches with point source dischargers of priority pollutants and water quality stations with priority pollutant data.

Information on the state's waters is summarized using the USGS hydrologic unit naming convention and the Reach Structure File (Version 1). Numerous databases were accessed and analyzed, including the Reach Structure and Reach Trace File (Version 1), industrial facilities discharge (IFD) file, STORET parameter file, PCS, and the STORET water quality file. The IFD file and PCS provided the facility information. Comparing information from both data sources identified active facilities and generated a complete list of facilities by their assigned reach numbers. Water quality data from STORET were summarized on reaches with priority pollutant monitoring data. Stations were retrieved with the following restrictions:

- Stations located within the state or hydrologic unit of interest.
- Ambient monitoring stations located on streams, lakes, or estuaries.
- Stations sampled for at least one priority pollutant in either water, sediment, or fish tissue on or after January 1, 1982.

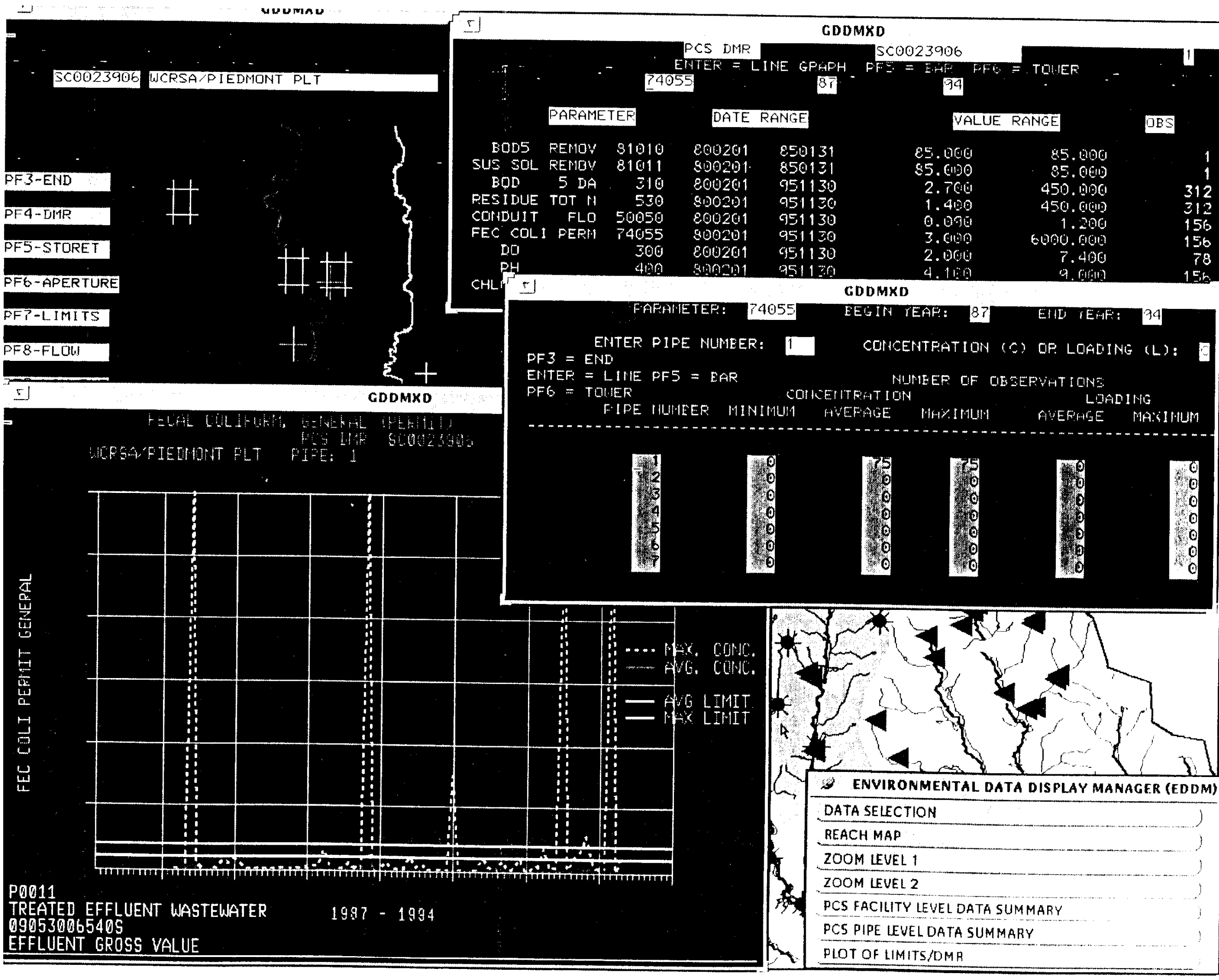


Figure 12. Using EDDM, PCS data is accessed for discharger SC0023906, Piedmont sewage treatment plant. The windows show plant location (upper left), facility and pipe summary data (upper right) and time series plot (lower left).

Recently, the RPA program was modified to output files compatible with GIS. An example of this is shown in Figure 15. The data in the table portion of this figure are written to two files as follows:

- For each reach in the hydrologic unit, the geographic coordinates are written to a file in ARC/INFO GENERATE format.
- The attributes associated with each reach (e.g., name, length, number of water quality stations) are written to a delimited ASCII file.

A third file is also automatically generated. This file is an AML that GENERATES the line coverage of reaches, defines and populates the INFO table of attributes, then joins the attributes in the INFO table to the line coverage of reaches. Once these three files are created, they are downloaded to the GIS workstation (via ftp) and processed by ARC/INFO. In ARC/VIEW, the user can identify a reach and determine:

- The number of water quality stations with priority pollutant monitoring data.
- The number and type of industrial facilities with priority pollutant discharge.
- The number of publicly owned treatment works (POTW) with and without indirect dischargers (see Figure 15).

In addition to this reach summary data, other tables (cross-linked to reaches, water quality stations, and dischargers) are produced that summarize the data by pollutant (see Figure 16). In this figure, each pollutant is cross-linked to the reach where it was detected, and the source of detection is also identified (i.e., water column, sediment, fish tissue, NPDES permit limit, Form 2(c) submittal) or

Step 4: Using the PCS/STORET INTERFACE for NPDES Permit Development—Analysis of Receiving Stream Flow Data

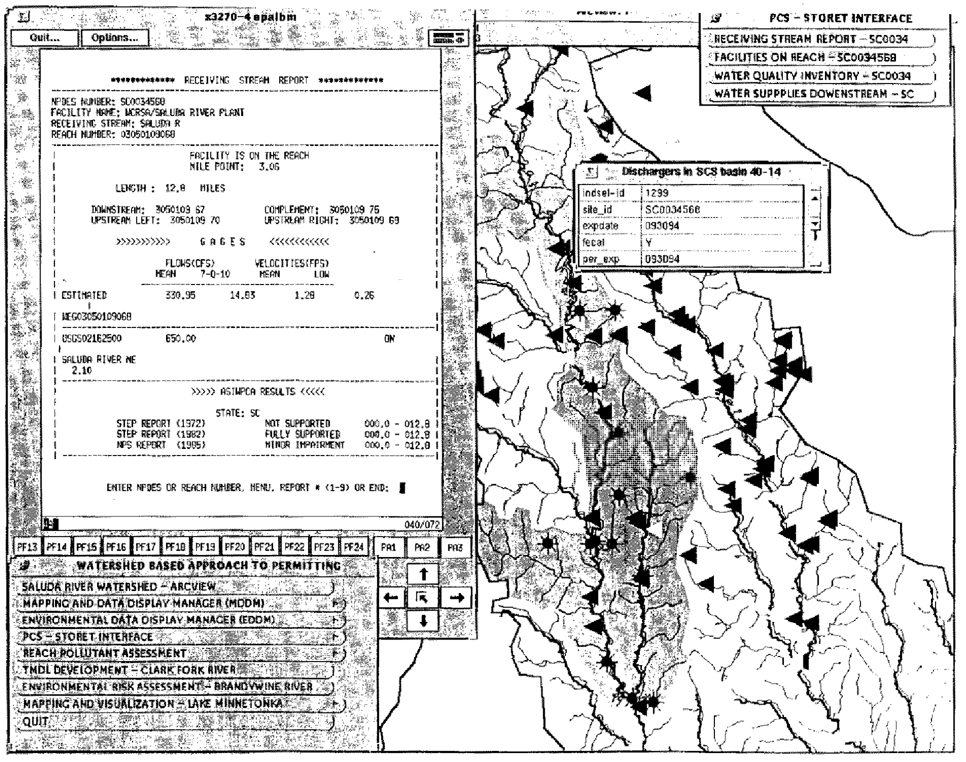


Figure 13. Using the PCS-STORET INTERFACE to access and compute design flows for a specific reach.

Step 4: Using the PCS/STORET INTERFACE To Determine Other Facilities on a Reach for Wasteload Allocation Purposes

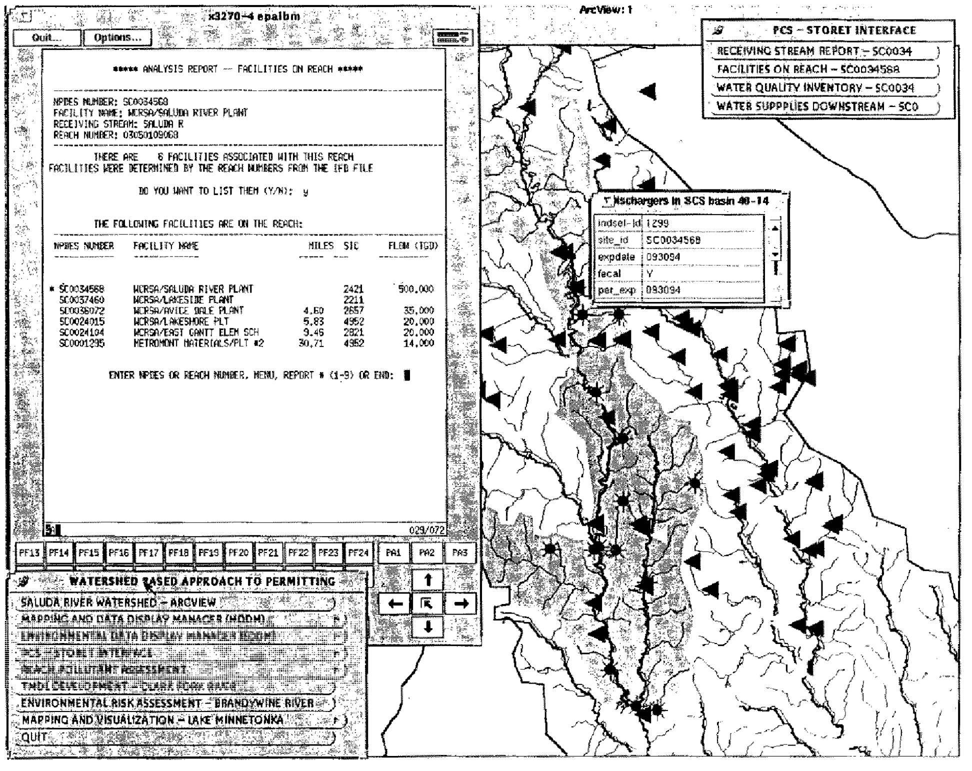


Figure 14. Using the PCS-STORET INTERFACE to list all facilities that discharge to a specific reach.

predicted to be in the discharger's effluent based on the standard industrial classification (SIC) code. More detailed information is also generated. For example, Figure 17 shows a detailed report for priority pollutants detected in the water column (similar reports are generated for sediment and fish tissue). Each pollutant is cross-linked to a reach and the specific monitoring station where it was detected. Basic summary statistics are also presented.

Figure 18 shows a detailed report for pollutants detected in the NPDES permit limit. In this figure, each pollutant is cross-linked to a reach and the specific NPDES discharger containing a permit limit. In addition, each discharger is identified as a major, minor, or POTW.

Figures 15 through 17 show the RPA output in the foreground and coverages displayed by ARC/VIEW in the background. Inspection of these figures shows that the Bush River is a priority pollutant reach containing seven industrial facilities, four of which discharge priority pollutants (one pulp and paper mill, three textile factories). Further examination of the data shows that cadmium has been detected in the water column and is also contained in the NPDES permit limit. It is also predicted to be in the discharge effluent based on SIC code. In the water column, cadmium was measured at 10 µg/L at two stations sampled in 1988. Finally, there is a limit for cadmium in NPDES facility SC0024490 (Newberry plant), a POTW on the Bush River. The RPA output, linked to GIS, can be used as a screening and targeting tool for identifying specific reaches within watersheds where toxic priority pollutants cause water quality degradation.

Summary and Conclusions

The proliferation of GIS workstations, the expansion of the Internet, and the development of X-Window-based graphics emulation software (X3270 and GDDMXD) has afforded analysts the opportunity to use the powerful analytical capabilities of GIS and the EPA mainframe databases and tools together on one desktop. Thus, a user performing a GIS watershed analysis can also have immediate and complete access to national on-line databases such as STORET and PCS by opening up a "window" to the EPA mainframe. This allows detailed queries to be performed that supplement the data already being analyzed at the local workstation. This capability allows users to easily visualize additional data without having to spend effort in retrieval, downloading, transforming, and reformatting to make it useful. By enhancing existing mainframe programs to create output in GIS format, the time spent importing data to the GIS is reduced and more time can be spent on analysis. An example of this capability is the RPA program.

Reach Pollutant Assessment: Priority Pollutant Detection in the Bush River

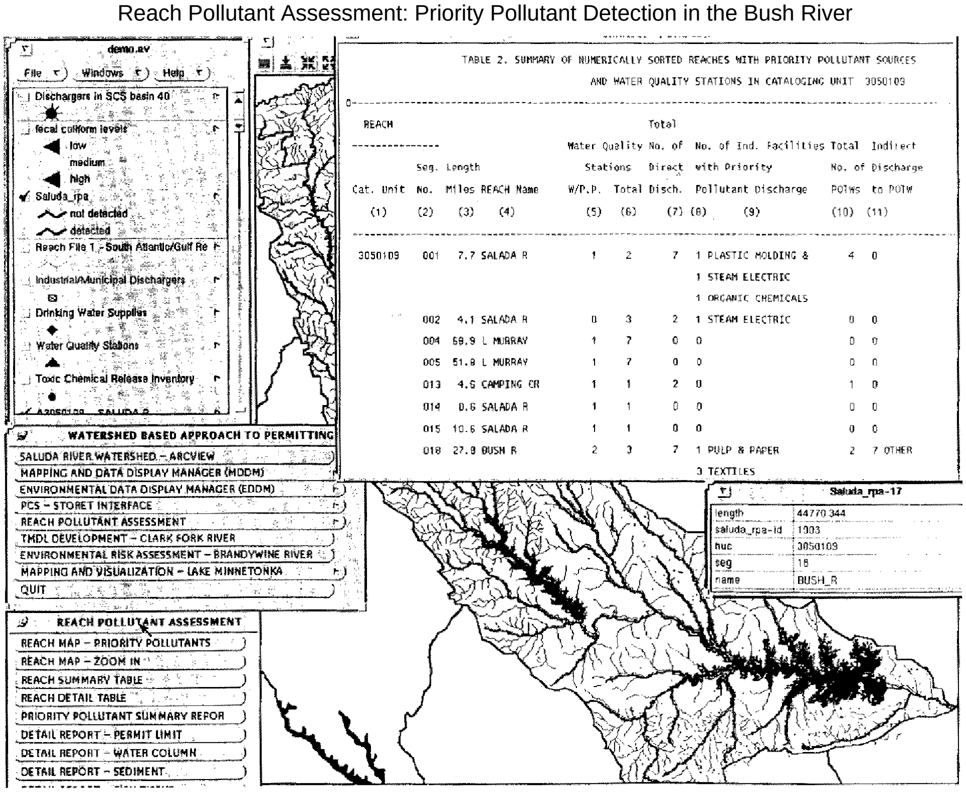


Figure 15. Using the RPA procedure to identify specific reaches with priority pollutants.

Reach Pollutant Assessment: Cadmium in the Bush River

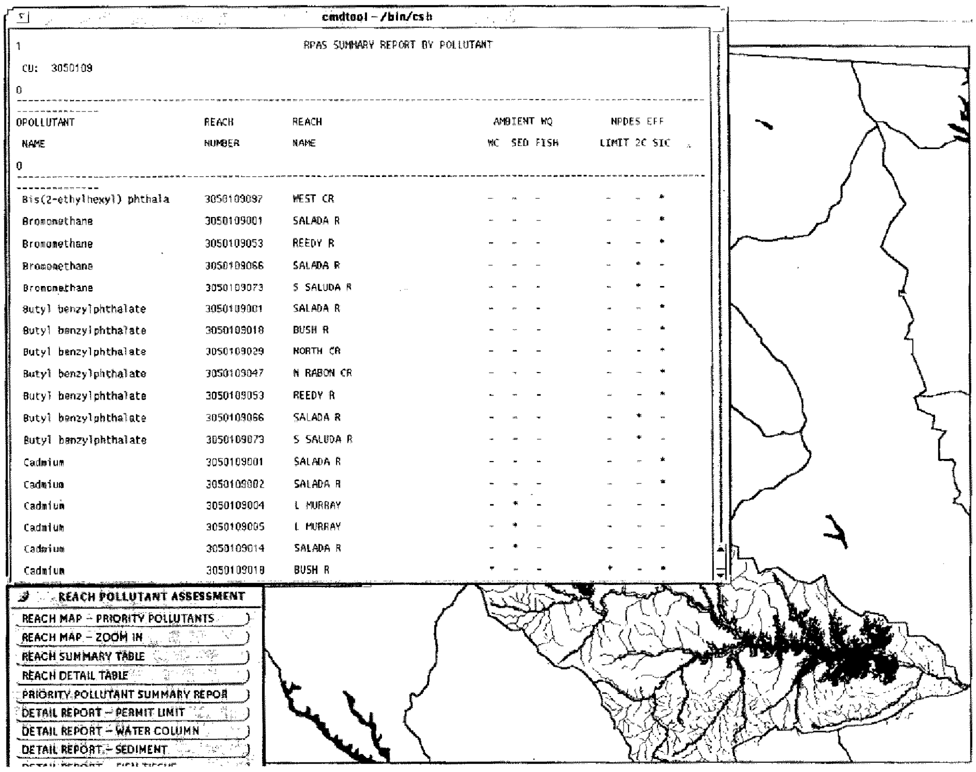


Figure 16. RPA summary report by pollutant.

Reach Pollutant Assessment: Bush River: Cadmium in the Water Column

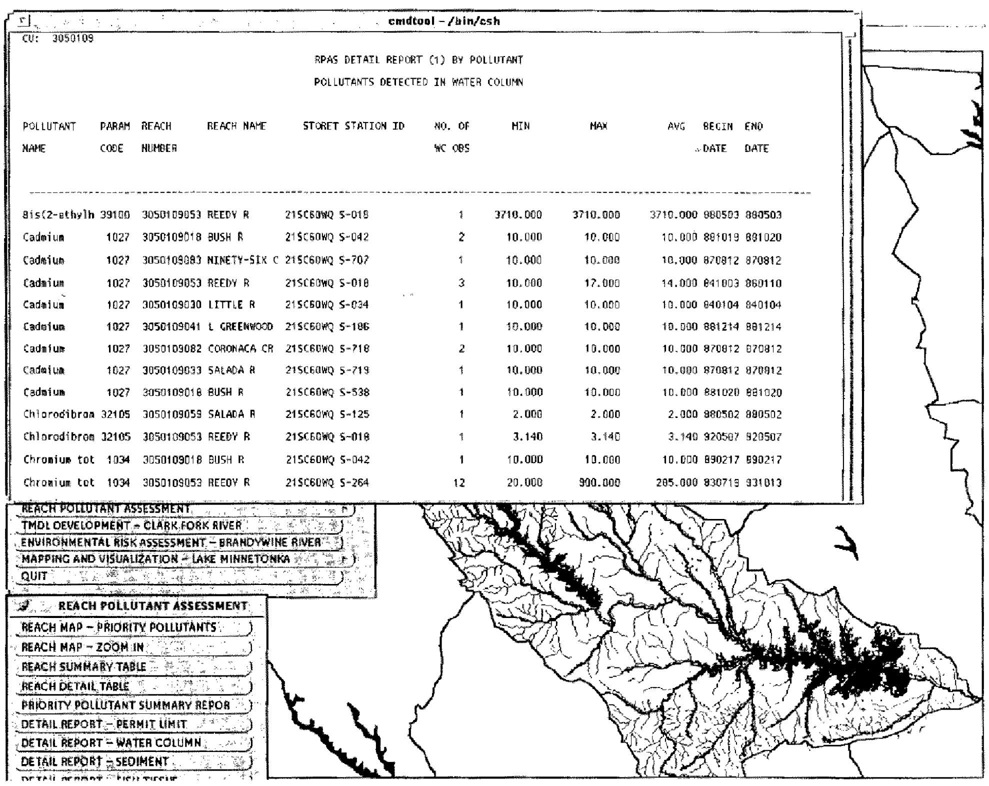


Figure 17. RPA detail report: pollutants detected in the water column.

Reach Pollutant Assessment: Permitted Industries for Cadmium—Bush River

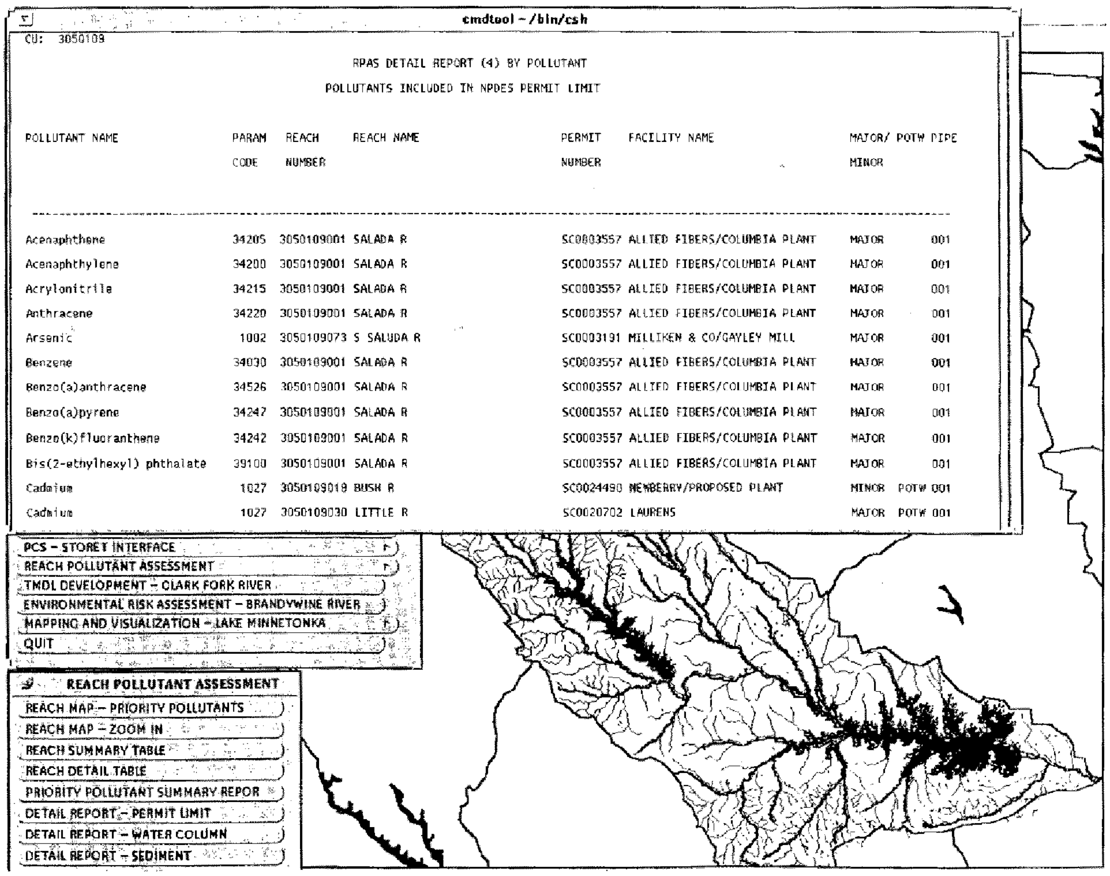


Figure 18. RPA detail report: pollutants included in the NPDES permit limit.

Two examples were presented as illustrations of how GIS and the mainframe databases and tools can work together.

In the first example, EPA's watershed-based approach to permitting, a four-step approach, was outlined, showing how a combination of local GIS functions and remote mainframe databases and tools were used in each step of the process. The end result was the targeting and prioritizing of watersheds of concern, and a detailed look at where and why water quality standards were not being met.

In the second example, the RPA program along with GIS was used to identify and map toxic priority pollutants and cross-link them to reaches, media (water column, sediment, fish tissue), NPDES dischargers, and monitoring stations. This analysis focused on what the toxic pollutant problems are and where they occur.

Acknowledgments

The authors wish to thank Jack Clifford, AWPD, OWOW, EPA, for providing data and ARC/INFO AMLs that were used in the South Carolina watershed example. The authors also extend their appreciation to Duane More and Tom Lewis (both with Martin Marietta), who provided support in setting up the X3270 and GDDMXD software.

References

1. Taylor, P.L., P. Evenhouse, L. Hoelman, T. DeWald, W.B. Samuels, and O. Hansen. 1988. STORET—Water quality analysis system. Presented at the State/EPA Data Management Conference, Raleigh, NC (June).
2. U.S. EPA. 1992. Office of Water environmental and program information systems compendium, information resources management: Tools for making water program decisions. EPA/800/B-92/001. April.
3. Samuels, W.B., P.L. Taylor, P.B. Evenhouse, T.R. Bondelid, P.C. Eggers, and S.A. Hanson. 1991. The environmental display manager: A tool for water quality data integration. Water Resources Bull. 27(6):939-956.
4. U.S. EPA. 1993. Geographic Resources Information and Data System (GRIDS): User guide. Office of Information Resources Management, National GIS Program.
5. Samuels, W.B. 1990. Reach pollutant assessment user's guide. McLean, VA: Science Applications International Corporation.
6. Washington State Department of Ecology (WDEC). 1993. Watershed approach to water quality management. F-WQ-93-029. October.
7. State of South Carolina. 1992. South Carolina code of regulations, Chapter 61, Regulation 68 and 69. Amended April 24, 1992.

You Can't Do That With These Data! Or: Uses and Abuses of Tap Water Monitoring Analyses

Michael R. Schock
Drinking Water Research Division, U.S. Environmental Protection Agency, Cincinnati, Ohio

Jonathan A. Clement
Black & Veatch, Cambridge, Massachusetts

Introduction

Linkage between human health and drinking water quality has been an area of interest in the United States for many years. Over the past approximately 10 years, drinking water monitoring requirements have expanded rapidly under the Safe Drinking Water Act (SDWA). Growing public and governmental interest in this environmental area makes the aggregation and consolidation of data on the occurrence and distribution of many organic and inorganic contaminants and background constituents of drinking water an important process. These data can then be made available for systemization and visualization to regulators, municipalities, water utilities, public interest groups, health researchers, consulting engineers, and water treatment scientists.

Given a sufficient number of data points and a convenient computerized database/mapping platform, a wide variety of maps can be generated to use in research and decision-making processes. The validity of doing so, however, rests inseparably upon the basis of the sampling plan and protocols, as well as the precision and accuracy of the analytical methods used for the constituents of interest. The well-known problem of matching the proper scale of the source data to that employed in the maps for interpretation is a critical problem with drinking water sampling, where many unappreciated small-scale variations render many, if not most, attempts to make generalizations inaccurate or meaningless.

This paper introduces and describes many concepts related to what generates or controls the concentrations of metals and other constituents in drinking water, ways in which the sampling protocol affects apparent levels of constituents, and the magnitude of temporal and spatial variability present in both municipal and private water supplies. Illustrations from water quality studies show in practical terms how generalizations must be kept to a minimum and how the data input into a geographic information system (GIS) for interpretation and evaluation must be carefully analyzed and screened to determine the appropriateness for various well-intended purposes. The discussion and examples show how many apparently significant trends and assessments of exposures or occurrences turn out to be merely artifacts of critical (yet subtle) inconsistencies or errors in the planning and execution of the sample collection process, or inconsistencies caused by the fact that regulatory (and not research) requirements govern the origin of the data.

The concepts this paper covers are equally valid in many other disciplines using or contemplating the use of GIS for interpretation of all kinds of "field" data.

Why Maps Are Useful for Drinking Water Studies

Maps and GIS databases could have wide applicability to drinking water studies. For example, they could provide the basis for investigating the occurrence of regulatory contaminants or related constituents, either to estimate the costs of compliance with a regulation or to estimate human health effects. Mapping could be useful to utilities and consultants investigating process changes for a utility or determining the effectiveness of some existing treatment such as corrosion control or chlorination. Use of GIS could also assist in assessing the feasibility and impact of system expansion. Another promising application would be GIS assistance in developing and implementing wellhead protection plans. Many other areas of application may be possible now, or will be discovered in the future, as GIS technology and regulatory requirements continue to develop.

Sampling Protocols for Data Usable in GIS

Several SDWA regulations have resulted or will result in the collection of geographically diverse drinking water quality data that may interest mappers. The Lead and Copper Rule, the Surface Water Treatment Rule, the proposed Information Collection Rule, and the Disinfection/Disinfection Byproduct Rule are but four examples. Many states have their own variations on federal drinking water regulations, so their data collection requirements may differ somewhat. Considerable data may also be collected for specific research studies of either academic or purely practical nature.

Chemical Factors in Constituent Behavior

For the purposes of this discussion, chemical constituents in drinking water may be classified as being generally reactive or nonreactive. Reactive constituents may change concentrations or chemical form for a variety of reasons, such as:

- A result of interaction with the background composition of the drinking water.
- By precipitation or dissolution reactions with pipe material used for the distribution system.
- By chemical reactions with disinfectants added at water treatment plants.
- By slow chemical reactions started at water treatment plants.

Nonreactive constituents may play an important role by providing a chemical background that indirectly influences the speed or extent of other chemical reactions and transformations. Table 1 gives a summary of many common constituents of drinking water and identifies whether they function essentially as reactive or nonreactive constituents.

Table 1. General Reactivity Trends for Common Drinking Water Constituents

Constituent	General Reactivity Tendency
pH	Highly reactive
Dissolved oxygen	Reactive
Calcium	Nonreactive (reactive when cementitious pipe linings are present)
Magnesium	Nonreactive
Total carbonate	Nonreactive
Total alkalinity	Reactive, particularly with pH changes
Chlorine residual	Reactive
Temperature	Either
Iron	Reactive
Copper	Reactive
Lead	Reactive
Zinc	Reactive
Silica	Nonreactive
Sulfate	Nonreactive
Orthophosphate	Reactive
Polyphosphate	Reactive
Total phosphate	Reactive
Nitrate	Nonreactive
Chloride	Nonreactive
Fluoride	Nonreactive
Trihalomethanes	Reactive
Haloacetic acids	Reactive

Reactive Constituents

Clearly, chemical species or compounds that can change in concentration or transform into other species or compounds during distribution make mapping on very large scales difficult to justify. Reactive constituents may also change concentration in the same place over time, such as water standing overnight in a home, school, or building, which is discussed in a later section. Some examples of reactions during water distribution follow:

- During lime softening processes at some central water treatment plants, a supersaturated state is used for the compound calcium carbonate to remove calcium (and sometimes magnesium) ions from the water. This condition is sometimes maintained into the distribution system as well to assist in maintaining chemical conditions useful for corrosion control of lead and copper. Thus, calcium levels, pH, and carbonate concentrations (and consequently, alkalinity) drop as water passes away from the plant (1, 2).
- The metals in pipe materials, such as iron, copper, zinc (in galvanized pipe), and lead, are oxidized by oxygen, free chlorine, chloramines, ozone, and other disinfectants, which renders them into a form that water can transport, unless other chemical conditions are such that a highly insoluble scale deposits on the pipe, immobilizing the metal (1, 3).
- Prolonged contact with chlorine disinfectant species converts a fraction of natural organic matter present in many distributed waters into regulated “disinfection byproduct” compounds, such as trihalomethanes, chloroform, and haloacetic acids (4, 5).
- Following the addition of chlorine or after increasing pH to enable some corrosion control for copper and lead, iron present in well waters in dissolved ferrous (Fe^{2+}) form oxidizes into Fe^{3+} form, which is much less soluble. Obnoxious “red water” results, as ferric oxyhydroxide precipitate forms and clouds the water.
- Polyphosphate chemicals added to “sequester” iron or manganese in well waters break down into simpler

polyphosphate forms of shorter chain lengths, plus orthophosphate. The orthophosphate frequently becomes present at high enough concentration to aid in controlling lead or copper (1, 6-8).

- Water passes through newly installed cement mortar-lined pipes, or aggressive water passes through older asbestos-cement pipes. Because of the particular chemical nature of the water, calcium carbonate and calcium hydroxide in the cement dissolve, raising the pH and hardness of the water (1).
- Free chlorine is added to disinfect water and is such a strong oxidant that it is unstable in water at normal concentrations. Additionally, it reacts with miles of unlined cast iron pipe, accelerating the decomposition of hypochlorous acid or hypochlorite ion to chloride. Consequently, the overall redox potential of the water supply and the effectiveness of disinfection decrease.
- A concentration of 1 milligram per liter (as PO_4) phosphoric acid is added to a distributed water at pH 7.5 to control lead corrosion. The orthophosphate reacts with exposed iron in the distribution main, however, and the residual concentration of orthophosphate decreases throughout distribution passage to the point where the level is no longer adequate to create the lead orthophosphate passivating film needed (1, 6, 8, 9).

Unless a constituent is known to be nonreactive, maps may be falsely generated under the premise that the concentration of a constituent is essentially a constant over some geographic area. Following the changes in concentration or chemical form of reactive constituents would also seem to be a useful application of GIS technology. One major restriction applies to the viability of that approach, however. Presuming that the analytical techniques used can adequately quantify the concentration and concentration changes observed, the scale of the variability or concentration change relative to the scale of the mapping perspective becomes critical to accurate mapping. A later section of this paper considers this critical factor in more specific detail.

Nonreactive Constituents

Almost no inorganic constituents in natural or drinking water are purely chemically inert. Under some conditions, and at some concentrations, significant reactions can occur. Some constituents that are actually reactive may act as if they are nonreactive constituents, however, because they are present in high enough concentrations relative to the extent of chemical reactions taking place that no discernible change in their concentration results. An obvious example is the dissolved inorganic carbonate (DIC = $H_2CO_3^* + HCO_3^- + CO_3^{2-}$) concentration (1, 10). Complexation and formation of passivating basic carbonate solid films of lead and copper by carbonate and bicarbonate ion dominate the corrosion control chemistry of copper(II) and lead(II) (1, 11). The concentration of DIC in water on either a molar or weight basis, however, is normally a factor of 500 to 10,000 higher than the lead or copper concentrations. Hence, changes in the DIC content from these reactions normally are analytically undetectable.

Another example is fluoride ion, which is often used as a distribution system water flow “tracer” because of its relative inertness. Actually, fluoride ion can form strong complexes with aluminum left in water following coagulation treatment with alum. The solubility of fluoride-containing solids with other major drinking water components (such as calcium and sodium) is very high, however, and fluoride reacts only weakly with metallic plumbing materials in the distribution system. Therefore, total fluoride concentrations tend to remain constant.

Relatively accurate maps of the occurrence and distribution of nonreactive constituents can be made, but their usefulness depends on the scale of the mapping relative to their occurrence and the particular question under investigation. All of this supports the need to ensure that the question asked can be answered correctly at the map scale.

Scale of Drinking Water Constituent Sources

More than 59,000 public water suppliers exist in the United States (12). Of these, approximately 660 are considered large water systems, which serve over 50,000 in population. These municipal systems use source water supplies that can be ground-water wells, “surface” waters (i.e., rivers, reservoirs, lakes), or a combination of both. Some water suppliers perform minimal water treatment of their own and purchase water from another water system or systems to satisfy their needs.

Surface Water Sources

Many water utilities use a single water treatment plant to treat surface waters, which could satisfy the entire water demand of the community all year. In many cases, however, utilities combine several surface water sources and use a different treatment plant to treat each water source. The water plants usually discharge into the distribution system at different points, and system hydraulics dictate the areas of the system in which waters mix. This is important because the water quality characteristics, which often differ among treatment plants, influence the corrosivity of the waters to various plumbing materials in the distribution system. Different water constituents also may affect the disinfection effectiveness of the treatment and the formation of unwanted disinfection byproducts.

For surface water systems, the chemical composition of the water depends on the upstream or watershed geochemistry, the seasonal nature of the water body used as the source, and the characteristics that the treatment imparts, such as coagulation with ferric sulfate or alum (aluminum sulfate), lime softening, filtration, pH adjustment, corrosion control treatment, chlorination, etc.

The scale of the source water chemical data, therefore, is large, driven by the geology, soil nature, land use, and climate. The chemical nature of the treated water, however, may differ significantly from that of its source.

Ground-Water Supplies

Many water utilities use multiple ground-water wells. A water supply of medium to large size usually uses multiple wells, instead of or in addition to the surface water supplies. Wells number from only two or three to more than 100 for very large water systems. Wells normally operate in different patterns, and only rarely do all wells operate at the same time. The yield of the wells and their water quality dictates the combination and number of wells used at a particular moment. The wells may or may not be from the same aquifer, and even if they are, local inhomogeneities frequently exist in water composition (especially with iron and manganese) that limit the usefulness of certain wells without substantial treatment.

Historically, utilities have treated some (but not necessarily all) wells with a chemical such as a polyphosphate or sodium silicate to sequester the iron and manganese from wells. Some utilities install physical removal processes such as ion-exchange softeners, reverse-osmosis plants, aeration systems for iron removal, air stripping towers for volatile organic compound or radon removal, or "greensand" filters for the removal of iron and manganese. These facilities sometimes exist at only certain well sites or at some point where water from multiple wells is combined.

The scale of chemical controls on ground-water supplies, therefore, becomes only hundreds of feet. Contaminants of raw waters, such as arsenic, nitrate, or chromium, are geologically and geochemically controlled. Therefore, their occurrence is geographically variable on even a small scale, and the variability exists vertically in the subsurface as well as horizontally. A municipality may use wells of different depths into different aquifers, or even approximately the same depth spread out over hundreds of feet to many miles in the same aquifer or a variety of geologic units.

The variability of individual ground-water wells over time (such as seasonally) is usually less apparent than with surface water sources, but the fact that many wells are frequently used in different combinations and for different lengths of time (hours to days, usually) makes characterizing "influent" water quality complicated. The same observation applies to water systems that allow different amounts of water to bypass treatment processes (e.g., ion-exchange, reverse osmosis) depending on the levels of targeted undesirable contaminants (e.g., nitrate, sulfate, arsenic).

These characteristics of the nature of chemical composition, use, and treatment of ground-water supplies clearly show that generalizations over areas such as states or geographic regions (e.g., New England, Upper Midwest) are at least very gross and uncertain and at worst, entirely misleading when decisions are to be made about risk and health assessments, or estimates of the necessity for certain treatments or economic impacts of different potential drinking water regulations.

Combination Systems

Some municipalities combine the use of surface water supplies and ground-water wells. Therefore, general water chemical characteristics vary throughout the system in a regular manner in response to the location and use of different sources, as well as relative amounts of water that the different sources produce and deliver.

Distribution System Mains

The next lower level of scale is the distribution system network of pipes and storage. Common materials used for distribution system piping include cast iron, ductile iron, cement mortar-lined iron, iron with organic coatings, asbestos-cement (A-C), and various forms of plastic. Pipe diameters range from about 4 inches to many feet, depending on size of the water utility and community, size of the neighborhood fed by the line, and distance of travel for the water. Here, because of the large volume of water involved relative to the pipe diameter, the major chemical interactions involve such constituents as hardness (calcium and magnesium) ions, pH, iron, bicarbonate and carbonate ions, and chlorine residual species, and possibly microbiological parameters such as total plate counts, heterotrophic plate counts, and assimilable organic carbon. Disinfection byproducts (DBPs) may change in concentration and type because of the time involved in the water traveling through the piping from the treatment plant. Trace metal contamination, such as lead and copper, is usually negligible from this source, unless it is present when distributed from the wells or water treatment plants.

Depending on prevalent economics and construction practices during periods of water system growth, the materials will not be either randomly or uniformly distributed geographically within system boundaries. Water flow often varies greatly within the distribution system, and water lines sometimes terminate in dead-end areas with minimal flow rates. Water quality often differs substantially in these dead ends from that in the fully flowing parts of the distribution system.

Household Service Lines

Service lines represent the connection between the house or building and the distribution main. Sometimes, the service lines are joined to the mains by a flexible, approximately 2- to 3-foot long pipe called a "gooseneck" or "pig-tail." Historically, this connector was often made of lead. Recently, copper has been the most widely used material, with plastic gaining in acceptance. Service lines for homes are usually 0.75 to 1 inch in diameter, with service lines for many commercial buildings or multifamily dwellings ranging in size from 1.5 to 3 inches in diameter. Service line for homes and buildings have usually been made of lead, brass, copper, galvanized steel, or plastic. The material used depends on the age of the water connection and the construction practices of the area involved. A recent report estimated that approximately 6.4 million lead connections ("goosenecks") still exist in the United States, and about 3.3 million lead service lines still exist (13). In many communities, old lead service lines remain a major source of lead in drinking water.

Like distribution system materials, service line materials may vary greatly within a distribution system by space and time. For instance, in large eastern cities, very old neighborhoods may have many (or even mostly) lead service lines. New neighborhoods likely have copper or plastic service lines. Galvanized steel or copper pipes may have been installed between the era when lead was used and modern times. With the exception of Chicago, where lead service lines were occasionally installed into the 1980s, the use of lead for service lines generally stopped in the late 1940s or early 1950s. An example of nonuniform distribution of service line materials is shown by Figure 1, a map indicating Cincinnati sampling sites for Lead and Copper Rule (14-17) monitoring. Erratic clustering of different service line materials is evident.

Rehabilitation of old houses or replacement of failed piping results in a mixture of new and old material in areas where houses are predominantly old. Following completion of the construction, maps of service line material would show many clusters representing prevalent plumbing codes and economics.

Interior Plumbing

Interior plumbing of buildings and houses reflects even more variability than service lines. This is the dominant contributor to lead and copper levels at most sites covered under the Lead and Copper Rule (14-17). Interior plumbing consists of piping, plus a large number of valves, connectors, fixtures, and perhaps soldered joints and a water meter. Any or all of these components are replaced at varying intervals as a result of failures or remodeling. Therefore, even generalizations within a small neighborhood are risky, unless the neighborhood is very new and uniformly constructed. When attempting to survey the composition of plumbing materials that might be the source of drinking water contamination, merely asking for the age of the house or building is insufficient. Questions must be asked to obtain the necessary precise information on the age and type of plumbing materials and components in the building.

Typical interior plumbing materials include lead, galvanized steel, copper, and different plastics for pipes. Some brass and black steel have been used for short times in some areas. Faucets are almost always made with either brass or plastic internal parts, which differ in composition from the exteriors, which are usually plated with chrome or other metal. Interior faucet volumes typically range from about 30 milliliters to 120 milliliters, depending upon design. Valves and meters are also frequently made of brass or bronze, which are copper/zinc alloys usually containing 2 percent to 6 percent lead. Until recently, solders used to join copper drinking water pipe sections were usually a tin and lead combination, containing 40 percent to 60 percent lead. Occasionally, connector lines to fixtures include copper, stainless steel, aluminum, or flexible plastic sections.

Private Water Systems

The many possible designs of domestic water systems originating from wells or cisterns are too numerous to illustrate. Figure 2 gives an example of one such system layout. Private systems share many features with domestic systems supplied by water utilities, however. Interior plumbing shares most of the same configurations and materials. For private water systems, additional plumbing that could cause contamination or water chemistry changes includes well casing material, submersible pump casing and fittings, pressure tank feed and control plumbing, and nonsubmersible pump interior materials. Therefore, problems with determining the frequency and distribution of levels of potential contaminants include those present for domestic situations in general, plus those complications arising from cycling of the pumps, pressure tank system, or both.

Water Samples Representing Distances

One of the most important fundamentals of understanding drinking water sampling is that volumes of water (e.g., 1-liter samples, 250-milliliter samples) represent the linear distance of plumbing material in contact with the water sampled. Because of water mixing and flow during use or sampling, they are also integrated samples of that volume. This understanding is at the heart of designing accurate water sampling programs and making viable interpretations of existing monitoring data that may be contained in (or mappable by) a GIS.

Table 2 summarizes some interesting and important relationships between pipes of different inside diameters

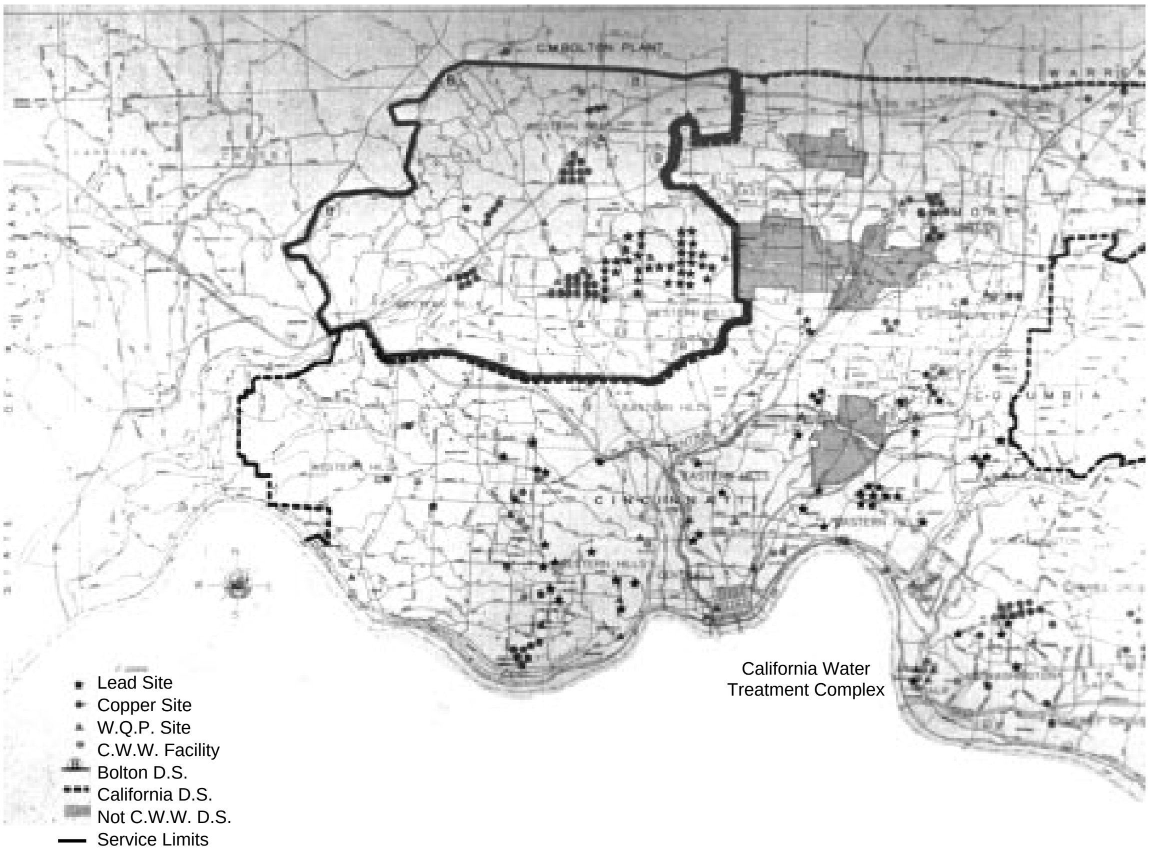


Figure 1. Cincinnati Water Works Lead and Copper Rule compliance monitoring, July to December 1992.

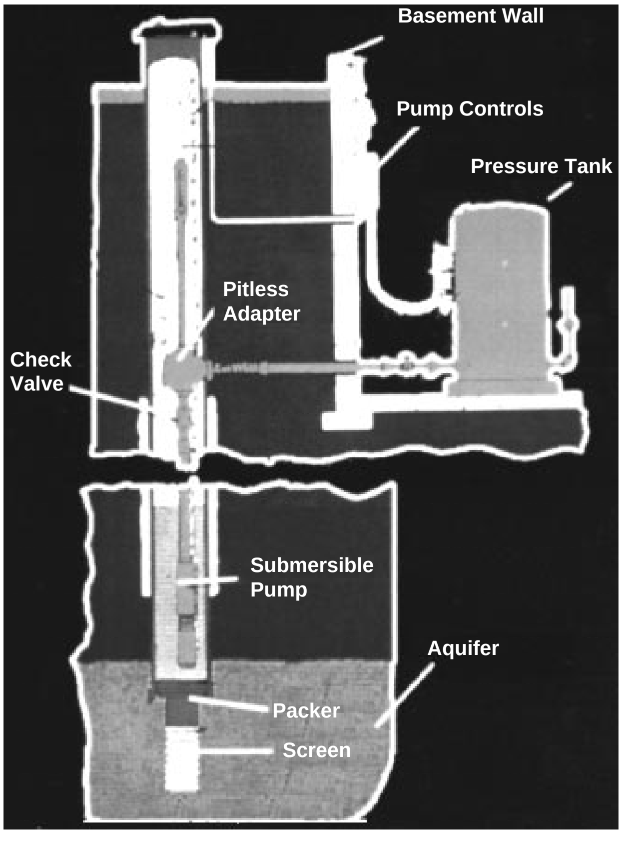


Figure 2. Distribution system.

(IDs) and the volumes of water they contain per unit of length (7). Much domestic interior plumbing has an ID of approximately 0.5 inches, depending upon the material.

Figure 3 shows schematically what parts of a plumbing system would likely be represented by samples of different volumes taken after water was allowed to stand in the pipe for many hours. Faucets, bubblers, and other terminating fixtures vary widely in volume. Kitchen-type fixtures usually contain from 60 to 120 milliliters of water. Bathroom-type fixtures may contain only about 30 to 60 milliliters of water. Bubblers, such as those frequently found on school or office drinking fountains, are smaller still. As can be seen schematically in Figure 3a, a small volume such as 125 milliliters captures the faucet and a short distance of pipe immediately leading to it. In many plumbing systems, this volume catches water in contact with numerous soldered joints. On the other hand, if a single 1-liter first-draw sample is taken, the water in the bottle represents a much longer distance back into the plumbing system. In a situation where the source of lead in drinking water is a new brass faucet, or soldered joints of lead-tin solder, this larger volume usually gives a lower lead concentration than the smaller volume because more water in the sample is not in intimate contact with materials containing lead.

Other sampling schemes logically follow. For instance, if examining copper pipe corrosion, discarding the first

Table 2. Interrelationships Among Pipe Length, ID, and Internal Volume for Selected Common Plumbing Materials and Pipe Sizes

Material	Identification/ Type	True ID	True OD	Length for 1,000 Milliliters (Feet)	Milliliters per Foot
Copper tubing	0.5-inch, type L, annealed	0.545	0.625	22	46
Copper tubing	0.5-inch, type L, drawn	0.545	0.625	22	46
Copper pipe	0.5-inch, schedule 40	0.622	0.840	17	60
Galvanized steel pipe	0.5-inch, schedule 40	0.616	0.840	17	59
Lead pipe or tube	0.5-inch ID, 0.25-inch wall	0.50	1.00	26	39
Lead pipe or tube	0.75-inch ID, 0.25-inch wall	0.75	1.25	11.5	87
PVC or CPVC pipe	0.5-inch, schedule 80	0.546	0.840	22	46

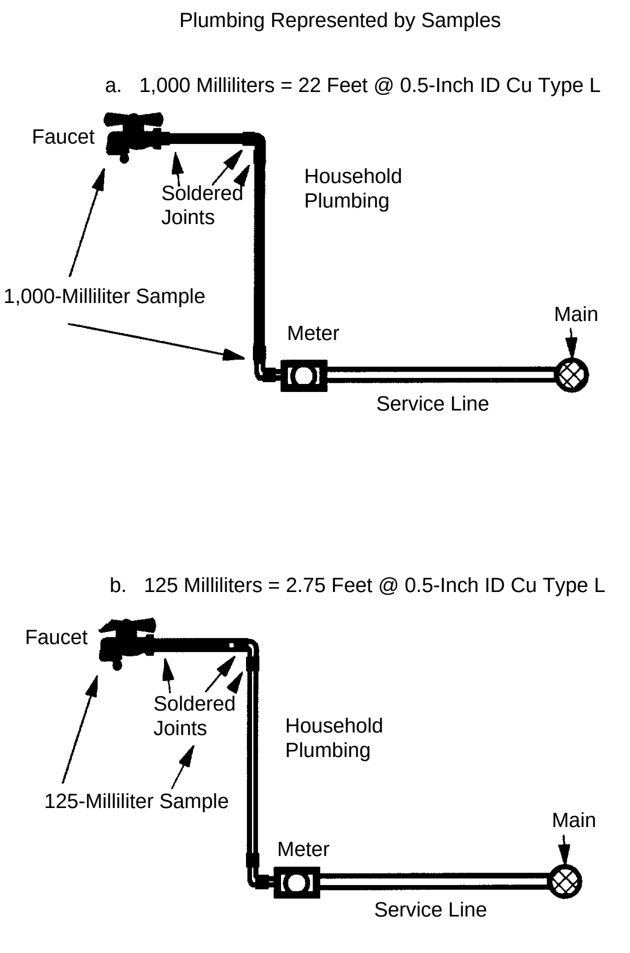


Figure 3. Schematic diagram of plumbing materials represented by sample volumes of a) 1 liter and b) 125 milliliters.

125 or 250 milliliters of water is likely to give more accurate information because it minimizes the effects of the faucet material as well as piping that connects the faucet to the interior line. Often, this connecting piping is not copper. If examining the corrosivity of the water to lead service lines, wasting a volume of water corresponding to the distance from the outlet to the service line better estimates the effect, although not without uncertainty (18). Many other sampling schemes are possible and useful, but users must be aware that the sampling protocol may have as much or more influence on the observed metal concentration than water quality or other variables. Hence, incorporation of monitoring data into a GIS database must be done only when the source represents equivalent samples.

Because of turbulent mixing during flow, local high concentrations of lead (or other contaminant) may become broadened and diluted by the time the water to be sampled reaches the sample collection bottle (18). In many cases, therefore, numerous small-volume sequential samples can be taken and used to profile a plumbing system to locate brass valves, connectors, soldered joints, etc. Figure 4 illustrates sequential sampling results for one room of a building. Peaks in the distribution of samples physically correspond to the location of a chrome-plated brass faucet and to a later concentration of fresh Sn:Pb soldered joints. Unfortunately, even small-volume sequential tap water samples must pass over other potentially contaminating or altering surfaces on the way through the sampling tap.

Kuch and Wagner have shown how water can dissolve large amounts of lead simply by traveling through long distances in lead pipes with small IDs (9, 19). Although this study specifically examined lead, the principle applies to other metallic piping materials. This phenomenon is inseparable from the aspect of time, which is the next subject.

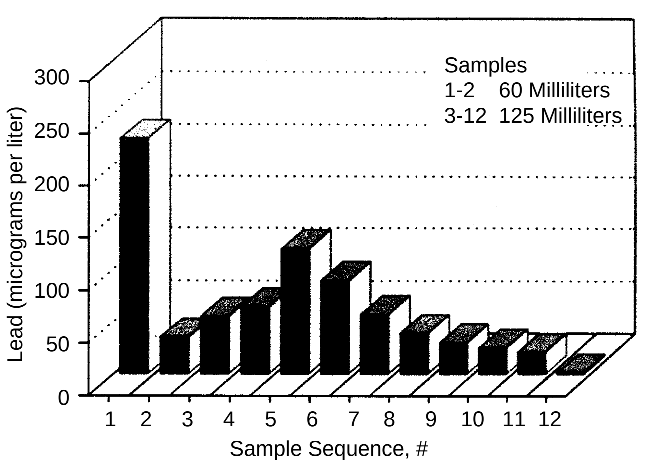


Figure 4. Sequential sampling results from a room on the ground floor of a building.

Effect of Time

While some chemical reactions are instantaneous, many dissolution and precipitation reaction steps that are important in controlling metal levels in water take many hours to many days to reach equilibrium. In fact, passivation films and scales on pipes that inhibit corrosion and reduce leaching of trace metals may take months to decades to develop substantially. Some other chemical transformations, such as creation of trihalomethanes from chlorination of natural organic matter during disinfection, or processes such as inactivation of pathogens, may occur over hours (20, 21).

Many steps in an overall chemical reaction process could be rate-limiting. Figure 5 shows how lead levels increase in 0.5-inch ID pipe given two different assumptions. The top curve shows how lead increases and levels off after about 8 to 12 hours (9, 19). This curve is closely applicable to any metal, as long as the limiting factor on the rate of metal migration into the water is the radial diffusion of the soluble metal species away from the pipe surface. The second curve shows, schematically, the effect of a diffusion barrier film (e.g., calcium carbonate, adsorbed iron hydroxide mixed with organic matter, aluminosilicate mineral deposits) or inhibition of metal oxidation rates on lead migration into water after different amounts of time.

In some water systems, significant chemical changes can occur while the water is standing that can drastically affect the oxidation or solubility of the plumbing material. For example, dissolved oxygen and free chlorine react quickly in new copper pipe or brass. If the water stands sufficiently long, their concentrations may become negligible, which would significantly alter the redox conditions governing metal solubility. In the absence of oxygen or chlorine species, the dominant form of copper in water and on plumbing material then becomes copper(I) instead of copper(II), resulting in different solubility characteristics after consumption of the oxidant than at initiation of the standing period (11).

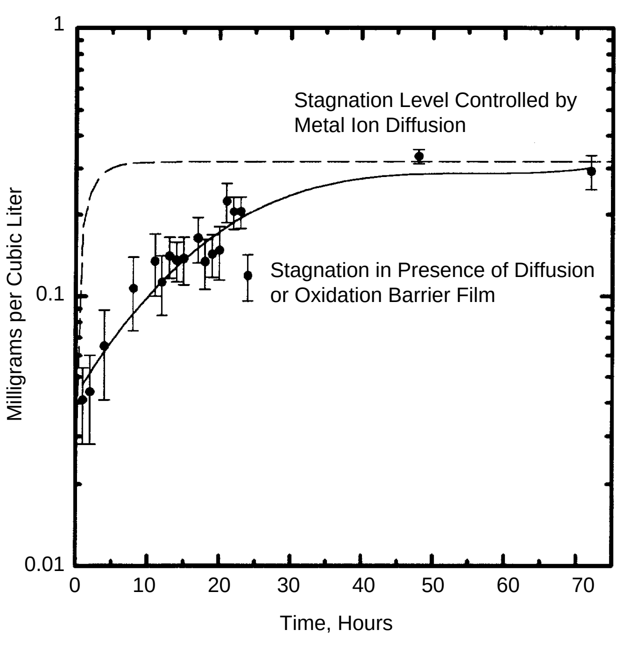


Figure 5. Comparison of lead concentrations that would be observed after water stands different amounts of time given different controlling chemistry factors.

Seemingly identical water samples collected from the same taps in houses, schools, or other buildings yield different metal concentrations, depending on the time the water was in contact with the faucet, solder, or piping material. Similarly, samples taken for disinfection byproducts after different chlorine contact times may produce different concentrations and different speciation (e.g., trihalomethanes, haloacetic acids). This factor causes considerable confusion in many investigations of contamination of school or building drinking water taps and water coolers and complicates estimating human exposure for health-effects studies.

Interconnectedness of Distance and Time

In innumerable situations, the effects of distance and time are impossible to separate. Some generalizations and examples follow.

Dead ends and slow rate areas produce long residence times for the distributed water. This results in long contact times with pipe materials, so reactive constituents can change considerably in concentration. The process is totally interactive, in that concentration changes of reactive constituents are in response to contact with the pipe materials, and in turn, the materials respond to the water composition. Water may take hours to days to reach a particular home or building and may traverse many miles of distribution system piping of the same or differing composition. Water thoroughly run through a household faucet for 5 or 10 minutes to purge the lines is “fresh” from the resident’s perspective, but may be “old” from the distribution system perspective.

The profile of the water line shown in Figure 4 was made on the basis of filling small volume (60 or 125 milliliters) sample bottles, one after another, without wasting any water. If the objective were to capture only the highest risk of lead contamination from a lead service line after some hours of stagnation, then the sampling process would be different. Instead of collecting all water between the tap and the service line, the water can be run until either a target volume is wasted (representing the linear plumbing distance to the service line) or can be run at a given rate until, after the appropriate length of time passes, the sample bottle can “intercept” the slug of water residing in the service line. Beware that differences may exist in the peak concentration of the contaminant and the “width” of the slug of elevated contaminant level, depending upon the rate of water flow before and during sampling (18, 22, 23).

Other Sources of Variability in Water Samples

Variability in water samples can stem from many sources aside from those discussed in this paper (18). The nature of the errors and their likely magnitude may vary with each episode of sampling and analysis and is far beyond the present scope of discussion. A brief listing to consider, however, when drawing conclusions from "field" data includes:

- Analytical imprecision or bias.
- Flow rate of the water during sampling.
- Temperature.
- Particulate erosion from plumbing materials during sampling.
- Effect of container material.
- Effect of air contact or other handling effects during sample collection and shipment.

A Case Study of Easy Misinterpretation

Interpretations of water quality problems based on aggregate monitoring data can be very misleading unless analysis is performed at the appropriate scale. The situation of one utility described below provides a good example of how using a GIS approach could have helped solve the problem but also highlights how carefully data would need to be matched and consolidated only at the proper scale if GIS were to be employed for evaluating some kinds of water quality problems.

The utility at Hopkinton, Massachusetts, found very high lead and copper levels exceeding the regulatory action levels under the Lead and Copper Rule (14-17). The 90th percentile copper levels even exceeded 6 milligrams per liter, compared with an action level requirement of only 1.3 milligrams per liter. Some sites with lead service lines are present in the system. Figure 6 shows a schematic representation of the distribution system for this utility. Five wells feed the system, and four (1, 2, 4, and 5) are used regularly.

The background hardness, alkalinity, and carbonate concentrations are fairly similar for all wells. The pH of the ground water from the wells is usually slightly above 6. Chlorine solution is dosed for disinfection. High iron levels are present in wells 1, 2, and 3, and high manganese is also present in well 3. Generally, high dosages of a polyphosphate chemical were added to wells 1 and 2 to respond to consumer complaints about the "red water" that results from iron oxidation and precipitation. Water from different wells mixes in the distribution system, but the water tends to partition into two zones as

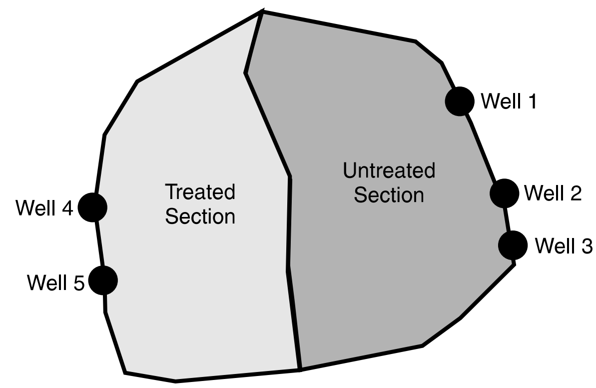


Figure 6. Schematic representation of the utility's distribution system.

Figure 6 indicates. The lead and copper levels tended to be distinctly lower in the section where the polyphosphate was dosed, marked as the "treated" part of the system. The ongoing research study has employed approximately 22 monitoring sites.

From the information presented thus far, the system clearly cannot be characterized by a discrete value for lead or copper contamination, as well as the chemical background of water throughout the system. Hence, putting data at the "whole system" scale into a statewide or countrywide data system would be tempting, but it could be very misleading in solving the treatment problem. Having accurate spatially distributed data for background water qualities, monitoring site characteristics, and metal levels at the subsystem scale, such as that which could be integrated into GIS, would have been extremely convenient, however. Yet, even more information at a smaller scale is necessary to understand and solve the whole treatment problem.

The utility initially observed that because the lowest lead and copper levels coincided with the area of the system fed by the polyphosphate chemical, that chemical likely caused the corrosion inhibition. Median lead levels, for example, were between about 200 and 300 milligrams per liter in the "untreated" section, compared with about 10 to 15 milligrams per liter in the "treated" section. Median copper levels were approximately 4 to 5 milligrams per liter in the untreated section, but only about 0.3 to 0.5 milligrams per liter in the treated section. The utility and the researchers wondered whether the polyphosphate chemical should also be added to the other wells. This is a matter of significant concern because some studies indicate polyphosphate chemicals can enhance lead corrosion (1, 6) and the subject has rarely been studied under statistically valid controlled conditions.

Additional site-by-site investigation, however, first revealed that the sites with lead service lines all lay in the untreated area of the distribution system. Because the research sampling program included two successive 1-liter samples, the additional contamination from the service lines was confirmed by higher lead levels in the second 1-liter sample than in the first in many cases. Therefore, physical reasons, in addition to chemical ones, explained the discrepancy in the lead levels. Further, when considering only the treated system sites, the lead levels were still high enough to be of concern.

Focusing on the copper sites resulted in the collection of more important and interesting small-scale information. Figure 7 shows the difference between average copper levels in the two sections of the system. Almost all sites in both parts of the system had copper interior plumbing with 50:50 or 60:40 Sn:Pb soldered joints and faucets with brass-containing internal materials. Though the chemical added for iron control was ostensibly a polyphosphate chemical, it also contained an initially present fraction of orthophosphate and also tended to partially break down to orthophosphate in the presence of iron and calcium (as most polyphosphates do). Figure 8 shows the orthophosphate concentrations in the two different parts of the system. While the levels of orthophosphate present in the treated section would be far too low to significantly inhibit lead leaching at the background pH (1, 6, 7, 9, 22), the orthophosphate plausibly may significantly inhibit copper dissolution, in concordance with recent research projections (11).

Having determined through detailed small-scale sampling and analysis that the chemistry affecting metal levels in the system is generally consistent with modern knowledge, a new treatment plan is being implemented to control copper and lead levels through pH adjustment in conjunction with iron control through a compatible sodium silicate/oxidation treatment. Incorporation of system and monitoring site physical characteristic data, plus monitoring results, into GIS could have saved considerable investigatory effort. The importance of this case history, however, is that the data must be of the appropriate scale and highly documented to be useful in problem-solving. Failure to use data meeting these requirements, as well as overgeneralization to a large mapping scale, can lead to ineffective if not damaging water treatment choices that could adversely affect public health.

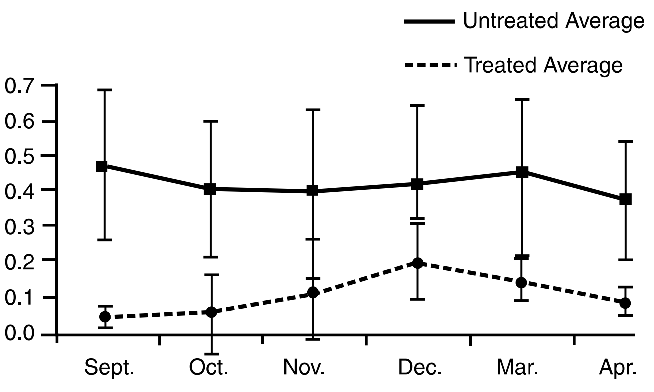


Figure 7. Average copper levels in treated and untreated sections of the system.

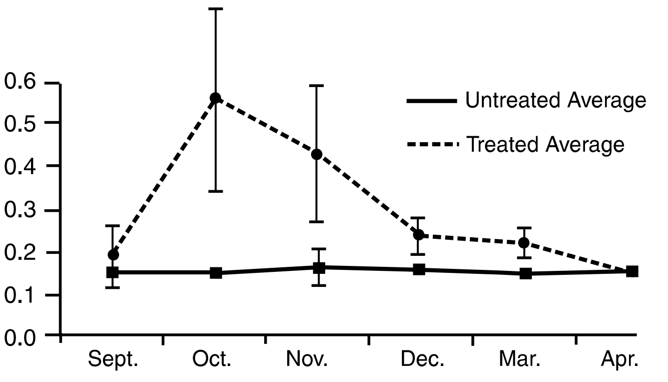


Figure 8. Average orthophosphate levels in treated and untreated sections of the system.

Conclusions

The examples and discussion above lead to several general conclusions about the use of GIS with drinking water monitoring data:

- Temporal and spatial variability stems from many causes, down to a very small scale.
- Sampling protocols must be keyed to the precise questions under investigation.
- Regulatory sampling, whose results are generally readily available, is usually inappropriate to assess human exposure to trace metals or other parameters of interest (such as DBPs).
- Generalizations on a large scale are often impossible because of the geology and water chemistry variations.

Additionally, some considerations apply to the types of mapping that could be employed by GIS. For example, a mapping technique such as contouring may be especially inappropriate for use with drinking water data. Major problems could result from:

- Discrete, small-scale (such as within an individual house) variability in distributions of certain contaminants, such as lead and copper.
- Physical constraints of the distribution system network.
- The small number of monitoring sites in relation to the size of the distribution network.
- Different chemical or hydraulic zones in the distribution system.

Employing GIS could be very useful in solving a variety of drinking water problems. Users must be extremely conscious of the nature of the source information, however, to avoid abusive extrapolations and generalizations.

Acknowledgments

Some of the examples cited were investigated as part of a U.S. Environmental Protection Agency cooperative agreement (Darren A. Lytle, Project Officer) with the New England Water Works Association. Jack DeMarco of the Cincinnati Water Works kindly provided Figure 1 on distribution of sampling sites for lead and copper monitoring.

References

1. Schock, M.R. 1990. Internal corrosion and deposition control. In: Water quality and treatment: A handbook of community water supplies, 4th ed. New York, NY: McGraw Hill.
2. Benefield, L.D., and J.S. Morgan. 1990. Chemical precipitation. In: Water quality and treatment: A handbook of community water supplies, 4th ed. New York, NY: McGraw Hill.
3. Glaze, W.H. 1990. Chemical oxidation. In: Water quality and treatment: A handbook of community water supplies, 4th ed. New York, NY: McGraw Hill.
4. Cotruvo, J.A., and C.D. Vogt. 1990. Rationale for water quality standards and goals. In: Water quality and treatment: A handbook of community water supplies, 4th ed. New York, NY: McGraw Hill.
5. Tate, C.H., and K.F. Arnold. 1990. Health and aesthetic aspects of water quality. In: Water quality and treatment: A handbook of community water supplies, 4th ed. New York, NY: McGraw Hill.
6. Schock, M.R. 1989. Understanding corrosion control strategies for lead. JAWWA 81:7:88.
7. AWWARF. 1990. Lead control strategies. Denver, CO: American Water Works Association Research Foundation and American Water Works Association.
8. AWWARF. 1994. Corrosion control in water distribution systems, 2nd ed. Denver, CO: American Water Works Association Research Foundation/Engler Bunte Institute.
9. Schock, M.R., and I. Wagner. 1985. The corrosion and solubility of lead in drinking water. In: Internal corrosion of water distribution systems. Denver, CO: American Water Works Association Research Foundation/DVGW Forschungsstelle.
10. Pankow, J.F. 1991. Aquatic chemistry concepts. Chelsea, MI: Lewis Publishers, Inc.
11. U.S. EPA. 1995. Effect of pH, DIC, orthophosphate and sulfate on drinking water cuprosolvency. Cincinnati, OH: Office of Research and Development. In press.
12. AWWA. 1993. Initial monitoring experiences of large water utilities under U.S. EPA's Lead and Copper Rule. Denver, CO: Water Industry Technical Action Fund/American Water Works Association.
13. AWWA. 1990. Lead service line replacement: Benefit to cost analysis. Denver, CO: Water Industry Technical Action Fund/American Water Works Association.
14. AWWA. 1991. Lead and copper: A working explanation of the Lead and Copper Rule. Denver, CO: American Water Works Association.
15. Federal Register. 1991. Lead and copper: Final rule correction. Fed. Reg. 56:135:32,112 (July 15).
16. Federal Register. 1991. Lead and copper: Final rule. Fed. Reg. 56:110:26,460 (June 7).
17. Federal Register. 1992. Lead and copper: Final rule correction. Fed. Reg. 57:125:28,785 (June 29).
18. Schock, M.R. 1990. Causes of temporal variability of lead in domestic plumbing systems. Environ. Monit. Assess. 15:59.
19. Kuch, A., and I. Wagner. 1983. Mass transfer model to describe lead concentrations in drinking water. Water Res. 17:10:1,303.
20. Ireland, J.C. 1993. Alternatives to chlorine for disinfection of drinking water. In: Strategies and technologies for meeting SDWA requirements. Lancaster, PA: Technomic.
21. Miltner, R.J. 1993. Pilot scale treatment for control of disinfection byproducts. In: Strategies and technologies for meeting SDWA requirements. Lancaster, PA: Technomic.
22. Sheiham, I., and P.J. Jackson. 1981. Scientific basis for control of lead in drinking water by water treatment. J. Inst. Water Engrs. and Scientists 35:6:491.
23. Heumann, D.W. 1989. Los Angeles Department of Water and Power: Solid lead gooseneck slug dispersion in consumer plumbing system. Proceedings of the AWWA Water Quality Technology Conference, Philadelphia, PA.

Evaluating Soil Erosion Parameter Estimates From Different Data Sources

G. B. Senay[1], B. Subramanian[2] and S. Cormier[2]
[1] SAIC/PAI, c/o USEPA, Cincinnati, OH
[2] US EPA, Cincinnati, OH

Abstract

Topographic factors and soil loss estimates that were derived from three data sources (STATSGO, 30-m DEM, and 3-arc second DEM) were compared. Slope magnitudes derived from the three data sources were consistently different. Slopes from the DEMs tended to provide a flattened surface with large areas having 0.0% values. The 3 arc second DEM generally produced a lower slope estimate than either the STATSGO or 30-m DEM. Slopes from the 30-m DEM fell between the 3-arc second and the STATSGO. Thus, the STATSGO database provided a higher slope estimate than either of the two DEM sources. However, slopes from the 30-m DEM and the STATSGO were more comparable than the slope estimates from 30-m and 3-arc second DEMs. For example, 0.0, 10.0, and 20.0% slope classes from the 30-m DEM showed mean values of 0.0, 4.0, and 6.0 %, respectively with the 3-arc second DEM, and 3.0, 18.0, and 24.0 %, respectively with the STATSGO. Along with the slope differences, potential erosion estimate trends varied between the data sets although the soil loss differences were higher than the slope differences. A recommendation was made to validate slope and erosion estimates using field data; however, it appeared that STATSGO may be more reliable than the two data sets for smaller slopes and either STATSGO or the 30-m DEM may be used for higher slopes. Due to the ease of GIS in combining data from various sources, the importance of a thorough understanding of data accuracy standards and their limitations and intended use were highlighted.

Introduction

One of the most difficult challenges in ecological hydrology is the paucity and accuracy of hydrologically relevant data on factors such as soils, topography, vegetation and stream ecology. The available data sources vary in their accuracy, accessibility, and data formats.

Stream ecology can be affected by pollution caused by the process of erosion and sediment transport. Erosion estimates are generally made using the Universal Soil Loss Equation (USLE) from the multiplicative nature of soil erodibility, topography, landcover, conservation practice and rainfall erosivity factors. Unlike the other factors, the topographic factor can be readily derived from different digital sources called Digital Elevation Models (DEMs) each with its own accuracy standard. Another widely used data set for this kind of study is the STATSGO (State Soil Geographic) data base that is distributed by the Natural Resources Conservation Service (NRCS).

Digital Elevation Model (DEM) data files are digital representations of cartographic information in a raster form. DEMs consist of a sampled array of elevations for a number of ground positions at regularly spaced intervals. These digital cartographic/geographic data files are produced by the U.S. Geological Survey (USGS) as part of the National Mapping Program and are sold in 7.5-minute (also known as 30 -m DEM), 15-minute, 2-arc-second (also known as 30-minute), and 1-degree units (3-arc second DEM). The 7.5- and 15-minute DEMs are included in the large scale category while 2-arc-second DEMs fall within the intermediate scale category and 1-degree DEMs fall within the small scale category (USGS, 1987).

The accuracy of a DEM is dependent upon its source and the spatial resolution (grid spacing) of the data profiles. One factor influencing DEM accuracy is source data scale and resolution. A dependency exists between the scale of the source materials and the level of grid refinement possible. The source resolution is also a factor in determining the level of content that may be extracted during digitization. For example, 1:250,000-scale topographic maps are the primary source of 1-degree DEMs while 1:24,000-scale maps are the primary source for 7.5 minute DEMs (USGS, 1987).Within a standard DEM, most terrain features are generalized by being reduced to grid nodes spaced at regular intersections in the horizontal plane. This generalization reduces the ability to recover positions of specific features less than the internal spacing and results in a de facto filtering or smoothing of the surface during gridding (USGS, 1987) .

STATSGO is a nationwide soils database that consists of soil maps and attributes that are assembled by generalizing more detailed soil survey data in accordance with the digitizing standards of the Natural Resources Conservation Service (NRCS). Soil maps were digitized by line segment format (vector) from a U. S. Geological Survey topographic quadrangle base map with a scale of 1:250,000. This level of mapping is designed to be used for broad planning and management uses covering state, regional, and multistate areas. The mapping area of the smallest soil polygon is about 1, 544 acres, i.e., soil map units less than this area are not represented in STATSGO (USDA, 1994).

Each STATSGO map is linked to the Soil Interpretations Record (SIR) attribute database. The attribute database gives the proportionate extent of the component soils and their properties for each map unit. The STATSGO map units consist of 1 to 21 components each. The Soil Interpretations Record database includes over 25 physical and chemical soil properties, interpretations, and productivity. Information that can be queried from the database include: land slope, available water capacity, soil organic matter, salinity, flooding, water table, bedrock, and interpretations for engineering uses, cropland, woodland, rangeland, pastureland, wildlife, and recreation development.

The objective of this study was to compare topographic factors and soil loss estimates that are derived from three data sources namely STATSGO, 30-m DEM (7.5 minute data), and 3-arc second DEM (1 degree data). The three spatial data sources differ in data collection techniques, spatial accuracy, and data formats.

Material and Methods

Study Site

This study compares spatial data sets for the state of Ohio. Although the two data sets (STATSGO and 3-arc second DEM) had complete coverage for the entire state, the 30-m DEM only covered a portion of the state since the rest of the area was still under “work-in-progress” by the USGS. The total area of Ohio is around 120,362.00 km^2 of which the 30-m DEM covered about 56%.

STATSGO

For this study, STATSGO data for the state of Ohio was downloaded from the NRCS web site. The file was formatted for use in a Geographical Information Systems software called Arc/Info 7.0 (ESRI, 1997). The spatial data was projected in the Albers Equal Area coordinate system.

Two soil/landscape parameters that are essential for soil erosion modeling were derived from the STATSGO database. These include the soil erodibility (K) and percent land slope (S). Since the attribute database is more detailed than the spatial mapping, the component percentage in each map unit was used to produce an area weighted average value that represents the various map components in a given spatial polygon (map unit). The attribute database contains information about soil components while the spatial coverage only identifies map units that are collections of soil components. Thus, the STATSGO attribute database and spatial map units were joined by a map unit ID called MAPUNIT in Microsoft Access, a relational database (Microsoft Corporation, 1996). Component percent values were used as a weighting factor to calculate the weighted average for each soil polygon. In STATSGO, generally, each soil parameter with a numerical value is represented by two values namely, the “low” and “high” range found in each soil component. Once the weighted average was calculated for the individual “low” and “high” range values, the average of the “high” and “low” was taken to represent a particular soil polygon (map unit).

In order to process the STATSGO arc coverage file in a grid based (raster) GIS (Arc/Info:Grid), together with the DEMs and for convenience during the USLE calculation, the vector data was rasterized in Arc/Info using the POLYGRID command at 90 m resolution, closer to the spatial resolution of the 3-arc second DEM. A finer resolution was not chosen in order to avoid increased file sizes and computation time. It was believed that there would not be a loss of significant information by rasterizing the data at 90 m since STATSGO is relatively coarse data with the smallest polygon area occupying 635 hectares which is equivalent to about 772 cells at the 90-m grid spacing.

<u>DEM</u>

For this study, land slope was also calculated from two Digital Elevation Models (DEMs) with spatial resolutions of 30-m and 3-arc second. The 30-m DEM (7.5 minute) data correspond to the USGS 1:24,000 and 1:25,000 scale topographic quadrangle map series for all of the United States and its territories. The 3-arc second DEM (1-degree) provides coverage in 1- by 1-degree blocks for all of the contiguous United States. The basic elevation model is produced by or for the Defense Mapping Agency (DMA), but is distributed by the USGS, in DEM data record format. The 3-arc second DEM data was available after being resampled to a uniform grid spacing of 102.0815 m in Albers Equal Area coordinate system.

While the 3-arc second DEM was readily available for the state of Ohio in Albers projection, the 30-m DEM required some pre-processing before it was used together with the other data layers. The 30-m DEM were purchased from the USGS in separate 7.5 by 7.5 minute topographic quadrangle sizes in Universal Transverse Mercator (UTM) projection. Using an automated procedure in Arc/Info's Arc Macro Language (AML), the individual data files (507) were first projected to an Albers projection, and elevation values were checked and converted to meters if they were not already in meters. Finally, the individual grids were mosaicked using the MOSAIC command in Grid-Arc/Info.

Generating Topographic Parameters

Land slope is a very important parameter in hydrological processes. Slope plays an important role in regulating the rate and amount of water flow on land surfaces. Particularly, slope is embedded in one of the multiplicative variables (topographic factor, LS) in the Universal Soil Loss Equation (USLE). The USLE is widely used to estimate the average annual soil loss from agricultural, range and forest lands. Wischmeier and Smith (1978) formulated the USLE as follows:

$$A = R * K * LS * C * P \tag{1}$$

Where,

A = average annual soil loss in tons/acre

R = rainfall and runoff erosivity index for a geographic location

K = soil erodibility factor

LS = slope steepness and length factor

C = cover management factor

P = conservation practice factor

While there are limited sources to obtain the values of most of the variables, the availability of DEMs at different accuracy levels and GIS overlay techniques allow the determination of the LS factor from DEMs plus STATSGO. Generally, the R factor is determined from weather data, requiring inputs of amount and pattern of rainfall for the area; the K factor is largely dependent on soil types with, for example, soils with high silt content being the most erodible. STATSGO provides K values for each soil polygon. The cover management factor (C) is generally dependent upon crop rotations and tillage sequences. For forest, rangeland and non-agricultural conditions, C is estimated from density of vegetation and the amount of residue on the soil surface. The P factor mainly accounts erosion control practices such as contouring, terracing and strip cropping whose effect vary with the land slope. For many applications of the USLE, no erosion control practice will be used, and the P factor will be just 1.0 (Ward and Elliot, 1995).

The objective of this study was to investigate the effect of using three data sources in the estimation of the LS factor and to evaluate its impact on the estimation of the average soil loss. The three data sources were: STATSGO, 30-m DEM and 3-arc second DEM.
The LS factor is a function of both the length (L) and the steepness of the land (slope, S). In most field applications, both factors are considered as a single topographic factor. LS is the expected ratio of soil loss per unit area from a field slope to that from a 72.6 ft length field of uniform 9-% slope under otherwise identical conditions.

The LS factor is calculated from the following formula (Wischmeier and Smith, 1978):

$$LS = (8 / 72.6)^m * (65.41 * \sin^2 2 + 4.56 * \sin 2 + 0.065) \qquad (2)$$

Where,
8= slope length in feet; 2= angle of slope; and
m = 0.5 if the percent slope is 5 or more; 0.4 on slopes of 3.5 to 4.5 percent; 0.3 on slopes of 1 to 3 percent; and 0.2 on uniform gradients of less than 1 percent.

For STATSGO soil polygons, slope was obtained from the attribute database. As described earlier, the component percentages in each soil polygon were used to calculate area weighted slope values for each soil map unit. The average of the ‘low’ (slope_low) and ‘high’ (slope_high) range values were used to represent each soil polygon.

For the DEM data sources, slopes were calculated using the SLOPE function in Grid Arc/Info. Slope identifies the maximum rate of change in value from each cell to its neighbors. An output “slope” layer was calculated as percent slope. Conceptually, the slope function fits a plane to the “z: elevation” values of a 3x3 cell neighborhood of a center cell. The slope for the cell is calculated from the 3x3 neighborhood using the average maximum technique (Burrough, 1986).

$$\%\text{-slope} = 100 * ((dz/dx)^2 + (dz/dy)^2)^{0.5} \qquad (3)$$

Where the deltas (differences) are using a 3x3 roving window; a through i represent the z values (elevation) in the window shown below:

a b c
d e f
g h i

$$dz/dx = (a + 2d + g) - (c + 2f + i) / (8 *x\text{-grid\_spacing}) \quad (4)$$

$$dz/dy = (a + 2b + c) - (g + 2h + i) / (8 * y\text{-grid spacing}) \quad (5)$$

Since both the x- and y-grid spacing are the same, the grid spacing can be replaced by the spatial resolution of the grid, i.e., simply 30-m for the 30-m DEM.

Once the slope layers were calculated from each DEM grid, the LS factor was calculated using Equation 2. A procedure was written in Arc Macro Language (AML) to execute equation 2 for each grid cell. The procedure mainly checks for the magnitude of the slope of a particular cell in order to vary the exponent 'm' in equation 2 accordingly. The length of the slope was taken to be equal to the length of a cell. For calculating LS from STATSGO, the rasterized, 90-m STATSGO layer was used; similarly, equation 2 was used to generate the LS values.

For estimating the average annual soil loss, the following assumptions and approximations were applied. First, the only variables varying from cell to cell were the LS and K factors. The rainfall and runoff erosivity index (R) was read from a nation wide chart by Wishemeier and Smith (1978). A value of 150 seemed a reasonable representative value for most of Ohio. On the other hand, a cover factor (C) of 0.42was used as a conservative estimate for corn and soybean fields that are managed in a conventional tillage practice in autumn (Cooperative Extension Service, OSU, 1979).A conservative value of 1.0 was assumed for the conservation factor for lack on any information on the application of any soil conservation measures. Although such estimates may not represent the actual soil loss on a cell by cell basis, the most important exercise in the study was to show the relative differences in such estimates due to using different data sources particularly for the topographic factor, LS.

The average annual soil loss was simply estimated by using equation 1 in Grid: Arc/Info. To develop the 'soil loss' GIS layer, two GIS layers (LS, K factors) and two constants (R and C) were multiplied on a cell by cell basis. Three soil loss layers were derived from each of the data sources which produced varying LS factors. Note that the K factor from STATSGO was also applied on the soil loss estimates from the DEM sources. This was possible since all the three data sources were projected to the same coordinate system in Albers Equal Area.

Data Manipulation and Extraction in Arc/Info

It was possible to estimate the average soil loss for the entire state of Ohio using STATSGO and a combination of STATSGO and the 3-arc second DEM. Only about 56% of the state could be evaluated using a combination of STATSGO and the 30-m DEM. STATSGO remains essential in all data sets since it also provides the K factor. Due to the limited area of coverage and its relative spatial accuracy, the 30–m DEM was used as reference data to which the results of the other two data sources were compared.

A recommended use of the LS equation is in land slopes not exceeding 20% (Wischmeier and Smith, 1978). A mask (20%-mask) was created to remove all areas with slopes greater than 20% in the 30-m DEM before any overlay analysis was made with the STATSGO and 3-arc second DEM data. The combined area of slopes with magnitudes greater than 20% covered 14.4 % of the available 30-m DEM coverage.

After creating the 20%-mask on the 30-m DEM, floating point slope values were rounded to the nearest integer so that distinct slope classes would be established based on which slope values from the other data layers can be queried and analyzed. For each of the slope classes, i.e., 1, 2, 3,...,20, the corresponding mean and standard deviation slope values were derived from STATSGO and the 3-arc second DEM using the ZONAL STATS command in Grid Arc/Info. The extracted slope values were again rounded to the nearest integer values.

Similarly, soil loss estimates from the three data sources were compared using a soil loss layer derived from the 30-m DEM as a reference. Only slopes less than or equal to 20% were used for the soil loss estimation as well. Soil loss estimates were first rounded to the nearest whole number for the reference data set (that derived from the 30-m DEM). Then, for each soil loss class (tons/ac/yr) the corresponding mean and standard deviation soil loss estimates were derived from the other soil loss estimate layers, i.e., those based on STATSGO and 3-arc second DEM. The mean soil loss estimates for the other layers were then rounded to the nearest whole number.

Scatter plots of class-mean slope and soil loss estimates were plotted using Excel, a spreadsheet and graphing software (Microsoft Corporation, 1996). For both cases, the estimates derived from the 30-m DEM were placed in the x-axis (the reference) and estimates from STATSGO and 3-arc second were placed in the Y-axis. Relative spreads of the data were quantified by the standard deviation and coefficient of variation of the class values.

Results

Slopes

It was shown that slopes derived from the three data sources were consistently different. The 3-arc second DEM showed a generally lower slope estimate than the STATSGO or 30-m DEM. Slopes from the 30-m DEM fell between the 3 arc second DEM and the STATSGO (Table 1, Figure 1). Thus, the STATSGO database generally provided a higher slope estimate than either of the two databases.

Generally, lower slope values occupied most of the area with higher slopes constituting a smaller percentage of the area. The most abundant slope was the 0.0% in both the 30-m and 3-arc second DEM, covering about 18% of the 30-m DEM coverage (56% of Ohio) and 44.5% of the study area, respectively. In STATSGO, the most abundant slope was the 2% slope covering 17% of the study area followed by the 4% slope and 1% (minimum for STATSGO) slope covering 11.4% and 10.6% of the study area, respectively. The maximum slope values were 245%, 61% and 41% for 30-m DEM, 3-arc Second DEM and STATSGO, respectively. However, the maximum slope values from the DEMs were more likely to represent outliers since they were only observed in one pixel. For STATSGO, the maximum slope covered about 0.9% of the study area. An upper quartile comparison of the slope distribution showed that about 75% of the area in each data set showed that slope values were less than or equal to 14%, 3%, and 16% for 30-M DEM, 3 arc second DEM, and STATSGO, respectively (data not shown).

Slopes from the 30-m DEM and the STATSGO appeared more comparable than the slope estimates from 30-m and 3 arc second DEMs, particularly in higher slope classes. For example, 0.0, 10.0, and 20.0% slope classes from the 30-m DEM showed mean values of 0.0, 4.0, and 6.0 %, respectively, with the 3-arc second DEM, and 3.0, 18.0, and 24.0 %, respectively with the STATSGO based data (Table 1, Figure 1).

One of the unique characteristics of the STATGO database is that the average between the high and low values will result in a condition where a zero data point will be eliminated. That is one of the reasons why the STATSGO minimum value does not start at zero. In the case of the 3-arc second DEM, the flat lands were accurately represented. On an average, slopes labeled as having 0.0% in the 30-m DEM were also labeled as having 0.0% slopes in the 3-arc second data set. This unique match in the two data sets can be explained by the general tendency of the 3-arc second to provide a flattened surface. However, the inverse was not true (data not shown) in that slopes identified as having 0.0% slopes in the 3-arc second DEM were not necessarily 0.0% in the 30-m DEM. This was because the 30-m DEM had a better capacity for resolving smaller slopes than the 3-arc second DEM.

Due to the generally depressed slopes observed from the 3-arc second DEM, the data also appeared less sensitive compared to the 30-m DEM, resulting in a step function pattern (Table1; Figure 1). For example, slopes varying from 1 to 3% in the 30-m DEM were labeled as 1% in the 3-arc second DEM. Similar effects are seen in other slope ranges such as6 to 9-% which were labeled as having 3% slope in the 3-arc second data set. A similar trend was observed in the STATSGO data, but only at the higher slope ranges. For example, slopes 13 and 14% were labeled as 21% and slopes ranging from 17 to 19% were labeled as 23 %.

A detailed investigation of the 30-m DEM with only slopes less than or equal to 20% showed that most of the study area was under smaller slopes with about 63 % of the total area having a slope of 5 % or lower (Table 1). Areas with 0.0% slope occupied the largest area with about 19% coverage, percentage slope area decreased with increasing slope (Table 1). The relatively large percentage of land areas with 0-% slopes indicates that even the 30-m DEM may not be adequate if a more accurate representation of a field is required in relatively flat to mild topography.

The spread of the slope values in each 30-m DEM slope-class shows that the 3-arc second DEM was more spread than the STATSGO data set (Table 1); with coefficient of variation (the ratio of standard deviation to mean) values ranging from 100 to 150% for 3-arc DEM and only from 37 to 125% for STATSGO.

Soil Loss Estimate

The differences in soil loss estimates between the three data sets were generally a reflection of differences in slope and slope length estimates since the topographic factor was the major difference between the different data sources. The combination of slope values and grid cell sizes (for slope length) were used to calculate the LS factor in the USLE equation. While the ® = 150) and © = 0.42) factors were constant, the “K” factor was spatially variable as obtained from the STATSGO data set. The “K” values had a minimum and maximum of 0.09 and 0.45, respectively, with a mean value of 0.348 for the entire state of Ohio.

From the 30-m DEM, the USLE calculation resulted in 119 soil loss classes ranging in values from 0 to 118 tons/ac/yr. The soils estimates are generally larger (up to 10 times) than is observed in agricultural fields and generally greater than the allowable soil tolerance limits that range from 3 to 20 tons/ac/yr, depending on the soil types (Ward and Elliot, 1995). This was mainly due to the assumption of a poor cover condition (high C factor) at the worst scenario. Table 2 shows soil loss classes in increments of 10 after the 10th class. The 0.0 soil class is due to rounding and it only represented values under < 0.5 tons/ac/yr that were obtained in a relatively small area (0.06%). As the slope extraction, the 30-m DEM soil classes were used to extract soil loss estimates from the 3-arc second DEM and STATSGO data sets (Table 2). The 3-arc DEM compared well with the 30-m DEM estimates on lower soil loss estimates up to 10 tons/ac/yr before beginning to underestimate by as much as about 3 times lower at higher soil loss estimates (Table 2, Figure 2). The reason for the better comparison between the two data sets at lower soil loss estimates was due to relatively small slope difference at the lower erosion estimates. For example, slopes 1 to 3 in the 30-m DEM were averaged as 1% in the 3-arc second DEM. Also, the LS factor tends to compensate (higher) for longer slopes as exhibited in the 3-arc second data set (102 m vs 30 m). Taking average conditions for K, R, and C factors, 10 tons/ac/yr corresponds to about 4% slope for the 30-m DEM. On an average, the 4% areas on the 30-m DEM are labeled as having 2% and 10% slope on 3-arc second DEM and STATSGO data sets, respectively (Table 1).

Taking into consideration the sensitivity (exponential type relationship) of the LS factor (Equation 2) on slope, erosion estimates generally will be increased more than the increase in slope. For example, the LS factor will be about 6 times higher when a 4% slope of 30-m DEM is replaced by a 10% slope from STATSGO. Thus, soil estimates from STATSGO were generally much higher than the 30-m DEM, reaching up to about 7 times higher (Table 2, Figure 2).This is certainly the result of the higher slopes observed in the STATSGO database even at lower slopes when compered to the 30-m DEM. In addition, the slope length of the STATSGO was three times higher than the 30-m DEM which would also lead to higher soil loss estimates.
Soil loss trends for both data sets (3-arc second, STATSGO) were similar to the increasing values from the 30-m DEM for most of the data set (Figure 2). Irregular trends were observed at higher soil loss estimates (from 90 to 118 tons/ac/or). This was due to the fewer number of grid

cells that were being averaged at high estimates. Table 2 shows that about 95% of the area had soils estimate values less than or equal to 90 tons/ac/or indicating that soil loss estimates beyond 90 tons/ac/or were found in smaller areas. This shows that although this kind of multi-source comparison seems to be reasonable at a regional scale, conclusions from such studies can be misleading in applying these studies at a local scale in smaller areas.

Conclusions

Slope and soil loss estimates from the three data sources varied greatly. Slope and soil loss estimates were ranked from high to low for STATSGO, 30-m DEM, and 3-arc second DEM in descending order. A large percentage (44.5%) of the 3-arc second DEM coverage produced slope values of 0.0%. While the 30-m DEM also contained a relatively large area (about 18%) under 0.0% slope, STATSGO showed that the 2% slope was dominant (about 17% of area) followed by 4 and 1% slope, each occupying about 11% of the area. The high percentage of 0.0% slopes in the DEM data sets indicated that these data sets were unable to resolve small relief differences between adjacent cells.

Slopes from the 30-m DEM and the STATSGO were more comparable than the slope estimates from 30-m and 3-arc second DEMs, particularly at higher slope values. This indicated that although both DEMs tend to underestimate slopes, the 30-m DEM was more reliable than the 3-arc second DEM at higher slopes.

Along with the slope differences, potential erosion estimates varied between the three data sources. Soil loss differences between the data sets were higher than the slope differences. An irregular trend in erosion estimates at the higher end was attributed to smaller class sizes, indicating the limitations of the this kind of study at smaller scales while suggesting its validity on larger scales.

A recommendation was made to validate the slope and erosion estimates using field data. However, it appears that STATSGO may be more reliable than the two other data sets for smaller slopes and either STATSGO or the 30-m DEM may be used for higher slopes. Further investigation will be necessary to recommend the applications and limitations of these data at different scales. This study showed that the three data sources resulted in widely varying estimates of topographic factors and soil loss estimates. Although GIS makes it easy to combine data from various sources, the compatibility of the data accuracy standards and their limitations and intended use should be thoroughly understood before scientific conclusions are drawn.

Table 1: Slope distribution (%-slope) of 3-arc second DEM and STATSGO based on 30-m DEM slope classes.

Class 30-m	Mean 3-arc	Mean STAT	Std. 3-arc	Std. STAT	Cum. %-Area
0	0	3	1	3	18.9
1	1	4	1	5	31.0
2	1	5	2	6	46.2
3	1	8	2	8	52.9
4	2	10	3	9	58.5
5	2	11	3	10	63.2
6	3	13	4	10	66.6
7	3	15	4	11	70.1
8	3	16	4	11	73.5
9	3	17	4	11	76.0
10	4	18	5	10	78.9
11	4	19	5	10	81.6
12	4	20	5	10	84.0
13	4	21	5	10	86.3
14	4	21	5	10	88.6
15	5	22	5	10	90.8
16	5	22	5	10	92.7
17	5	23	6	9	94.7
18	5	23	6	9	96.7
19	5	23	6	9	98.3
20	6	24	6	9	100.0

30–m: 30-M DEM; 3-arc: 3-arc second DEM; STAT: STATSGO; Std.: Standard deviation; Cum.: Cumulative area occupied by the 30-m DEM class values.

Table 2: Soil loss estimate (tons/ac/or) distribution of 3-arc second and STATSGO based on 30-m DEM slope classes.

Class 30-m	Mean 3-arc	Mean STAT	Std. 3-arc	Std. STAT	Cum. %-Area
0	1	7	1	0	0.1
1	2	7	3	21	11.4
2	3	17	5	35	24.5
3	4	22	8	53	32.2
4	4	25	9	51	41.5
5	6	39	12	69	46.7
6	7	52	15	86	49.6
7	8	51	15	80	52.5
8	8	58	16	89	55.1
9	9	70	18	96	57.2
10	10	72	19	96	58.8
20	16	130	27	123	71.5
30	20	160	31	122	77.9
40	23	179	34	120	83.3
50	25	183	37	116	87.8
60	28	196	39	112	91.4
70	30	206	43	114	94.5
80	30	205	42	113	97.1
90	32	238	48	108	99.0
100	37	175	54	81	99.7
110	18	20	22	19	100.0
118	5	18	2	0	100.0

30-m: 30-m DEM; 3-arc: 3-arc second DEM; STAT: STATSGO; Std.: Standard deviation; Cum.: Cumulative area occupied by the 30-m DEM class values.

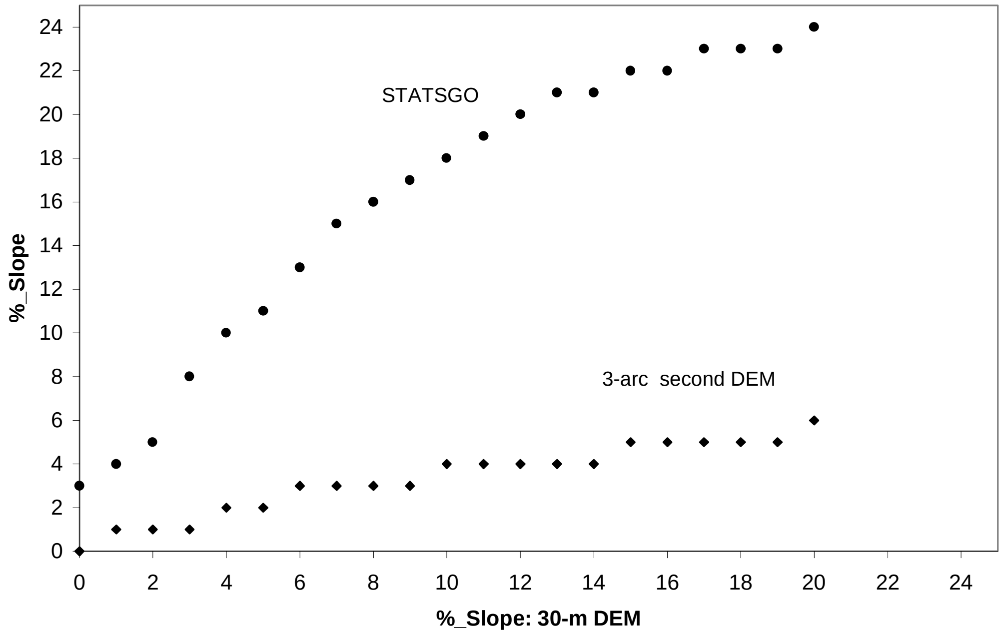


Figure 1: Percent land slope comparison between three data sources with 30-m DEM as a reference.

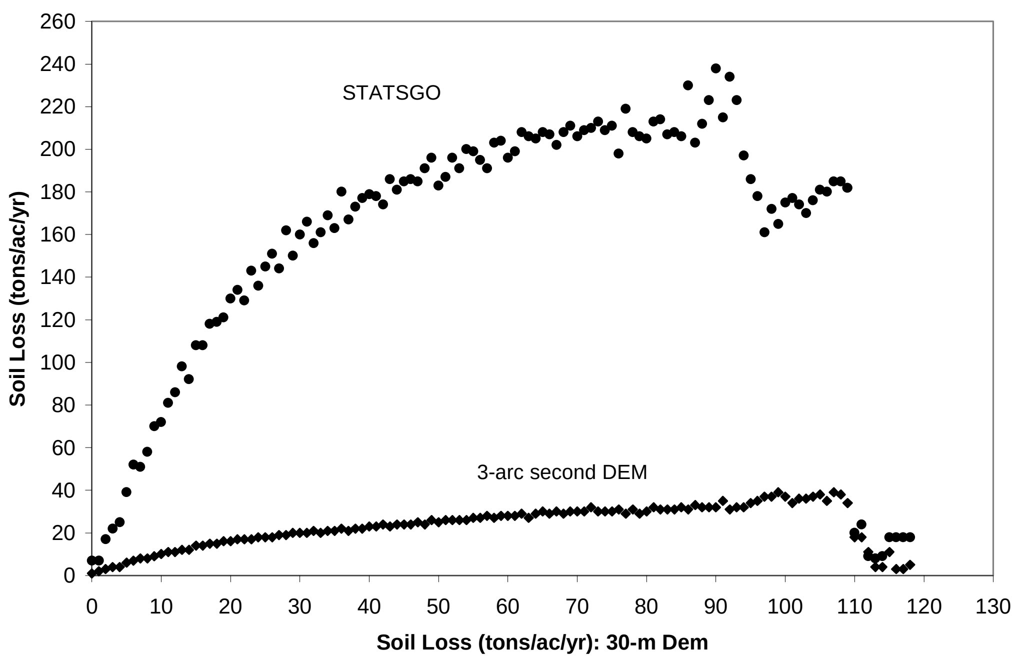


Figure 2: Soil loss estimate comparison between three data sources with 30-m DEM as a reference. Data fluctuations at the end are caused by a small number of data points.

References

Burrough, P. A., 1986. Principles of geographic information systems for land resources assessment. Oxford University Press, New York, NY, 50pp.

Cooperative Extension Service and the Ohio State University, 1979. Ohio erosion control and sediment abatement guide. Columbus, OH, 24pp.

ESRI, 1995. Arc/Info version 7.0.

Microsoft Corporation, 1996. Office 97 spreadsheet, graphing and database software.

U.S. Department of Agriculture, 1994. State soil geographic (STATSGO) data base--data use information, miscellaneous publication number 1492: Fort Worth, TX, Natural Resources Conservation Service.

U.S. Geological Survey, 1987. Digital Elevation Models. Data Users Guide 5. Reston, VA.

Ward, A. D. and W. J. Elliot, 1995. Environmental Hydrology. Lewis Publishers, NY, 462pp.

Geology of Will and Southern Cook Counties, Illinois

Edward Caldwell Smith
Illinois State Geological Survey, Champaign, Illinois

Introduction

The Silurian dolomite aquifer is the primary source of ground water in northeastern Illinois. It is overlain by glacially derived sands and gravels or tills. The sands and gravels within the glacial drift hydrologically interact with the fractured and creviced dolomite bedrock.

The purpose of this study was to define the extent of major glacial drift aquifers and their relationship to the shallow bedrock aquifer surface. The study succeeded in identifying two principal sand and gravel aquifers: an "upper" drift aquifer within the glacial tills and a "basal" drift aquifer overlying the bedrock. Bedrock topography, drift thickness, thickness of the Silurian dolomite, and thickness of major sand and gravel units were mapped to help define the geologic and hydrologic system and the interaction of the upper bedrock aquifer and the drift aquifers.

The data collected to create the various maps came from well records, engineering borings, oil and gas tests, and structure tests on file at the Illinois State Geological Survey (ISGS). Reviewing published reports, manuscripts, and unpublished reports on open file at the ISGS provided an overall perspective of the geology of the study area. Previously, no detailed studies of the hydrogeology of the entire area had been conducted. Incorporating water well and other data into a computer database greatly facilitated map construction. Preliminary maps were developed using Interactive Surface Modeling (ISM) software and a geographic information system (GIS).

Past regional geologic studies of the northeastern Illinois area that have encompassed this study area include Thwaites (1), Bretz (2), Bergstrom et al. (3), Bretz (4), Suter et al. (5), Hughes et al. (6), and Willman (7). Bogner (8) and Larsen (9) included interpretive maps of the surficial geology of the area as a part of planning studies for northeastern Illinois.

Map Construction

Creating the database used in the construction of the maps for this project entailed inputting information from well driller's logs into a PC-based computerized spreadsheet (Quattro Pro). Well logs were primarily from water wells and engineering borings. Data items input into the spreadsheet included:

- Well identification (ID) number
- Owner name
- Location of well
- Thickness of drift
- Depth to top and bottom of the bedrock
- Depth to top and bottom of each sand unit

The ground surface elevation of each well was interpolated from United States Geological Survey (USGS) 7.5-minute quadrangles. Elevations of the top of bedrock and top and bottom of sand bodies were calculated based on the interpolated elevations. Locations were verified wherever possible using plat books by matching either landowner names or the address location from the well log. After compilation, the data were converted to ASCII text and transferred into an ARC/INFO (Versions 5.0.1 and 6.0) database on a SUN SPARC workstation. ARC/INFO is a product of Environmental Systems Research Institute, Inc., of Redlands, California.

Of the more than 10,000 records reviewed for this project, over 5,100 were input into the database. Subsequently, numerous data quality checks ensured that duplicate well ID numbers were corrected, locations were corrected, thicknesses were checked so that the sand thickness data reported did not exceed drift thickness, and elevations were checked so that elevation of a sand body was not below the bedrock surface. After running the data quality checks and removing questionable data from the database, approximately 5,000 records remained.

ISM, a contouring package from Dynamic Graphics, Inc., of Alameda, California, helped to create two dimensional grid representations of:

- Surface topography
- Drift thickness

- Bedrock topography
- Bedrock isopach
- Intermediate sand body isopach
- Basal sand isopach

ISM also allowed for the creation of contoured output of the grids. Grids are regularly spaced rectangular arrays of data points (nodes) that allow for efficient mathematical calculations and contouring. ISM uses a minimum tension gridding technique, allowing for the curvature (change in slope) of the surface to be spread throughout the surface rather than being concentrated at the input data points. The ISM program uses a biharmonic inverse cubic spline function (algorithm) to assign data values to grid nodes. This function assumes that for any grid node assignment, input data points farther away from the node being evaluated have less influence on that node's value than nearer data points. To determine each grid node value, ISM calculates an average value from the surrounding scattered input data (up to 15 input data points) and finds the standard deviation. ISM continues to refine the values of the grid nodes until the standard deviation is minimized (10).

Several grid spacings were reviewed to determine which would best represent the density of the data. The grids that ISM uses, as described above, determine the fineness to which the data control the resultant contours. Experimentation was necessary to determine a grid spacing that adequately represented the data. Too fine a grid spacing can exaggerate or overly weight individual points, causing the resultant contours to be overly jagged. With too large of a grid spacing, the contours can become overgeneralized and become much less data dependent because the calculated grids are overaveraged.

The two-dimensional grid of the land surface topography was based on surface topography lines and spot elevations digitized from USGS 7.5-minute quadrangles. The linework for each quadrangle was converted to ASCII files of data points. The ASCII files contained x and y coordinates and the elevation value of each data point. After inputting the ASCII files into ISM, a two-dimensional grid for each quadrangle was created. ISM also generated contour lines from each grid. Comparing plots of the generated lines with USGS 7.5-minute topographic maps allowed for the correction of errors and ensured that the grid elevation values were within 10 feet of the elevations shown on the USGS maps. An ISM two-dimensional grid of the entire area's surface topography was created by combining the grids. After creating a contoured surface of the grid, an ARC/INFO coverage of the output was produced. ARC/INFO was used to edit the coverage and produce the final map.

The two-dimensional grid of the bedrock surface topography was based on data from water well and engineering boring logs, ISGS field observations of outcrop locations, and previous ISGS mapping (9). An ASCII file of x and y coordinates and the elevation of each bedrock top was input into ISM. Subtracting the bedrock topography grid from the land surface grid produced a grid of the drift thickness. A contoured output of the grid was produced, and an ARC/INFO coverage of the output was created. Again, ARC/INFO was used to edit the coverage and produce the final map.

Creating the isopach maps entailed subtracting the top and bottom elevations of each unit to calculate the thickness of each unit. ASCII files of the x and y coordinates and the thickness values for each data point were input into ISM. ISM then created two-dimensional grids of each isopach. Contoured output of each grid was produced, which allowed for the creation of ARC/INFO coverages of the output. ARC/INFO was used to edit the coverages and produce the final maps.

Bedrock Geology of the Study Area

All the sedimentary bedrock units are of the Paleozoic Era. The Paleozoic bedrock comprises sequences of sandstones, dolomites, limestones, and shales. The stratigraphic column of Figure 1 illustrates the vertical succession of the bedrock. Major tectonic activity of the area includes the formation of the Kankakee Arch in Ordovician time (11) and faulting along the Sandwich Fault Zone. Faulting along the Sandwich Fault Zone (see Figure 2) may have occurred coincidentally with the formation of the Lasalle Anticlinorum in early Pennsylvanian time (12). No further faulting has been noted since deposition of glacial sediments. Bedrock units gently dip to the east (7). The majority of the area lies on the Niagara cuesta, a south and west facing scarp that comprises the resistant Silurian strata that have an eastward dip of roughly 15 feet per mile (13). The Silurian strata are absent west of the Kankakee River as well as in an area west of the Des Plaines River in west-central Will County (see Figure 2). This study relates to the hydrogeology of the Silurian strata and the drift materials, and details only the uppermost bedrock units. The report, however, does briefly summarize units below the Maquoketa Group using information from Hughes et al. (6) and Visockey et al. (14).

Precambrian Bedrock

Granites or granitic rock compose the Precambrian basement of northern Illinois. Few details about the nature of the basement rocks are known because few wells have completely penetrated the sedimentary bedrock of the region. The elevation of the top of the Precambrian basement probably stands at 4,000 feet below mean sea level in the study area.

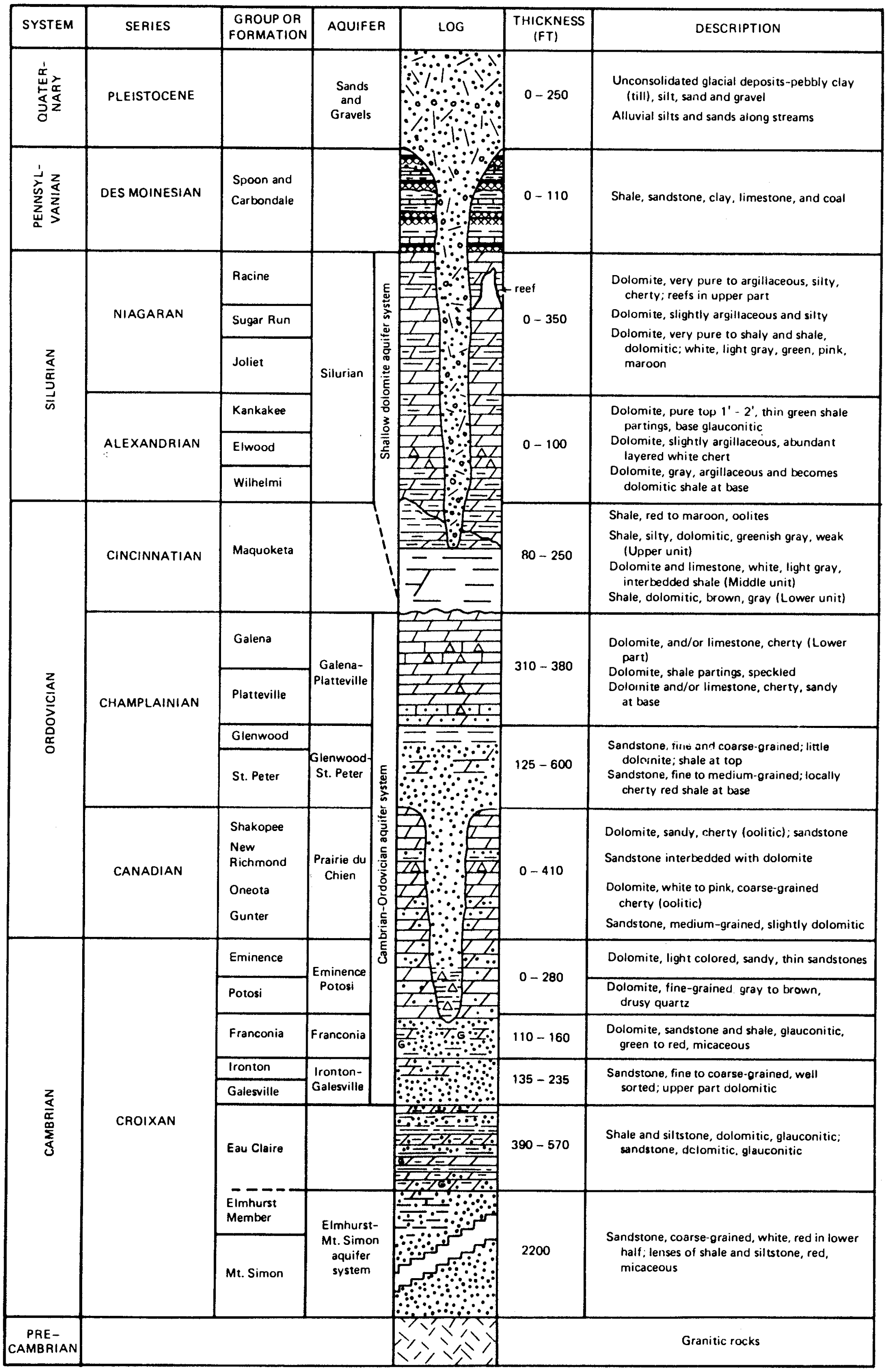


Figure 1. Generalized stratigraphic column of rock units and aquifers in northern Illinois (prepared by M.L. Sargent, ISGS).

Cambrian

The Elmhurst-Mt. Simon Sandstone comprises the oldest sedimentary units in Illinois and consists of medium-grained sandstones. It has a total thickness of approximately 2,500 feet. The upper part of this unit has acted as an aquifer in the Chicago region in the past; ground-water mining of the aquifer (a nonreplenished lowering of the static water level), however, has led to a discontinuation of its use for that purpose. The Eau Claire Formation, the Basal Sandstone Confining Unit (14), consists of dolomitic shale and siltstone with thin beds of sandstone. It has a thickness of 300 feet to 400 feet and separates the Elmhurst-Mt. Simon aquifer from the Ironton-Galesville Sandstones. The Ironton-Galesville Sandstones have a thickness of 150 feet to 250 feet and serve as a source of ground water in northern Illinois (6). The Galesville Sandstone is fine-grained, while the Ironton Sandstone is coarser grained and contains more dolomite. The Knox Megagroup, the Middle Confining Unit (14), comprises all the bedrock units between the Ironton-Galesville Sandstones and the Ancell Group. It includes the:

- Cambrian Franconia Formation
- Potosi Dolomite
- Eminence Formation
- Jordan Sandstone
- Ordovician Prairie du Chien Group

The Knox Megagroup is primarily dolomitic in composition, though it contains thin sandstones. Its thickness ranges from 400 feet in the northern portion of the study area to about 700 feet in the southernmost tip of Will County. The sandstones tend to be somewhat discontinuous and, where present, offer a localized source of ground water. The group as a whole acts as a confining unit between the Ironton Sandstone and the Ancell Group.

Ordovician

The Ancell Group, which contains the St. Peter Sandstone and Glenwood Sandstone, has a thickness of roughly 200 feet throughout the study area except in north-central Will County where it is over 400 feet. The thickness of the Ancell Group varies considerably in northern Illinois because it rests on an erosion surface. The Ancell Group is the shallowest aquifer present in this area below the Silurian dolomite aquifer. The elevations of the top of the Ancell Group range from just over sea level in the northwest corner of Will County to 500 feet below mean sea level in the southwestern corner. The Galena and Platteville Groups provide a sequence of carbonate rocks that are primarily dolomitic in composition. The Platteville Group conformably overlies the Ancell Group. The two units have a combined thickness of 350 feet throughout this part of the state. The Galena and Platteville Groups, combined with the overlying Maquoketa Shale Group, act as an aquitard between the Ancell aquifer and the Silurian dolomite aquifer.

Maquoketa Shale Group

The study area has three subaerially exposed bedrock units. The oldest of these that this report details are Ordovician-aged strata comprising the Cincinnatian Series Maquoketa Shale Group. The thickness of the Maquoketa Group ranges from 260 feet in eastern Will County to 120 feet in the northwestern corner of Will County and is unconformably overlain by Silurian strata (15). The Maquoketa Group comprises four formations:

- Scales Shale
- Fort Atkinson Limestone
- Brainard Shale
- Neda Formation

The Scales Shale forms the lowermost unit and consists of gray to brown dolomitic shale. Thin layers with phosphatic nodules and pyritic fossils occur near the top and base of the unit. The Scales Shale may attain a thickness of up to 120 feet in this region (15). The Fort Atkinson Limestone, a coarse-grained crinoidal limestone to fine-grained dolomite, may range up to 60 feet thick (15). The Brainard Shale comprises greenish gray dolomitic shale and has a thickness of generally less than 100 feet (16). The Neda Formation, the youngest formation in the Maquoketa Group, is relatively thin with a thickness of usually less than 10 feet. In some places, it may attain a maximum thickness of 15 feet. The Neda is exposed along the Kankakee River, and the Silurian-aged Kankakee Formation typically overlies it. The Neda Formation consists mostly of red and green shale with interbedded goethite and hematite oolite beds (7, 16).

Silurian System

Silurian-aged rocks consist almost solely of dolomites and dolomitic limestones. The Silurian is divided into the Alexandrian and Niagaran Series. The Alexandrian Series is about 25 feet thick and is represented by the Kankakee, Elwood, and Wilhelmi Formations. These formations are a fine- to medium-grained, white, gray to pinkish gray dolomite. The Kankakee Formation is exposed along the Kankakee River in southern Will County (17).

The Niagaran Series comprises much of the bedrock surface of this area and includes three formations. The Joliet Formation has a lower member of dolomite with interbedded red and green shale, and two upper members with an increasing purity of dolomite toward the top of the formation (7). The Sugar Run Formation, formerly termed the Waukesha Formation (17), is an argillaceous, fine-grained, medium- to thick-bedded, brownish gray dolomite (7). The Racine Formation is the thickest unit in the Niagaran Series, attaining a thickness of as much as 300 feet (17). The Racine Formation contains large reefs that are as high as 100 feet and consist of vugular gray dolomite. The inter-reef rock consists of dense, cherty gray dolomite. The Racine Formation is exposed in the bluffs along the Des Plaines River from Joliet to Blue Island, Illinois (17).

Figure 2 is an isopach of the Silurian dolomite indicating the thickness of the unit in the study area and the boundary of the Silurian rocks. The Silurian dolomite aquifer has a maximum thickness of just over 500 feet in the southeast corner of Will County and becomes thicker to the east and south. It rapidly increases in thickness from its margin along the western border of Will County, where it has eroded. The contact between the Silurian dolomite and the underlying Maquoketa Shale Group has relatively little relief. Thus, the major differences in thickness of the unit result from erosion of the bedrock surface. Joints and fracture patterns within the upper bedrock have a dominantly northwest-southeast and northeast-southwest orientation (18).

Pennsylvanian System

Pennsylvanian-aged bedrock is found in the southeastern portion of Will County west of the Kankakee River with an outcropping at the confluence of the Des Plaines and Kankakee Rivers. The lowermost unit, the Spoon Formation, is very thin and consists of clay beds with scattered occurrences of coal formed in channel-like depressions (19). The Spoon Formation overlies the Maquoketa Shale Group. The overlying Carbondale Formation may attain a thickness of over 100 feet in the southwestern corner of Will County. The Carbondale Formation consists of shale with thin limestone beds. The lowermost unit, the Colchester (Number 2) Coal Member, outcrops in this area and attains a thickness of up to 3 feet. It has been extensively mined along the Will-Grundy-Kankakee County border where large areas of strip-mined land are evident. Most of the available coal has been mined out, and numerous gob piles exist

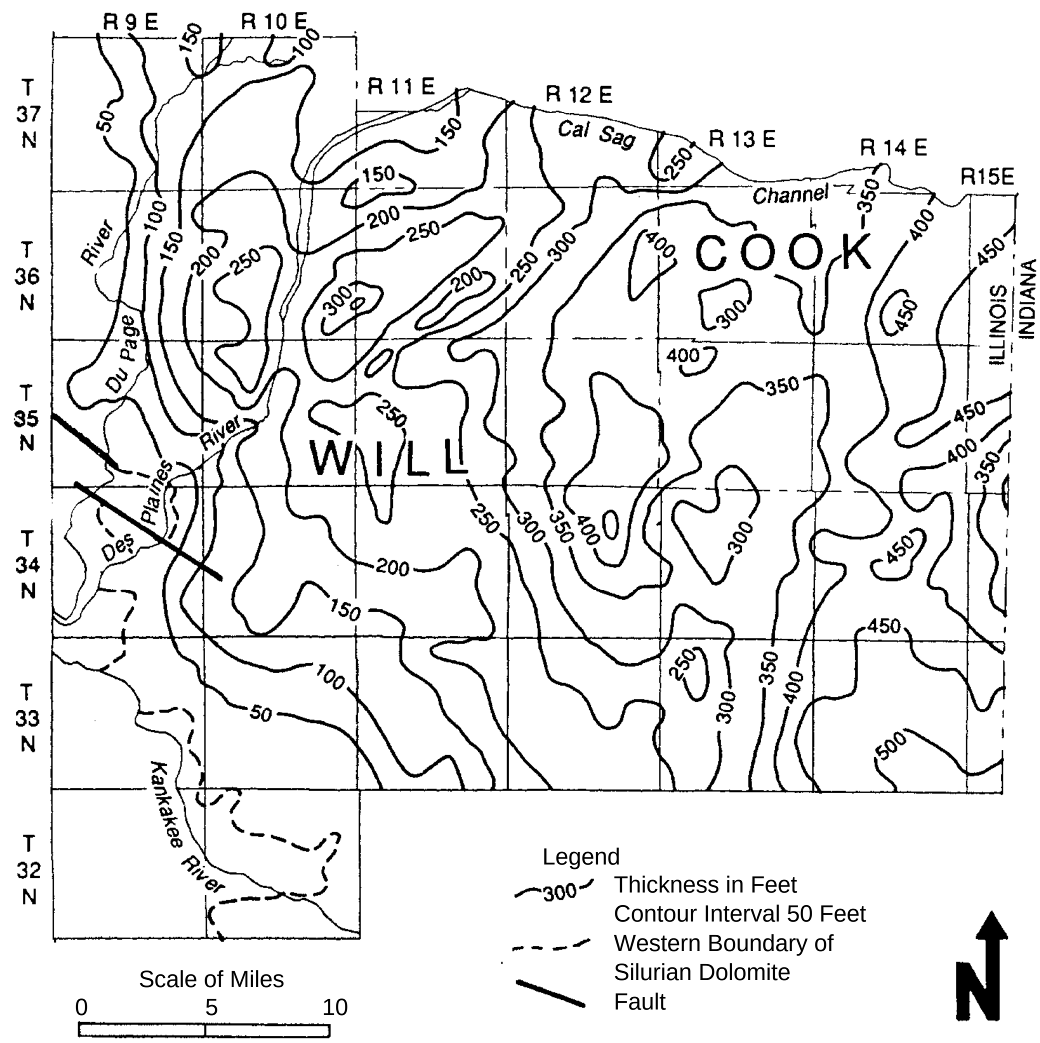


Figure 2. Thickness of the Silurian dolomite.

in the area of Braidwood. The Francis Creek Shale Member, which overlies the Number 2 coal, constitutes the remainder of Pennsylvanian units in the study area. The Francis Creek Shale is gray with numerous flattened concretions that contain the Mazon Creek flora of Pennsylvanian-aged fossils (19). Weathering of the mine slag materials may have exposed fossiliferous concretions in the gob piles (7).

Bedrock Topography

The highest bedrock elevations are in east-central Will County where the bedrock rises to over 700 feet above mean sea level (see Figure 3). Bedrock uplands occur as a broken curved ridge from the southeast to the northwest with bedrock elevations consistently rising over 650 feet above mean sea level. The bedrock surface slopes from the bedrock upland high westward to the Des Plaines River. It also has a regional downward slope to the south into Kankakee County, Illinois, to the northeast into the Lake Michigan basin, and to the east into Indiana. West of the Des Plaines River, the bedrock surface rises to over 650 feet above mean sea level in northeastern Kendall County. Elsewhere, the surface has relatively low relief.

The dominant features of the bedrock surface are the river valleys. The Des Plaines River valley is better expressed than the Kankakee River valley. This is true, in part, because it is older and acted as a drainageway for glacial meltwater where it may have become entrenched in the present valley. The Kankakee River valley may be less expressed partly because of the amount of scouring that occurred over a large area during the Kankakee flood event such that the river is not entrenched in most places. Also, smoothing of the study's contour maps has generalized some of the detail.

The buried Hadley Bedrock valley, described initially by Horberg and Emery (20), probably existed prior to glaciation and concurrently with the preglacial Des Plaines River. The valley may have acted as a drainageway for glacial meltwaters until the time that glacial debris buried it. Glacial scouring was originally believed to have formed the valley, but evidence presented by McConnel indicated a fluvial origin of the valley (21). Also, the base of the Hadley valley does not overhang or lie much below the Des Plaines valley but rather joins it at a smooth juncture.

The bedrock surface contains a number of sinkholes or closed depressions that are expressions of karst

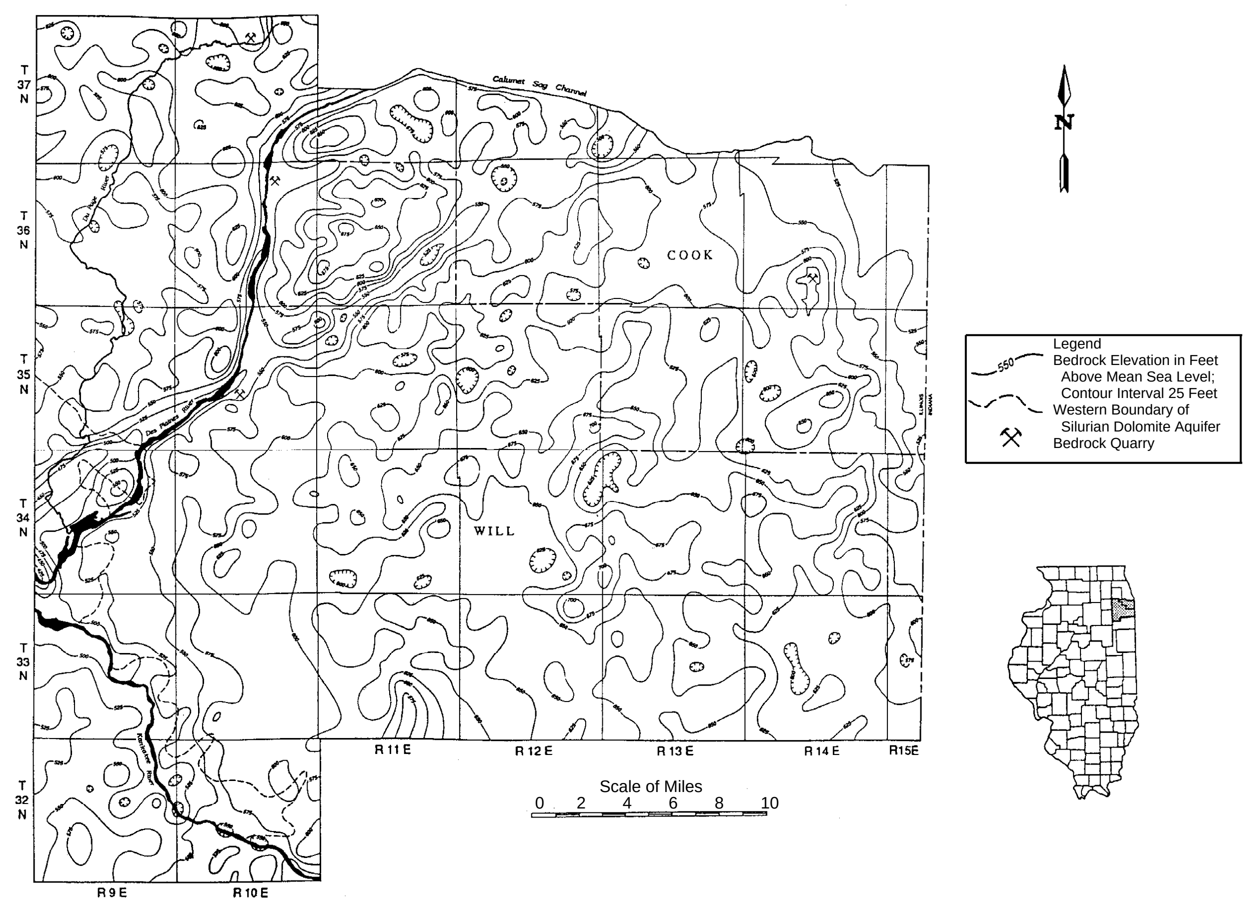


Figure 3. Topography of the bedrock in the study area.

development that formed prior to continental glaciation. Karst, a terrain developed on limestone or dolomite by solution or dissolving of the rock, is characterized by closed depressions and cavity development along joints and fractures. Fischer (22) first noted karst features in the Joliet area where early Pennsylvanian sediments of shale and clay filled cavities in the upper bedrock. Buschbach and Heim (23) indicated closed depressions in the Silurian dolomite surface in their bedrock topography map for the Chicago region. They speculated these depressions were expressions of karst development. McConnel (21) demonstrated the existence of sinkholes in the area of the buried Hadley Bedrock valley northeast of Joliet by using seismic refraction survey data.

Glacial Geology

The sediments overlying the bedrock comprise tills, sands and gravels, lacustrine deposits from glacial lakes, and surficial eolian deposits of loess and sand. The unconsolidated deposits are over 150 feet thick along the crest of the Valparaiso Morainic System. Figure 4, adapted from Willman (24), indicates the principal moraines. In the area where the Hadley Bedrock valley is present, the deposits attain a thickness of over 175 feet. Bedrock is mainly exposed along the Des Plaines River valley and its tributaries. It is also exposed in isolated areas in southeastern Cook County. The drift thickness map (see Figure 5) indicates the distribution of the earth materials overlying the bedrock and the locations of bedrock outcrops. The bedrock outcrop information for this map was derived from Piskin (25) and Berg and Kempton (26).

Erosion of the glacial sediments was a major factor in controlling the drift thickness of the area. Succeeding glaciers scraped off previously deposited sediments, but glacial meltwaters, which came from the east and north along the river channels, caused much of the erosion. Both the Kankakee River and Des Plaines River acted as meltwater channels as the glaciers melted. The Du Page River acted as a minor drainageway and was most active during large-scale flooding events. The thickness of the drift varies in the area also because of the topographic control that the bedrock on the overlying sediments exercises. The crest of the Valparaiso moraines coincides with the topographic high in the bedrock

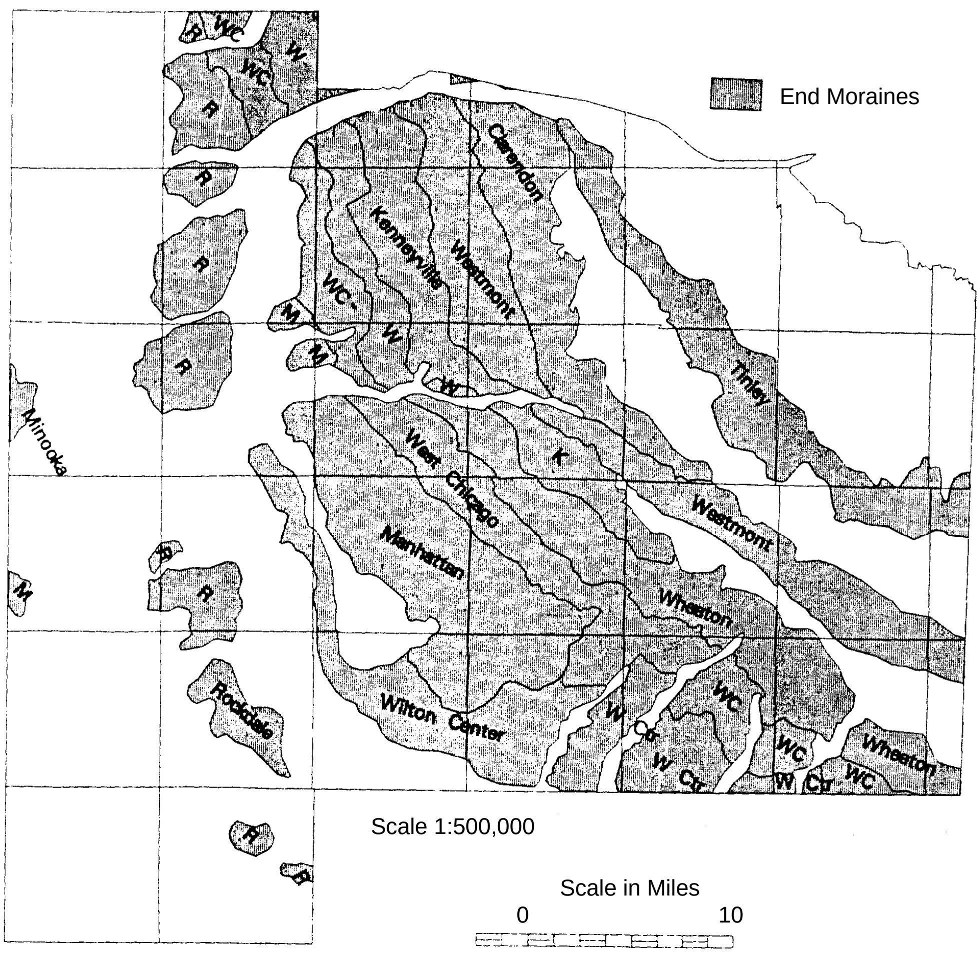


Figure 4. End moraines (late Wisconsin) in Will and southern Cook Counties, Illinois.

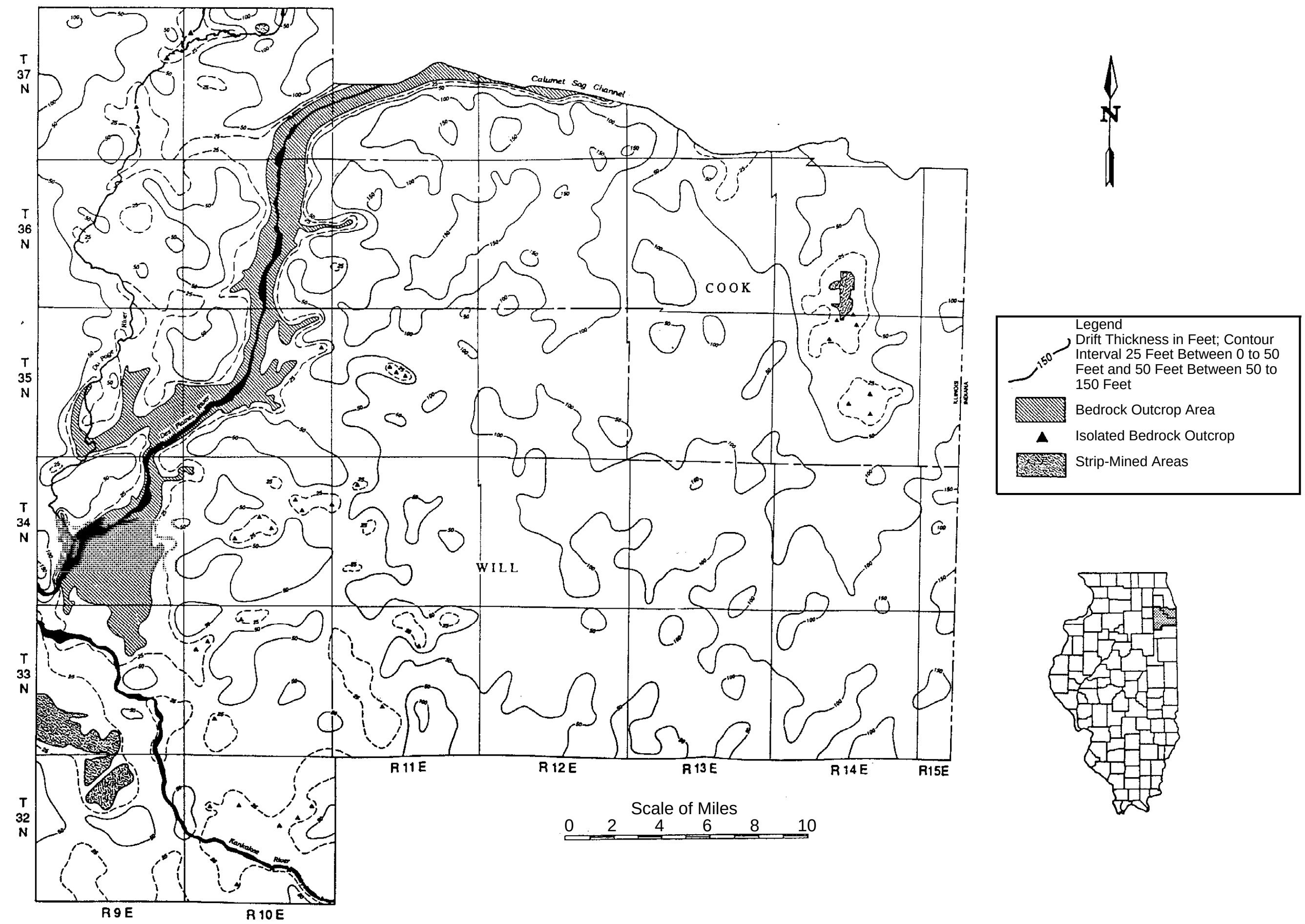


Figure 5. Thickness of the glacial drift in the study area and bedrock outcrop information (25, 26).

surface. The cross sections in Figure 6 also show this. The bedrock high may have caused late Woodfordian glaciers to stall repeatedly in the same area, causing moraines to build atop one another sequentially (27).

Descriptions by well drillers note few variations in the character of the unconsolidated sediments; therefore, we did not attempt to correlate these deposits. The drift materials present in the study area are late Wisconsinan or younger. Though this region experienced glaciating repeatedly prior to the Wisconsinan glaciation, no Illinoisan or pre-Illinoisan deposits have been identified (28). The drift units divide into three main units (29):

- The Lemont Drift
- The Yorkville Drift
- The Wadsworth Drift

The three drift units are all part of the Wedron Formation of Wisconsinan age. The Lemont Drift has a dolomitic character because the source material for the diamicton was glacially eroded Silurian dolomite. The Lemont Drift is the oldest of the three units and is found only underlying the Wadsworth Drift. The Yorkville Drift is the only drift unit present west of the Valparaiso Morainic System boundary within the study area. It overlies the bedrock surface wherever the basal sand unit is not present. The Wadsworth Drift comprises silty and clayey diamictons and is the youngest of the drifts (29). It overlies the Lemont Drift and the upper sand unit. In the cross sections (see Figure 6), where the upper sand unit is present, it roughly indicates the boundary between the Lemont and Wadsworth Drifts. The gradation between the different drift units at the Valparaiso System boundary is not well defined. The Wadsworth Drift appears to grade into the Yorkville Drift because they are very similar in composition near the boundary (9).

The large Kankakee flood left extensive deposits of sand and gravel and lacustrine sediments along the Kankakee River and Des Plaines River. The flood occurred as glacial meltwaters built up behind a constriction at the Marseilles Morainic System to the west (30). Large glacial lakes, which developed during the flood, subsequently emptied into the Illinois River valley after a breach in the moraines developed. The force of the flood waters eroded the glacial deposits along the river valleys, flattened the surface of the drift, and, in places,

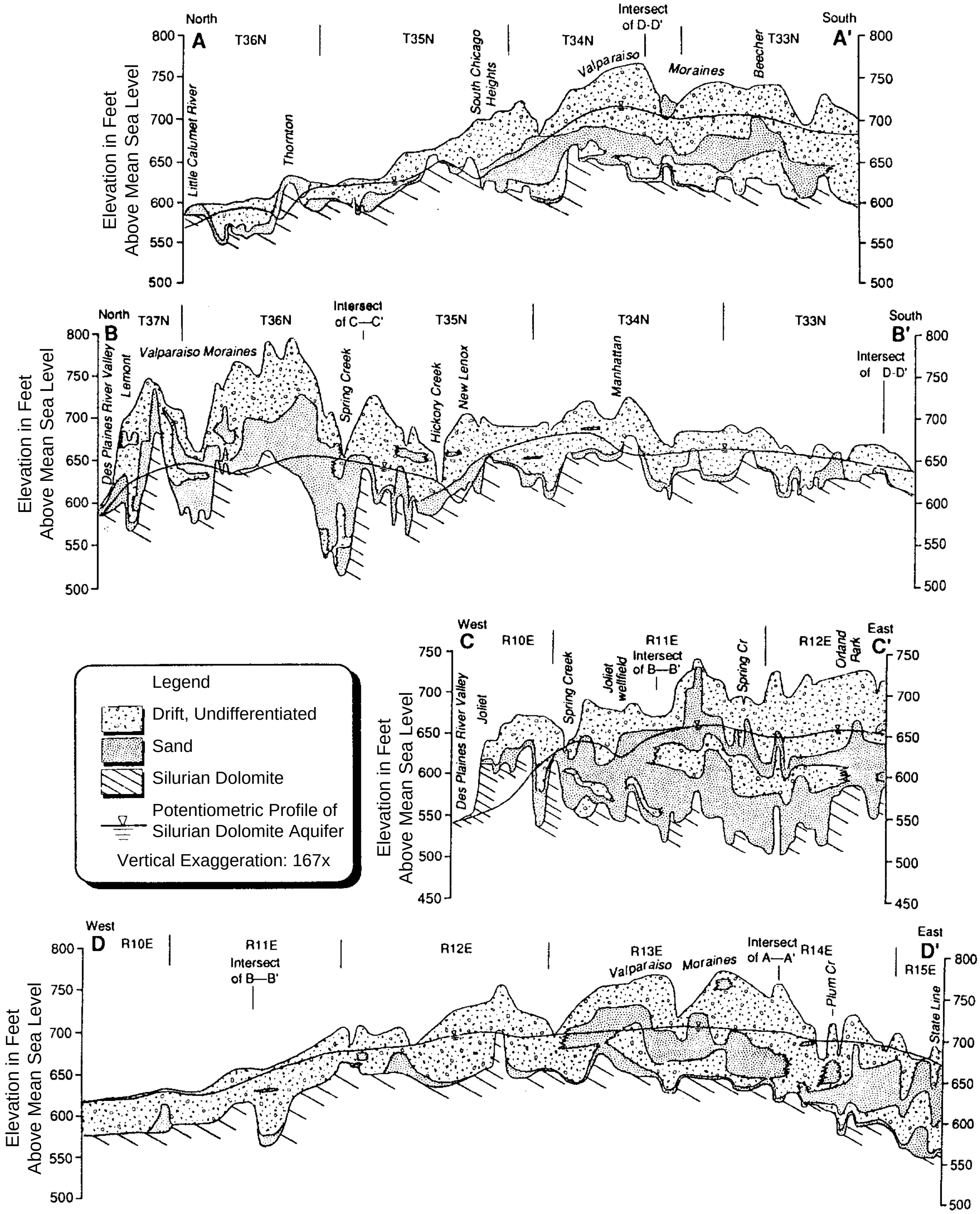


Figure 6. Geologic cross sections of the glacial drift and potentiometric profile of the Silurian dolomite aquifer.

exposed the underlying bedrock. The flood event formed thin, dispersed lake plain deposits of silt, clay, and sand in southwestern Will County. Some lacustrine deposits lie between morainic ridges in southern Cook County where small glacial lakes developed as the Valparaiso Moraines were being deposited (8).

Figure 7 shows the locations of some of the surficial materials. Sands and gravels were also deposited along tributary creeks and in abandoned channels that once connected the Du Page River and Des Plaines River north of their present juncture. Wind has reworked the surficial sand deposits forming low dunes along the Kankakee River in southern Will County. Masters (31) classified the sand and gravel deposits of the area by their origin, indicating that most of the deposits present in the valley of the Des Plaines River formed as well-sorted valley train deposits. In the Kankakee River valley, the sands and gravels were primarily deposited as riverine sediments during the Kankakee flood event.

Sand and Gravel Isopachs

The sand and gravel isopach maps (see Figures 8 and 9) indicate the variations in thickness of the upper and basal sand and gravel units. The most extensive deposits of both exist throughout the area overlain by the Valparaiso Morainic System. This may be associated with bedrock control on the formation of the moraines and associated deposits referred to earlier.

The thickest deposits lie in the buried Hadley Bedrock valley where thicknesses of both units can exceed over 100 feet. The upper sand unit may be found in the glacial drift within a wide range of elevations. For mapping purposes, we defined the upper sand unit as a sand unit greater than 1 foot thick that occurs between two fine-grained layers. The basal sand unit includes all coarse-grained materials that overlie the bedrock surface. Most of the basal sands present west of the Des Plaines River were formed as valley train deposits along the river

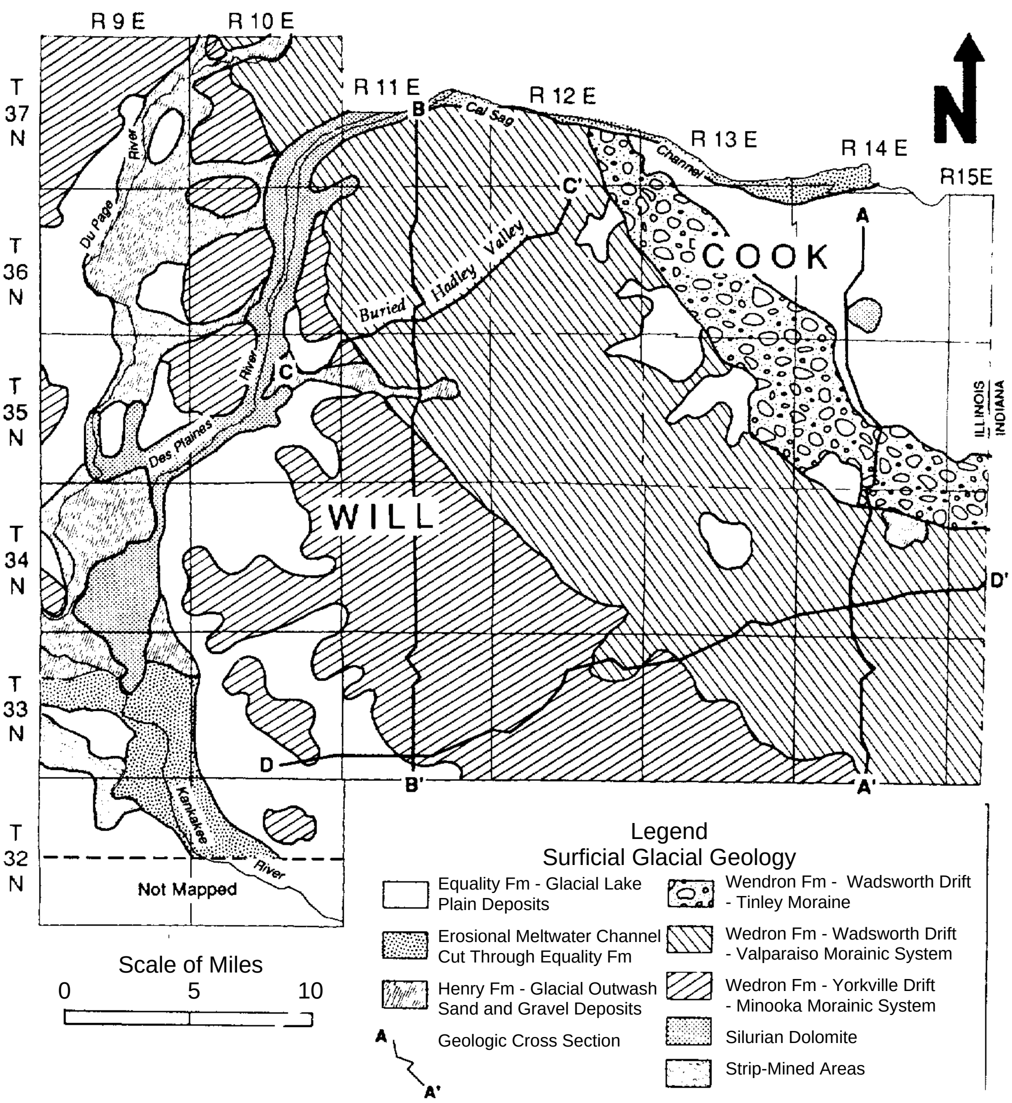


Figure 7. Surficial glacial geology (23).

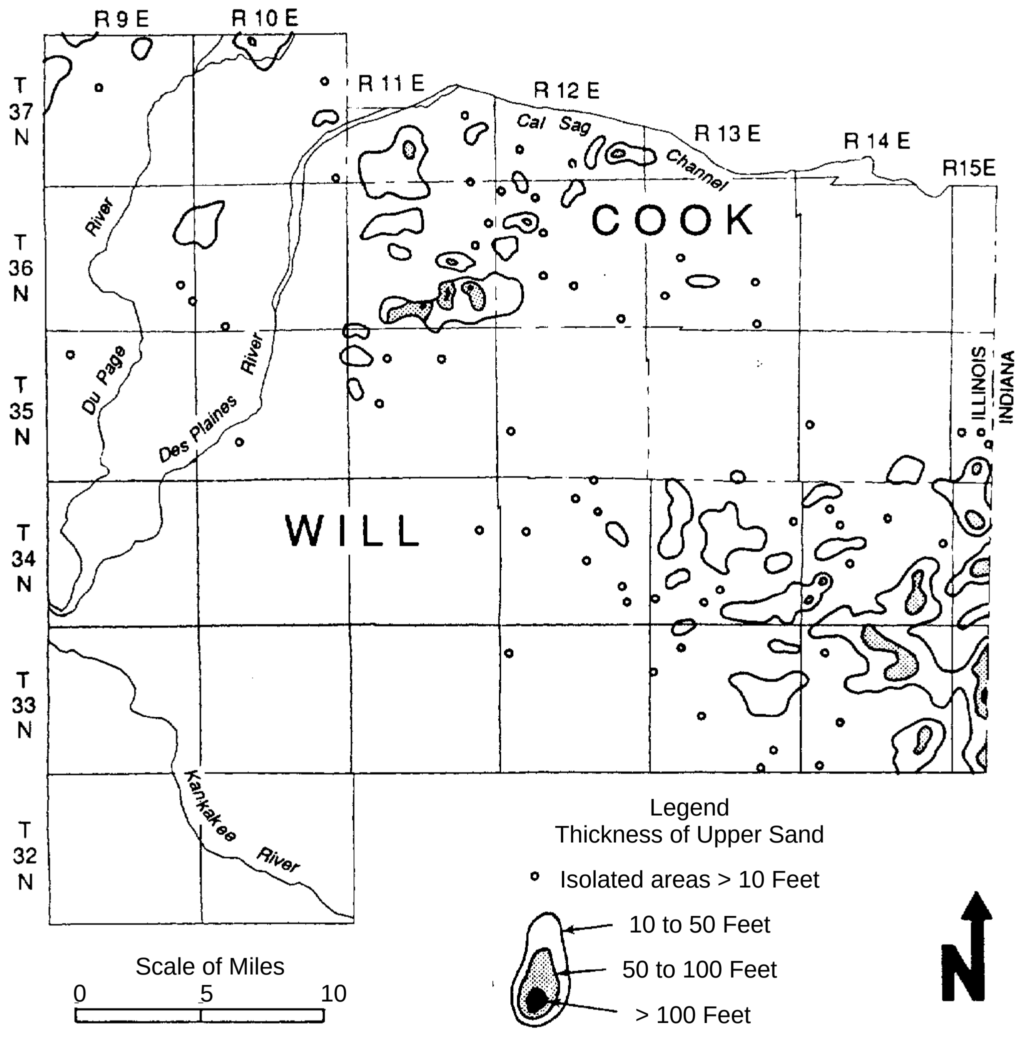


Figure 8. Thickness of the upper sand unit.

channels as the glaciers melted back. The origin of the extensive deposits underlying the Valparaiso Moraines is not clear. They may have been formed during early Wisconsinan glacial events as outwash plain deposits or they may have been deposited subglacially. The cross sections (see Figure 6) can reveal the variability and complexity of the sand and gravel layers as they occur within the drift. The sand and gravel deposits very seldomly act as aquifers in this region because almost all wells are completed in the Silurian dolomite aquifer. Clearly, Figures 8 and 9 indicate that some ground-water resource potential may exist within these deposits.

References

1. Thwaites, F.T. 1927. Stratigraphy and geologic structure of northern Illinois. Illinois State Geological Survey Report of Investigations 13.
2. Bretz, J.H. 1939. Geology of the Chicago region, Part I: General. Illinois State Geological Survey Bulletin 65.
3. Bergstrom, R.E., J.W. Foster, L.F. Selkregg, and W.A. Pryor. 1955. Ground-water possibilities in northeastern Illinois. Illinois State Geological Survey Circular 198.
4. Bretz, J.H. 1955. Geology of the Chicago region, Part II: The Pleistocene. Illinois State Geological Survey Bulletin 65.
5. Suter, M., R.E. Bergstrom, H.F. Smith, G.H. Emerich, W.C. Walton, and T.E. Larson. 1959. Preliminary report on the ground-water resources of the Chicago region, Illinois. Illinois State Geological Survey and Illinois State Water Survey Cooperative Ground-Water Report 1.
6. Hughes, G.M., P. Kraatz, and R.A. Landon. 1966. Bedrock aquifers of northeastern Illinois. Illinois State Geological Survey Circular 406.
7. Willman, H.B. 1971. Summary of the geology of the Chicago area. Illinois State Geological Survey Circular 460.
8. Bogner, J.E. 1976. Geology for planning in northeastern Illinois, Vol. V: Cook County. Illinois State Geological Survey Open-File Report.
9. Larsen, J.I. 1976. Geology for planning in northeastern Illinois, Part VI: Will County. Illinois State Geological Survey Open-File Report.

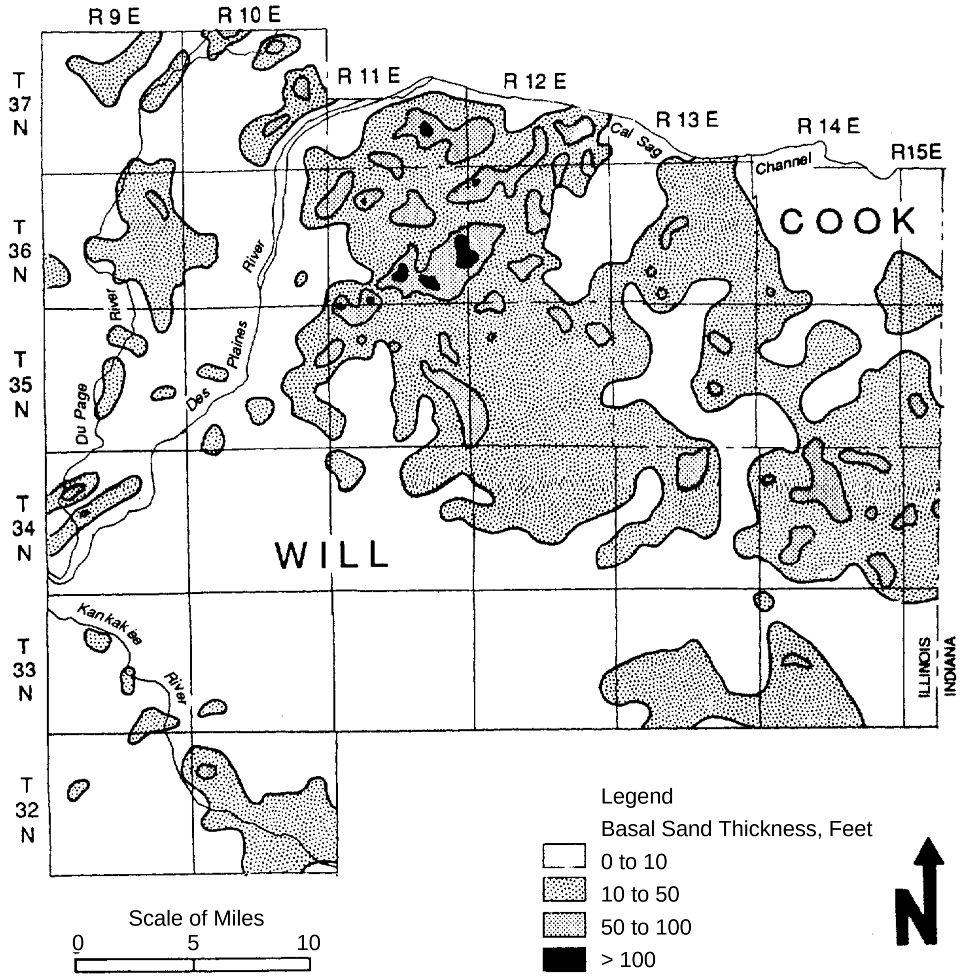


Figure 9. Thickness of the basal sand unit.

10. Dynamic Graphics, Inc. 1991. Interactive surface modeling users guide. Alameda, CA.

11. Ekblaw, G.E. 1938. Kankakee Arch in Illinois. Illinois State Geological Survey Circular 40.

12. Kolata, D.R., T.C. Buschbach, and J.D. Treworgy. 1978. The Sandwich Fault Zone of northern Illinois. Illinois State Geological Survey Circular 500.

13. Horberg, L. 1950. Bedrock topography in Illinois. Illinois State Geological Survey Bulletin 73.

14. Visocky, A.P., M.G. Sherrill, and K. Cartwright. 1985. Geology, hydrology, and water quality of the Cambrian and Ordovician systems in northern Illinois. Illinois Department of Energy and Natural Resources Cooperative Groundwater Report 10.

15. Kolata, D.R., and A.M. Graese. 1983. Lithostatigraphy and depositional environments of the Maquoketa group (Ordovician) in northern Illinois. Illinois State Geological Survey Circular 528.

16. Willman, H.B., et al. 1975. Handbook of Illinois stratigraphy. Illinois State Geological Survey Bulletin 95.

17. Willman, H.B. 1973. Rock stratigraphy of the Silurian system in northeastern and northwestern Illinois. Illinois State Geological Survey Circular 479.

18. Foote, G.R. 1982. Fracture analysis in northeastern Illinois and northern Indiana. Master's thesis, University of Illinois.

19. Smith, W.H. 1968. Strippable coal reserves of Illinois, Part 6: La Salle, Livingston, Grundy, Kankakee, Will, Putnam, and parts of Bureau and Marshall Counties. Illinois State Geological Survey Circular 419.

20. Horberg, L., and K.O. Emery. 1943. Buried bedrock valleys east of Joliet and their relation to water supply. Illinois State Geological Survey Circular 95.

21. McConnel, D.R. 1972. Bedrock topography and paleogeomorphology northeast of Joliet, Illinois. Master's thesis, University of Illinois.

22. Fischer, D.J. 1925. Geology and mineral resources of the Joliet quadrangle. Illinois State Geological Survey Bulletin 51.

23. Buschbach, T.C., and G.E. Heim. 1972. Preliminary geologic investigations of rock tunnel sites for flood and pollution control in the greater Chicago area. Illinois State Geological Survey Environmental Geology Notes 52.

24. Wilman, H.B., and J.C. Frye. 1970. Pleistocene stratigraphy of Illinois. Illinois State Geological Survey Bulletin 94.

25. Piskin, K. 1975. Thickness of glacial drift in Illinois. Illinois State Geological Survey 1:500,000 scale map.

26. Berg, R.C., and J.P. Kempton. 1988. Stack-unit mapping of geologic materials in Illinois to a depth of 15 meters. Illinois State Geological Survey Circular 542.

27. Hansel, A.K., and W.H. Johnson. 1987. Ice marginal sedimentation in a late Wisconsinan end moraine complex, northeastern Illinois, USA. In: van der Meer, J.J.M., ed. Tills and glaciotectonics. Proceedings of an INQUA Symposium on genesis and lithology of glacial deposits.

28. Horberg, L. 1953. Pleistocene deposits below the Wisconsinan drift in northeastern Illinois. Illinois State Geological Survey Report of Investigations 165.

29. Johnson, W.H., and A.K. Hansel. 1985. The Lemont section. In: Johnson, W.H., A.K. Hansel, B.J. Socha, L.R. Follmer, and J.M. Masters, eds. Depositional environments and correlation problems of the Wedron Formation (Wisconsinan) in northeastern Illinois. Illinois State Geological Survey Guidebook 16.

30. Gross, D.L., and R.C. Berg. 1981. Geology of the Kankakee River system in Kankakee County, Illinois. Illinois State Geological Survey Environmental Geology Notes 92.

31. Masters, J.M. 1978. Sand and gravel resources in northeastern Illinois. Illinois State Geological Survey Circular 503.

Application of GIS for Environmental Impact Analysis in a Traffic Relief Study

Bruce Stauffer
Advanced Technology Solutions, Inc., Lancaster, Pennsylvania

Xinhao Wang
University of Cincinnati, Cincinnati, Ohio

Abstract

This paper presents an application of a geographic information system (GIS) in a traffic relief study. Traffic congestion has severely affected the environmental quality and the quality of life for residents in the study area. A team of planners, environmental specialists, historians, landscape architects, traffic engineers, and GIS professionals organized to solve the problem. The team has evaluated the environmental and socioeconomic impacts of highway alignments from the very first step through every major decision for the duration of the project.

The GIS professionals have played a crucial role in maintaining constant and active interactions among members of the project team, federal and state agencies, and the public. GIS has helped to develop a natural and cultural resource inventory, identify contamination sources, assess environmental constraints, and evaluate proposed highway alignment alternatives. GIS provides an ideal atmosphere for professionals to analyze data, apply models, and make the best decisions. The high-quality map products that GIS creates enhance the quality of public presentations and reports. The authors feel that, as this project has progressed, more people have realized the benefit of using GIS.

Introduction

A traffic relief study, as one type of transportation project, aims to resolve traffic congestion problems through a combined strategy of upgrading existing infrastructure, building new infrastructure, and controlling traffic demand using congestion management strategies (CMS). This type of study proceeds through at least the following steps:

- Problem identification
- Data collection
- Preliminary design
- Environmental impact analysis
- Final design
- Construction

The process heavily involves federal, state, and local government agencies, as well as the public. The goal of the project is to develop an environmentally sound solution to the traffic congestion, which also happens to promote economic development and improve quality of life for people in the local area and the region. Environmental, social, and economic issues must be equally addressed from the very first step through the final design. Federal and state regulations generally require an environmental impact statement (EIS) when constructing new infrastructure or upgrading existing road systems. Preparing an EIS is a requirement for such a project and demands a significant commitment of time, money, staff, and technical resources.

A geographic information system (GIS) has the ability to process spatially referenced data for particular purposes. Along with the development of computer hardware and software, GIS has progressed from pure geoprocessing, to management of geographic information, to decision support (1). This paper presents the application of GIS in an ongoing traffic relief study in Marshalls Creek, Pennsylvania. The GIS function in this study has had various purposes:

- Inventory data compilation
- Spatial data analysis
- Map production
- Traffic modeling support
- Public presentations

The study shows that GIS can play an important and innovative role in transportation studies.

Project Description

The study area is in the Pocono region, located in northeastern Pennsylvania (see Figure 1). The Pocono Mountains and Delaware Water Gap National Recreation Area possess a wealth of natural and cultural resources. The area is famous for providing year-round vacation activities. Attractions include fishing, canoeing, and whitewater rafting in the summer and downhill and cross-country skiing, snowmobiling, snowboarding, and ice fishing in the winter. The area includes quiet woodland trails past a rushing waterfall and scenic settings for camping. In the fall, the area is ablaze with the brilliant colors of foliage. Various scenic sights, recreational sites, and national historic sites make the area ideal for attracting people to come for a day, a weekend, or a longer vacation.

Although tourism brings people to the Pocono area and promotes economic growth, it also brings a traffic congestion problem to the community. In addition, the influx of new home owners from New Jersey and New York adds the problem of commuter traffic to the area. The most troublesome section is in the vicinity of Marshalls Creek where U.S. Route 209 intersects with Pennsylvania (PA) Route 402. Two intersections are only about 500 feet apart. The traffic tieups can extend up to 3.5 miles on northbound Route 209, all the way back to Interstate Highway I-80. Emergency response times on U.S. Route 209 can be up to 20 minutes during peak traffic. The heavy traffic volume results in traffic accidents exceeding state averages on secondary roads as motorists seek alternative routes to avoid congestion. Through traffic traveling to and from New England using U.S. Route 209 as a connector between I-80 and I-84 makes the problem even worse. The year-round outdoor activities perpetuate the constant traffic problems that have severely affected the quality of life for residents in and around the Marshalls Creek area.

In response to the problems, the Pennsylvania Department of Transportation (PennDOT) selected a project team in February 1993 to conduct a traffic relief study in the Marshalls Creek area. The project team consists of individuals from seven firms and represents a wealth of experience in the variety of disciplines necessary to successfully complete this project. The team members include land use and traffic planners, biologists, historians, traffic and environmental engineers, surveyors, and GIS and global positioning system (GPS) professionals.

In addition to PennDOT, the funding agency, several federal and state regulatory agencies periodically review the development of the EIS to ensure that it meets regulations. These agencies are the Federal Highway Administration (FHWA), the U.S. Environmental Protection Agency (EPA), the U.S. Army Corps of Engineers, the Pennsylvania Department of Environmental Resources (DER), the Pennsylvania Historic Museum Commission, and the Pennsylvania Fish and Game Commission. Local planning commissions and citizen representatives also actively participate in advisory capacities. A series of agency coordination meetings, public meetings, and public information newsletters coordinates the activities of all participants over the course of the project.

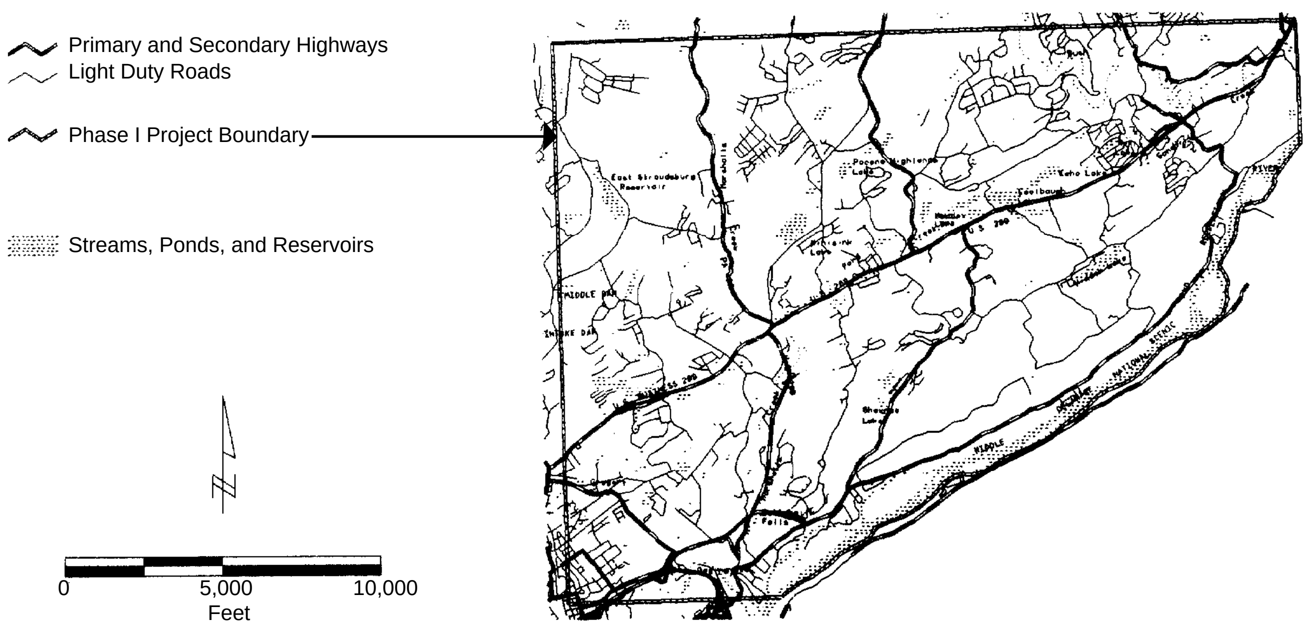


Figure 1. Phase I project area.

After collecting traffic data and performing traffic demand modeling, the team realized that adopting strategies to control traffic demand and upgrading existing roads through widening and intersection improvements would not suffice to meet demand projections for the design year 2015. Consequently, the team has determined a new road is needed to alleviate congestion in Marshalls Creek.

The study aims to identify alternatives to relieve traffic congestion along U.S. Route 209, PA Route 402, and Creek Road, and eliminate backups onto I-80 from U.S. Route 209. The alternatives should also improve air quality by reducing fuel consumption and vehicle emissions and facilitate travel through Marshalls Creek for local and through traffic. The improvements must comply with federal and state regulations. The study team must consider county and local government goals and objectives so that the traffic capacity improvements will be compatible with planned local development.

The project is being conducted in two phases. Phase I, which is complete, was an investigation broad in scope. It used inventory of secondary data to describe the environmental characteristics of the area. A traffic demand model identified the area for detailed study after a preliminary analysis of a wide range of baseline data. In Phase II, the team analyzes both primary and secondary data and delineates alternative alignments or transportation upgrading options that meet the need and minimize impacts. Analysis of environmental and engineering factors assists both in the determination of the most practical alternative and in preparation of the final EIS.

The nature of the study requires the analysis of a variety of data at different scales by different professionals. Through field investigation, the project team also constantly updates and adds new data to the existing database. The new data may be attribute data about some geographic features or may be locational data. The project team has found GIS to be an appropriate tool to meet the challenge of better conducting the study.

The team has used GIS extensively in both Phase I and Phase II studies. The two phases vary in data requirements, scales, and purposes of spatial analysis. With the support of GIS, the team has been able to quickly assemble data at adequate scales and present data in formats that are familiar to different professionals. GIS's data manipulation power distinguishes the different requirements of the two phases and at the same time, clearly depicts the linkage between the two phases. The following sections describe GIS applications that have helped facilitate the study and coordinate project team members, public agencies, and citizens.

GIS Application

GIS contains powerful tools to process spatially referenced data. These processes and their results are meaningless, however, without a clearly defined objective. Many professionals point out the importance of focusing GIS on practical problems. Using GIS is not an end; it is a means to represent the real world in both spatial and temporal dimensions. The benefits of using GIS can be summarized in three aspects:

- GIS helps to portray characteristics of the earth and monitor changes of the environment in space and time (2).
- GIS helps us to more deeply understand the meaning of spatial information and how that information can more faithfully reflect the true nature of spatially distributed processes (3).
- GIS helps us to model alternatives of actions and processes operating in the environment (2), to anticipate possible results of planning decisions (4), and to make better decisions.

This project demonstrates the advantages of applying GIS to solve practical problems from the above three aspects. An EIS requires extensive data about natural resources, land uses, infrastructure, and distribution of many interrelated socioeconomic factors. The accuracy and availability of required data depend on the scope of a study and the size of the study area. Our study shows that GIS, with its data retrieval, analysis, and reporting abilities, significantly improves the analysis. GIS helps to collect data at various scales, store data, and present data in forms that allow the project team to carry out the study in an innovative way.

Phase I Study

Phase I of the traffic relief study was completed in 1993. The goal of Phase I was to acquire understanding of the general features of the area and to use a traffic demand model for delineating an area for detailed study. The study area is approximately 52 square miles. To provide data for the preliminary analysis and the traffic demand modeling, the team developed baseline data inventory with GIS. Data were primarily secondary data that came from several different sources in different formats. For example, the U.S. Census Bureau 1990 population data were in TIGER format, the U.S. Fish and Wildlife Service National Wildlife Inventory (NWI) files in digital line graph (DLG) format, the U.S. Soil Conservation Service Monroe County Soils in DLG format, and the U.S. EPA Monroe County Natural Areas in ARC/INFO format. The majority of data sources were at scales between 1:15,000 and 1:24,000. With GIS tools, the team integrated these baseline data into a common presentation scale and projection. This process ensured an effective and comprehensive spatial analysis in the study area.

The team arranged and stored data in layers according to themes. Examples of the data layers included:

- Road center lines.
- Twenty-foot elevation contours.
- Utility lines.
- Water features.
- Subdivision boundaries.
- 1990 Census population by Census Tracts and Blocks.
- Flood plains.
- Geological formations.
- Public facilities, including schools, churches, and cemeteries.
- Political boundaries.
- Hazardous waste locations.
- Potential archaeological areas.

With these data layers, the team generated a series of 17 thematic maps to describe the features of the study areas. All maps were plotted on E-sized papers (48 inches x 36 inches) with the same map layout. The general reference map served as a base map for the other themes. It included several data layers to provide geographic references to the study area.

In addition to the base map, each individual theme map showed only one theme at a time, such as soils, subdivisions, and wetlands. Some theme maps showed derived data from the original data layers. For example, in developing a slope theme map, the team first built a three-dimensional surface from the 20-foot elevation contours, then calculated slope in degrees and aggregated areas based on a 10-degree interval. The slope theme map showed the result from the data processing. In addition, the project team created summary statistics tables to help team members gain knowledge about the study area. Table 1 is an example of the summary statistics for land use categories.

Table 1. Phase I Statistical Summaries for Land Use Categories

Land Use	Acres
Urban	10,628
Agricultural	1,183
Rangeland	12
Woodland	20,957
Water	1,394
Wetland	1,428
Transitional	422

During preparation of the Phase I inventory data and summary statistics, traffic planners performed traffic demand modeling to determine new road connections that would provide a minimum acceptable level of service in the year 2015. The modeling result was loaded into the GIS and converted into the same format and projection as other inventory data. Figure 2 displays the boundary that the traffic demand model delineated and the actual Phase II boundary. The two boundaries were not the

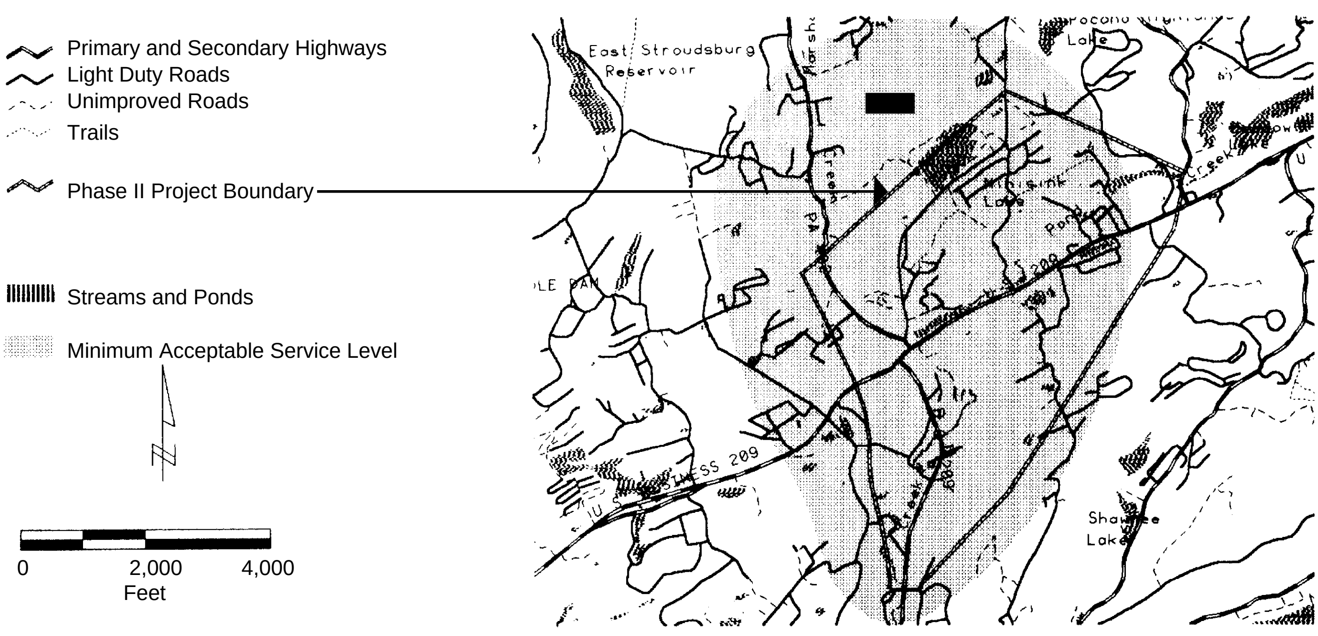


Figure 2. Phase II project area.

same. The thick line enclosed a Phase II study area that was delineated based on the traffic demand modeling and the team's understanding of environmental and other factors in the project area.

Phase II Study

The Phase II study is still ongoing. The area for the Phase II study is much smaller than that for Phase I. It is about 3.2 miles by 2.5 miles, or approximately 4 square miles.

The objective of the Phase II study is to conduct a detailed analysis for delineating a full range of feasible highway alignment alternatives. The alternatives must meet the needs of relieving traffic congestion and minimizing its impact on environmental and cultural resources.

Because the accuracy of the Phase I data was not sufficient for the Phase II study, the project team has collected data using different approaches to develop a similar set of baseline data at a finer scale. The major data source has been the photogrammetry data provided by PennDOT at a 1:2,400 scale. The data include:

- Road cartways
- Five-foot elevation contours
- Utility lines
- Water features
- Buildings footprints
- Bridges

The team directly digitized tax parcel boundaries from Monroe County Tax Assessor's maps that range in scales from 1:1,200 to 1:4,800. After digitizing each map sheet separately, the team merged them together to create a continuous parcel layer. The data layer has been adjusted to fit with the PennDOT photogrammetry data although the two data sets do not seem to match exactly. In addition, digital orthophotographs at 5-foot pixel resolution were also obtained for the project.

From these baseline data, the team has constructed Phase II data layers in four different ways.

The first approach digitizes from compilations on project base maps. For example, the team creates a land use data layer from the digital orthophotographs and infrared photography. The GIS group first plots the digital orthophotographs on a set of 1:2,400-scale map sheets. Road cartway and water features are plotted on top of the orthophotographs. Then land use specialists delineate land use boundaries with fine color markers and code land uses on maps according to the Anderson land use classification. In the end, the GIS group digitizes the land use boundaries from the compilations to create a land use data layer. Similarly, the 100-year flood plain data layer is delineated from compilations on project base maps with 5-foot contours and digitized.

The second approach derives new data layers from existing data. Buildings and structures are plotted at a 1:2,400 scale. Both historians and environmental engineers use the plots in their field investigations. After historians identify historic-eligible buildings on the plots, the GIS group develops an attribute data file that links the historic inventory data to the building geometry. Similarly, field investigations identify buildings and structures associated with contamination sites. The system stores types of contamination as building attributes. By overlaying the building data layer with the tax parcel data layer, the team can identify properties on which historic buildings or contamination sites are located.

The third approach constructs data layers by referencing Phase I data. For example, Phase II subdivision boundaries are derived from the digitized tax parcel boundaries by referring to the Phase I subdivision boundaries. Phase I subdivisions were manually compiled at 1:24,000 using approximate location, which did not align very well with the more accurate tax parcel boundaries. Using a 1:7,200-scale plot that shows both Phase I subdivision boundaries and the Phase II tax parcel boundaries, planners can verify and indicate properties associated with each subdivision. These properties are dissolved to create new boundaries for subdivisions that precisely fit with tax parcels. The same approach is used to refine public parks and private recreation areas.

The fourth approach obtains spatial data with GPS. The GPS surveyors collect accurate locational data about key features, such as boundaries of wetlands, site locations for hazardous waste, and locations of archaeological field samples. GPS data also supplement existing data, such as delineating footprints of new buildings to update the PennDOT baseline data. The integration of GIS and GPS provides the project team with accurate and up-to-date data.

Phase II data layers have provided much richer information for a detailed study of environmental features. They are merged in many different combinations to show the spatial distributions of different factors from different perspectives. Table 2 lists some of the map themes created for the Phase II study. All maps are plotted at a 1:7,200 scale.

In addition to using GIS as a data library and map production tool, we use GIS to support decision-making in two ways. First, the creation of a composite data layer has revealed the impact of alternative alignments on several composite constraints. The composite data layer is an overlay of several inventory layers and shows the various factors coincident at any location, and the relative importance of these factors.

Table 2. Selected Phase II Map Themes

Theme	Description
General reference	Road networks, hydrographic features, churches, municipal boundaries, and utilities
Parcels	Tax parcel boundaries
Community facilities	Public parks and private recreation, cemeteries, and public buildings
Flood plains	100-year flood plains
Land use	Current land uses
Subdivision	Approved subdivisions
Slope	Areas delineated with 5-degree slope intervals
Wetlands	Wetlands
Historic resources	Historic buildings and properties
Hazardous wastes	Hazardous waste sites and related buildings

These constraints have been identified by citizens, government agencies, and the project team. The major composite constraints include wetlands, historic properties, steep slopes (slope greater than 15 degrees), public parks and private recreation areas, 100-year flood plains, potential archaeological areas, prime agricultural land, subdivisions, and existing buildings and structures.

Two maps have been created from the composite constraints layer. One map shows the number of coincident constraint layers that occurs in any one location (see Figure 3). The other map shows the composite relative importance of coincident features. Both maps present a "sensitivity surface" view of the project area.

Secondly, the team uses GIS to perform interactive summary statistics for each alternative alignment. The project team analyzes the impact of the alignment alternatives on each individual constraint layer.

Two approaches define the impact areas for comparison. The first set impact areas are the areas enclosed by the footprint that traffic engineers delineated for each alternative. The second set impact areas are 300-foot buffers on both sides of the alignment delineated by the traffic engineers. The boundaries of the impact areas overlay on constraint data layers. Figure 4 displays wetlands crossed by alignment alternative ROW1B.

A set of summary statistics are calculated for each alignment. In the end, we compare the statistics for each alignment in a matrix (see Table 3). The matrix arranges constraints as rows and alignment alternatives as columns. The statistics include acres of selected features within each impact area, such as wetlands or high-quality watersheds, and total counts of features, such as historic-eligible buildings. The summary statistics also include listings of building names for businesses or public facilities within the impact areas.

The team has repeated the summary statistics several times as alignments shift. This procedure ensures that the final selected highway alignment minimizes environmental impacts, best meets project needs, and is the most cost-effective alignment to construct. The statistical matrix of impacts versus alignments is one of the critical evaluation criteria for comparing alignments and ultimately for selecting the final alignment.

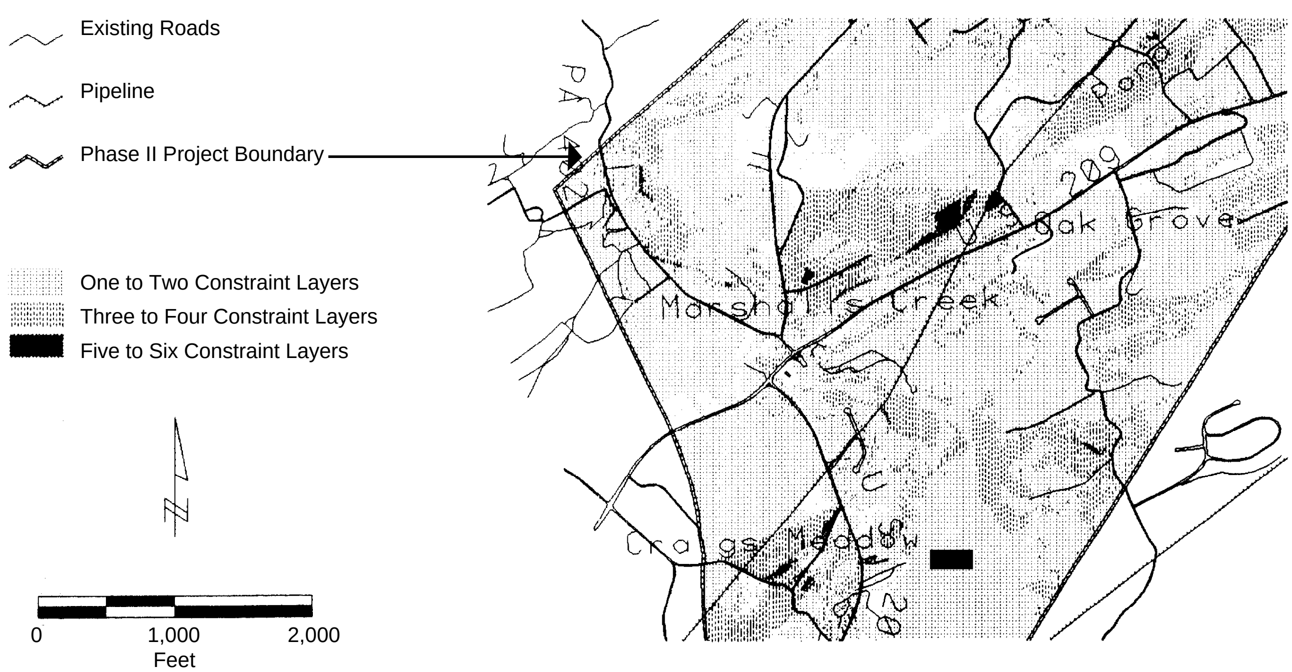


Figure 3. Composite constraints.

Table 3. Phase II Summary Statistics by Alternative

Constraints	Alignment Alternatives				
	ROW1A	ROW1B	ROW2A	ROW2B	ROW3
Wetlands					
Acres	2.76	3.27	3.25	1.8	13.64
Count	13	14	17	15	17
High-Quality Watersheds					
Acres	66.18	85.62	71.92	94.58	82.90
Hazardous Waste Parcels					
Acres	99.50	116.77	98.25	120.68	215.38
Count	8	5	11	10	6
Parcels With Historic-Eligible Buildings					
Acres	14.20	0.10	0.14	0.14	3.46
Count	3	1	2	2	2
Historic-Eligible Buildings					
Count	1	0	1	1	1

Benefit of Using GIS

In recent years, many federal, state, and local agencies have been actively acquiring and automating digital data (5). These databases provide various types of information at scales that are appropriate for a preliminary study covering a large area.

A more detailed study, which usually covers a smaller area, often requires more accurate data to describe the spatial distribution of relevant factors. GIS is flexible, allowing use of data at the scale and accuracy appropriate to the study purpose. The team has found that this feature helps improve the efficiency of the project without sacrificing the accuracy. This study has required two sets of scales ranging from 1:24,000 scale for Phase I, which required projectwide socioeconomic and environmental assessments, to 1:2,400 scale for Phase II, which requires detailed analysis for design of alternative alignments.

GIS has served as a digital database manager to assemble environmental, traffic, geographic, socioeconomic, and other data into a centralized project database. Data analyzed in this study originate from a variety of sources, such as PennDOT, U.S. Geological Survey (USGS), U.S. Census Bureau, Monroe County, and field survey. They are in different formats, including digital data in ARC/INFO, INTERGRAPH, and AutoCAD formats, GPS data, digital images, paper maps, and tabular data. Many of the public agencies and private organizations involved in this project already have digital data that GIS could easily use for this specific project. This has helped to reduce the overall costs of data collection and conversion.

This study demonstrates that GIS can support the information needs of many disciplines within a common framework and provide powerful, new tools for spatial analysis. Aside from technological considerations, GIS development initiates a higher-order systematization of geographic thinking (3), which is crucial to the success of a transportation project. In this project, GIS helps to determine the total impacts of alternative alignments on identified constraints. The team accomplishes this by superimposing the alternative alignments on constraint

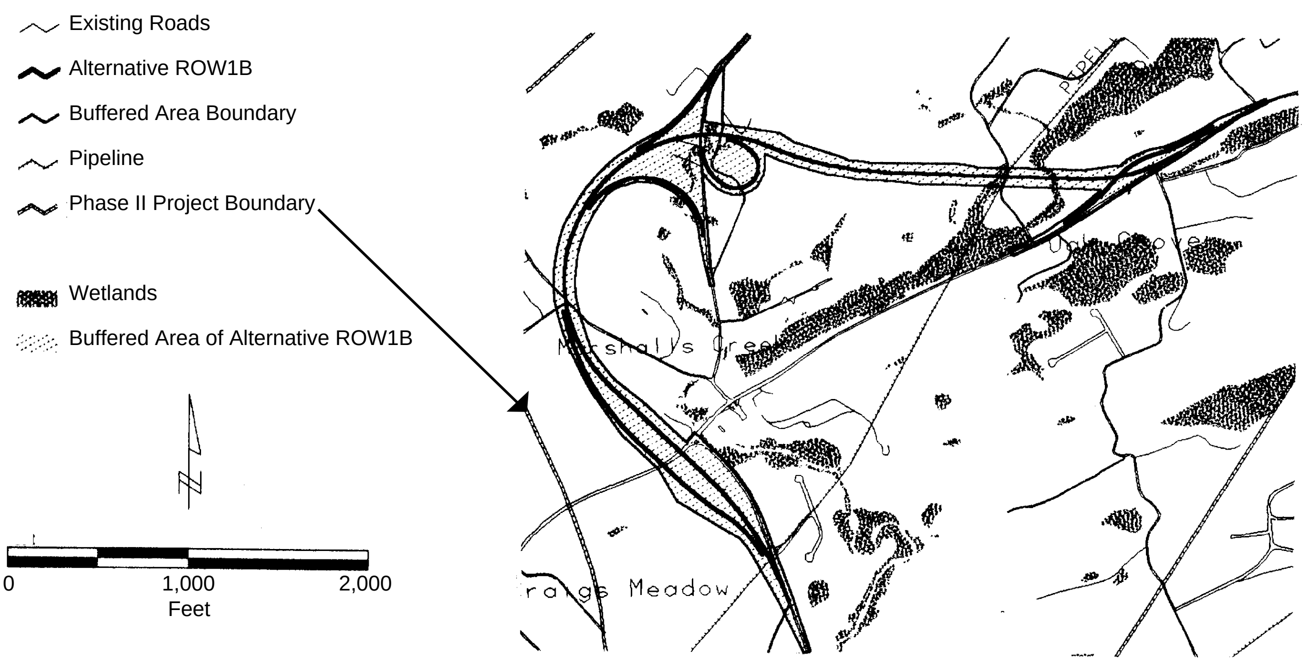


Figure 4. Wetlands crossed by an alternative alignment.

data layers to determine the total amount of each constraint layer that each alignment encounters. This approach supports the analysis of multiple alignments across the constraint surfaces for a variety of alternative scenarios. The spatial analysis tools and statistical function embedded in GIS prove to be very useful in such study.

Digital data that the GIS stores are used to summarize environmental, social, and economic data in many different ways. These functions include summing total acreage, listing entities of special interest, and counting numbers to provide useful baseline statistics for various alignments. Within a day, GIS accomplished what would have required several months of staff labor; GIS summarized impacts of 11 alternative alignments on all constraint layers. In addition, GIS creates composite constraints from individual data layers. A composite constraint data layer is created through a series of overlays to illustrate geographic coincidence of inventory themes.

Conventionally, engineers in a project such as this first delineate alternatives for alignments from the engineering perspective; they often consider factors such as steep slopes and costs. Then, other professionals, such as environmental specialists, historians, and planners, evaluate the alternatives from each point of view. GIS makes possible an early integration of environmental and engineering activities, ongoing communication with funding agencies and the public, and continual integration of a multidisciplinary team.

GIS helps to maintain high-quality data for the project. It allows for error checking and quality control of multiple data layers that would not be possible with conventional mapping. The team always compares a new data layer with other data to check for conflicts. Making check plots allows for quick identification of errors and missing data.

For instance, in the process of assigning building-use attributes to existing buildings, the GIS team first plotted buildings on a map and created a table with building identifiers. The field team then used the unique identifiers shown on the plots when noting building names and building uses in data collection tables. After relating the data table with the building data layer, the team found some buildings that did not have building use data. Moreover, some buildings were assigned uses that were out of range or seemed out of place, such as a residential building surrounded by several commercial buildings. The team highlighted the data for those buildings and sent them to engineers for verification. Through this data cleaning process, the team was able to obtain complete building use information.

Data quality directly affects project quality. Without GIS, this type of study often involves using and comparing maps at different scales, which frequently introduces serious errors. For each phase of this study, project team members have used a fixed-base map scale for all compilations. During data entry and data transformation, the team has kept accurate registration between data layers, ensuring data of the same resolution. In addition, the team has used FREQUENCY, one of the tools that GIS provides, to look for data values that are out of range as well as missing data. GIS tools have also helped to derive new relationships for features. For example, dissolving parcels has helped to create subdivision outlines, or overlaying historic buildings with parcels has helped to find the parcels on which they are located.

Planners, environmental specialists, historians, and landscape architects on the project team are responsible for field data collection, verification, and if necessary, compilation of field data into the standard project database. Wherever possible, the team has used GPS to eliminate the task of manual compilation and to improve accuracy of locating data. A fundamental requirement in applying the technology appropriately is to understand its capabilities, requirements, and limitations. Because several members manage inventory attributes, they need to know how to maintain unique identifiers for features so they can link up to the geometry. Because AutoCAD data transfers occur routinely with engineers, it is necessary to structure how the AutoCAD drawing files can be organized, as well as how certain attributes can be transferred by line color, layer name, or line width. The GIS group coordinates closely with other team specialists to identify quick, accurate, and cost-effective methods of data collection, data analysis, and presentation. In conjunction with the progress of the project, specialists from different fields have become familiar with the concept, requirements, and use of GIS. They now feel comfortable discussing alternatives while looking at results displayed on a computer screen.

GIS has created high-quality map products for public presentations and reports. GIS has also been used in several agency coordination meetings to display data and alternative alignments. GIS has allowed for different data layers to be displayed on the screen with specific combinations of features at various scales. Public agencies and citizens have been impressed by the clear and friendly graphic response to their questions. They have expressed interest in using the technology in future projects.

Once the study is complete, digital data assembled in this project will be an excellent resource for future projects in the study area. For these reasons, GIS provides more cost-effective project support for gathering, managing, and using data than that provided by paper and mylar maps.

Summary

For this project, the importance of innovation based on a solid scientific foundation cannot be overstated. In the current economic and regulatory climate, sound GIS methods are emerging as the only convincing and cost-effective means for locating, designing, and gaining approval for major public and private infrastructure projects.

The technology offers new and exciting tools for transportation planning.

The methodology that this project team has used can be successfully applied to other projects that require environmental assessment. The team has found GIS to be an extremely useful tool as users continue to learn its capabilities and the multiple tools that it offers. The regulatory agencies have repeatedly made favorable comments on how GIS can offer interactive viewing in a show-and-tell environment. This project is one of the first EIS projects to use GIS in a PennDOT-funded project. PennDOT appears convinced that GIS is an important component for conducting EIS for highway studies. More EIS projects probably will demand GIS services as part of the project approach, in part, because many federal and state regulatory agencies are increasingly using GIS.

References

1. Fischer, M.M. 1994. From conventional to knowledge-based geographic information systems. Computer, Environment, and Urban Systems 18(4):233-242.

2. Star, J., and J. Estes. 1990. Geographic information systems: An introduction. Englewood Cliffs, NJ: Prentice Hall, Inc.

3. Bracken, I., and C. Webster. 1990. Information technology in geography and planning, including principles of GIS. London, England: Routledge.

4. Burrough, P.A. 1986. Principles of geographic information system for land resources assessment. Oxford, England: Clarendon Press.

5. Hendrix, W.G., and D.J.A. Buckley. 1992. Use of a geographic information system for selection of sites for land application of sewage waste. J. Soil and Water Conserv. 47(3):271-275.

A Watershed-Oriented Database for Regional Cumulative Impact Assessment and Land Use Planning

Steven J. Stichter
Division of Coastal Management, State of North Carolina Department of Environment, Health, and Natural Resources, Raleigh, North Carolina

Introduction

In 1974, North Carolina passed the Coastal Area Management Act (CAMA) to guide growth and development in the state's coastal zone. Today, the Division of Coastal Management (DCM), under the direction of the governor-appointed Coastal Resources Council, implements CAMA. DCM's jurisdiction covers the 20 counties that border either the Atlantic Ocean or the Albemarle-Pamlico estuary.

This coastal region comprises a diverse set of human, animal, and plant communities. A broad array of coastal plain ecosystems occurs in this area, from the barrier dunes and maritime forests of the outer banks to cedar swamps and large pocosin complexes of interior areas. This area includes some of the state's fastest growing counties and some that are losing population. Urban centers such as Wilmington do exist, but the region remains primarily rural.

In recognition of the 20th anniversary of the passage of CAMA, the governor designated 1994 as the "Year of the Coast." Associated celebrations, panels, and studies highlighted the unique features of the North Carolina coast, successes of coastal management in the state, and unresolved problems and concerns. Problems remain despite protection efforts by various agencies. For instance:

- Fish landings have dropped dramatically of late.
- Shellfish Sanitation recently closed a set of shellfish beds located in outstanding resource waters.
- Shellfish statistics show that the quality of the state's most productive coastal waters continues to decline.

Because coastal North Carolina as a whole is growing more rapidly than any other section of the state, pressures on coastal resources can only continue to increase.

Declining water quality and associated sensitive habitats, resources, and animal populations have prompted several state agencies to develop new approaches to environmental protection that incorporate a broader, natural systems perspective. The North Carolina Division of Environmental Management is developing river basin plans to guide point and nonpoint water pollution control efforts. The DCM has begun work to assess and manage the cumulative and secondary impacts of development and other land-based activities by using coastal watersheds as the basis for analysis. The goal of this work is to expand the regulatory and planning programs in order to better address cumulative impacts. This paper describes the approach that DCM has developed for cumulative impacts management, with special emphasis on the use of a geographic information system (GIS). The project described here is scheduled for completion by the fall of 1996.

Cumulative Impacts Management

The concept of cumulative impacts management (1) is not new to North Carolina's coastal program. CAMA requires the consideration of cumulative impacts when evaluating development permits within defined areas of environmental concern. A permit must be denied if "the proposed development would contribute to cumulative effects that would be inconsistent with the guidelines. . . ." Cumulative effects are defined as "impacts attributable to the collective effects of a number of projects and include the effects of additional projects similar to the requested permit in areas available for development in the vicinity" (2). Despite this directive, few permitting actions have been denied because of cumulative effects; the existence of limited impact data and a dearth of viable analysis approaches have restricted application of this rule.

Since the passage of the National Environmental Policy Act (NEPA) in 1969, many attempts have been made to define and assess cumulative impacts. The Council on Environmental Quality developed the most familiar definition in its guidelines for NEPA implementation. It defines cumulative impact as:

> . . . the impact on the environment which results from the incremental impact of the action when added to other past, present, and reasonably foreseeable future actions, regardless of what agency or person undertakes such other actions. Cumulative impacts can result from individually minor but collectively significant actions taking place over a period of time (3).

Although this definition focuses the discussion of cumulative impacts, it provides little guidance on how to carry out such an analysis. Selecting both an appropriate time frame for the assessment (how far into the past and future to carry the analysis) and appropriate boundaries for the study (municipal or county boundaries, watersheds, ecoregions) are but two of the questions that require answers to successfully investigate cumulative impacts. Such decisions become even more complex when incorporating the limits imposed by available data and existing management structures.

Rigorous cumulative impact analysis is a difficult proposition. It requires identification of all sources of degradation that affect a given resource. The next step involves assigning relative significance to each of these sources along with any impacts that result from additive or synergistic interaction between sources. Assessment of the impacts of a pier on surrounding sea grasses, for instance, must include not only impacts related to the structure, such as shading and wave or current changes, but also such ambient impacts as natural wave and wind effects, upland runoff, and varying salinity.

All these investigations require the availability or collection of baseline environmental data at an appropriate spatial and temporal scale. Quantifying all the sources and causal pathways that affect a resource is extremely complicated in all but the simplest of systems. Assigning proof of significant impact is difficult unless the cause is clear and direct.

Because of the difficulties associated with assigning cause in cumulative impact analysis, especially in a regional review, DCM has chosen a different approach. It is focusing instead on locating areas at high risk to cumulative impacts. Impacts management studies and responses can then target the areas at greatest risk of degradation. Changing the scale of analysis from the site to the region requires applying some simplifying assumptions. The first assumes that any existing resource degradation results from the cumulative impact of all sources within the system boundaries. Locating such areas is relatively straightforward because most natural resource fields have developed measurements and indicators for locating degraded resources. The second assumption claims that a sufficiently intensive concentration of activities within a limited area will result in cumulative impacts on the affected system. Determining a threshold beyond which impacts cause degradation is much harder than locating already degraded resources because the level of such a threshold depends upon both the strength or spatial concentration of the impacts and the sensitivity of the resource.

Working from these simplifying assumptions, DCM's first step in assessing regional cumulative impacts is to identify areas within coastal North Carolina that exhibit symptoms of resource degradation, contain a concentration of activities that affect resources, or contain a concentration of sensitive resources. The use of categories of resources and impacts have helped to focus this search. These eight cumulative impact, high-risk area categories are:

- Impaired water quality
- High potential for water quality impairment
- Sensitive ground-water resources
- Impaired air quality (present or potential)
- Historical rapid growth
- Anticipated high growth
- High-value resources
- Productive and aesthetic resources

This set of categories is presently under public review. The next step is to develop indicators of the presence of impacts or resources appropriate to each of these categories. These indicators, when applied to a database of information about the study area, will help identify those locations at high risk as defined by the eight categories.

The regional cumulative impacts assessment approach that DCM developed is a hybrid of various assessment techniques. The overall approach is grounded in the theory and methods of site-specific cumulative impact assessment. Determination of high-risk categories and appropriate indicators and indexes is closely associated with both relative risk assessment procedures and geographic targeting. By focusing on known causes and effects of cumulative impacts on terrestrial and aquatic natural resources instead of attempting to quantify all impact pathways, available data and analysis techniques can help assess relative risk of cumulative impacts.

A Watershed Database for Cumulative Impact Assessment

High-risk categories and indicators of degradation or sensitivity are useless without information on the location of sensitive resources and impact sites. Consequently, a comprehensive database of information about coastal

North Carolina is central to cumulative impacts management in this area. The form of any database determines what types of questions to ask it; the selection of boundaries has been central to this study.

County boundaries constitute the most typical reporting unit in DCM operations. Counties determine the boundaries of DCM's jurisdiction, and the great majority of statistics used in planning and assessment are available primarily or solely by county. County size, heterogeneity, and the small number of counties available for comparison, however, have made county boundaries inappropriate for this project. Because the study focuses on impacts on natural resources, clearly the most appropriate boundaries would relate more directly to those resources.

Although using a single set of boundaries may not be appropriate for assessing impacts on all resources, management constraints limit the choice to one boundary type. Because the primary resources of concern are water based, watersheds were considered most appropriate. Surface waters receive the integrated effects of activities within a watershed; such boundaries fit intuitively with the concept of cumulative impact assessment. The number of water-related resources of concern also supported this choice. This analysis used small watersheds (5,000 to 50,000 acres) delineated in 1993 by the Soil Conservation Service for the entire state of North Carolina.

The Population, Development, and Resources Information System (PDRIS), which was designed for this project, is a PC-based, watershed-oriented database that contains the following information about the coastal area:

- Natural resources
- Population and housing
- Agricultural activities
- Economic activities
- Development activities

Table 1 includes a list of database fields. The presence and extent (or absence) of each of the features that this database represents will be available for each coastal watershed. The small watershed orientation of this study is only possible because of the availability of GIS; the volume and complexity of the watershed boundaries preclude any other assessment tool. In fact, 348 of these watersheds fall wholly or partially in the 20-county region. Figure 1 shows a map of these small watersheds. This map indicates county boundaries and shorelines in solid lines and the watershed boundaries in gray.

Data Needs and GIS Analysis

Over the past 5 years, North Carolina has actively collected a large amount of natural resource and base map information in GIS form. Research and funding associated with the Albemarle/Pamlico Estuarine Study (APES), a national estuary program study, spurred much of this data development in the coastal area. The state maintains a central repository for geographic data at the North Carolina Center for Geographic Information and Analysis (CGIA). Table 2 lists the general types of information available from the state database. The availability of data in GIS form is but one criterion for selecting a data set for use in this analysis. To be useful, the scale and accuracy of the data must be appropriate to the analysis.

Data Scale

The majority of data in the state's GIS database was collected at a scale of 1:100,000. Broader use and interest will probably urge the development of data layers at finer scales. A recently released layer of closed shellfish waters, for instance, was created at 1:24,000 scale. This prompted an update of the associated shoreline coverage to the same base scale. A handful of state departments and divisions, including Coastal Management, now use global positioning systems to collect even more precise locational information. This scale suits DCM's regional cumulative impacts scan, which is based on summary values for entire watersheds. More detailed intrawatershed planning and analysis would require finer scale data. A scale of 1:24,000 delineated the watershed boundaries in this project.

Mixing these 1:24,000 boundaries with the 1:100,000 data sets, however, can cause problems. For instance, a number of watersheds were designated for the large open water areas in Albemarle and Pamlico sounds. Although these should comprise exclusively water, overlay analysis of these watershed boundaries on the TIGER-derived census boundaries (1:100,000 scale) resulted in the assignment of small population counts to some of these watersheds. Individually locating and correcting such discrepancies is necessary.

Database Accuracy

Data layer accuracy problems are difficult to identify and assess. Because other agencies developed the majority of data used in this project, these source agencies must be relied upon for accuracy assessment of the source data. CGIA, steward of the state GIS database, adheres to National Map Accuracy Standards for all GIS data that it maintains. CGIA delivers metadata reports with any data; these reports include the source agency, collection date, and scale for the information used to derive the GIS layer. Descriptions of data lineage (collection and processing procedures), completeness, and positional

Table 1. Population, Development, and Resource Information System: Database Fields

Agriculture: Livestock and Poultry
Beef feedlots (< 300 head, > 300 head)
Dairy farms (< 70 head, > 70 head)
Hog farms (< 200 head, > 200 head)
Horse stables (< 200 head, > 200 head)
Poultry farms (< 15,000 birds, > 15,000 birds)

Agriculture: Farming
Land in farms (acres, % of HU)
Land with best mgmt. practices (acres, % of HU)
Land w/o best mgmt. practices (acres, % of HU)
Land in conservation tillage (acres, % of HU)
Land w/o conservation tillage (acres, % of HU)
Harvested cropland (acres, % of HU)
Hay crops (acres, % of HU)
Irrigated land (acres, % of HU)
Pasture land (acres, % of HU)
Row crops (acres, % of HU)

Primary
Estuarine waters (acres, % of HU)
Freshwater lakes
HU name
Receiving HU
Receiving water body
Primary water body
Secondary water body
Shoreline
Waterways w/vegetated buffers (miles, % of HU)
Population 1970
Population 1980
Population 1990
Population growth 1970 to 1980
Population growth 1980 to 1990
Counties
Total HU size
Land area (acres, % of HU)
Water area (acres, % of HU)
Stream length (miles)
Stream order (miles, % of stream length)

Development
Building permits—all residential
Building permits—amusement/recreation
Building permits—multifamily residential
Building permits—one-family residential
Building permits—hotels and motels
Building permits—retail
Building permits—industrial
Highway mileage:
  Total (miles)
  Primary (miles, % of total)
  Secondary (miles, % of total)
  Paved (miles, % of total)
  Unpaved (miles, % of total)
Rail lines (miles)
Increase of primary & secondary roads (miles, %)
Increase of paved vs. unpaved roads (miles, %)

Economic
Ag-related business (number, employees, income)
Farms (number, employees, income)
Fisheries business (number, employees, income)
Forestry/wood-using business (number, employees, income)
Lodging establishments (number, employees, income)
Manufacturing establishments (number, employees, income)
Marinas (number, employees, income)
Mining establishments (number, employees, income)
Recreation business (number, employees, income)
Restaurants (number, employees, income)
Retail establishments (number, employees, income)

Ground Water
Ground-water contamination incidents
Ground-water class (acres, % of HU)
Ground-water contamination area (acres, % of HU)
Ground-water capacity use areas (acres, % of HU)

Land and Estuarine Resources
Anadromous fish streams (miles, % of streams)
Coastal reserve waters (acres, % of HU)
Coastal reserve lands (acres, % of HU)
Federal ownership:
  National parks (acres, % of HU)
  National forests (acres, % of HU)
  Military reservations (acres, % of HU)
  USFWS refuges (acres, % of HU)
  Federal ownership—other (acres, % of HU)
State ownership:
  Game lands (acres, % of HU)
  State parks (acres, % of HU)
  State forests (acres, % of HU)
  State ownership—other (acres, % of HU)
Natural heritage inventory sites (count)
Primary nursery areas (acres, % of water area)
Private preservation (acres, % of HU)
Secondary nursery areas (acres, % of water area)
Threatened/endangered species habitat
Water supply watersheds (acres, % of HU)

Land Use
Total wetland area (acres, % of HU)
High-value wetlands (acres, % of HU)
Medium-value wetlands (acres, % of HU)
Low-value wetlands (acres, % of HU)
Predominant land cover

Population and Housing
Average seasonal population
Peak seasonal population
Units without indoor plumbing
Units with septic tanks
Units on central water systems
Units on central sewer
Units with wells

Permits
Air emission permits—PSD
Air emission permits—toxic
CAMA minor permits
CAMA general permits
CAMA major permits
CAMA exemptions
CWA Sect. 404/10 permits
Landfill permits—municipal
Landfill permits—industrial
Nondischarge permits
NPDES permits—industrial
NPDES permits—other
NPDES permits—POTW
Stormwater discharge permits
Sedimentation control plans
Septic tank permits

Shellfish
Shellfish waters (acres, % of HU)
Shellfish closures—permanent (acres, % of HU)
Shellfish closures—temporary (acres, % of HU)

Water Quality—Open Water
Class B waters (acres, % of water area)
Class C waters (acres, % of water area)
HQW waters (acres, % of water area)
NSW waters (acres, % of water area)
ORW waters (acres, % of water area)
Swamp waters (acres, % of water area)
SA waters (acres, % of water area)
SB waters (acres, % of water area)
SC waters (acres, % of water area)
WS-I waters (acres, % of water area)
WS-II waters (acres, % of water area)
WS-III waters (acres, % of water area)

Table 1. Population, Development, and Resource Information System: Database Fields (Continued)

Water Quality—Streams
Class B streams (miles, % of streams)
Class C streams (miles, % of streams)
HQW streams (miles, % of streams)
NSW streams (miles, % of streams)
ORW streams (miles, % of streams)
Swamp water streams (miles, % of streams)
SA streams (miles, % of streams)
SB streams (miles, % of streams)
SC streams (miles, % of streams)
WS-I streams (miles, % of streams)
WS-II streams (miles, % of streams)
WS-III streams (miles, % of streams)

Key
HU = hydrologic unit
PSD = point source discharges
POTW = publicly owned treatment work
NPDES = National Pollutant Discharge Elimination System
HQW = high-quality waters
NSW = nutrient-sensitive waters
ORW = outstanding resource waters
SA = saltwater classification A
SB = saltwater classification B
SC = saltwater classification C
WS1 = water supply classification 1
WS2: water supply classification 2
WS3: water supply classification 3

Water Quality—Use Support
Algal blooms (count, extent/severity)
Fish kills (count, extent/severity)
Streams fully supporting (miles, % of streams)
Streams support threatened (miles, % of streams)
Streams partially supporting (miles, % of streams)
Streams nonsupporting (miles, % of streams)
Waters fully supporting (acres, % of water area)
Waters support threatened (acres, % of water area)
Water partially supporting (acres, % of water area)
Waters nonsupporting (acres, % of water area)

accuracy are not available from these standard metadata reports, however.

DCM's cumulative impacts analysis also incorporates information not available from the state GIS database. Some of this information, such as business locations, is available from private data providers. Other information, especially agricultural statistics, does not presently exist in GIS form. Non-GIS formats include county statistics, voluntary compliance databases with self-reported coordinates, and other tabular databases. Typically, little quality control has been performed on any coordinate information. When the data originate from other state agencies, DCM is often the first user of the data outside of the source agency.

GIS Analysis Procedures

This study involves no sophisticated GIS analysis procedures. GIS helps to generate summary statistics by watershed for each of the database features. GIS drawing and query operations allow analysis of database accuracy. If the data are acceptable, the next step requires overlaying the watershed boundaries on the feature and assigning the appropriate watershed codes to all features that fall within the study area. Statistics can then be generated on the number of points, length of lines, acreage of polygons, or a total of any other numeric field in the feature coverage. Finally, the resulting summary file is converted to the format that PDRIS requires. A macro has been developed to complete this analysis. This macro generates a page-size reference map, performs the overlay, generates the watershed summary statistics, and converts the statistics to the final PC format.

Extra steps are necessary to analyze any information that does not already exist as a GIS coverage. Typically, these are tabular summaries associated with a specific boundary layer, such as county or U.S. Census statistics. These cases entail overlaying the watershed boundaries on the reporting unit boundaries; the data are distributed to the watershed in direct proportion to the percentage of the unit that falls into the watershed.

For instance, if a census tract falls 30 percent into watershed A and 70 percent into watershed B, 30 percent of the total tract population will be assigned to watershed A and the remainder to B. After performing all assignments, summary statistics are again generated by watershed. This procedure assumes that the distribution of the feature is even across each reporting unit. Rarely is this a valid assumption, but when the units are considerably smaller than the watersheds, as is the case with census tracts and blocks, this assumption introduces only limited errors. Watershed estimates based solely on county statistics, however, can be grossly inaccurate. When working with county information, therefore, using covariate information that ties more precisely to specific locations is necessary. Cropland location derived from the LANDSAT land cover layer, for instance, can be used to better distribute county-level agricultural statistics.

Database and Analysis Documentation

Data documentation is essential to this project. Given the large number of fields in the final database and the

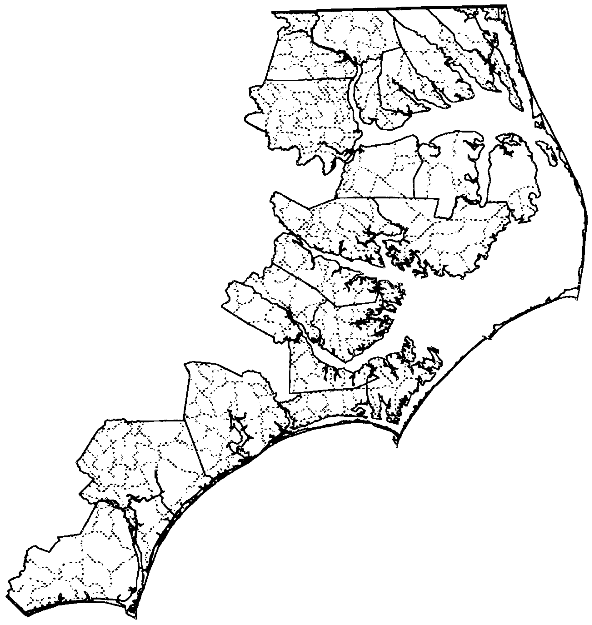


Figure 1. Watersheds in the North Carolina coastal area.

correspondingly large number of data types and sources, such documentation is key to understanding the quality of the individual database components as well as easing future database additions and updates. Because the results of this cumulative impact analysis exercise will be used to extend DCM's resource management efforts, documenting data sources and analyses will be critical if any decisions made based on this information are disputed.

A metadata database has been developed to document PDRIS data sources and analysis procedures. For each database entry, fields exist for a description and contact, collection methodology, and geographic extent of the source data. Data selection, overlay, and conversion procedures are also documented, along with any assumptions made in the analysis. In addition, recording data source, analysis procedure, and the date facilitates future database updates. Fields also record accuracy assessments for positional and attribute accuracy, logical consistency, and completeness.

The restrictions listed above regarding source data accuracy assessment, however, have limited their use. Once an entry is made to the PDRIS, all project team members receive metadata reports along with a reference map for a final review of completeness of the source data, data selection, and analysis logic. Figure 2 shows an example of a blank metadata worksheet.

Status of the Cumulative Impacts Assessment

DCM is presently gathering, verifying, and analyzing information for entry into the PDRIS. Although each source data layer was checked for accuracy before use, the logical consistency of each of the database entries relative to the other components also needs addressing. One example of such database inconsistencies is the watersheds that are covered entirely by water but also, according to the database, support a resident population. These inconsistencies could result from problems related to scale, differing category definitions, data inaccuracies, or errors in the GIS conversion or analysis at DCM. Database precision is essential for an accurate analysis and for general support of DCM's cumulative

Table 2. A Sample of North Carolina GIS Database Contents

Type	Examples	Coverage
Natural resource	Fishery nursery areas	Coastal North Carolina
	Natural heritage sites	Statewide
	Detailed soils	Varied
	Closed shellfish areas	Coastal North Carolina
	Water quality use classes	Statewide
	Detailed wetlands maps	Varied
Base data	Hydrography (24K, 100K)	Statewide
	Roads/transportation	Statewide
	County and city boundaries	Statewide
	LANDSAT-derived land cover	Coastal North Carolina
Ownership	Federal and state ownership	Statewide
Permits, waste sites	NPDES permit site	Statewide
	Landfills, hazardous waste, Superfund	Statewide
Cultural, population	TIGER boundaries, census information	Statewide
	Historic register sites, districts	Statewide

General Information:
Field
Description
Database — Units
Definition — 0 0

Source Data Description:
Contact
Data
Scale
Sample Method
Geographic Extent

Database Entry:
Procedures
Assumptions

Accuracy Assessment:

	Overall	Positional	Attribute	Logical Consistency	Completeness
Rating					
Logic Test					
Error Comment					
Value Range					

Update Procedure:
Next Update — Source
Procedure

Figure 2. Population/Development database: Data dictionary

impacts approach. Because the watershed database produced for this project will be widely available, errors and inconsistencies will undermine support for the rest of the project. Careful documentation of data sources, limitations, and analysis assumptions and procedures will provide useful support should problems or concerns arise.

Once database development is sufficiently complete (the database encompasses much dynamic data and could be constantly updated), indexes describing each of the cumulative impact, high-risk areas must be finalized. Applying these indexes to the database will allow identification of the watersheds at highest risk to cumulative impacts. Discussions held concurrently with index development will determine which management responses are appropriate to each high-risk category. Possibilities include strengthened land use planning requirements, new permit standards, or the designation of a new type of environmental critical area.

Although the data-intensive approach that DCM has chosen relies heavily on a GIS, the greatest challenges in this project do not lie in the GIS analysis. Applying this watershed-based analysis to existing political jurisdictions will be a more difficult undertaking. A convincing demonstration of the importance of including a natural systems perspective into a development permitting system, land use plan, or even economic development strategy, will ultimately contribute more to environmental protection in coastal North Carolina than any individual regulation that emanates from this project.

Summary

Twenty years after the passage of CAMA, DCM has developed a framework for a consistent approach to the problem of cumulative impacts of development. The approach and PDRIS database combine existing natural resource management techniques to locate areas of the coast at greatest risk of serious impairment from cumulative impacts. The availability of natural resource data at an acceptable scale (1:100,000) eases the development of the database essential to this analysis. The simultaneous development of a set of comprehensive small watershed boundaries for the state, along with the initial planning of this project, provided the final critical component to DCM's approach.

Perhaps more importantly, both DCM and individual local governments will have a large volume of information on natural units, which will provide an important, new perspective on the problems and prospects for local governmental action.

This project will not solve all problems related to cumulative impacts. The PDRIS will provide little support for site-specific or within-watershed cumulative impacts analysis; such an analysis at fine scales requires a much more precise database. By providing a broader-scale framework for this discussion, however, DCM's regional cumulative impacts study will hopefully further discussion, understanding, and management of cumulative and secondary impacts on natural systems.

References

1. Wuenscher, J. 1994. Managing cumulative impacts in the North Carolina coastal area. Report of the Strategic Plan for Improving Coastal Management in North Carolina. North Carolina Division of Coastal Management.
2. North Carolina General Statutes (NCGS) 113A-120.
3. 40 CFR §1508.7.

Wetlands Mapping and Assessment in Coastal North Carolina: A GIS-Based Approach

Lori Sutter and James Wuenscher
Division of Coastal Management, State of North Carolina Department of Environment, Health, and Natural Resources, Raleigh, North Carolina

Introduction

The coastal area of North Carolina covers 20 counties and over 9,000 square miles of land area, about 20 percent of the state (see Figure 1). It also includes over 87 percent of the state's surface water. The North Carolina Coastal Management Program (NC CMP) is responsible for managing this area to meet the goals set forth in the Coastal Area Management Act (CAMA) (North Carolina General Statute [NCGS] 113A, Article 7). These goals provide a broad mandate to protect the overall environmental quality of the coastal area and to guide growth and development in a manner "consistent with the capability of the land and water for development, use, or preservation based on ecological considerations" (NCGS 113A-102(b)(2)).

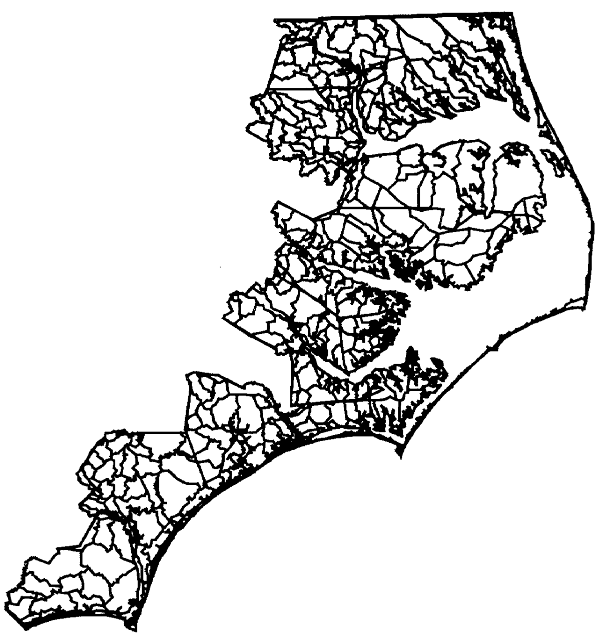


Figure 1. HU and county boundaries in the North Carolina coastal area.

Much of the North Carolina coastal area consists of wetlands, which, in many areas, constitute nearly 50 percent of the landscape. These wetlands are of great ecological importance, in part because they occupy so much of the area and are significant components of virtually all coastal ecosystems, and in part because of their relationships to coastal water quality, estuarine productivity, wildlife habitat, and the overall character of the coastal area.

Historically, close to 50 percent of the original wetlands of the coastal area have been drained and converted to other land uses (1-3). Although agricultural conversion, the largest historical contributor to wetlands loss, has largely stopped, wetlands continue to disappear as they are drained or filled for development. Conflicts between economic development and wetlands protection continue to be a major concern, with many coastal communities considering wetlands protection to be a major barrier to needed economic development.

Because wetlands are such a dominant part of the coastal landscape and are vitally important to many aspects of the area's ecology, their management and protection is a major concern of the NC CMP. The State Dredge and Fill Act (NCGS 113-229) and the CAMA regulatory program stringently protect tidal wetlands, or "coastal wetlands" as law and administrative rules call them. Coastal wetlands are designated areas of environmental concern (AECs), with the management objective "to give highest priority to the protection and management of coastal wetlands so as to safeguard and perpetuate their biological, social, economic and aesthetic values; and to coordinate and establish a management system capable of conserving and utilizing coastal wetlands as a natural resource essential to the functioning of the entire estuarine system" (15A NCAC 7H .0205).

North Carolina law does not, however, specifically protect nontidal freshwater wetlands. State protection of freshwater wetlands is limited to the regulatory authority provided under federal laws for state agency review of federal permits; in this case, §404 permits granted by the U.S. Army Corps of Engineers. Under §401 of the Federal Water Pollution Control Act (33 USC 1341), a Water Quality Certification from the North Carolina Division of Environmental Management (DEM) is required for a 404 permit to discharge fill material into wetlands. Section 307 of the federal Coastal Zone Management Act (CZMA - 16 USC 1451 et seg.) also requires that 404 permits be consistent with the enforceable rules and policies of the NC CMP. The standards for consistency are the use standards for AECs and wetlands policies stated in the applicable local land use plan. Other than AECs, the NC CMP has no consistent policies regarding wetlands. A few local land use plans include policies to protect freshwater wetlands, but most do not.

Wetlands Conservation Plan

In 1991, the CZMA §309 Assessment of the NC CMP revealed NC CMP's weakness in protecting nontidal wetlands (4). The assessment demonstrated that both opponents and proponents of wetlands protection considered the current system inadequate. Economic development interests found the 404 regulatory program to be unpredictable and inconsistent, often resulting in the loss of needed economic growth in coastal counties. Environmental interests felt that the program allowed the continued loss of ecologically important wetlands. As a result, the assessment identified wetlands management and protection as one of the primary program areas in need of enhancement.

The North Carolina Division of Coastal Management (DCM) developed a 5-year strategy (5) for improving wetlands protection and management in the coastal area using funds provided under the Coastal Zone Enhancement Grants Program established by 1990 amendments to §309 of the federal CZMA. The Office of Ocean and Coastal Resources Management (OCRM) in the National Oceanographic and Atmospheric Administration (NOAA), U.S. Department of Commerce administers the §309 program. Funds provided under this program, particularly Project of Special Merit awards for fiscal years 1992 and 1993, supported the work reported in this paper. A grant from the U.S. Environmental Protection Agency (EPA) for a Wetlands Advance Identification (ADID) project in Carteret County, North Carolina, also funded this work.

The key element of DCM's strategy for improving wetlands protection is the development of a wetlands conservation plan for the North Carolina coastal area. The plan has several components:

- Wetlands mapping inventory
- Functional assessment of wetlands
- Wetlands restoration
- Coordination with wetlands regulatory agencies
- Coastal area wetlands policies
- Local land use planning

The obvious first step in developing a wetlands conservation plan is to describe the wetlands resource. An extensive geographic information system (GIS) wetlands mapping program is helping to accomplish this first step by producing a GIS coverage of wetlands by wetland type for the entire coastal area. The GIS coverage allows generation of paper maps for areas within any boundaries available in GIS format. This is the subject of the first part of this report.

One weakness of the 404 program is that, for individual permits, it attempts to apply the same rules and procedures equally to all wetlands, regardless of the wetland type and location in the landscape. This approach can result in permits being granted for fill of wetlands of high ecological significance or permits being denied to protect wetlands of little significance. Neither outcome is desirable because the result may be the loss of either vital wetland functions or beneficial economic activity. This is an unsatisfactory way to manage wetland resources in an area such as the North Carolina coast, where:

- A high proportion of the land is wetlands.
- Many of the wetlands are vital to the area's environmental quality.
- Economic stimulation is sorely needed.

To help overcome this weakness in the current wetland regulatory framework, the Wetlands Conservation Plan includes an assessment of the ecological significance of all wetlands to determine which are the most important in maintaining the environmental integrity of the area. This will result in a designation of each wetland polygon in the GIS coverage as being of high, medium, or low functional significance in the watershed in which it exists. The procedure by which this occurs is the subject of the second part of this report.

The remaining components of the Wetlands Conservation Plan comprise the means by which the results of the wetland mapping and functional assessment steps will be used to improve wetland protection and management. Close coordination with other state and federal agencies involved in wetlands protection and management has been an important component of the entire effort. Agency representatives have been involved in developing the methods used, and the agencies will receive copies of the resulting maps for use in their own planning and decision-making. Policies for protection of wetlands of varying functional significance will be proposed to the Coastal Resources Commission to serve as the basis for consistency review of 404 permit applications. Wetland maps and functional assessment results will also be provided to local governments for use in local land use planning, and DCM will work with local governments to increase local involvement in the wetlands regulatory structure.

While the wetland maps themselves are useful for land use planning and helping to find suitable development sites, simply knowing where the wetlands are located is insufficient information for many purposes. Any area for which a 404 permit application is in process has been officially delineated as a wetland by the Corps of Engineers. The value of wetland maps to the regulatory review agencies at this stage is limited to determining the relationship of the site to other wetlands in the area. While, ideally, all wetlands should be avoided in planning development, avoiding wetlands completely in the coastal area is difficult, and avoiding all wetlands in any extensive development is virtually impossible.

The results of the functional assessment will provide additional information about the ecological significance of wetlands. This information will be valuable to wetland regulatory review agencies in determining the importance to an area's environmental integrity of protecting a particular site for which a permit to fill has been requested. It will also enable development projects to be planned so as to avoid, at all reasonable costs, the most ecologically important wetlands. An accurate functional assessment of wetland significance, then, is the most valuable component of the Wetlands Conservation Plan.

Wetlands Mapping Inventory

An important, initial step in developing a comprehensive plan for wetlands protection is to understand the extent and location of wetlands in the coastal area. When developing mapping methods, DCM quickly realized that the more than 9,000-square-mile coastal area was too large for any mapping effort in the field (see Figure 1). To complete this task in an accelerated timeframe, DCM needed to use existing data compatible with GIS. Reviewing the existing data revealed that most are not applicable for one of two reasons: (1) available wetlands data are based on older photography, and (2) more recent data are not classified with the intent of wetlands mapping. These data types, used independently, are inappropriate for use in a coastal area wetlands conservation plan. In addition, the classification schemes used in the existing methods are too complex or not focused on wetlands.

While several data sets were believed to be inappropriate if used exclusively for wetlands mapping in coastal North Carolina, each contained useful components. DCM elected to combine three primary layers of data and extract the most pertinent information from each layer. DCM selected the National Wetlands Inventory (NWI) because its primary purpose is to map wetlands. Unfortunately, these maps were based on photography from the early 1980s in coastal North Carolina, and many changes have occurred in the landscape since that time. NWI also omitted some managed wet pine areas from its maps; DCM wished to include these areas because they are important to the ecology of the North Carolina coastal area. DCM also selected detailed soils lines for use in its mapping efforts. While soils alone should not be used to identify wetlands, soils can be very useful in identifying marginal areas. Finally, DCM also employed thematic mapper (TM) satellite imagery in its methods. This data layer was not developed as a wetlands inventory; however, the imagery is more recent than the soils and NWIs. DCM desired to incorporate the benefits of each of these data sources into its mapping techniques.

The information provided by this mapping exercise will be useful to county and municipal planners in helping guide growth away from environmentally sensitive areas. For this reason, DCM elected to pursue mapping on a county by county basis. In addition, a single county allowed DCM to focus methodology development to a limited geographic area to refine its methods. Carteret County was selected as a methods development laboratory because data were available for the area and because Carteret has a large number of representative wetlands.

Data Descriptions

The U.S. Fish & Wildlife Service produces the NWI for all wetlands in the country. For the coastal North Carolina area, these vector data were developed from 1:58,000-scale color infrared photography taken during the winters of 1981, 1982, and 1983. Photointerpreters delineated wetland polygons on clear stabilene mylar taped over the photographs. After an initial scan of the photographs to identify questions or problem signatures, the photointerpreters reviewed areas in the field. They performed approximately one-half to one full day of field verification per quadrangle (quad) (6). Features were compared with U.S. Geological Survey (USGS) topographic maps for consistency. Following completion of the 'draft' paper maps, the Regional Coordinator reviewed the data. After approval as a final map, each quad was digitized. Initially, the North Carolina Center for Geographic Information and Analysis (CGIA) digitized the coastal North Carolina NWI maps, and later, the NWI Headquarters in St. Petersburg, Florida, who subcontracted the task, digitized them. Digital maps were obtained initially from 1/4-inch tape transfer and later from direct access to NWI via the Internet.

CGIA provided digital, detailed soil lines, which also are vector data based on 1:24,000 quads. County soil scientists delineated soil boundaries on aerial photographs based on slope, topography, vegetative cover, and other characteristics. This process occurs in any soil survey. After appropriate personnel approved the lines, a qualified soil scientist recompiled them onto orthophoto quads. CGIA scanned or manually digitized these lines. The coverage incorporated databases describing soil characteristics, which were then released for use.

The Landsat Thematic Mapper (TM) imagery was classified as part of the Albemarle-Pamlico Estuarine Study (APES). To provide complete coverage for the southernmost region of DCM's jurisdiction (Onslow, Pender, Brunswick, and New Hanover Counties), DCM contracted with CGIA and the North Carolina State University (NCSU) Computer Graphics Center to have that area processed identically to the APES region. These data provide a raster-based coverage of approximately 30-meter pixel resolution. Some of the imagery was taken at high tide, which precludes some near-water wetlands from appearing in certain areas. Using ERDAS, imagery specialists grouped similar spectral signatures into one of 20 classes. DCM used these data in two formats: filtered and unfiltered. The unfiltered information was vectorized with the ARC/INFO GRIDPOLY command. To remove some of the background noise in the coverage, it was filtered using ERDAS 'scan' with a Majority filter of 5 by 5 pixels, then vectorized with the ARC/INFO GRIDPOLY command.

Methods

Within each county, mapping is based on 1:24,000 USGS quads. After completion, each quad is assembled into a countywide coverage, which eventually is assembled into a coastal area coverage. The initial step in the mapping process is to ensure completion of the base layers described previously. Reviewing for errors at early stages prevents confusion in correction later in the process; therefore, the importance of the preliminary techniques cannot be overemphasized. The NWI data are first inspected to ensure complete coverage. If parts of the quad are missing, the error is investigated and corrected. Omissions may be areas of severe cloud cover on the photography or areas neglected during the digitization process. Next, the coverage is reviewed for missing label points. Any omissions are corrected based on the finalized version of the published NWI paper map. Appropriate NWI staff are contacted for the necessary information. At this time, labels are verified for typographical misentry. If not corrected, these errors could lead to confusion later in the mapping process.

Once the label errors are detected and corrected, the polygons are reviewed for completion. Verifying every line in the areas of coastal North Carolina densely populated with wetlands is impossible, but the lines are reviewed for completeness. NWI staff again must provide necessary information for any omissions. When the map is approved, technicians ensure projection of the quad to the State Plane Coordinate System. If this has not been completed, the ARC/INFO PROJECT command is employed.

The soils information is prepared in a similar manner to the NWIs, with questions being directed to the county soil scientist. Prior to the steps described previously, soils must be verified for completeness. Because soils are mapped by county boundaries and DCM maps by quad, some files must be joined in quads that intersect county boundaries. At this time, the quad must be checked for differing abbreviations between counties. Discrepancies are handled on a case-by-case basis. When an abbreviation describes different soils in different counties, a temporary abbreviation is created for one of the counties. If a single soil is described by two abbreviations across counties, both abbreviations are incorporated into the classification scheme.

The Landsat data do not require additional verification. Review of this layer is often helpful, however, to ensure that the geographic boundaries match. Cases where landforms do not appear to match require investigation of the discrepancies. If the area is misregistered, this layer might be omitted from the analyses. To date, no area has been mapped without this imagery.

The hydrogeomorphology of a wetland is unique in defining the wetland's function (7). Because these maps serve as the base for additional wetland projects (as described later in this report), an accurate determination of this characteristic is essential. Prior to the overlay procedure, technicians add a new item, hydrogeomorphic (HGM), to the NWI coverage. Because DCM considers both vegetation and landscape position in its classification (discussed later), riverine, headwater, and depressional wetland polygons are assigned an HGM of 'r,' 'h,' or 'd,' respectively. The digital line graphs (DLGs) of hydrography are essential in this step of the procedure.

All wetlands that are adjacent to streams or rivers are considered in the riverine HGM class and are designated as riverine polygons. This class should include all bottomland hardwood swamps and some swamp forests. It rarely includes any of the interfluvial wetland types. If it does, it is a small section of a large interfluvial flatwood from which a small stream emerges. Only the polygons adjacent to the stream are considered riverine. Headwaters are defined as linear areas adjacent to riverine areas that do not have a stream designated on the hydrography data layer. Because these unique systems form the transition between flatwoods and riverine wetlands, they are treated specially. Finally, polygons that exist on interfluvial divides are designated as flats or depressional wetlands. This class should not include any wetlands along streams.

The complete data coverages are overlaid to create a new, integrated coverage that often approaches 100,000 polygons. Each polygon has many characteristics assigned to it, including the Cowardin classification assigned by the NWI, the soil series provided by the detailed soil lines, the unfiltered land use/land cover code, the filtered land use/land cover code from the Landsat TM imagery, and the HGM classification assigned in the previous step.

Based on these characteristics, each polygon is assigned to one of DCM's classes through an automated ARC/INFO model using an arc macro language (AML). Personnel from the NWI and the North Carolina Department of Environment, Health, and Natural Resources Division of Soil and Water Resources have reviewed the classification of the Cowardin types into DCM wetland types. The classes that DCM currently recognizes are upland, salt/brackish marsh, estuarine shrub scrub, estuarine forest, maritime forest, pocosin, bottomland hardwood, swamp forest, headwater swamp, hardwood flatwoods, piney flatwoods, and managed pinelands. DCM also classifies soils as hydric or nonhydric based on List A of the U.S. Soil Conservation Service (SCS) List of Hydric Soils.

The base of the map is the NWI polygon coverage. Some NWI polygons are omitted from the DCM maps because they are temporarily flooded, but on nonhydric soils or because recent TM imagery indicates these areas are currently bare ground. The managed pineland wetland group on DCM maps includes areas that NWI considers uplands, identified as pine monocultures on the imagery, and that occur on hydric soil.

In addition, DCM also provides a modifier to some of these polygons. DCM notes if NWI has determined that the area has been drained or ditched. Areas designated as wetlands at the time of the NWI photography that currently appear as bare ground on the TM imagery are designated as 'cleared' on the maps. Many of the cleared areas would no longer be considered jurisdictional wetlands. These modifiers are useful indicators of the impacts wetlands sustain from human activities.

Initiation of an interactive session follows completion of the automated procedure. This session considers landscape characteristics that are not easily described to a computer model in correcting the classification. This is especially important in distinguishing bottomland hardwood swamps from hardwood flats. Both contain deciduous, broad leaf species of trees and can be temporarily flooded. The hydrology of these systems, however, is completely different. All bottomland hardwood swamps, for example, must be adjacent to a river where they receive seasonal floodwaters from the channel. Conversely, hardwood flatwoods should be located on interfluvial flats and not adjacent to any streams. Water is not introduced into hardwood flatwoods via a channel; rather, precipitation and ground water provide the water for this system. Polygons that are adjacent to rivers or estuaries but do not have a distinct channel designated in the hydrography coverage are considered headwater swamps.

During the course of methodology development, staff members visited at least 371 sites in the field. As staff members encountered new Cowardin classes, they would verify that the polygons were being placed into the correct DCM categories. If they determined that a particular Cowardin class was systematically misidentified, they updated the algorithm for automation. While this method does not provide for a usable accuracy assessment, it allowed development of the most accurate methods.

The accuracy of these data is unknown at this time. An accuracy assessment of the data is anticipated in the near future. This assessment will allow map users to understand the strengths and limitations of the data. It also will provide an overall summary of data error.

Functional Assessment of Wetlands

Certain initial considerations shaped the approach and methods used in developing a wetlands functional assessment procedure. The procedure needed to fit within the context and objectives of the Wetlands Conservation Plan for the North Carolina coastal area as described above. This context, and the opportunities and limitations it imposed, had considerable influence on the specific procedure developed.

Because we are dealing with a large geographic area with many wetlands, we recognized from the outset that we needed a method we could apply to large land areas without site visits to each individual wetland. This ruled out the many site-specific functional assessment methods that were applied in other contexts. Almost of necessity, a GIS-based approach was chosen. That meant we would have to use information available in GIS format and make use of GIS analytical techniques. The wetland mapping on which the functional assessment is based was performed using GIS, so the basic digital data were available.

The primary objective was to produce information about the relative ecological importance of wetlands that would be useful for planning and overall management of wetlands rather than to serve as the basis for regulatory decisions. While we could not visit every wetland, the goal was to predict the functional assessment value that a detailed, site-specific method would determine. We wanted to be able to predict in advance what the wetland regulatory agencies would determine as a wetland's significance so that the resulting maps would identify those wetlands where a 404 permit would be difficult or impossible to obtain. The resulting information would then be useful in determining where not to plan development. This would benefit potential permit applicants by preventing ill-advised plans that would be unlikely to receive permits and simultaneously serve to protect the most ecologically important wetlands. The result of the procedure, then, is not a substitute for a site visit in making regulatory decisions, but a predictor of what a site visit would determine.

A primary consideration was that the procedure be ecologically sound and scientifically valid, based on the best information available about the functions of wetlands. It needed to be based on fundamental principles of wetlands and landscape ecology rather than on arbitrary or subjective decisions.

Finally, the procedure was to be watershed-based. This requirement was primarily because consideration of a wetland's role in its watershed is the soundest basis for determining its ecological significance, but also because the other components of the Wetlands Conservation Plan, including wetland mapping and restoration planning, are based on watershed units. The watersheds being used are 5,000- to 50,000-acre hydrologic units (HUs) delineated by the SCS as illustrated in Figure 1. The North Carolina coastal area comprises 348 of these HUs. Watershed units of any size, however, could be used without changing the validity of the watershed-based considerations used in the procedure.

These initial considerations result in a summary definition of the functional assessment procedure. It is a GIS-based, landscape scale procedure for predicting the relative ecological significance of wetlands throughout a region using fundamental ecological principles to determine the functions of wetlands within their watersheds.

The functional assessment procedure is meant to be used with GIS data for regional application. It is not a field-oriented, site-specific method that involves visiting individual wetlands and recording information. A GIS-based procedure is the only practical approach for dealing with a large geographic area with many wetlands in a limited amount of time.

This GIS-based approach can make information on wetland functional significance available for broad regions in advance of specific development plans. The information is then available for planning to help avoid impacts to the most ecologically important wetlands. In this sense, the North Carolina procedure is unlike other functional assessment techniques that are designed for use in a regulatory context or that require field data for each wetland.

Data Requirements

Because the procedure uses GIS analysis, it requires digital information in GIS format. GIS data layers used in the procedure include:

- Wetland boundaries and types (the topic of the first section of this report).
- Soils maps.
- Land use/land cover.
- Hydrography.
- Watershed boundaries.
- Threatened and endangered species occurrences.
- Estuarine primary nursery areas.
- Water quality classifications.

In the North Carolina coastal area, these data layers either already existed and were available from the CGIA or were developed as part of the Wetlands Conservation Plan. Because other projects funded most of the data acquisition and digitization, developing the necessary GIS databases was not a major cost.

The soils coverage consists of digitized, detailed county soils maps produced by SCS and digitized by CGIA. The soils coverage allows identification of the soil series underlying a wetland, and the properties of the series are used to determine soil capacity for facilitating the wetland's performance of various functions.

The land use/land cover data layer was produced for the APES from interpretation of satellite TM imagery (8). It is used to determine land cover and uses surrounding each wetland and in the watershed.

The basic hydrography coverage consists of 1:24,000-scale USGS DLGs. Because the functional assessment procedure uses stream order as an indicator of watershed position, stream order according to the Strahler system was determined manually and added to the DLG attribute files.

As described previously, the watersheds used in the procedure are relatively small HUs delineated by SCS. DCM contracted with CGIA to have these boundaries digitized for the coastal area. During the digitization process, the watershed boundaries were rectified to USGS and DEM boundaries of larger subbasins to ensure that the HUs could be combined into larger watershed units.

A data layer produced by the North Carolina Natural Heritage Program is used to identify threatened and endangered species occurrences. The North Carolina Division of Marine Fisheries maintains the coverage of primary nursery areas, and the Division of Environmental Management developed a map of water quality classifications that was digitized by CGIA.

The ways in which these data layers are used to determine values for various parameters in the functional assessment procedure are described later in this report. The GIS procedures have been automated using ARC/ INFO

AML on a Sun workstation. The AML programs are available from DCM to anyone planning to use the procedure elsewhere.

Because the assessment procedure was designed for GIS analysis, the choice and expression of individual parameters have been shaped to some extent by the GIS data available and the capabilities and limitations of ARC/INFO techniques and AML automation. DCM was fortunate to have a relatively large amount of GIS data readily available. For use in other areas, the procedure could be modified to use different GIS coverages. At least the first five databases listed above, however, are essential to its basic propositions.

Classification Considerations

The HGM classification system for wetlands (7) classifies wetlands into categories based on landscape position (geomorphic setting), water sources, and hydrodynamics (direction of water flow and strength of water movement). It is being increasingly used as the basis for wetland classification and functional assessment systems. HGM classification focuses on the abiotic features of wetlands rather than on the species composition of wetland vegetation as do most traditional wetland classification schemes.

Several features of the HGM classification system make it a useful starting point for an assessment of wetland functions. Because the HGM system is based on geomorphic, physical, and chemical properties of wetlands, it aggregates wetlands with similar functions into classes. The HGM class of a wetland, in itself, indicates much about the ecosystem functions of the wetland. The HGM approach also forces consideration of factors external to the wetland site, such as water source. This helps relate the wetland to the larger landscape of which it is a part and puts consideration of the wetland's functions in a landscape and watershed context.

Three HGM classes are used as the starting point for the North Carolina functional assessment procedure. All wetlands are first classified as one of the following:

- Riverine
- Headwater
- Depressional

Riverine wetlands are those in which hydrology is determined or heavily influenced by proximity to a perennial stream of any size or order. Overbank flow from the stream exerts considerable influence on their hydrology. Headwater wetlands exist in the uppermost reaches of local watersheds upstream of perennial streams. Headwater systems may contain channels with intermittent flow, but the sources of water entering them are precipitation, overland runoff, and ground-water discharge rather than overbank flow from a stream. Depressional wetlands, including wet flats and pocosins, generally are not in direct proximity to surface water. While they may be either isolated from or hydrologically connected to surface water, the hydrology of depressional wetlands is determined by ground-water discharge, overland runoff, and precipitation.

The functions of wetlands in these different HGM classes differ significantly. Riverine wetlands regularly receive overbank flow from flooding streams and, thus, perform the functions of removing sediment and pollutants that may be present in the stream water and providing temporary floodwater storage. Headwater and depressional wetlands cannot perform these functions because they do not receive overbank flow. Headwater wetlands occur at landscape interfaces where ground water and surface runoff coalesce to form streams. Headwater wetlands provide a buffer between uplands and stream flow so they can perform significant water quality and hydrology functions. While depressional wetlands do not perform buffer functions, they often store large amounts of precipitation or surface runoff waters that otherwise would more rapidly enter streams. Wetlands in all HGM classes can perform important habitat functions.

Because the wetlands in these different HGM classes are functionally different, their functional significance is assessed using different, though similar, procedures. If the same procedure were used for all HGM classes, depressional wetlands would always be considered of lower functional significance simply because they are not in a landscape position to perform some of the water quality and hydrologic functions of riverine and headwater wetlands.

In addition to HGM classes, wetland types identified by dominant vegetation are used at several points in the functional assessment. This reflects a recognition that the biologic properties of a wetland site considered together with its hydrogeomorphic properties can provide a more detailed indication of its functions than either taken alone. The HGM class of a wetland, as a broad functional indicator, determines which assessment procedure to use. Within each HGM class and corresponding assessment procedure, wetland type determines the level or extent of specific parameters.

The wetland types used are those typical of the North Carolina coastal area. They result from a clumping of the Cowardin classes used on NWI maps into fewer types with more intuitively obvious type names (e.g., swamp forest, pocosin), as described previously. These wetland types are used in the wetland maps that form the starting point for the functional assessment.

Wetland types are used in the procedure as indicators of functional characteristics. Correlations between wetland type and wetland functions were determined from statistical analysis of field data from nearly 400 sites. At each site, the presence or absence of a list of functional indicators was recorded. Dr. Mark Brinson of East Carolina University developed the functional indicators lists, in part. Dr. Brinson served as primary scientific consultant in developing the HGM classification system and the field sampling methodology.

Wetland types differ in other areas, so their inclusion in this procedure limits its use in its current form to the southeastern coastal plain. Adaptation of the procedure for use in other areas would require either extensive field sampling as was performed in coastal North Carolina or a more arbitrary clumping of wetland types based on best professional judgment. Other methods of wetland classification could be used, provided wetlands are classified in such a way that functional characteristics of the wetland types are constant and can be determined by field sampling, literature values, and/or professional judgment. The procedure could be applied directly to NWI polygons if these are the only wetland map base available.

In addition to wetland type, several other parameters are used as indicators of the existence or level of specific wetland functions. These include both site-specific parameters, such as wetland size and soil characteristics, and landscape considerations, such as watershed position, water sources, land uses, and landscape patterns. GIS analysis determines values for these parameters based on the data layers discussed above. They could be determined manually, but the process would be very labor intensive.

Unlike assessment procedures that depend solely on information that can be collected within a wetland, this procedure relies heavily on factors external to the wetland site itself. Relationships between a wetland and the landscape within which it exists are integral considerations in determining wetland functional significance. Characteristics of the landscape surrounding a wetland are often more important determinants of its functional significance than are the characteristics of the wetland itself. Of the 39 parameters evaluated in the procedure, 21 are landscape characteristics, and 18 are internal characteristics of the wetland itself.

While we believe this emphasis on a wetland's landscape context is a more ecologically sound approach to functional assessment than site-specific methods, it requires a great deal more information than could be collected within the wetland itself. The procedure is based on GIS data and analysis, not only to make it suitable for regional application, but because GIS provides the most practical way to analyze the spatial relationships of landscape elements and their properties.

Structure of the Assessment Procedure

The assessment procedure uses a hierarchical structure that rates individual parameters and successively combines them to determine the wetland's overall functional significance. The complete hierarchical structure is illustrated in Figure 2. It consists of four levels:

- Overall functional significance of the wetland.
- Specific functions and risk of wetland loss.
- Subfunctions.
- Parameters evaluated to determine the level and extent of functions.

The objective of functional assessment is to determine an individual wetland's ecological significance in its watershed and the larger landscape. The highest hierarchical level, or end result of applying the procedure, then, is the wetland's overall functional significance.

The second hierarchical level includes the four primary factors that are considered in determining the wetland's functional significance (see Figure 3). The overall ecological significance of a wetland is determined by the degree to which it performs, or has the capacity to perform, specific functions. The broadest grouping of wetland functions includes water quality functions, hydrologic functions, and habitat functions. The nature of the landscape and the water characteristics of the watershed in which a wetland functions also determine ecological significance to some extent. These factors determine the potential risk to watershed and landscape integrity if the wetland functions were lost. Including a "risk factor" as a basic consideration in functional assessment also provides a means of considering cumulative impacts and the practicality of replacing lost functions through mitigation in determining a wetland's overall significance.

Each primary function of wetlands is actually a combination of separate, more specific subfunctions. Water quality subfunctions include the removal of nonpoint source pollutants from surface runoff and the removal of suspended or dissolved pollutants from flooding streams. Hydrology subfunctions include storage of precipitation and surface runoff, storage of floodwater from streams, and shoreline stabilization. Habitat subfunctions include providing habitat for both terrestrial species and aquatic life. Several considerations that, while not truly wetland functions, are called subfunctions for parallelism also determine risk factor. The subfunction levels of the assessment procedure are illustrated in Figures 4 through 7.

Properties of the wetland and its surrounding landscape determine the extent to which a wetland performs these different subfunctions. The assessment procedure refers to these properties as "parameters." Parameters

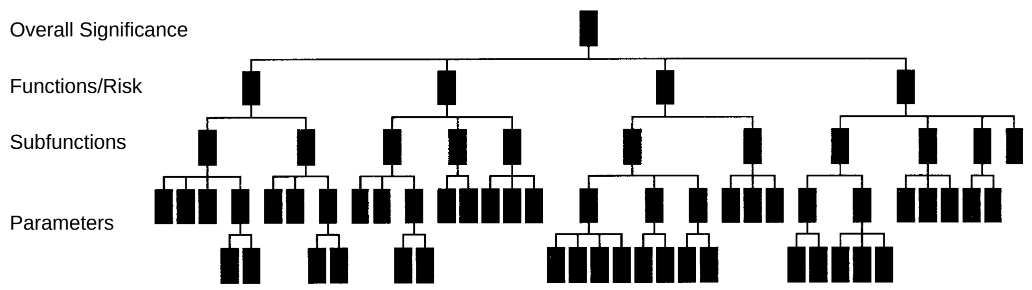


Figure 2. Overall hierarchical structure of the functional assessment procedure.

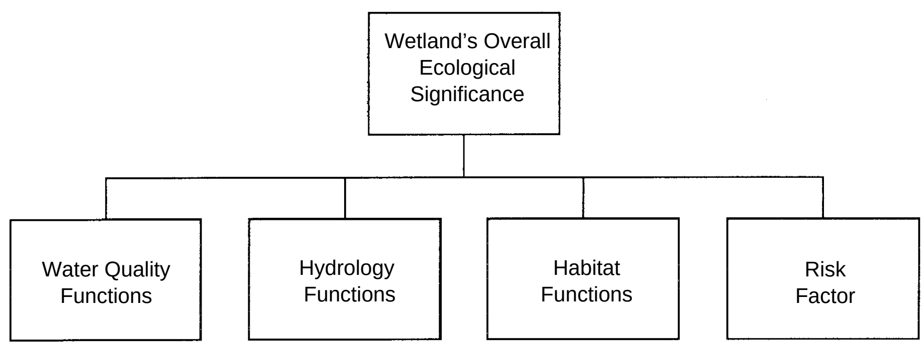


Figure 3. Assessment level two: Primary wetland functions and risk factor.

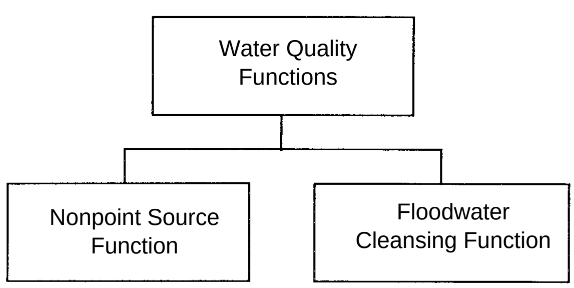


Figure 4. Water quality subfunctions.

make up the levels in the hierarchical structure that are actually evaluated based on fundamental ecological considerations. Parameter values, in turn, are combined to produce ratings for the subfunctions. Future reports will explain in detail all parameters evaluated in the assessment procedure and document them for scientific validity. This paper discusses only the parameters under the nonpoint source removal subfunction of the water quality function for illustration (see Figure 8).

The first parameter determining a wetland's significance in removing nonpoint source pollutants from surface runoff water is whether the water contains sediment, nutrients, or toxic pollutants in significant quantities. This is evaluated in the "proximity to sources" parameter based on the land uses surrounding the wetland. If agricultural fields or developed areas from which pollutants are likely to enter surface runoff largely surround the wetland, the wetland's potential for removing nonpoint source pollutants is high. If, on the other hand, natural vegetation from which runoff water is likely to be largely unpolluted mostly surrounds the wetland, its potential for removing significant pollutants is low.

Proximity to sources is an "opportunity" parameter. That is, it determines whether a wetland has the opportunity to remove pollutants from surface runoff by considering how likely the runoff water is to be polluted. The other parameters for this subfunction are "capacity" parameters that measure the wetland's ability to perform the function if the opportunity is present. Opportunity and capacity parameters are treated differently in determining a wetland's overall significance to prevent a wetland from being rated lower simply because present opportunity does not exist. This is discussed in more detail below.

The second parameter considered in determining a wetland's significance in nonpoint source removal is its proximity to a surface water body. If runoff entering a wetland would otherwise directly enter surface water, the wetland's significance as a filter is greater than if the wetland is far removed from surface water. In that case, pollutants in runoff could either settle out or be removed by other means before they enter surface water as pollutants.

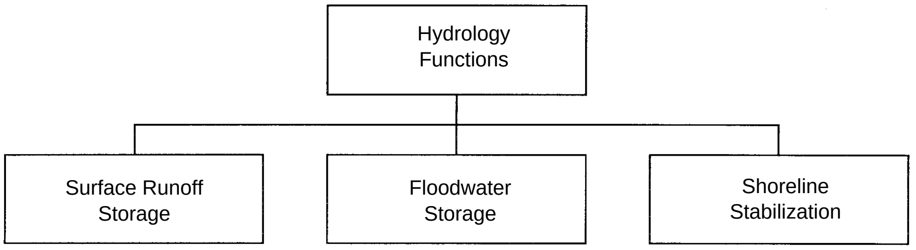


Figure 5. Hydrology subfunctions.

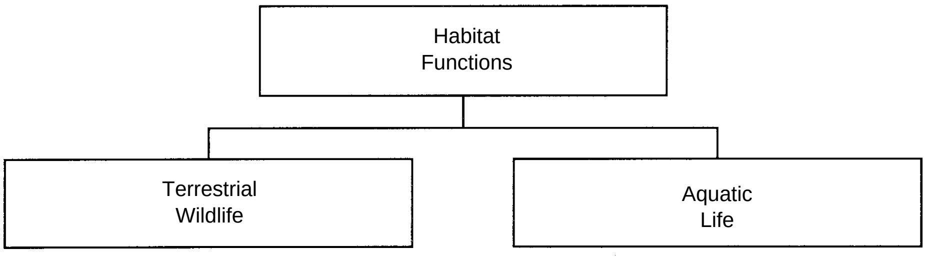


Figure 6. Habitat subfunctions.

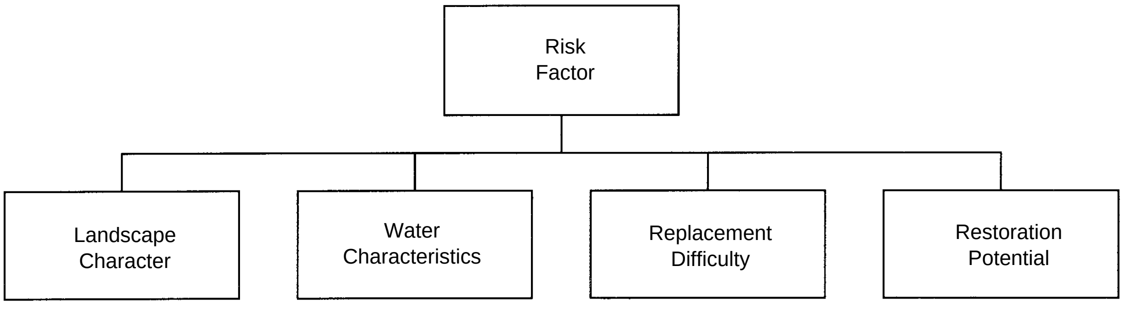


Figure 7. Risk factor subfunctions.

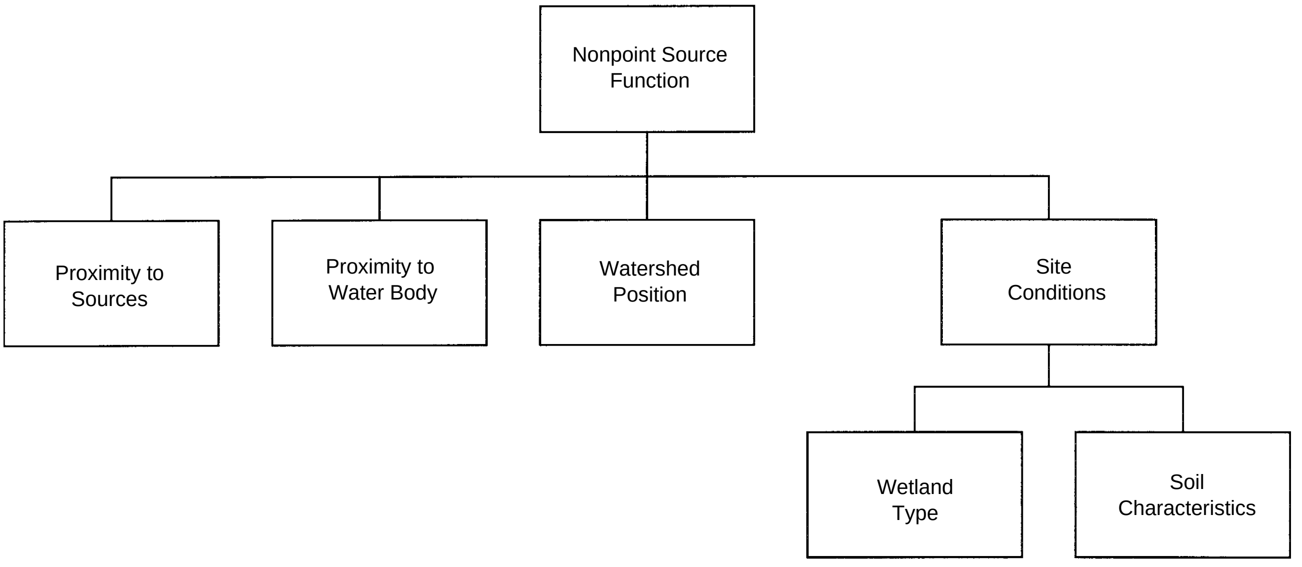


Figure 8. Parameters evaluated under nonpoint source pollutant removal subfunction.

The third parameter is the position of the wetland in its watershed. Several studies have documented that headwater wetlands are most effective in removing nonpoint source pollutants (9-11). Thus, the higher in its watershed a wetland is located, the higher is its significance in nonpoint source removal.

Two subparameters, wetland type and soil characteristics, determine the value of the fourth parameter, site conditions. By virtue of their typical microtopography, hydrology, and vegetative structure, some wetland types more effectively retain and filter surface runoff than do other types. Some soil series are more effective than others in retaining and chemically transforming pollutants. Each subparameter is rated, and their combined values produce a rating for the site conditions parameter.

A similar evaluation of specific parameters is performed to derive significance ratings for other wetland subfunctions. In all cases, GIS analysis determines parameter values based on the data layers described above. Some parameters, such as wetland type in the nonpoint source illustration, are surrogates or indicators of other wetland properties that actually determine the wetland's functional capacity. The limitations of GIS data and techniques necessitate the use of indicator parameters.

Evaluation Procedure

The objective of the assessment procedure is to determine an individual wetland's ecological significance in the watershed in which it exists. Ecological significance is divided into three broad classes (high, medium, and low) rather than attempting to derive a specific numerical "score." This is partly because of the procedure's initial application in an EPA ADID project performed by DCM in Carteret County, North Carolina. Standard ADID procedure is to classify wetlands into three groups:

- Areas generally unsuitable for the discharge of dredged or fill material.
- Areas that require a project-by-project determination.
- Possible future disposal sites for dredged or fill material.

These groups correspond to the H, M, and L used in the assessment procedure.

The approach of classifying wetlands into three broad functional significance classes is also used, however, because it is feasible with our current understanding of wetland function. Attempting to assign a specific value along a numeric continuum of functional significance greatly exaggerates the precision with which we can realistically apply current knowledge. The three significance classes used in the assessment procedure provide the information necessary to meet the procedure's objectives without going beyond the realm of reasonable scientific validity.

As explained above, the basic evaluation is performed at the parameter level. An H, M, or L value is assigned to each parameter as it relates to the performance of the wetland subfunction being considered. For example, if the soils underlying a wetland have properties that are highly conducive to the function being considered, the soil characteristics parameter is rated H; if soil properties are less conducive to performing the function, the parameter is rated M; and if soil properties are not at all conducive to the function, the parameter is rated L. All individual parameters under a given subfunction receive similar ratings.

The individual parameter ratings are then combined to give an H, M, or L rating for each subfunction. The subfunction ratings are combined into a rating of the wetland's significance in performing each of the primary wetland functions. Finally, the ratings for primary functions are combined into an overall rating of the wetland's functional significance.

The process of successively combining ratings up the structural hierarchy is the most complex aspect of the assessment procedure. The combining, as well as the evaluation of individual parameters, is based on fundamental ecological principles about how wetlands and landscapes function. Because the ecological processes themselves interact in complex ways, combining ratings is much more complex than a simple summation of individual ratings. Some parameters are normally more important than others in determining the level at which a wetland performs a specific function and, thus, must be weighed more heavily in determining the combined value. In some cases, different combinations of individual parameter ratings result in the same level of functional significance. Each possible combination of parameters must then be considered.

The automated version of the assessment procedure maintains all individual parameter ratings and combinations in a database. Because the combining process is complex, the reason a wetland receives an overall H, M, or L rating may not be intuitively obvious. The database makes it possible to trace through the parameter, subfunction, and primary function ratings that result in a wetland's overall rating.

This database also allows consideration of specific wetland functions individually. For example, in a watershed targeted for nonpoint source pollution reduction, one management objective may be to give the highest level of protection to wetlands most important in performing this function. The database allows examination of each wetland for its significance in nonpoint source removal and production of a map of wetlands rated according to their significance for this single function.

Individual function ratings in the database can also be used to improve planning, impact assessment, and mitigation for development projects that affect wetlands. If alternative sites are available, such as alternative corridors for a highway, the alternative with the least impact on the wetland function considered most important in the watershed can be identified. Rather than simply minimizing acres of wetland impact, the objective would be to minimize impacts to the most important wetland functions. Environmental assessment of wetland impacts can identify specific functions to be lost. Mitigation can be improved by giving priority to sites with the highest potential for performing the same functions.

Future reports will explain detailed procedures for evaluating individual parameters and combining them into functional ratings. This paper illustrates only the water quality nonpoint source removal subfunction. The rating system for this subfunction is summarized in Figure 9.

Four parameters are evaluated to determine the significance of the nonpoint source removal subfunction. Because (d), the site conditions parameter, has subparameters below it, it is first evaluated using a relatively simple procedure. If conditions typical of the wetland type and characteristics of the underlying soil are both highly conducive to removal of pollutants in runoff water entering the wetland, the site conditions parameter is H. If either the wetland type or the soil is not at all conducive to pollutant removal and the other subparameter is no more than somewhat conducive, the site conditions parameter is L. Any other combination results in an M.

Parameters

(a) Proximity to Sources

(b) Proximity to Surface Water

(c) Watershed Position

(d) Site Conditions

  (1) Wetland Type

  (2) Soil Characteristics

Evaluation Procedures

Site Conditions

H Both Parameters H

M Other combinations

L One parameter L and neither H

NPS Subfunction

H (a) & (b) H and (d) at least M *or*
(c) & (d) H and (b) at least M

M Other combinations

L Two of (b), (c), & (d) L

Figure 9. Parameters evaluated under nonpoint source pollutant removal subfunction.

Following evaluation of all parameters, they are combined to evaluate the significance of the wetland in removing nonpoint source pollutants. Two combinations result in the wetland being evaluated as highly significant in performing this function. First, if the wetland is adjacent to both a significant source of polluted runoff (a = H) and a permanent surface water body into which the runoff would flow if the wetland were not there (b = H), *and* has site conditions that are at least reasonably efficient in catching, holding, and removing pollutants from the runoff (d at least M), it receives an H. Alternatively, even if the wetland is not adjacent to a pollutant source, it receives an H if it is in the headwaters of the watershed (c = H), site conditions are highly conducive to pollutant removal (d = H), and it is at least close to an intermittent stream (b at least M).

On the other hand, if any two of parameters (b), (c), and (d) are evaluated L, the significance of the wetland for nonpoint source pollutant removal is **L**ow. That is, the wetland is evaluated as L for this function if any of the following conditions exist:

- The wetland is not close to surface water (b = L) and downstream in the watershed (c = L).
- The wetland is not close to surface water (b = L), and its site conditions are poor for pollutant removal (d = L).
- The wetland is downstream in the watershed (c = L) and has poor site conditions (d = L).

Any combination of parameter evaluations other than those resulting in an H or L results in the wetland being evaluated as of moderate significance for removing nonpoint source pollutants. This example is typical of evaluation procedures used for all subfunctions. More often than not, the evaluation procedures are complex and multifarious in their reasoning and application. Hopefully, though, they are scientifically valid based on current knowledge of wetland ecology.

Opportunity and Capacity

The concepts of opportunity and capacity for a wetland to perform a given function were briefly discussed above. For a wetland to actually perform a function, it must have both the opportunity and the capacity for the function. In terms of the nonpoint source example, a source of potentially polluted runoff must enter the wetland to provide an opportunity, and the wetland must have the internal capacity to hold the runoff and remove the pollutants before releasing the water. Factors external to the wetland usually determine the opportunity to perform a function, while properties of the wetland itself along with its landscape position determine the capacity to perform the function.

Because the assessment procedure is a landscape scale procedure that evaluates the functions a wetland performs in relation to its surroundings, essentially every subfunction includes opportunity parameters. A functional assessment that is too heavily dependent on opportunity parameters, however, is static and rapidly becomes invalid as land uses change. A wetland that is bordered by natural forest today can be bordered by a young pine plantation or a subdivision under construction by next year. The fact that a wetland does not have the opportunity to perform certain functions today does not mean that it will not have the opportunity in the future. If an assessment of wetland significance is to remain valid over time in a landscape subject to change, opportunity parameters alone cannot be determinative.

The evaluation procedure for the nonpoint source subfunction explained above is an example of how the assessment procedure handles this situation. The opportunity for a wetland to receive polluted runoff water from surrounding lands (a = H) *can* result in an evaluation of H for this subfunction if other properties are also present, but it does not *have* to be present for a wetland to be evaluated H. Other parameters (c and d = H, and b at least M) that give a wetland a high capacity to remove nonpoint source pollutants can also result in an H. Conversely, lack of present opportunity (a = L) does not result in an evaluation of low significance for this function. At least two of the other parameters must be L for the wetland to be evaluated as L.

These conventions hold throughout the procedure. A present high opportunity to perform a function can result in an evaluation of high significance for the function, but high capacity can also result in an H evaluation even if present opportunity is lacking. Lack of present opportunity alone never results in an evaluation of low significance for a function. High opportunity is treated essentially as a "bonus" consideration that can result in a higher evaluation for a wetland than its capacity alone would indicate but that will never result in a lower evaluation because of its absence.

Overriding Considerations

Several considerations are of such importance in the North Carolina coastal area that their presence alone will result in a wetland evaluation of high significance. These parameters are evaluated first as either true or false, and if one or more of them is true, the rest of the evaluation procedure is not performed.

The first overriding consideration is whether the wetland is a salt or brackish marsh meeting the definition of "coastal wetland" as set forth in North Carolina statutes (NCGS 113-229(n)(3)) and rule (NCAC 7H .0205(a)). Coastal wetlands in North Carolina are designated by law as highly significant. Consequently, the assessment procedure evaluates them automatically as H and includes no considerations for differentiating among the functional significance of these wetland types.

The second overriding consideration is whether the wetland is adjacent to an officially designated primary nursery area (PNA). All designated PNAs are included in "areas of environmental concern" in the NC CMP and are protected by a specific set of regulations. They are areas where initial postlarval development of finfish and crustaceans takes place and, thus, are critical to estuarine fish and shellfish populations. Wetlands adjacent to PNAs are highly important in maintaining water quality and appropriate salinity gradients in these critical areas and are automatically evaluated as of high functional significance.

The third overriding consideration is whether the wetland contains threatened or endangered species. If a known threatened or endangered plant or animal species on either federal or state lists is present, the wetland is evaluated as highly significant. The determination is based on information obtained from the North Carolina Natural Heritage Program.

The fourth overriding consideration is whether the wetland includes all or part of a critical natural area as designated by the North Carolina Natural Heritage Program. If so, the site is considered of high significance. GIS data layers maintained by the Natural Heritage Program also help make this determination.

Verification

Throughout the development and initial application of the assessment procedure, we have checked and verified its validity. Parameter evaluations and combination procedures are based on the best wetland science available in the scientific literature. The validity and accuracy of the GIS databases used to apply the procedure have been verified to the extent possible. Following sections of this report fully document any assumptions made about wetland ecology, GIS data, or GIS analytical techniques.

An advisory panel of wetland scientists familiar with the wetlands of coastal North Carolina and representatives of several state and federal wetland-related agencies reviewed every step of the procedure's development. While their review does not represent an endorsement of the procedure or its results by the agencies or individuals included, it does indicate the level of peer review the procedure has received.

During development of the procedure, field visits were made to nearly 400 wetland sites to gather data on functional indicators. On these same site visits, a field-based functional assessment procedure, the Wetland Rating System developed by the North Carolina Division of Environmental Management, was applied. This provides the basis for a field verification of the assessment procedure.

Discussion

As we continue to understand more about the role of wetlands in maintaining a healthy environment, the usefulness of wetlands locational data continues to grow in importance. Spatial data can assist county planners in guiding development away from environmentally sensitive areas. Landowners now have the capability to look at a map and realize very quickly that wetlands exist on a given area of land. In addition, economic development councils can use this information to plan development in areas attractive to a particular industry. If a new business or industry wishes to locate in an area positioned such that the wetlands permitting process could be avoided, maps showing lands void of wetlands could be a significant tool to the economic development council. The representation of these wetlands' ecological significance dramatically increases the utility for these data.

While paper maps can be distributed to all interested parties, digital data also are available to public agencies who have GIS capabilities. In Carteret County, for example, a publicly installed workstation will be made available with these data installed. The county government will be able to view wetlands in the context of cadastral boundaries that already are on GIS. Information about sensitive resources made available prior to any development will, hopefully, lead development away from environmentally sensitive areas.

References

1. Hefner, J.M., and J.D. Brown. 1985. Wetland trends in the southeastern United States. Wetlands 4:1-12.
2. Dahl, T.E. 1990. Wetlands losses in the United States 1780s to 1980s. U.S. Department of the Interior, Fish and Wildlife Service, Washington, DC.
3. DEM. 1991. Original extent, status, and trends of wetlands in North Carolina. Report No. 91-01. North Carolina Department of Environment, Health, and Natural Resources, Division of Environmental Management, Raleigh, NC.
4. DCM. 1992. Final assessment of the North Carolina coastal management program. Report to the Office of Ocean and Coastal Resource Management, NOAA, U.S. Department of Commerce, performed under the Coastal Zone Enhancement Grants Program (January 10). Section 309, CZMA, North Carolina Division of Coastal Management, Raleigh, NC.
5. DCM. 1992. Final strategy for achieving enhancements to the North Carolina coastal management program. Proposal to the Office of Ocean and Coastal Resource Management, NOAA, U.S. Department of Commerce, performed under the Coastal Zone Enhancement Grants Program (March 25). Section 309, CZMA, North Carolina Division of Coastal Management, Raleigh, NC.
6. Hefner, J.M., and K.K. Moorhead. 1991. Mapping pocosins and associated wetlands in North Carolina. Wetlands 2(Special Issue):377-389.
7. Brinson, M.M. 1993. A hydrogeomorphic classification for wetlands. U.S. Army Corps of Engineers Waterways Experiment Station Report WRP-DE-4. Washington, DC.
8. Khorram, S., H. Cheshire, K. Siderelis, and Z. Nagy. 1992. Mapping and GIS development of land use/land cover categories for the Albemarle-Pamlico drainage basin. Report No. 91-08-NC. Raleigh, NC: North Carolina Department of Environment, Health, and Natural Resources.
9. Leopold, L.B. 1974. Water: A primer. San Francisco, CA: W.H. Freeman and Co.
10. Whigham, D.F., C. Chitterling, and B. Palmer. 1988. Impacts of freshwater wetlands on water quality: A landscape perspective. Environ. Mgmt. 12(5):663-671.
11. Novitski, R.P. 1979. The hydrologic characteristics of Wisconsin wetlands and their influence on floods, streamflow, and sediment. In: Greeson, P.E., J.R. Clark, and J.E. Clark, eds. Wetland functions and values: The state of our understanding. Lake Buena Vista, FL: American Water Resources Association.

A GIS Strategy for Lake Management Issues

Michael F. Troge
University of Wisconsin, Stevens Point, Wisconsin

Abstract

Lake management plans are crucial to the sustained life of a lake as it experiences pressures from human as well as environmental activities. As proven in the past, geographic information systems (GIS) can meet the needs of most if not all environmental entities. Applying GIS to lakes and lake management, however, is a fairly new concept because most previous work focused on the terrestrial realm. Future studies must address problems relating to dimension, but adopting certain methods (i.e., cross-sectional coverages) can help lake managers plan for critical lake issues. By using sufficiently planned coverages, lake quality data management coverages can increase storage and/or analysis efficiency. After evaluating certain management criteria, a lake management plan can be derived and set up as a coverage. These criteria can then correspond collectively to form management zones within a lake. Each of these zones has its own set of management goals to which all lake users must strictly adhere.

Introduction

The importance of maintaining lake quality has long concerned recreationalists and ecologists. The multifaceted interrelationships of the lake environment, however, usually make proper assessment and analysis of lake quality information difficult. Over the last decade, assessment has become easier due to the increased use and acceptance of geographic information systems (GIS). This computer-based tool has allowed successful integration of water quality variables into a comprehensible format.

One area of the environmental sciences that has neglected GIS is lake management. This paper presents an alternative method for using a traditional two-dimensional GIS for viewing, querying, and displaying three-dimensional information—in this case, lake quality information and lake management criteria. Lakes, unlike geologic entities, offer a three-dimensional realm that humans can fully penetrate without a great amount of effort. Lakes also contain a complete aquatic environment of physical, chemical, and biological entities that humans can effectively observe and analyze. This paper does not discuss the issue of dimension; however, future studies, primarily those relating to the creation of three-dimensional GIS, should address this issue.

GIS allows incorporation of a multitude of environmental variables (e.g., water chemistry, geologic strata) into a synergism of the many coexisting variables of the lake environment. The ability of a GIS to "capture, manipulate, process, and display spatial or georeferenced data" is now well known and accepted (1). Surprisingly though, GIS is rarely used for lake management databases and associated water quality analysis.

The few examples that exist include Schoolmaster's Texas Water Development Board System (2), which examined water use on a county basis, and RAISON GIS (3), which is an expert computer system implementing proper application of hydrologic principles to a particular lake. Many other systems are simply database collection storehouses of lake information, such as the Galveston Bay National Estuary Program (4).

One notable example is the LAKEMAP program (5). This extensive and comprehensive GIS spans the entire United States covering approximately 800,000 lake sampling sites from the U.S. Environmental Protection Agency's (EPA's) STORET system. LAKEMAP uses both a database management and mapping display system, allowing retrieval of information for specific sites or aggregation of regional areas. This GIS is unique because it examined the creation of standards that could be used across the country in database development and the presentation of that data.

Using GIS in a lake management study inspires many questions because of the lack of existing research and the absence of any true standards. For example, how should one create a lake quality database for general purpose management? Is visualizing the integration of several variables within the lake ecology possible? Can one examine temporal changes in pH? Many technical and logistical GIS questions therefore existed when the Legend Lake study began.

Background of the Legend Lake Study

Legend Lake (see Figure 1) is located in Menominee Reservation, which is in Menominee County in northeastern Wisconsin. Legend Lake is a 1,230-acre impoundment comprising eight natural drainage lakes that a single stream once connected. In the late 1960s, a plan was introduced to convert this area into an impoundment/recreational area, and construction soon began. The ecology and hydrology had not been seriously evaluated since the development was finalized, hence the Legend Lake project was designed in cooperation with EPA, Wisconsin Department of Natural Resources (WDNR), Menominee Reservation and County, Legend Lake District, and the Legend Lake Property Owners Association. This intensive study spanned the qualitative and quantitative aspects of surface water, ground water, sediment, and aquatic plants, as well as human influences on the Legend Lake watershed and surrounding areas.

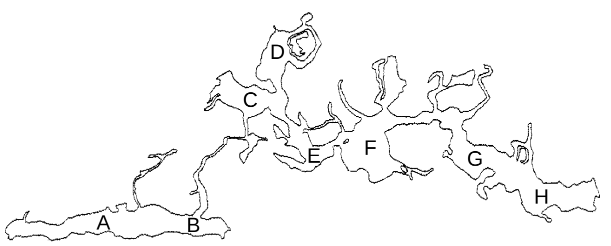


Figure 1. Legend Lake with basin identifiers.

Initially, questions needed to be answered regarding aquatic plants and sediment and their influences on lake management strategies. Because the scope of the Legend Lake study included subjects such as surface- and ground-water chemistry, land use and development, septic system impacts, and recreational stress, all these factors needed to be considered in determining optimum management strategies. In addition, the study addressed the question of GIS's ability to alleviate some technical aspects of deriving and presenting a lake management plan. Given these questions, the goal was to integrate collected data into a GIS database to create a prototype standard for future lake studies, as well as to present new techniques for visualization and analysis of lake quality data.

GIS/Database Creation

Several techniques were used to best manage the data for a three-dimensional system: cross-sectional coverages (described later), data summary coverages, and multidate coverages. The latter two coverage types are traditional coverages that contain general lake information excluding water column data. These techniques work fairly well for data storage and visualization; however, they were not sufficient for determining the geographic areas *within* the lake that required intensive management decisions as opposed to areas that needed little attention.

During formulation of a new coverage design, the study focused on the littoral zone, which is usually defined as that area of the lake with a depth of 15 feet or less. Most management concerns deal with the littoral zone because most recreational and ecological activities occur in this zone. One of the most pressing issues in lake management is aquatic plant growth. The fact that most aquatic plant growth is confined to the littoral zone reinforced the decision to use the littoral zone as the primary sink for potential management decisions.

To curtail the dimensional problem, depth was basically ignored. This allowed for easier delineation of areas within the lake. This, in turn, facilitated classifying areas into management zones to implement varying degrees of activity, ranging from casual to intensive efforts. Thus, management recommendations for a particular zone were the same at a depth of 2 feet as at a depth of 12 feet. This greatly reduced complication of the model and centered effort on the areal extent of the lake. It also facilitated visualization of management decisions by professionals and lay people.

Problems can arise when combining depths for management considerations, as this technique did. For example, a littoral zone that contains a gentle slope usually does not receive the same attention as a littoral zone with a very abrupt slope because the littoral zone with the gentle slope contains more area. Thus, if a lake manager recommends restricting boat traffic in a management zone (a section of the littoral zone) that has a gentle slope to promote wildlife habitat, the amount of area available for boaters would decrease significantly. In this situation, primary activity would be most crucial in shallow areas where wildlife or waterfowl predominate rather than in more open water areas where boaters predominate. Situations like these may require the creation of management subzones when setting up a lake management GIS and database to increase efficiency of lake area use and increase support by lake users.

Many criteria can affect the decisions made for a management zone. For instance, if the lake manager recognizes excessive, unhealthy weed growth in a particular zone, the lake manager may recommend extensive weed harvesting to neutralize the situation. An adjacent zone may have very little weed growth and may not require weed harvesting. Criteria such as these must be recognized before constructing the GIS. Table 1 lists some common criteria to consider. Generally, management criteria include anything that is influent on the

Table 1. Criteria To Consider When Creating Lake Management Plans

Environmental Variables	Artificial Variables
Aquatic plants	Adjacent land use
Sediment	Septic systems
Ground water	Development
Surface water	Construction sites
Wildlife/waterfowl	Fuel leaks
Fisheries	Shoreland zoning
Climate	Population density
Wind	Primary lake use (recreational)
Geology/geography	Nutrient loading
Adjacent natural land cover	Visitor use
Natural nutrient loading	
Hydrologic characteristics	

shoreline and lake itself and any influence the lake has on the shoreline or adjacent shorelands.

A primary concern when accessing lake data from a computer database is being sure to query the correct lake. For example, searching a database for all data on “Sand Lake” would be a legitimate action, except that the database may include close to 200 lakes with the name of Sand Lake. This is one of the main reasons why the WDNR developed a system known as the Water Mileage System (6). Based on logical criteria, each water body (e.g., lakes, streams, sloughs) receives a unique six- or seven-digit number called the waterbody number. Thus, if the number 197900, assigned to Sand Lake near Legend Lake, is the query subject, then the output should include all data for this particular Sand Lake. Because having a unique identifier for each specific entity in a GIS database is ideal, the waterbody number was used, and all sampling performed on this lake will be linked with this number.

Examples of Types of Coverage

Management Zone Coverages

Figure 2 and Table 2 together show how a potential lake management GIS and plan might work. The lake manager can easily manage and frequently update this system if necessary, or the system can serve as a long-term plan to consult for all decision-making. A plan of this sort specifically emphasizes areas that need intensive management over areas that may need frequent monitoring. It provides specific instructions for plan implementation, leaving little guess-work for the manager. This technique is also visually informative to the lake user because the user can easily discern areas of concern. This example is hypothetical, but a plan is being formulated based on the information collected during the Legend Lake study.

Cross-Sectional Coverages

Another technique currently included in the Legend Lake study entails the z dimension. Cross-sectional views of each lake basin, derived from 1992 lake contour maps, provided a more detailed description of the lake bottom. These cross sections were then digitized and transformed into GIS coverages. For each lake basin, 22 tests were conducted on the deepest part of the lake at several different depth intervals along the water column. These data provided valuable information on the way various chemical and biological attributes react to depth. Using GIS, a point could represent each depth where data were collected. These points could actually act as labels for polygons based on depth.

For example, if performing a series of analyses on a lake (maximum depth of 10 feet) at 3-foot, 5-foot, and 8-foot depth intervals, the labels on the cross-sectional coverage would be placed at these respective depths. Thus, labels would be positioned at depths of 3, 5, and 8 feet. Because these labels represent cross-sectional polygons, the 3-foot depth label may represent a polygon with boundaries at 0 feet and 4 feet. The 5-foot depth label may represent a polygon with boundaries at 4 feet and 6 feet, and the 8-foot depth label may represent a polygon with boundaries at 6 feet and the lake bottom (10 feet). Those who are familiar with lake ecology understand that no clear-cut boundaries distinguish where chemical values jump from one measurement to another without a gradual transition. All users of a lake quality GIS must be made aware of these types of inaccuracies (see Figure 3).

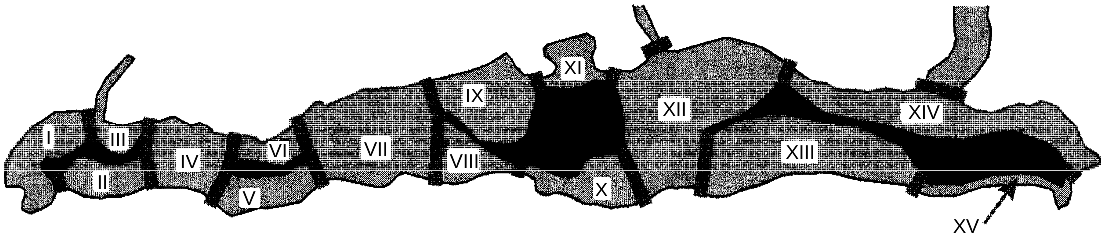


Figure 2. Hypothetical management zones for a section of Legend Lake that correspond to the management plans in Table 2; black areas indicate depths greater than 15 feet.

Table 2. Hypothetical Management Plans for a Section of Legend Lake (see Figure 2)

Management Zone	Management Plan (brief explanations)
I	High recreation area, high plant growth, frequent harvesting; frequently monitor water quality
II	Moderate recreation; manage for fish habitat
III	Moderate recreation; manage for fish habitat
IV	Open water, high recreation/possible fish habitat; consider subzoning
V	High-grade wildlife habitat; restrict human contact
VI	Moderate recreation; manage for fish habitat
VII	Open water, high recreation, increased shoreland development; frequently monitor water quality
VIII	Adjacent to high recreation area, possible fish habitat; manage for fish and aquatic habitat
IX	Adjacent to high recreation area, shoreland development; monitor water quality
X	Increased development; frequently monitor water quality
XI	[a]Prime wildlife/waterfowl habitat adjacent to high recreation area; restrict human presence (hot spot)
XII	Open water, high recreation area; frequently monitor water quality
XIII	Open water, high recreation area, possible fisheries and wildlife habitat; consider subzoning
XIV	Excessive aquatic plant growth (species listed) choking out preferred species; potential wildlife/waterfowl habitat, fisheries potential; continual harvesting; restrict human presence
XV	High-grade slope with little plant growth, potential for increased sedimentation; monitor shoreland development

[a]Critical area between good habitat and high recreation area; monitor extensively.

Figure 3 shows a cross section from one of the larger basins, Basin F, in the Legend Lake system (see Figure 1). The shades of gray represent ranges of temperature, with the lightest being the coldest and the darkest being the warmest. The thermocline can be located roughly in the middle. Each colored section represents a depth range where certain chemical attributes were collected. Using ARC/VIEW, the user can choose these areas with a pointer (mouse) and gain access to the database that contains all the sample results for this depth range.

Conclusion

These ideas are still preliminary as the Legend Lake study analysis concludes. Clear-cut discussions and recommendations will become available at a later date, although some observations can be made at this point. First, future research should focus on creating and developing a three-dimensional GIS, not to be confused with three-dimensional or cartographic models. Ideally, lake management plans should consider depth. This paper did not include depth because of the dimensional factor. Depth was not compatible with our two-dimensional GIS, and the presentation quality was not sufficient to relay our results in the form of management zones. We must develop a three-dimensional GIS to address the problems of the three-dimensional environment in which we live.

Lastly, maps are the best form for communicating this information to professionals and the public. Maps can also easily confuse people, however. Proper cartographic techniques are a necessity (7). Significant effort must be devoted to map creation to ensure a successful plan and successful relationships between lake managers and lake users. GIS and map-making are closely related. Both the planning stages and the database development phase of the lake quality GIS should emphasize this point. At an early stage of the process, management criteria should be determined, and all players or potential players must be included. A poorly planned project can lead to a failed GIS.

Creativity may offer new ideas in map development. For instance, animation (8) has some unique traits. Trend analysis using animation may produce the best visual results. Techniques such as these augment our methods of communication, and some are very revolutionary. Remember, however, that cartographic principles still must apply.

Acknowledgments

I would like to thank Dr. Keith Rice of the University of Wisconsin, Stevens Point, for his review efforts as well as Dr. Byron Shaw and Steven Weber for their technical support on lake management issues. Also, I would like

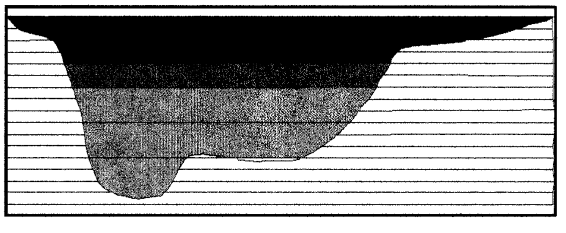


Figure 3. Cross section from one of the larger basins, Basin F, in the Legend Lake system.

to thank members of the University of Wisconsin, Stevens Point, Geography/Geology Department for their support and for the use of their equipment.

References

1. Fedra, K. 1993. Models, GIS, and expert systems: Integrated water resources models. Proceedings of the HydroGIS '93: Application of Geographic Information Systems in Hydrology and Water Resources Conference, Vienna. International Association of Hydrological Sciences Publication No. 211.
2. Schoolmaster, F.A., and P.G. Marr. 1992. Geographic information systems as a tool in water use data management. Water Resources Bull. 28(2):331-336.
3. Lam, D.C.L., and D.A. Swayne. 1993. An expert system approach of integrating hydrological database, models and GIS: Application of the RAISON system. Proceedings of the HydroGIS '93: Application of Geographic Information Systems in Hydrology and Water Resources Conference, Vienna. International Association of Hydrological Sciences Publication No. 211.
4. Rifai, H.S., C.J. Newell, and P.B. Bedient. 1993. GIS enhances water quality modeling. GIS World 6(8):52-55.
5. Samuels, W.B. 1993. LAKEMAP: A 2-D and 3-D mapping system for visualizing water quality data in lakes. Water Resources Bull. 29(6):917-922.
6. Fago, D. 1988. Retrieval and analysis system used in Wisconsin's statewide fish distribution survey, 2nd ed. Report 148. Wisconsin Department of Natural Resources Bureau of Research.
7. Ahner, A.L. 1993. Modern cartography plays essential role in GIS. GIS World 6(10):48-50.
8. Loucks, D.P. 1993. Interactive multimedia, GIS, and water resources simulation. Proceedings of the HydroGIS '93: Application of Geographic Information Systems in Hydrology and Water Resources Conference, Vienna. International Association of Hydrological Sciences Publication No. 211.

Sustainable Developments: Definition, Location, and Understanding

Michael E. Troyer, Ph.D.
USEPA, ORD, NRMRL

ABSTRACT

"Sustainable development" is an important concept, but not well defined in an operational context. In this study, sustainable development was ideally defined as a positive relationship between ecological integrity and human welfare within a specific area over time. Watersheds were chosen as the areal unit of interest because human welfare ultimately relies on the natural resources and life support systems provided by ecological rather than political systems. Furthermore, the physical boundaries of watersheds are unambiguous relative to many other ecological systems and regions. Although no scientific consensus on defining and measuring "ecological integrity" and "human welfare" exists, partial metrics for each were found in the literature. Data constraints limited the study to one time period (1988-1994). Therefore, "sustainable watersheds" were operationally defined as those with significant, above average levels of ecological <u>and</u> human conditions relative to other watersheds. In the analysis, three such watersheds were identified on the outskirts of a single metropolitan area and their locations assisted in developing partial explanations about how their "sustainable" conditions were achieved. Other watersheds with above average human conditions were also clustered around metropolitan areas, but had only average or below average ecological conditions. Watersheds with below average human conditions were all located in less accessible or intensively farmed areas of Ohio. These results suggested that higher levels of well-being in Ohio (i.e., in terms of educational attainment, employment, income, and lack of poverty), even in "sustainable watersheds," were tied to metropolitan areas with connections to national and global economies rather than isolated or "self-sufficient" economies based on local resources. In other words, modern sustainability appears to depend more on maintaining open and accessible economic systems rather than closed systems. Theoretically, economic relations between different places (e.g., cities, watersheds, countries) may range from mutualistic to competitive interactions, but mutualistic, along with cooperative and commensal relations, are preferred if sustainable developments are meant to be spatially non-exclusive.

Key words: Sustainable Development, Ecological Integrity, Human Welfare or Well-Being, Geographic Information System, Watersheds.

Introduction

Areas on earth are capable of sustaining various durations, frequencies, intensities, spatial scales and types of human activity. However, in many parts of the world, human activities continue to over-tax natural processes; degrade aquatic, terrestrial, and atmospheric resources; cause irretrievable losses of biological diversity; and increase uncertainty about the well-being of current and future human populations (Vitousek et al. 1986, Goodland 1991, Wilson 1988, Daly and Cobb 1989). In response, most member countries of the United Nations have pledged to develop in a more sustainable way (Goodland, et al. 1991), but have not specified what this means in an operational context.

Although perceived by some as a politically (or intentionally) vague concept, “sustainable development” does call attention to the critical question of how humans can continue to thrive on Earth in a more equitable and ecologically viable manner. Definitions of sustainable development commonly begin with the “Brundtland Report” written by the United Nation’s World Commission on Environment and Development (WCED 1987).[1] In 1987, the WCED defined sustainable development as development that meets the needs of the present without compromising the ability of future generations to meet their own needs. However, in the literature, sustainable development embodies many other concepts, objectives, and constraints such as those listed in Table 1.

The Research Problem

Systems theory suggests that political units (e.g., a particular city, state, or nation) may attempt to maintain or improve human welfare and environmental conditions within their jurisdictions by importing or exporting “sustainability,” or by making substitutions for depleted local resources. In other words, political units may increase their reliance upon external sources for energy, information, or raw materials; transport their wastes and pollution beyond their boundaries; and/or adopt new innovations in technology in order to sustain themselves over time. Thus, theoretically, political units at various hierarchical levels may behave or compete in ways which result in spatial variations of sustainable and unsustainable developments across the landscape. However, few, if any, attempts have been made to map and analyze such spatial

[1] The WCED was established in 1983 by the General Assembly of the United Nations for the purpose of developing long-term, environmental strategies for achieving sustainable development by the year 2000 and beyond.

variations, particularly at sub-national scales or within the context of ecologically-relevant systems.

Table 1. Examples of Concepts, Objectives, and Constraints Commonly Found in Studies of Sustainable Development

Concepts	Objectives	Constraints
Economic development (not "throughput" growth)	Survival of humans	Biogeophysical system
Equitable distribution of wealth within and between generations	Welfare of humans	Technology
Remaining supplies of natural resource stocks	Satisfying needs	Self-imposed laws, regulations, taxes, policies, and/or treaties
Environmental quality	Ecological vitality	
Carrying capacity	Biodiversity	

Modified from: Braat and Steetskamp 1991

In this light, this research attempts to measure, detect, and understand spatial variations of sustainable and unsustainable developments at a watershed scale. However, in doing so, it is recognized that sustainable development remains a complex and difficult concept to quantify, especially for researchers working within the theoretical confines of one discipline. Thus, an interdisciplinary approach is used. Several limitations prevent any single researcher from implementing such a definition. Most important, perhaps, is the lack of scientific consensus on defining and measuring broad concepts like "ecological integrity" and "human welfare" (e.g., consider Figures 1 and 2). However, the literature does suggest several partial metrics which can serve as a starting point. With the exception of remotely-sensed data, another limitation is the relative paucity of comparative ecological data at regional scales over time. As a result, this paper focuses on the relative difference between watersheds at one point in time. The questions addressed include the following:

- Question 1: Which watersheds had above average ecological <u>and</u> socio-economic (i.e., "sustainable") conditions in the early 1990s?
- Question 2: Does the spatial pattern or location of these watersheds suggest how their "sustainable" conditions were caused?

Sustainable Development as a
Relationship Between Ecological Integrity and Human Welfare

The field of Ecology offers several terms for characterizing the relationship between two organisms. For example, positive relationships may be mutual or cooperative, and negative relationships may be parasitic or competitive. It is suggested here that these same concepts can also be applied for defining a range of “sustainable” versus “unsustainable” conditions by characterizing observable relations between ecological integrity and human welfare within specific areas. In brief, relationships between co-existing or interacting species (or conditions of ecological integrity and human welfare within a particular place) may be characterized as shown in Table 2. In general, commensal, cooperative, and mutual relations can be thought of as successive stages of relatively benign to symbiotic interactions or, in a sense, increasingly "sustainable” relations. Neutral or non-interactions may exist between specific organisms, or between humans and specific ecosystems, but neither humans nor ecosystems exist alone. Finally, amensal, parasitic, and competitive interactions can be viewed as negative associations which may lead to the ill-health or demise of one or more organisms or ecosystems if internal or external resources cannot be found to sustain them.

It is important to note that the type of interaction occurring between organisms, or between humans and their environment may change over time as conditions change or as they develop through different stages. For example, Odum (1983, 395) suggests that “mutualism seems to replace parasitism as ecosystems evolve toward maturity, and it seems to be especially important when some aspect of the environment is limiting” (also see Bormann’s 1985 discussion on redundancy in mature forests). In such cases, Odum implies that mutualism has a strong selective advantage over other relations because it can lead to the emergence of more resilient properties within a system. Such properties are important, as Arrow, et al. (1995) note, because without ecosystem resilience: (1) discontinuous changes in ecosystem functions may result in a sudden loss of biological productivity and a reduced capacity to support human life, (2) irreversible changes may also occur to the set of options open for sustaining current and future generations, and (3) changes from familiar to unfamiliar states will tend to increase uncertainties associated with the environmental effects of economic activities.

Thus, if sustainable development is defined within a context of humans and other species interacting within a larger biotic and abiotic environment, then an ecological argument can be made for defining sustainable developments in watersheds, or other ecologically-relevant areas as instances where mutual, cooperative, or commensal relations between human welfare and ecological integrity can be detected over time (Figure 3). In areas with lower human welfare and higher levels of ecological integrity, nature may be managed as a preserve or be relatively unaffected by human activity for various reasons such as lying within less navigable or less economically productive terrain. Conversely, in areas where human welfare has increased or remained high, and ecological integrity has decreased or remained low, humans will have likely dominated the area in an amensal or predatory/parasitic manner. Competitive interactions, where society and nature are both inhibited, can also happen and may occur in areas where production has degraded or used-up all the local natural resources. From a systems point of view, these latter three types of interaction result in areas becoming more dependent upon external resources for sustaining high levels of welfare.

Table 2. Positive, Neutral, and Negative Interactions

<u>Positive Interactions</u>	
Mutualism	Both benefit and become totally dependent on each other.
Cooperation	Both benefit, but are not dependent upon each other.
Commensalism	One benefits and the other remains unaffected or no worse off.
<u>Non-interactions</u>	
Neutralism	Neither affects the other.
<u>Negative Interactions</u>	
Amensalism	One is inhibited, but the other remains unaffected or no worse off.
Predation/Parasitism	One benefits at the expense of the other.
Competition	Both inhibit each other directly or indirectly when a common resource is limited.

Source: Odum 1983.

Sustainable development, defined as positive interactions between society and nature, appears to coincide with thoughts for a new paradigm of environmental management which suggests the following definition for sustainability:

> Sustainability is a relationship between dynamic human economic systems and larger, dynamic, but normally slower-changing ecological systems, such that human life can continue indefinitely, human individuals can flourish, and human cultures can develop—but also a relationship in which the effects of human activities remain within bounds so as not to destroy the health and integrity of self-organizing systems that provide the environmental context for these activities (Norton 1992).

However, by formally adding concepts of interaction, it appears that sustainable development can be defined in a more testable or operational context as:

> ecologically-relevant areas (e.g., watersheds, ecoregions, etc.) where ecological integrity and human welfare both increase (i.e., exhibit mutual or cooperative relations) over time, or where either the ecological or human welfare condition remains constant (but not in a degraded state) while the other increases over time (i.e., a commensal relationship) (Figure 3).

In instances where temporal studies are not possible (as in this study), sustainable developments may also be defined in a relative sense as:

> ecologically-relevant areas with significant, above average levels of ecological integrity and human welfare with respect to other comparable areas at a specific point in time (Figure 4).

Conceptual Model and Data

The entire state of Ohio was used as a study area in order to obtain an adequate sample size of watersheds for statistical analysis. The state of Ohio was also chosen to limit any potential travel or data collection costs, and because aquatic biological data, not commonly found in other states, was available. Furthermore, Ohio contains a variety of ecological, economic, cultural, and physical features which combine to make a diverse geography (Peacefull 1996, Omernik 1987) and thus, a likely landscape for finding variations in both human and ecological conditions in watersheds across the state.

Watersheds became the ecologically-relevant areal units of interest in this study because: (1) no consensus on a standard system for classifying habitats, biological communities, nor ecosystems exists, and developing one would be difficult and contentious (Orians 1993); and (2) watersheds are aquatic systems with relatively undisputed physical boundaries. Most of Ohio's watersheds were assumed to be composed of a number of competing local economies, not "cooperatives" looking out for the best ecological and economic interests of the entire watershed. It was also assumed that State-level policies may impact or constrain certain activities at lower economic and ecological levels and so on. Figure 3 Sus. Dev. Over Time

For purposes of this study, Ohio was assumed to have many ecologically-relevant areas defined as watersheds and to have many state laws, regulations, and policies that potentially affect the ecological and socio-economic conditions of these watersheds. Any particular watershed, in turn, was thought capable of maintaining only a certain level of human and ecological activity at any given time. Watersheds were also assumed to be composed of open and complex systems including an interacting mixture of human settlements (each with its own local authorities governing land use, commerce, and trade), and a variety of terrestrial and aquatic ecosystems. Ecological components considered in this study, however, only focused on the physical conditions of rivers and forests (the latter with respect to forest-interior breeding birds), and the biological conditions of aquatic invertebrates and fish.

Watersheds having available data for this study (Figure 5) were delineated using Arc/Info (ESRI 1992-1998) and the following three coverages:

1) Point coverages derived from OhioEPA monitoring data measuring aquatic invertebrate, fish, and physical habitat conditions in Ohio rivers and streams during 1988-1992,

2) Line coverages of Ohio streams based on USGS 1:100,000-scale Digital Line Graphs (Wessex 1997), and

3) 1:250,000-scale Digital Elevation Models (DEM) of Ohio (USGS 1997).

All coverages used in this study were transformed into a consistent Albers Equal Area Conic projection prior to any overlays and analysis.

Census tracts (U.S. Census Bureau 1992, Wessex 1997) were used to aggregate socio-economic data at the watershed scale. This is because watersheds are typically composed of many census tracts, and these tracts represent the best resolution data available for comparatively assessing the social-economic statistics collected for this study.

Metrics of ecological and of human condition were selected based on the following criteria:

1. Can support for the measure be found in the scientific literature?

2. Does a data set based on this metric already exist, or can the metric be derived from existing data for the entire state of Ohio?

3. Are the data relevant to the time period of study?

4. Are the data at an acceptable spatial resolution (i.e., at a watershed or lesser spatial scale)?
5. Are values comparable across space?
6. Does the metric quantify a unique aspect of ecological or human condition?

Ecological variables designed to measure aquatic conditions and appearing to meet all of the above criteria included a number of Ohio EPA's aquatic monitoring data (i.e., the Index of Biological Integrity, the Modified Index of Well-Being, the Invertebrate Community Index, and the Qualitative Habitat Evaluation Index, see Ohio EPA 1988). In general, these indices focus on the biological and physical components of aquatic ecosystems and together characterize several interacting trophic levels. Data limitations included the fact that several sample years of Ohio EPA data (1988-1992) were necessary to get an adequate spatial coverage of values for this study. This was a minor concern because biological and physical data tend to reflect longer term aquatic conditions than chemical measures.

Terrestrial conditions within the watershed were measured in order to supplement the aquatic indices. Specifically, several landscape metrics of forest conditions relevant to forest-interior breeding birds were identified from peer-reviewed literature. These metrics included: percent forest cover, percent core forest, mean forest patch size, fractal dimension, and a measure of "connectedness" versus fragmentation called the "mean proximity index" (McGarigal and Marks 1995). In brief, these landscape metrics were derived from an existing remote sensing coverage of Ohio classified by Anderson Level 1 land use codes (Ohio DNR 1994).

Metrics of socio-economic condition were also chosen to represent different aspects of human well-being. Variables commonly used in the literature for measuring "human development," "sustainable economic welfare," or "quality of life" include: infant mortality, longevity, the level of educational attainment, the degree of economic diversity, real income or gross domestic product per capita, and income equity (Liu 1975; Daly and Cobb 1989; Cobb and Cobb 1994; Cobb, Halstead, and Rowe 1995; Miller 1996; and UNDP 1997). Data used in this study for deriving human welfare metrics at the watershed scale were obtained from the 1990 Census at the census tract level (as compiled by Wessex 1997). Infant mortality (a general indicator of health) which is available from “The Vital Statistics Annual Report” for 1990 at the county and city level (Ohio DOH 1997) was not included because county level data were too spatially coarse for aggregation to the watershed scale, and using it would constitute an ecological fallacy (i.e., “inferring characteristics of individuals from aggregate data referring to a population” (Johnston et al. 1986, 115)). One approximation considered though was to use the city data and point interpolation techniques as described by Lam (1983). However, discussions with Ohio’s Department of Health indicated that the data points were limited, predominately urban, not well distributed throughout the state, and thus, not amenable to such an effort (Dorothy Myers, pers. comm., 17 April 1998). For the spatial analysis here, final variables selected included: educational attainment (percentage of 18+ year-olds with a high school degree or higher), percent employment (where people live, not work), measures of economic diversity and evenness, the percentage of persons above poverty thresholds (all ages), median household and per capita income, and several measures of income equity.

Mean values of available aquatic samples were used to estimate watershed scale conditions. All socio-economic data at the census tract level were aggregated to the watershed scale using an area-weighted mean procedure. As a result, values of aquatic ecological and human condition were merely estimates, not actual measurements of the whole. Landscape measures of forest conditions were calculated using Arc/Info Grid (ESRI, 1992-1998) and Fragstats, a software program for spatial pattern analysis (McGarigal and Marks 1995).

Spatial Analysis of Watersheds with Varying Ecological and Human Conditions

In brief, each watershed was first ranked against all others with respect to both its aggregate ecological conditions and aggregate human conditions around 1990. Next, each watershed was plotted on a graph with coordinate axes and 99% confidence intervals representing average, aggregate ecological (X) and human (Y) conditions for all watersheds (Figure 6). Third, watersheds having above average ecological and human conditions were identified as those lying in the upper right hand (+ +) quadrant of Figure 6. These watersheds were then mapped and analyzed from a spatial perspective. Each of these steps are described in more detail below. Watersheds classed as either “+ -,” “- +,” or “- -” in Figure 6 were also identified and mapped, and are also discussed in this paper.

- Calculating Aggregate Ecological Conditions and

- Aggregating Human Conditions for Each Watershed

Aggregate scores for each watershed were derived by considering variables which were:

1) not significantly correlated with watershed area, and
2) minimally collinear (i.e., representing different and independent components of ecological or human watershed conditions).

All variables were converted to ordinal ranks with increasing values indicating better ecological (or human) conditions within the watershed, and lower ranks indicating poorer conditions. Thus, watersheds with higher summed ranks of ecological (or human) variables meeting the two criteria above were assumed to have higher overall ecological (or human) conditions, and vice versa.

Although not originally intended, principal components analysis (PCA) was explored to help select a few, independent "components" which explained most of the variance in the ecological and human welfare data sets. PCA and factor analysis were both considered as possible data reduction techniques for this purpose. However PCA was chosen because it simply focuses on data reduction and makes no assumptions about possible underlying causal structures responsible for covariation within a data set (Hatcher and Stepanski 1994, 503). It should be emphasized that PCA is a large sample procedure based typically on Pearson correlation coefficients derived from raw data measured on an interval or ratio scale. However, SAS does allow for the direct input of correlation matrices, including those derived from Spearman's non-parametric test. As such, PCA was used to help supplement the visual inspection of matrices in order to identify variables which empirically "hung together." PCA also helped to identify minimally collinear and important components, in terms of explained variance, to include in the aggregate rankings. In order to focus on solutions not affected by watershed area, correlation matrices derived from a sub-sample of watersheds with similar area (i.e., those with mean watershed areas $\pm$ 99% confidence interval) were used in the SAS routine.

After the principal components analyses, the next step was to use the resulting information to help rank and plot watersheds based on their relative, aggregate ecological and aggregate human conditions. Given the two criteria stated before, remaining variables for calculating aggregate scores included the following:

Aggregate Score for
Ecological Condition = Three Components = (MPS + MPI) + M_ICI + M_MIwb

Aggregate Score for
Human Conditions = One Component = (EA + EMP + P_AP + PCI + MEDHHI)

where:

MPS	=	Mean forest patch size (hectares) within the watershed.
MPI	=	Mean proximity index of forest patches (search radius = 2 kilometers).
M_ICI	=	Mean Invertebrate Community Index (aquatic invertebrates).
M_MIwb	=	Mean Modified Index of Well-Being (fish community).
EA	=	Educational attainment. Percentage of persons 18 years and older in the watershed with a high school degree or higher.
EMP	=	Percent employment in the watershed.
P_AP	=	Percentage of persons above poverty thresholds, all ages.
PCI	=	Per capita income in the watershed.
MEDHHI	=	Median household income in the watershed.

In other words, aggregate scores were summations of independent components consisting of variables not affected by watershed area and representing the majority of variance within each of the data sets. Specifically, the rankings and plots incorporate ecological data on terrestrial forest conditions and two aquatic trophic levels versus one human welfare component describing inter-correlated aspects of human capital. Because the principal components analyses were based on correlation matrices rather than raw, interval-ratio data, SAS could not derive optimally-weighted component scores. Therefore, factor-based scoring (a summation of ranks of variables loading meaningfully on a given component) was used in the aggregate score equations above.

Plotting, Classifying, and Mapping Watersheds Based Upon Their Relative Ecological and Human Conditions

Once aggregate scores were calculated for each watershed, all watersheds were then plotted and classified based upon which quadrant they occupied in Figure 6. Watersheds falling outside the 99% confidence intervals of the mean X and Y values were classified as having either:

1. above average ecological and human conditions (+ +),
2. below average ecological and human conditions (- -),
3. above average ecological conditions and below average human conditions (+ -), or
4. below average ecological conditions and above average human conditions (- +).

Maps were then prepared to show where watersheds within each of these classes were located and how they were spatially distributed in Ohio (see Figures 7 through 10).

Analysis

Results from the spatial analysis showed that watersheds with relatively lower poverty, and higher levels of education, employment, and incomes were generally located on the outskirts of metropolitan areas (e.g., Cincinnati, Lima, Columbus, Cleveland, and Akron in Figures 7 and 10). In contrast, watersheds with lesser welfare conditions were typically located in rural or more inaccessible areas apart from any Metropolitan Statistical Area (MSA) or significant laborshed (e.g., the Paint Creek and Appalachian watersheds in Figures 8 and 9, respectively). This led to the conclusion that higher levels of welfare in Ohio in the early 1990s were more dependent upon larger-scale economies with more extensive connections to the outside world, than small, self-sufficient economies based on local natural resources. In general, watersheds near metropolitan areas were well situated for residents to take advantage of employment opportunities and the goods and services provided by urban and suburban economies. They were also distant from any potential inner city problems. However, except for the three “sustainable” watersheds shown in Figure 7, most watersheds with relatively higher welfare conditions had only average or lower ecological conditions in response to varying intensities of suburban development and/or agriculture.

Watersheds with the relatively highest ecological conditions measured in this study were primarily found in less developed areas of Ohio. For example, the unglaciated and rugged part of Ohio– the Appalachian Plateau Province– where soil fertility is moderate-to-low, pastures and farmlands are generally limited to hilltops or lower-slope areas, and forest is the predominate land cover (Figure 9). Relatively lower levels of human condition were measured in this region and corroborated by supplemental literature reporting on limited economic opportunities due, in part, to declines in demand for Appalachian reserves of high-sulphur coal (Stephens 1996).

Comparatively better ecological conditions were also observed in three “sustainable” watersheds (as defined in this study) located within the Chagrin River watershed system in Northeast Ohio (Figure 7). In contrast to the Appalachian watersheds, human conditions in these watersheds were relatively high and likely supported by their proximity to Cleveland. Historically, more focused developments in Cleveland rather than in these outlying watersheds may be why above average ecological conditions remained there in the early 1990s. Another interesting feature was their location within a special survey area known as the Connecticut Western Reserve. In brief, this part of Ohio was originally surveyed into slightly smaller townships (i.e., 5 by 5 miles rather than 6 by 6 miles found elsewhere), but not subdivided into sections (Wilhelm and Noble 1996). Also of interest was the emergence of non-government organizations (NGOs) committed to protecting natural resources within the watershed, and informing citizens about local land use decisions and economic and environmental policies affecting the watershed. Similar organizations were also found elsewhere in Ohio where valued natural resources were experiencing increased anthropogenic stress (e.g., suburban sprawl in the Big Darby Watershed west of Columbus and in the Little Miami River Watershed east and northeast of Cincinnati).

Conclusions

"Sustainable development" is not directly observable. It needs to be defined before it can be visualized and studied. The unique approach in this study was to explicitly define sustainable development as a positive relationship between "ecological integrity" and "human welfare." In other words, the thrust was to explore whether mutualistic or cooperative relationships existed at a watershed scale; to locate watersheds with such relations (if any); and to suggest how these relations may have been caused.

The methodology was straightforward and made use of only a few select variables which, although limited in number and scope with respect to Figures 1 and 2, went a long way towards answering questions like: "Can society <u>and</u> nature thrive together? If so, where and how?" Overall, the methodology was successful in identifying watersheds with variable ecological <u>and</u> human conditions, and knowing their locations assisted in developing hypotheses about how these conditions were achieved. In this paper, the method identified and located a relative range of "sustainable" to "unsustainable" watersheds for one time period. Watersheds observed with above average human <u>and</u> ecological conditions (i.e., "+ +" conditions) were significant because they differed from the general inverse relationship between society and nature observed in Figure 6. Their identification in Figures 6 and 7 demonstrated that apparent harmonies of society and nature may still be found and studied with respect to their environmental endowments and histories of human occupation.

Few (if any) attempts have been made to map and analyze spatial variations of sustainable development at sub-national scales or within the context of ecologically-relevant systems. Therefore, this research differed from others focusing on political units and much larger spatial scales. Reasons for looking at ecological rather than political areas stemmed from the assumption that human welfare ultimately depends upon the natural resources and life support systems provided by healthy ecosystems (WCED 1987, U.S. EPA 1990). It also stemmed from the fact that aquatic and terrestrial ecosystems are often degraded or destroyed by the cumulative, seemingly "independent" decisions made by cities, towns, and villages within a particular watershed (e.g., see Ohio EPA 1996).

State-of-the-art geographic information systems (Arc/Info and Grid), statistical and spreadsheet software (SAS and Excel), and landscape metric programs (Fragstats) made this study possible, but most of the work, in terms of time and effort, still centered on data aggregation and input. Specifically, this involved aggregating point, polygon, and raster data to the watershed scale, and converting data from their existing form into others which could be transferred among these different tools of "automated geography" (see Dobson 1993 and Aronoff 1993, 42).

The analysis in this paper could be improved by considering changes over time (e.g., recall Figure 3). This may be done after the Year 2000 census and would assist in developing a richer understanding of human and ecosystem responses to stress over time and space. Others may find the method useful as a "broad-brush" approach for highlighting specific watersheds for more detailed studies (e.g, studies like those conducted by Braat and Steetskamp 1991, and Rees 1995). Similarly, the approach could also assist in prioritizing watersheds in need of conservation, remediation, or other environmental management efforts. Where sufficient data and interdisciplinary expertise exists, the method could also be improved by: including better or more comprehensive sets of ecological and human welfare measures (e.g., additional measures of chemical pollutants or human health), considering other types of ecological systems or boundaries (e.g., ecoregions), or by developing predictive models of ecological integrity and human welfare and plotting and mapping their outputs similar to that done in Figures 6 through 10.

Future researchers may have better and more continuous coverages of data than used here for observing spatial patterns of sustainable development across a landscape. Flows of people, money, information, and/or natural resources may also become more apparent, as well as, links between areas of sustainable and unsustainable developments. Furthermore, researchers may develop or have better tools for modeling or analyzing such flows at multiple scales (e.g., multiple political, economic, and ecological scales). However, humility in quantification will always need to remain, particularly when attempts are made to address the many important, qualitative aspects of ecological integrity and human welfare (e.g., recall Figures 1 and 2).

Bibliography

Aronoff, S. 1993. *Geographic information systems: A management perspective*. Ottawa, Ontario: WDL Publications.

Arrow, K., B. Bolin, R. Costanza, P. Dasgupta, C. Folke, C. S. Holling, Bengt-Owe Jansson, S. Levin, Karl-Göran Mäler, C. Perrings, D. Pimentel. 1995. Economic growth, carrying capacity, and the environment. *Science* 268: 520-521.

Bormann, F. H. 1985. Air pollution and forests: An ecosystem perspective. *Bioscience* 35(7):434-441.

Braat, L. C. and I. Steetskamp. 1991. Ecological-economic analysis for regional sustainable development. In *Ecological Economics: The science and management of sustainability*, ed. R. Costanza, 269-288. New York: Columbia University Press.

Cobb, C. W. and J. B. Cobb, Jr. 1994. *The green national product: A proposed index of sustainable economic welfare*. New York: University Press of America.

Cobb, C. W., T. Halstead, and J. Rowe. 1995. *The Genuine Progress Indicator: Summary of data and methodology*. San Francisco, CA: Redefining Progress.

Daly, H. and J. B. Cobb, Jr. 1989. *For the common good: Redirecting the economy toward community, the environment, and a sustainable future*. Boston: Beacon Press.

Daly, H. 1991. From empty-world economics to full-world economics: Recognizing an historical turning point in economic development. In *Environmentally Sustainable Economic Development: Building on Brundtland*, ed. R. Goodland, H.Daly, S. El Serafy, and B. von Droste, 29-38. Paris: UNESCO.

Dobson, J. E. 1993. The Geographic Revolution: A Retrospective on the Age of Automated Geography. *Professional Geographer* 45(4): 431-439.

Environmental Sytems Research Institute, Inc. (ESRI). 1992-1998. Arc/Info, Version 7.2.1. Redlands, CA.

Goodland, R. 1991. The case that the world has reached limits: More precisely that current throughput growth in the global economy cannot be sustained. In *Environmentally Sustainable Economic Development: Building on Brundtland*, ed. R. Goodland, H. Daly, S. El Serafy, and B. von Droste, 15-26. Paris: UNESCO.

Goodland, R., H. Daly, S. El Serafy, and B. von Droste, ed. 1991. *Environmentally Sustainable Economic Development: Building on Brundtland*. Paris: UNESCO.

Hatcher, L. and E. J. Stepanski. 1994. *A step-by-step approach to using the SAS systems for univariate and multivariate statistics*. Cary, NC: SAS Institute, Inc.

Johnston, R. J., D. Gregory, and D. M. Smith, ed. 1986. *The dictionary of human geography*. Second edition. Oxford, UK: Basil Blackwell Ltd.

Lam, N. 1983. Spatial Interpolation Methods: A Review. *The American Cartographer* 10(2):129-149.

Liu, Ben-Chieh. 1975. *Quality of life indicators in U.S. metropolitan areas, 1970*. Washington, DC: U.S. Environmental Protection Agency, Washington Environmental Research Center.

McGarigal, K. and B. J. Marks. 1995. *FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure*. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. Gen. Tech. Rep. PNW-GTR-351.

Miller, G. T. 1996. *Living in the environment*. Ninth Edition. Belmont, CA: Wadsworth Publishing Company.

Norton, B. G. 1992. A new paradigm for environmental management. In *Ecosystem health: New goals for environmental management*, ed. R. Costanza, B. G. Norton, and B. D. Haskell. Washington, DC: Island Press.

Odum, E. O. 1983. *Basic ecology*. New York: Saunders College Publishing.

Ohio Department of Health (Ohio DOH). 1997. *Catalog of electronic products and publications* [available on-line]; Internet; http://www.state.oh.us/healthe/ cat9719.pdf; accessed 23 February 1998.

Ohio Department of Natural Resources (Ohio DNR). 1994. *State of Ohio land cover*. Division of Real Estate and Land Management. Columbus, OH.

Ohio Environmental Protection Agency (Ohio EPA). 1988. *Biological criteria for the protection of aquatic life: Volume II: Users manual for biological field assessment of Ohio surface waters*. October 30, 1987 (Updated January 1, 1988). Columbus, OH: Ohio Environmental Protection Agency, Division of Water Quality Monitoring and Assessment, Surface Water Section.

________. 1996. *Ohio water resource inventory. Executive summary: Summary, recommendations, and conclusions*. Columbus, OH: Ohio Environmental Protection Agency, Division of Surface Water, Monitoring and Assessment Section.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77(1): 118-125.

Orians, G. H. 1993. Endangered at what level? *Ecological Applications* 3(2): 206-208.

Peacefull, L., ed. 1996. *A geography of Ohio*. Kent, OH: Kent State University Press.

Rees, W.E. 1995. Revising Carrying Capacity: Area-Based Indicators of Sustainability. *Population and Environment: A Journal of Interdisciplinary Studies* 17(3)[article on-line]; available from http://dieoff.org/page110.htm; Internet; accessed 26 January 1998.

Stephens, D. T. 1996. Population Patterns. In *A Geography of Ohio*, ed. L. Peacefull, 127-145. Kent, OH: Kent State University Press.

United Nations Development Programme (UNDP). 1997. *Human development report*. Cary, NC: Oxford University Press [Summary available on-line]; Internet; http://www.undp.org/ undp/hdro/97.htm; accessed 19 March 1998.

United States Bureau of the Census (U.S. Census Bureau). 1992. Census of population and housing, 1990: Summary Tape File 3 Technical Documentation, Appendix B on CD-ROM [machine-readable data files]. Washington, DC: U.S. Bureau of the Census [Available on-line]; Internet; http://www.census.gov/td/stf3/ append_b.html; accessed 21 September 1998.

United States Environmental Protection Agency (U.S. EPA). 1990. *Reducing risk: Setting priorities and strategies for environmental protection*. Washington, DC: U.S. Environmental Protection Agency, Science Advisory Board, EPA SAB-EC-90-021.

United States Geological Survey (USGS). 1997. 1:250,000-scale digital elevation models for Ohio. Earth Sciences Information Center, Reston, VA, [Available on-line] http://www.usgs.gov.

Vitousek, P. M., P. R. Ehrlich, A. H. Ehrlich, and P. A. Matson. 1986. Human appropriation of the products of photosynthesis. *Bioscience* 34(6): 368-373.

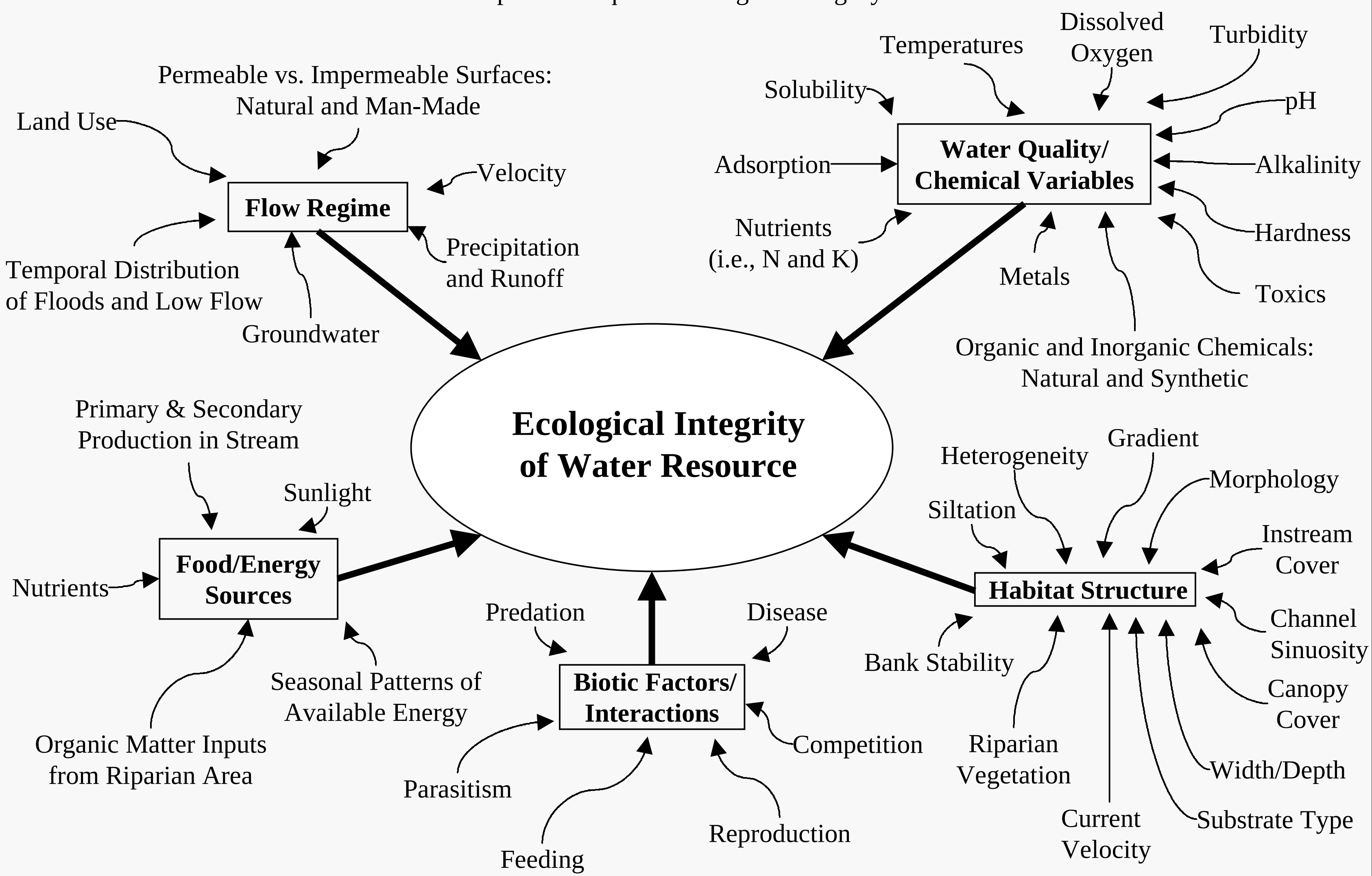
Wessex, Inc. 1997. *First St*. [CD-ROM]. Winnetka, IL.

Wilhelm, H. G. H. and A. G. Nobel. 1996. Ohio's settlement landscape. In Peaceful, L., ed. *A geography of Ohio*. Kent, OH: Kent State University Press.

Wilson, E. O., ed. 1988. *Biodiversity*. Washington, DC: National Academy Press.

World Commission on Environment and Development (WCED). 1987. *Our common future*. Oxford: Oxford University Press.

Figure 1
Aspects of Aquatic Ecological Integrity



Modified from: Karr et al. 1986, Karr 1991, and Yoder 1995

Figure 2
Aspects of Human Welfare

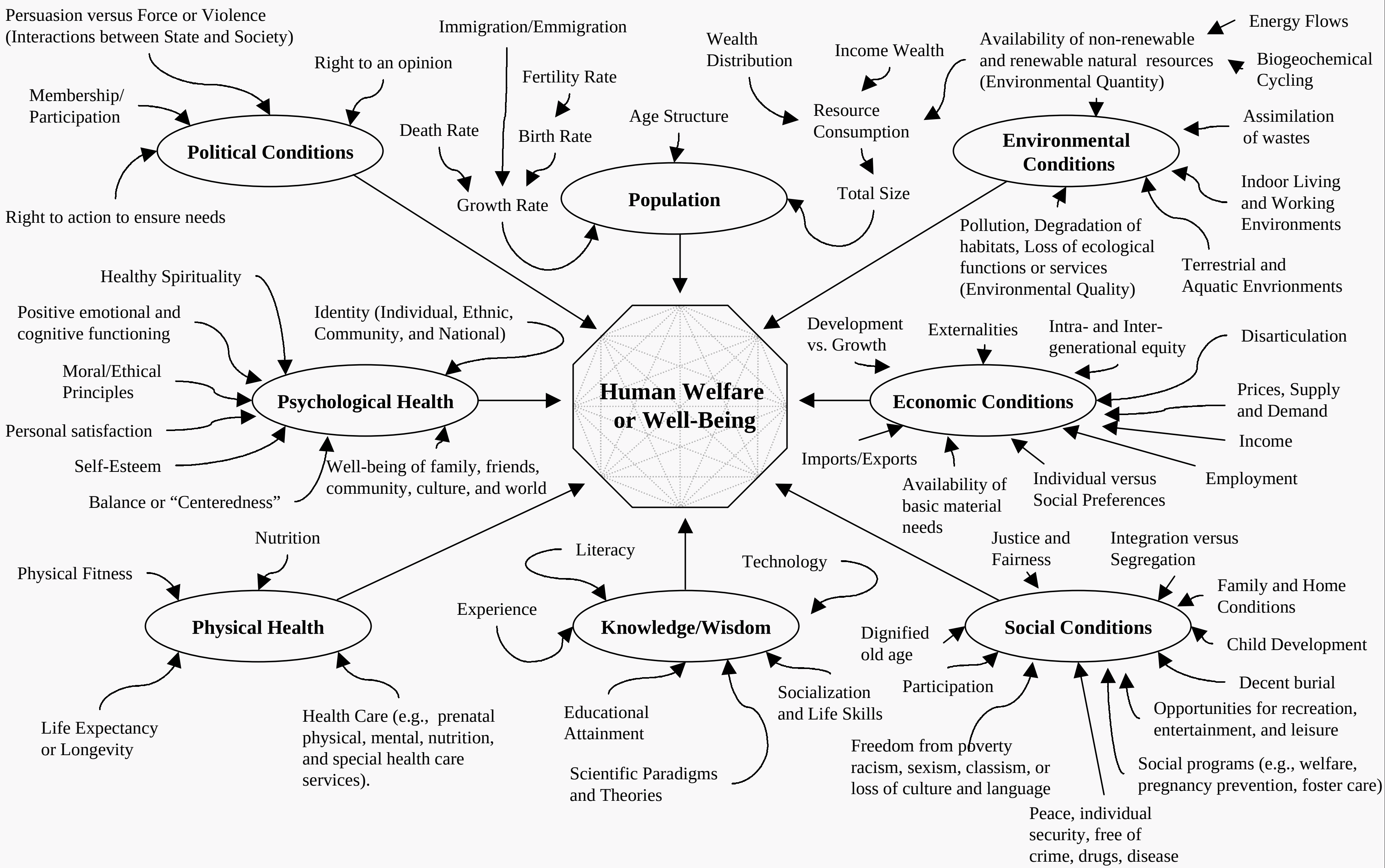


Figure 3
Hypothetical Sustainable Developments Over Time

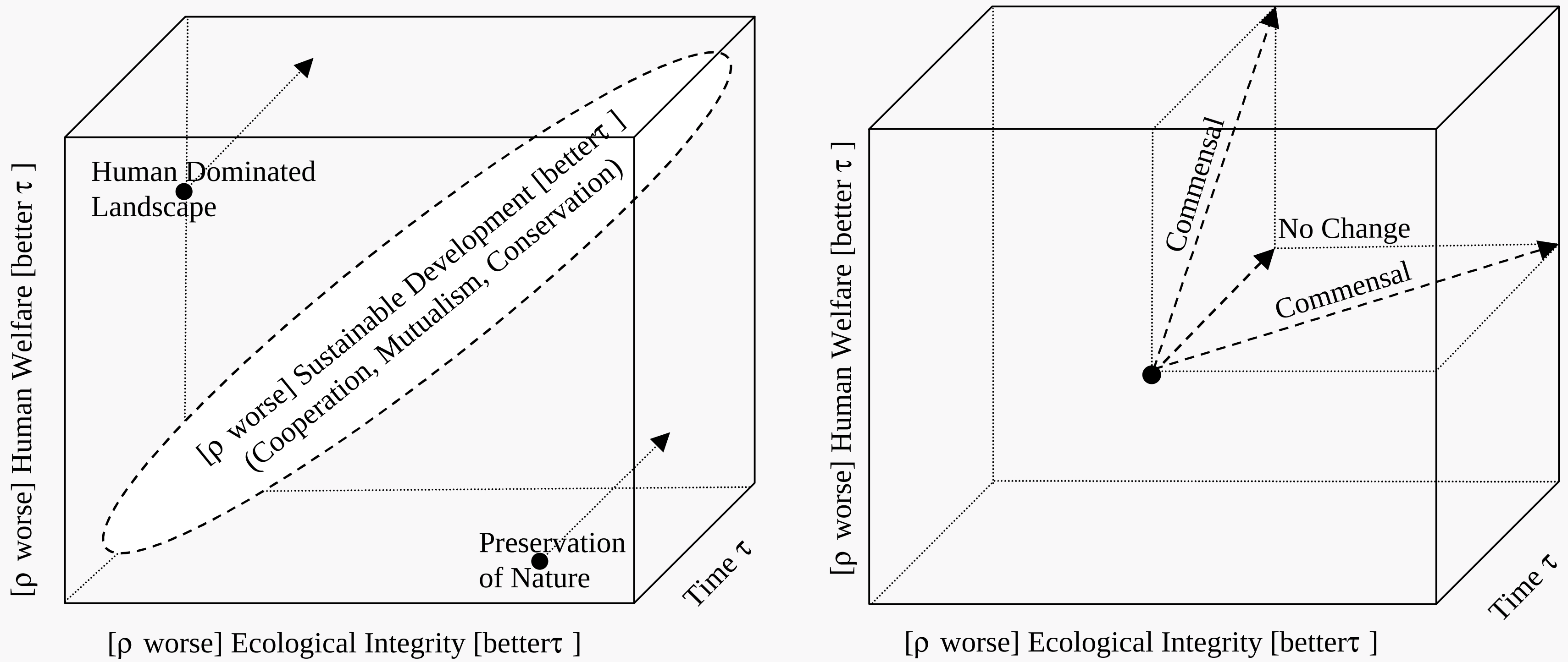


Figure 4
Hypothetical, Relative Sustainable Development At One Point In Time

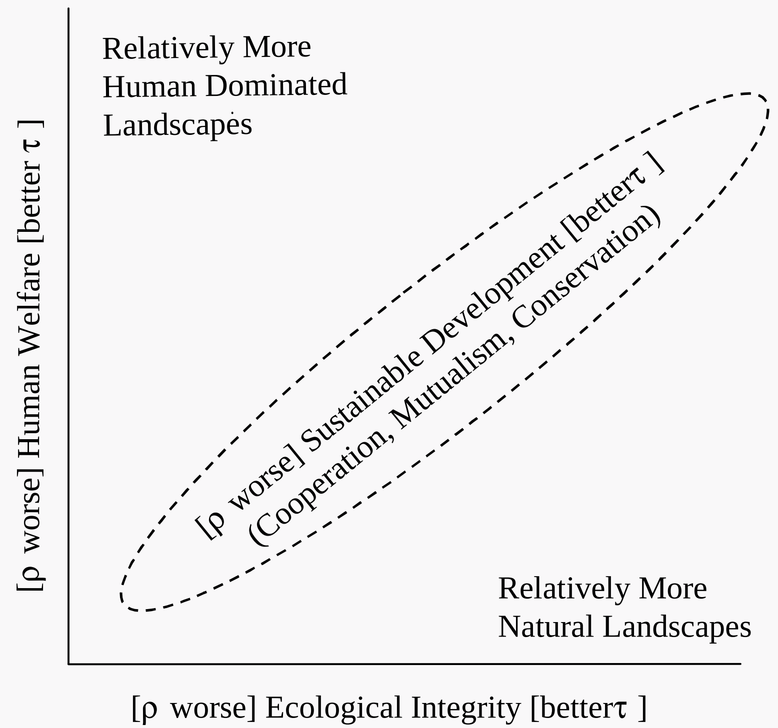


Figure 5
Watersheds Analyzed in this Study
(N = 57)

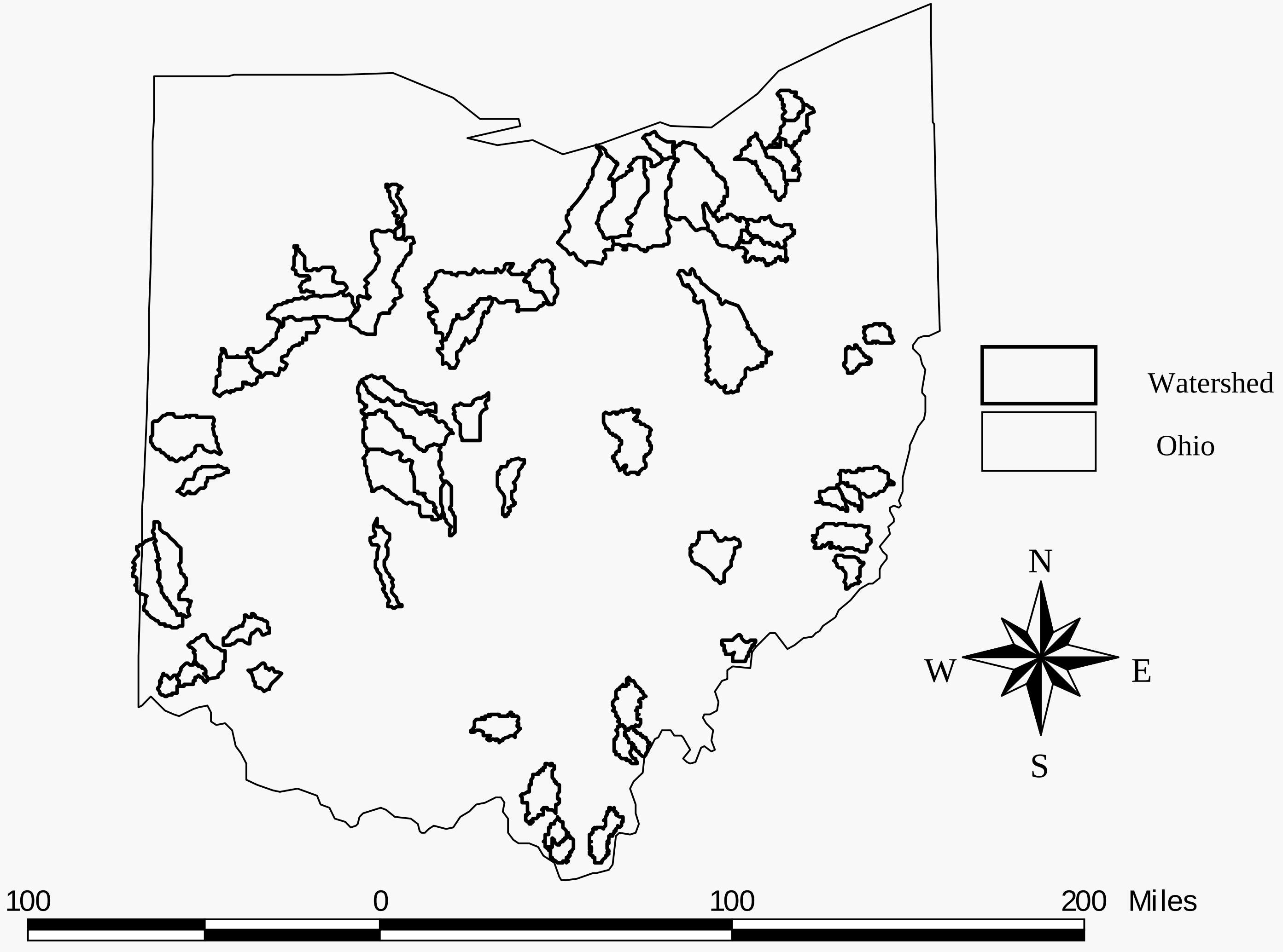


Figure 6
Plot of Watersheds Based on Their Aggregate
Ecological and Aggregate Human Conditions
(N = 57)

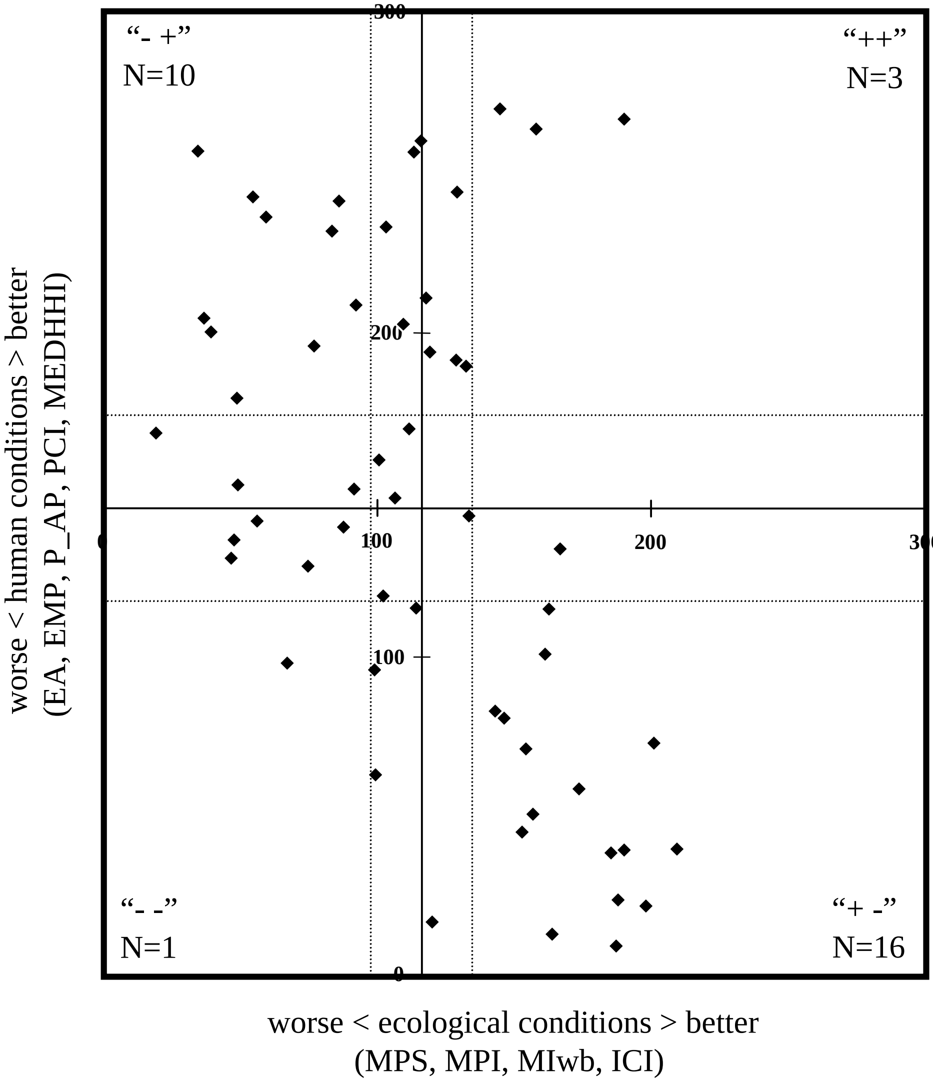


Figure 7
Map of “Sustainable” Watersheds Relative to Others Studied:
Watersheds With Above Average Ecological and Human Conditions, Circa 1990

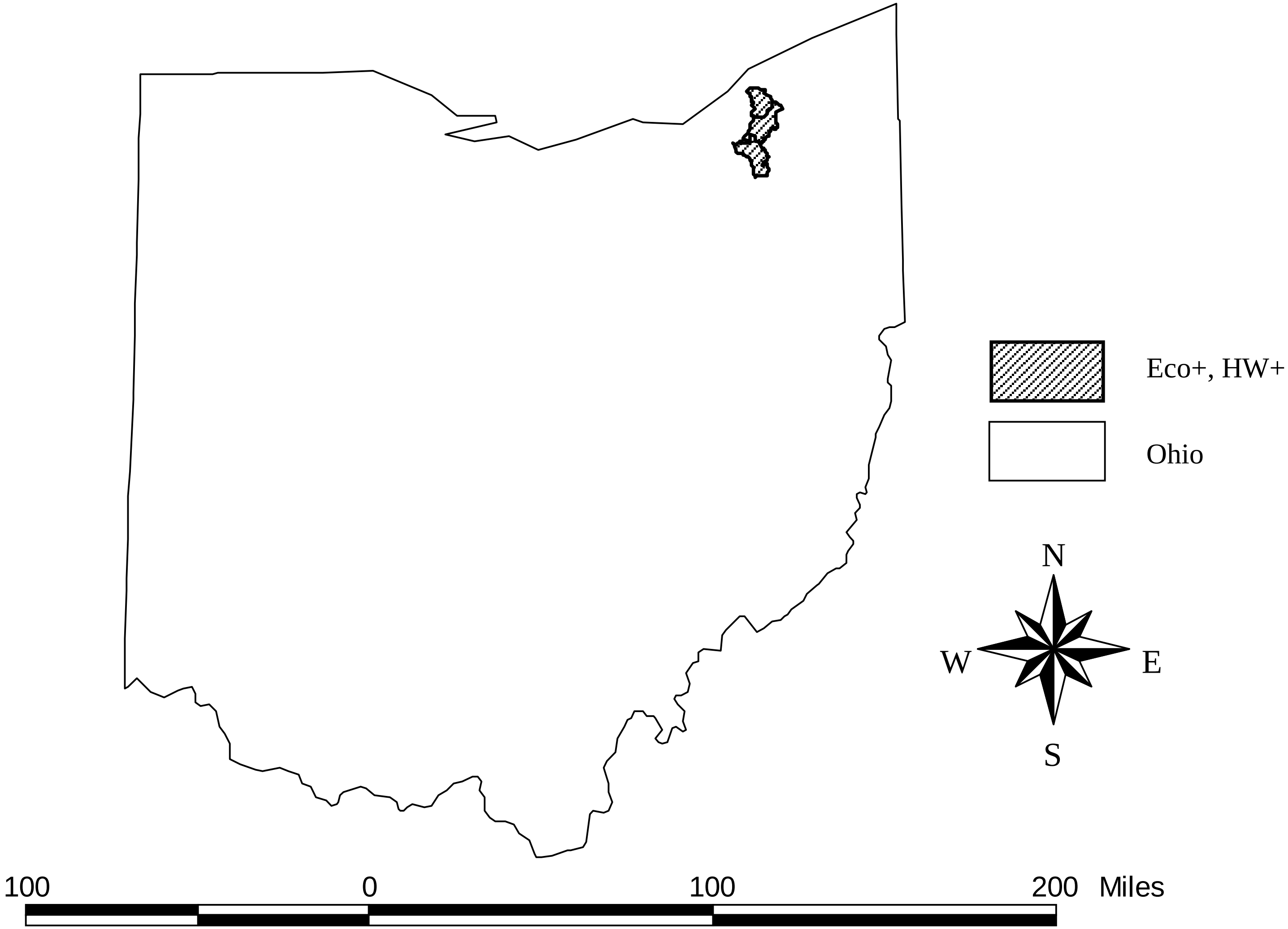


Figure 8
Map of Watersheds Studied With Below Average Ecological and Human Conditions, Circa 1990

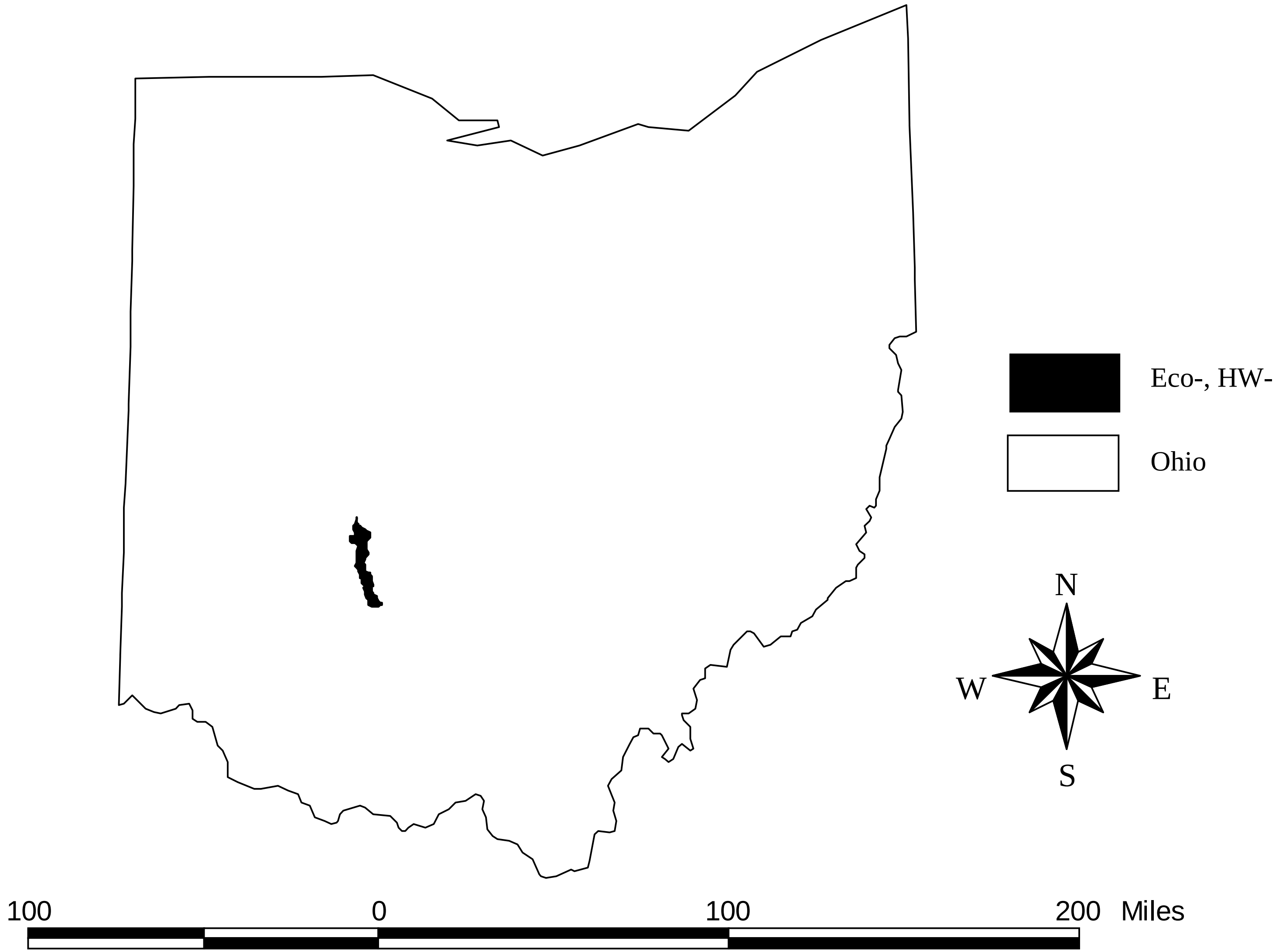


Figure 9
Map of Watersheds Studied With Above Average Ecological Conditions and Below Average Human Conditions, Circa 1990

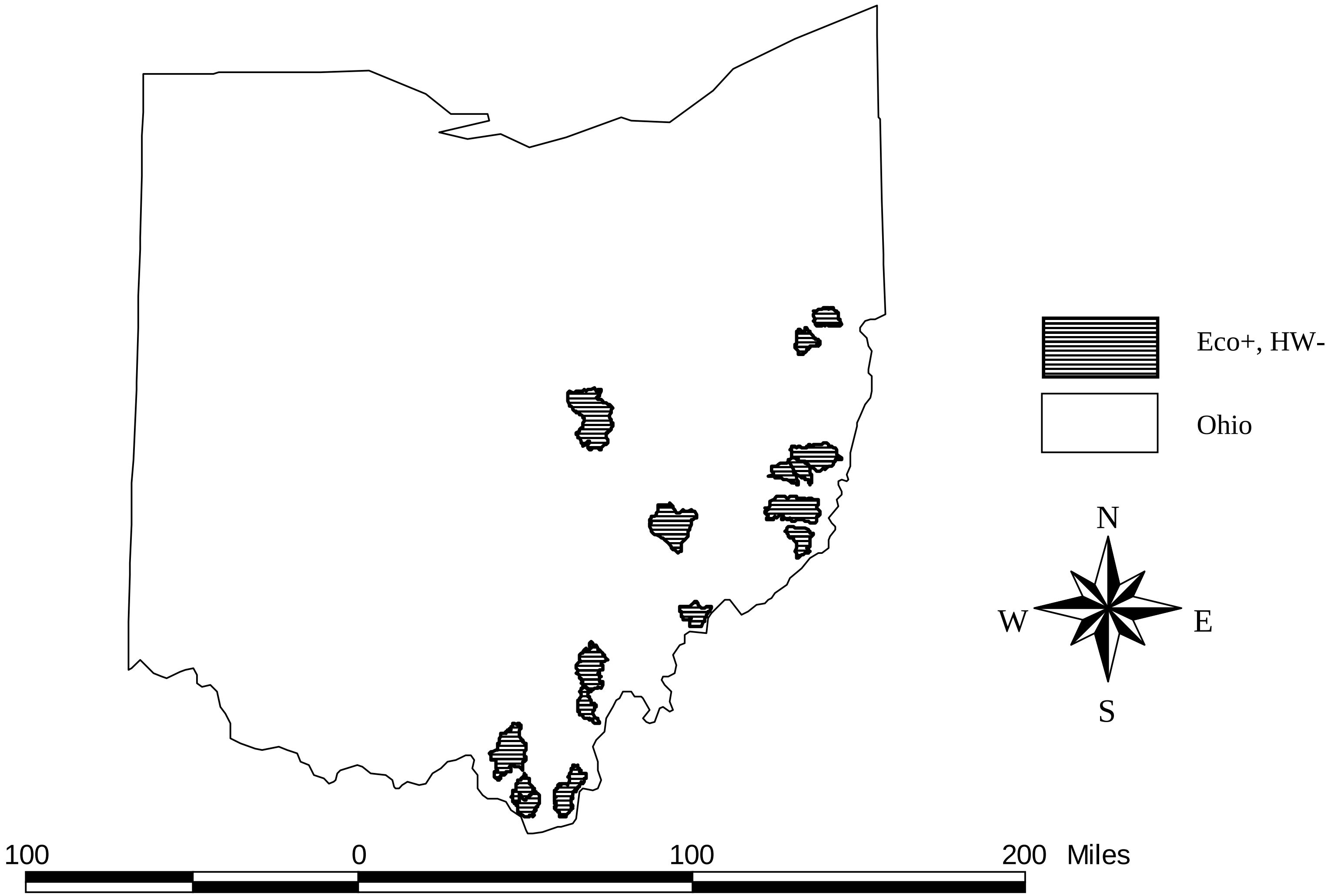
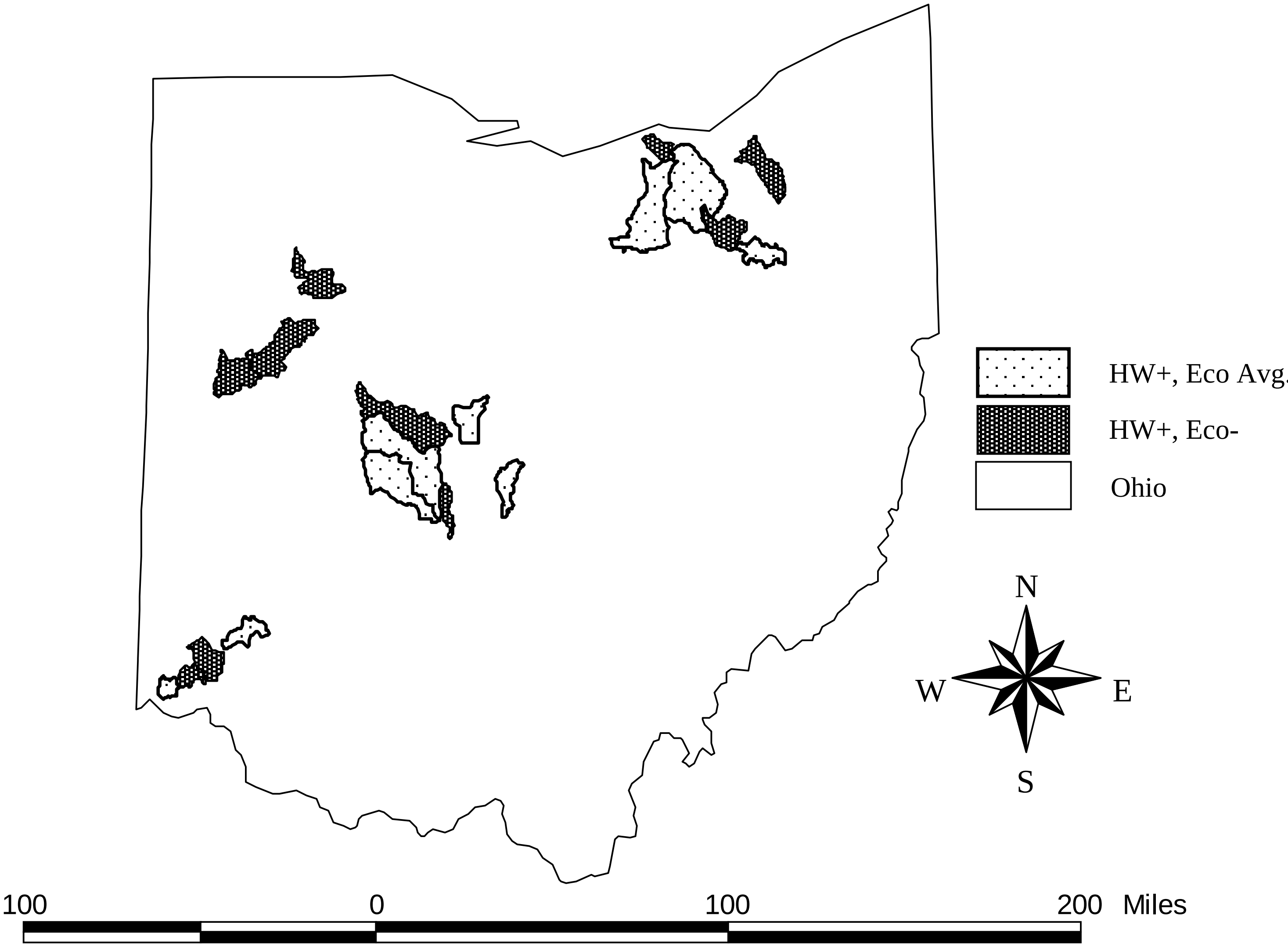


Figure 10
Map of Watersheds Studied With Above Average Human Conditions, and Average or Below Average Levels of Ecological Condition, Circa 1990



Enhancing the Spatial Comparison of Multiple Environmental Databases using the Prototype NY/NJ Harbor Environmental Data Management System (HEDMS)

Raymond M. Valente, Peggy M. Myre, and David C. Inglin[1]
Science Applications International Corporation

James Lodge, U. S. Army Corps of Engineers[2]

1.0 Introduction

The New York/New Jersey harbor area is one of the busiest ports on the East Coast. It has served as a reservoir of contaminants as a result of decades of pollution from multiple sources, including runoff from watersheds, shore-based industrial activity, and atmospheric input. Environmental data have been collected under the auspices of multiple national and regional monitoring programs, providing some of the information needed for environmental management of the harbor. The Contaminant Assessment and Reduction Program (CARP) is a new program involving the collection of multiple data types. This large-scale monitoring effort is intended to address the distribution of contaminants in the water, sediments and biota of the harbor, with the ultimate goal of source reduction. However, data comparability problems exist because the historical data, as well as the new CARP data, have been collected using different programmatic objectives, sampling protocols and laboratory data quality objectives.

The New York District of the U.S. Army Corps of Engineers has developed a prototype environmental chemistry database with a geographic-based (GIS) user interface for New York/New Jersey Harbor. The prototype is the first phase of the Harbor Environmental Data Management System (HEDMS) that ultimately will include multiple environmental databases (chemical, geological, biological and physical), as well as tools for querying and analyzing the data in a geographic context. The primary objective of the HEDMS prototype is to allow the user to readily select different data types, import them into a geographic interface, and conduct spatial data analyses while reducing common problems associated with chemical data

[1] Raymond M. Valente, Peggy M. Myre, and David C. Inglin, Science Applications International Corporation, 221 Third Street, Newport, RI 02840 **(**rvalente@mtg.saic.com; pmurray@mtg.saic.com: dinglin@mtg.saic.com**)**

[2] James Lodge, U. S. Army Corps of Engineers, New York District, 26 Federal Plaza, New York, NY (james.lodge@nan02.usace.army.mil)

comparison.

The system provides a user interface which includes utilities to increase comparability of databases collected for different programs using different analytical methods. One of the key requirements for HEDMS is the inclusion of extensive data documentation, including laboratory method and quality control information, spatial metadata, and overall program metadata (including program objectives and sampling design).

2.0 Description of HEDMS Features

The primary purpose of the HEDMS tool is to allow the user to readily select different data types, import them into a geographic interface, and conduct spatial data analyses while reducing common problems associated with chemical data comparison. The current system provides a graphic user interface (GUI) that includes utilities to increase comparability of databases collected for different programs using different analytical methods (Figure 1). The system graphically displays georeferenced sample locations from specific projects, NOAA coastlines, major rivers and tributaries, and other relative boundaries. Another key feature of HEDMS is the inclusion of extensive data documentation, including laboratory method and quality control information, spatial metadata, and overall programs metadata (including program objectives and sampling design).

Software/Hardware requirements: The HEDMS interface is a development within the ESRI ArcView® GIS program. The actual HEDMS program acts as an ArcView® Extension and requires a licensed copy of the ESRI ArcView® program. An Extension (Figure 2) is an add-in to ArcView® that adds custom capabilities. Extensions allow the newly created functions to be added to or removed from ArcView® with ease. The database is implemented using Open Database Connectivity (ODBC), which allows the user to access the database readily without having the original program that the database was created in, such as MS Access. The system is delivered on a single CD-ROM and requires either a Microsoft Windows95, 98 or NT platform. Minimum system requirements are a processor speed of 166 MHz, 32 megabytes of RAM, and at least 50 megabytes of free hard drive space.

Data types: The prototype HEDMS represents the first phase of a longer-term system development effort. It has been developed specifically for the NY/NJ harbor estuary area and includes multiple chemical databases from this area. Presently, the system is populated mainly with base map data and data on the concentrations of suite of chemical contaminants in surface sediments. The HEDMS system will eventually hold multiple environmental databases (chemical, geological, biological and physical) and will allow display of not only surface sediment chemistry data, but also subsurface sediment chemistry from cores, tissue chemistry, and biological testing data (e.g., toxicity and bioaccumulation).

The first-phase base map data include NOAA medium-resolution coastline, channels, navigation information, watersheds of NY/NJ, and jurisdictional boundaries, as well as a limited amount of supporting data, gleaned largely from the EPA's BASINS database (e.g., locations of pollution point sources like Superfund sites and NPDES discharge permit holders). Marine sediment chemistry data from the following monitoring programs also are included in the present system.

- U.S. Environmental Protection Agency (EPA) Environmental Monitoring and Assessment Program for Estuaries (EMAP-Estuaries): sediment chemistry data from sampling in the Virginian Province for the years 1990 through 1993. Data from throughout the NY/NJ Harbor Estuary system and the Hudson River.
- U.S. EPA Regional Environmental Monitoring and Assessment Program (R-EMAP): sediment chemistry data from throughout the NY/NJ Harbor Estuary system for the years 1991 and 1993.
- National Oceanic and Atmospheric Administration (NOAA) National Status and Trends and Program (NS&T). Sediment chemistry data from both the "Mussel Watch" and "Benthic Surveillance" components of this long-term marine monitoring effort.
- NOAA Study of Sediment Toxicity in the Hudson-Raritan Estuary in 1994.

Data selection: The program was designed to provide users with speed and flexibility in querying the database for the purpose of choosing what data and which parameters they want to examine. The "picker screen" utility provides an interactive window to allow the user to select data by program, data set, date, or data type (Figure 3). Once these basic specifications have been set, the user can then specify a particular analyte group, such as metals, organic contaminants such as PAHs, PCBs, or pesticides, or physical parameters such as sediment grain size (Figure 4). The data query can be further refined using specific numeric qualifiers (e.g., above or below a specified value). The remaining functionality parameters for this window allow the user to either view or export data to a table or text format (Figure 5), add data to a map view as an ArcView® theme, and finally to plot just station locations into a view window.

When the user chooses to export data to a table, they can choose exactly what columns they want to view in the table by clicking on desired headers (Figure 6).

Data assessment and comparison: The system contains a number of functions to facilitate environmental assessments and comparisons across data sets.

- Under the "preferences" menu, the user can specify how "below detection limit" (BDL) chemistry values are to be treated (Figure 7). The system provides the following options for handling BDL data in any subsequent calculations or comparisons: substitute zero, substitute one-half of detection limit, substitute detection limit, or ignore.
- A menu function allows the user to sum the results for individual analytes into commonly-employed "aggregate" classes. For example, the sediment concentrations of 15 to 20 individual PAH compounds commonly are summed to produce an aggregate value called "total PAHs." Likewise, "total PCBs" are calculated based on summing the values for individual PCB congeners. Different investigators traditionally have used slightly different approaches to defining/calculating aggregate values like total PAHs and total PCBs. In recognition of such differences, the HEDMS program allows the user to identify which of the individual compounds will be summed to produce the aggregate value (Figure 7). This ensures a common basis for comparison among different monitoring programs.
- A menu screen is provided to allow the user to "normalize" the sediment chemistry data by either grain size or total organic carbon (TOC) content.
- Data on sediment metal concentrations can be normalized by the sediment concentration of acid volatile sulfide (i.e., SEM/AVS ratio).
- The "compare to reference" function allows the user to calculate and display the results of comparisons to different, commonly-employed "reference" values (e.g., Effects Range-Low/Effects Range-Median values (ERL/ERM), EPA Sediment Quality Criteria (SQC), Threshold/Probable Effects Levels (TEL/PEL)).

Data display: The GIS application to this system is a very useful tool, in that it allows the user to create a view and choose the area of interest to be displayed. The area is then displayed in the state plane projection reference. The view can be set for a number of projection changes such as latitude/longitude or UTM units, and the user can also change the view's distance units. Once the view has been given the desired projection, the user will add the data of interest and can start examining the data spatially (Figure 8).

The HEDMS program also contains a preference feature with functions that allow the user to personalize a work session for future uses. The user can actually save the chosen preference configuration to an individual preference file that can be re-loaded. The functions for dealing with “below detection limit” values and for the calculation of aggregate values are included under the preferences menu (Figure 7). The preferences feature is intended to make the program a more flexible to a larger group of users. For example, if there were only one computer that was available with the HEDMS program installed, a number of different users could have their own preference files containing the features that they want to see, as opposed to having the user start anew each time they began a work session.

Quality Assurance/Quality Control: Chemistry data typically require a careful assessment of quality before use in a study or spatial evaluation. Many times this quality assessment is lost when a data set is incorporated into a larger database. The HEDMS has been designed to store, access, and use documentation and quality control data provided by the source laboratory and also provide a framework for documentation of the decisions made in the evaluation of data quality. For each sample result, the system is capable of producing a QA/QC report for the associated quality control data such as matrix spikes, duplicates, Standard Reference Materials, and blanks. In addition, the system allows users to access program-level metadata for viewing or downloading in simple text file format. The metadata provides a complete set of documentation on each data set, including information on program sponsors/participants, objectives, sampling and analytical methods, analyte lists, and data quality objectives.

3.0 Summary

The HEDMS has been designed to enhance the utility of historical databases (“data mining”), increase the statistical confidence (power) of environmental data interpretation by integrating multiple databases, increase the longevity of newly collected data, and heighten comparability of disparate data sources.

This system is the first step in a plan to create a useful tool to manage large environmental data sets of varying complexities. The main focus of the existing prototype is chemical data from the NY/NJ harbor estuary area. However, future versions of the software are not limited to chemical data nor is the area of interest a limitation. HEDMS incorporates utilities for integration and comparison of multiple databases, which can be done using a variety of data types. A primary system function is the automated features included for complex chemical data analysis. The system provides rapid spatial analysis due to the integration of the program into a GIS platform. The flexibility of the system and the use of the existing ArcView® platform allow for a efficient, user-friendly analysis tool designed to save resource managers and other decision-makers valuable time.

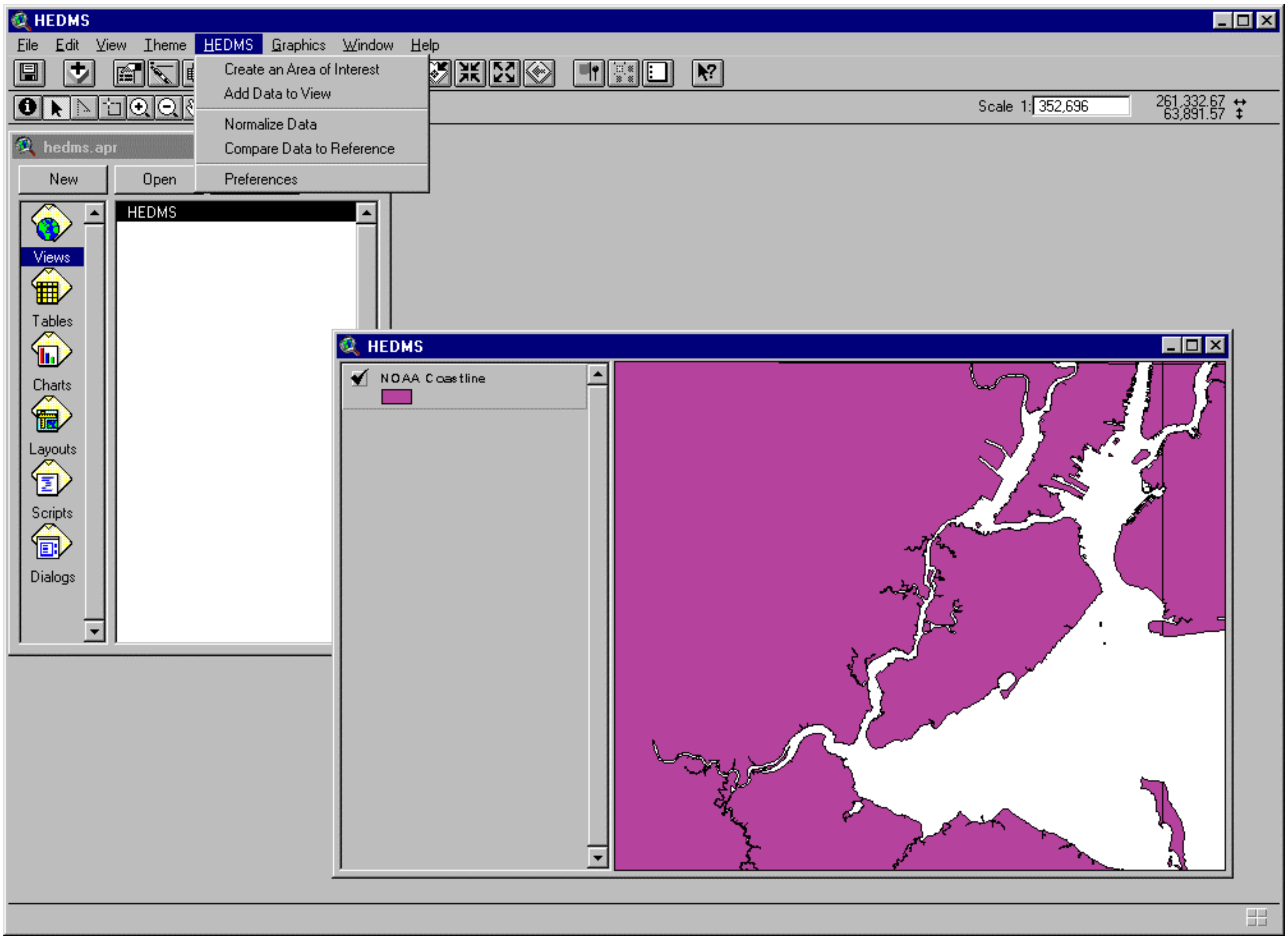


Figure 1. Graphic User Interface (GUI) of the HEDMS program that users will most commonly start from when creating a new view.

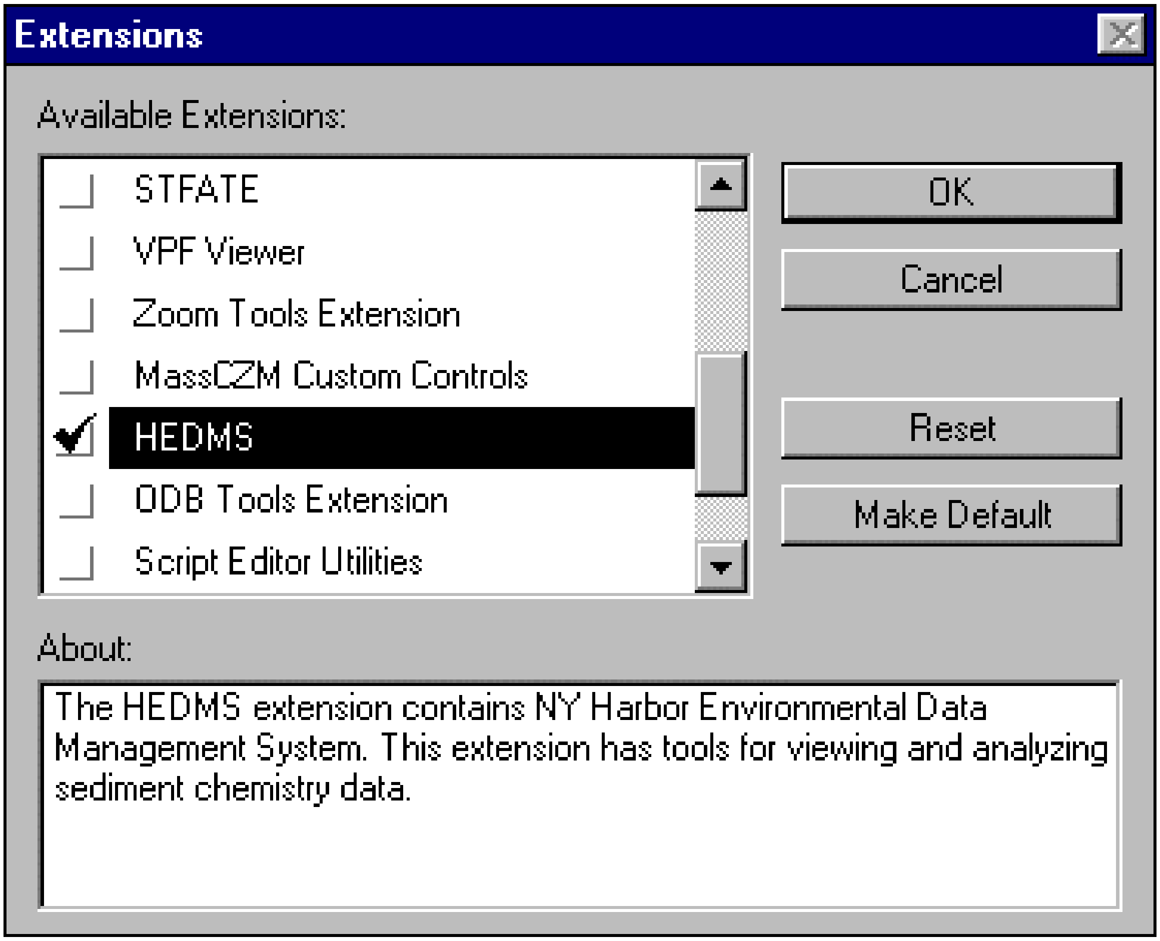


Figure 2. The HEDMS has been designed to run as an ArcView® Extension.

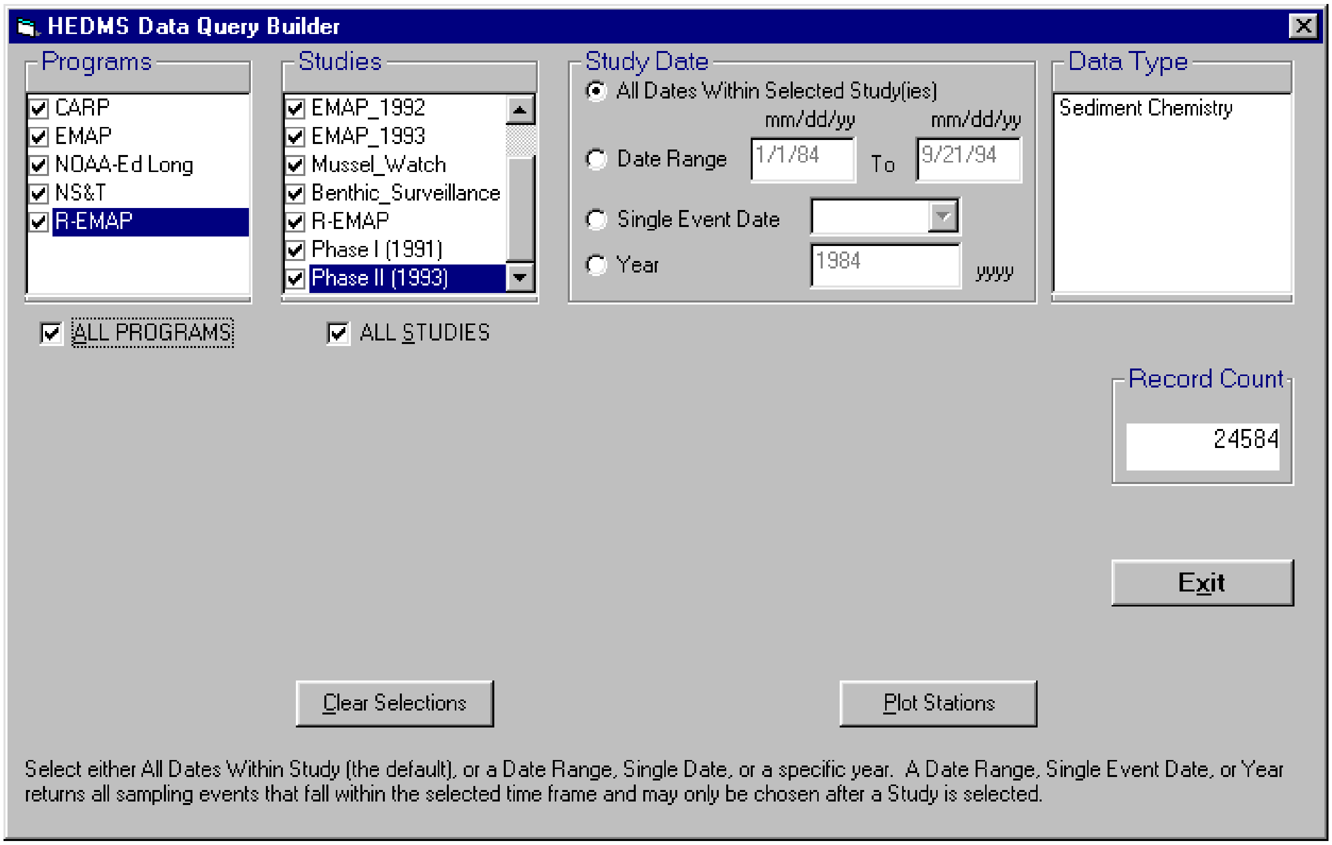


Figure 3. The initial "picker screen" window, which allows the user to select data by program, data set, date, or data type.

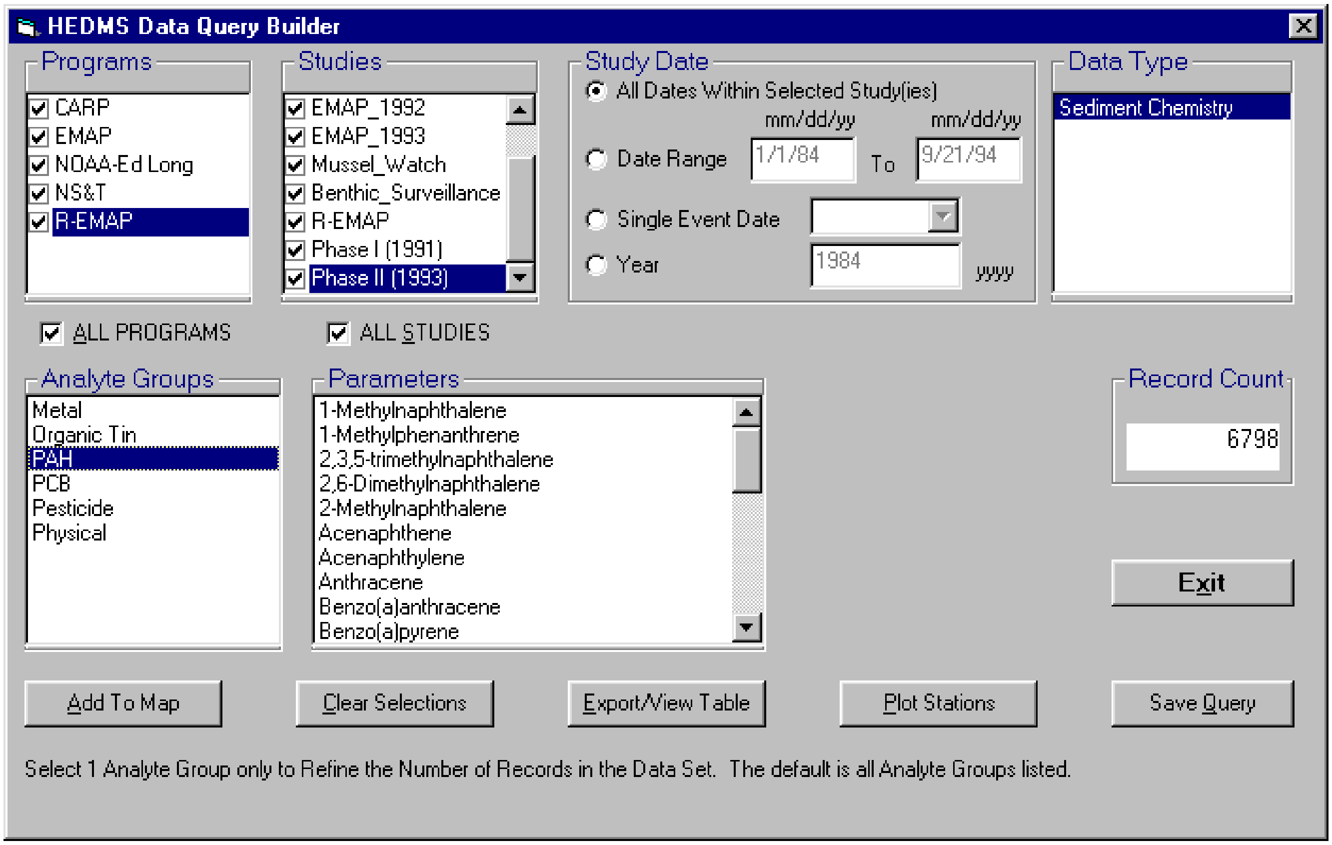
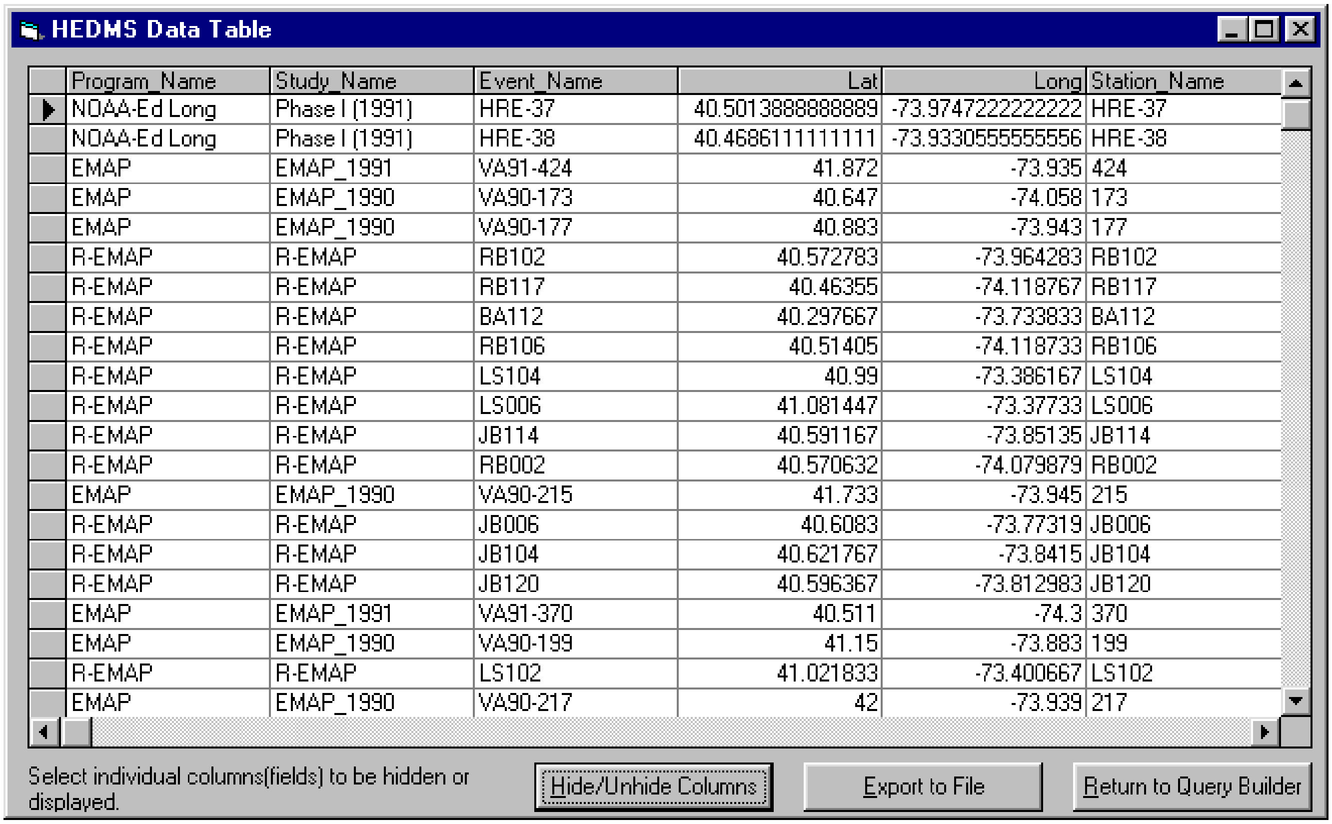


Figure 4. After the user chooses the first set of screening options, additional parameters appear to allow for more specific screening.



Program_Name	Study_Name	Event_Name	Lat	Long	Station_Name
NOAA-Ed Long	Phase I (1991)	HRE-37	40.5013888888889	-73.9747222222222	HRE-37
NOAA-Ed Long	Phase I (1991)	HRE-38	40.4686111111111	-73.9330555555556	HRE-38
EMAP	EMAP_1991	VA91-424	41.872	-73.935	424
EMAP	EMAP_1990	VA90-173	40.647	-74.058	173
EMAP	EMAP_1990	VA90-177	40.883	-73.943	177
R-EMAP	R-EMAP	RB102	40.572783	-73.964283	RB102
R-EMAP	R-EMAP	RB117	40.46355	-74.118767	RB117
R-EMAP	R-EMAP	BA112	40.297667	-73.733833	BA112
R-EMAP	R-EMAP	RB106	40.51405	-74.118733	RB106
R-EMAP	R-EMAP	LS104	40.99	-73.386167	LS104
R-EMAP	R-EMAP	LS006	41.081447	-73.37733	LS006
R-EMAP	R-EMAP	JB114	40.591167	-73.85135	JB114
R-EMAP	R-EMAP	RB002	40.570632	-74.079879	RB002
EMAP	EMAP_1990	VA90-215	41.733	-73.945	215
R-EMAP	R-EMAP	JB006	40.6083	-73.77319	JB006
R-EMAP	R-EMAP	JB104	40.621767	-73.8415	JB104
R-EMAP	R-EMAP	JB120	40.596367	-73.812983	JB120
EMAP	EMAP_1991	VA91-370	40.511	-74.3	370
EMAP	EMAP_1990	VA90-199	41.15	-73.883	199
R-EMAP	R-EMAP	LS102	41.021833	-73.400667	LS102
EMAP	EMAP_1990	VA90-217	42	-73.939	217

Figure 5. One option available from the data picker screen (Figure 4) is export of the selected data to a table. The user has the ability to select/deselect different table columns (see Figure 6) and export the table into a text file.

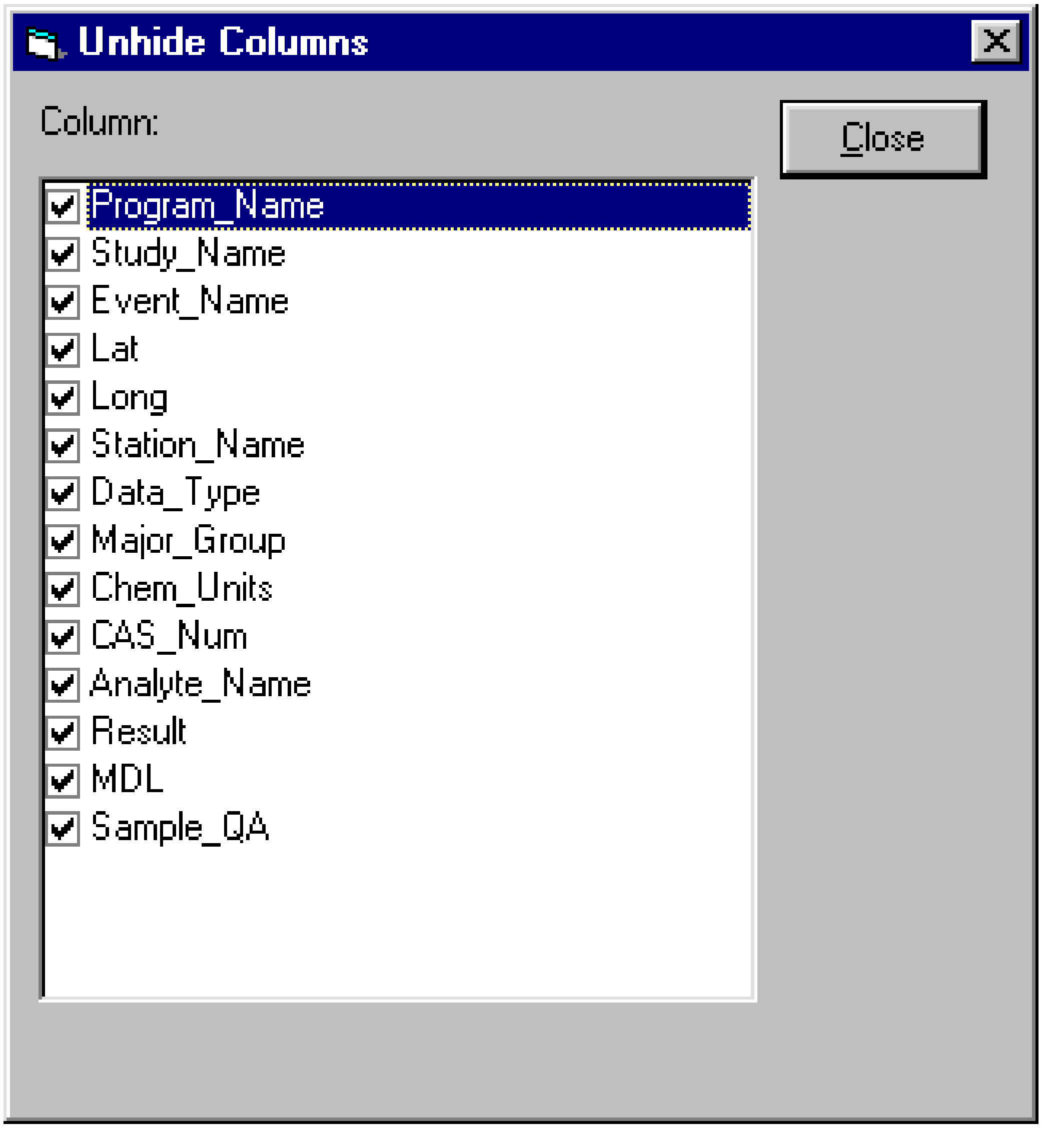


Figure 6. The "unhide columns" window, which allows the user to choose the data to be displayed in table format.

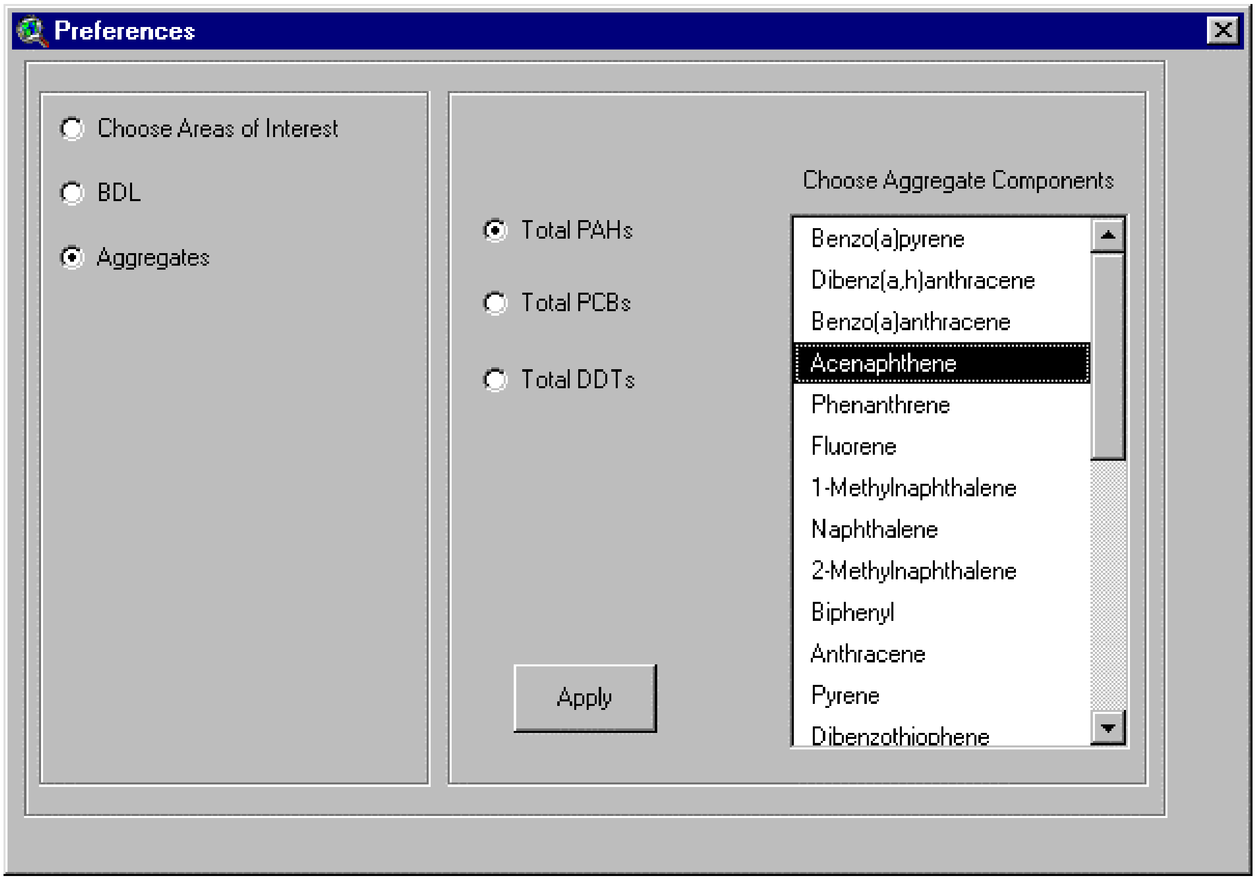


Figure 7. The preferences menu allows users to choose how below detection limit values are handled and how aggregates are to be calculated.

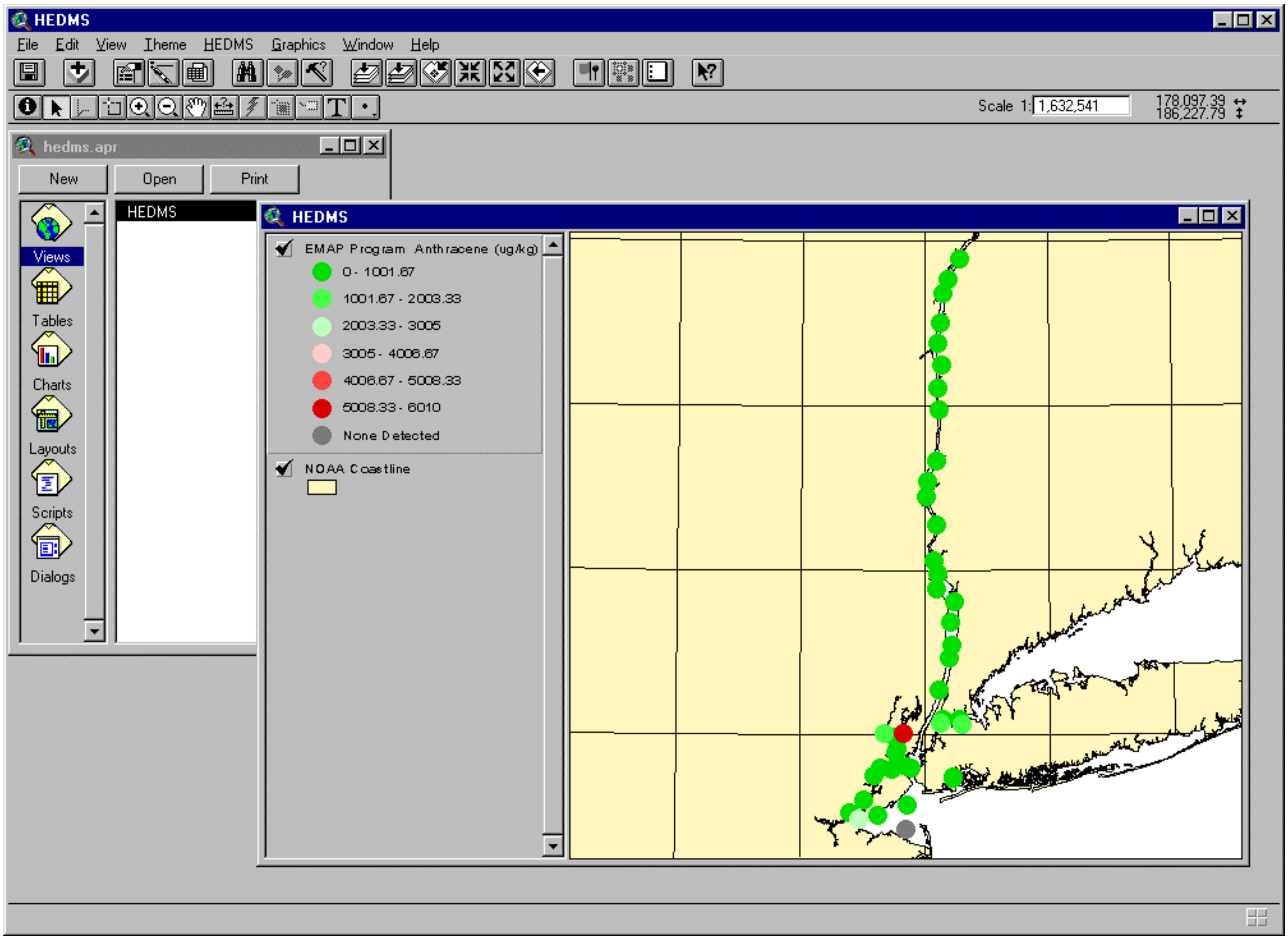


Figure 8. Example of spatial data display within HEDMS. The user can add or remove as many themes or layers as they choose in performing spatial evaluation of the data.

Nonpoint Pollutant Loading Application for ArcView GIS

Laurens van der Tak, Mike Miller, Cody Zook, Cheri Edwards[1]

I. Pollutant Loading Application Overview

An ArcView GIS tool was developed that calculates nonpoint source pollutant loads for watersheds and sub-watersheds. The application is called PLOAD. Currently the application is set up to estimate nonpoint sources (NPS) of pollution on an annual average basis, for any user specified pollutant. The user has the option to calculate the NPS loads with either the Simple Method, or with areal export coefficients. The program was designed to be generic so that it can be applied as a screening tool in typical NPDES stormwater permitting projects, watershed management projects or reservoir protection projects, and readily modified for special needs. It is designed to be a useful analytical tool for end users (primarily water resources engineers and planners). Therefore, it was programmed in a desktop GIS environment, ArcView, using the ArcView Avenue programming language.

The tool requires the following input data:

- GIS landuse data.
- GIS watershed data.
- GIS BMP site and area data (optional) showing the location as a point coverage with the BMP type and service area, or as a polygon coverage with BMP type and the polygon boundaries delineating the service area. This layer is optional, if BMPs are included in the analysis.
- Pollutant loading rate data tables for each land use type showing either event mean concentrations (EMCs, in mg/L of a given pollutant), or export coefficients (e.g. pounds per acre of a given pollutant).
- Impervious data tables.
- Pollutant reduction BMP data tables (optional).

[1] Vice President and Sr. Water Resources Engineer, Sr. Gis Analyst, GIS Analyst, and Water Resources Engineer, CH2M HILL. Address questions to: CH2M HILL, 13921 Park Center Rd., Suite 600, Herndon VA 20171. Tel.: 703-417-1441.

Several output options are available:

- Tabular summaries of pollutant loads by basin (lbs.), EMCs by basin (mg/L), aerial loading rates by basin (lbs./acre/yr.), pollutant loads by land use (lbs.).
- Graphical summaries of pollutant loads by basin (lbs.), EMCs by basin (mg/L), and aerial loading rates by basin (lbs./acre/yr.)

This paper presents the theory and application of the PLOAD tool. The PLOAD application input data, pollutant evaluation equations, products, installation instructions, and general use guidelines are described in detail below. A case study is presented showing input and output screens, and illustrating how the tool was used to develop nonpoint source pollutant loads for NPDES stormwater permit annual reports for the Cities of Chesapeake, Hampton, Newport News, Norfolk, Portsmouth, and Virginia Beach, Virginia. These six cities are working together as part of the Hampton Roads Planning District Commission (HRPDC) Regional Stormwater Management Committee (RSMC).

II. Input Data

A variety of GIS and tabular source data is accessed by the PLOAD application. This section describes the required and optional input data components. Note that the GIS data must be developed as either ESRI Arc/Info coverages or ArcView shape files, while the tabular data may be prepared as Excel, comma delimited text, dBASE, or INFO files.

II.1 GIS Data

Watershed basin and landuse GIS data coverages are required for PLOAD. The basins define the areas for which the pollutant loads are calculated. The basin coverage must have a code field containing unique identifiers for each basin. The landuse file is essential for calculating the pollutant loads. The landuse coverage must also have a code field identifying the landuse types, but these types need not be unique. Prior to calculating the pollutant loads, PLOAD will spatially overlay the basin and landuse coverages in order to determine the areas of the various landuse types for each basin. The landuse coverage should encompass the entire basin coverage.

Digital watershed basin were available for the six Virginia cities. However, current landuse data were developed for several of the cities by digitizing planning maps, or converting real estate parcel information and infrared aerial photography.

Best Management Practices (BMPs) serve to reduce pollutant loads using natural processes (settling, filtration, biological uptake) for the BMP area of influence. PLOAD will account for the influence of either site or areal BMPs. Site BMPs represented as point GIS files must contain attribute codes describing the BMP type and area of influence. Areal BMPs must be delineated as polygon files coded for BMP type only. The polygon boundaries define the area of influence. BMP input is optional because they may not exist for the area of evaluation or be desired for analysis.

II.2 Tabular Data

Pollutant loading rate, impervious factor, and BMP efficiency information must be compiled in tabular files for use in the PLOAD application. The three files of tabular input data can be provided in one of four formats: Excel spreadsheet, comma-delimited text, dBASE, or INFO database tables.

The pollutant loading tables consist of the either the event mean concentration (EMC) or the export coefficient. The user can choose to use either form of pollutant loading rates (EMCs and aerial export coefficients) because data are typically available for the EMC and export coefficient tables for urban and rural landuse types, respectively. Pollutants commonly evaluated include: TSS, Total Nitrogen, TKN, Nitrate plus Nitrite, Lead, Zinc, BOD_5, COD, Ammonia, Total Phosphorus, and Fecal Coliform.

The impervious factor table identifies percent imperviousness for each landuse type. The BMP table identifies the percent efficiency for reducing pollutant loads for each BMP type. Multiple versions of each type of table may be generated to simulate alternative conditions. A description of each lookup table is provided below.

II.2.1 Export Coefficient Table

The export coefficient table lists loading rates for each pollutant type by landuse type. The table may contain any number of landuse and pollutant types. There should be loading rates for each landuse and pollutant type in the evaluation area, otherwise the load for the area will be zero. The rates in the export coefficient table are measured in lbs./acre/year and are typically used to calculate the pollutant loads for rural landuse types. An example of an export coefficient table follows:

Landuse Type	Total Nitrogen (lb./acre/year)	Total Phosphorus (lb./acre/year)	TSS (lb./acre/year)	Lead (lb./acre/year)
Agriculture	21.2	3.5	0	.18
Pastures	5.9	0.5	0	.06
Forests	2.4	0.0	0	.02

II.2.2 Event Mean Concentration Table

The event mean concentration (EMC) table is identical to the export coefficient table, except the EMC values are measured in mg/L and they are typically used to calculate the pollutant loads for urban landuse types. The following is an example of an EMC table:

Landuse Type	Total Nitrogen (mg/L)	Total Phosphorus (mg/L)	TSS (mg/L)	Lead (mg/L)
1-Family Res	2.00	.26	52.60	.037
Rural Res	2.00	.26	52.60	.037
Industrial	1.62	.35	79.25	.031
Commercial	2.77	.38	56.50	.027

II.2.3 Impervious Factor Table

The impervious factor table identifies the percent imperviousness for each landuse type. It is used to calculate the EMC runoff coefficient. The table may contain any number of landuse types, but there should be impervious percentages for each landuse type in the evaluation area. If there is no impervious factor in the table for a particular landuse type, then the EMC runoff coefficient will default to .05 for areas with that landuse. The names describing the landuse types must be the same in the table as they are in the GIS landuse file.

The following is an example of an impervious factor table:

Landuse Type	% Imperviousness
Residential	18
Commercial	84
Industrial	71

II.2.4 BMP Efficiency Table

The BMP table contains percent efficiency multipliers for each BMP type that are used to calculate pollutant load reductions. The first record (row) of the table identifies the field names starting with BMP type followed by the pollutants under evaluation. The table may contain any number of BMP types. The pollutant types without percent efficiency multipliers will not reduce the pollutant load for the BMP type.

The BMP table was developed by water resource engineers by using literature values, or by analyzing local monitoring data comparing pollutant loads entering and leaving BMPs.
The following is a simple example of a BMP efficiency table:

BMP Type	Total Phosphorus (% removal)	Total Nitrogen (% removal)	TSS (% removal)
Dry Pond	30	20	50
Basin	45	25	60
Grass Filter Strip	30	50	20

III. Pollutant Loading Calculation Equations

Annual pollutant loads may be calculated for each watershed basin using either the pollutant export coefficient or simple methods. Optionally, the pollutant loads derived from these methods may be refined based on the remedial effects of BMPs. Descriptions of the equations used to calculate the pollutant loads follows:

III.1 Export Coefficient Method

If the export coefficient method is designated for calculating pollutant loads in PLOAD, then the loads are calculated for each specified pollutant type by basin using the following equation:

$$L_P = \Sigma_U(l_{PU} * A_U)$$

Where: L_P = Pollutant load, lbs.;

L_{PU} = Pollutant loading rate for landuse type u, lbs./acre/year; and

A_U = Area of landuse type u, acres

The loading rates are derived from the export coefficient tables, while the landuse areas are interpreted from the landuse and basin GIS data.

III.2 Simple Method

If the Simple Method is designated for calculating pollutant loads in PLOAD, then two equations are required to calculate the loads for each specified pollutant type. First, the runoff coefficient for each landuse type must be derived with the equation:

$$R_{VU} = 0.05 + (0.009 * I_U)$$

Where: R_{VU} = Runoff Coefficient for landuse type u, $inches_{run}/inches_{rain}$

I_U = Percent Imperviousness

Percent impervious is extracted from the impervious factor table.

The pollutant loads are then calculated with the following equation:

$$L_P = \Sigma_U (P * P_J * R_{VU} * C_U * A_U * 2.72 / 12)$$

Where: L_P = Pollutant load, lbs.

P = Precipitation, inches/year (default = 40.86)

P_J = Ratio of storms producing runoff (default = 0.9)

R_{VU} = Runoff Coefficient for landuse type u, $inches_{run}/inches_{rain}$

C_U = Event Mean Concentration for landuse type u, milligrams/liter

A_U = Area of landuse type u, acres

The precipitation and storm ratio values are entered by the PLOAD user interactively. The loading rates are derived from the EMC tables, while the landuse areas are interpreted from the landuse and basin GIS data.

III.3 BMP Computations

BMPs serve to reduce pollutant loads and PLOAD has an option to calculate loads based on the remedial effects of the various BMP types. This section describes the equations that are used to calculate pollutant loads influenced by BMPs. BMP types may be represented as either area or site features, but the approach for both is similar. After the raw pollutant loads are calculated using the export coefficient or simple methods, three equations are used to recalculate the pollutant loads.

First, the percent of the basins area serviced by BMPs are determined using the following equation:

$$\%AS_{BMP} = AS_{BMP}/A_B$$

Where: $\%AS_{BMP}$ = Percent area serviced by the BMP, decimal percent

AS_{BMP} = Area serviced by the BMP, acres

A_B = Area of basin, acres

The BMP and basin areas are derived from the BMP and basin GIS data. Next, the pollutant loads for each BMP are calculated:

$$L_{BMP}= (L_P * \%AS_{BMP}) * [1- \%EFF_{BMP}/100]$$

Where: L_{BMP} = BMP load, lbs.

L_P = Raw basin load, lbs.

%EFF = Percent load reduction of BMP, percentage

The raw basin pollutant loads are derived from the results of the export coefficient or simple methods, while the percent load reduction comes from the BMP efficiency tables.
Finally, the total pollutant loads accounting for BMPs are computed by basin. Each basin load is a cumulative total of areas which are and are not influenced by BMPs.

$$L = (\Sigma_{BMP} (L_{BMP})) + L_P * (A_B - (\Sigma_{AS} (AS_{BMP}))$$

IV. Optional User Input Parameters

Many input parameter options have been built into the PLOAD tool that the user must specify. Several of the more important ones are listed as follows:

- Specify watershed basin and landuse GIS data files.
- Select single, multiple, or all watershed basins from basin file for evaluation.
- Specify either the export coefficients or the simple method for calculating pollutant loads.

- If simple method is specified, then enter annual precipitation and ratio of storms producing runoff value to override defaults.
- Evaluate pollutant loads with or without BMPs.
- When BMPs are evaluated, identify whether they are derived from point or polygon GIS data.
- Select output products.
- Save file of input data sources and parameter settings that may be used to rerun PLOAD at a later date, with or without input modifications.

V. Output Product Options

After the pollutant loads have been determined, PLOAD may be used to generate a variety of graphic plot and tabular report products. Listed below are the product options. See Appendix I – Graphic and Tabular Product Examples.

- Total Pollutant Loads by Basin – Map and Table
- Pollutant Loads Per Acre by Basin – Map and Table
- Event Mean Concentration (EMC) by Basin – Map and Table
- Pollutant Loads by BMP, Landuse and Basin - Table
- PLOAD Data Sources and Parameters - Table

VI. Potential Future Enhancements

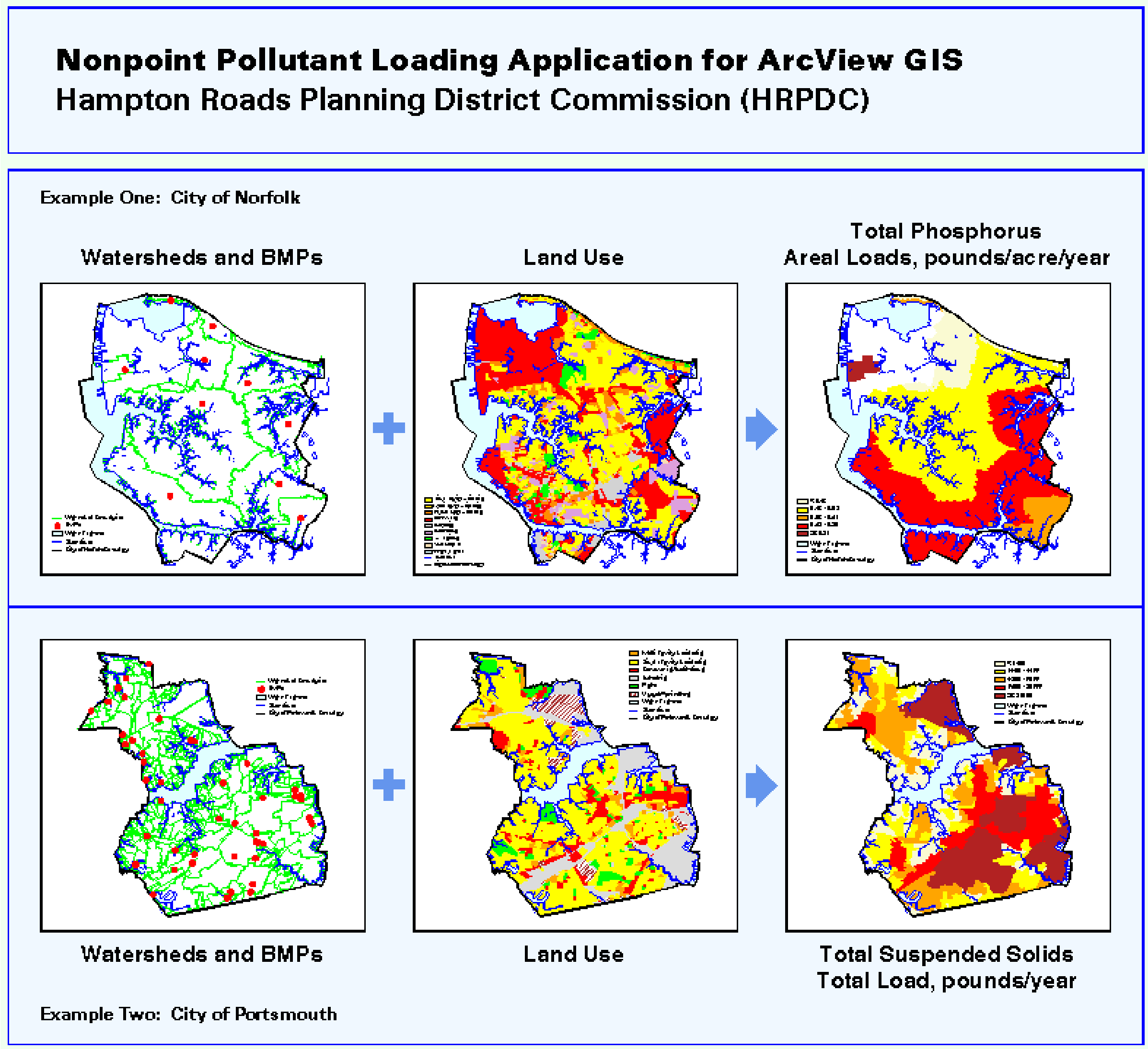
It is hoped that the PLOAD application will continue to be enhanced as needed on projects with related needs. Possible future enhancements include:

- Addition of point sources of pollution to account for stream baseflow background concentrations from groundwater or treatment plant discharges.
- Addition of pollutant transport or sediment delivery ratio concepts, to account for reductions in loads in-stream due to deposition or other mechanisms.
- Addition of other types of BMPs, such as those in series and source control type BMPs.
- Evaluate simple and export coefficient methods simultaneously to reflect both urban and rural landuses.
- Account for BMPs whose remediation properties overlap within a basin.

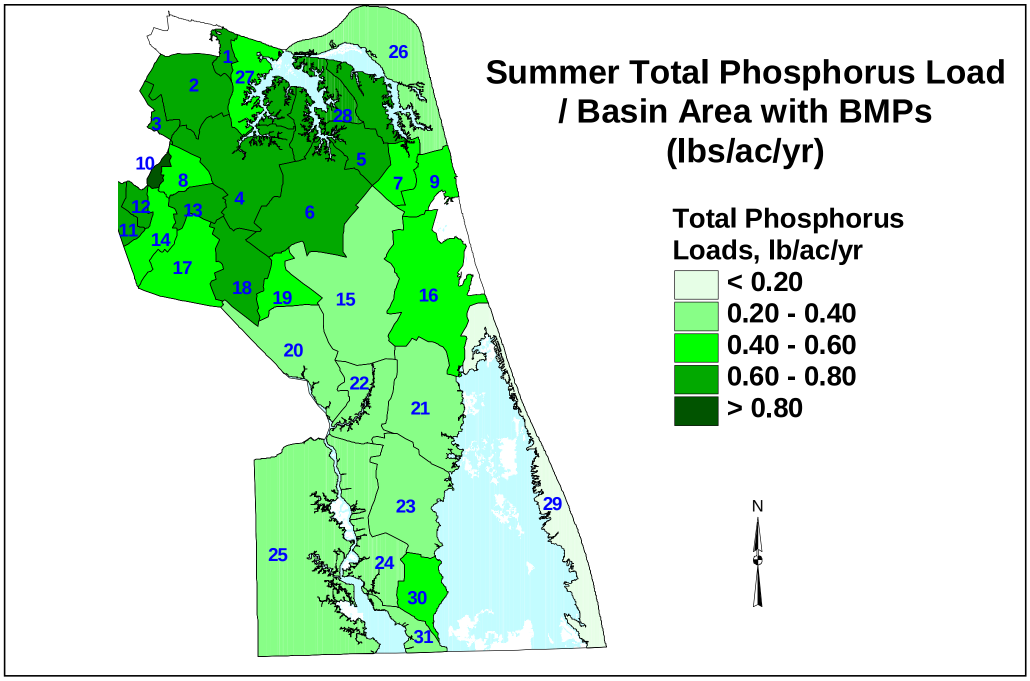
Appendix I

Graphic and Tabular Product Examples

The following examples are simplified versions of the PLOAD products.

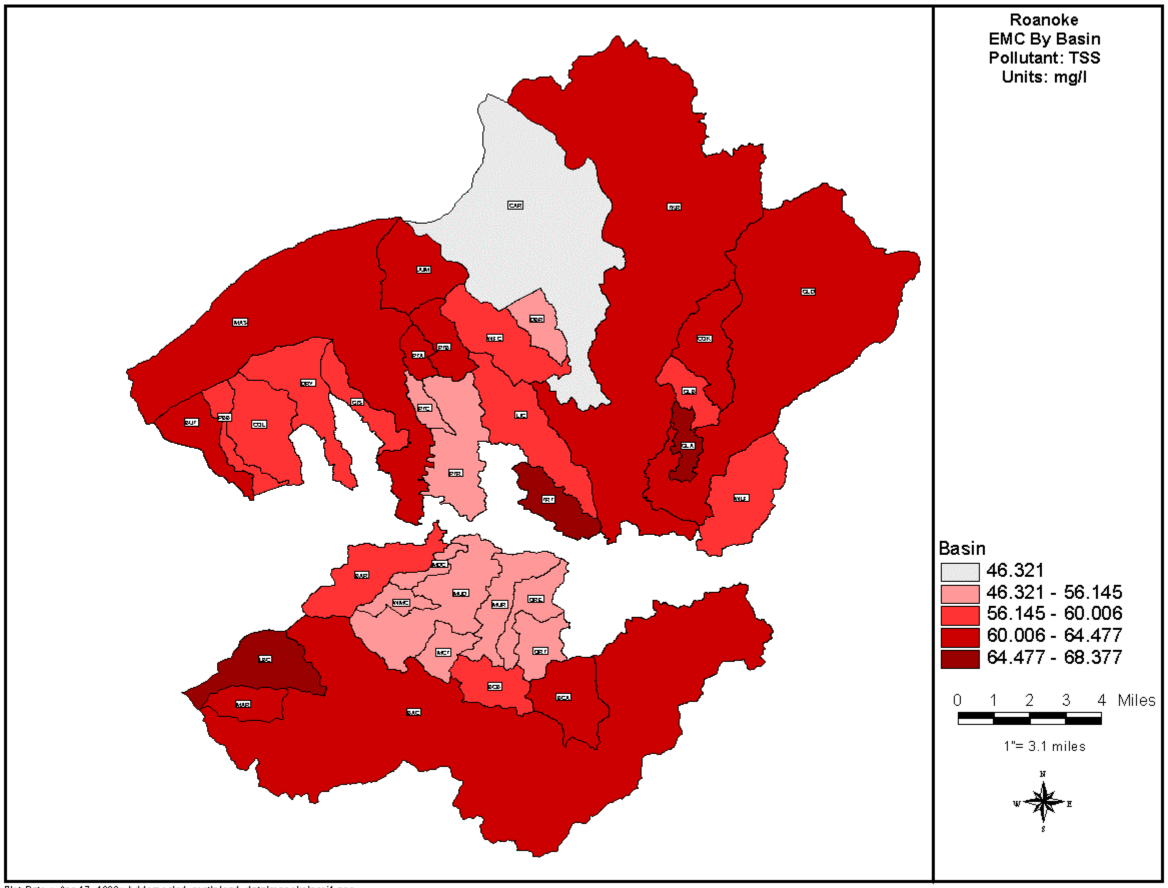


Pollutant Loads Per Acre by Basin - Map



Plot Date : June 15, 1999; j:\ploadapp\vabeach\w_bmps\summer\vb_summer_bmp.apr

Event Mean Concentration by Basin - Map



Plot Date : Apr 17, 1999; h:\demos\cd_root\pload_data\roanoke\proj1.apr

Pollutant Loads, Event Mean Concentrations, and Pollutant Loads by Basin Area – Table

Watershed	Count	Area	Total Load (lb./yr.)		Event Mean Conc. (mg/L)		Areal Load (lb./acre/yr.)	
			BOD	TSS	BOD	TSS	BOD	TSS
1	10	4948	102555	665717	8	53	21	135
10	15	12153	150439	1034060	7	49	12	85
11	9	4019	37662	360819	4	38	9	90
12	10	4894	106127	676898	8	54	22	138
13	11	4065	74297	530243	9	65	18	130
14	10	6576	74682	635766	4	35	11	97
15	8	2180	37024	231704	5	34	17	106
16	9	6248	56488	540617	4	41	9	87

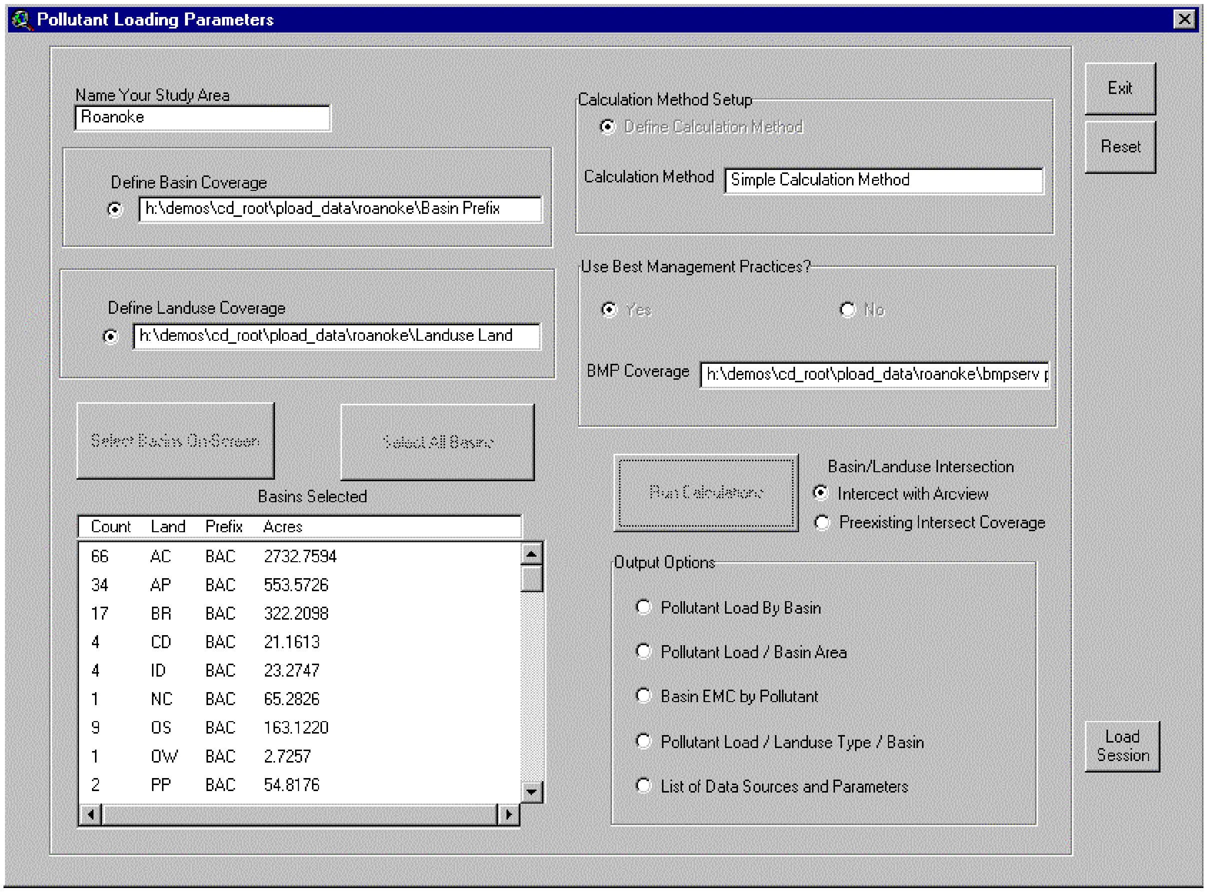
Pollutant Loads by BMP, Landuse and Basin - Table

	1	2	3	4	5
1	Roanoke AreaSum.dbf Load in lbs/yr				
2	Merge	Prefix	Land	Acres	LD_TSS
3	BAC AC no	BAC	AC	2732.759	189519.2
4	BAC AP no	BAC	AP	553.5726	35991.32
5	BAR OS no	BAR	OS	193.2443	11726.45
6	BAR OW no	BAR	OW	11.7574	0
7	BAR PP no	BAR	PP	21.3951	8067.278
8	BAR R1 no	BAR	R1	21.6817	2015.318
9	BAR R12 BASIN	BAR	R12	1.362	61.3881
10	BAR R12 no	BAR	R12	473.8863	53397.58
11	BAR R13 no	BAR	R13	86.0562	11394.71
12	BAR R14 no	BAR	R14	189.1171	31011.06
13	BAR R18 no	BAR	R18	18.6515	5045.605
14	BAR R2 no	BAR	R2	15.0539	924.0412
15	BAR WF BASIN	BAR	WF	10.0423	226.3437
16	BAR WF no	BAR	WF	1427.735	80449.39
17	BCA AC no	BCA	AC	141.9152	9841.941
18	MAS AC no	MAS	AC	563.3961	39072
19	MAS AP no	MAS	AP	76.5106	4974.446
20	MAS BR DRY POND	MAS	BR	0.0225	0.5461
21	MAS BR no	MAS	BR	26.5379	1610.373
22	MAS CD DRY POND	MAS	CD	2.5089	380.9389
23	MAS CD no	MAS	CD	470.054	178426.7
24	MAS ID no	MAS	ID	60.7858	27666.18
25	MAS OS DRY POND	MAS	OS	9.8503	239.0944
26	MAS OS no	MAS	OS	468.2329	28413.31
27	MAS PP DRY POND	MAS	PP	2.0393	307.577
28	MAS PP no	MAS	PP	105.4346	39755.38
29	MAS R1 no	MAS	R1	335.8436	31216.72
30	MAS R12 no	MAS	R12	105.9188	11934.95
31	MAS R13 no	MAS	R13	14.6245	1936.432
32	MAS R14 DRY POND	MAS	R14	0.2368	15.532

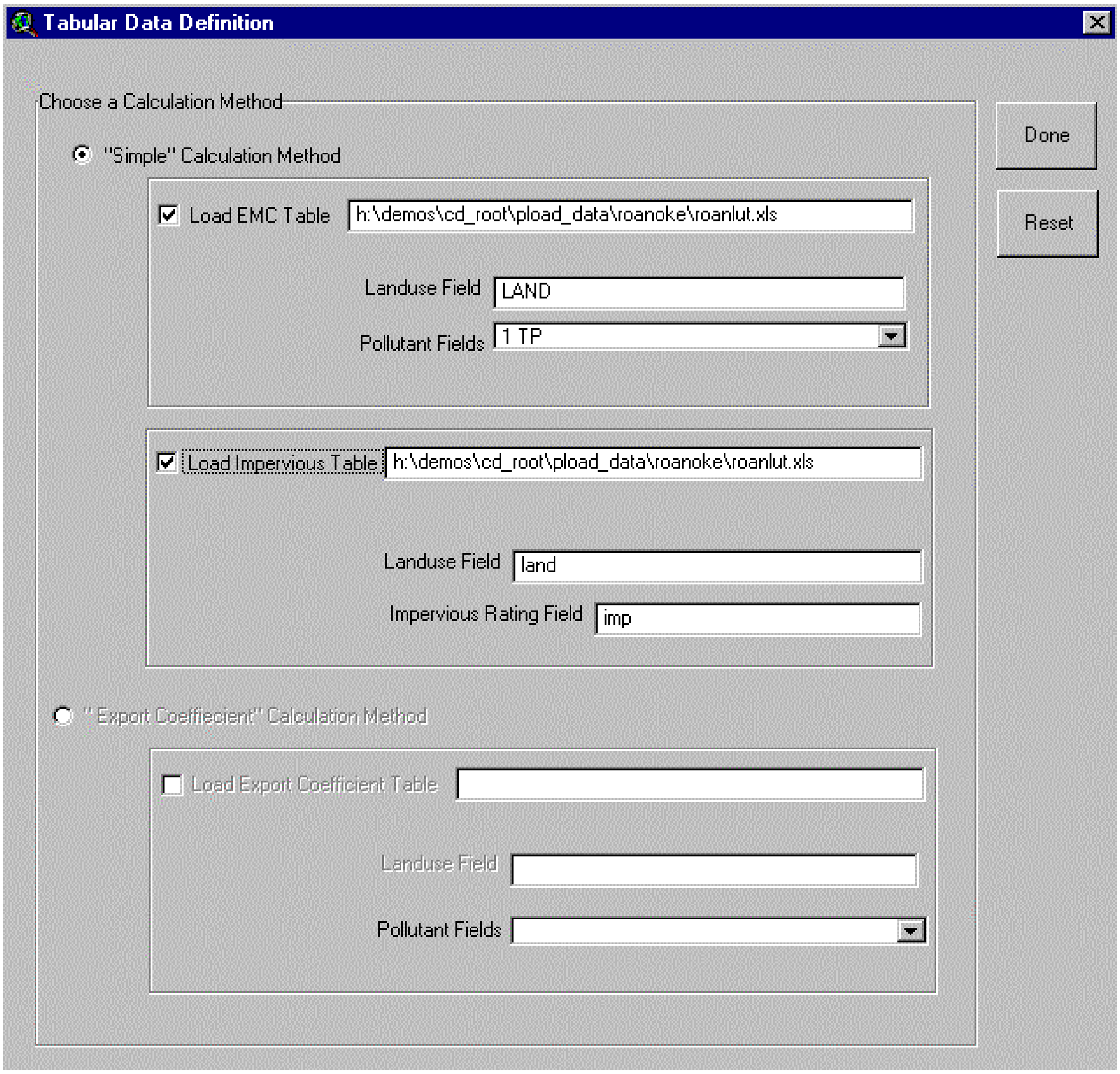
Appendix II

PLOAD Dialogue Menus

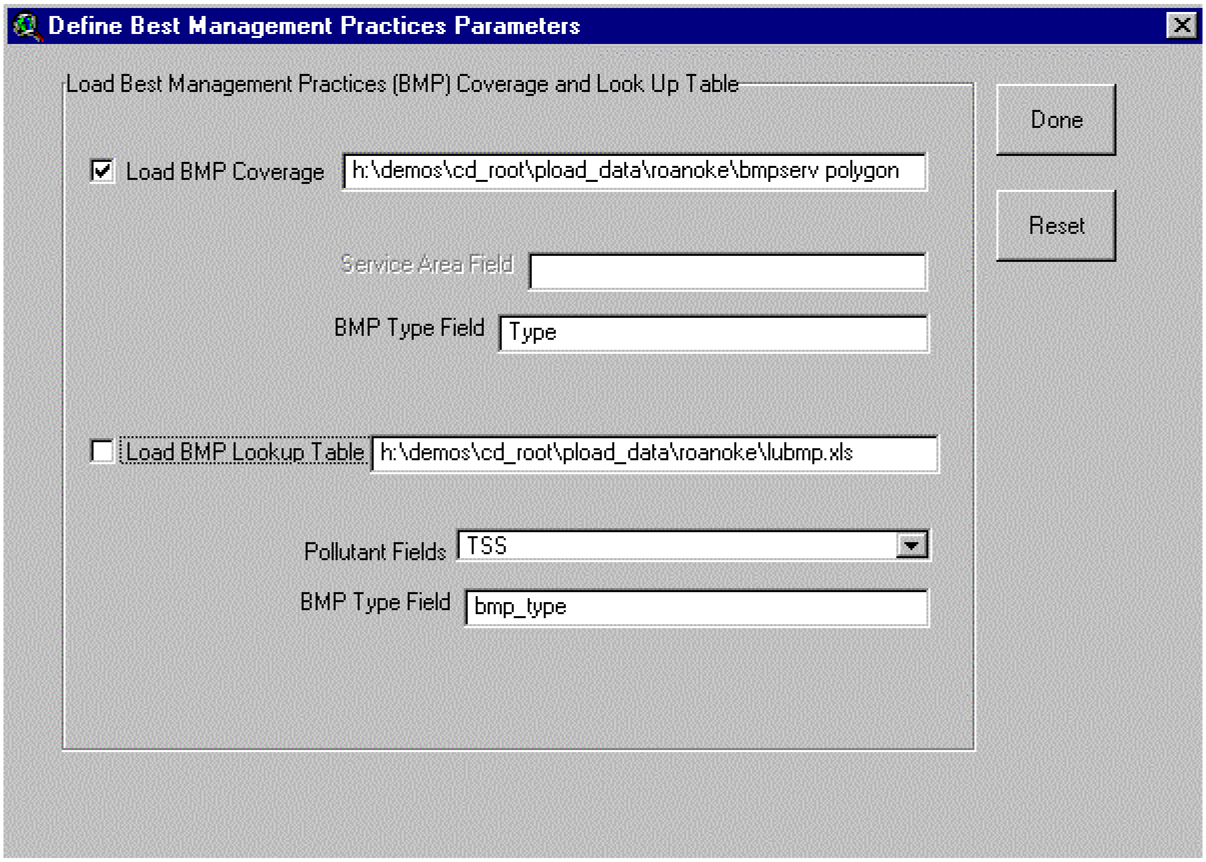
Pollutant Loading Parameters - Dialogue Design Menu



Tabular Data Definition - Dialogue Design Menu



Define Best Management Practices Parameters - Dialogue Design Menu



Using GIS to Analyze the Spatial Distribution of Environmental, Human Health, and Socio-Economic Characteristics in Cincinnati

Christopher Auffrey and Xinhao Wang
School of Planning, University of Cincinnati

ABSTRACT

The hazardous byproducts of commercial processes are frequently released into the environment at operations sites. There are two ways that such sites may affect the health of nearby residents. The first is a direct threat to human health from the hazardous materials released purposely or inadvertently from the sites. The second is an indirect threat, involving the perception of a hazard that affects both stress levels and property values, which in turn lead to a lower quality of life for people living nearby. Sites where the likelihood of substantial hazardous releases is relatively high, or where environmental degradation has taken place have become a major concern of environmental planners, especially when they are located near densely populated urban neighborhoods. These urban sites may not be evenly distributed spatially, but rather, create an uneven pattern of resident exposure. Therefore, assessment of the environmental effects of such sites is inherently geographic and spatial analysis can be used to reveal the patterns, and assist efforts to understand the processes by which the patterns have evolved.

This paper presents a study analyzing the spatial distribution of potential pollution sources, mortality rate and other socio-economic indicators by census block groups within the city of Cincinnati, Ohio. Geographic Information Systems (GIS) and statistical analysis tools are combined to assess the spatial and temporal variations. The sites included are those included on the Ohio Environmental Protection Agency's (OEPA) Master Sites List (MSL) and the U. S. Environmental Protection Agency's (USEPA) Toxic Release Inventory (TRI). The annual release data from TRI sites were also retrieved for analysis. The results show no correlation of the annual mortality rates with annual amounts of toxic chemical releases into the air from the TRI sites. The block groups closest to MSL sites were found to have significantly higher age-adjusted mortality rates than those farther away, though no significant difference was found etween the census block groups closer to the TRI sites and those farther away. When both the MSL sites and the TRI sites were included in the spatial analysis, the mortality rates for the

census block groups closest to the sites were higher than those farther away, at higher significance levels than was found for just the MSL sites. Furthermore, census block groups closer to the potential pollution sites were found to be poorer, less educated and have a larger proportion of minority residents. Race and median rent were found to be significant predictors of the proximity of a census block group to a hazardous material site, after controlling for other socio-economic factors. After controlling for other socio-economic factors, race and proximity to MSL sites were significant predictors of age-adjusted mortality, though proximity to TRI sites was not.

The study demonstrates while the direct impact from hazardous sites on human health was not observed, the spatial connection can not be ignored. Although the associations do not establish causality, they provide the basis for further study to assess how the propinquity of poor and minority communities to hazardous commercial wastes contributes to their significantly higher mortality rates. Further, an investigation is indicated for the role of institutional factors in siting low-income and minority housing near environmental hazards. The findings here will support better design, implementation and evaluation of policy alternatives to improve the quality of life for urban residents. This study also demonstrates the usefulness of GIS for assessment of potential environmental hazards.

INTRODUCTION

The social and health implications of environmental degradation have drawn increasing attention among scholars and policy makers in environmental, social, and human health studies. Investigations of hazardous waste sites have demonstrated health effects in exposed persons, including low birth weight, cardiac anomalies, headache, fatigue, miscarriages, respiratory problems, neurobehavioral problems and higher cancer rates (Berry & Bove 1997, Geschwind et. al. 1992, Washington 1994, NRC 1991). Clearly, there are reasons to be concerned about the effect of hazardous waste on human health. Although there is widespread agreement that exposure to hazardous waste may be adding to our disease burden in significant, although as yet not always precisely defined, ways (USDHHS 1980) the social and health effects of exposure to environmental contamination have been the subjects of considerable controversy. Sites where substantial amounts of hazardous waste are released, or where environmental degradation has taken place have become a major concern of environmental planners, especially when they are located near densely populated urban neighborhoods. Spatial analysis may be used as a tool to investigate if these sites are evenly distributed spatially or create an uneven pattern of exposure for certain group of people. The spatial pattern, or lack of it, may provide an important step in describing problems and in formulating and testing hypotheses about possible links between environment hazards and quality of life and lead to studies that can help direct action to areas of greatest need.

Beyond the consideration of health effects of hazardous emissions on urban populations generally, the disproportionate exposure of poor and minority populations raises important issues of environmental justice. Environmental justice has been defined as the equitable sharing of the adverse effects of pollution across the spectrum of social, economic and political power (Bryant & Mohai 1992; Bullard 1993; Xia et al. 1997). This implies that public policies and regulations, including the siting of polluting industries or the permitting of toxic releases into the environment, should not disproportionately expose minorities or the poor to environmental hazards (Bullard 1997). Unfortunately, we still lack full knowledge of many of the relationships involved, especially spatial relationships. Studies investigating environmental justice issues have generally concluded that minorities and the poor are likely to have greater exposure to toxic landfills, waste incinerators, hazardous industrial facilities and other environmentally detrimental activities (Bullard 1993, Bryant & Mohai 1992, Buntin 1994; Ortolano 1997; Ringquist 1997; Bullard 1997, Burby 1997, Vos 1997). Other work has found that commercial hazardous waste treatment, storage and disposal facilities (TSDFs) were not more likely to be located in poor and minority communities (Oakes et al. 1996; Anderton et al.1994), though a more recent study came to an opposite conclusion (Boer et al. 1997). Several researchers have raised questions about the methods used in some of the early environmental justice studies. Results were found to be different depending on whether a large (county or zip code) or small (census tract) geographic unit was analyzed (Bowen et al. 1995, Anderton et al.1994). The effective handling of spatial information is essential to facilitate an appropriate public policy response. Planners, public officials and residents concerned with quality of life and environmental justice require analytical tools that enable them to identify and initiate responses to potential threats.

These tools must allow for determining the possibility of a significant health threat and the need for more in depth epidemiological, environmental or land use analyses. The combination of geographic information systems (GIS) and statistical analysis comprises just such a tool to analyze health, environmental and demographic data. GIS allows the construction of maps, identification of nearest neighbors, and display of spatial relationships. A series of patterns, each for a different variable of interest, may be created and combined to reveal correspondences and disparities. GIS functions, such as the storage, retrieval and manipulation of spatially related data also allow the aggregation of data collected from varying sources. This data, in turn, can be used to develop a composite description, and to explore associations, such as the use of distance buffers to identify how the effects of a potential hazard may change as one moves father away (Wartenberg 1993).

It is within this context that this study uses GIS and statistical analysis to describe and analyze the spatial relationships between human health, race, socioeconomic status (SES) and the proximity to potential environmental hazardous sites in Cincinnati. The potential environmental hazards used in this study are sites of reported industrial releases for toxic chemicals or where improper hazardous materials management has resulted in environmental contamination. Mortality rates calculated at the 1990 census block group level were used as the human health indicator. Socioeconomic status was also evaluated at the census block group level. This study describes the spatial distribution of environmental, human health, and socioeconomic characteristics in Cincinnati. The relationship between the mortality rates, SES and proximity to hazardous sites is also addressed.

DATA

The study area was the City of Cincinnati, Ohio (1990 city population 364,040; 1990 consolidated metropolitan statistical area population 1,744,124). Census block groups from the 1990 U.S. Census were used as the unit of analysis. To integrate the data for this study, five types of digital data files were compiled: 1) death records, including location of residence, date and causes of death for all persons who died in Cincinnati from January 1, 1986 to December 31, 1994; 2) locations of potential environmental hazardous sites; 3) 1990 U.S. census block group population and housing and other socioeconomic characteristics for Cincinnati; 4) census block group boundary for Cincinnati; and 5) streets in Cincinnati.

Records of all deaths reported to the Cincinnati Health Department for the years 1979 to 1994 were obtained from the Cincinnati Health Department. The data file contained 35 data items for each of the death records, including last residence of the diseased, cause of death, birth and death dates, and limited socioeconomic information such as race, number of years of education, and state of birth (CHD 1994). Records for Cincinnati residents who may have died outside the city were not included. Also, we excluded the records for non-Cincinnati residents who died in Cincinnati since the study area was limited to the City of Cincinnati.

Two types of environmental contamination and hazardous waste release data were collected: the 1994 Master Sites List (MSL) from the Ohio Environmental Protection Agency's (OEPA) Division of Emergency and Remedial Response (DERR); and the 1992 Toxic Release Inventory (TRI) Annual Report, from OEPA's Division of Air Pollution Control. The MSL is a database of sites in Ohio where there is evidence of, or it is suspected that improper hazardous waste management has resulted in the contamination of air, water or soil, and there is a confirmed or potential threat to human health or the environment. The MSL includes a diversity of sites of varied environmental concerns. In addition to those sites in the USEPA's Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS) prior to 1989, sites were added to the MSL listing by DERR staff based on inter-program referrals, citizen complaints, or DERR's discovery efforts. The DERR updates the MSL annually and sites may be delisted if formal remediation has been completed (DERR 1994). Examples of the MSL sites are chemical companies where spills or improper storage or disposal have taken place, and closed landfills that pose a threat because of poor or antiquated design and hazardous contents. The MSL database records contain a field with the street address for each site.

The TRI report contained annually compiled data on the quantity and location of industrial releases for approximately 300 toxic chemicals and 20 chemical categories (Bowen et al. 1995). Manufacturing firms subjected to Title III, Section 313 of the federal Emergency Planning and Community Right-to-Know Act of 1986 are included in the TRI list. These firms are required to report the location and amount of toxic chemicals released to the air, water, or land. In Ohio, the TRI Program within the Division of Air Pollution Control of the OEPA coordinates the collection, digitizing and distribution of TRI data. The TRI sites include a broad range of industrial facilities, from manufacturing and food processing, to chemical plants. Like the MSL database, the locations of the TRI sites are stored in a street address field.

Census population data were directly extracted from the U.S. Census Bureau's 1990 Summary Tape File (STF) 3A. The census data are considered the most reliable source for population information by geographic area. Data used in this study include population by age group for each census block groups within the city of Cincinnati and population by age group for the state of Ohio. The total numbers of deaths in Ohio by age group were obtained from the Ohio

Department of Health (ODH 1992). The Census block group boundary data were extracted from the First Street geographic data files produced by Wessex Inc. The First Street files were compiled and enhanced from the U. S. Census Bureau's 1992 Topographically Integrated Geographic Encoding and Referencing (TIGER) files. The files contained graphic data (maps) defining census block group boundaries and associated attribute data. There are 417 census block groups in the City of Cincinnati, of which 22 had no residents in 1990. These were eliminated, thus providing a total of 395 census block groups for this study. The street address data used in the study were the 1994 TIGER files. The street files contained street name and address ranges for street sections in Hamilton County.

METHODS

Data files were integrated based on their spatial location. GIS functions, including geocoding, buffering, and overlay analysis, were used to complete the tasks. The location of TRI and MSL sites were identified with GIS geocoding function based on the street addresses in the data file and the TIGER street file. After the distance from census block groups to the nearest hazardous sites were calculated, the block groups were divided into six buffer zones based on an 800-meter (0.5 miles) interval to the nearest MSL and TRI sites, respectively. Then, statistical tools were used to calculate mortality rates for each census block group and identify significant predictors of mortality rates from among proximity to hazardous site variables and selected socioeconomic status (SES) variables. ArcView, a GIS software program (Environmental System Research Institute, Inc., Redlands, CA) was used for spatial analyses. SPSS, a statistical analysis software program (SPSS Inc., Chicago, IL), was used for statistical analysis. To control for the age effects when comparing mortality rates, age-adjusted rates were calculated for each census block group using the direct age-adjustment method (Friedman 1994) based on nine-year average crude mortality rates and population age cohorts for the state of Ohio.

Socioeconomic status indicators were selected to reflect known mortality risk factors, as well as to provide insight to conditions found in each block group. The following indicators were selected: length of residence in unit (percent of households living in unit more than 10 years); median household income; median housing unit value (owner-occupied units); median household rent (renter-occupied units); percent of persons 25 years or older with less than a ninth grade education; and percent of population that is African American. A table of bivariate correlation coefficients was constructed for the relationships between age-adjusted mortality rates, distance of the census block group from a hazardous site and the SES indicators.

Multiple linear regression analysis (ordinary least squares) was used to identify statistically significant predictors for proximity to MSL and TRI sites. Based on the correlation table of SES indicators, four were selected for inclusion in the regression analysis (1) percent of residents in home for more than 20 years, 2) median rent, 3) percent of population African American, and 4) percent of population over age 25 with less than nine years of schooling). To avoid multi-collinearity among the predictor variables, median income and median home value were not included in the regression. Median household income was strongly correlated with median home value and median rent, and moderately correlated with percent African American. The Durbin-Watson statistic was calculated and plots of residuals analyzed to check for autocorrelation and heteroscedasticity. No evidence of these problems was found.

RESULTS AND DISCUSSION

Among the 1979-1994 death records in the Cincinnati Health Department's data file, 96,440 were identified as Hamilton County residents. Of these, 96,235 were geocoded (99.8%) based on the recorded street addresses. After excluding accidental causes of death, this study used records for 31,526 decedents who were Cincinnati residents at the time of their death and died between January 1,1986 and December 31, 1994.

The locations for 75 MSL sites and 131 TRI sites in Hamilton County were geocoded based on the street addresses of the sites. Only 12 sites were found on both the TRI and MSL lists. The MSL sites were scattered within the central part of the city while there were no MSL sites in the northwestern and southeastern portions of the city (Fig. 1). Several MSL sites were located in the southern part of the city, just to the west of the central business district. Most MSL sites were near major highways or close to waterways. The census block groups within 800 meters (approximately 0.5 miles) of the MSL sites, taken as a whole, were almost entirely contiguous, except for a few in a small area in the western part of the city.

The distribution of TRI sites shows a similar pattern to that of the MSL sites (Fig. 2). There were only a few TRI sites in the western portion of the city, mostly located along the narrow corridor in the southwestern corner. Similarly, there were only a few TRI sites in the southeastern portion of the city. The TRI sites were even more concentrated along the Mill Creek and two major highways, I-71 and I-75, than were the MSL sites. It should be noted that a number of TRI sites were found in the two communities (Norwood and Elmwood Place) surrounded by the city of Cincinnati. Those sites were included in the analysis since their impact would not stop at the political boundaries.

The age adjusted mortality rates by census block groups in Cincinnati are displayed in Fig. 3. Mortality rates are higher in the older, less affluent areas in the center and west portions of the city. The newer areas to the east generally have lower mortality rates. Figures 4 and 5 show the percent African American by block group and median household income by block group, respectively. Most block groups with higher percentage of African American population are in the center part of the city. To a large degree, the low median household income block groups correspond to the high percentage of African American population.

Table 1 shows the mean values for age-adjusted mortality and the selected SES characteristics for each of the six MSL zones. The zones further away from MSL sites tend to have lower mortality rates, illiteracy and percent African American and higher income, rent, and home value. Similar pattern can be found for the percent of residents in their home more than 20 years, but to a lesser extent. When the similar comparison is made to the TRI sites, the results are similar to, though less pronounced than, those for the MSL zones (Table 2).

The regression results for predicting the proximity of a census block group to a MSL or TRI site are presented in Table 3. Percent African American, median rent and percent adult illiterate were significant predictors at a 0.05 level or better for proximity to MSL sites. Length of residence was nearly significant (p. =0.0977). This model predicts a block group that has 75 percent African American residents, and citywide average values for the other SES variables will be located 0.85 miles from a MSL site, compared to 1.15 miles for a block group that has 25 percent African American residents. Percent African American, median rent and percent of residents residing for more than 20 years were significant predictors at a 0.05 level or better for proximity to TRI sites. The model predicts a block group that has 75 percent African American residents, and citywide average values for the other SES variables will be located 0.8 miles from a TRI site, compared to 1.05 miles for a block group that has 25 percent African American residents.

The regression results for predicting age-adjusted total mortality and age 35-64-cancer mortality are presented in Table 4. Percent African American, median rent, length of residence and MSL zone were significant predictors at an 0.05 level or better for predicting total mortality. The model predicts a block group located in MSL Zone 1 will have a 24 percent higher mortality rate than average. For predicting age 35-64 cancer mortality, percent African American and median rent were significant predictors at a 0.05 level or better. This model predicts a block group located in MSL Zone 1 has a 26 percent higher mortality rate than average.

The results show that the proximity of a census block group to a MSL site is a significant predictor of age-adjusted mortality rates, even after controlling for socioeconomic risk factors. However, the reasons for this association are not clear. Given that an association has been confirmed, a natural expansion of the research will be to explain the nature of the relationships involved. There may be direct causal relationships, or there may be synergistic effects involving several factors. Also, it is likely there are other confounding factors which are related to both proximity and health status. The effect of social stress on community mortality differentials has been receiving increasing attention, and may be a factor here. Closer analysis of the variation in mortality rates among subgroups may provide additional insight, as may scrutinize other factors such as property values, health care access and utilization, and health-related personal habits.

Environmental justice can be considered to entail two components: distributive justice that relates to the status of spatial relationships, and procedural justice that relates to the processes by which spatial injustices came about. The results here show that elements of distributive injustice are clearly present in Cincinnati. The market economy in the U.S. dictates where people live based on their ability to pay. Consequently, it might be argued that the disparity identified is simply the result of “the invisible hand” of the market. However, there are at least two reasons to suggest that non-market institutional factors may be major contributors to the apparent bias against the poor, less educated and minority. First, several of the largest residential concentrations closest to the hazardous sites are public housing complexes built over the past forty years. Decision-makers may have determined that the nearby industrial sites posed no hazards, or, more likely, simply failed to consider the issue. That the real estate was relatively inexpensive, and separated from existing neighborhoods made these sites politically attractive for situating public housing. In either case, hindsight suggests this decision making was flawed and biased against the current public housing residents. Second, many of the hazardous facilities were built or substantially expanded over the past forty years, often on existing industrial sites that had been in use for a century of more. While the industry may have been there first, the residences may have been built when the quantity and toxicity of the contamination and releases of nearby industry posed a greatly reduced threat. In such a situation, the residences have remained the same, but the hazards posed by the nearby industry have increased substantially. In such cases it would seem that environmental permitting and land-use regulations have failed to adequately protect residents from the evolving hazards of modern industrial processes. To the residents it makes little difference whether their increased exposure to environmental hazards was the result of a bias in permit decisions (Boer et al. 1997) or simply a reflection of the institutional neglect fostered by existing power structures. Both explanations recommend changes in planning processes that would give greater consideration to the hazards resulting from the way in which industrial land is used. For example, performance zoning might be used, based on an enlightened, more comprehensive view of the hazards of industrial use, rather than the current system that assumes all permitted hazardous waste releases pose no hazard to nearby residents.

CONCLUSIONS

After examining the spatial distribution of hazardous waste sites, mortality rate and six socioeconomic indicators at the census block group level in Cincinnati, we found associations between the distance to hazardous waste sites and the mortality rate and selected socioeconomic indicators. The study has demonstrated the need for more in-depth investigations to elucidate the processes by which the offending spatial relationships evolved so that the residential areas closest to potential environmental hazards are poorer, less educated, and minority. An interdisciplinary team from urban and environmental planning, public health, and environmental engineering is needed to address the complex issues involved. Clearly, developing a better understanding of the fundamental nature of the relationships would be of great benefit to planners assisting a wide-range of land use-related decisions.

The locations of hazardous waste sites should be considered in making decisions regarding the appropriate land use for nearby properties, as well as environmental remediation and pollution prevention. Planners and communities must take steps to ensure that hazardous sites are managed in effective ways in order to prevent the exposure of nearby residents to dangerous levels of toxins. Industrial facilities cannot be permitted for hazardous releases, even in small quantities, if the cumulative effect of multiple small releases creates health hazards. Steps must be taken to avoid multiple small hazardous releases that are effectively equivalent to the hazards posed by more seriously contaminated individual sites. Further, while the synergistic effects of likely chemical combinations are not well understood, explicit consideration of the potential for such effects would seem prudent, and should become an integral part of standard land use planning. Also, when resources are allocated for environmental remediation, higher priority generally should be given to sites that individually pose more serious problems, yet the impact of other nearby sites must also be considered.

In this study, we integrated GIS and statistical analyses to examine the spatial distribution of environmental, human health, and socio-economic characteristics in Cincinnati. For this project, the importance of innovation based on a solid scientific foundation cannot be overly stressed. In the current economic and regulatory climate, sound GIS methods are emerging as convincing and cost effective means for analyzing multiple factors in a project like this one. An important issue in environmental research concerns the development and use of analytic tools that can provide rapid, reliable and valid assessment for use in planning, implementation and evaluation of policy alternatives. These tools should give both researchers and policy-makers new insight into the consequences of environmental interventions, and facilitate a clear view of the social, economic and political context for which such interventions are proposed. Just as important as the technological advances is the higher-order systematization of geographical thinking that GIS development initiates.

REFERENCES

Anderton, D, A Anderson, P Rossi, M Oakes, and M Fraser. 1994. "Hazardous waste facilities: Environmental equity issues in metropolitan areas." *Evaluation Review*, 18(2):123-140.

Been, V. 1994. "Unpopular neighbors: Are dumps and landfills sited equitably?" *Resources*, Spring, 16-19.

Berry, M and F Bove. 1997. "Birth Weight Reduction Associated With Residence Near a Hazardous Waste Landfill." *Environmental Health Perspectives*, 105(8):856-861.

Boer, JT, M Pastor, JL Sadd, and LD Snyder. 1997. "Is there environmental racism? The demographics of hazardous waste in Los Angeles County." *Social Science Quarterly*, 78(4):793-810.

Bowen, WM, MJ Salling, KE Haynes and EJ Cyran. 1995. "Toward Environmental Justice: Spatial Equity in Ohio and Cleveland." *Annals of the Association of American Geographers,* 85:641-663.

Bryant, B and P Mohai (eds). 1992. *Race and the incidence of environmental hazards: A time for discourse*. Boulder, CO: Westview Press.

Bullard, R. (Ed), 1993. *Confronting environmental racism: Voices from the grassroots*. Boston: South End Press.

Bullard, R. 1997. *Dismantling Environmental Racism in the Policy Area: The Role of Collaborative Social Research*. College Station, TX: Texas A&M University Press.

Buntin, B. 1994. *Environmental Justice: Issues, Policies, and Solutions.* Washington, DC: Island Brothers.

Burby, R. 1997. "People and pollution: What can government and industry do to foster environmental justice?" Paper prepared for presentation at *the Annual Meeting of the Association of Collegiate Schools of Planning*, Ft. Lauderdale, FL, November 6-9, 1999.

CHD (Cincinnati Health Department). 1994. Unpublished internal documents provided by Joseph Geisland, Data Center Manager, 1525 Elm Street, December 29th.

DERR (Division of Emergency and Remedial Response). 1994. *Master Sites List 1994*. Columbus, OH: Ohio Environmental Protection Agency.

Friedman, G. 1994. *Primer of Epidemiology*. New York: McGraw-Hill, Inc.

Geschwind, SA, JA Stolwijk, M Bracken, E Fitzgerald, A Stark, C Olsen, and J Melius. 1992. "Risk of Congenital Malformations Associated with Proximity to Hazardous waste sites." *American Journal of Epidemiology*, 135(11):1197-1207.

Heiman, MK. 1996. "Race, waste and class: New perspectives on environmental justice." *Antipode* 28(2):111-121.

Lichtveld, MY and BL Johnson. 1993. *Public health implications of hazardous waste sites in the United States, paper presented at Hazardous Wastes and Public Health Approaches: Hazardous Waste Conference.* Washington DC: Agency for Toxic Substances and Disease Registry.

NRC (National Research Council). 1991. *Environmental epidemiology: public health and hazardous wastes*, vol. 1. Washington, D.C.: National Academy Press.

Oakes, JM, DL Anderton, and AB Anderson. 1996. “A longitudinal analysis of environmental equity in communities with hazardous waste facilities.” *Social Science Research*, 25:125-148.

ODH (Division of Vital Statistics, Ohio Department of Health). 1992. *Annual Report: Vital Statistics*. Columbus: Ohio Department of Health.

Ohio EPA (Ohio Environmental Protection Agency). 1997. *Master Sites List*, Columbus, OH: Division of Emergency and Remedial Response.

Ortolano, L. 1997. *Environmental Regulation and Impact Assessment*. New York: John Wiley and Sons.

OTA (Office of Technology Assessment). 1983. *Technologies and Management, Strategies for Hazardous Waste Control.* Washington, D.C.: Government Printing Office.

Penderhughes, R. 1996. “The impact of race on environmental quality: An empirical and theoretical discussion,” *Sociological Perspectives* 39(2):231-248.

Ringquist, EJ. 1997. “Equity and the distribution of environmental risk: The case of TRI facilities.” *Social Science Quarterly*, 78(4):811-829.

USDHHS (U.S. Department of Health and Human Services). 1980. *Health Effects of Toxic Pollutants*. Report prepared for the U.S. Senate by the Surgeon General, Serial No. 96-15. Washington, DC: U.S. Government Printing Office.

USEPA (U.S. Environmental Protection Agency). 1990. "Hazard ranking system", *Federal Register,* 55:51532.

USEPA (U.S. Environmental Protection Agency). 1997. *1995 Toxic Release Inventory Public Data Release*, Washington D.C.: http://www.epa.gov/opptintr/tri/pdr95/drover01.htm, last updated: April 14th.

Vos, J. 1997. “The role of local planners and decision-makers in the occurrence of environmental injustice.” Paper prepared for presentation at the *Annual Meeting of the Association of Collegiate Schools of Planning*, Ft. Lauderdale, FL, November 6-9.

Wartenberg, D. 1993. "Use of geographic information systems for risk screening and epidemiology." paper presented at *Hazardous Wastes and Public Health Approaches: Hazardous Waste Conference*, Washington DC: Agency for Toxic Substances and Disease Registry.

Washington, R. 1994. "Environmental Equity: Dilemmas and Challenges for Public Health and Social Work for the 1990s." *Journal of Health and social Policy*, 62(2):1-17.

Xia, H, BP Carlin, and LA Waller. 1997. "Hierarchical models for mapping Ohio lung cancer rates." *Environmetrics*, 8(2):107-120.

Burby, R., and D. Strong. 1997. "Coping with Chemicals: Blacks, Whites, Planners, and Pollution," *Journal of the American Planning Association* 63(4):469-80.

Szasz, A., and M. Meuser. 1997. "Environmental Inequalities: Literature Review and Proposals for New Directions in Research and Theory," *Current Sociology* 45(3):99-120.

United Church of Christ, Commission for Racial Justice. 1987. *Toxic Waste and Race in the United States*. New York: United Church of Christ, Commission for Racial Justice.

U.S. Environmental Protection Agency, Office of Environmental Justice. 1995. Environmental Justice Strategy: Executive Order 12898. EPA/200-R-95-002. Washington DC: USEPA.

Wartenberg, D. 1993. "Use of Geographic Information Systems for Risk Screening and Epidemiology," *Hazardous Wastes and Public Health Approaches: Hazardous Waste Conference 1993*. Washington DC: Agency for Toxic Substances and Disease Registry.

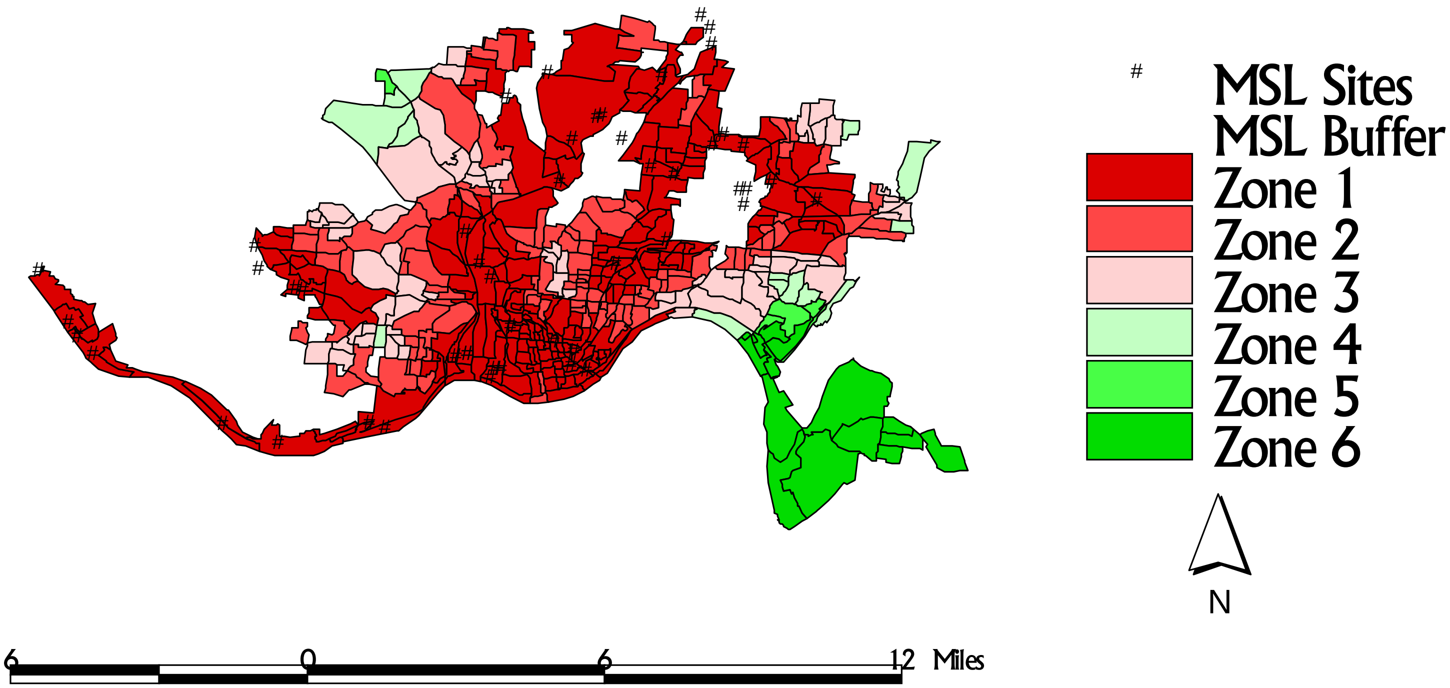


Figure 1. Block Groups by MSL Buffers

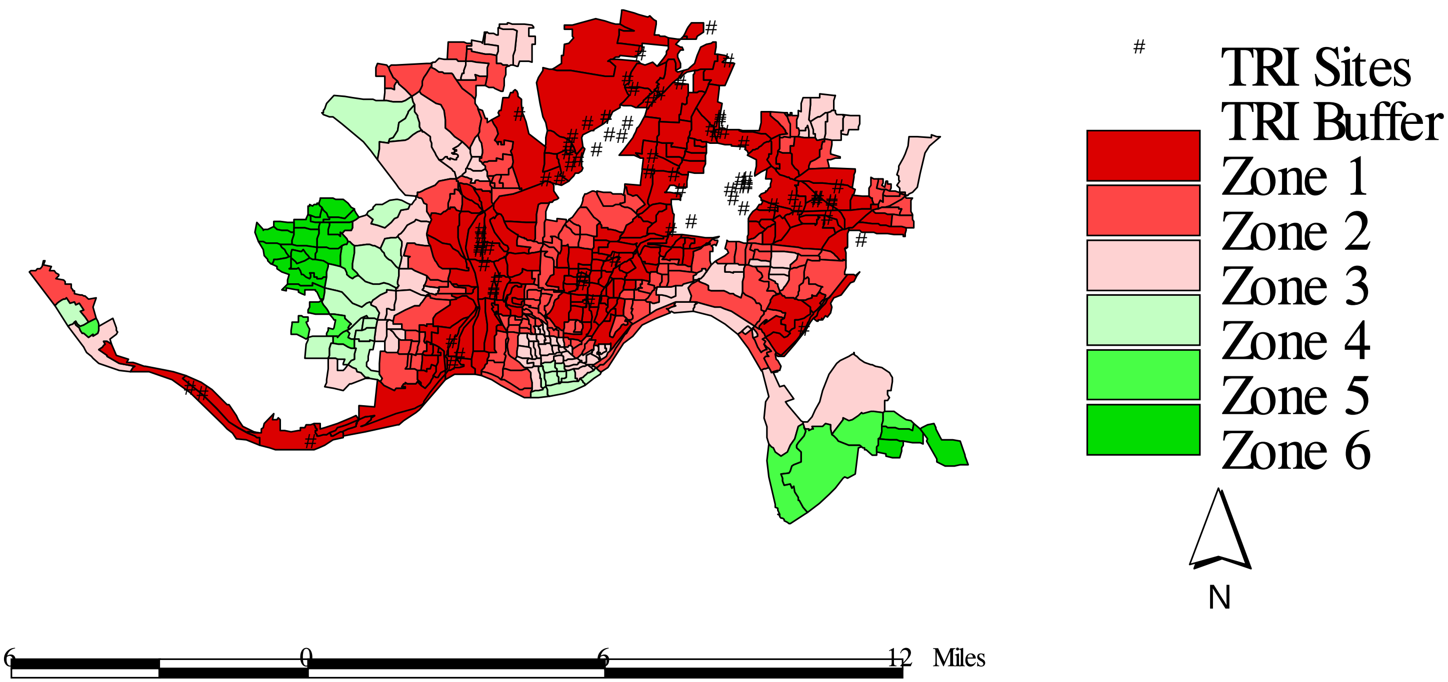


Figure 2. Block Groups by TRI Buffers

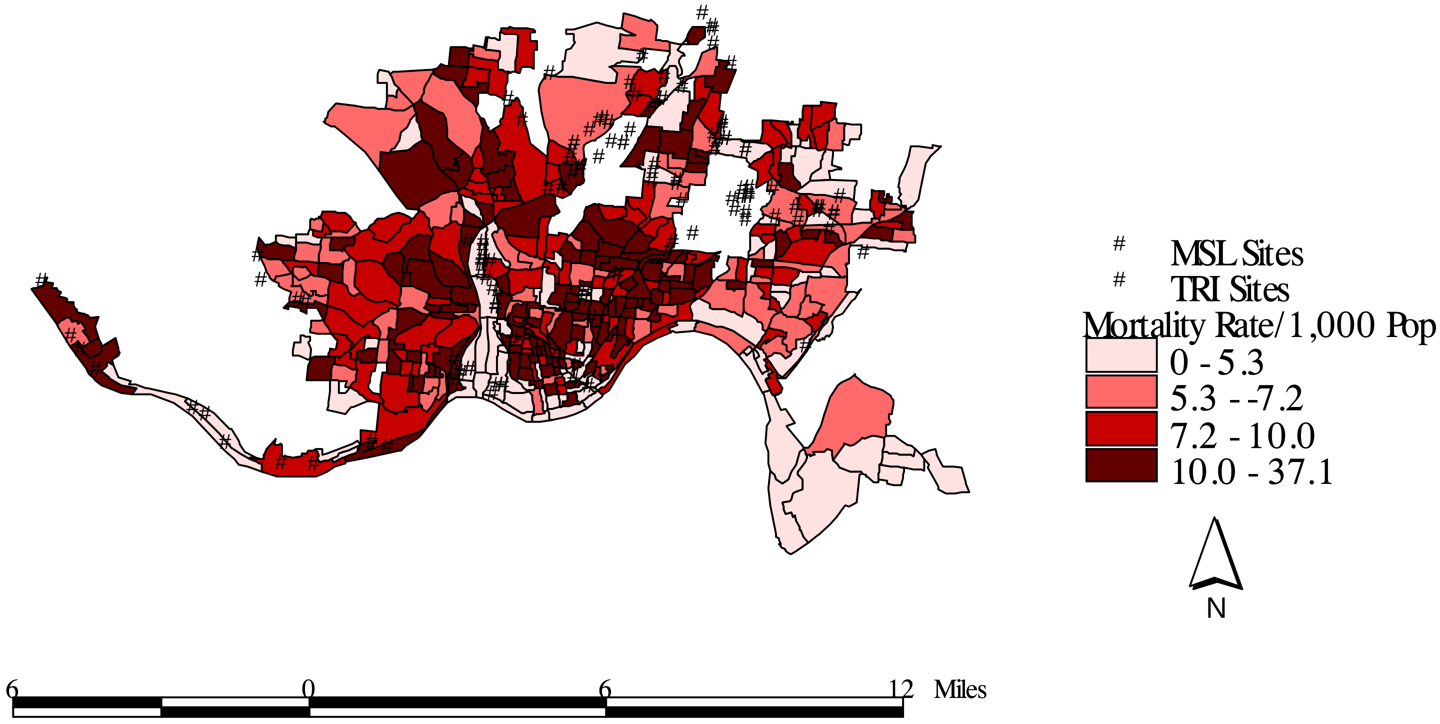


Figure 3 Total Mortality Rate by Block Group

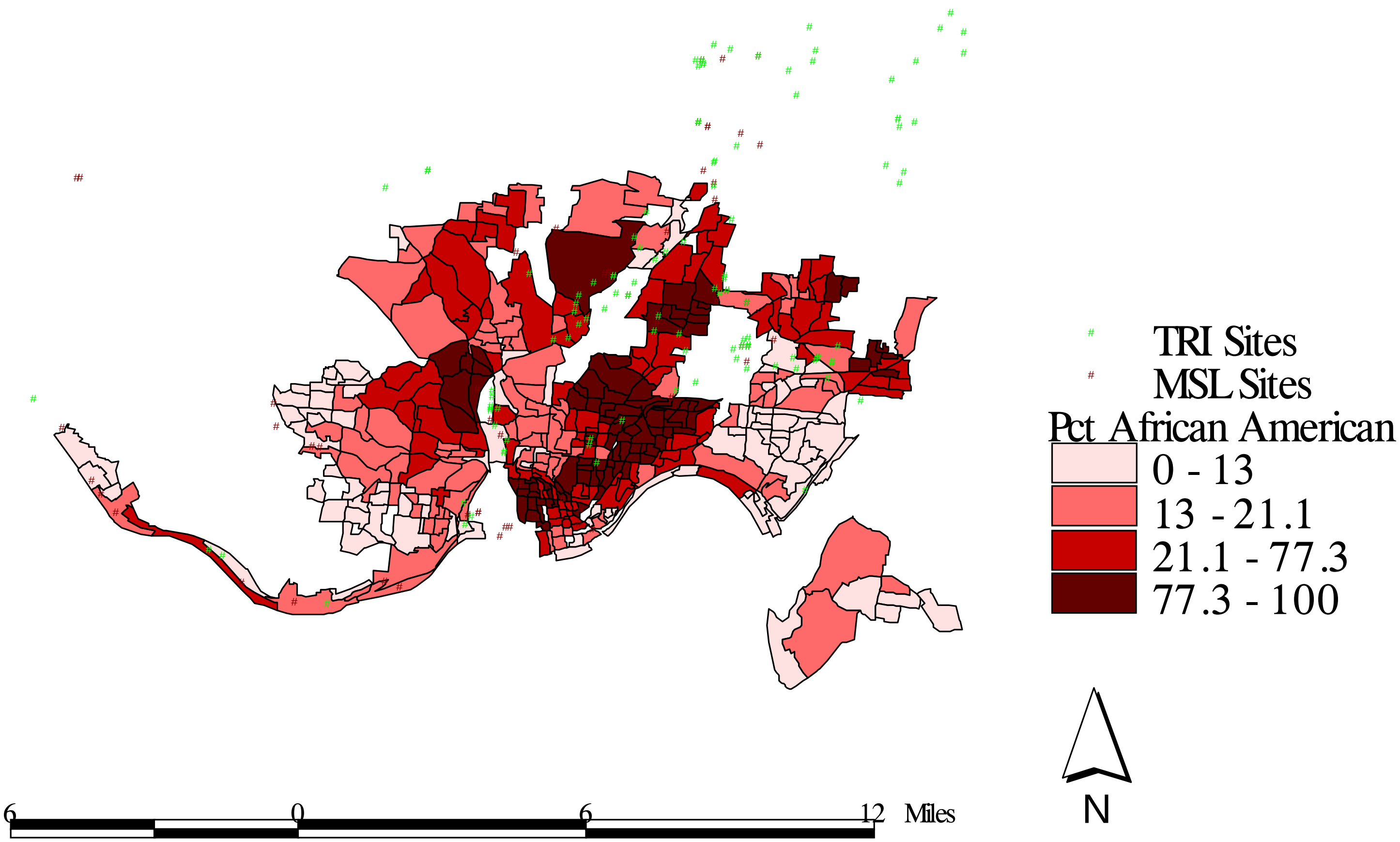


Figure 4. Percent of African American by Block Group

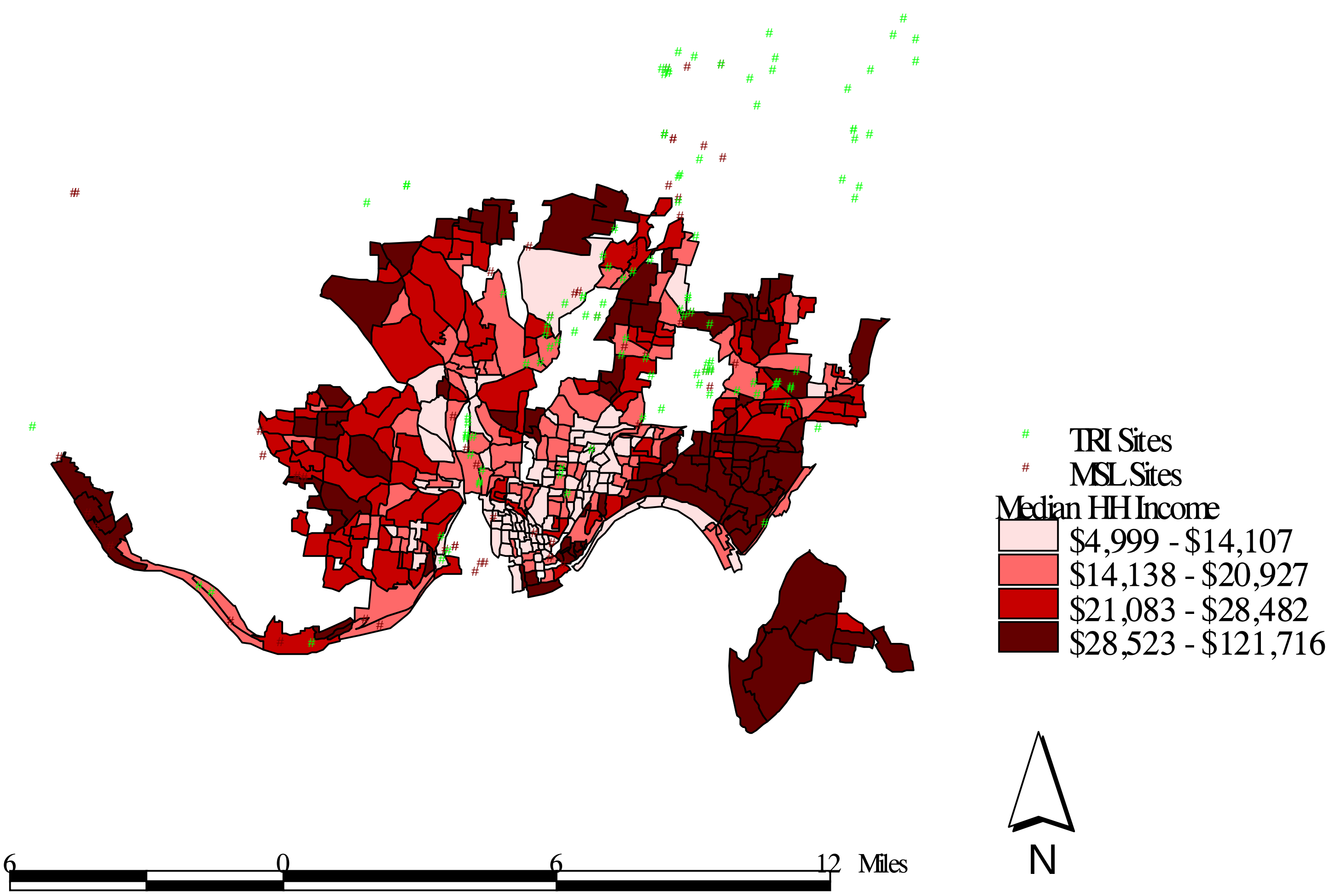


Figure 5. Median Household Income by Block Group

Table 1 Differences in Mortality and Socioeconomic Characteristics by MSL Zone

Variable	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6
Mortality	9.9	9.6	8.9	5.0	6.9	3.4
Income	$18,586	$22,210	$27,667	$34,952	$36,159	$29,639
Rent	$305	$358	$376	$428	$398	$389
Home Value	$59,561	$65,792	$88,895	$87,350	$94,880	$69,599
Illiteracy	14%	11%	9%	8%	11%	10%
20+ Yrs. Residence	16%	17%	17%	24%	16%	22%
African American	47%	39%	20%	17%	1%	2%

Table 2 Differences in Mortality and Socioeconomic Characteristics by TRI Zone

Variable	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6
Mortality	9.6	9.8	9.4	7.3	5.3	6.5
Income	$19,696	$22,199	$24,150	$26,650	$31,408	$27,947
Rent	$330	$352	$323	$389	$367	$389
Home Value	$56,753	$75,097	$92,643	$66,381	$68,222	$68,495
Illiteracy	14%	11%	12%	8%	6%	8%
20+ Yrs. Residence	17%	18%	13%	13%	23%	23%
African American	46%	40%	35%	10%	2%	1%

Table 3 Multiple Linear Regression Predicting Proximity to Hazardous Sites

Independent Variables	Dependent Variable: Proximity to MSL Sites		Dependent Variable: Proximity to TRI Sites	
	Coefficient	p	Coefficient	p
African American	-1.2290	0.0000*	-0.9243	0.0000*
Rent	-0.0014	0.0284*	0.0159	0.0068*
20+ Years Resident	0.9121	0.0977	1.2279	0.0122
Illiteracy	-1.8607	0.0141*	-0.6793	0.3109
Constant	3.1618	0.0000	1.6205	0.0000
Sig F	0.0000		0.0000	
Adj. R^2	0.1247		0.1555	
N	389		389	

* Differs from zero at 0.05 level of significance.

Table 4 Multiple Linear Regression Predicting Mortality

Independent Variables	Dependent Variable: Total Mortality		Dependent Variable: 35-64 Cancer Mortality	
	Coefficient	p	Coefficient	p
African American	0.0028	0.0001*	9.4241E-04	0.0088*
Rent	-7.1930E-06	0.0014*	-4.4588E-06	0.0001*
20+ Years Resident	-0.0078	0.0001*	1.5067E-04	0.8823
MSL Proximity	-5.0456E-04	0.0182*	-1.0440E-04	0.3467
TRI Proximity	-1.9533E-04	0.2985	-9.2487E-05	0.3413
Constant	0.01343	0.0000	0.0042	0.0000
Sig F	0.0000		0.0000	
Adj. R^2	0.1879		0.0942	
Sample Size	389		392	

* Differs from zero at 0.05 level of significance.

Verification of Contaminant Flow Estimation With GIS and Aerial Photography

Thomas M. Williams
Clemson University, Georgetown, South Carolina

Abstract

Estimation of contaminant movement in ground water requires interpolation of data from sampling wells that represent a very small sample of aquifer volume. Spatial statistics and kriging provide the best unbiased estimator of interpolated concentrations. Hurricane Hugo provided an opportunity to compare these estimators with actual forest mortality caused by saltwater inundation associated with the tidal surge. During the 9- to 15- month period after the hurricane, salt from the tidal surge moved within the shallow water table aquifer, causing widespread tree mortality on Hobcaw Forest in eastern Georgetown County, South Carolina. A small watershed (12 acres) was instrumented with 24 multilevel sampling wells. Piezometric potential and samples for salt concentration were collected for 12 months (months 18 to 30 after the tidal surge). These data produced three-dimensional estimations of flow directions and two-dimensional maps of chloride concentration. These maps led to the identification of important heterogeneities in the water table aquifer. Apparently, the infiltrated salt water moved to the bottom of the aquifer (15 feet) and emerged, killing the forest, where aquifer heterogeneity resulted in upward movements of ground water.

Georgetown County implemented a geographic information system (GIS) for tax mapping in 1988 and prepared 1:400-scale orthophotographs of the entire county with true ground accuracy of less than 5 feet. Color infrared aerial photographs were taken from a Cessna 150 platform annually after the hurricane. ERDAS GIS software and the accurate photo base allowed removal of scale irregularities and distortion that resulted from using a small aircraft. Scanned images, using a 10-square-foot pixel, were compared with kriged chloride concentration maps, also using a 10-foot cell size. Grid cells with estimated chloride concentration of more than 500 milligrams per liter also exhibited low reflectance in the infrared-enhanced color band, indicating tree mortality. Here, a small number of sampling wells accurately predicted ground-water movement of a contaminant (NaCl), and GIS and remote sensing verified this movement.

Introduction

Estimating contaminant flow in ground water is difficult because we cannot “see” the aquifer. We know that aquifers comprise sediments that vary from place to place, that changes in hydraulic conductivity determine the rate of water movement, and that the spatial variability of the aquifer sediment determines the hydraulic conductivity. Our inability to accurately represent spatial variability of the aquifer limits our ability to predict ground-water flow and, thereby, contaminant transport.

A large variety of prediction models are available (1-3), and stochastic methods of estimating spatial heterogeneity have been developed (4, 5) and tested (6, 7). On well-characterized field sites, these techniques can produce predictions of tracer movements that accurately predict experimental plumes in terms of mass behavior. Even at these research sites, the spatial distribution of hydraulic conductivity is not known well enough to predict behavior at any particular point.

Ground-water measurements generally derive from wells that are single points. To understand movement of an entire plume, these single point samples must be extended to represent areas. The geostatistical approach allows quantitative estimation of the spatial variation of point estimates (8). Kriging is a technique that uses spatial covariance to estimate values at points where no measurement exists (9). It produces the best linear unbiased estimator of nonmeasured points (8).

Following Hurricane Hugo, these techniques were used to study saltwater movement in the water table aquifer in forested stands of eastern Georgetown County, South Carolina. Clemson University received a grant from the U.S. Forest Service to examine forest mortality and regeneration success within the forest zone covered by salt water during the tidal surge. In this study, we used a small sample of ground water to estimate the direction and concentration of salt moving in the aquifer. Geographic information systems (GIS) proved to be a useful tool to verify conclusions based on the small sample size. Onsite sampling, aerial photography, vector and raster GIS, and spatial statistics were combined into one analysis system. The system estimated and verified directions of salt movement within the aquifer. GIS and remote sensing of forest mortality produced an independent indicator of salt movement that could be compared with the geostatistical technique.

Problem Statement

The main goal of the research project was to evaluate problems for forest regeneration in areas covered by salt water during the hurricane. In many of these areas, the mature trees died during the summer following the hurricane. These areas have very low elevation, little relief, and abundant rainfall, causing the water table to remain near the soil surface. The hypothesis was that salt movement within the aquifer killed the mature trees and could limit regeneration success. We divided the problem into three tasks: to determine if salt concentrations in the aquifer were high in areas where mature trees died, to determine pathways of salt movement within the aquifer that could explain high salt concentrations, and to predict regeneration success from the pattern of salt movement.

GIS contributed both to testing the initial hypothesis and to extending predictions to areas not initially studied. GIS has been used primarily to store and display spatial data in a way that preserves and presents the spatial relationships as well as the data. For this project, we collected two dissimilar data types. To determine salt movement, we measured salt concentrations and piezometric pressures in a series of wells. GIS had to represent the well data and the domain of the kriging procedures in a coordinate system compatible with the mortality data. We determined forest mortality data from infrared-enhanced color aerial photography. GIS also had to allow separation of the infrared signature of the photography, transform the signature into data that were comparable with the well data, and ensure that the coordinate systems of the well data and the mortality data represented the same true ground positions.

To use the GIS ability for this project, we needed to choose several GIS parameters. In this case, we estimated ground-water chloride concentration using kriging, which produced data on a grid comparable with mortality interpreted from photographs. A raster GIS representation could be compared with individual grid chloride values. Each grid cell was 10 square feet so that each cell would be within a single tree crown.

Methods

Study Location

The study was located on 12 acres of a small watershed located on the eastern side of Hobcaw Forest, an experimental forest managed by Clemson University, Department of Forest Resources. Hobcaw Forest is located on the end of a peninsula between the Winyah Bay and the Atlantic Ocean in eastern Georgetown County, South Carolina. The study watershed is located immediately west of the salt marsh and barrier island separating the forest from the Atlantic Ocean and is in Pleistocene-aged beach sediment. Watershed divides were created by former low dune lines, and the stream is within a small depression between these former dune lines. Divides are from 7 to 8 feet above sea level and the stream from 4 to 6 feet above sea level.

The study watershed is 50 miles northeast of Charleston, South Carolina, where the eye of Hurricane Hugo struck the U.S. coastline. Along this portion of the South Carolina coast, the tidal surge was approximately 10 feet above mean sea level (10), covering the entire watershed. After the hurricane, shallow auger holes contained water with sodium concentrations of 4,000 milligrams per liter (11). The hurricane winds did little damage to the watershed forest, but 25 percent of the large oaks were windthrown (12). Beginning in the spring of 1990, however, many hardwoods and pines began dying. By the winter of 1990 and 1991, a large portion of the forest on the watershed had died. Tree mortality did not correspond with high salinity measured by the initial auger-hole method, suggesting movement in the water table aquifer.

Well Installation

The water table aquifer is about 20 feet thick, consisting of fine sand similar to the present beach, with thin beds of shells 10 feet beneath the stream. The bottom of the aquifer is a bed of clay up to 3 feet thick over a leaky artesian aquifer composed of shell and sand. Local rainfall recharges the water table aquifer. Recharge for the lower aquifer is provided by leakage from the water table aquifer beneath the center of the peninsula, about 2 miles west of the watershed, where land elevations are 15 to 25 feet above sea level. Piezometric potential in the lower aquifer is generally a few inches above the water table aquifer, making it only weakly artesian (13).

We installed 24 multilevel ground-water samplers (14) in the water table aquifer. Five samplers were located in regeneration measurement plots (15) placed within the stream. Two samplers, one at each edge of the hardwood wetland, formed a line perpendicular to the stream at the regeneration plot. Two more samplers were located along these lines near the watershed divides on each side of the watershed. The 24 samplers formed five transects across the stream (see Figure 1). Piezometric potential and ground-water chloride concentrations were measured from these samplers from March 31, 1991, through April 1, 1991. Williams (16) provides a complete description of samplers, sampling procedures, and laboratory analysis.

GIS Implementation and Measures

A GIS system for Hobcaw Forest management was developed in 1987 using Environmental Systems Research Institute's PC ARC/INFO software (17). The initial system consisted of forest stand boundaries digitized onto 1:100,000 digital line graphs (DLGs) purchased from the U.S. Geological Survey. These relatively crude maps were combined with stand records and used for management decisions that did not require exact locations of stand boundaries. Later, management of the endangered red-cockaded woodpecker's habitat required that mapped stand lines be closer to true ground locations than the original DLG data scale allowed. A program of ground surveys and aerial photography was conducted in the late 1980s to locate stand boundaries more accurately (18).

In 1988, Georgetown County began a program to convert county tax mapping to computer-based systems. The first step was to acquire survey grade orthophotography. Copies of 1:400-scale orthophotographs, with guaranteed ground accuracy of plus or minus 5 feet,

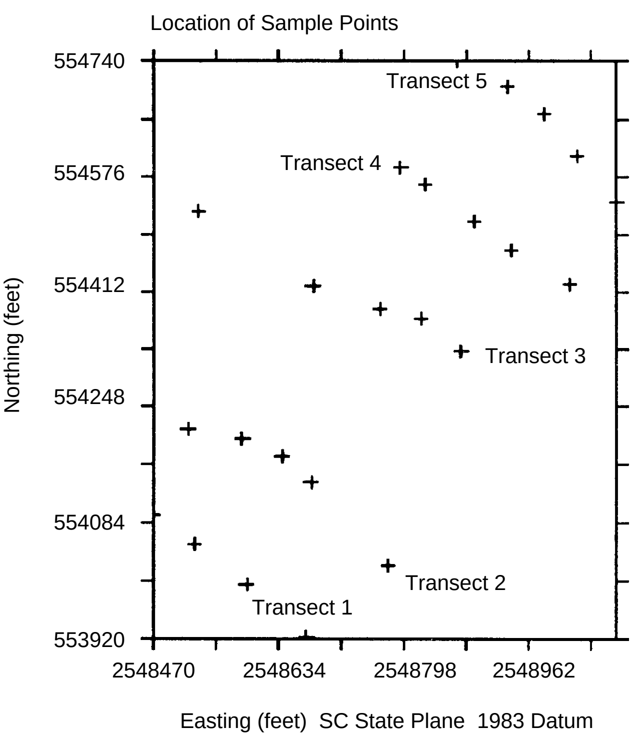


Figure 1. Position of sampling wells in rectangle defined for estimation of salinity movements.

became available in 1990. Roads and stand boundaries were digitized from these photographs into the PC ARC/INFO database. The new, accurate map was combined with stand records of previous coverages to create stand record coverages on a map that was true to the ground within 5 feet, plus or minus 0.3 percent.

In 1991, the GIS programs ERDAS VGA and LIVE LINK were obtained. ERDAS VGA programs allow image processing and raster GIS to be done on personal computers with VGA and some Super VGA monitor adapters. The LIVE LINK program allows display of both ARC/INFO and ERDAS images on the same monitor screen. Orthophotographs were scanned using 5-square-foot pixels and rectified with less than 1-pixel mean error, giving accurate ground locations of plus or minus 8 feet.

Ground surveys from a nearby benchmark provided accurate locations of the sampling wells. PC-TRAVERSE performed coordinate geometry from the survey notes, which was then plotted on the ARC/INFO forest stand database. Accuracy was checked by plotting the recognizable points on the survey with the scanned orthophotograph using the LIVE LINK software.

In February 1991, the Hobcaw Forest was photographed with infrared-enhanced color film. In this film, the red layer is sensitive to near infrared radiation that is strongly reflected by chlorophyll. Red colors in resulting prints indicate living vegetation. The color photography was not corrected for scale variation (from small fluctuation in aircraft altitude) or for distortion (caused by slight variations in the aircraft attitude). One photograph (1:1,320 scale) covering the study watershed was scanned into the ERDAS program. This image was rectified to the 1988 orthophotograph image using control points visible on both. A 10-square-foot pixel was used to sample individual tree crowns. The mean error of rectification was 1.5 pixels for a ground location, plus or minus 15 feet.

Ground-water chloride values at any one point varied over three orders of magnitude in all three dimensions and over two orders of magnitude with time. Annual averages of piezometric potential and chloride concentration, however, yielded interpretable results that were also statistically significant (16). Averaged values could then be combined with the surveyed sampler locations in the ARC/INFO system. The GIS also calculated the corner coordinates for a rectangle that would include all the sampler locations.

Geostatistical calculations were performed using the GS+ software. The data input to this program is an ASCII file of sample point locations in x,y coordinates and data values. The program allows calculation of semivariograms with various combinations of active and maximum lag distance and fitting of various model types to best fit the semivariogram (19). An active lag of 65 feet and a Gausian model produced the best fit:

$$\tau(h) = 0.001 + 1.337\ (1 - \exp\ [-h^2/20736]),\ r^2 = 0.616$$

where $\tau(h)$ is the semivariance at lag distance h.

This best fit model was then used in a block kriging (9) procedure. The procedure used eight nearest neighbors and calculated average values for 10- by 10-square-foot blocks. Values were calculated within the rectangle defined by the corner coordinates from the ARC/INFO procedure described above.

Finally, the rectangle defined around the sampler position formed the region of comparison between the rate of mortality, as sampled by infrared reflection, and estimated average chloride concentrations. The first comparison involved mapping reflection in the red band of the aerial photograph as a gray-scale map and comparing it with the contoured map of chloride concentration. The rectangle coordinates were used in the ERDAS software to create a subset of the scanned aerial photograph that included only the red band in the 5,145 pixels defined by the rectangle surrounding the samplers. In this subset, the infrared reflection was scaled as a gray-scale value between 0 and 255 for each 10- by 10-square-foot block defined in the concentration map. In addition to mapping, a regression of chloride concentration to gray-scale value was performed using the individual blocks.

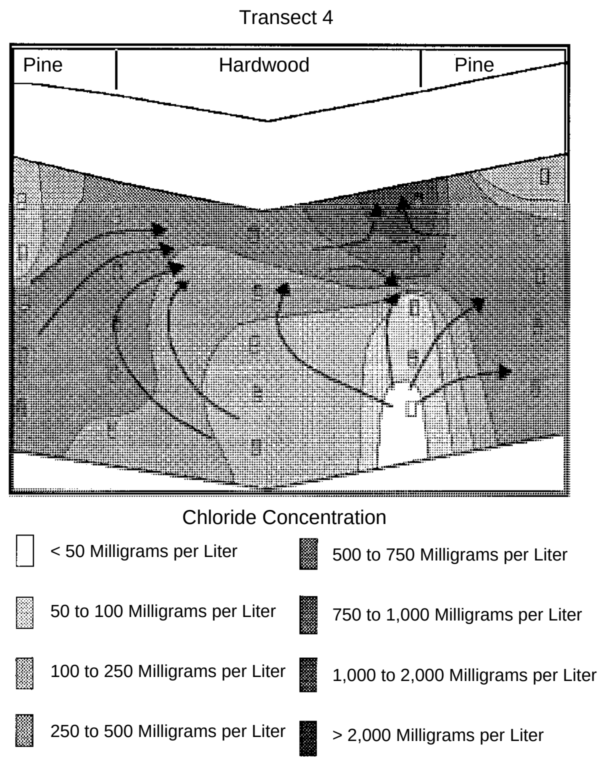


Figure 2. Cross section of aquifer at transect 4; gray scale represents chloride concentration, boxes at sampler positions represent 95-percent confidence limits of average chloride concentration in same gray scale, and arrows are perpendicular to contours of piezometric potential and represent two-dimensional vectors of streamlines.

Results

Ground Water

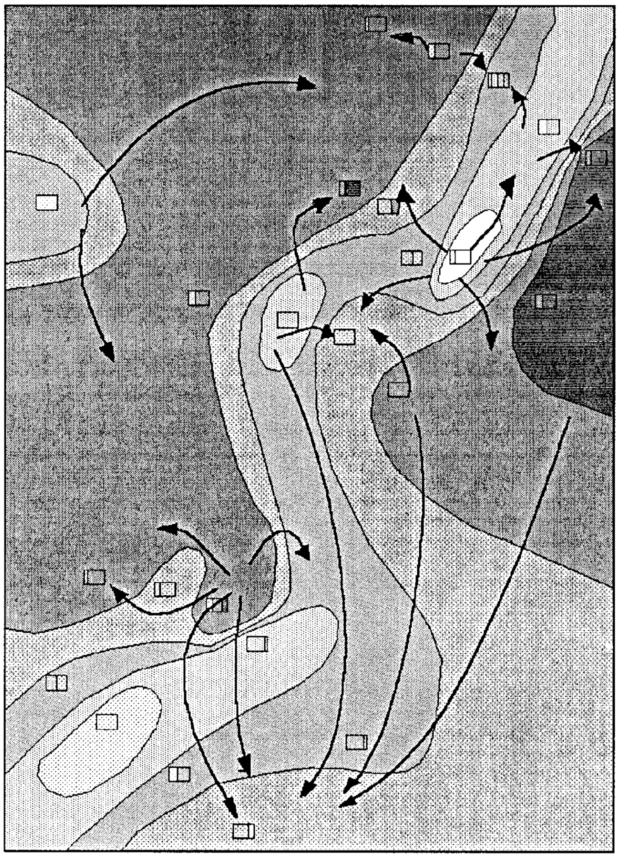
Ground-water chloride data reflected a consistent explanation of salt movement. Initial auger-hole data collected within a month of the hurricane indicated most salt was near the surface of the pine ridges, where, presumably, salt water had filled the aquifer to the soil surface during the hurricane. Data collected 30 months after the hurricane indicated the bulk of the salt had moved to the bottom of the aquifer under the pine ridges. Figures 2 and 3 represent the most significant results interpreted from the piezometric potential and chloride concentration measurements.

Figure 2 represents a cross section of the chloride concentrations and directions of ground-water flow in transect 4, the second most northern transect. The common information represented in this cross section was the west to east movement of ground water, representing the regional flow toward the forest edge. Also, there is an area of upwelling just east of the stream at the bottom of the aquifer, representing a leaky spot in the underlying clay layer. Upwelling causes the west to east streamlines to rise toward the surface along the western edge of the wetland. Chloride concentrations indicate large reservoirs of salt beneath each of the pine ridges and small pockets of fresher water near the surface, probably the result of rain infiltration during the 30 months since the hurricane. The water upwelling from the artesian aquifer was consistently fresh. Flow passing through the concentrated zone beneath the western ridge was pushed to the surface beneath the stream, where evaporation caused chloride concentrations to average more than 1,000 milligrams per liter.

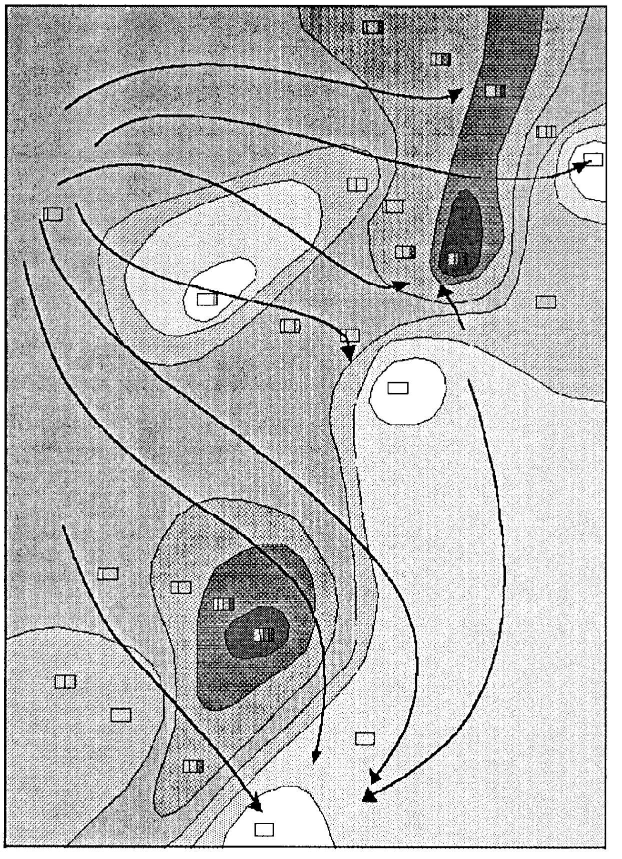
Figure 3 represents two plan views of the site at depths of 4 and 12 feet below the surface. The 12-foot plan view indicated that the east to west flow in transect 4 is only the east vector of a southeast flow. Other areas of upwelling exist beneath the stream. Chloride concentrations are highest beneath both ridges and lowest in the stream center. At the 4-foot depth, the east and southeast flows are also obvious. Also, as deeper flows from the western ridge are turned to the surface, high concentrations of chloride are present near the surface. High concentrations within the wetland result from water being carried to the surface due to upwelling within the wetland.

GIS Evaluations

The ground-water interpretations show a consistent explanation of salt movement. These interpretations are



Plan View 12-Foot Depth



Plan View 4-Foot Depth

Figure 3. Plan views of aquifer at depths of 4 and 12 feet below the surface with same chloride scale and two-dimensional plan vectors of streamlines.

based on only 24 sample points. Interpolation was linear using the nearest neighbor. The samples removed from the aquifer represent only 0.000014 percent of the aquifer volume. Interpretation of a three-dimensional flow regimen from such small sampling does not produce great confidence in the validity of the interpretation.

Salt at the 4-foot depth is most likely to interact with tree roots, and concentrations at this depth were used for kriging. Kriged results, mapped in the same manner as the aerial photography, show general agreement of high mortality and predicted chloride concentrations over 500 milligrams per liter (see Figure 4). A regression of chloride concentration (chloride) to gray-scale value (G) for the individual points yielded a significant negative correlation. The regression line G = 169 - 0.12 (chloride) explained only 27 percent of the variation in gray-scale value, however.

Conclusions

GIS was successfully used to verify interpretations of ground-water flow and salt movements in a shallow water table aquifer. A variety of computer software combined to create a system of analysis that allowed integration of field sampling, aerial photography, vector and raster GIS, and spatial statistics. Using this system, we compared chloride movement, measured by subsurface samplers, with remotely sensed tree mortality caused by soil salinity. The overall pattern of mortality was predicted by a 500-milligram-per-liter chloride contour estimated by kriging averaged concentrations. Estimation of mortality on a single tree basis was less successful, with a regression of chloride to infrared reflection explaining only 27 percent of the variation in reflection. The regression did not fit values of high reflection well but did predict reflection values of 100 or below (regions of high mortality) for concentrations above 500 milligrams per liter.

The most important factor in the success of this project was the availability of large-scale orthophotography. Georgetown County's investment in accurate mapping allowed creation of a map base that made scale correction of less costly aerial photography possible. Without assurance that pixels on the aerial photograph corresponded to the same locations as the subsurface samplers, correlations would have been meaningless.

Another factor that contributed greatly to the research was the fact that most of the computer software exported or imported data from simple ASCII files. The standard (x coordinate, y coordinate, data value) format in ASCII allowed files to be manipulated with spreadsheets

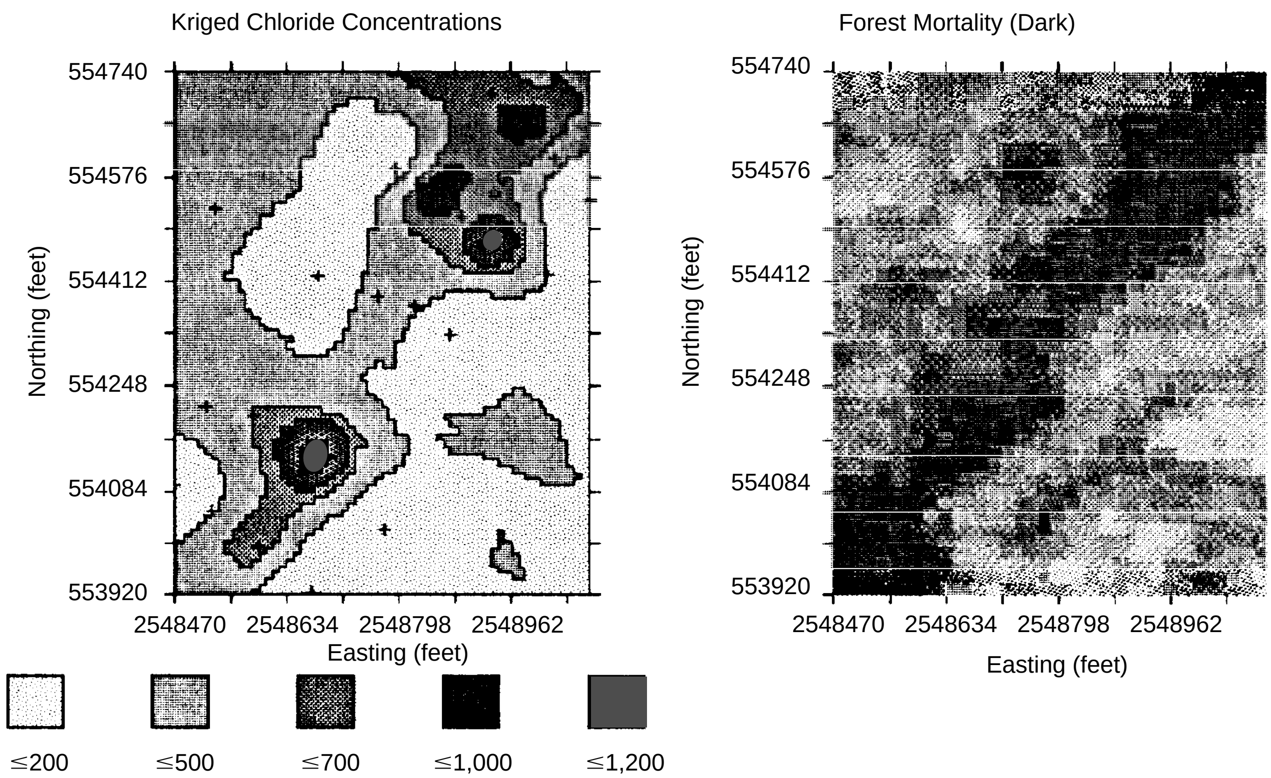


Figure 4. Plan view of chloride concentrations from krig analysis and gray scale of infrared reflection from aerial photograph. Lighter tones represent greater infrared reflection and less forest mortality.

or word processors. Creation of headers, positioning of columns, or changing order of rows or columns could be done for import into the next program. Although more difficult than point and click file transfers of the modern software, the simple standard format creates freedom to use the software in ways not anticipated by the software developers.

Finally, a clear problem statement aided in selecting the most applicable GIS techniques. GIS software allows several methods of data representation. In this example, we chose a raster representation with a cell the size of a tree crown. Criteria for choosing these parameters included physical dimensions of the phenomenon of interest, dimensions of GIS accuracy, and a desire for automated determination of values for individual comparisons. A careful review of the problem to be solved, data available, and capabilities of the GIS software are all necessary ingredients for a useful problem statement.

References

1. Van Genutchen, M.T., and W.J. Alves. 1982. Analytical solutions of the one-dimensional convective-dispersive solute transport equation. U.S. Department of Agriculture Technical Bulletin 1661.

2. Wang, H.F., and M.P. Anderson. 1982. Introduction to groundwater modeling: Finite difference and finite element methods. San Francisco, CA: W.H. Freeman.

3. Bear, J., and A. Verruijt. 1987. Modeling ground-water flow and pollution. Dordrecht, The Netherlands: D. Reidel Publishing Co.

4. Gelhar, L.W. 1993. Stochastic subsurface hydrology. Englewood Cliffs, NJ: Prentice Hall.

5. Dagan, G. 1987. Theory of solute transport by ground water. Ann. Rev. Fluid Mech. 19:83-215.

6. Garabedian, S.P., D.R. LeBlanc, L.W. Gelhar, and M.A. Celia. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 2. Analysis of spatial moments for a nonreactive tracer. Water Resour. Res. 19(1):1,387-1,397.

7. Rehfield, K.R., J.M. Boggs, and L.W. Gelhar. 1992. Field study of dispersion in a heterogeneous aquifer, 3. Geostatistical analysis of hydraulic conductivity. Water Resour. Res. 28(12):3,309-3,324.

8. de Marsily, G. 1986. Quantitative hydrogeology. San Diego, CA: Academic Press.

9. Krige, D.G. 1966. Two-dimensional weighted moving average trend surfaces for ore-evaluation. J. South African Instit. Mining and Metallurgy 66:13-38.

10. Coch, N.K., and M.P. Wolfe. 1991. Effect of Hurricane Hugo storm surge in coastal South Carolina. J. Coastal Res. SI 8:201-208.

11. Gardner, L.R., W.K Michner, E.R. Blood, T.M. Williams, D.J. Lipscomb, and W.H. Jefferson. 1991. Ecological impact of Hurricane Hugo—salinization of a coastal forest. J. Coastal Res. SI 8:301-318.

12. Gresham, C.A., T.M. Williams, and D.J. Lipscomb. 1991. Hurricane Hugo wind damage to southeastern U.S. coastal tree species. Biotropica 23:420-426.

13. Nittrouer, P.L. 1988. Ground-water flow patterns in a forested beach ridge-scale system adjacent to a South Carolina salt marsh. Master's thesis, Department of Geological Sciences, University of South Carolina.

14. Williams, T.M., and J.C. McCarthy. 1991. Field-scale tests of colloid-facilitated transport. In: National Research and Development Conference on the Control of Hazardous Materials, Anaheim, California. Greenbelt, MD: Hazardous Materials Control Research Institute.

15. Conner, W.H. 1993. Artificial regeneration of bald cypress in three South Carolina forested wetland areas after Hurricane Hugo. In: Brissette, J.C., ed. Proceedings of the Seventh Southern Silvicultural Research Conference. U.S. Department of Agriculture Forest Service, General Technical Report SO-93. Southern Forest Experiment Station, Mobile, AL.

16. Williams, T.M. 1993. Salt water movement within the water table aquifer following Hurricane Hugo. In: Brissette, J.C., ed. Proceedings of the Seventh Southern Silvicultural Research Conference. U.S. Department of Agriculture Forest Service, General Technical Report SO-93. Southern Forest Experiment Station, Mobile, AL.

17. Lipscomb D.J., and T.M. Williams. 1988. A low-cost, map-based data retrieval system for forest managers. In: Proceedings of the American Congress of Surveying and Mapping/American Society for Photogrammetry and Remote Sensing, Fall Convention. Falls Church, VA: American Congress of Surveying and Mapping.

18. Lipscomb, D.J., and T.M. Williams. 1990. Developing a GIS for forest management in the 1990s. In: Resource technology 90: Second international symposium on advanced technology in natural resource management. Bethesda, MD: American Society of Photogrammetry and Remote Sensing.

19. Rossi, R.E., D.J. Mulla, A.G. Journel, and E.H. Franz. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecol. Monographs 62:277-314.

Modeling Combined Sewer Overflow (CSO) Impact:
The Use of a Regional GIS in Facilities Planning

Michael Witwer, Jeffrey Amero, Steven Benton,
David Bingham, and Molly Hodgson (Metcalf & Eddy)
Betsy Yingling (NEORSD)

ABSTRACT

This paper presents how a regional geographic information system (GIS) was used to assess the impacts of combined sewer overflows (CSOs) in Cleveland, Ohio during a recent facilities planning project. A hydraulic model was used to evaluate the performance of the sewer system at the present time and with alternative facilities in place. A detailed system and receiving water quality study was also conducted to assess impacts of CSOs and non-point source pollution on the receiving waters during wet weather. Using GIS for integration of the database and model allowed a more accurate and detailed analysis. The GIS interface with the hydraulic and receiving water quality models during facilities planning presented a significant time and cost savings. The quality of the data generated was improved through the incorporation of accepted source information that was repeatable and justifiable. Significant planning prior to the initiation of a GIS-based effort was necessary to ensure that all data needs and quality objectives were sufficient for the project.

Keywords: Modeling, CSO, water quality, and pollution

BACKGROUND

The Northeast Ohio Regional Sewer District (NEORSD) provides wastewater collection and treatment services for the greater-Cleveland metropolitan area through its Westerly, Southerly and Easterly wastewater treatment plants. NEORSD owns and operates the treatment plants and the major interceptor sewers. The remainder of the collection system is owned and maintained by the communities in which it resides. Large portions of each of the three districts consist of combined sewers. The Easterly district, on which this paper will focus, has a total area of approximately 49,300 acres, of which about 20,000 acres are served by combined sewers. The study area is shown in Figure 1.

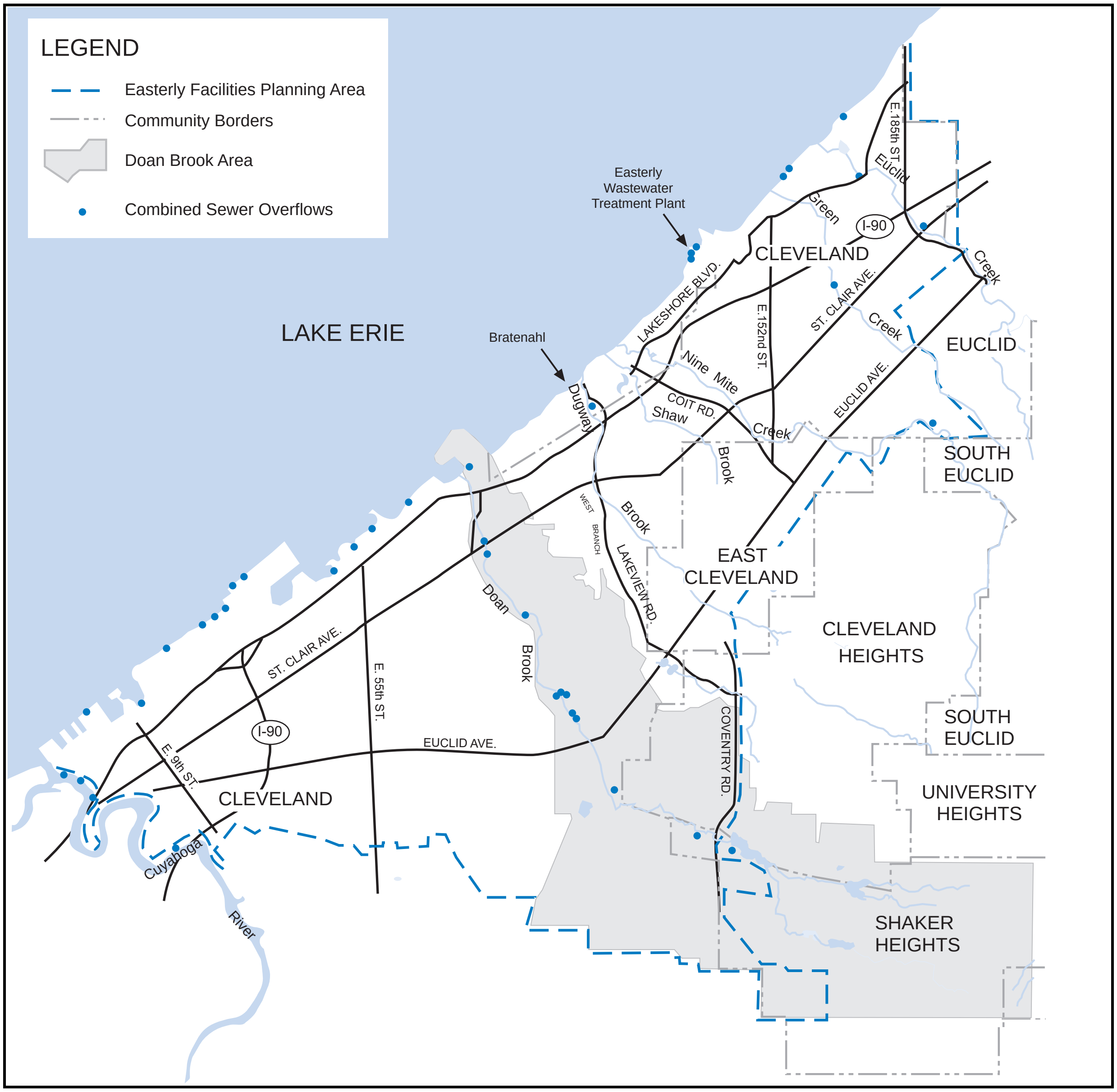


Figure 1. Easterly District Study Area

The Easterly district consists of approximately 2,368,700 linear feet of combined sewers, ranging in size from 6-inches to 13.5-feet in diameter. Of these, approximately 333,700 feet (14%) comprise the interceptor sewers and CSO conduits owned by NEORSD. The 224 combined sewer regulator structures, owned and maintained by NEORSD, provide hydraulic relief and dictate the point at which a CSO event occurs. Facilities planning was initiated to assess the frequency, volume and magnitude of CSOs and their impact on urban streams, the Cuyahoga River and Lake Erie.

The objective of facilities planning activities is to evaluate and recommend alternatives to control CSOs to meet USEPA CSO Policy and OhioEPA CSO Strategy, and to improve the quality of receiving waters after wet-weather events. To begin this process, a hydraulic model of the collection system was created using MOUSE™, produced by the Danish Hydraulic Institute DHI, and calibrated using data obtained from temporary flow monitors and rainfall gauges. Sampling of CSO, storm water runoff and receiving waters was conducted during discrete storm events throughout the temporary flow monitoring period. This was performed to quantify pollutant loadings to the receiving waters and evaluate receiving water response. The use of a GIS allowed the efficient creation of the hydraulic model network and establishment of model input parameters. Having a visual link with the hydraulic model allowed problem areas to be identified easily and alternatives to be evaluated quickly.

AVAILABLE INFORMATION AND DATA

Many sources of information, in both digital and hard copy formats, were available about the physical and operational aspects of the collection system. In conjunction with the facilities planning, NEORSD conducted a complete inspection and evaluation of their facilities within the Easterly service area. This inspection provided spatial and physical condition data, which was used to develop interceptor and CSO conduit portions of the hydraulic model. Record drawings were obtained for sewers owned by individual communities and not inspected by NEORSD.

Spatial data, available from State and Federal agencies, was used to obtain parameters used in the hydraulic model. TIGER files, produced by U.S. Department of Commerce Bureau of Census, were used to define population density within sewer catchments. Land use, used to determine imperviousness of land areas, was obtained from the Ohio Department of Natural Resources (ODNR) and interpreted using 1981 aerial photography. Watershed boundaries for urban streams were also delineated using information available from ODNR.

Extensive information was available from Cuyahoga County, which aided facilities planning efforts. Digital orthophotography (1 inch = 200 feet scale) existed for the entire service area. From these orthophotos, the Cuyahoga County Engineer’s office had created digital planimetric drawings (AutoCAD format) providing topographic contours (1 foot interval), roads, road names, buildings and waterways, which were easily converted to drawing exchange format (.dxf), and used as coverages within the GIS. Additionally, an address referenced parcel coverage was available, which allowed the visual distribution and spatial analysis of flooding complaints to be utilized.

DATA NEEDS AND ACQUISITION METHODS

It was necessary to acquire physical data about the collection system, major storm conveyance routes, receiving waters and water quality to complete the GIS database for planning efforts. The data acquired for the collection system consisted of manhole location surveys, pipe and manhole condition assessment, internal inspection and surveying, CSO and stormwater effluent sampling. This data was complimented by sewer flow and rainfall data collected over a 55-day period. Figure 2 illustrates the data sources and interaction used during the project.

Approximately one-third of the sewers to be included in the collection system model were surveyed and physically inspected. Inspection data was incorporated directly into the GIS by querying the necessary information into consistent data table structures. Each manhole (node) was given a unique numeric attribute. This attribute was used to relate the spatial data defining each node to the endpoints of sewer lines (arcs). Maintaining this strict arc to node topology allowed network connectivity and other quality control checks to be executed directly from ARC/INFO.

Incorporating the remainder of the sewer lines and manholes into the GIS presented a unique challenge. These sewers, not owned by NEORSD, consisted primarily of major trunk sewers greater than 30 inches in diameter and conveyed flows significantly impacting the hydraulic capacity of the sewers. Therefore, it became necessary to incorporate approximately 582,000 feet of sewer in an efficient and cost-effective manner. The incorporation of these sewers proceeded in a three-step process. First, the spatial location of manholes (model nodes), and the sewer lines connecting them, were created in the GIS by digitizing existing sewer system plans. These plans typically did not contain invert and ground surface elevations. In some instances, it was necessary to verify pipe material and size as well. The second step involved obtaining record drawings from the respective owners. This information served as a redundant quality control check of pipe size and material. Invert and ground surface elevations were also obtained from these record drawings. In isolated instances, neither plan information nor record drawings were available for certain reaches of sewer. In these locations, inspection and survey crews were dispatched to obtain the necessary information.

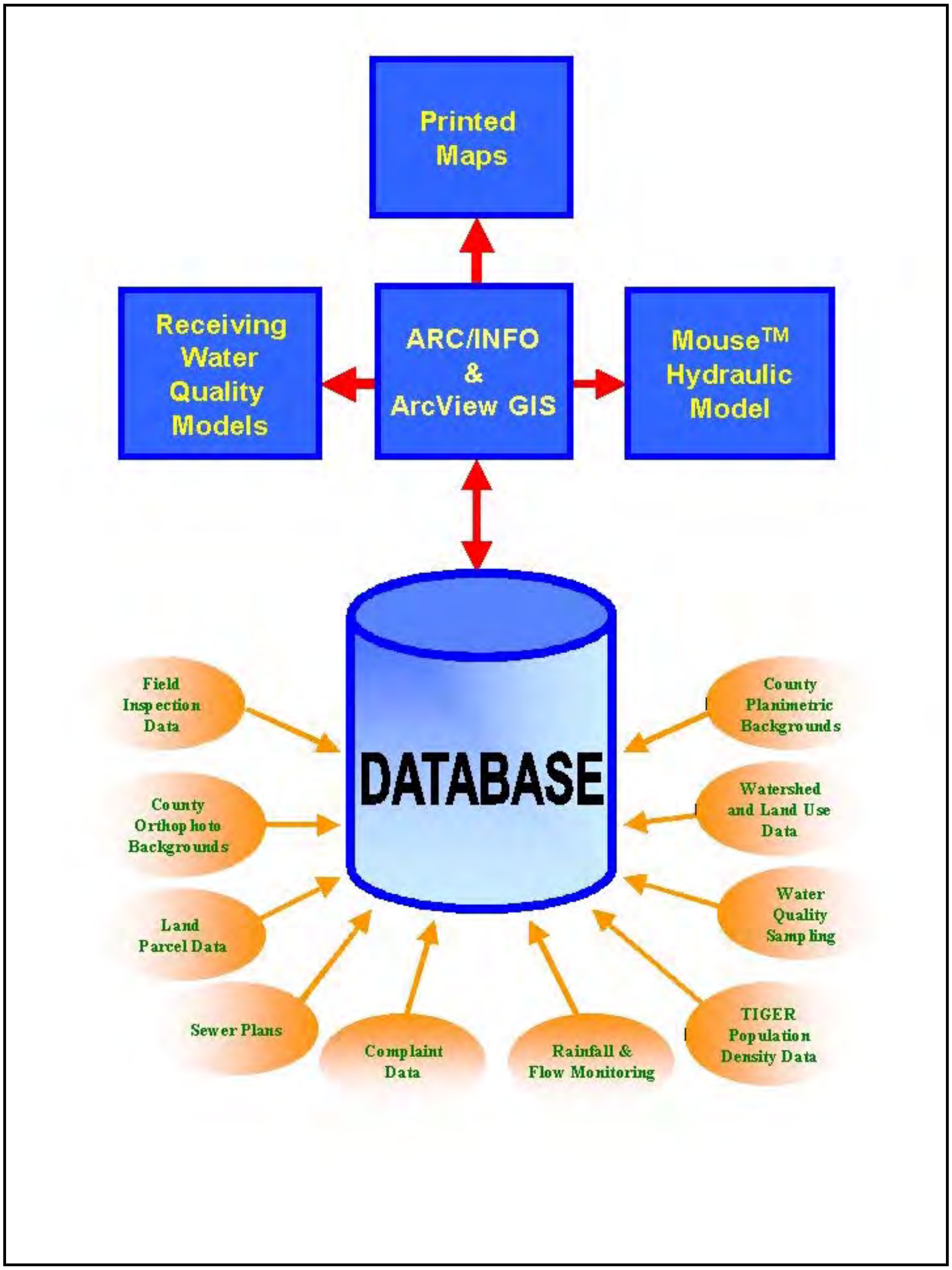


Figure 2. Project Data Sources and Interaction

Flows in the sewer collection system, storm conveyance culverts, receiving streams and the treatment plant headworks were monitored for a 55-day period to evaluate daily dry-weather flow patterns and wet-weather hydraulic responses. Each location was equipped with a monitor, which recorded Doppler velocity, ultrasonic and pressure-based depth at 5-minute intervals.

This allowed the calculation of flow rates and quantities and provided energy and hydraulic grade line data for model calibration. The spatial location, unique identifier and operational characteristics of each monitor were entered in the GIS database for planning use.

Efforts to identify sewer system surcharging were made using community flooding complaint records and address-referenced land parcel GIS coverage data. Using ArcView GIS to view the spatial distribution of the flooding complaints was the first step. This was accomplished by joining a database table containing both the complaint address and type of complaint with the parcel attribute table in ArcView, with the address being the common field. At this point, complaints with the designation “water in basement” were queried out and viewed geographically. Three analysis methods were employed to identify flooding areas, with varying degrees of effectiveness. First, the overall complaint distribution was studied to identify visual trends or trouble spots. This proved inconclusive, as the complaints had a wide spatial distribution. Next, complaints adjacent to NEORSD facilities and very large local facilities were identified and used as quality assurance to verify hydraulic model coverage in such areas. Results from this analysis, although useful, did not provide a wide distribution throughout the combined sewer service area. Therefore, a final spatial analysis of the flooding complaints was performed. The number of complaints per unit area of each sewer catchment basin was calculated using the GIS by overlaying the basin coverage and summing the complaints then dividing by the acreage. This served to normalize the complaint totals by area thus creating a “complaint density.” Basins with the highest density of complaints were evaluated for excessive sewer system surcharging and other operational problems with the hydraulic model.

CSO and storm water discharges to receiving water bodies were also sampled to determine average pollutant concentrations. During wet-weather events, automated samplers were equiped to collect water samples at 15-minute intervals for up to six hours, whenever flow was adequate. Samples were laboratory composited and analyzed for ammonia, BOD, CBOD, hardness, metals, TSS, total phosphorus, fecal coliform and *E. Coli*. Analytical results were incorporated into the database and combined with flow calculations to develop loadings to receiving waters. Field monitoring for dissolved oxygen and pH was performed, which included continuous monitoring in selected streams to observe wet-weather dissolved oxygen sag curve response.

MODELING

Hydraulic modeling was enhanced and accelerated through the use of GIS. The MOUSE^{TM} hydraulic model has an available extension called MOUSE GIS. This extension allows the user to import GIS data directly. Once all model nodes and pipes were loaded into MOUSE, it was necessary to create drainage basins, or catchments. Basins were defined by one of three downstream control points being either a regulator structure, flow monitoring location or major sewer junction. Tracing the collection system upstream from each control point delineated basin areas. Boundary lines were extended outward from the sewer lines to the adjacent basin boundaries or the edge of the service area.

Geometric and spatial calculations were performed with the GIS when basin delineations were complete. The area of each basin was calculated along with the coordinates of the basin centroid. Using the basin centroid coordinates, and the coordinates of each of twenty rainfall monitors, the rainfall in each basin was approximated by weighting the rainfall by the distance from the basin centroid to the monitor. This was accomplished using the inverse-distance method.

Several values for each basin were necessary as beginning values for model calibration. Population and land-use values were required for each basin. Base population values were calculated by overlaying 1990 census data with basin boundaries. Census data GIS coverages are presented as areas of varying population density. A composite basin population was calculated by weighting the percentage of basin area attributed to each density value and multiplying by the area of each density. Land-use, as provided by the ODNR, was used to calculate an average percent imperviousness for the basin. GIS coverages initially identified approximately 58 different types of land use. Land uses were consolidated into 9 broader groups based on similar imperviousness values ranging from 0 (large open flat land areas) to 90 (urban industrial land) percent. Imperviousness values were varied by up to 15% during model calibration to simulate actual rainfall runoff response. Population values were used as a check for sanitary flow rates. Imperviousness values were varied to calibrate storm flow runoff and time to concentration.

GIS was also used to enhance water quality modeling. Watersheds for streams receiving CSO flows, and tributaries receiving separate storm flow, were delineated using watershed boundaries obtained from the ODNR. However, the Easterly service area has four streams that are culverted throughout the entire combined sewer service area. The streams not only receive storm water flows from separate sewer areas, but CSO flows from the combined sewer regulators. Additional separate storm water inputs to the culverted streams within the combined sewer area were defined as individual basins.

Once all hydraulic and hydrologic data were incorporated into the GIS and MOUSE, the model was calibrated to field monitoring data collected during a 55-day period in April and May of 1998. The calibrated model was then used to predict CSO flows and loads for 4-month, 6-month, 1-year, 2-year and 5-year design storms. Flows were calculated for each permitted CSO point directly from model output. Loadings for each point were calculated by multiplying the CSO flow by contaminant event-mean concentration. Contaminant concentrations vary according to the percentage of flows coming from storm and combined sewers. Therefore, event-mean concentrations were weighted based on the source of the flows. The weighted CSO loadings were used as input to water quality models for major receiving streams, the Cuyahoga River and Lake Erie.

Water quality models were developed for Lake Erie, the Cuyahoga River and major creeks and streams that receive discharges from Easterly District CSOs. The objective of receiving water modeling was to provide a tool to evaluate potential improvements in receiving water quality resulting from the reduction of CSOs. Five streams flow through the Easterly District, four of which are contained in culverts for most of their length. All five streams flow to Lake Erie. The Cuyahoga River is the western boundary of the Easterly District, from which it receives a small number of CSO discharges. Water quality in Lake Erie was modeled using a modification of the Lake Erie Information System (LEIFS) hydrodynamic model, which had been developed by Ohio State University. The LEIFS model was modified to provide a finer resolution in near-shore areas, which will permit a more detailed evaluation of the flow-splitting effect of the breakwall that surrounds downtown Cleveland at the mouth of the Cuyahoga River. Euclid Creek, which is located at the eastern end of the Easterly District, is not culverted throughout its length. Euclid Creek receives CSO and other wastewater discharges from the Easterly District and numerous upstream sources. Euclid Creek, along with three of the culverted streams, were modeled using the transport block of the USEPA Storm Water Management Model (SWMM) to route flows to Lake Erie, and to account for the timing of pollutant loads to the lake. The fourth culverted stream, Nine Mile Creek, is culverted for most of its length, but has a significant open-water section at the mouth near Lake Erie. Therefore, continuous modeling of both bacteria and dissolved oxygen was performed for Nine Mile Creek using SWMM for transport, and USEPA Water Quality Analysis Simulation Program (WASP5) to simulate contaminant fate. The Cuyahoga River was modeled using the Cuyahoga River Navigation Channel Model previously updated for work on the Westerly District CSO Planning Study conducted by NEORSD. In each instance, the GIS was used to provide spatial data and physical characteristics used as input to the respective models and represented a central data warehouse structure of acceptable quality-controlled data.

GIS data obtained from ODNR was used to define the stream drainage basins, which were further refined by overlaying two-foot elevation contours and stormwater pipe information obtained from county planimetric maps and collection system mapping, respectively. Future facilities planning and alternatives were readily visualized using GIS. The MOUSE GIS module enabled users to visualize areas where CSO and sewer system surcharging readily occurred. Alternatives including infiltration/inflow (I/I) reduction, in-system storage, off-line storage, conveyance tunnels, flow routing and rapid treatment were considered for CSO reduction. GIS allowed new facility siting to be expedited and visualized quickly.

The GIS database was also a key factor in the automated mapping of the sewer system for the Easterly area. Approximately 140 1-inch = 200-feet scale sheets of the system, were generated using ARC/INFO software. The planimetric coverages were used along with the sewer system database to provide a complete set of sewer maps for the client. Several routines were used to verify flow direction, pipe type and size, and also for assigning unique manhole identifiers. These map sheets will be used by NEORSD for planning, system interpretation, and field maintenance crews, by serving as a reference point for the GIS. As the database is updated, new maps will automatically reflect sewer system changes in that area.

DATA QUALITY AND REPEATABLE RESULTS

Data quality assurance was a primary goal in the development of the GIS database. This goal was considered achieved when data sources were documented, accepted and produced repeatable results. The use of a GIS allowed parameters to be calculated by standard routines that did not vary according to user or individual interpretation. This allowed the users to focus efforts on the legitimacy and accuracy of the results, rather than data compilation.

In addition to using documented source information, the GIS was used to check the collection system spatial data before incorporation into the model. Network connectivity, pipe sizes, and material of construction was checked through automated routines. Using the ArcNetwork™ module of ARC/INFO™, sewer system tracing was initiated from the treatment plant. Sewer lines (arcs) are given a direction within ARC/INFO. The tracing routine, developed using Arc Macro Language (AML), identified all manholes (nodes) with flow connected to the treatment plant. The result of this routine was to identify sewers not connected to the collection system within the GIS. Incidentally, CSO outfall pipes, by their nature, were also identified by this procedure.

Data gaps existed for both pipe size and material of construction even after review of inspection data, plans and record drawings. In most instances, missing sewer segments were the cause of data gaps. To address this, another AML routine was developed. Tracing was initiated on a reach of sewer to identify the missing segment data. The pipes immediately upstream and downstream of the missing data were examined for the missing parameters. If all sizes and materials were the same, the missing attributes were automatically populated in the database for the segment. Segments with generated data were given a qualifier indicating a lower confidence level than documented record data.

TIME AND COST SAVINGS

The GIS provided a significant time and cost savings in the facilities planning process. These savings did not only include the calculation of model input parameters, but the initial development of the model network. While field inspection and survey data would be desirable for all sewers within the hydraulic model, the physical inspection of urban sewers is expensive. It was desired to incorporate information about these sewers in a manner that had a better benefit to cost ratio. Utilizing the data acquisition and quality assurance / quality control (QA/QC) procedures outlined earlier allowed the creation of a database comparable to complete inspection data.

PLANNING AND LESSONS LEARNED

Significant planning was necessary before beginning a GIS-based modeling and planning effort. This ensured that all data needs and quality objectives were sufficient for the project. All sources of GIS data should be reviewed for accuracy, completeness and acceptability prior to their use. Choosing the right software was also an integral part of the planning process for both

GIS functions and modeling. In the case of the GIS, ARC/INFO was chosen because of its broad capabilities as an analytical tool. Network analysis, which could not be performed in desktop software by cost-effective means, was completed with success. Completed ARC/INFO coverages easily converted to shapefile format, which were native to ArcView GIS software. ArcView shapefiles, which the MOUSE hydraulic model can use as data input, were used by GIS analysts, managers and casual users. The advantages of using ArcView were its user-friendly windows-based interface, ability to access documents, images, tables, spreadsheets and CAD drawings, and its simplicity to query the database.

To expedite model network creation, the collection system layout should be created within the GIS. This allows basin delineation and parameter calculation to be completed once, thus avoiding changes and recalculations. It is very easy to lose sight of the ultimate goal of facilities planning if the use of a GIS becomes a project within itself. A GIS is designed to be a tool to reduce costs, expedite planning and visualize data and alternatives. Using a GIS is a definite enhancement to the assessment and control of wet weather water pollution.

Building a Shared and Integrated GIS to Support Environmental Regulatory Activities in South Carolina

Guang Zhao[1], Zhumei Qian, Jeannie Eidson, Paul Laymon, Hsiu-hua Liao, Rob Devlin, Dakin MacPhail, Steve Dennis, and Derek Graves
South Carolina Department of Health and Environmental Control

Abstract

Environmental managers in South Carolina often encounter difficulty in making scientifically sound decisions involving complex spatial data and analysis without the benefit of GIS technology. To facilitate the decision making process, an agency-wide shared and integrated geographic information system (SIGIS) program was developed and an enterprise SIGIS database was constructed at the Office of Environmental Quality Control (EQC), South Carolina Department of Health and Environmental Control (SCDHEC). The goal of the project was to provide managers and policy makers with decision support systems to applications that enable them to effectively analyze spatial information related to environmental regulation and management. The SIGIS database contains 30 environmental layers and 32 baseline layers which are housed in an ESRI's ARC/INFO, ARCVIEW, and MapObject environments. The majorities of the environmental GIS data layers are state-wide in scope and were collected and updated annually using GPS equipment. Their accuracies are mostly sub-meter or meters which satisfy EPA's 25-meter standard. Enterprise applications such as the SIGIS interface to the environmental facility information system (EFIS) and a Data Dictionary have been developed using the SIGIS databse. The SIGIS/EFIS interfaces are web-based and built using ESRI's MapObject and MapObject Internet Map Server in the Microsoft Visual Basic and NetScape JavaScript environments. Environmental facility information system is a management information system housed in the Oracle environment. The integration of EFIS with SIGIS enables environmental managers to easily query and analyze spatial information for environmental permitting and other regulatory activities. The SIGIS data dictionary was developed to document standards, procedures, and GIS data layers stored in the database. It is updated annually and available through the Internet. While these enterprise applications are

[1] Contact author's current address: Department of Health and Environmental Control, Environmental Quality Control Administration, Information Technology Section, 2600 Bull Street, Columbia, SC 29201. Phone (803) 898-3653, E-mail: zhaog@columb20.dhec.state.sc.us

being completed, departmental GISs still function to perform front-end data analysis and develop project-specific GIS applications.

Introduction

Environmental and public health managers in the South Carolina Department of Health and Environmental Control (SCDHEC) often encounter difficulty in making scientifically sound decisions when complex spatial problems are involved. Since early 1990s, GIS technology has been gradually introduced into the agency. In the past nine years, there have been three major ways in which GIS systems have been developed: project GIS, departmental GIS, and enterprise GIS. In the project stage, GIS is used to support an individual project's needs. For example, the community-based environmental project (CBEP) utilized GIS to help delineate polluted sites and their neighboring demographics. Once the project is finished, GIS support ends. In the departmental approach, GIS is used within a single department to support the department's business functions. For example, the Bureau of Water GIS Lab has generated numerous maps and reports to support the bureau's major permitting and regulatory activities. The growing project and departmental GISs, however, resulted in an inefficient sharing of limited resources, distribution of data, management of the integrity of the database, and service to GIS customers. Moreover, at the agency level there was no consistent GIS support to the agency's major business functions. GIS activities have not been in alignment with the strategic directions of the agency. It is preferable to establish a generic, enterprise-wide GIS infrastructure including hardware, software, network, database, and applications to in order to meet the needs for multiple application developments at both the department and the agency levels.

In March 1996, the Research and Planning Division (RPD) of the the Office of Environmental Quality Control developed and implemented an agency-wide, shared and integrated geographic information system (SIGIS) program. The goal was to build an enterprise SIGIS database and platform to provide internal employees of both the environmental management and public health services with decision support systems that enable them to effectively analyze spatial information in supporting their duties. Under such an umbrella, departmental GISs still function to perform data processing and develop project-specific applications, while enterprise applications are developed at the agency level to support the agency's critical functions. GIS information, meanwhile, is also available to the general public through the SIGIS Internet module.

The enterprise GIS approach has been used in recent years by several other state environmental protection agencies. For example, Florida EPA (http://www.dep.state.fl.us/gis/), West Virginia DEP (http://www.dep.state.wv.us/mapping.html) and New Jersey DEP (http://www.state.nj.us/dep/gis/) have developed such an approach. An enterprise method is an organizational effort to develop and implement a shared and integrated GIS database and applications to meet the needs of various user groups across organizational boundaries (ESRI, 1996; Dueker, K.J. 1998; Peng et al., 1998). It focuses on long-term consistency of GIS support, integration of applications for major business functions, and improvement of an organization's business operations.

This paper examines the general design issues for an enterprise SIGIS database. It describes the application examples at both the enterprise as well as the department levels for a state environmental protection agency.

SIGIS Database Development

The development of an enterprise GIS database follows a systematic and functional approach. That is, the GIS database reflects the functional components of the organization. Generally, the following four major steps are followed during the initial development: identification of agency's major business functions, their logical translation into the GIS counterparts, development of system standards, and development of database standards. Overall, the agency's business functions are the backbone for the database development. The existing organizational structures are only considered as indirect factors.

A functional decomposition process is employed to define functional equivalents, which are directly translated from the current organizational functions. The functional equivalents are then translated into SIGIS database counterparts (GIS data layers and associated applications) by combining with constraints such as users' needs, data density, scale, and locational accuracy (Figure 1).

For example, administratively Environmental Quality Control (EQC) is the entity in SCDHEC responsibile for regulating air quality, water pollution, drinking water quality, solid and hazardous waste, and coastal resources in South Carolina. EQC it is organized into five program areas (bureaus), with each area responsible for a particular set of permitting and regulatory functions. Table 1 categorizes the program area and their major functions. Through the functional decomposition process, 30 environmental thematic layers have been developed. These are the GIS equivalents translated from the EQC's business functions. Table 2 lists these layers and their descriptions. Cultural, imagery, and baseline data layers are also collected, generated and incorporated into the databases with similar data standards to allow spatial display and analysis.

The SIGIS database is then integrated using ESRI's standards for database design. The components and factors below are included and considered during the integration process:

- File system organization
- Naming conventions
- Spatial data automation standards
- Coordinate system and scale
- Thematic layer organization and description
- ARC/INFO database standards
- Standard symbol and legend
- Spatial data management
- Access and security
- Maintenance and updates
- Metadata and Data sharing

After conforming to the same standards, all GIS data stored in the SIGIS database are consolidated to form a common platform to support all users' needs, instead of only specific users or applications. Most problems associated with data redundancy, inconsistency, dissemination, and management are eliminated. Many applications at the departmental as well as the enterprise levels now share a common database using different hardware and software particular to their needs. That is, an open and integrated framework is built independent of any one application.

Enterprise Application Development

a. SIGIS/EFIS Interfaces

Two enterprise GIS applications have been developed since 1996: the SIGIS Intranet and Internet Interfaces with environmental facility information system (EFIS) and SIGIS Database Dictionary.

Currently EQC's Environmental Facility Information System (EFIS) is the only management information system (MIS) that interfaces with the SIGIS database to support EQC's permitting and regulatory functions. EFIS resides in an Oracle environment to record and process environmental regulatory information. The integration of EFIS and SIGIS enables EFIS users to display and analyze spatial information about environmental facilities consistently across EQC's organizational boundaries.

The SIGIS/EFIS interfaces are web-based, built using ESRI's MapObject and MapObject Internet Map Server (IMS) in the Microsoft Visual Basic and NetScape JavaScript environment. The Intranet version of the interfaces is embedded in the EFIS Oracle application in a client/server environment, while the Internet version uses an HTML setting. EFIS clients including internal users and public browsers can use the interface to view a map displaying the locations of permitted facilities associated with facility permit information. The basic functions provided by the interfaces include: displaying multiple layers, panning and zooming, spatial query, address matching, buffering, and identification. Since the interface is designed to be an independent module of the EFIS application, it can be fully applied to other management information systems at the enterprise level. Figure 2 shows the basic design of the interfaces.

b. SIGIS Data Dictionary

The Data Dictionary is another enterprise product developed through the SIGIS program. Its purpose is to provide, at the agency level, both internal and external GIS professionals and database users with the background information, procedures, guidelines, standards, and metadata used in the creation and maintenance of the SIGIS database and the state-wide environmental, public health, and baseline GIS data. Specifically, it describes all data layers included in the database, their intended uses, coverage characteristics, associated files, source contacts, and a detailed documentation of attribute information. It is updated annually and is available through the Internet.

The Data Dictionary serves as a general reference for internal GIS developers in the continued development and use of the SIGIS database. Along with the dictionary, a SIGIS database CD is also available for those internal users who do not have Internet access. An ArcExplorer interface was also developed (and is included on the CD) to allow users to browse the database on their computers without additional software.

Department Application Development

As the enterprise applications may not satisfy all users needs, departmental GISs still have their advantages in processing data analysis and developing project-specific applications. Specifically, all applications developed at the project and department levels share the same enterprise SIGIS database and are integrated into the same framework. Below is an example application. Table 3 lists additional departmental applications using the enterprise SIGIS database.

a. Groundwater 305(b) Report

The goal of the GIS project in the Bureau of Water was to develop a statewide groundwater network of critical sites for monitoring groundwater quality. This network will contain both locational and quantitative data: 1) locations and types of known groundwater contamination, 2) sample locations and aquifer-specific background values for selected dissolved natural chemicals and other water-chemistry characteristics, and 3) locations and water quality for all operating public water supply wells.

The application will link directly with the SIGIS database containing layers of known groundwater contamination sites, public water supply wells, and ambient groundwater network sites. The application facilitates the compilation of summary reports and will be used in well-site selection and permitting, wellhead protection planning, and basin- or area-wide assessments or inventories. It also will help identify sources or potential sources of problems at public water supplies, and will be used in initial and updated assessments of all contamination sites.

Summary

An enterprise-wide Shared and Integrated GIS (SIGIS) has been built to support the environmental regulatory effort at the South Carolina Department of Health and Environmental Control (SCDHEC). Enterprise GIS applications such as the SIGIS interfaces to environmental facility information system (EFIS) and a Data Dictionary have been developed to support the agency's major business functions. Department GIS applications were also developed using the enterprise SIGIS database to support individual projects and departmental GIS needs.

Acknowledgement

Special thanks are due to W. Stephen Vassey, Manager, Application Systems and Mike Rowe, Director, Research and Planning Division, Environmental Quality Control (EQC) Administration for their strong encouragement, direction, and critical comments through the development of the SIGIS program.

References

ESRI. 1996. Managing a GIS: planning and implementation. ESRI, Redland, CA.

Dueker, K.J. and J.A. Butler. 1998. GIS-T enterprise data model with suggested implementation choices. URISA Journal, Vol. 10(1):12-36.

Peng, Z.R., J.N. Groff and K.J. Dueker. 1998. An enterprise GIS database design for agency-wide transit applications. URISA Journal, Vol. 10(2).

Table 1. EQC's Program Areas and Major Functions

EQC Program Areas	Business Functions
Bureau of Air Quality (BAQ)	Regulates air emissions including: Air quality permitting Asbestos
Bureau of Ocean and Coastal Resources Management (BOCRM)	Regulates coastal resources including: Critical area permitting Coastal zone consistency certification
Bureau of Land and Waste Management (BLWM)	Regulates solid and hazardous wastes including: Hazardous waste facility permitting Hazardous waste transporter permitting Infectious waste facility permitting Radioactive waste facility permitting Solid waste landfill permitting Solid waste handling facility permitting Mining and reclamation permitting Certificate to explore for minerals Terminal facility registration Oil and gas exploration, drilling, transportation, and production
Bureau of Water (BW)	Regulates water pollution and drinking water including: Waste water discharge permitting (NPDES) and land application permitting State construction permitting Storm water NPDES permitting Section 401: Water quality certification State dams and reservoirs safety act permitting Navigable waters permitting State storm water management and sediment reduction act permitting Shellfish sanitation - certificates and permits Public water system construction and operation permits Interbasin transfer permitting program Underground injection control permitting Groundwater use permitting Recreational waters construction and operating permitting
Bureau of Environmental Services (BES)	Manages EQC's environmental laboratories Environmental lab certification program

Table 2. Environmental Thematic Layers in SIGIS Database

Thematic Layer (Code)	Description
AGWQS	Ambient Ground Water Quality Stations
AMS	Air Monitoring Stations
ARF	Air Regulated Facilities
BIOSTAT	Biological Monitoring Stations
BTRP	Boat Ramps
CERCLA	Comprehensive, Environmental Response, Compensation and Liability Act of 1980
CUW	Capacity Use Wells
DAMS	Dams
DSWL	Domestic Solid Waste Landfills
ESTAURY	Estuaries
HWTSD	Hazardous Waste Treatment, Storage, and Disposal Facility Sites
ISWI	Industrial Surface Water Intakes
ISWL	Industrial Solid Waste Landfills
KGWCS	Known Ground Water Contamination Sites
MRNA	Marina Locations
MSWL	Municipal Solid Waste Landfills
NPL	National Priority List Sites
NPDES	National Pollutant Discharge Elimination System Permit/Discharge Locations
ORPM	Ocean and Coastal Resources Management Permitted Sites
PORTS	Port Locations
PWSW	Public Water Supply Wells
RSWL	Regulated Solid Waste Landfills
SMS	Shellfish Monitoring Stations
SWI	Surface Water Intakes
TRI	Toxic Chemical Release Inventory Facility Sites
USTS	Underground Storage Tank Sites
WQMS	Water Quality Monitoring Stations
SFCLASS	Shellfish Classification (Standard)
WHPA	Wellhead Protection Area
NERR	National Eastuarine Research Reserves

Table 3. Example Departmental GIS Applications Using the SIGIS Enterprise Database.

Application	Department
208 Water Quality Management Plan	Water
303(d) Priority Ranked Waterbodies	Water
305(b) Report - Groundwater	Water
305(b) Report - Surface/Shellfish	Water
Air Monitoring and Regulated Facility Information System	Air
Charleston Harbor Project	Coastal Resource
Health Resources and Local Economic Development Planning Zones	Health Service
Linking Vital Health Statistics and Census Data Using GIS	Coastal Resources
OCRM Natural Dune Baseline Determination Project	Coastal Resources
OCRM Post-Hurricane Recovery GIS Project	Coastal Resources
Orthophoto Production Project (OPP)	Coastal Resources
Savannah River Site Federal Facility Agreement	Land and Waste
Shellfish Sanitation Program	Water
Spatial Distribution of Tuberculosis in South Carolina	Health Service
Targeting Public Health Outreach for Immunization	Health Service
Vital Health and Census Data Integration System	Biostatistics
Vital Records Birth Certificate Data	Biostatistics
Watershed Water Quality Management Strategy	Water
Wetland 401 Certification Decision Analysis System	Water

Figure 1. A Functional Decomposition Process for Developing GIS Thematic Layers

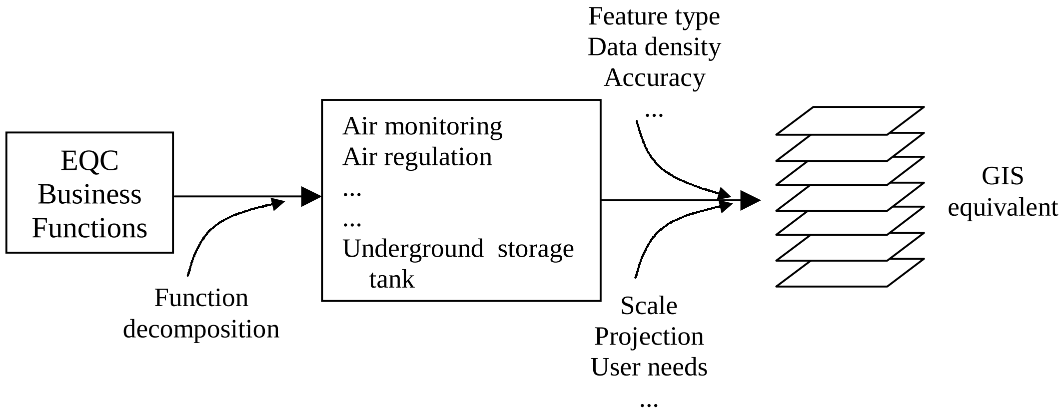
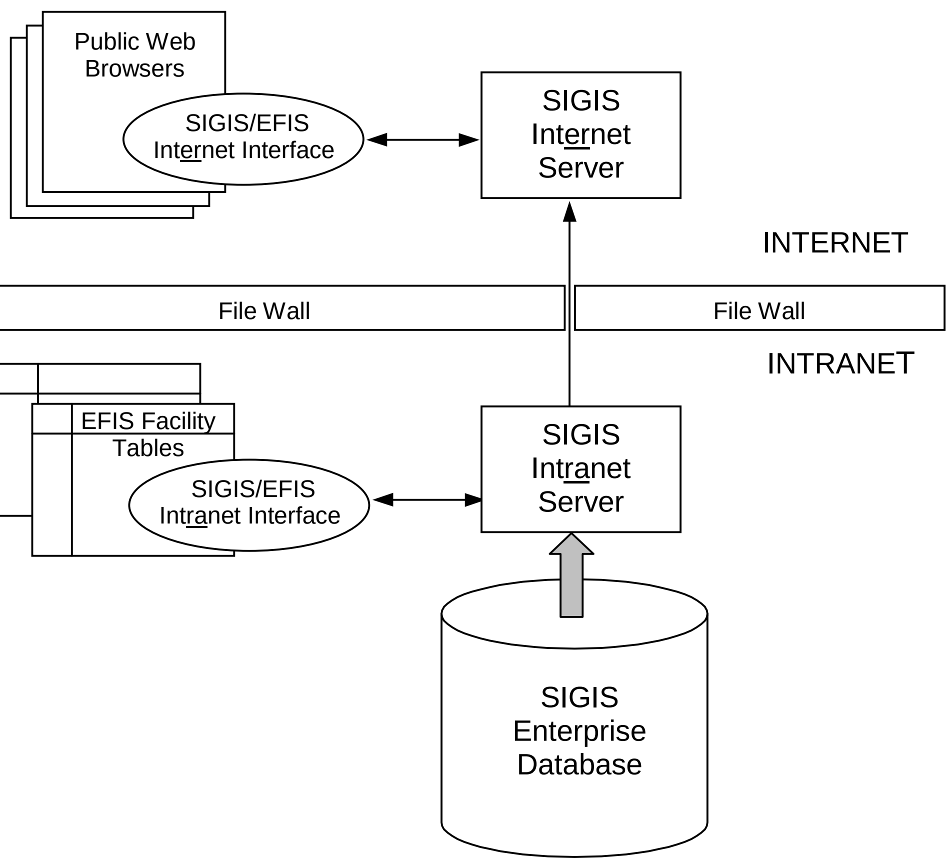


Figure 2. An Enterprise SIGIS/EFIS Interface Design



XGRCWP, a Knowledge- and GIS-Based System for Selection, Evaluation, and Design of Water Quality Control Practices in Agricultural Watersheds

Runxuan Zhao, Michael A. Foster, Paul D. Robillard, and David W. Lehning
Penn State University, University Park, Pennsylvania

Abstract

The Expert GIS Rural Clean Water Program (XGRCWP) integrates a geographic information system (GIS), a relational database, simulation models, and hypertext mark language documents to form an advisory system that selects, evaluates, sites, and designs nonpoint source pollution control systems in agricultural watersheds. Its major features include:

- Customized GIS functions to obtain spatial and attribute data and feed them to a rule-based expert system for selecting feasible control practices.
- A user interface for examining the field-specific conditions and recommended control practices on the screen by clicking on the displayed field boundary map.
- A direct linkage between the GIS spatial data and the relational attribute data, which allows users to examine data on the screen interactively.
- A graphic user interface to GIS functions, which enables users to perform routine watershed analyses.
- Linkage to hypertext reference modules viewable by Mosaic Internet document browser.
- Dynamic access to other models such as the Agricultural Nonpoint Source Simulation Model.

The software environment of XGRCWP is GRASS 4.1 and X-Windows on SUN OS 4.3.1. Its major functions have been tested for the Sycamore Creek watershed in Ingham County, Michigan.

Introduction

In 1981, the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Agriculture (USDA) initiated the Rural Clean Water Program (RCWP) in 21 agricultural watersheds. This program represents the most intensive water quality monitoring and implementation and evaluation of nutrient, sediment, and pesticide reduction practices ever undertaken in the United States (1). More than a decade of research efforts has resulted in a wealth of experiences and lessons on selection, siting, and evaluation of nonpoint source control practices.

The storehouse of knowledge gained from RCWP is of little use, however, unless it is properly integrated and packaged in an easily accessible form. Technology transfer of this knowledge is therefore critically important. To integrate and synthesize the lessons learned from RCWP, Penn State University initiated an RCWP expert project. The hypertext-based version of the RCWP expert system, completed in 1993, can select and evaluate nonpoint source control systems at a single site. Although the hypertext-based version is still suitable for users who do not have access to geographic information systems (GIS) data, it is inadequate for the comprehensive selection and evaluation of control systems on a watershed basis. It does not provide the user the spatial reference of a site and requires the user's subjective judgment for the model input.

The Expert GIS Rural Clean Water Program (XGRCWP) is the UNIX and X-Window version of the RCWP expert system, which integrates GIS and the RCWP expert system to provide decision support at multiple spatial scales from single fields to subwatersheds to the watershed scale. This paper presents the major features of XGRCWP, including design of the expert system, interface to GIS functions, and linkages to a relational database and simulation models.

Overview

XGRCWP comprises five major components (see Figure 1):

- An expert system for recommending control practices based on site-specific information.

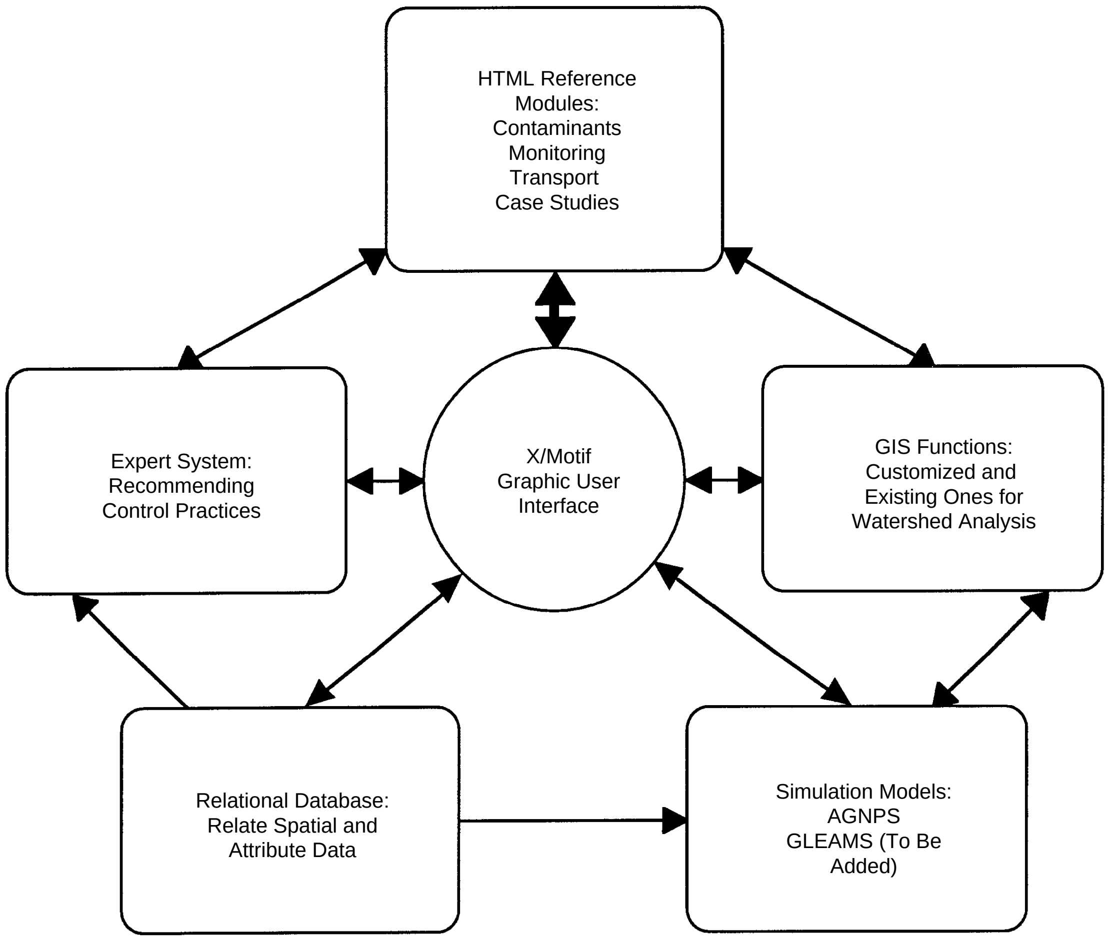


Figure 1. Major components of XGRCWP and their relationships.

- Custom and existing GIS functions for watershed analysis and estimation of contaminant loading potential.
- Linkage to fields, soils, and land use databases.
- Linkage to the Agricultural Nonpoint Source Simulation Model (AGNPS) (2).
- Hypertext mark language (HTML) reference modules.

The X/Motif graphic user interface (GUI) integrates the five components and allows the user to navigate flexibly among them. The components are also internally connected in different ways. For example, the expert system can use the customized GIS functions to retrieve site-specific information from Geographical Resource Analysis Support Systems (GRASS) (3) data layers and INFORMIX relational database tables. In addition, the expert recommendations of control practices can be displayed and examined using GRASS functions. Finally, the GIS functions can help generate input to the AGNPS model, and its output can be converted to GIS format for additional analyses.

Design of the Expert System

The objective of the expert system is to recommend feasible control systems (i.e., complementary sets of control practices to reduce nonpoint source pollution based on site-specific conditions). One distinct feature of this system is the combination of two modes of data acquisition: direct user input and GIS functions. XGRCWP also has two modes for deriving the expert recommendations: batch or interactive. This section discusses these aspects of the expert system as well as its knowledge base.

Rules for Control Practice Selection

The knowledge base of the expert system includes the following six site-specific characteristics:

- Contaminant of interest and its adsorption characteristic.
- Potential level of contaminant loading (low, medium, or high).
- Potential level of contaminant leaching (low, medium, or high).
- Soil hydrologic group (A, B, C, or D).

- Time of year (during or outside the growing season).
- Type of land use (cropland, animal waste, or critical area).

The user first chooses a contaminant of interest from a list consisting of four kinds of pesticides (strongly, moderately, or weakly adsorbed, and nonadsorbed) and eight other contaminants (ammonia, bacteria, sediment, total nitrogen, total phosphorus, nitrate, orthophosphorus, and viruses). The values of other characteristics, some of which vary with the contaminant specified, can then be input either directly by the user or by custom GRASS functions as discussed in the section of this paper on data acquisition.

The RCWP used 14 general categories of control practices (see Table 1). Many suitable conditions were established for each general category. For example, conservation tillage is recommended to reduce runoff for cropland under conditions otherwise favoring loss through sediment transport, such as a contaminant strongly adsorbed to the soil (e.g., total phosphorus), the nongrowing season, and soils with a relatively high runoff potential (soil group C or D) (see Figure 2). Each general category includes several specific control practices. When a general practice category is recommended, the user must decide which specific practice within that general category to evaluate further by consulting the nonpoint source database (NPSDB) for the reported research data about this practice or by running the AGNPS simulation model.

Data Acquisition

The expert system recommends one or more control systems based on site-specific conditions that are either directly input by the user or calculated by the customized GRASS functions. The user always specifies the contaminant of interest and the season, while a GRASS function (R.HYDRO-GRP) always determines the soil hydrologic group of each field. For the other factors (loading potential, leaching potential, and application class), however, the user has two alternative ways to decide input values. For example, after the user selects a contaminant of interest, the program displays the contaminant loading potential window (see Figure 3). The potential level of the selected contaminant can be indicated if the user knows it. Otherwise, the user can let the GRASS functions derive loading potential from existing field data.

Table 1. The Best Management Practices Used in the Rural Clean Water Program

Source control practices	Nutrient Management (NUTR) Pesticide Management (PEST)
Structural control practices	Animal waste systems (AWS) Diversion systems (DIV) Sediment retention and water control (SED) Terrace systems (TERR) Waterway systems (WATW)
Vegetative control practices	Conservation tillage (CT) Critical area treatment (CAT) Cropland protection systems (CPS) Grazing land protection (GLP) Permanent vegetative cover (PVC) Stream protection (SP) Stripcropping (SCR)

The direct input option can also be used to help the user address "what-if" questions. When the user selects the GIS functions to determine the loading potential, XGRCWP makes a series of calls to appropriate customized GRASS functions according to the current contaminant of interest. For example, if the contaminant is total nitrogen, the functions R.MANURE, R.FERT, and R.B.CONCENTRATION are called to estimate total nitrogen from manure, fertilizer, and soil base concentration, respectively. Another GRASS function, R.NP.LOADING, is then called to translate the quantitative measure of loading potential into the qualitative classification (low, medium, or high) as input to the expert systems. These GRASS functions generate the inputs by searching and converting the data from INFORMIX relational data tables that are associated with the GRASS spatially referenced data, such as field boundary and soil map. Table 2 lists the customized GRASS functions developed for data acquisition.

Control System Recommendation

XGRCWP derives the expert recommendations for control systems in two ways: in a batch mode for every field in a watershed and in an interactive mode for a user-specified field.

In batch mode, an existing GRASS function, R.INFER, is used to create a raster data layer for each general practice category of control practice according to a rule-set prepared for that general category. For example, the contents and formats for the conservation tillage practice are documented in Table 3. The raster data layer for representing the conservation tillage recommendations (CT.rec) is generated by running R.INFER with the appropriate rule. The category value of CT.rec is 1 at each point in the data layer where the conservation tillage is recommended, or 0 otherwise. The R.INFER function is similarly called for other general practice categories. Additional GRASS functions can then display or further analyze the resulting map layers. The batch mode provides the user the overall picture with a watershed-wide view of feasible control systems.

In the interactive mode, the field boundary map is displayed and the user can specify any field of interest by clicking the mouse on it. The recommendations and the site-specific conditions of the field are displayed on the right half of the screen. The recommended control practices are also displayed in a popup window for further

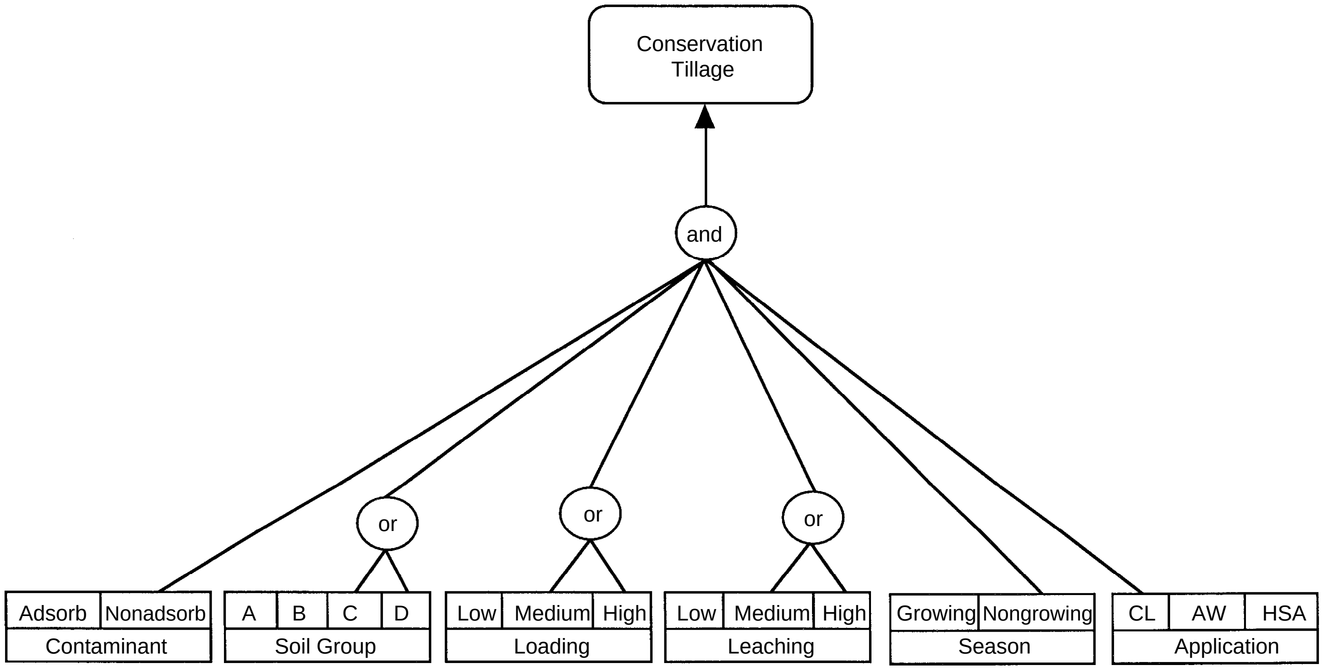


Figure 2. Dependency network (AND-OR diagram) for site-specific recommendation of conservation tillage.

examination, such as the specific practices within each general category, the feasible control systems for nonpoint source pollution control, and research data on the practices. The interactive mode is implemented through the integration of a Bourne shell script, structured query language (SQL) commands, a customized GRASS function (R.RCWP.EXPERT), and GRASS display functions with the Motif GUI. Interactive mode is intended for detailed consideration of a specific farm.

Interface to GIS Functions

XGRCWP provides a GUI to most of the customized GRASS functions and some of GRASS's existing functions (see Figure 4). This interface shields the user from complex syntax so the user can focus on the subject matter. The GUI makes it easier for the user to perform routine operations such as estimation of contaminant loading, identification of critical areas, erosion and runoff calculation, and other watershed analysis tasks. It also helps the user make full, effective use of all custom and some existing GRASS functions.

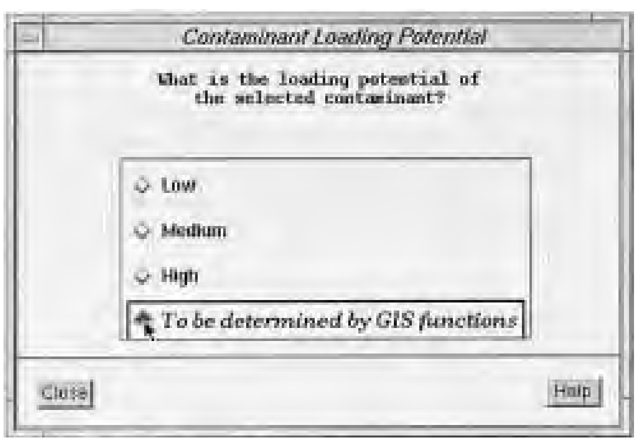


Figure 3. The popup window for the potential level of contaminant loading.

Linkages to Database and Other Models

Data Structure

The GRASS functions used to generate inputs for the expert system use the same soils and fields relational databases as the Water Quality Model/GRASS Interface under development by the Soil Conservation Service (SCS) (4). XGRCWP and our custom GRASS functions were tested for the Sycamore Creek watershed, Ingham County, Michigan. In this data structure, spatial data (e.g., field boundaries, watershed boundaries, soils map unit boundaries, and elevation data) are saved as GRASS raster data layers while attribute data (e.g., crop information, fertilizing schedule, soil information) are stored in INFORMIX relational database tables. Each field or soil map unit is assigned a unique identification (ID) number. The field attribute (INFORMIX) data also contain this ID number. The linkage between the GRASS raster map and the INFORMIX data is accomplished with a GRASS category label (see Figure 5).

Linkage to Database

To allow the interactive examination of field data from GRASS raster layers and the associated relational da-

Table 2. Summary of the Customized GRASS Functions Developed by Nonpoint Source Agricultural Engineering Research Group at Penn State University To Generate Inputs for the RCWP Expert System

Name	Descriptions
R.FERT	Produces raster maps of total nitrogen or total phosphorous from the scheduled fertilizer applications for different crops by dynamically retrieving information from a GRASS data layer and INFORMIX data tables
R.MANURE	Calculates the total manure on each farm according to animal numbers and types (e.g., dairy cow, beef cow, horses, swine), allocates manure to the fields on a farm by a user specified strategy (uniformly spreading or inverse distance weighted distributing method), and finally estimates nitrogen and phosphorous loading from manure application rate, conversion factor, percentages of transportation losses, and volatile losses
R.B.CONCENTRATION	Estimates nitrogen and phosphorous concentration in parts per million within different types of soils according to the organic matter contents
R.NP.LOADING	Classifies the loading potential of nitrogen or phosphorous into three categories (low, medium, and high) based on the actual loading from fertilizer and manure and the N or P concentration in soils
R.EROSION	Obtains a relative measure of soil erosion severity by dividing the amount of erosion by the tolerance values of the soils and then reclassifying them into three categories (low, medium, and high)
R.LEACHING.P	Estimates leaching index from soil hydrologic group and annual and seasonal precipitation and classifies it into three categories (low, medium, and high)
R.HYDRO-GRP	Retrieves soil hydrologic group from the INFORMIX database and reclassifies the soil map into soil hydrologic groups

tabase tables, XGRCWP calls our custom function, D.WHAT.FIELD.SH, a Bourne shell script that dynamically links GRASS raster layers and the INFORMIX database tables. When the user clicks on a field, for example, this function extracts field-specific information from INFORMIX tables such as field information, fertilization schedule, crop operation schedule, and soil information. The D.WHAT.FIELD.SH function then displays all related soils and fields information for the given field. It also marks the field boundary map to remind the user which fields have already been examined.

Linkage to Reference Modules

At any stage of the selection, evaluation, siting, and design procedure for control practices, the user can consult reference modules that provide information, guidance, and data about contaminant properties, transport variables, and examples of applications from RCWP projects. Four reference modules are available in the Macintosh version of RCWP expert system: contaminants, monitoring, transport, and case studies. We are currently converting these reference modules into Mosaic-viewable HTML documents so that they can be accessed from XGRCWP. Mosaic is a public domain, Internet-aware document browser that is available for X-Windows, Macintosh, and Microsoft Windows.

All four modules use graphics to demonstrate design procedures and contaminant control processes. The contaminant module provides information about 11 categories of contaminants cited in RCWP projects and their impacts on surface and ground-water resources. The monitoring module describes different aspects of water quality sampling and analysis systems. The transport module describes contaminant pathways in surface and ground water. The case studies module presents detailed examples from key RCWP projects. These examples cover both practice selection and implementation aspects of control systems. The reference modules serve as a complementary component of XGRCWP.

Linkage to AGNPS

AGNPS is a distributed-parameter, storm event-based model that estimates runoff, sedimentation, and nutrient loss in surface runoff within agricultural watersheds (2). The prototype version of the Water Quality Model/GRASS Interface developed by SCS conveniently generates an AGNPS input file for all cells in a watershed from the spatial and relational soils and fields databases. The UNIX version of AGNPS can then use this input file. XGRCWP can call AGNPS directly from its X-Window interface and convert standard AGNPS model outputs for all cells in the watershed into GRASS raster format for display and analysis.

Discussion

The literature on software systems for managing nonpoint source pollution in agricultural watersheds is diverse and rapidly growing. With few exceptions (5-7), these decision support systems are purely model-based, GIS-based (8), or hybrid systems with models running within a GIS framework (9-14). The addition of expert system components can overcome some of the difficulties in primarily model-based systems:

- Overly intensive input data requirements.
- Inability to handle missing or incomplete data.
- Requirements that all inputs be numerically expressed.

Table 3. The Rule File for Recommending Conservation Tillage

IFNOTMAP app.class 3	!application class is not high-source area
ANDIFMAP contam.feature 2	!contaminant is nonadsorbance
ANDIFMAP leaching.p 1	!leaching potential is low
ANDIFMAP soil.g 1 2	!soil groups are A or D
ANDIFMAP contam.load 1	!contaminant loading is low
ANDIFMAP season 2	!nongrowing season
THENMAPHYP 1 yes, CT	
is recommended	
!	
IFNOTMAP app.class 3	!application class is not high-source area
ANDIFMAP contam.feature 2	!contaminant is nonadsorbance
ANDNOTMAP leaching.p 3	!leaching potential is not high
ANDNOTMAP soil.g 1	!soil groups are not A
ANDNOTMAP contam.load 3	!contaminant loading is not high
ANDIFMAP season 1	!growing season
THENMAPHYP 1 yes, CT	
is recommended	
!	
IFNOTMAP app.class 3	!application class is not high-source area
ANDIFMAP contam.feature 1	!contaminant is strong adsorbance
ANDNOTMAP contam.load 3	!contaminant loading is not high
ANDIFMAP season 1	!growing season
THENMAPHYP 1 yes, CT	
is recommended	
!	
IFNOTMAP app.class 3	!application class is not high-source area
ANDIFMAP contam.feature 1	!contaminant is strong adsorbance
ANDNOTMAP leaching.p 3	!leaching potential is not high
ANDIFMAP soil.g 3 4	!soil groups are C or D
ANDNOTMAP contam.load 3	!contaminant loading is not high
ANDIFMAP season 2	!nongrowing season
THENMAPHYP 1 yes, CT	
is recommended	

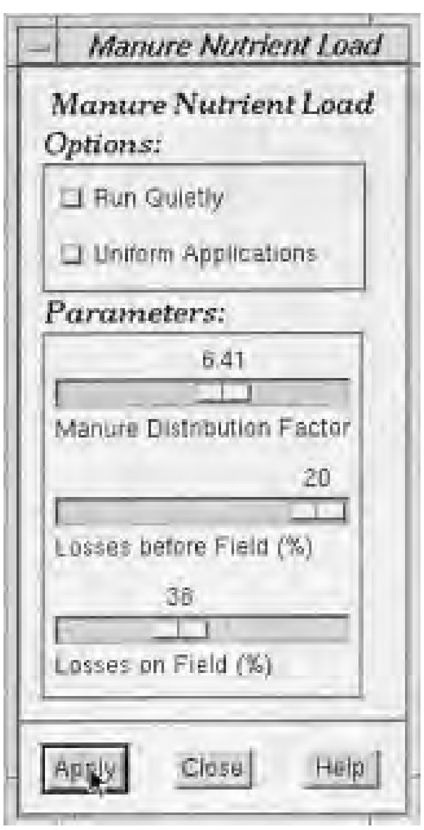


Figure 4. The GUI to the R.MANURE function.

- High degree of expertise needed to structure model input and explain model output relative to the user's problem context.

The expert system component of XGRCWP also reduces the number of model runs needed for decision support through preliminary, rule-based screening of control systems at each site of interest in the watershed.

Conclusions

XGRCWP incorporates several kinds of expertise for the user's benefit:

- Subject matter expertise in siting and selecting nonpoint source control systems in agricultural watersheds.
- Expertise in configuring AGNPS model input from the soils and fields databases.
- Expertise in interpreting, explaining, and visualizing expert system and model input.

The integration of the expert system and the GRASS GIS makes input to the expert system easier and more objective. It enhances the expert system's capability for recommending effective control practices at the field level to achieve watershed contaminant loading objectives. XGRCWP is designed as an open structured

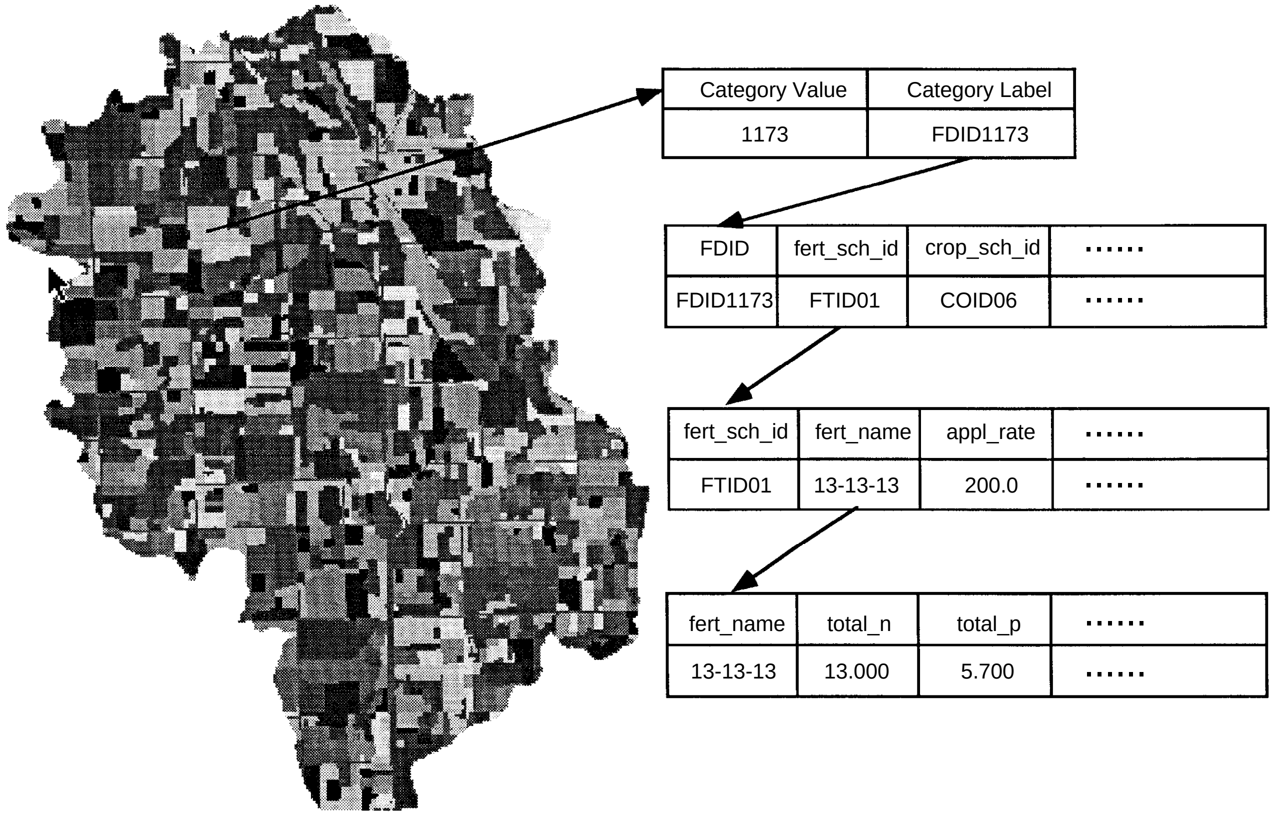


Figure 5. Data structure of Sycamore Creek watershed, Ingham County, Michigan.

program and has great potential to be improved easily and continually according to users' feedback. Ongoing efforts to enhance the program include:

- Developing more rules that incorporate topographical factors such as slope and slope length for the expert system so that more site-specific control practices can be recommended.
- Adding dynamic hypertext-based help and reference to the program.
- Establishing intelligent linkages among the expert system, the GIS functions, and other simulation or design models for nonpoint source control practices.

Acknowledgments

We thank Vicki Anderson, Ruth Shaffer, and Brent Stinson of the Michigan USDA SCS and John Suppnick of the Michigan Department of Natural Resources for their assistance in working with the underlying databases in the Sycamore Creek watershed. This project is supported by U.S. EPA Grant No. X-818240.

References

1. U.S. EPA. 1992. Extending the RCWP knowledge base to future nonpoint source control projects. In: Proceedings of the National RCWP Symposium. EPA/625/R-92/006. Washington, DC.
2. Young, R.A., C.A. Onstad, D.D. Bosch, and W.P. Anderson. 1989. AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. J. Soil and Water Conserv. 44(2):168-173.
3. USACERL. 1993. GRASS 4.1 users reference manual. Champaign, IL: U.S. Army Corps of Engineers Construction Engineering Research Laboratory.
4. USDA. 1993. Users' guide to water quality model/GRASS interface. Fort Collins, CO: U.S. Department of Agriculture, Soil Conservation Service, Technology Information Systems Division.
5. Barnwell, T.O., L.C. Brown, and W. Marek. 1989. Application of expert systems technology in water quality modeling. Wat. Sci. Tech. 21:1,045-1,056.
6. Ford, D.A., A.P. Kruzic, and R.L. Doneker. 1993. Using GLEAMS to Evaluate the Agricultural Waste Application Rule-Based Decision Support (AWARDS) computer program. Wat. Sci. Tech. 28(3-5):625-634.
7. Yakowitz, D.S., J.J. Stone, L.J. Lane, P. Heilman, J. Masterson, J. Abolt, and B. Imam. 1993. A decision support system for evaluating the effects of alternative farm management systems on water and economics. Wat. Sci. Tech. 28(3-5):47-54.

8. Hamlett, J.M., D.A. Miller, R.L. Day, G.A. Peterson, G.M. Baumer, and J. Russo. 1992. Statewide GIS-based ranking of watersheds for agricultural pollution prevention. J. Soil Water Cons. 47(3):399-404.

9. Srinivasan, R., and B.A. Engel. 1994. A spatial decision support system for assessing agricultural nonpoint source pollution. Water Resources Bull. 30(3):441-452.

10. Srinivasan, R., and J.G. Arnold. 1994. Integration of a basin-scale water quality model with GIS. Water Resources Bull. 30(3):453-462.

11. Engel, B.A., R. Srinivasan, J. Arnold, and S.J. Brown. 1993. Nonpoint source (NPS) pollution modeling using models integrated with geographic information systems (GIS). Wat. Sci. Tech. 28(3-5):625-690.

12. He, C., J.F. Riggs, and Y. Kang. 1993. Integration of geographic information systems and a computer model to evaluate impacts of agricultural runoff on water quality. Water Resources Bull. 29(6):891-900.

13. Kiker, G.A., G.M. Cambell, and J. Zhang. 1992. CREAMS-WT linked with GIS to simulate phosphorus loading. ASAE Paper No. 92-9016. St. Joseph, MI: American Society of Agricultural Engineers.

14. Tim, U.S., S. Mostaghimi, and V.O. Shanholtz. 1992. Identification of critical nonpoint pollution source areas using geographic information systems and water quality modeling. Water Resources Bull. 28:877-887.