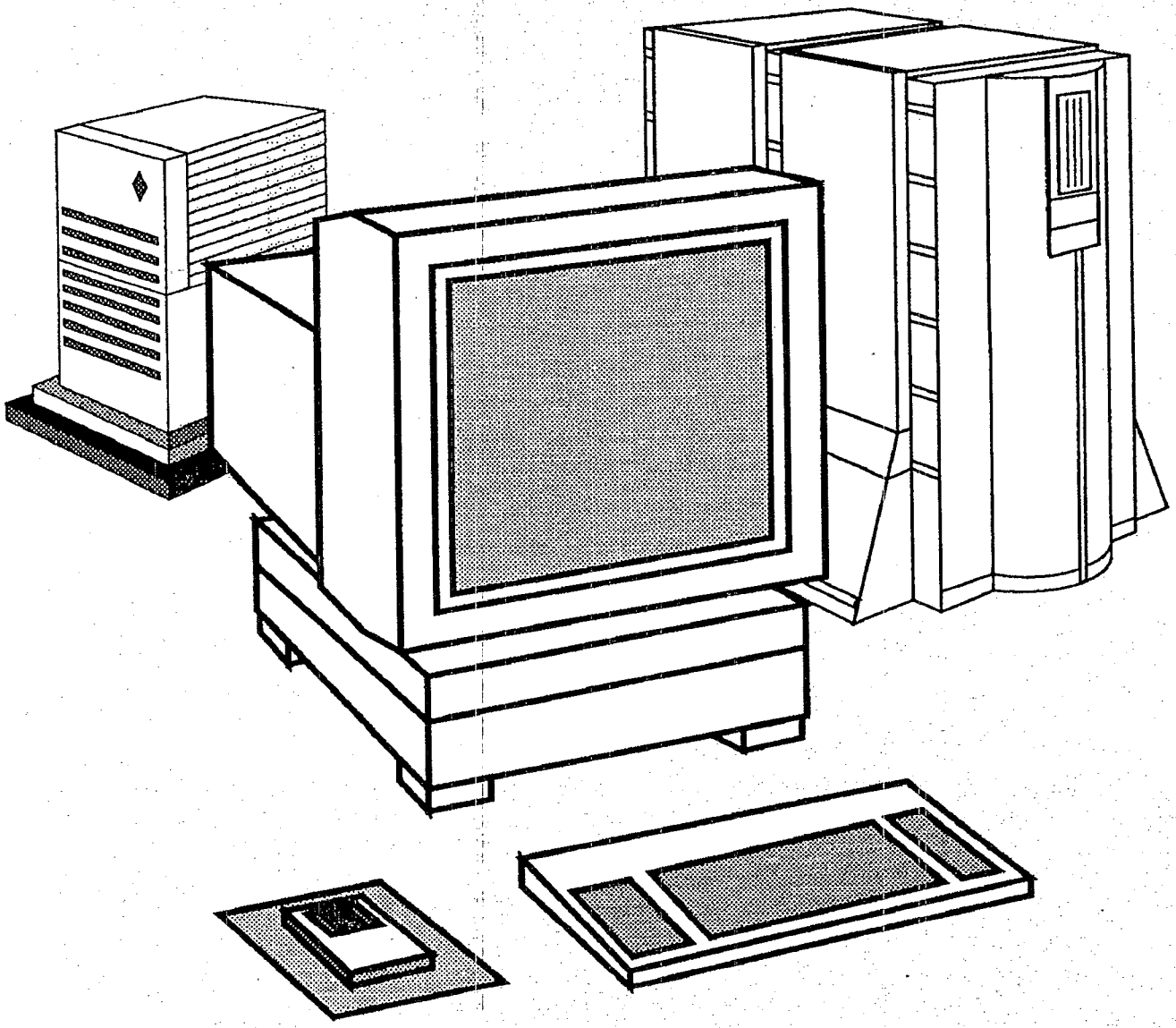
United States
Environmental Protection
Agency

National Environmental
Supercomputing Center (NESC)
135 Washington Avenue
Bay City, MI 48708-5845

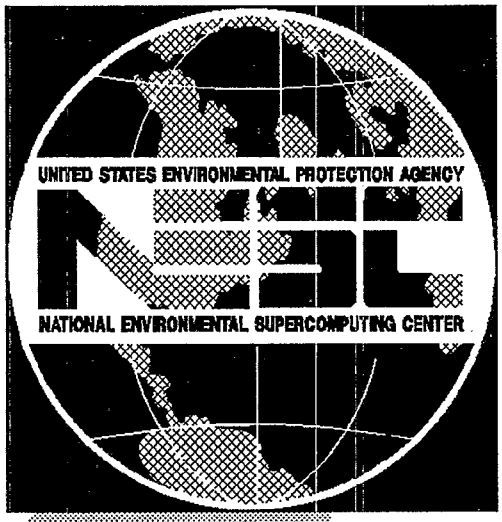
EPA 208-K-93-001
December 31, 1993

EPA

FY 1993 National Environmental Supercomputing Center (NESC) Annual Report



The NESC Mission:



The mission of the NESC is to provide compute-intensive processing for scientific applications that have a high priority within the EPA's overall mission. These applications will come from the EPA researchers, regulatory staff, and support personnel. In addition, the NESC will service those universities, agencies, and private companies external to the EPA having grants, cooperative agreements, and memoranda of understanding with the EPA, in which their science qualifies as compute-intensive and has a high priority within the EPA.

The computational services of the NESC include:

- Management of a wide range of hardware, software, and services into an integrated supercomputing service.
- Centralized compute-intensive processing facility.
- Data communications networks.
- Consultation services relating to the NESC.

A secondary mission of the NESC is to promote environmental modeling and computational science within local schools, by means of academic-year and summer programs.

Contents

Introduction to the NESC FY1993 Annual Report[1]

The National Environmental Supercomputing Center (NESC) is the EPA's latest investment to assure that science of the highest quality is conducted to support environmental protection. The benefits of the NESC to the public and EPA programs center on fostering collaborative efforts among scientists to bring the results of scientific research to the decision making process.

To achieve success at the NESC, four tightly integrated programs are maintained.

- First, operation of a supercomputing resource to provide the maximum amount of computing time to researchers is the backbone of the NESC mission.
- Second, a strong computational science support effort for Agency scientists is essential to ensure the efficient use of resources and the improvement of mathematical models.
- Third, an aggressive and innovative program in visualization and presentation of scientific information is directed toward scientists, environmental decision making officials, and the public.
- Fourth, collaborative efforts among all groups are strongly supported through a nationwide telecommunications network, workshops and seminars, and educational programs such as EarthVision: EPA's Grand Challenge for High Schools.

In its first year of operation, the NESC has become the central resource for carrying out the research programs that are vital to large-scale studies of ecological and biological systems. Its continued success in supporting these efforts depends upon the collaboration among scientists, managers, and the public.

The NESC remains dedicated to providing not only supercomputing resources, but also to providing the collegial environment necessary for that collaboration.

---

[1] Walter M. Shackelford, Director of Scientific Computing, U.S. EPA, National Data Processing Division (NDPD), RTP, NC.

“

It is our task in our time and in our generation to hand down undiminished to those who come after us, as was handed down to us by those who went before, the natural wealth and beauty which is ours.

”

John Fitzgerald Kennedy 1917-1963, Address at the dedication ceremonies of the National Wildlife Federation Building [March 3, 1961]

Message from the NESC Director [1]

Overview

The National Environmental Supercomputing Center (NESC) was established in 1992 by the United States Environmental Protection Agency (EPA). The NESC focuses its resources on solving critical problems in the environment by utilizing the power of the most advanced supercomputers and visualization techniques available.

In 1992, an eighty year old building, in Bay City, Michigan, was refurbished and transformed into a high technology center. On August 25, 1992, the first EPA owned supercomputer, a Cray Research Y-MP 8I/232, was delivered. The Cray Y-MP is a two processor computer with 32 million words of primary memory and 32 gigabytes (billions of characters) of disk storage.

The NESC's computer runs Cray's standard operating system, UNICOS, which is an acronym formed by the words UNIX Cray Operating System. As the first part of the acronym suggests, it is a UNIX System V-based system with University of California at Berkeley extensions. The second three letters indicate that the UNIX internals have been heavily modified to make it viable for use on a supercomputer.

To balance our computing facility, we procured a StorageTek Automated Cartridge System (Silo) with one terabyte (trillions of characters) of automated cartridge tape storage. Because of the storage demands made by large environmental data files, a second silo was procured giving the system two terabytes of storage.

The silos are integrated into the Cray system using Data Migration Facility (DMF) software. DMF manages the total storage pool (disks and tapes) and coordinates the movement of customer files among the devices as needed.

Our first year of operation (Fiscal Year 1993) was very successful. Immediately upon installation, customers from around the country began using the center. By the end of calendar year 1993, the resources were being fully utilized.

NESC Customers

The NESC's customers include EPA scientists and those agencies, universities, or private companies having grants, cooperative agreements, memoranda of understanding, or contracts with EPA. The models run at the NESC support the full panoply of environmental work that demands compute-intensive processing. Examples of programs supported include the Clean Air Act and the Clean Water Act. There are also models dedicated to particular areas, such as the Great Lakes, Chesapeake Bay, and San Joaquin Valley. Other areas of research include molecular modeling of various types, such as CFC studies.

By the end of FY93, the NESC's supercomputer was being used by more than 200 active customers running over forty projects. Customers are scattered from coast to coast. Representative customer locations include: Grosse Ile, MI; Duluth, MN; Research Triangle Park, NC; Las Vegas, NV; Athens, GA; Ada, OK; Washington, DC; Seattle, WA; Kansas City, KS; San Francisco, CA; University of Michigan; University of Pennsylvania; and State University of New York.

Computing at the NESC

All scientific computers at the NESC use a form of the UNIX operating system and

the TCP/IP communication protocol. Virtually all supercomputer centers use this arrangement, since the user interface is the same regardless of whether one is using a Data General or Silicon Graphics workstation or a Cray supercomputer. This makes user access easy and minimizes the time required to learn a new operating system. Any scientific customer who knows UNIX can immediately use the NESC Cray.

Since all customers are remote to the NESC, telecommunications play an essential role in the NESC's effectiveness. One T3 (45 Mbps) line and one T1 (1.5 Mbps) line connect the NESC to the EPA's National Computing Center (NCC) in Research Triangle Park. Remote laboratories and regions connect to the NESC through the NCC. Plans are being formulated to have the remote sites connect directly to the NESC.

The NESC also has an Internet connection through a cooperative agreement with the Consortium for International Earth Science Information Network (CIESIN). Many universities and laboratories gain access to the NESC's supercomputer using the Internet.

Visualization

Visualization is an integral component of high-performance computing. The NESC has established a visualization laboratory, which acts in conjunction with the visualization laboratory at RTP. Virtually all customers of the NESC use some form of visualization.

Visualization training classes have been held for customers. A highly successful visualization conference was held in July, and featured speakers and attendees from the EPA as well as external to the EPA, (e.g., Jet Propulsion Lab).

Customer support

The NESC's customers are supported by a special scientific group dedicated to customer satisfaction. Staff members include scientists with advanced degrees in physics, chemistry, and computer science. This group helps scientists port their codes to the supercomputer and optimize existing codes in order to meet customers' missions. This group coordinated a computational chemistry workshop in September, which had speakers and attendees from the EPA as well as external to the EPA (e.g., Dow Chemical and Midland Molecular Institute).

Future plans

Because of the increasing demand for computation cycles, plans have been formulated for the replacement of the Cray Y-MP 8I/232 with a more powerful Cray C94/264. The C94 will have two processors, each two to three times more powerful than our existing processors. The primary memory, which has been a limiting factor on the Y-MP, will be doubled to 64 million words. The increased memory will facilitate those jobs that require a large primary memory space. Examples are the Regional Oxidant Model (ROM) and the Regional Acid Deposition Model (RADM), which are required for compliance with the Clean Air Act. In addition to primary memory, the NESC's high-speed disk storage will be almost tripled in size.

EarthVision

EarthVision, EPA's educational program, is administered through a cooperative agreement with a local university, Saginaw Valley State University. This is a competitive program, whereby high schools submit proposals. If selected, the school's students take part in a tutorial program that takes place on Saturdays during the academic year. The high school teams then submit another proposal and a winner is selected for a three-week Summer Research Institute.

It is during this institute that the schools work on their accepted projects. They continue to work on their projects during the next academic year, and at the end of the year produce a report on the research. The winning proposal for FY93 was *Uptake and Food Chain Transfer of Polychlorinated Biphenyls (PCBs) in the Zebra Mussel (Dreissena polymorpha).* (Editors note: Research articles from both EarthVision teams appear on page 105 and page 109.)

---

[1] Arthur G. Cullati, Director, National Environmental Supercomputing Center (NESC), 135 Washington Avenue, Bay City, MI 48708.

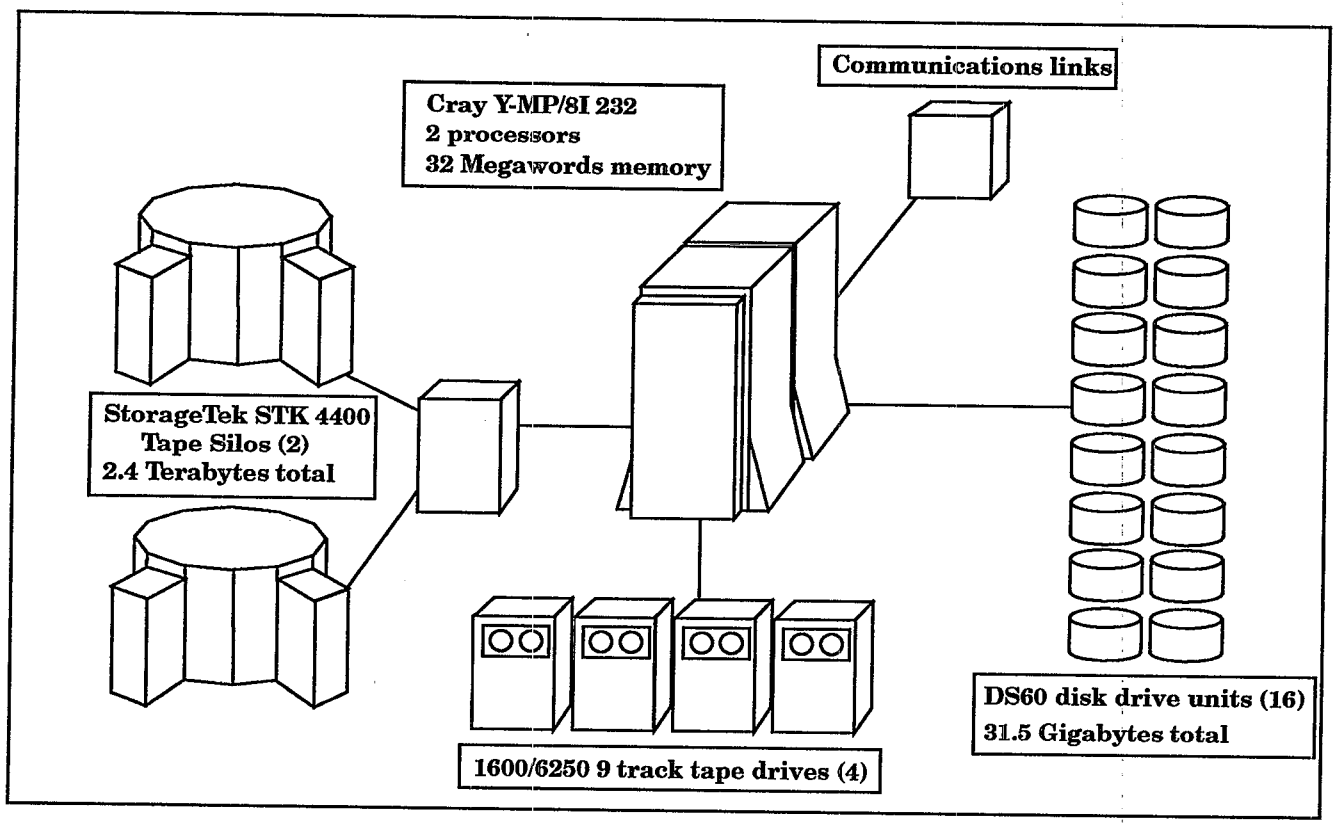


Figure 1: Hardware Configuration Overview, National Environmental Supercomputing Center (NESC) - September 1993.

Mass data storage

In addition to disk drives, the NESC has two StorageTek 4400 robotic tape silos. Each silo contains about 6,000 tape cartridges and is capable of storing 1.2 trillion bytes of information. The two units combined provide a total storage capacity of 2.4 terabytes (2,400,000,000,000) of data. All tape handling is performed by robotic arms and is completely automatic and "transparent" to the user. Two 3 MB/second data lines connect the silos with the Cray.

Traditional "round" tape facilities are also available by special request. Other data transfer media may be available. Contact the NESC for further information.

Telecommunications

In order to meet its mission, the NESC must serve customers throughout the United States. From its location in Bay City, the NESC uses a sophisticated telecommunications network to serve customers at EPA sites around the country.

The NESC's telecommunications network is illustrated in Figure 2 and Figure 3 (page 11). The network consists of both a Local Area Network (LAN) and a Wide Area Network (WAN). Each is described in greater detail in the following paragraphs.

The LAN, shown on Figure 4 (page 12), is responsible for communications inside the NESC. It consists of four Ethernet

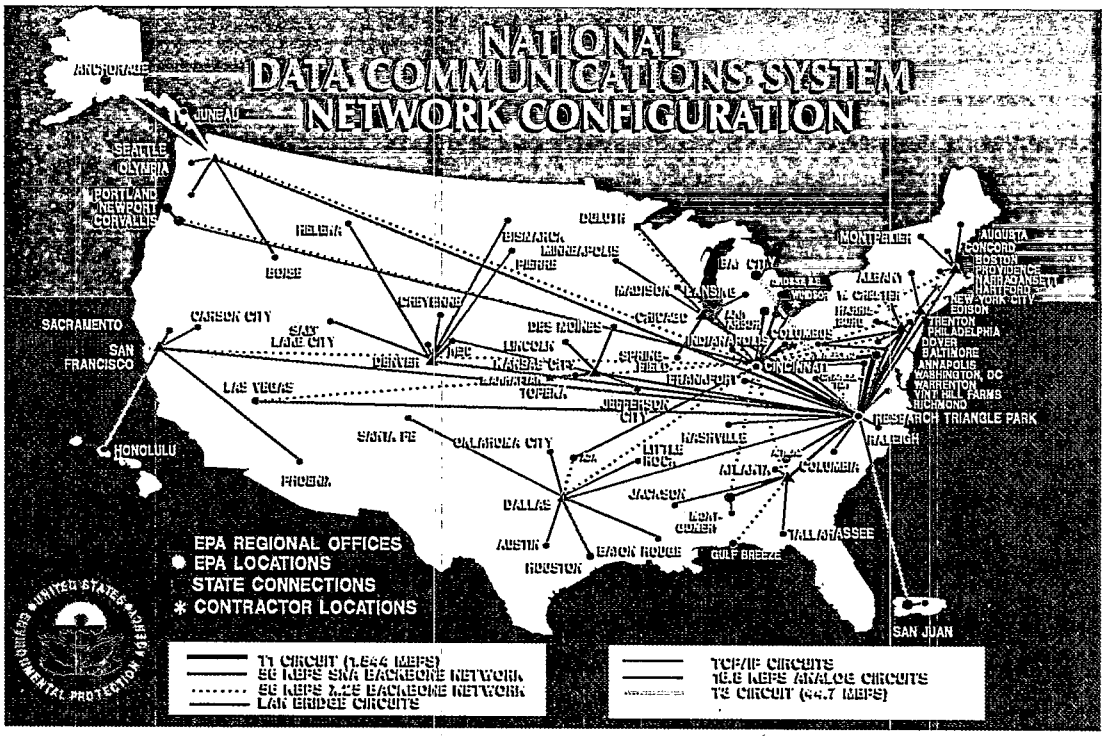


Figure 2: A sophisticated telecommunications network connects the NESC with EPA researchers throughout the United States.

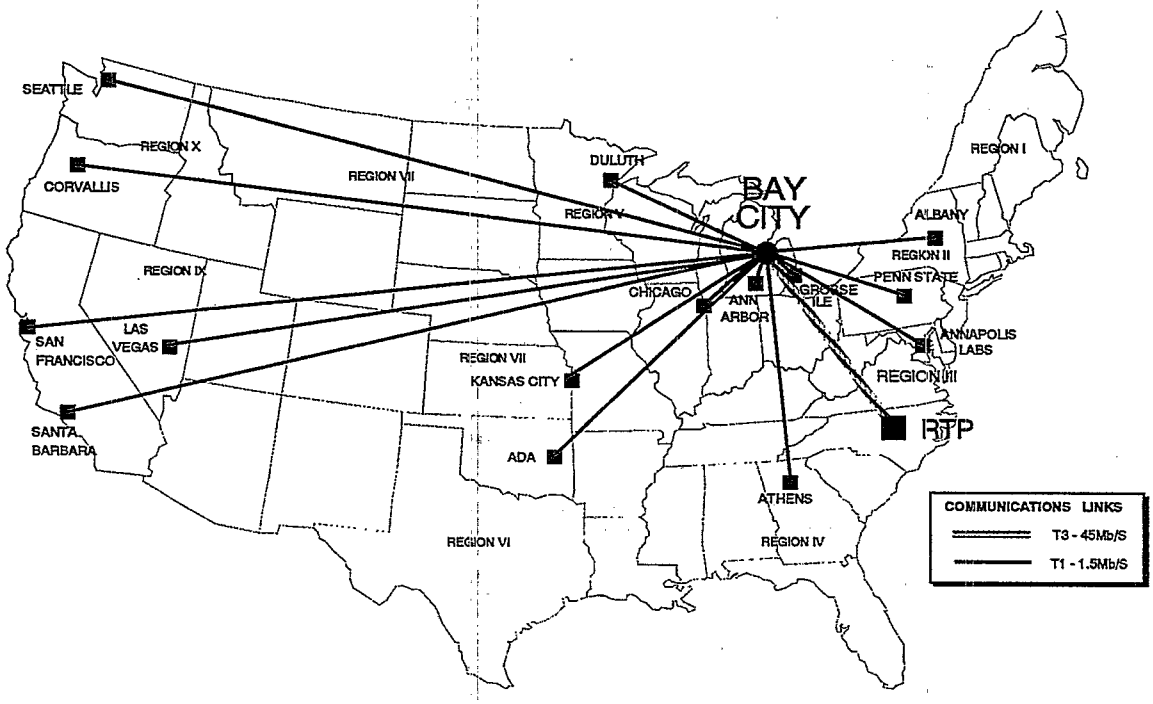


Figure 3: EPA Researchers use the NESC's supercomputer from throughout the United States via dedicated and Internet communications links.

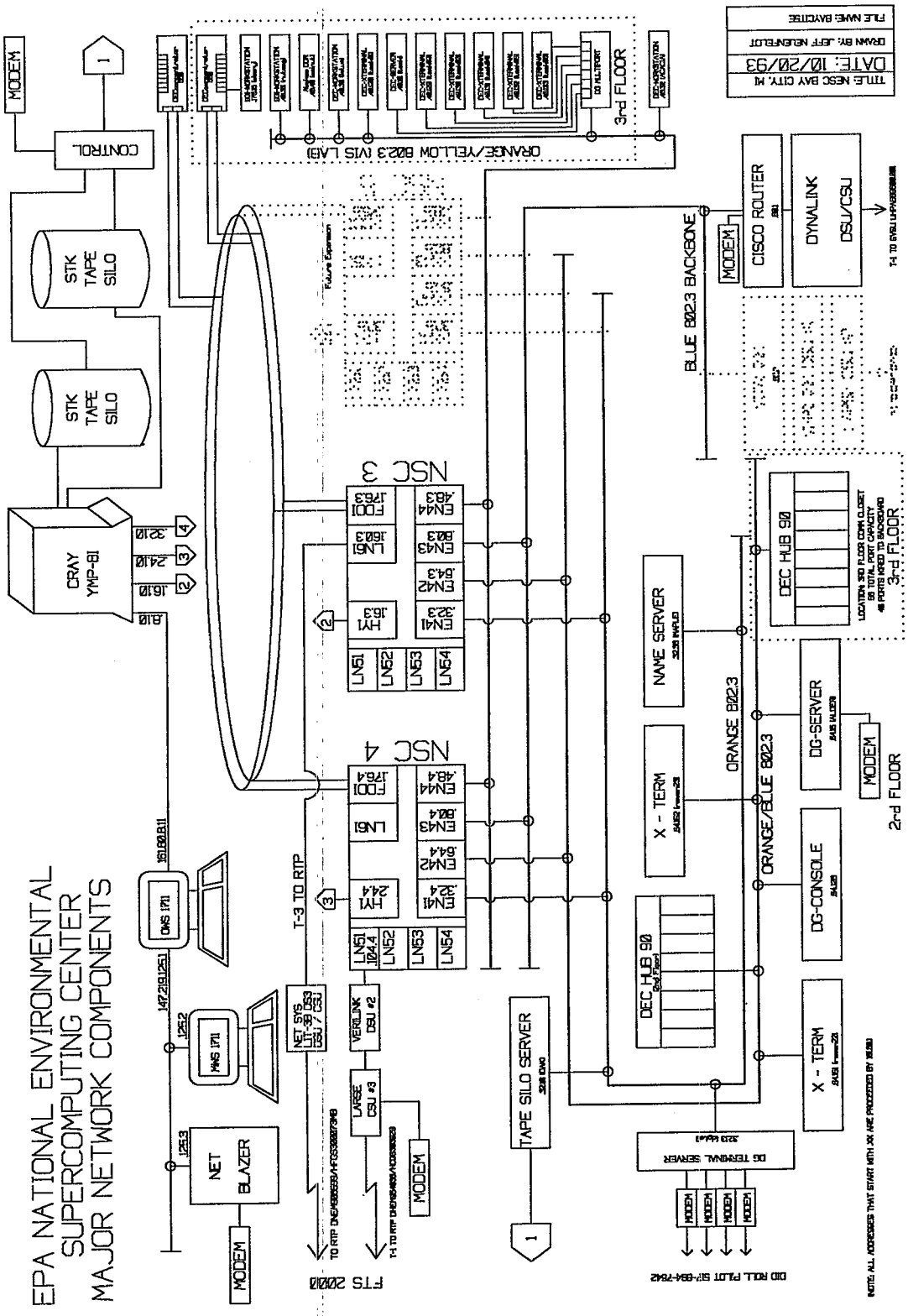


Figure 4: NESC Local Area Network (LAN)

backbones, capable of transmitting data at a rate of 10 million bits per second (MbS). In addition, there is a single Fiber Distributed Data Interface (FDDI), which moves data at 100MbS. These networks are responsible for moving data within the NESC.

The WAN is responsible for moving data between the NESC and its customers. The WAN consists of one T3 transmission link and two T1 transmission links. The T3 link, which is capable of a peak transmission rate of 45MbS, connects the NESC with the EPA's largest research facility in Research Triangle Park (RTP), North Carolina.

One T1 link, which has a peak data transmission rate of 1.5MbS, also links the NESC with the EPA's RTP facilities. The second T1 link connects the NESC with MichNet which, in turn, connects the NESC with the NSF Net and the Internet.

Telecommunications routing is handled through two NSC high speed routers. These routers are fully redundant, with each router capable of managing all telecommunications traffic. Two 12MB/second data lines connect the routers to Sequoia.

Through the use of TCP/IP protocols and the File Transfer Protocol (FTP), researchers can log into Sequoia, move data to and from Sequoia and their local network, and obtain the results of their research. The network's high speed and bandwidth offer the NESC's customers the same speed and throughput as if Sequoia was in the same room.

Facilities

A significant amount of infrastructure is required to operate and maintain a supercomputing center. As with an iceberg, much of this equipment lies below the surface. The NESC is no exception. The two major supercomputing support areas are electrical power and equipment cooling.

Electrical Power

Supercomputers, such as the Cray Y-MP 8I, require considerable quantities of properly conditioned electricity at the correct voltage. To support that need, the NESC has two 2,500 KVA utility feeds, each one of which is capable of supplying all necessary power. A motor/generator unit provides both voltage transfer and conditioned power for essential systems and support equipment.

In the event of a power failure, the NESC has two 750 KVA uninterruptable power supplies (UPS), which are capable

Table 1: Software Applications currently available on Sequoia (as of September 1993.)

Discipline	Application
Chemistry	Amber 4.0 AMSOL 3.0 CHARMm 22 DISCOVER 2.9.0 DMol 2.3 GAUSSIAN 92 MOPAC 6.0.2
Mathematics / Statistics	BCSLIB/BCS-EXT Release 12 IMSL 2.0 NAGlib mark 15 LIBSCI
Data Exchange	netCDF
Graphics / Visualization	AVS 5.0 NCAR Graphics 3.1.3a

of operating the center for fifteen minutes, sufficient time to permit an orderly shutdown of the computer. The UPS are powered by a bank of 360 storage batteries. Finally, in the event of a prolonged power interruption, a 150 KVA natural gas-powered generator is available to support the NESC's non-supercomputing functions.

Cooling

A by-product of supercomputing is the generation of considerable heat by the computer's densely-packed circuitry. Without sufficient cooling, the Cray Y-MP 8I would be subject to extensive thermal damage.

To keep Sequoia functioning, the NESC has three 110-ton chilling units operating through two 175-ton cooling towers. In the event of a total failure of the chilling units, a 2,000-gallon chilled water reservoir provides up to 15-minutes of emergency cooling capacity.

Software

To complement the NESC's supercomputing hardware, the NESC supports EPA researchers and scientists with specialized scientific application software packages. Table 1, page 13, lists the software applications that are available to researchers as of September 1993.

Sequoia uses an operating system based on UNIX. By using a UNIX-based operating system, researchers can easily move their programs and applications between their local environment and that of the NESC. Once a user becomes familiar with UNIX, those skills are transferrable across a number of hardware platforms, including that of the Cray.

Another advantage of UNIX is its adaptability to Distributed Computing. Be it through Massively Parallel Processing (MPP) or some form of distributed computing, UNIX permits the NESC to easily embrace future trends in large-scale scientific computing.

Visualization

In addition to "crunching numbers", the power and speed of a supercomputer supports the extensive use of graphical visualization. EPA scientists can call upon state-of-the-art graphical visualization and computer-modeling capabilities to augment their research. These visualization techniques permit the NESC's users to "see the unseeable".

Using graphically-based scientific workstations, environmental researchers develop complex mathematical models of air pollution, atmospheric conditions, the chemical components of pollution, and other Grand Challenges. The speed and data handling capabilities of supercomputers allow environmental scientists to model the interaction of the complex variables that, until now, could not be tested in the laboratory.

Another important aspect of the NESC's visualization support is in the vital area of Computational Chemistry. This rapidly developing branch of chemistry permits chemists to use a supercomputer in place of their more traditional test tubes and flasks. Computational Chemistry experiments are intuitive, fast, and cost-effective.

The NESC features a state-of-the-art visualization laboratory staffed by experts in scientific visualization. EPA researchers are encouraged to use the laboratory and its staff to transform their research data into strikingly meaningful graphical images.

The NESC's visualization group includes six specialists located at RTP. They are available to serve EPA's researchers with personalized service and advice.

The NESC staff

In addition to its hardware and software, a world-class supercomputing center requires considerable talent and expertise on the part of its staff. The NESC's staff includes experts in supercomputing operations, planning, computational science, and related fields. The NESC staff is organized into the following functional areas:

- Management
- Operations and Systems Support
- Computational Science Support
- Facilities
- Visualization
- Documentation
- Telecommunications

The NESC's staff is dedicated to supporting the users. Staff expertise is available to assist researchers with questions about computer systems, UNIX, code optimization, application porting, and documentation. User contacts and inquiries are encouraged.

NESC Usage

Almost from its dedication in late 1992, EPA researchers have made extensive use of the NESC's computational resources. For the fiscal year, more than 99% of available CPU hours were available to our users. Figure 5 shows the CPU-hour utilization for the fiscal year.

Plans are underway to increase the NESC's computing resources. With the addition of increased computational power scheduled for the spring of 1994, the NESC will continue to be a resource dedicated to meeting our customers's compute-intensive needs.

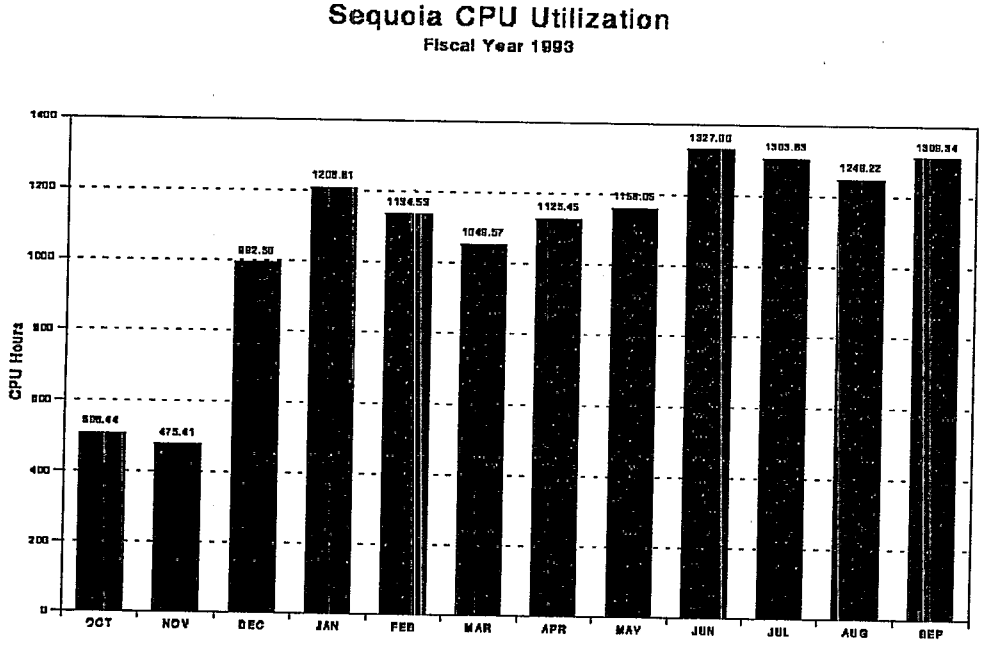


Figure 5: FY1993 CPU utilization - Sequoia

“Now ’tis the spring, and weeds are shallow-rooted; Suffer them now and they’ll o’ergrow the garden.”

William Shakespeare 1564-1616
King Henry the Sixth, Part II [1591],
Act: III, Scene: i, Line: 31

Great Lakes Environmental Visualization Conference Overview

The United States Environmental Protection Agency's (EPA) National Data Processing Division (NDPD) and the Great Lakes National Program Office (GLNPO) sponsored the Great Lakes Environmental Visualization Workshop, July 15–16, 1993, in Cleveland's Marriott Society Center. Approximately 100 people, representing EPA research and policy-making activities, other Federal agencies, the International Joint Commission, the Great Lakes Commission, State environmental agencies, and local colleges, attended the two-day event.

Day 1 - July 15

Dr. Walter M. Shackelford, NDPD Director of Scientific Computing, set the stage for the workshop, and Dr. Arthur G. Cullati, Director of the EPA's National Environmental Supercomputing Center (NESC), provided an overview of high-performance computing and visualization technologies at the NESC.

The workshop was designed to provide EPA and associated environmental researchers and policy analysts with an opportunity to learn about scientific visualization tools. The focus of the workshop was on the application of visualization and high-performance computing technologies to environmental research problems. With that goal in mind, three visualization experts shared their experiences and perspectives on working with environmental data sets.

Bill Hibbard, University of Wisconsin at Madison, gave a presentation on using visualization as a diagnostic tool for troubleshooting complex environmental models; Kevin Hussey, NASA Jet Propulsion Laboratory, discussed creating high-end animation sequences for the environmental sciences; and, Lloyd Treinish, IBM Thomas J. Watson Research Center, talked about maintaining data integrity while transporting environmental data sets into visualization toolkits.

These visualization experts also conducted hands-on, 30-minute workshops on visualization tools their groups have developed. Hibbard provided an overview of Viz AD; Hussey presented a new software environment (Surveyor) under development; and, Treinish demonstrated IBM Data Explorer.

Martin Marietta Technical Services, Inc. staff taught sessions on visualizing environmental data with the Flow Analysis Software Toolkit (FAST) and the Application Visualization System (AVS).

Day 2 - July 16

An important component of visualization technology transfer includes computer graphics education programs. On the second day of the workshop, Dr. Acha Debela, chair of the Historically Black Colleges computer graphics education effort of the Association for Computing Machinery's Special Interest Group on Computer Graphics (ACM/Siggraph), and members of his committee examined education issues. Gloria Brown-Simmons, Jet Propulsion Laboratory Visiting Scientist at Central State University in Ohio, gave a talk on establishing the Center for Scientific Visualization at Central State University.

The day's technical program included presentations by EPA researchers using visualization technology to solve environmental science problems and by EPA

managers responsible for implementing high-performance computing technologies within the Agency.

James Giatinna, Deputy Director of GLNPO, discussed the crucial function of high-performance computing in understanding the Great Lakes ecosystem; William Richardson, Station Chief of the Large Lakes Research Station at Grosse Ile, Michigan, discussed changing the future of environmental research with the use of scientific visualization at the Large Lakes Research Station; and, Robin Dennis, Senior Program Manager at AREAL and Co-Chair of the Agency's High-Performance Computing and Communications (HPCC) Program, highlighted the Agency's HPCC Program and components related to multi-environmental media visualization and collaborative computing.

Pranas Pranckevicius, Chief of Data Integration at GLNPO, concluded the workshop with a presentation on GLNPO's support of ecosystem protection in the Great Lakes with visualization.

Also on July 16, Bob Beltran from GLNPO and Barry Bolka with the GIS Management Office in Region 5 demonstrated working with visualization tools in the MS-DOS and 486 PC environment.

NDPD and GLNPO look forward to making the visualization workshop an annual event for sharing high-performance computing and visualization technology activities taking place within EPA.

EPA Computational Chemistry Workshop

The United States Environmental Protection Agency's (EPA) National Data Processing Division (NDPD) and the Environmental Monitoring Systems Laboratory (EMSL) - Las Vegas jointly sponsored a Computational Chemistry workshop. The workshop was held from September 27 through 29, 1993 at the National Environmental Supercomputing Center (NESC) in Bay City, Michigan.

This workshop was the first major event held at the NESC. Topics of discussion included some of the major work being done using the NESC's supercomputer. This was a unique opportunity to hear in-depth details about this new and exciting scientific discipline. The study of challenging problems such as global warming, ozone depletion, air and water pollution, and acid rain has been tremendously accelerated as a result of the implementation of computational chemistry approaches. Instead of conducting experiments with flasks and test tubes, the computational chemist works with computers, usually a workstation, mainframe, and/or supercomputer. These experiments, called simulated experiments, are intuitive, fast, and cost-effective.

Approximately 60 people, representing various EPA related institutions and others, attended the three-day workshop. Prominent scientists from local institutions, such as Dow Chemical, Michigan Molecular Institute, the University of Michigan, Saginaw Valley State University, and Wayne State University, participated in the scientific presentations and discussions. Other speakers included scientists from national and international universities specializing in the field of computational chemistry.

The main objective of this workshop was to survey the scientific objectives and achievements of the EPA in the field of computational chemistry. These objectives include studying problems such as: global warming, ozone depletion, air and water pollution, acid rain, radiation hazards, and chemical toxicity. Experts in areas such as Quantitative Structure Activity Relationship (QSAR) for chemical exposure and risk assessment, Computational Analytical Chemistry, Water and Atmospheric Chemistry Modeling, Toxicity Prediction, and Database Design, led the scientific presentations and subsequent discussions.

Of particular interest was that, for the first time, scientists from the regulatory wing and the research wing of the EPA met and exchanged ideas and discussed issues of mutual interest. Thus, this workshop at the NESC served as a melting pot for the Agency's long term research and regulatory objectives in the field of Computational Chemistry. The local press and scientific journals carried news items about the workshop. The coordinators of this workshop are pleased to note that the NESC workshop met and exceeded all expectations.

Day One - September 27

The opening session was chaired by Ben Bryan, Martin Marietta/NESC Manager. Dr. Walter M. Shackelford, NDPD Director of Scientific Computing, opened the workshop by welcoming the participants. He specifically thanked the Environmental Monitoring Systems Laboratory (EMSL) - Las Vegas for jointly sponsoring this workshop. A welcome note from Wayne Marchant, Director, Environmental

Figure 1: Computational Chemistry Workshop session

Monitoring Systems Laboratory (EMSL) - Las Vegas, was subsequently read by Don Betowski of EMSL - Las Vegas.

Dr. Arthur G. Cullati, Director of the NESC, then addressed the audience and gave them a brief introduction and history of the NESC. The next speaker was Dr. George Delic, Martin Marietta/NESC Computational Science Services Manager. Dr. Delic presented an overview of his group's activities, including this workshop, as an outreach activity. Dudley Bromley, Martin Marietta/NESC Visualization Support Manager, then presented an overview of his group's activities. Dr. Don Betowski, one of the workshop coordinators, delivered the vote of thanks.

The keynote speaker was Gilles Klopman, Professor and Chairman, Department of Chemistry, Case Western Reserve University, Cleveland, Ohio. He spoke on his years of research quantifying the activities of chemicals in terms of their structures. Klopman has been using computers in his research since 1958, and he acknowledged that "the structure-bioactivity challenge" still continues, whether it is in toxicity prediction or in drug design.

The next speaker was Dr. Steven Bradbury, Acting Associate Director for Research Operations, U. S. Environmental Research Laboratory, Duluth, MN. In his talk, Dr. Bradbury stressed the importance of key experimental data for the development of computer models for toxicity prediction. Dr. Stephen C. DeVito, Program Manager, U. S. Environmental Protection Agency, Headquarters, Washington, DC, gave an overview of various problems experienced by the regulatory wing of the EPA and stressed the importance of computational chemistry methods in coping with those issues.

The next session was titled, "Structure-Activity: Computational Methods." The speakers for this session were: Prof. Bernard Schlegel of Wayne State University, Prof. Donald Aue of the University of California-Santa Barbara, Dr. Gilda Loew of Molecular Research Institute, Dr. George Famini of the U. S. Army, Dr. Ann M.

Richard of the U. S. EPA-RTP, Dr. Krish Namboodiri of Martin Marietta/NESC, and Dr. Scott DePriest of Tripos Associates. They spoke on various computational chemistry methods used for solving a number of environmentally-related problems.

In the evening, the panel session that followed the reception attracted almost all of the workshop's participants. The main discussions centered around the strengths and weaknesses of various models and the sharing of information between the regulatory and research wings of the EPA. Dr. Steven Bradbury of ERL-Duluth chaired this exciting session.

Day Two - September 28

The second day opened with the session theme: "Molecular Basis of Toxicity." Prof. Roman Osman of Mt. Sinai Medical School, NY, presented a talk on the molecular changes/damages due to radiation effects on the most important biomolecule, deoxyribonucleic acid (DNA). Dr. James R. Rabinowitz of the U. S. EPA-RTP and Prof. George Pack of the UIC College of Medicine presented their research on explaining carcinogenesis based on the interaction of toxic molecules with DNA. Prof. Robert Pearlman of the University of Texas - Austin, Prof. Robert Bach of Wayne State University, and Dr. Christopher Waller of Washington University presented many approaches to the study of the basis of biological activity (e.g. toxicity) by extensively using workstations and supercomputers.

The next session was devoted to one of the most recent offshoots of computational chemistry, Computational Analytical Chemistry. Prof. Einar Uggerud of the University of Oslo, Norway, presented a method of calculating mass spectra of toxic molecules, a tool with plausible applications in characterizing toxic waste dump sites. Dr. Don Betowski, Prof. Don Aue, and Dr. Kathleen Robins presented their research on predicting the infrared spectra of toxic molecules, which has tremendous applications in remote sensing. The final session of the second day was "Computational Chemistry - Selected Topics." Bob Hunter of the NRRI-Duluth, talked on neural net approaches for classifying toxic chemicals. Dr. Robert Lipnick, of the U. S. EPA, Headquarters, Washington, DC, spoke of the necessity for using advanced computational chemistry tools for regulatory purposes. And Dr. Sandy Sillman, of the University of Michigan, Ann Arbor, presented a talk on numerical methods used in the prediction of acid rain and ozone depletion processes.

Day Three - September 29

The last day of the workshop was specially devoted to vendor presentations and hands-on sessions. The day started with four short presentations: Dudley Bromley, Martin Marietta/NESC, Visualization Support Manager, presented the video on the Center for Ecological Research and Training (CERT); Dr. Aileen Alvarado-Swaisgood presented various computational chemistry software tools available from Biosym Technologies, Inc.; Dr. Eric Stahlberg, of Cray Research, Inc., presented that firm's software, Unichem; Scott Hutton and Scott DePriest presented various molecular modeling tools available from Tripos Associates, Inc.; and Dr. Peter Grant of Molecular Simulations, Inc., presented computational chemistry tools available from his company. Following those presentations, the vendors demonstrated their software to the attendees on an individual basis.

Workshop summation

Using computers in chemical research adds a new dimension to the traditional experimental approach. For example, it is difficult to comprehend the nature of ozone depletion based on certain rare and costly experimental results. By incorporating computers and graphic visualization into this research, a whole new world of graphical and three-dimensional images emerges. This not only improves comprehension, but also saves time and money, while permitting more aggressive and innovative scientific research.

Queueing Theory Analysis of Service Level for NQS on the Cray Y-MP 8I/232

Abstract

The NQS complex on the National Environmental Supercomputing Center's Y-MP8I/232 has been analyzed to determine service levels as required by the U.S. EPA's National Data Processing Division under the existing contract with Martin Marietta Technical Services (MMTS). To meet this contractual requirement, elementary results from queueing theory have been applied in the analysis of jobs submitted to ten public and four private queues on Sequoia over a nine month period during the fiscal year 1993. The analysis shows that the probability density function of queue wait times and service (or wall-clock) times are hyperexponential distributions and that process rates and mean times are easily extracted after a fit to the empirical data. The queueing theory analysis has enabled the installation of a program (`qperf`) on the Cray that indicates the expected queue wait time and service time for the queue appropriate to the user specified CPU and memory requirements. While past performance is no guarantee of future results, stability of the analysis is only suspect if the character of the whole job population changes drastically at some future time. Clearly, the longer the time interval of the sample the more stable the prediction and therefore the analysis will be updated on a quarterly basis at NESC.

Introduction

It is a requirement of the U.S. EPA's National Data Processing Division under the existing contract with MMTS that users of the NESC have access to a quantitative measure of service levels for jobs submitted to the Cray resource at the NESC. At first sight, this seems a difficult requirement to meet because of the non-deterministic fashion in which resources are allocated to jobs submitted to the Cray under UNICOS. However, detailed information on job processing is available from Cray Standard Accounting (CSA) and, at the NESC, locally written code extracts job-level transaction data from the CSA super-record on a daily basis. This process tabulates (for each batch job) the CPU, service (or wall-clock) and queue wait times. In addition to these, the I/O activity (4K blocks moved) and memory integral (KW-minutes) is also recorded. This sample has been accumulating since inception and provides a valuable resource for analysis with a view to determining the quality of service that users enjoy at NESC.

The NESC NQS System and Job Level Data Collection

Sequoia at NESC has ten public and four private queues as given in Table 1, page 24, which also shows the sample sizes for the respective queues. The period of the sample represents nine months of throughput on Sequoia during which no major system changes occurred that could drastically affect throughput characteristics. The sample represents a valuable resource and this report summarizes how the service and queue times are analyzed. While no detailed statistical analysis is present here, Table 1 does show some important basic results. In most cases, for both of these times, the sample standard deviation is larger than the mean which indicates the likelihood of a hyperexponential distribution. Also, it is

Table 1: Sample statistics for CPU, service and queue times for the ten public and four private queues of the NESC NQS. Queue limits are shown in million words (memory) and seconds (CPU time). The sample size for each queue is shown in the column labelled 'N'.

	Mem MW	CPU secs	N	CPU time (min)		Service time (min)		Queue time (min)	
				Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
small_short	4	600	1377	1.1938	2.3719	21.175	164.93	74.93	299.41
medium_short	12	600	255	1.0492	2.036	6.4031	11.065	8.8683	40.389
large_short	18	600	11	2.4914	3.7456	12.32	17.61	309.02	240.44
small_normal	4	3600	1440	10.022	13.307	33.24	52.43	58.03	172.73
medium_normal	12	3600	1268	13.235	16.14	41.6	54.57	80.789	169.22
large_normal	18	3600	629	9.834	13.128	39.378	131.83	159.95	215.19
small_long	4	100000	1197	85.586	142.31	212.9	420.59	117.05	349.7
medium_long	12	100000	1355	106.33	197.81	241.47	452.62	758.69	1777.53
large_long	24	999999	494	226.18	466.77	493.64	1114.26	261.92	598.46
night	24	999999	42	213.86	346.16	449.53	775.1	600.09	616.1
areal1	10	999999	570	93.678	85.545	204.77	209.73	642.9	960.74
areal2	16	999999	505	70.496	58.162	228.58	233.69	287.85	626.94
areal_romdp	6	1830	388	0.2947	0.3353	15.31	22.286	9.8729	26.0712
areal_radmp	4	1830	22	0.7498	0.4968	75.484	272.76	4.4242	7.4717

interesting to observe that the mean CPU times are invariably significantly smaller than the queue CPU time limits.

Elements of queueing theory and analysis of NESC data

Once a job has been dispatched to a particular queue by NQS, it can be viewed as residing in a single server queue system. Queueing theory[1] then applies and treats these times as random variables or observables which do not have individually predictable values but whose values show statistical regularity. In particular, a random variable, X, is completely described by a probability distribution function, $F(t)=Prob\{X \le t\}$, or by the corresponding probability density $f(t)=dF/dt$. The latter is a frequency distribution and may have various possible shapes depending on the details of the queueing system. However, in this analysis it is found that the distribution of queue and service times is dominated by exponential shapes such as $f(t)= \mu e^{-\mu t}$, $t>0$, or combinations of exponentials (hyperexponential). Therefore, the analysis[2] requires the determination of the rate parameter μ from which the probability distribution is computed as $F(t)=1-e^{-\mu t}$. The function $F(t)$ is also known as the cumulative probability because it represents the "area" under the density curve $f(t)$. Since both job queue wait and service times have been recorded at NESC, each is analyzed in four simple steps: (1) a sort into bins as is done in a

histogram plot, (2) the fitting of the resulting distribution with $\mu e^{-\mu t}$ to determine μ, (3) computation of the mean as $1/\mu$ (property of the exponential distribution), (4) computation of `F(t)`. Results of this procedure are shown in Figure 1, page 26, for the service time in the small_normal queue. The distribution is characteristic of a hyperexponential function where the forward peak is described by an exponential with the service rate of 0.03218 jobs/minute and the tail is described by another exponential with a service rate of 0.004532 jobs/min. The mean service times are simply the inverses of the respective rates for each exponential distribution. The fact that there is a hyperexponential distribution shows that the jobs in service are of (at least) two types. If p assigns a probability that the job is of either type, then Figure 2, page 27, shows a sequence of probability densities corresponding to different values of p. The large dots connected by the dotted line represent the same data shown in Figure 1, page 26.

Figure 3, page 27, shows the probability distribution functions computed from the two component exponential distributions of Figure 2, page 27, for various values of the probability, p. Each curve shown here represents the accumulated area under the corresponding probability density distribution of Figure 2. The cumulative probability distribution in Figure 3 shows (on the vertical axis) the likelihood that a given job has the predetermined total service (or wall-clock) time chosen (on the horizontal axis). While a table of the mean rates (μ) and mean service times (σ) is shown for the tabulated choices of the probability, p, the mean service time (σ) is not the best indicator of service level. A job with a value `p=0` has a 60% chance (lowest curve) of receiving a service time of 200 minutes, whereas, a job with a `p=1` value has a 60% chance of receiving a service time of approximately 25 minutes. Stated in another way: the expectation that the same job would complete in 200 minutes changes from 60% to nearly 100% if p moves closer to 1. Yet another way to read the probability distribution of Figure 3 is to obtain the percentiles of the sample. As an example, using the `p=1` curve for the above-mentioned case, 60% (vertical axis) of the sample of 1440 jobs had a service time of approximately 25 minutes (horizontal axis) or less.

Conclusions

A nine month sample of job level NQS data on Sequoia has been analyzed by a simple single server queueing model. The resulting analysis demonstrates that there are at least two job classes. The first job class shows shorter queue wait and service times while the second job class has longer times. The second job class represents a small fraction of the measured sample and some reasons for large values can be given. In the case of queue wait time, users tend to submit multiple jobs, but the system will only service two jobs from the same user at any given instant. Therefore, the queue wait time increases for later jobs in a single batch from the same user. In the case of service (or wall-clock) time, large values can result from several causes, such as: waiting on migrated files, availability of requested memory, heavy I/O activity, or competition with interactive usage for either memory or CPU time.

A program called `qperf` has been installed on Sequoia and made available to users of NESC. A simple command generates a tabular form of the probability distribution function for both queue wait and service times observed for the sample. The user specifies the CPU time and memory requirement and the appropriate queue is selected. Because no model exists at present for estimating p in the range `0<p<1` and selecting the job class

characteristics *a priori*, `qperf` uses the results for the analysis of the first job class only since this class represents the major fraction of the sample.

---

[1] H. Kobayashi, *Modeling and Analysis: an Introduction to System Performance Evaluation Methodology*, Addison-Wesley Publishing Company, Reading, MA, 1978.

[2] J. Banks and J. S. Carson, *Discrete-Event System Simulation*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1984.

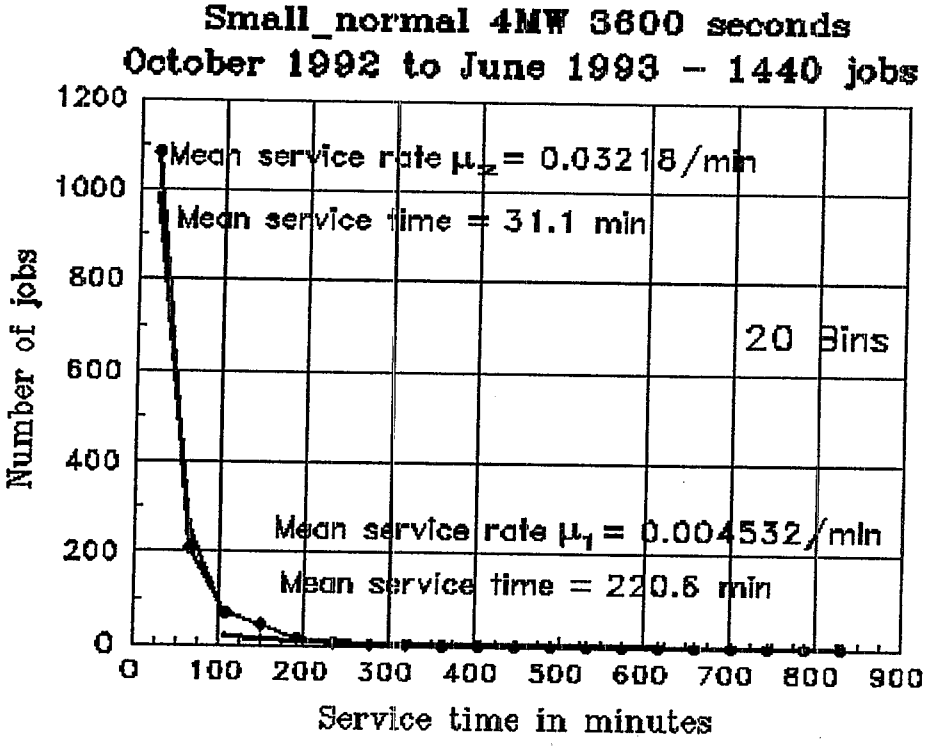


Figure 1: Example of the probability density function for the service time of the small_normal queue which shows the number of jobs versus the amount of service time required.

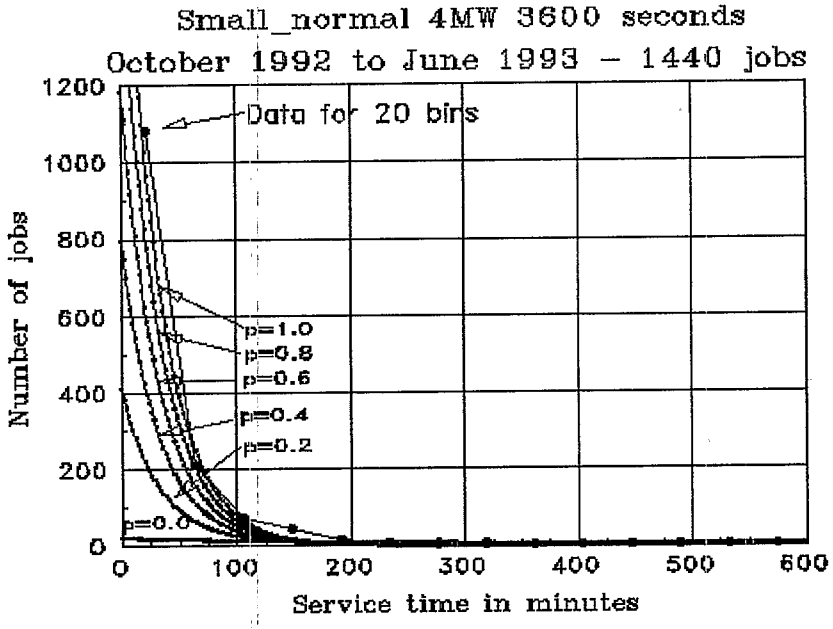


Figure 2 : Probability density functions computed from the two component exponential distributions of Figure 1 for various values of the probability, p, which measures if the job is likely to be in the tail (p=0) or the peak (p=1) of the empirical distribution.

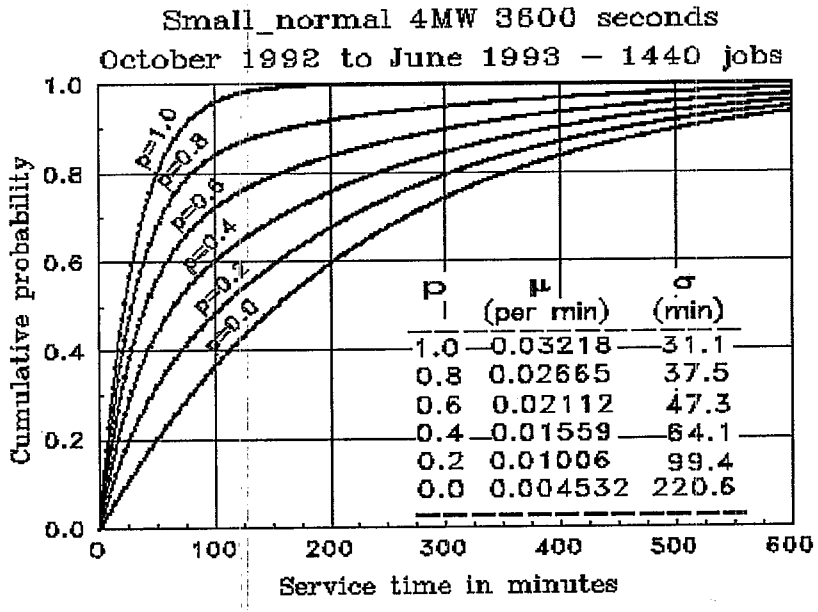


Figure 3: Probability distribution functions computed from the two component exponential distributions of Figure 1 for various values of the probability, p, which measures if the job is likely to be in the tail (p=0) or the peak (p=1) of the empirical distribution. Each curve shown here represents the accumulated area under the corresponding probability density distribution of Figure 2.

“All the mathematical sciences are founded on relations between physical laws and laws of numbers, so that the aim of exact science is to reduce the problems of nature to the determination of quantities by operations with number”

James Clerk Maxwell 1831-1879
On Faraday's Lines of Force [1856]

Optimization of Programs on a Supercomputer

Introduction

The NESC's Cray Y-MP 8I supercomputer (Sequoia) has been installed and running for just over a year. During that time, many EPA programs have been executed on it. As we see an increasing number of different program types running on Sequoia, we are increasing our understanding of those program's performance.

Many of the programs running on Sequoia were ported from other types of machines, such as Digital Equipment VAX, and IBM 3090. Some users of the Cray Y-MP have made the assumption that their programs will execute with supercomputer speed once they are running on the Cray Y-MP. Unfortunately, this assumption is simply not true.

Due to its unique design, the Cray Y-MP can execute user programs over a large range of performance levels. How effectively these programs perform, depends on their optimal use of the machine's special capabilities. The difference in performance between a well-optimized program, and an unoptimized program, can easily be an order of magnitude (as much as 448 times faster as described later in this article.). That is, if the program, unchanged after porting from another machine, takes ten hours of execution (CPU) time to run, a small effort at optimization could change it to run in only one hour. The example that follows shows how dramatically performance can be improved on a Cray Y-MP machine.

Why make the effort to optimize?

While optimization may require some effort, there are several major benefits that far outweigh that investment.

- If researchers can get their model results back sooner, they will have more time to "do their science." Otherwise, they may be waiting for those results that will tell them what modeling changes to try for the next run.
- Naturally, if a program runs faster, it will not deplete the scientist's time resource allocation on Sequoia as quickly as an unoptimized program.
- If the program can be made to run faster, the scientist using it may be able to run more iterations of the model. Faster execution may also lead the way to further refinements of the model, such as a smaller grid with more gridpoints.

It is definitely to the advantage of all Sequoia users to optimize their programs for improved Cray execution. The Computational Science Support (CSS) staff provides optimization help and advice.

"Some users of the Cray Y-MP have made the assumption that their programs will execute with supercomputer speed once they are running on the Cray Y-MP. Unfortunately, this assumption is simply not true."

The IDAHO4 Optimization Effort

By making certain changes to the IDAHO4 program, NESC's CSS staff was able to speed it up by a factor of 448. This was accomplished using the Cray Y-MP machine at the National Environmental Supercomputing Center (NESC), located at Bay City, MI. The code was provided by Ross Kiester, a scientist at the Forest Research Station, which is associated with the Environmental Research Laboratory, EPA, Corvallis, OR (ERL-Corvallis).

The program originally required 42,142 seconds (11.7 hours) of computation (CPU) time to execute on the Cray Y-MP machine at the North Carolina Supercomputer Center (NCSC). The program now requires only 94 seconds (1.6 minutes) of CPU time, running on the NESC machine. Thus, this scientist could perform 448 runs of his model in the time previously required to execute only one run.

It took about seven hours to achieve this speedup. While this is a dramatic example of the benefits of optimization, even a small speedup in a long-running program can be of help to the user of that program.

Technical Issues

The IDAHO4 program is written in the C language. The problem this program solves is a combinatorial study. There are 404 distinct sections (quads) of data, that must be compared in groups of four. Every possible combination of four must be tried. This causes the number of comparisons to "explode" to slightly more than one billion. Each section of data (quad) contains 359 separate flags. Each flag can take on a value of 1 (true) or 0 (false).

The IDAHO4 program attacks this problem using brute-force methods. Every one of the 359 flags is logically combined with every other flag in the group of four being checked. Thus, for each of the one billion combinations tested, 1436 boolean calculations, and 359 boolean tests must be performed.

The first optimization issue was not one of vectorization, but rather the volume of work performed. The amount of work performed in the hardware was 64 times more than the program required, since each boolean flag occupied one Cray word. This fact led the optimizer to pack the flags into a 64-bit Cray word (64 flags per word), a better utilization of the Cray Y-MP hardware.

Note that the Cray Y-MP machine can not only perform boolean operations on 64 bits at a time, but there are two parallel boolean functional units available in the machine's CPU. Thus, this configuration can allow 128 simultaneous boolean operations to take place in every machine cycle (six nanoseconds). Given in rate terms, this means that an ideal code could perform over 21 billion boolean operations per second per CPU.

Once the amount of work was lessened, it was possible to continue the optimization process by improving the vectorization of the loops that the program performs.

Note however, that vectorization alone cannot always improve a code's performance. In the case of the IDAHO4 program, it was necessary to first reduce the total amount of work performed by the vector functional units of the Cray Y-MP. The final execution time figure (94 seconds) represents a boolean rate of 9.8 billion operations per second (the peak theoretical rate of the Cray Y-MP is 21 billion per second). This rate is 97 times higher than the boolean rate of the original code, which was 101 million per second.

EarthVision: EPA's Grand Challenge for High Schools

What is EarthVision?

EarthVision is a professional development and educational program for teams of high school teachers and students. Each team is composed of two teachers and four students. It is a joint venture between the United States Environmental Protection Agency (EPA) and Saginaw Valley State University (SVSU). EarthVision helps high schools develop environmental research programs using computational science and access to high speed computers.

EarthVision is the first computational science education program to concentrate solely on environmental issues, offer Saturday tutorials, and provide multi-tiered outreach activities.

EarthVision Components

Saturday Tutorials

During the academic year, teams of high school teachers and students learn the specific skills required to conduct environmental research using computational science. Each team member is given substantial hands-on experience. Each team is supported by mentors, who are professional researchers and/or specialists in computational science, and by a member of the SVSU outreach support team. These sessions are held at the National Environmental Supercomputing Center (NESC) in Bay City, Michigan. After the completion of the Saturday tutorials, the teacher/student teams write two proposals to apply for admission to the Summer Research Institute: (1) a research proposal and (2) an education plan proposal to introduce environmental research and computational science into their high school program.

Summer Research Institute

During the Summer, competitively selected teams participate in a three-week educational program at SVSU. Team selection is based on proposals for a research project to be conducted during the academic year and an education plan. Summer Research Institute participants receive instruction in conducting an environmental research project of their own design. Each team is supported by mentors and by an SVSU outreach support team. Each high school participating in the Summer Research Institute is supplied with a scientific workstation, and a telecommunications link to the National Environmental Supercomputing Center (NESC) in Bay City. The Summer Research Institute prepares the teams for the next part of EarthVision, conducting environmental research at their school.

"The students are not only excited, they are very committed! They're going the extra mile, they're taking their spare time to learn as much as they can ... It's been a surprise and excitement for us all that they're far beyond where we originally thought they'd be at this time."

Lynne Petterson, Project Officer, EarthVision: EPA's Grand Challenge for High Schools, National Data Processing Division (NDPD), U.S. EPA.

Environmental Research at High School(s)

During the academic year, the participants from the Summer Research Institute conduct environmental research activities at their high schools. They will use their scientific workstation, and the supercomputer at the National Environmental Supercomputing Center in Bay City, to analyze data, conduct environmental modeling and use scientific visualization to implement their research. The teams will prepare a paper describing their research. The mentors advising each team will continue to provide support. The SVSU outreach team visits the schools which have participating teams and assists them in their research and in establishing an educational program in environmental research and computational science.

Benefits to the General Public

The general public will benefit by exposure to student research projects and presentations, enhanced awareness of new technologies, and new ways of conducting scientific research. Additionally, the general public will have opportunities to see the types of research conducted on a supercomputer.

The EarthVision teams will present information about their environmental research, their experience and their curriculum plans at local and national conferences such as the IEEE SuperComputing 1993 and the International Conference on Scientific Visualization. They have also given numerous local presentations to their colleagues in their own schools, other schools and community groups. Additional presentations and publications are

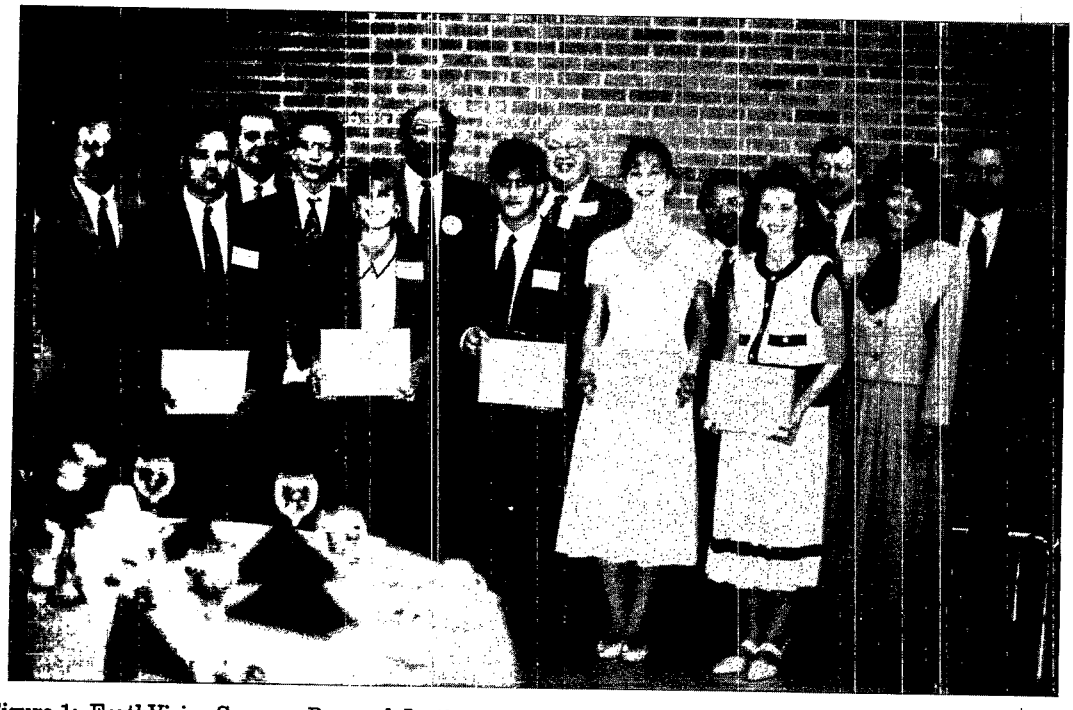


Figure 1: EarthVision Summer Research Institute 1993 Course Completion Ceremony, Bay City Central High School Team. First row from left to right Mark Neal, teacher; Jill Bisel, teacher; Navid Mazloom, student; Kim Kukla, student; Erin Gatza, student; Lynne Petterson, U.S. EPA; Second row from left to right Jon Whan, Superintendent Bay-Arenac ISD; Don Fulford, U.S. EPA; Jason Schroeder, student; George Charles, Principal; Willis Greenstreet, U.S. EPA; Joe Gonzales, Superintendent, Bay City Public Schools; Walter Shackelford, U.S. EPA; The Honorable James Barcia, U.S. House of Representatives.

being planned and prepared. During the Summer Research Institute, the students had several opportunities to present information about their research to various constituencies. It was found that the experience of preparing for presentations was a valuable pedagogical tool in assisting them in focusing on specific elements of their research.

"The teams worked very hard and it was not a very easy task for them because they had to combine everything they had been taught: the science, the computers, the operating systems, the visualizations. All of this had to be put together in such a way that they could go back and look at the question they wanted to ask and determine whether or not they could actually ask that in the way proposed. Watching them do this and turn it into a visualization was absolutely spectacular."

Ken Flurchick, Chief Scientist North Carolina Supercomputer Center

How has EarthVision developed beyond its original intent?

Environmental Research during the Summer Research Institute

Two of the three 1992-93 EarthVision teams that finished the Saturday tutorials also completed a three-week Summer Research Institute. The original intent was simply to prepare the teams to begin conducting their research during the following academic year. Both teams actually began conducting their research during the Summer Research Institute. They made progress in formulating their models, mathematically writing code to represent those models, and preparing initial visualizations of their data.

During the Summer Research Institute, lecture and hands-on sessions were lead by experts in scientific research, modeling and visualization. These experts were from organizations such as: the North Carolina Supercomputer Center, U.S. EPA Visualization Lab, Martin Marietta Technical Services, Inc., and Saginaw Valley State University.

Assistance from experts in diverse fields enabled the EarthVision teams to integrate complicated views, methods and obtain initial results from their research projects. Initially, it was not anticipated that the diverse representation of expertise would promote team building among the participants, but it was clear that the mentors, working together to assist the teams, served as role models of team functioning. Role differentiation on the part of the participants was a key factor in facilitating their enhanced progress. Each individual had specific tasks to complete to contribute to the team's research effort. The individual achievement made possible the team's success.

Cognitive Apprenticeships

In addition to expert lecturers, the EarthVision teams were supported by mentors working on environmentally-related research at organizations such as: the EPA's Large Lakes Research Station at Grosse Ile, MI, Dow Chemical Company, Alma College, HydroQual/Manhattan College, and the EPA's Environmental Research Laboratory, Duluth, MN. These mentors donated their time to help the teams refine and focus their approach, find additional data, and construct computa-

tional models. They will continue to work with the teams during the 1993-94 academic year until the teams complete their research projects. In addition to donating their time, mentors provide a more valuable commodity; they share their own thought processes, problem solving methods and serve as role models. Thus, EarthVision teams learn environmental science in a cognitive apprenticeship with scientists solving current environmental challenges, very much like the traditional master-apprentice learning format.

Peer Reviews and Collaborations

A multitude of currently available communication methods help the teams keep in touch with other student-teacher teams with similar missions as well as with mentors. In addition to traditional face-to-face interactions and telephone conferences; electronic mail, videoconferences, and collaborative visualization

"These kids are very surprising in a lot of ways and in some cases very frightening. They are extremely bright ... Their enthusiasm continues to challenge anybody who works with them. It's very difficult at times to keep up with them. They're like a sponge ... they tend to challenge and push anybody dealing with them because they're so interested in the way these things work."

Ken Flurchick, Chief Scientist North Carolina Supercomputer Center

Figure 2: EarthVision Summer Research Institute 1993 Project Peer Review between EarthVision teams and NCSA SuperQuest participants, using compressed video teleconference.

techniques are some of the currently available means of communication the EarthVision teams can employ.

The NESC's videoconference room has been used to contact experts and review research projects with other students nationwide. One such peer-review videoconference took place during the summer of 1993, at which the teams shared ideas with colleagues at the National Center for Supercomputing Applications (NCSA) SuperQuest project. The teams, particularly the students, recognized the importance of communication with others doing similar work. The teleconference with NCSA was a means by which the value of peer review and collaboration was demonstrated. For high school students, recognizing that they need not operate in an isolated environment was a revelation. The students believed that collaborations and cognitive apprenticeships were highlights of the Summer Research Institute.

Collaborations among the two current EarthVision teams, mentors and other students in national projects, were motivating for students and helped them prepare for the multidisciplinary-collaboration-based research environment of the future.

"You get to meet students from other schools with interests really similar to your own. And you get to meet (field) scientists and computational scientists who love to share as much information with you as they can."

Natasha Sefcovic, EarthVision student, Center for the Arts and Sciences, Saginaw, MI

"I think these students will help redefine the way we do science in America, and I think that's very exciting. Ten years, twenty years from now, I think we will be hearing again from these EarthVision students."

Lynne Petterson, Project Officer, EarthVision: EPA's Grand Challenge for High Schools, National Data Processing Division (NDPD), U.S. EPA

How Are Opportunities Translated into Remarkable Results?

Motivated and Dedicated Teachers and Students

The excitement and motivation was apparent from the extensive hours teachers and students spent working at the computers. They spent most of their recreational time, as well as regular instructional and hands-on sessions, working on their research projects.

The teacher members of the EarthVision teams played key roles in the success of the Summer Research Institute. While they were closely monitoring their students' work at the computers, advising them on methods and approach, they did not have enough time to become as knowledgeable as the students on some of the new applications, such as the Explorer visualization software. The teachers worked in the SVSU Advanced Learning Technologies Laboratory on their own time late Saturday evenings and Sundays in order to keep up with the students.

The original EarthVision plan incorporated

"We have just been thrilled by what's been reported to us about the students' reaction and their working 12 to 14 hour days. They're totally committed to the project and the things that they are doing and it's great to see."

Don Fulford, Director, National Data Processing Division (NDPD), U.S. EPA

"I was somewhat apprehensive that we were offering too much material for the teams to try to work on and perhaps too much work and not enough fun but I understand the work to them is the fun."

Walter Shackelford, Director of Scientific Computing,, National Data Processing Division (NDPD), U.S. EPA

"We would hope that one or two of these kids might see fit to find their research careers in the environmental business. I don't think we're going to run out of environmental problems in the near future. So, it is a good future for them."

Willis Greenstreet, Director, Office of Administration and Resources Management/RTP, U.S. EPA

enough funding to support one team during the 1993 Summer Research Institute. This support provided stipends for teachers and students and a scientific work-station and telecommunications link to the supercomputer at NESC for their school. The Bay City team was chosen to receive this support. The EarthVision planning team invited the other Saturday Tutorial teams to participate in the Summer Research Institute without financial support, since they worked so hard and had excellent proposals. The Saginaw Center for the Arts and Sciences team chose to participate in the Summer Research Institute at their own expense. Both the Bay City team and SVSU volunteered the use of their workstations to the Saginaw team during the following academic year. The Saginaw team participated with a great deal of enthusiasm and were significant contributors to the success of the summer experience.

Experiencing the Excitement of Environmental Science

Near the end of the Summer Research Institute, the EarthVision teams demonstrated their current accomplishments at an open house for parents, siblings, friends, and invited guests. This proved to be an excellent opportunity for parents and the general public to understand more about how research and computational science can contribute to the solution of some of today's most pressing environmental problems. Even at this stage in the program, before the teams' research is done, it appears that the experience is having an influence on the students career choices.

What is Next for EarthVision?

A New Annual Cycle for EarthVision

The EarthVision planning team is preparing for a new group of participants to begin the Saturday tutorials in the 1993-94 academic year. The enthusiasm and excitement of the students and teachers partici-

Figure 3: EarthVision Summer Research Institute 1993 Course Completion Ceremony, Saginaw Center for the Arts and Sciences team. First row from left to right Chuck Rohde, student; Chris Stark, student; Natasha Sefcovic, student; Brian Weeden, student; Lynne Petterson, U.S. EPA. Second row from left to right Don Fulford, U.S. EPA; Willis Greenstreet, U.S. EPA; The Honorable James Barcia, U.S. House of Representatives; Dan Sealey, teacher; Gary Barker, teacher; Walter Shackelford, U.S. EPA; Burris Smith, Assistant Superintendent Saginaw Public Schools.

pating in the 1993 Summer Research Institute was demonstrated by their volunteering to come back during the 1993-94 Saturday tutorials to assist the new teams.

Invitations to participate in the EarthVision project have been sent to high schools in the State of Michigan. Six new teams will be selected this year for the Saturday tutorials, and four schools with the best research and education proposals will be selected for the Summer Research Institute. All four EarthVision teams will receive a scientific workstation to be used for the environmental research at the teams' schools. Participation in the Saturday tutorials is not a prerequisite for the Summer Research Institute and all research and education proposals will have equal opportunity in the competition.

"When I first heard the hours of the Summer Research Institute ... I thought that's way too much time to ask. Yet there were times that it got to be 9 o'clock and I wasn't really ready to go home. There were things I wanted to get done, and I'd go home and log on my computer and I'd work until I couldn't stay awake anymore."

Jill Bisel, teacher, Bay City Central High School

"For the longest period of time, I have wanted to exploit the computers' power. This program not only gives me the opportunity, but when I heard about the supercomputer plus experts coming in and giving shape to some of the learning I had, it gave me an idea about how to bring it all together. It was an opportunity I was excited about."

Dan Sealey, teacher, Saginaw Center for the Arts and Sciences

"I was kind of undecided between Pediatrics and Environmental Engineering. This program has helped me get a better idea of what's involved in dealing with the environment and the technology behind it."

Kim Kukla, student, Bay City Central High School

"In a lot of ways I believe we are all born scientists.... We try to figure out how the world works. Scientists try to carry it a little further and these kids are great examples of that. They want to know HOW, how everything fits together and how they can go about understanding that!"

Ken Flurchick, Chief Scientist, North Carolina Supercomputer Center

Intern and Team Role Development

The current teams have surprised the planning team by their amount of contact and level of participation. After the Summer Research Institute course completion ceremony, the project planning team assumed they would have a break after three intensive weeks of activity. However, on the following Monday morning, the teams were back at SVSU working as diligently as they had during the Summer Research Institute. This has continued into the academic year even after school has started. They have demonstrated a strong desire to continue and maintain a very close relationship with the EarthVision project, the instructors, the mentors and the planning team. They have helped the planning team to redefine the role of participants. They wish to be a part of the 1993-94 Saturday Tutorial program, and they have moved easily from the role of student to that of colleagues and life-long learners. We plan to use teacher and student interns from past EarthVision teams to interface with new EarthVision teams and provide demonstrations to the general public.

Research Interns (graduate and undergraduate students) have been working as part of the EarthVision planning team to help with the implementation of the outreach program, assisting school teams in their research and curriculum plans, providing technical support, and networking with mentors. The Research Interns will also conduct their own environmental-computational-science research projects.

Projected program growth

The original plan for EarthVision identified a fourteen county area around the NESC from which three teams would be drawn for Saturday tutorials and one team for the Summer Research Institute

Figure 4: EarthVision students work on their project.

"One of the things over the years that we've noticed is that math and science is very, very competitive. If you're going to enter math and science as a female, you are going to be competing with people internationally who have a serious attitude about learning, maybe more serious than some of the Americans. ... This project has just been a wonderful opportunity."

Kathlene Sefcovic, parent of EarthVision student

for the 1992-93 project year. Three teams were invited to participate in the 1993 Summer Research Institute with two actually participating. During the 1992-93 project year, grants were received from the Michigan Department of Natural Resources and the Michigan Department of Education to train high school teachers to use the EPA's STORET dataset in environmental research to be conducted with their students at their schools. The ten schools that participated were each given Zenith 486 computers with modems. This project was designed to build capacity at the schools by providing professional development for teachers, building a computing infrastructure and introducing environmental research. As such, it was designed to prepare schools for participation in EarthVision.

The proposed 1993-94 project year included three teams for the Saturday

tutorials and four teams for the Summer Research Institute drawn from the entire state of Michigan. The planning team decided to invite six teams to participate in the 1993-94 Saturday Tutorial program. The projection for the 1994-95 project year called for three teams in the Saturday tutorials and eight teams participating in the Summer Research Institute, drawn from the eight Great Lakes States. It is expected that six teams will participate in the Saturday Tutorial phase and plans are being formulated to design an urban EarthVision initiative in Cleveland, Ohio, that would become a national model for implementation of EarthVision in large city environments.

What Have the Teams Accomplished in their Research?

EarthVision teams have studied two questions that have social implication today. The Bay City Central High School team has been researching the possible role of zebra mussels in the accumulation of contaminants in Great Lakes fish. The Saginaw Center for the Arts and Sciences team has been researching the transport of heavy metal ion contaminants in Saginaw Bay. The two research papers detailing the approach of the teams, current accomplishments and future plans are included elsewhere in this Annual Report (the Bay City team's report starts on page 105 and the Saginaw team's report starts on page 109).

NESC Research Reports

Estimation of Global Climate Change Impacts on Lake and Stream Environmental Conditions and Fishery Resources [1,2,3]

Abstract

Mathematical models have been developed for estimating the effects of climate change (changed meteorological conditions) on lake and stream thermal structure and dissolved oxygen concentrations and on fishery resources. Regional impact analyses require the development of lake and stream classification systems to define waterbody types, which in turn require the availability of extensive regional data bases for these resources. Fishery resource response predictions require the development of large field temperature and fish distribution data bases from which species and guild thermal requirements can be derived. Supercomputing capabilities are being utilized in the development and manipulation of the large data bases to integrate the various data and program modules, and to make the calculations required to perform regional impact estimates.

EPA Research Objectives

According to a 1979 report by the U.S. National Academy of Sciences, and supported by several general circulation models of ocean atmosphere heat budgets, doubling atmospheric concentrations of CO_2 could increase global mean air temperatures by 1.5° to 4.5° Celsius in the next 100 years. This is likely to have many environmental consequences, such as changes in water temperature and dissolved oxygen concentrations, which in turn are likely to affect fish populations. The fact that such changes are occurring many times faster than expected has resulted in requests for information on the causes, effects, and response options to the projected climate changes. The Environmental Research Laboratory - Duluth and the University of Minnesota have initiated a cooperative study to determine the impacts of global warming on lake and stream environmental conditions and fishery resources. In order to continue with this study, fish thermal requirements need to be estimated using a historical fish presence/temperature record data base.

Background/Approach

The Fish Temperature Distribution Management System (FTDMS) is a national data base system that spatially and temporally associates discrete fish sample records with water temperature data. Recent efforts have concentrated on the expansion of data base content by assembling information from a multitude of sources, including federal agencies (e.g. EPA/STORET and USGS) and private museum and university collections. The assimilation of data from many sources necessitates the need for automated spatial and temporal matching of a fish record with water temperature data. Prior versions of several program modules have been converted from a PC data base platform to C and have been run on the NESC's supercomputer. Program performance has been closely monitored and several steps are being taken to increase performance.

Scientific Accomplishments

A modeling approach has been developed for estimating the effects of global warming on lake and stream environmental conditions and fisheries resources. The

initial phase of the work was partially supported by the EPA Office of Policy, Planning, and Evaluation, and the results are currently being applied to an economic impact analysis of global climate impacts on the United States. One ongoing program is a component of the FCCSET Committee on Earth and Environmental Sciences, Global Climate Research Program. At last count, more than ten publications, most in the form of technical journal articles, had been produced by the project.

Results

The program module that calculates raw temperature data into weekly mean values has been run on the NESC's supercomputer for sixteen states. The results have been used to calculate the maximum (warmest throughout the year) 95th percentile temperature where a fish species was collected for 32 species of North American freshwater fish. This temperature is used as an approximation of the lethal limit for that species and allows us to estimate the distribution of fish after global climate change.

The speed by which the weekly mean temperatures are calculated on the supercomputer makes it possible to perform the temporal matching in a number of different ways. This has the potential for improving estimates of thermal requirements for fish. For example, the southern range of distribution of cool-water fish is generally near 40° latitude. Maximum weekly mean values from south of this parallel would be expected to provide a better estimate of thermal tolerances than values from all of North America. Data manipulation on the basis of geographic regions is, therefore, desirable. Temporal matching criteria can be restricted in other ways (fish and temperatures sampled in the same year, season, or month). Comparing "monthly" and "yearly" datasets provides a means of examining the importance of the temporal relationship of temperature and fish records.

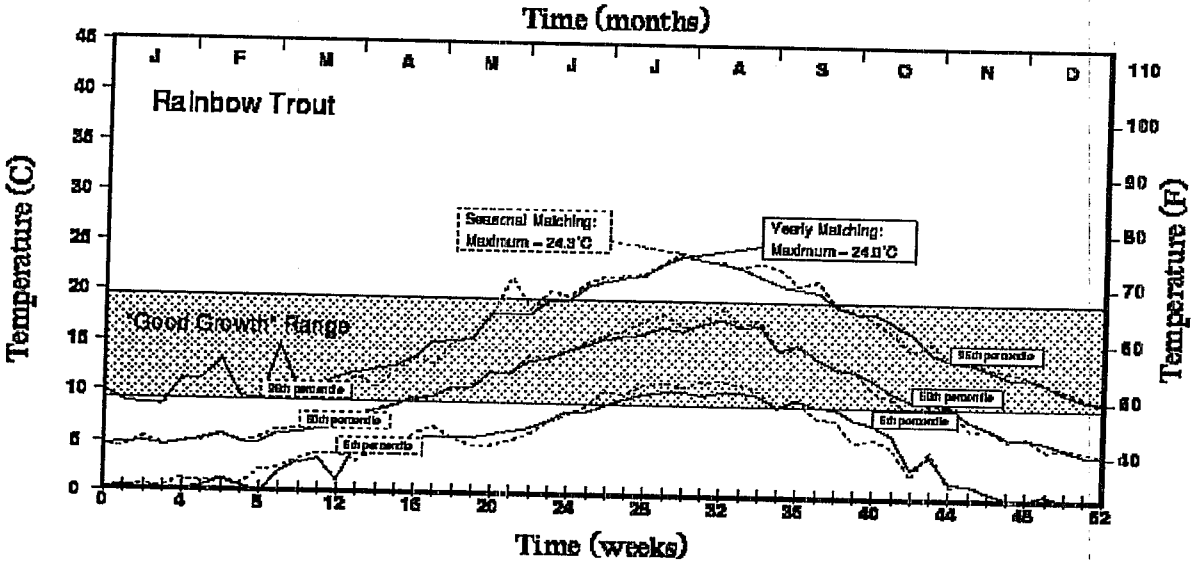


Figure 1: Graphic representation of raw temperature data

Future Objectives

Currently, the weekly mean temperature is used to describe surface water conditions for the week in which a fish sample was taken. In the near future, the weekly maximum temperature will be calculated and daily means and daily maxima will be matched to fish collections to re-calculate the maximum 95th percentile temperature. These values will provide a unique and valuable look at the relationship between various expressions of a fish's thermal regime and its geographic distribution. Most laboratory-derived measures of thermal tolerance, the source of most past temperature effects information, have employed constant temperature exposure conditions when short-term peaks might be as much or more important in nature. So far, studies of fish thermal requirements and global climate effects have been focused on species in the central United States. A template has been developed for extending these analyses to other regions of the United States and beyond. The relationship of cold temperatures to the distribution of warmwater fishes has only been examined superficially. The data storage and manipulation requirements and modeling demands will be greatly increased, just for dealing with this single environmental component (fishes). Research is underway to incorporate functional ecosystem responses (e.g. system productivity) into models projecting climate change impacts. The area of ecological processes and effects research, only as related to aquatic ecosystems, is obviously huge and can benefit greatly from enhanced computational capabilities.

Publications

Hokanson, K.E.F., B. Goodno, and J.G. Eaton, *Evaluation of field and laboratory derived fish thermal requirements for global climate warming impact assessment*, USEPA ORD "A" milestone report, p. 56, 1990.

Hondzo, M. and H.G. Stefan, *Water temperature characteristics of lakes subjected to climate change*, University of Minnesota, St. Anthony Falls Hydraulic Lab, Project Report No. 329, p. 156, 1992.

Stefan, H.G., M. Hondzo, B. Sinokrot, X. Fang, J.G. Eaton, B.E. Goodno, K.E.F. Hokanson, J.H. McCormick, D.G. O'Brien, and J.A. Wisniewski, *A methodology to estimate global climate change impacts on lake and stream environmental conditions and fishery resources with application to Minnesota*, University of Minnesota, St. Anthony Falls Hydraulic Laboratory, Project Report No. 323, p. 222, 2nd edition, March, 1992.

[1] J. G. Eaton, U.S. EPA, ERL-Duluth, 6201 Congdon Blvd., Duluth, MN 55804, 218-720-5557.
[2] H. G. Stefan, St. Anthony Falls Hydraulic Lab, Dept. Civil & Mineral Eng., University of Minnesota, Mississippi R. & 3rd Ave. S.E., Minneapolis, MN 55414-2196, 612-627-4010.
[3] D. G. O'Brien, Computer Sciences Corporation, ERL-Duluth, 6201 Congdon Blvd., Duluth, MN 55804, 218-720-5718.

“For at least two million years men have been reproducing and multiplying on a little automated Spaceship Earth.”

Prospect for Humanity [1964]

Modeling Sediment and Contaminant Transport and Fate in Aquatic Systems: [1,2,3]

Buffalo River Sediment and Contaminant Transport Modeling, Saginaw River Sediment and Contaminant Transport Modeling, and Fox River / Green Bay Sediment and Contaminant Transport Modeling

Overview

Contamination of sediments by toxic chemicals is a problem common to most of the forty-two "Areas of Concern" identified in the Great Lakes. Assessment and remediation of contaminated sediments have become high priorities for applied research within EPA, as State and Federal agencies struggle with remediation decisions involving great uncertainty, potential significant risks, and high cost. Through involvement in the Assessment and Remediation of Contaminated Sediments (ARCS) and Green Bay/Fox River Mass Balance Studies, LLRS and USCB are developing and applying models for understanding and predicting the transport and fate of contaminated sediments. These models are being applied to support remediation decisions in several major Great Lakes tributaries suffering extensive sediment contamination. The models will be used as tools to determine how various hydrologic events may affect contaminated sediments and to determine the potential effects of proposed remediation projects. The development of such models will offer predictive tools appropriate for assessing the remediation of contaminated sediments in the Great Lakes and nationwide.

Numerical models of the transport and fate of sediments and hydrophobic contaminants in the Buffalo and Saginaw rivers and the Fox River/Green Bay system have been developed to achieve this objective. The models are used to run real time simulations of high flow events as well as low flow periods. The models predict the movement of contaminants in the river and the resuspension of in-place pollutants from the sediment bed. They also predict the movement of pollutants from the rivers into the Great Lakes. The Buffalo, Fox and Saginaw Rivers are each major Great Lakes tributaries; they were chosen for modeling because of their differences. The Buffalo River is a small, narrow, winding river that has been dredged from bank to bank to allow large ship movement. This dredging has significantly increased the cross sectional area of the river in most places. Much of the time the river is almost stagnant due to flow rates less than 1 m^3/s, with flushing of the river dominated by seiche motion. The few high flow events contribute the majority of sediment movement. The Fox River is controlled by a series of dams, which stabilizes the flow response to hydrologic events. Pools created above the dams also function as reservoirs for contaminated sediments. The Saginaw River is a much larger, wider, more dynamic river. Although dredged to allow shipping, this dredging has not significantly increased the cross sectional area of the river in most areas. The median flow rates and velocities are much greater on the Saginaw.

The area around each river is urbanized and heavily industrialized. For many years, industrial and municipal wastes, including PCBs, PAHs, and heavy metals have been dumped into these rivers. These toxics have deposited in the river bed and are now buried deep in the river sediments. Resuspension and transport of these contaminants is of primary concern. Many contaminants, such as PCBs, adhere to fine-grained sediments. There-

fore, prediction of the erosion, transport and deposition of sediments is essential for predicting the movement of contaminants. Screening-level model calculations conducted in the Fox River indicate that bottom sediment resuspension is, in fact, the largest source of PCBs to the river during floods.

Research Objectives

The purpose of this study is to develop quantitative, predictive models describing the transport of fine-grained sediments and associated hydrophobic contaminants in the Buffalo and Saginaw rivers and the Fox River/Green Bay system. These models provide the high spatial and temporal resolution necessary to predict the variability and dynamics inherent to sediment and associated contaminant transport. The models also estimate the transport of sediments and contaminants from the rivers into the Great Lakes. Such estimates are essential to predict the impact of contaminated sediments and their remediation upon receiving water quality.

Approach

The SEDZL model, a numerical model of the resuspension, deposition, and transport of fine-grained cohesive sediments and associated contaminants, is used in this study. The model consists of two-dimensional, vertically integrated, time dependent hydrodynamic and transport submodels coupled with a three-dimensional, time dependent submodel of the sediment bed and its properties. The three-dimensional sediment bed model is necessary to accurately model bed properties which vary with age and depth as well as location beneath the river. Contaminant partition coefficients, sediment settling speeds, sediment resuspension parameters and sediment bed properties needed in the model were determined from laboratory and field tests. The need to determine *in situ* sediment bed properties has led to the development of a portable flume for measuring the resuspension properties at depth in cores collected from the sediment bed. Measurements of flow rates, suspended solids, and PCB concentrations from each river were used to estimate input loadings for the model, while data from sediment core samples were used to develop initial sediment bed contaminant concentrations.

Accomplishments

Model simulations for each river were performed to predict sediment and contaminant transport over numerous time periods varying from three to six months in duration. The sediment deposition/erosion patterns were then compared to changes in bathymetry measured at transects across each river at the beginning and end of each time period. By comparing the bathymetry measurements and model predictions for several time periods, as well as by comparing total suspended solids (TSS) predictions to data at the river mouth, we were able to calibrate and verify the models for sediment transport. Contaminant modeling predictions are also being analyzed and compared to existing data, especially on the Fox River where extensive PCB water column measurements were conducted.

Analysis of the transect measurements indicates that very little sediment transport occurs in the rivers during low flow periods and almost all erosion and deposition occurs during high flow events. Therefore a method was developed for running long term simulations in which each event was modeled in detail and the low flow periods were modeled statistically. This greatly decreased the CPU time necessary for the three to six month simulations and allowed us to simulate the much longer periods that are relevant for persistent, in-place pollutants. Currently we are modeling 25 year scenarios

to determine the long-term decrease in load of contaminants from each river and the erosion and burial of contaminants in the sediment bed. These scenarios will then be modified to test the effectiveness of proposed remedial alternatives, such as dredging or capping of contaminated sediments, or discontinuation of navigation dredging.

Cray Usage

To accurately predict sediment and contaminant transport on a river requires both a complex set of equations and a fine grid. Both the Saginaw and Buffalo rivers require 900 grid elements to accurately define the river; over 2000 grid elements were required for the Fox River. This makes the SEDZL code computationally intensive. Workstations like the Sun SPARCstation 2 require several CPU hours to simulate one day. Therefore, the only realistic way to do long term simulations (even with statistical averaging of low flow periods) is to use a vectorized code on a supercomputer. All long term simulations were done on the NESC Cray, with only short duration test runs being done on workstations. The CPU time necessary for a 1-day simulation on, for example, the Buffalo River, is 0.3 CPU hours. A three-month simulation without averaging of low-flow conditions would require 30 CPU hours. By using the statistical averaging procedure, the same simulation can be reduced to only 4-5 hours. The 25-year forecasts, with statistical averaging, require about 70 CPU hours. This is a considerable quantity of supercomputer time; however, this is the relevant time scale for evaluating transport and fate of persistent toxic chemicals. Such computations would be virtually impossible without supercomputers.

Future Plans

Our research team will continue to verify the sediment and contaminant transport model predictions on the Buffalo and Saginaw Rivers, and the Fox River/Green Bay system, using existing data as well as information gathered from additional sampling. In the Fox River, the model is being used to identify and prioritize sediment sampling efforts for Wisconsin Department of Natural Resources. Also, long-term forecasts will be performed with the models for specific remediation design alternatives to assist decision makers.

Our plans also include modeling sediments and contaminants in Green Bay and Lake Michigan using a three-dimensional version of the SEDZL model. The ultimate goal of these efforts is to develop and demonstrate a predictive model that can be applied anywhere in the Great Lakes, which can be run with minimal field data from the area of study. The model will be based on an accurate understanding and description of sediment and contaminant transport and fate processes, the former including particle settling, aggregation/disaggregation, and resuspension. For PCBs, the contaminant processes include sorption, sediment/water and air/water transfer.

Publications:

Gailani, Joe, C. Kirk Ziegler, and Wilbert Lick (1991) *Transport of Suspended Solids in the Lower Fox River*, Journal of Great Lakes Research. 17 (4), pp. 479-494.

---

[1] Douglas Endicott, U.S. EPA - Large Lakes Research Station, Grosse Ile, MI.

[2] Mary Cardenas, Kirk Freeman, and Wilbert Lick, Department of Mechanical and Environmental Engineering, University of California, Santa Barbara, CA.

[3] Joseph Gailani and Mark Velleux, A S & I at U.S. EPA - Large Lakes Research Station, Grosse Ile, MI.

“The only solid piece of scientific truth about which I feel totally confident is that we are profoundly ignorant about nature. It is this sudden confrontation with the depth and scope of ignorance that represents the most significant contribution of twentieth-century science to the human intellect.”

The Medusa and the Snail [1979]
On Cloning a Human Being

Development, Calibration and Evaluation of a User-Friendly Littoral Ecosystem Risk Assessment Model [1,2]

Overview

The goal of this research is to define the domain of application of the Littoral Ecosystem Risk Assessment Model (LERAM) using a number of littoral enclosure and pond field studies where different types of pesticides and toxic chemicals were tested. Once the best uses of LERAM have been delineated and its ecological risk assessment capabilities defined, it will be made available to the regulatory and risk management communities. It is anticipated that LERAM will be used by the Office of Prevention, Pesticides, and Toxic Substances (OPPTS) during the registration of pesticides and industrial chemicals and by risk managers for predicting changes in ecological risk associated with watershed management options.

The NESC supercomputing facilities will be used for three tasks:

1- Calibrate the parameters of the Littoral Ecosystem Risk Assessment Model (LERAM). This is done using an optimization algorithm to minimize a function that measures the difference between the LERAM simulation of population biomasses and biomass data from a littoral ecosystem.

2- Create a user-friendly interface for LERAM that will allow the user to make predictions of ecological risk to a littoral ecosystem from exposure to specified stressors and display the results in real time.

3- Evaluate LERAM by simulating the effects of a number of concentrations of selected chemical stressors and comparing these simulations to field data. The supercomputing environment will be used to run 500 (or more) Monte Carlo iterations of the model at each treatment concentration for the purpose of sensitivity and uncertainty analysis.

EPA Research Objectives

At the present time, laboratory tests and mathematical models form the basis of the ecological risk assessment paradigm used by the United States Environmental Protection Agency (EPA). Until the early 1980s, single species tests were used almost exclusively to provide hazard assessments of chemicals. At that time, the National Academy of Sciences (1981)[3] and others (Levin and Kimball, 1984)[4] documented the need for supplementary information from field tests, microcosm experiments, and mathematical models to better assess chemical hazards for different geographic regions, seasons, application methods, spatial scales, and levels of biological organization. Along with the increased interest in using field tests, microcosm experiments, and mathematical models to predict system responses to perturbations, it became apparent that little was known about the accuracy of predictions made by these techniques. The EPA's objectives for the research proposed here include evaluating and refining one ecological risk assessment technique using field data from controlled experiments in natural systems.

Background/Approach

In order to reach the EPA's objectives, Lake Superior Research Institute (LSRI) and Environmental Research Laboratory - Duluth (ERL-D) researchers began the

development of the Littoral Ecosystem Risk Assessment Model (LERAM) in June, 1989. LERAM is a bioenergetic ecosystem effects model that links single species toxicity data to a bioenergetic model of the trophic structure of an ecosystem in order to simulate community and ecosystem level effects of chemical stressors. It uses Monte Carlo iterations of these simulations in order to calculate probabilistic ecological risk assessments of chemical stressors. To date, LSRI and ERL-D researchers have developed LERAM to the point where it models the unperturbed behavior of a littoral ecosystem (i.e., the "behavior" of control communities), and the response of that system to the insecticide chlorpyrifos, with a high degree of both accuracy and precision (Hanratty and Stay[5], submitted to J. Appl. Ecol.; Hanratty et al.[8], 1992). Current work using data from the 1988 esfenvalerate study (Lozano et al., 1989)[6] appears even more promising. Further work is required, however, to continue refinement of the model, to validate its output using data from other littoral enclosure studies, and to develop more accurate model predictions.

The LERAM parameters have been calibrated to the control enclosures from two different littoral enclosure field experiments (Brazner et al., 1988[7]; Lozano et al., 1989)[6] using the North Carolina Supercomputing Center's (NCSC) Cray. The simulations resulting from these new parameter sets are much closer to the field data than the simulations earlier calibration methods produced. In many cases, the model simulation overwrites the field data (figure 1, page 52). The NCSC personnel worked on improving the vectorization of the LERAM code by vectorizing over the Monte Carlo iterations of the

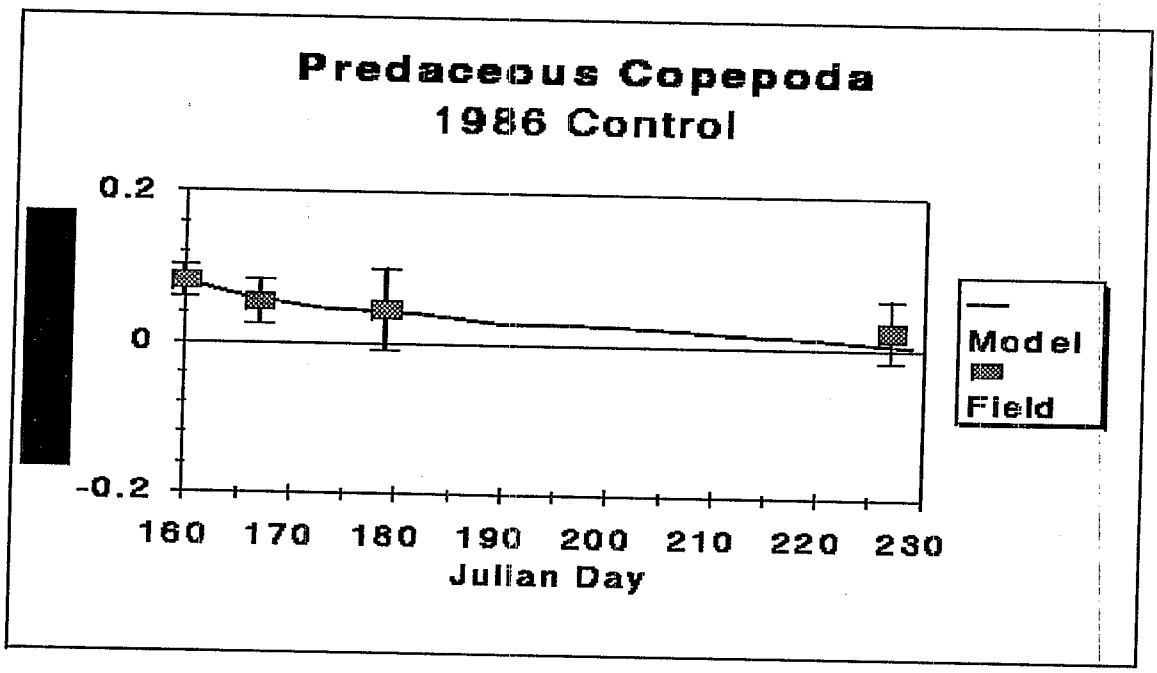


Figure 1: One example of the fit obtained by calibrating the LERAM parameters using field data. The solid line represents the simulated biomass for predaceous Copepoda and the squares with error bars represent the biomass measured in the field. The field data is the mean predaceous Copepoda biomass in the control enclosures in the chlorpyrifos littoral enclosure study (Brazner, et al., 1988).

model. In the future, we will work on improving the efficiency of the calibration algorithm and continue to improve the vectorization of the LERAM code as much as possible.

Comparison with pre-Cray Results

When the calibration program is run on other computers, such as a DEC VAX or a 486 PC, it takes so long that the program becomes impractical to use. The model itself can be run on other computers, but one loses the advantage of being able to sit and wait about a minute for the results, which allows the user to concentrate on the ecological risk assessment or modeling problem at hand. When the simulations are performed on other computers, the model can take anywhere from 20 minutes to three hours, depending on the computer and the number of Monte Carlo iterations performed.

Future Objectives

The following tasks are proposed for evaluating and refining LERAM using the NESC supercomputing facilities.

Fiscal year 1994:

1- Port the LERAM code and input files to the NESC supercomputer and establish communication with NESC personnel.

2- Enhance the vectorization of the LERAM code in the supercomputing environment.

3- Investigate methods for improving the minimization algorithm that is used to calibrate the LERAM parameters.

4- Use the above minimization algorithm to find the best parameter sets for representing the control experimental ecosystems measured in various field studies.

5- Simulate the effect of a number of chemical stressors and evaluate the model's accuracy using field data from littoral enclosure and pond studies.

Fiscal years 1995-96:

1- Create user-friendly interfaces for LERAM.

2- Calibrate the LERAM parameters using the control experimental units in future field studies.

3- Simulate the effect of the chemical stressors used in future field studies in order to further evaluate LERAM and define its domain of application.

Relevant Reports and Publications

Bartell, S.M. 1987. *Technical Reference and User Manual for Ecosystem Uncertainty Analysis* (EUA):

1- The Pascal PC Demonstration Program.

2- The Standard Water Column Model (SWACOM).

3- The Comprehensive Aquatic System Model (CASM). U.S. EPA Office of Toxic Substances Report.

Bowie, G.L., W.B. Mills, D.B. Porcella, C.L. Campbell, J.R. Pagenkopf, G.L. Rupp, K.M. Johnson, P.W.H. Chan, S.A. Gherini, *Tetra Tech, Inc. and C.E. Chamberlin. 1985. Rates, Constants, and Kinetics Formulations in Surface Water Quality Modeling* (Second Edition). EPA 600/3-85/040. U.S. Environmental Protection Agency, Athens, Georgia.

Hanratty, M.P., F.S. Stay, and S.J. Lozano. 1992. *Field Evaluation of LERAM: The Littoral Ecosystem Risk Assessment Model, Phase I.* U.S. EPA Environmental Research Laboratory-Duluth, Minnesota.

Hanratty, M.P. and S.J. Lozano. In prep. *Field evaluation of modeling and laboratory risk assessment methods.*

Press, W.H., B.P. Flannery, S.A. Teukolsky, W.T. Vetterling. 1989. *Numerical Recipes in Pascal: The Art of Scientific Computing*. Cambridge University Press, New York.

[1] Michael P. Hanratty, Lake Superior Research Institute, University of Wisconsin-superior WI 54880. Phone: 715-394- 8315. Email (Internet): *mhanratt@uwsuper.edu.*

[2] Francis S. Stay, U.S. EPA Environmental Research Laboratory-Duluth, 6201 Congdon Blvd., Duluth, MN 55803. Phone: 218-720-5542.

[3] National Academy of Sciences, 1981. *Testing for Effects of Chemicals on Ecosystems.* National Academic Press, Washington, DC.

[4] Levin, S. A. and K. D. Kimball. 1984. *New Perspectives in Ecotoxicology.* Environmental Management 8:375-442.

[5] Hanratty, M.P. and F.S. Stay. Field evaluation of the Littoral Ecosystem Risk Assessment Model's predictions of the effects of chlorpyrifos. Submitted to J. Appl. Ecol.

[6] Lozano, S.J., J.C. Brazner, M.L. Knuth, L.J. Heinis, K.W. Sargent, D.K. Tanner, L.E. Anderson, S.L. O'Halloran, S.L. Bertelsen, D.A. Jensen, E.R. Kline, M.D. Balcer, F.S. Stay, and R.E. Siefert. 1989. *Effects, Persistence and Distribution of Esfenvalerate in Littoral Enclosures.* EPA 600/X-89/292. U.S. EPA Environmental Research Laboratory-Duluth MN.

[7] Brazner, J.C., L.J. Heinis, and D.A. Jensen. 1989. *A Littoral Enclosure Design For Replicated Field Experiments.* Environmental Toxicology and Chemistry 8:1209-1216.

[8] Hanratty, M.P., F.S. Stay, and S.J. Lozano. 1992. *Field Evaluation of LERAM: The Littoral Ecosystem Risk Assessment Model, Phase I.* U.S. EPA Environmental Research Laboratory-Duluth, Minnesota.

Integration of Computational and Theoretical Chemistry with Mechanistically-Based Predictive Ecotoxicology Modeling [1,2,3,4]

Background

In the field of environmental toxicology, and especially aquatic toxicology, first generation interpretive methods, such as Quantitative Structure Activity Relationships (QSARs), have developed as scientifically credible tools for predicting the ecological effect and fate of chemicals when little or no empirical data are available. The proper application and continued acceptance of these techniques requires that methods and models be developed to systematically assign chemicals, with an estimate of certainty, to the appropriate QSARs. For example, failure to use the correct model to predict the toxicity of a given compound can result in a toxicity prediction error of 10 to 1000. In addition to using quantitative models, analog selection techniques are also employed whereby data associated with "structurally-similar" chemicals are used to estimate risk levels of compounds for which no data is available. Failure to identify and select correct structural analogs can also lead to order of magnitude errors. The lack of sophisticated techniques to select structural analogs and appropriate quantitative relationships represents a significant area of uncertainty in the ecological risk assessment of chemicals.

As the use of first generation interpretive models have become well established in ecological hazard identifications, their use in ecological hazard and exposure assessments has increased. Coupled with these new application issues, it is necessary to address several areas of research to provide the scientific basis whereby second generation methods can be applied to these emerging needs. Current QSAR methods reliably estimate acute lethality and chronic no-effect levels for 75 to 85 percent of the industrial chemicals; however, these methods underestimate the risks of reactive chemicals by several orders of magnitude. Consequently, QSAR methods must be developed to accurately predict the toxicity of electrophiles and free radicals. In addition, the current inability to quantify xenobiotic metabolism can result in significant error when predicting toxicity. For example, failure to identify and quantify metabolic activation can lead to toxicity predictions that underestimate hazards of xenobiotics to aquatic organisms by several orders of magnitude. Conversely, typical bioconcentration models, which assume no xenobiotic metabolism, can result in estimates that are overly conservative by several orders of magnitude.

Research Objectives

To reduce uncertainties in ecological risk assessments of chemical stressors, a second generation of advanced predictive modeling techniques is required. The second generation of models must be based on fundamental principles of chemistry, biochemistry and toxicology, and designed in such a way to efficiently assess the thousands of chemicals in commerce. Research must be directed towards establishing quantitative means of assessing chemical similarity to reduce uncertainties in selecting chemical analogs and QSARs. Research must also be directed towards the development of mechanistically-based QSARs for reactive chemicals to improve toxicity predictions and estimates of metabolism.

Approach

Based on the areas of uncertainties previously described, two major projects are being undertaken. The research within these projects is designed to resolve critical scientific concepts that will result in the development of advanced interpretive methods for assessing the ecological risk of chemical stressors. The overall approach is designed to assess hypotheses by integrating computational and theoretical chemistry with toxicological principles. Through this integration, advanced QSARs for reactive toxicants and xenobiotic metabolism will be developed, as well as techniques to quantify chemical similarity in the context of toxicological properties.

For well established physico-chemical and ecotoxicological databases maintained within the U.S. EPA Environmental Research Laboratory-Duluth (ERL-Duluth) QSAR system, a variety of similarity metrics and artificial intelligence systems, including neural networks, are being evaluated. A critical component in this research effort will be to develop approaches whereby variation in a physico-chemical or toxicological property can be related to variability in a given distance metric used in a multi-dimensional chemical structure space. An important area of investigation has been to determine the means whereby chemicals with common modes of toxic action can be classified. Efforts to date have been addressing the use of neural networks as a complement to a currently used approach whereby a classic expert system, based on substructural rules, is used to assign chemicals to a given mode of toxic action for analog or QSAR selection.

Concurrent with establishment of topologically-based chemical similarity approaches, efforts have been initiated to generate global stereoelectronic parameters for chemicals within the industrial chemical inventory. Initial efforts have focused on specified subsets to assess computational needs to generate atomic and molecular parameters, which will be followed by an intensive computing exercise. Generation of these data will be combined with topological indices described above to re-evaluate chemical similarity, especially in the context of assessing the probability of metabolic transformations and the toxicity of reactive chemicals. Initial chemicals selected for study have included those currently in the ERL-Duluth ecotoxicity database and additional compounds for which reliable metabolic maps and rate constants are available. Using these chemical datasets, exploratory QSARs are being established to evaluate the nature of global and local parameters that are needed to identify reactive centers for predicting metabolism and toxicity.

Results and Future Objectives

Neural Networks: During the past year a prototype PC-based neural network, designed to classify chemical structures by modes of toxic action, was installed on the National Environmental Supercomputer. Modifications to the software are currently underway to improve computing efficiency. After completion of these software modifications, additional efforts will be undertaken to explore attenuation algorithms, to incorporate Monte-Carlo subsampling techniques, and further develop training algorithms that are consistent with evolving chemical similarity metrics. As these research findings progress, neural networks will be employed to characterize and assess the industrial chemical inventory in the context of ecotoxicological endpoints.

Chemical Reactivity: As previously discussed, the broad objective of this effort is to establish an efficient and mechanistically-relevant means to incorporate stere-

oelectronic parameters in the next generation of QSAR models. This effort is, in part, addressing the means whereby the three dimensional structures of chemicals currently in commerce (ca. 60,000 compounds) can be quantified and stored for subsequent modeling. During the past year, a subset of 50 chemicals have been used to examine techniques whereby approximated three dimensional structures can be ported to the National Environmental Supercomputer, for subsequent molecular structure optimization using Gaussian-92. This effort is establishing a procedure to generate molecular structures for large files of compounds and is providing estimates of the resources needed for large scale calculations.

In related efforts, studies have been undertaken to explore specific toxicological processes and associated chemical reactivity parameters to establish a mechanistically-based approach to screen compounds and stereoelectronic indices and thereby focus future three-dimensional calculations. Based on hypotheses concerning toxic mechanisms and metabolic activation pathways, several studies have been undertaken to initially explore the use of stereoelectronic descriptors to identify potentially reactive toxicants. Descriptors of soft electrophilicity and one electron reduction potentials have been calculated for a diverse group of aromatic compounds and used to discriminate the narcosis mode(s) of toxic action from mechanisms associated with covalent binding to soft nucleophiles in and oxidative stress, respectively. These studies are providing some insights into how a mechanistically-based strategy may be developed for selecting and using electronic indices in QSARs for biochemical and cellular toxicity.

---

[1] Steven Bradbury and Gilman Veith, U.S. EPA, Environmental Research Laboratory - Duluth, 6201 Congdon Blvd., Duluth, MN 55804.
[2] Ovanes Mekenyan, Lake Superior Research Institute, University of Wisconsin-Superior, Superior, WI.
[3] Robert Hunter, Natural Resources Research Institute, University of Minnesota-Duluth, Duluth, MN.
[4] Eric Anderson, Computer Sciences Corporation, Duluth, MN.

“

It is false dichotomy to think of nature and man. Mankind is that factor in nature which exhibits in its most intense form the plasticity of nature.

”

Alfred North Whitehead [1861-1947]
British philosopher

The Supercomputer in Medicine: Interpretation of Gamma Camera Composition [1, 2, 3]

Abstract

The efficacies of inhaled pharmacologic drugs in the prophylaxis and treatment of airway diseases could be improved if particles were selectively directed to appropriate sites. In the medical arena, planar gamma scintillation cameras may be employed to study factors affecting such particle deposition patterns within the lung. However, the value and versatility of such instruments are compromised by the limited resolution of their images. Specifically, it is not possible to determine the composition of their central (C) or large airway, intermediate (I), and peripheral (P) or small airway zones. Herein, an analytical model is presented *to assist the clinician in the systematic analysis and interpretation of gamma camera images.*

Using a Cray Y-MP supercomputer, human lung morphologies have been mapped to function as templates to be superimposed upon scans. The model is intended *to complement laboratory regimens* by providing a heretofore unavailable method to define the C, I, and P zones of the human lung on an airway generation-by-generation basis. A quantitative value can now be assigned to the degree of overlapping that exists in the images. For example, in the lung morphology consisting of 16,777,215 airways (total), the *C zone* itself may contain 1,608,246 airways of which 1,595,940, or more than 99%, are *alveolated* airways. By identifying composition our intent is to integrate the model into future aerosol therapy protocols and thereby assist procedures for targeted delivery of airborne pharmaceuticals.

Introduction

Numerous clinical studies have been performed in attempts to quantitate the deposition patterns of radiolabeled test aerosols and pharmacologic drugs as functions of: (1) aerosol characteristics (e.g., particle size and density, the degree of polydispersity) and (2) ventilatory parameters (e.g., tidal volume, breathing frequency, breath-hold time). The experimental investigations (Barnes, et al., 1978[5], Newman et al., 1991[6], Ruffin et al., 1978[7], and Smaldone et al., 1989[8]) have demonstrated that the efficacies of inhaled pharmacologic drugs could be enhanced when particles are targeted to regional areas of the lung ("large bronchial airways", "small central airways", etc.).

The benefits of such selective deposition processes may be further improved with the increased spatial resolution of deposition among the airways of the lung; that is, if drug delivery were to be attempted on a more focused basis, perhaps even targeting airway bifurcations.[9, 10] Attempts to relate deposition and response, indirectly, have been made via pulmonary function testing. However, direct observations of deposition patterns can be achieved via gamma scintillation cameras, and PET and SPECT instrumentation.

The clinical observations cited above may be explained, at least partially, in terms of the spatial dispersions of receptors and nerve endings within the lung. Their respective distributions could vary from being heterogeneous (e.g., localized) to homogeneous (e.g., uniform). Let us address the former case. The roles of

afferent neural pathways were studied by Karlsson et al[9]. They determined, for example, that upper tracheobronchial (TB) bifurcations were very sensitive to cough stimulation. The results suggested that rapidly adapting stretch receptors (RARs) could function as "cough receptors" commensurate with their relatively high concentrations in the upper TB airways and the superficial locations of RARs within the mucosa. This would be consistent with the documented fast blocking effects of topically applied and aerosolized anaesthetics.

Let us now consider the second case. The first mappings of β-receptor distributions within the lung were reported by Barnes et al.[5] Using autoradiography techniques, they detected a dense labeling of smooth muscle that was greater in the smaller bronchioles than the larger (i.e., cartilaginous) bronchi. The strong bronchodilator effect of β-agonists, therefore, might be attributed to the rather widespread (but not uniformly so) dispersion of β-receptors throughout the lung.

Gamma camera images cannot unambiguously portray the morphology of the lung. That is, the respective central (C), intermediate (I), and peripheral (P) zones will be of mixed composition. For example, deposition within the alveolated region will be superimposed upon that within the larger bronchial airways. This inherent and indeterminate character of such instruments is an unresolved problem which compromises their effectiveness and versatility. Therefore, *the mathematical model presented herein was developed expressly to complement laboratory investigations* by determining deposition on an airway generation-by-generation basis. The function of the model is twofold: (1) to aid the clinician in the analysis and interpretation of experimental data; and, thereby, (2) to assist in the design of future drug delivery tests.

Methods

In the medical arena, intersubject variabilities of patient lung morphologies and interlaboratory differences in experimental techniques must be acknowledged. In our work we have focused upon matters related to morphology. The major factors that inherently affect and complicate the use and intercomparison of gamma cameras are the spatial discrimination of an image and the radioactive dissipation which occurs prior to reaching an instrument's detector. These issues are addressed, albeit briefly, below.

Human Lung Morphology

The human adult lung morphology has been described by Soong et al.[12] (see their Table 2). The branching angle of a daughter airway is the deviation between the longitudinal axes of it and its parent. The branching angles measured by Horsfield et al.[13] have suggested a mean value of 70° to be appropriate.

In the human lung network described by Soong et al.[12] there are 2^n airways in each generation, n. For instance, there are 2^{23} = 8,388,608 alveolar sacs in generation n = 23. This numbering system is an important element in our later analyses.

Our ultimate goal was to superimpose a well-defined format such as used in the clinical environment (see the next subsection entitled *Geometric Definition of a Scan*, page 61) for the C, I, and P regions of gamma scans upon the lung. Therefore, it was incumbent upon us to develop an unambiguous technical representation of the lung's labyrinth of passages.

A model for the structure of the lung from Soong et al.[12] was made using the geometric parameters (lengths and branching angles). To define each lung, the pathway to each alveolar sac from the trachea was first calculated. Then, the

coordinates for the midpoint of each individual airway in the scheme were determined. Therefore, the lung became, literally, a matrix of points. Subsequently, the gamma camera scan format selected for use in our studies (i.e., that of McMaster University) was superimposed upon the matrix and the composition of the C, I, and P partitions determined directly by sheer itemization. The problem was thereby reduced to one of very extensive bookkeeping.

Geometric Definition of a Scan

To study aerosol deposition patterns with gamma cameras, it is essential that the boundaries of the lung be delineated. Agnew et al.[11] discussed two techniques used to characterize regional lung deposition using planar images. Newhouse et al.[14] and Smaldone et al.[8] defined the outer contours of the lung using xenon-127 and xenon-133, respectively, whereas Agnew et al.[15] used krypton-81*m*. The advantage of using xenon rather than krypton gas is that the former can be delivered to the whole of the ventilated lung volume. More recently, submicronic aerosols of ^{99m}Tc-DTPA have been used successfully to measure ventilation in both normal patients and those with a variety of chest diseases. The distribution of the aerosol in the lung is similar to xenon (Coates et al., 1983[16].) Alternatively, a perfusion scan can be performed, using a sub-clinical dose of ^{99m}Tc-MAA which will allow the lung edge to be defined as well as providing a correction for chest wall and tissue attenuation for technetium-99*m*.

In Figure 1, the methodologies employed at St. Joseph's Hospital, McMaster University, to examine gamma camera scans are displayed. The protocols have been described in other literature[7, 17, 18] and will be used herein to provide technical examples of how our analytical model may be applied in the medical arena. The perimeter of the lung is initially determined from the xenon-127 gas ventilation image. This template is then transposed to the aerosol deposition scan and the lung regions defined as follows. From the outer contour of the lung, a one (1.0) inch interval is offset to define the P zone. Subsequently, another one (1.0) inch concentric segment is delineated to identify the I zone. The balance of the image is the C zone. Radioactive markers placed on the chest can be used to identify the trachea.

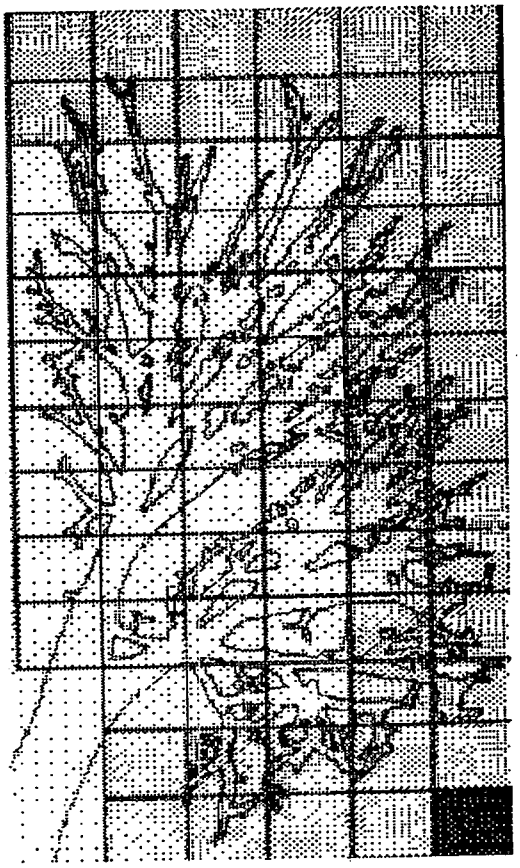


Figure 1: Diagrammatic outline of right lung illustrating the three-zone lung model. Squares 1.25 *cm* × 1.25 *cm* in original represent areas from which data were obtained. Dark squares represent the peripheral zone, light-shaded squares, intermediate zone; and white squares, perihilar zone; (reprinted with permission of Sanchis et al.) [17]

Results and Discussion

The mapping routine is quite simple in concept and theoretically straightforward, but requires numerous and repetitive computational procedures to track the positions of all airways within the lung. It is an ideal exercise, therefore, for the type of vector programming most efficient with a supercomputer.

In Figure 2 (page 63), the tracheobronchial tree of the adult human lung is presented. The systematic branching network of the $0 \leq n \leq 16$ generations is clearly demonstrated. The airways making up the balance of the lung (airway generations 17 through 23, inclusive) are not depicted simply because of the extent of overlapping that occurs. That is, those alveolated airways (if displayed) would be superimposed upon the currently visible structure and, as a result, no distinct airway network could be discerned. It should be noted that the TB network of Figure 2 is drawn to a 0.75 scale; that is, the corresponding generation-by-generation dimensions are systematically reduced by one quarter.

Table 1: The airway generation-by-generation composition of the central (C), intermediate (I), and peripheral (P) zones of a gamma camera image. The description is based upon the dimensions of an adult human lung (Soong, et. al.[12]).

Lung Generation	Partition of Scan		
	Region C	Region I	Region P
1	0	2	0
2	4	0	0
3	8	0	0
4	16	0	0
5	4	28	0
6	4	60	0
7	12	116	0
8	24	232	0
9	42	316	154
10	90	446	488
11	182	794	1072
12	372	1372	2352
13	754	2542	4896
14	1528	4964	9892
15	3054	9902	19812
16	6212	19530	39794
17	12440	38818	79814
18	25098	77148	159898
19	50412	154250	319626
20	100702	309416	638458
21	201338	619260	1276554
22	402072	1240374	2551858
23	803878	2483770	5100960
Cumulative	**1608246**	**4963340**	**10205628**
% of Lung	**9.59**	**29.58**	**60.83**

In Table 1, the thicknesses of the associated P and I concentric gamma camera zones were one (1.0) inch as prescribed in the subsection entitled *Geometric Definition of a Scan* (page 61). The data of Table 1 itemize the compositions of the C, I, and P components of the respective gamma camera scans. Specifically, those partitions are explicitly broken-down into their *constituent* elements. To our knowledge, our paper is the first to accomplish such an unambiguous spatial definition of the adult human lung.

The tabulated information is quite self-explanatory. For brevity, we shall make only a few comments to orient the reader. Regarding use of the results in the clinical laboratory, some critical points are the following. The C partition contains a total of 1,608,246 airways of which 1,595,940 are alveolated passages from generations $17 \leq n \leq 23$. Thus, more than 99% of the C airways are from the pulmonary compartment of the lung.

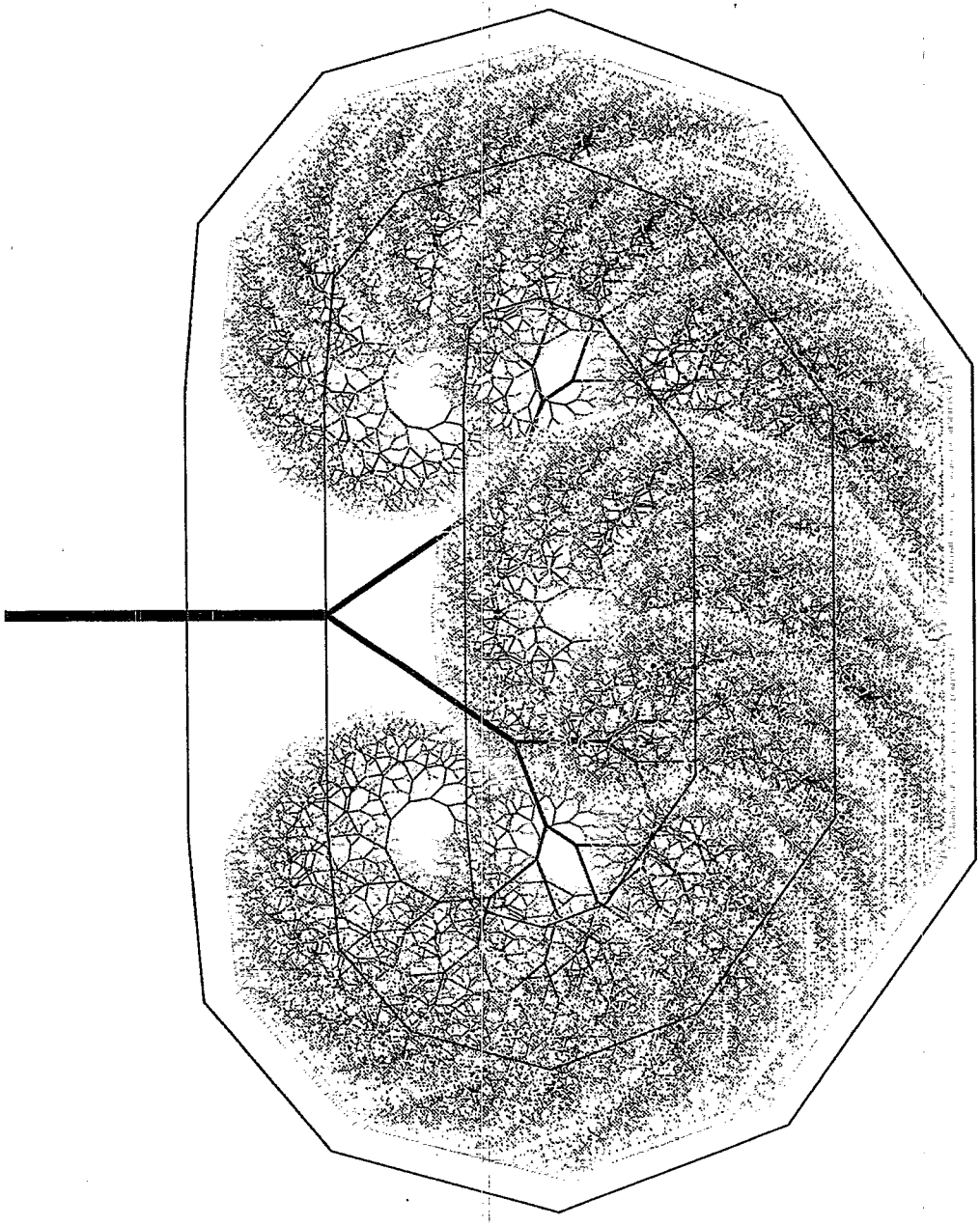


Figure 2: Mapping of the adult human lung achieved with the Cray Y-MP supercomputer. For clarity, only the tracheobronchial airways inclusive of the trachea (generation $n = 0$) through the terminal bronchioles (generation $n = 16$) are displayed. The outer contour and concentric regions of the mapping which correspond to the central (C), intermediate (I), and peripheral (P) zones of a gamma camera image are identified.

It may be prudent to view the data from Table 1 from another perspective. The total number of all alveolated airways (i.e., the entire pulmonary compartment of the human lung) is 1,6646,144; of that total, between 9-10% are in the C partition.

In summary, regarding the difficulties involved in standardizing such methodologies (i.e., the subsection entitled *Geometric Definition of a Scan*, page 61.) between laboratories, Agnew et al.[11] submitted a list of proposals to be incorporated into future investigations to facilitate interpretation of images and intercomparison of data. The resolution of such instruments, however, will still not permit deposition to be determined on an airway generation-by-generation basis. Therefore, our mathematical model was derived to complement such laboratory investigations.

Disclaimer

Although the research described in this article has been supported by the United States Environmental Protection Agency, it has not been subjected to Agency review and, therefore, does not necessarily reflect the views of the Agency, and no official endorsement should be inferred. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

---

[1] T. B. Martonen - Health Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711 and Division of Pulmonary Diseases, Department of Medicine, University of North Carolina, Chapel Hill, NC 27599.

[2] Y. Yang - Center for Environmental Medicine and Biology, University of North Carolina, Chapel Hill, NC 27599.

[3] M. Dolovich - Department of Nuclear Medicine, Health Sciences Centre - 1P, McMaster University, Hamilton, Ontario L8N 3Z5.

[4] Supported by funds provided by the U.S. Environmental Protection Agency under Collaborative Agreement CR817643 on Health Effects of Exposure to Air Pollutants through the Center for Environmental Medicine and Biology, University of North Carolina, Chapel Hill, NC.

[5] Barnes P.J., Basbaum C.B., Nadel J.A., Roberts J.M. Localization of β-adrenoreceptors in mammalian lung by light microscopic autoradiography, *Nature* 1978, 299:444-447.

[6] Newman S.P., Clark A.R., Talaee N., Clarke S.W. Lung deposition of 5 mg Intal from a pressurised metered dose inhaler assessed by radiotracer technique. *Inter. J. Pharm.* 1991; 74:203-208.

[7] Ruffin R.E., Dolovich M.B., Wolff R.K., Newhouse M.T.. The effects of preferential deposition of histamine in the human airway. *Am. Rev. Respir. Dis.* 1978; 117:485-492.

[8] Smaldone G.C., Walser L., Perry R.J., Ilowite J.S., Bennet W.D., Greco M. Generation and administration of aerosols for medical and physiological research studies. *J. Aerosol Med.* 1989; 2:81-87.

[9] Karlsson J.A., Sant'Ambrogio G., Widdicombe J. Afferent neural pathways in cough and reflex bronchoconstriction, *J. Appl. Physiol.* 1988; 65:1007-1023.

[10] Martonen T.B. Aerosol therapy implications of particle deposition patterns in simulated human airways. *J. Aerosol Med.* 1991; 4:25-40.

[11] Agnew J.E. Characterizing lung aerosol penetration. *J. Aerosol Med.* 1991; 4:237-249.

[12] Soong T.T., Nicolaides P., Yu C.P., Soong S.C. A statistical description of the human tracheobronchial tree geometry. *Respir. Physio.* 1979; 37:161-172.

[13] Horsfield K., Dart G., Olson D.E., Filley G.F., Cumming G. Models of the human bronchial tree. *J. Appl. Physiol.* 1971; 31:207-217.

[14] Newhouse M.T., Wright F.J., Ingham G.K., Archer N.P., Hughes L.B., Hopkins O.L. Use of scintillation camera and 135xenon for study of topographic pulmonary function, *Respir. Physio.* 1968; 4:141-153.

[15] Agnew J.E., Lopez-Vidriero M.T., Pavia D., Clarke S.W. Functional small airways defense in symptomless cigarette smokers. *Thorax* 1986; 41:524-530.

[16] Coates G., Dolovich M., Newhouse M.T. A comparison of submicronic technetium aerosol with xenon-127 for ventilation studies. In: Raynaud C ed. *Nuclear Medicine and Biology: Proceedings of Third World Congress of Nuclear Medicine and Biology, Vol 11.* Elmsford, New York: Pergamon Press, 1983:91-96.

[17] Sanchis J., Dolovich M., Chalmers R., Newhouse M. Quantitation of regional aerosol clearance in the normal lung. *J. Appl Physiol.* 1972; 33:757-762.

[18] Dolovich M.B., Sanchis J., Rossman C., Newhouse M.T. Aerosol penetrance: a sensitive index of peripheral airways obstruction. *J. Appl. Physiol.* 1976; 40:468-471.

Regional Oxidant Model Sensitivity Analysis [1]

EPA Research Objectives

The use of regional scale (~1000-2000 km) air quality simulation models, such as EPA's Regional Oxidant Model (ROM), for analysis of regional air quality and for assessing potential future air quality control strategies has become more common over the last decade. The specific objective of this work is to assess the sensitivity of ROM ozone predictions to key input variables. The sensitivity case outlined here is that of mobile source (automobile) emissions.

Overview of Project

There are several classes of variables that are required to drive air quality models such as the ROM. These classes include meteorological and emissions variables, the principal forcing functions, as well as specification of initial and boundary conditions and terrain and land-use data. There are uncertainties associated with the specification of every variable within each class; generally speaking, the emissions variables tend to be more uncertain than the others. The concentration predictions from the air quality model will be sensitive, to one degree or another, to each of the input variables. We are most concerned with those variables which contain the greatest uncertainty in their specification, and for which the model is also quite sensitive. To fully understand the model response to these variables, we must first characterize the degree of model sensitivity. The Atmospheric Research and Exposure Assessment Laboratory (AREAL) of EPA/ORD is attempting to characterize the ROM's sensitivity to a subset of chemical and physical input variables to the model by direct simulation using perturbations of the input variables. We focus our discussion here on one of the sensitivity tests in our series.

Background and Approach

Photochemical oxidant pollution is a frequent and widespread phenomenon across the eastern U.S. in the warm-weather months of the year. Three-dimensional numerical grid models, such as the ROM, are being developed, tested, applied, and evaluated to handle the air quality simulation needs of the pollution problem. These computationally demanding models simulate all of the necessary physical and chemical processes responsible for regional photochemical smog. Once evaluated, the models are used by EPA to assess potential future emissions control strategy options to reduce the levels of high ozone concentrations across the U.S.

As previously discussed, our confidence level in the results of the ROM is, in part, a function of how well we can characterize the model's sensitivity to key input variables. As an illustration of this process, we focus here on the ROM's sensitivity to the mobile source component of the source emissions inventory. Automobiles emit significant quantities of hydrocarbons and nitrogen oxides, precursor chemicals to the formation of ozone and photochemical smog. Subjectively, we expect the model to be responsive to this emissions component. To perform this test, we remove the automobile emissions from the inventory feeding the ROM and compare the results of an air quality model simulation with the altered inventory with that using a "base case" inventory. Specifically, for this test all emissions from light duty cars and trucks were set to zero, as well as associated emissions from gasoline marketing and storage. The emissions changes were

made in the U.S. portion of the modeling domain only.

Scientific Results and Relevance to EPA Mission

The meteorological data used for these simulations were taken from the period July 2-10, 1988, a time period characterized by high heat and air mass stagnation over the eastern U.S. This entire time period was simulated with the ROM, using the model's three vertical layers extending through the earth's boundary and cloud layers, and with 18.5-km horizontal resolution. We present here the results of the model simulations using the maximum hourly ozone concentration predicted by the ROM in each of the model's lowest level grid cells over the July 2-10 period. Figure 1, page 68, displays the base case ozone results using the original emissions inventory. Figure 2, page 69, displays the results of the simulation using the inventory without the mobile source emissions, and Figure 3, page 70, shows the concentration differences between the simulations.

Inspection of these figures reveals that the removal of the mobile source emissions component does indeed dramatically reduce the ozone concentrations over the eastern U.S. Regionally, concentrations drop by 5 ppb to greater than 50 ppb over their base case values, with the greatest decreases occurring over and downwind of the urban and industrialized portions of the model domain. In Figure 3, note especially the concentration differences throughout the eastern seaboard region, industrial Midwest, and Great Lakes region. Downwind influences of the emissions change are seen over the Atlantic Ocean and into interior New England as well. It is also worth noting that the removal of the mobile source emissions did not fully alleviate the ozone problem in the model domain. Figure 2, page 69, reveals estimated ozone concentration areas still in excess of the 120 ppb Federal ambient air quality standard to exist in portions of the domain.

Implications of the result of this study indicate that the emissions reductions incurred through programs such as the Federal Motor Vehicle Control Program, which aim to reduce mobile source emissions of hydrocarbons and nitrogen oxides, should serve to reduce ozone levels, but the gains may be modest in many areas. Other sources of ozone precursors, such as naturally-emitted hydrocarbons and nitrogen oxides from industrial and power generation facilities can also combine to form significant amounts of ozone. Therefore, while the motor vehicle emissions control program is likely to be beneficial, other source categories for targeted control should also be considered to reduce high ozone concentrations levels in the eastern U.S.

Advantages of Using NESC Cray

Earlier ROM simulations had been performed on scalar mainframes, such as the IBM-3090. Resource limitations on these machines prevented our simulating the entire Eastern U.S. with the ROM. Typically smaller domains were used, such as the Northeastern U.S. only, about one-fourth the size of the current model domain depicted in Figures 1, 2, and 3. With the availability of the NESC CRAY, and the vector optimization of the ROM code, we are now able to efficiently simulate the larger domain, providing a more comprehensive picture of estimated air quality throughout eastern North America.

Future Plans

We plan to continue our series of sensitivity tests with the ROM, including more emissions sensitivities as well as a series of tests for meteorological sensitivities. Results of these tests will help interpret

the significance of future air quality modeling analyses with ROM and other regional air quality models.

Publications

Roselle, S.J., T.E. Pierce, and K.L. Schere. *The Sensitivity of Regional Ozone Modeling to Biogenic Hydrocarbons.* Journal of Geophysical Research, Vol. 96, pp. 7371-7394.

Roselle, S.J., K.L. Schere, and S.H. Chu. *Estimates of Ozone Response to Various Combinations of NO_x and VOC Emission Reductions in the Eastern United States.* Proceedings of 1992 Quadrennial Ozone Symposium, June 1992.

---

[1] Thomas E. Pierce and Kenneth L. Schere, U.S. EPA, Office of Research and Development, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC.

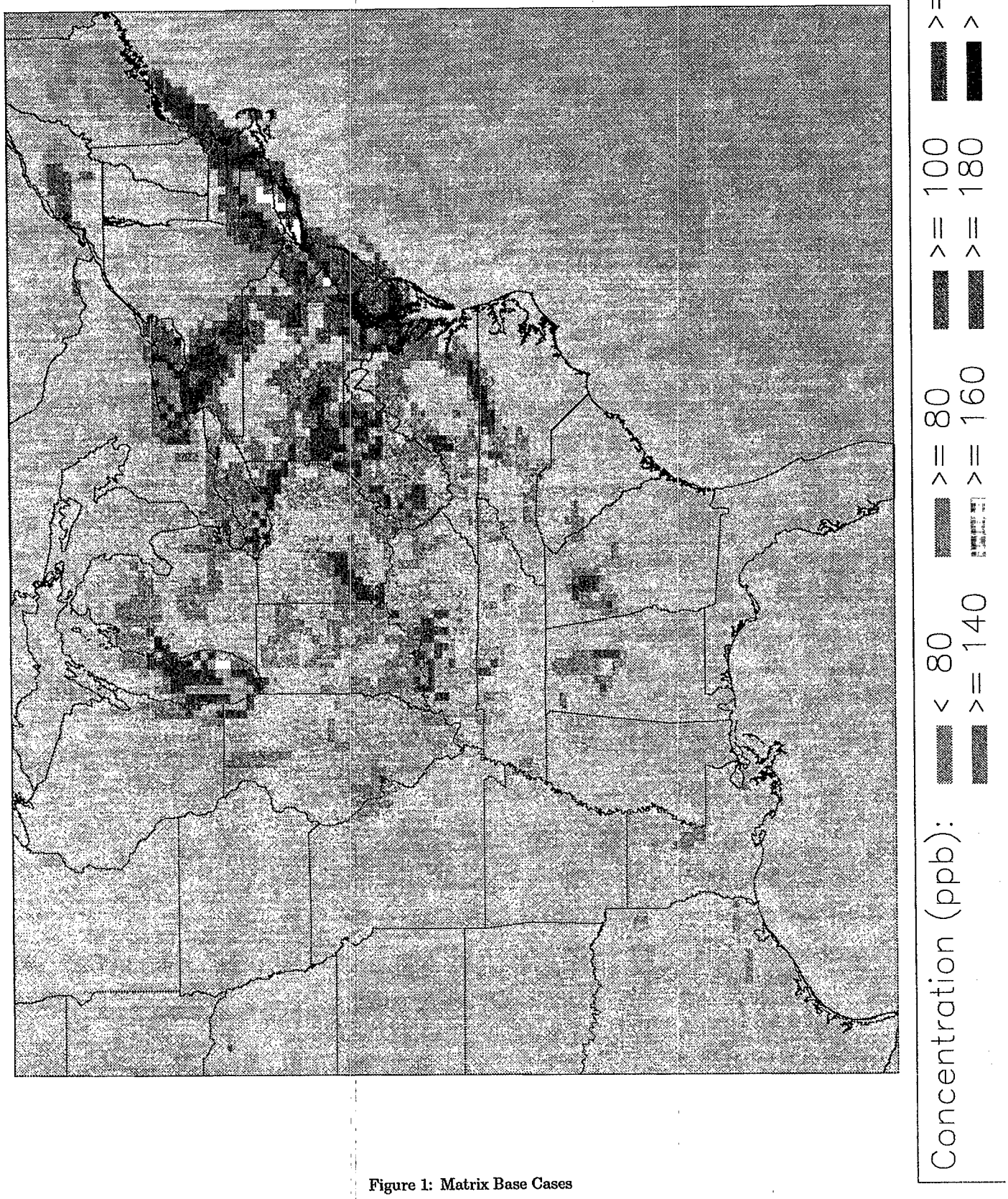


Figure 1: Matrix Base Cases

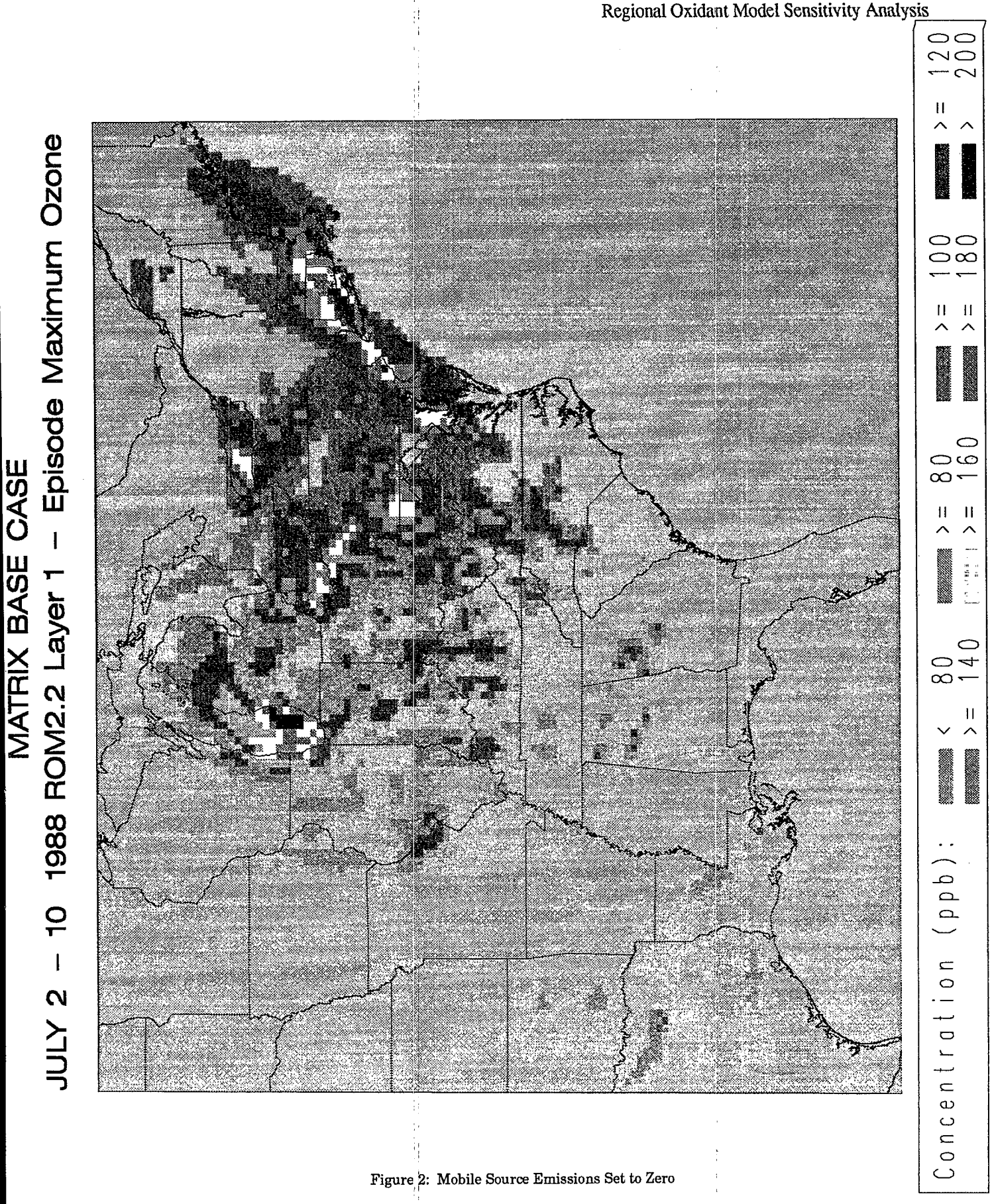


Figure 2: Mobile Source Emissions Set to Zero

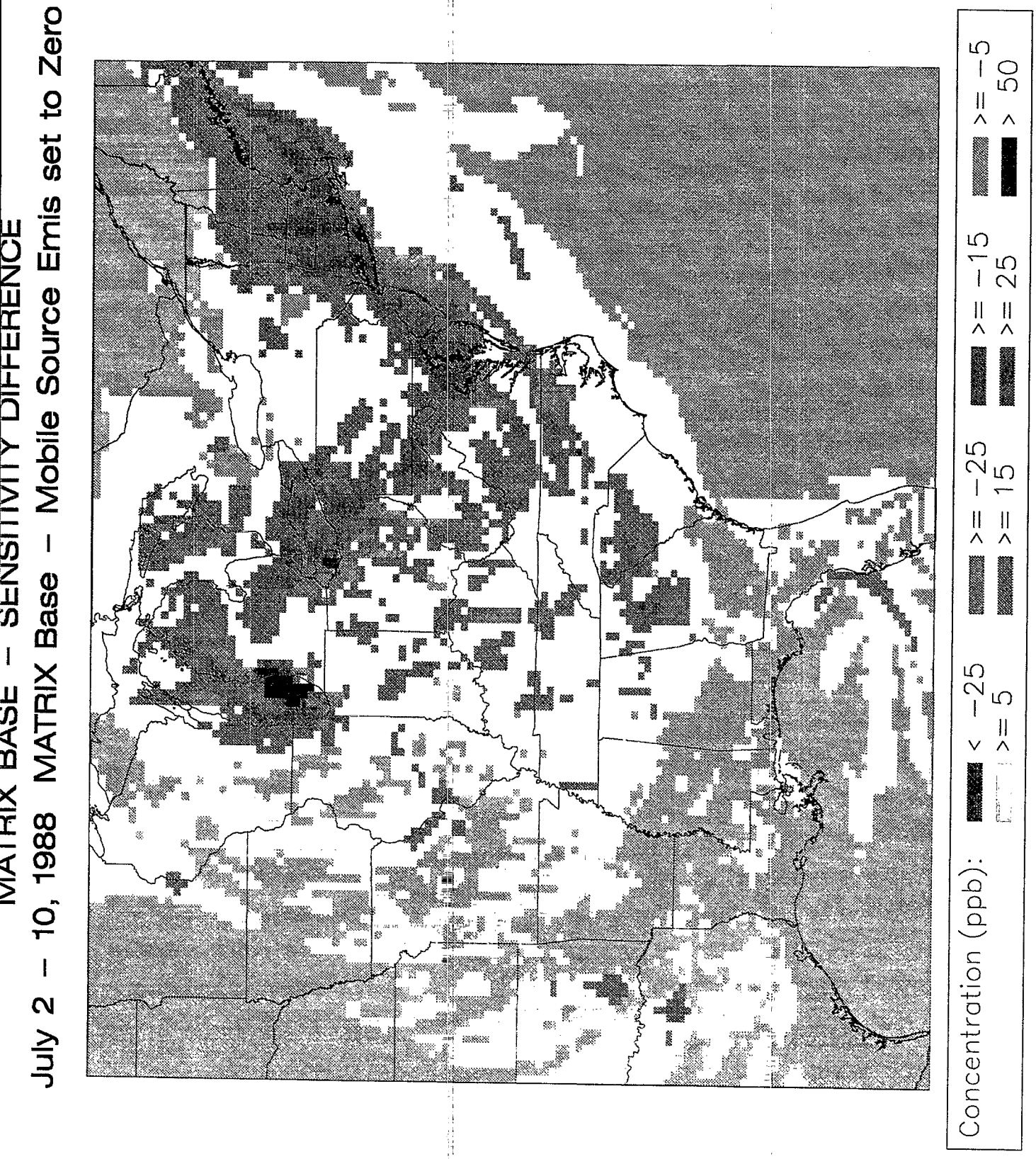


Figure 3: Matrix Base - Sensitivity Difference

Meteorological and Photochemical Grid Modeling in Urban Domains [1]

Abstract

Photochemical grid modeling is essential to predicting the impacts of prescribed emissions changes on predicted ozone concentrations and other secondary pollutant species. An intensive effort has been underway to develop improved meteorological model algorithms and an updated version of the Urban Airshed Model (UAM) for use in urban-scale ozone applications. Simulations with an efficient, computationally fast diagnostic meteorological model and a computationally rigorous dynamic meteorological model, which numerically solves the set of governing equations for atmospheric motion, temperature, and moisture, are being performed to generate meteorological data files for use by an updated UAM model. The dynamic model code also features a four-dimensional data assimilation technique which continuously adjusts modeled values toward observations during the simulation period. The updated UAM model is being exercised with various meteorological and emissions inputs for selected urban domains in order to investigate the sensitivity of predicted spatial pattern and peak concentrations due to different inputs and to evaluate the model results against measurements.

Research Objectives

An important goal is to perform simulations with the meteorological models and the research version of the UAM photochemical grid model code in order to evaluate the modeled concentrations against experimental measurements in urban environments. Test simulations with the models are being conducted in various urban domains where complex airflows often occur due to complicated topography and/or land-water boundaries. Model intercomparisons of results from the updated research UAM and the standard UAM model is to provide useful information about the effects of new scientific algorithms and revisions in updated UAM on predicted concentrations. Numerous UAM model simulations with variations in anthropogenic and biogenic emission inventories are being designed to assess the impact on peak ozone concentrations within urban domains.

Overview and Relevance of the Project

This research and development project has been underway to update scientific components and the software of the current Urban Airshed Model (UAM) system, the EPA's preferred regulatory ozone model. The UAM system is receiving widespread use in ozone attainment demonstrations in many urban areas in order to meet mandates in the 1990 Clean Air Act Amendments. The Urban Airshed Model (UAM) is an Eulerian photochemical grid model which treats the transport, dispersion, chemical transformation, and dry removal processes impacting ozone and other pollutant species in urban areas. Key input data sets for a model simulation over a typical two day period include gridded input data sets on initial and boundary conditions, hourly wind flow fields and related meteorological parameters, and emissions from point and area sources of nitrogen oxides (NOx) and volatile organic compounds (VOCs).

Due to the crucial role of the three-dimensional wind flow and other

meteorological parameter fields on pollutant concentrations and their spatial patterns, considerable effort has been expended to design and to test updated meteorological algorithms that can provide compatible input data files for the UAM. Consequently, both diagnostic and dynamic meteorological modeling approaches are being applied in order to assess their strengths and weaknesses in different urban environments. While a diagnostic model, relying on available observations and interpolative techniques to generate gridded data, is computationally fast, complicated circulations on smaller spatial and temporal scales cannot be resolved. In contrast, a dynamic meteorological model is capable of simulating physically-realistic land-sea breeze circulations and terrain-induced flows. However, the short computational time step necessitates much longer CPU time and results are not always expected to reproduce observations. Therefore, a four-dimensional data assimilation (FDDA) technique requiring little additional computational time has been incorporated into the dynamic model code in order to provide more accurate winds at grids cells near data sites and produce more realistic flows in data-sparse areas. Simulations with and without the FDDA feature are needed to assess key variables involved in the approach for weighting the observations and modeled values.

Accomplishments and Preliminary Results

Test simulations with both meteorological model codes are successfully being performed on the NESC Cray system in order to generate meteorological data files for use in simulations with the updated UAM model code. Urban domains being modeled include the greater New York City metropolitan area with limited observational data and the Los Angeles basin where numerous measurements were obtained during an intensive experimental field study. An assessment of differences in wind fields and other meteorological parameters between these models and analysis of the impact with different meteorological input data on simulated spatial patterns and peak concentration of ozone are underway.

Simulation results of both meteorological models from the Los Angeles domain reveal the existence of a sea breeze from the Pacific Ocean developing during daytime hours which transports pollutants inland toward mountain ridges and through gaps surrounding the urban basin. A statistical evaluation of simulated pollutant concentrations from the updated UAM code versus measured concentrations at sites throughout the urban domain is underway to assess model performance using different meteorological model inputs.

A series of UAM sensitivity test simulations are also underway for a case study day in the New York City domain to investigate the impact on peak ozone concentrations from different meteorological inputs and algorithm changes. Initial results with a new horizontal advection scheme designed to reduce numerical diffusion in the updated UAM indicate a somewhat narrower and more elongated urban ozone plume downwind of the New York metropolitan area than produced by the existing horizontal advection technique in the standard UAM code.

Future Plans

Numerous simulations are being planned with the updated UAM to investigate differences in model results due to prescribed changes in key input parameters, and a series of model simulations with selected variations in emissions are anticipated in order to assess the effect on ozone concentrations and spatial patterns. Additional test simulations with different wind fields,

mixing height fields, and vertical eddy diffusivity coefficients are expected to yield valuable information about the sensitivity of modeled ozone concentrations.

---

[1] James M. Godowitch, Principal Investigator (JUG), U.S. EPA Office of Research and Development, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC.

“

Physics does not change the nature of the world it studies, and no science of behavior can change the essential nature of man, even though both sciences yield technologies with a vast power to manipulate their subject matters.

”

B[urrhus] F[rederic] Skinner [1904-]
Cumulative Record [third edition, 1972], ch. 5

Regional Particulate Modeling [1,2]

Research Objective

The objective of this project is to investigate the behavior of atmospheric aerosol particles on the regional scale. Such particles result from human activity as well as natural processes. A major portion of the particles resulting from human activity are acidic sulfates formed in the atmosphere from the chemical transformation of sulfur dioxide, a by-product of fossil fuel combustion. Other anthropogenic particles result from the chemical transformation of reactive organic gases (ROG's) and oxides of nitrogen.

Approach

A three-dimensional air quality model (RADM) has been extended to include the chemistry and dynamics of aerosol particles. The model has a full photochemical mechanism which transforms emitted compounds from the 1985 NAPAP emissions inventory into atmospheric oxidants and particle forming species, such as sulfates, nitrates and organics. The aerosol dynamics and chemistry codes explicitly recognize processes such as particle growth and coagulation, as well as the equilibration with atmospheric relative humidity. Particle size distributions are treated in two ranges or modes, both of which are log normal. The mean diameters and geometric standard deviations of these two modes are time-dependent variables.

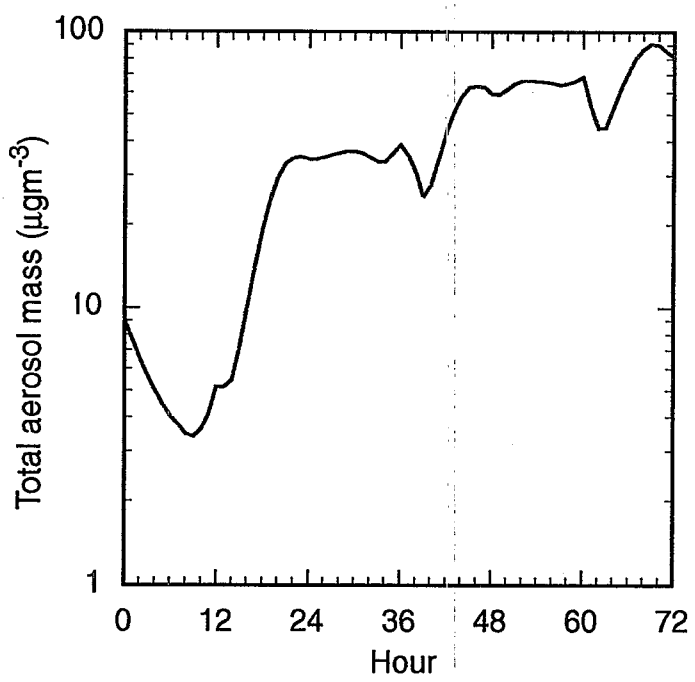


Figure 1: Total aerosol mass (μgm^{-3}) for a high emission grid cell

Accomplishment Description

The behavior of sulfate aerosol particles has been studied for cloud-free conditions. Three sensitivity studies for a typical summer 72 hour meteorological episode have been conducted. The first was a base case using the full inventory. The second used the same inventory but omitted

ammonia. The third used the full inventory for all species except sulfur which was reduced by a factor of two. These studies indicated that the model behaved in a qualitatively correct manner. As expected the water content of the particles is inversely related to the degree of neutralization of the acidity by ammonia. Thus in general, the more acidic the particles are, the higher the water content, and the larger the particles become.

Significance

Aerosol particles are of great interest to EPA for several reasons. Major reasons are the health effect of inhaled particles, the role of these particles in acidic deposition to sensitive ecosystems, and the reduction in visibility by the interaction of the particles with light. Related to this last item is the role of aerosols in the global climatic system. The model can be used to investigate the impact of proposed control scenarios required under the 1990 Clean Air Act Amendments upon aerosol particle behavior.

Future Plans

The model is constantly evolving as new features are added. Current plans call for the addition of cloud interactions with particles during FY1994. Semi-volatile organic compounds will be added in FY1995, thus allowing the model to be used to study toxic species. Extensive model evaluation is planned using existing and future field study data sets.

---

[1] Francis S. Binkowski, Principal Investigator - U.S. EPA, Office of Research and Development, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC.

[2] Uma Shankar, Co-Investigator - MCNC, Research Triangle Park, NC.

Visualizing Environmental Data at the EPA[1,2,3]

Overview

Environmental visualization aims to make visible the unseen. This helps researchers and policymakers alike. Here we examine the unique issues and problems involved.

Scientists at the U.S. Environmental Protection Agency pursue a wide range of research interests. One group develops models for the transport and deposition of airborne pollutants. Policy makers use this information to develop control strategies for managing air pollution, such as the Clean Air Act. Another group evaluates the positions of air quality monitoring sites with respect to the distribution of pollutants. Other researchers collaborate with investigators at NASA Langley and Lawrence Livermore National Labs to examine global climate change. Still others study water quality and sedimentation in the Great Lakes region (see Figure 1), electrical properties of carcinogens (see Figure 2, page 78), subsurface contamination of waste disposal sites, and the air flow through and around buildings. In this article we describe some of these research efforts.

Environmental Data

Environmental data sets are large, up to 4 Mbytes for a single time step. They often contain many time steps, representing changing conditions. For example, one data set consists of a set of ozone levels across the world for each day of a 13-year period. Another data set shows pollutant concentrations for each hour of a 15-day period. Still another data set describes lake sedimentation levels over the course of a storm.

Data sets arise from a wide range of sources, including computationally intensive models, atmospheric monitoring stations, satellites, laboratory experiments, and meteorological records. Data formats also vary, compounding the handling problem. Data can occupy either a 2D or 3D spatial domain, and we might need both surface and volume representation techniques to handle them.

A major problem afflicts environmental data: it is seldom regularly gridded, yet visualization software expects regular grids. For example, water quality data might come on a curvilinear grid from monitoring sites at points scattered irregularly across the nation.

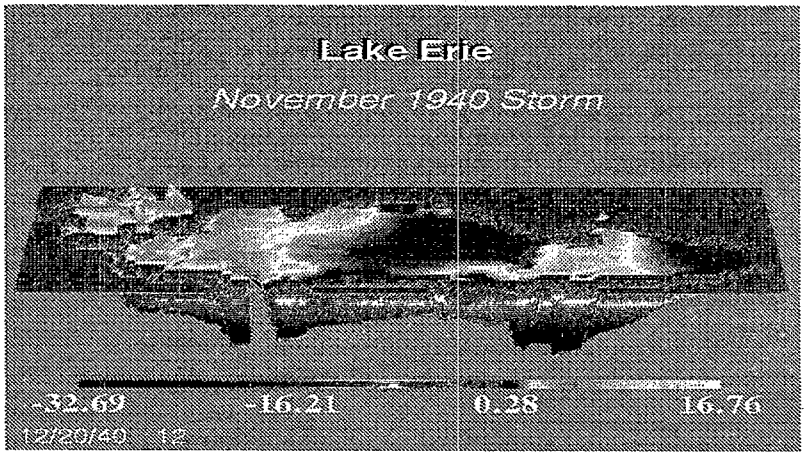


Figure 1. Environmental researchers study water quality using visualization techniques. This image shows sediment concentrations in Lake Erie for 1940 storm data obtained from a collaboration between the University of California at Santa Barbara and U.S. EPA Large Lakes Research Station.

Moreover, pollutant totals might be associated with county-sized areas of space.

Visualization Requirements

Environmental visualization attempts to represent abstract environmental data using concrete visual metaphors. Several factors influence the choice of representation: type of data, relationships among different parts of the data, placement of data in a spatial and temporal context, and interpretation of the data.

Environmental processes can be complex, involving many chemical species, atmospheric conditions, and geographical factors. Scientists frequently want to see multiple data sets or data variables displayed in the same visualization. For example, the Regional Oxidant Model (described below) calculates airborne pollutant transport at three different atmospheric levels. Sometimes scientists want to see those three layers displayed together.

To validate the model, other visualizations display data produced by environmental models and data collected by monitoring stations together. Data sets displayed together can have different sources, different data types, and different internal coordinate systems. Sometimes, we must first register the various data sets so that we can represent them in a single display coordinate system. Unfortunately, these time-intensive and error-prone procedures often dominate the real applications.

Adding information to place the data in its spatial and temporal context can make visualization of environmental data more valuable. To add spatial context, for example, we can take pollutant levels across the U.S. and overlay them with a map showing state borders (see Figure 3, page 80). We can show fluid flow through and around a building in conjunction with a simple representation of the building.We can show the electrical charge distribution of a molecule relative to the atoms and bonds of the molecule. These spatial landmarks, while not part of the data itself, provide valuable cues for interpretation. We can provide temporal context by showing the progression of a distribution over a time series. Showing a time series, rather than a single frame, emphasizes the development and progression of environmental processes. Labeling a visualization with its date and time also provides temporal context.

Configuration, General Strategies

The EPA uses a heterogeneous computer architecture for scientific computing. It supports a range of platforms and tools, including workstations and networked supercomputers.

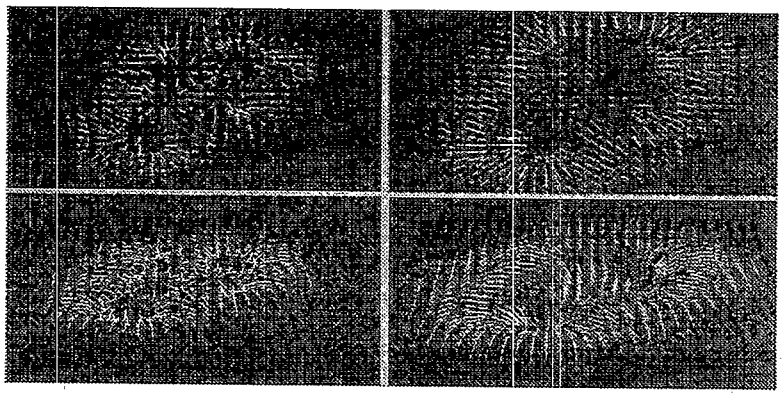


Figure 2: U.S. EPA environmental researchers evaluate molecular properties of carcinogens from pollution sources. This images shows electric field vectors for the benzopyrene molecule. This is based on data from the U.S. EPA Health Effects Research Laboratory.

The Visualization Centers software configuration is similarly heterogeneous, using a wide range of tools and packages on the various platforms. Despite the desirability of agency-wide standardization on a single visualization environment, none of the tools we use provides all the functionality we need. For example, while one tool provides photorealistic rendering and specialized functionality for generic gridded data sets, our environmental researchers find its script-driven environment unnatural. Another package offers interactive control and a variety of features useful for representing computational fluid dynamics data, but it runs only on one platform and requires researchers to convert data into a specific format for visualization.

The visual programming and multiple-platform capabilities of tool kits are attractive for the EPA's heterogeneous hardware configuration. However, for a tool kit to provide the required functionalities, the EPA visualization team needs to develop, support, and inventory customized modules that support environmental research requirements.

Through necessity or by design we have adopted a few general strategies for visualization. First, we introduce scientists to the possibilities of visualization, and then we let their needs and vision drive the research product. Second, we push software packages beyond their intended limits, then supplement and extend them with customized modules. Third, we seek a balance between image quality and interactive control. We can interactively manipulate small- to medium-sized data sets. Fourth, we support the EPA's nationwide heterogeneous hardware and software environment.

Regional Oxidant Model

The Regional Oxidant Model is an air quality model designed to examine the transport and deposition of airborne pollutants. EPA scientists use the model to develop control strategies for the Clean Air Act. The Regional Oxidant Model simulates most of the significant chemical and physical processes responsible for photochemical production of ozone over a 100-kilometer domain for episodes lasting 15 days. These processes involve

- horizontal transport.
- atmospheric chemistry and subgrid-scale chemical processes.
- nighttime wind shear and turbulence associated with the low-level nocturnal jet.
- effects of cumulus clouds on vertical mass transport and photochemical reaction rates.
- mesoscale vertical motions included by terrain and the large-scaleflow.
- terrain effects on advection, diffusion, and deposition.
- emissions of natural and anthropogenic ozone precursors.
- dry deposition.

Three separate atmospheric layers are associated with every Regional Oxidant Model run, and the output provides the geographic distribution of three chemical species (ozone, NOX, and ROG). A typical geographic domain of the model is the northeast corridor of the United States (see Figure 4, page 80).

We use visualization techniques to examine the inputs generated for the Regional Oxidant Model, such as sequences of wind inputs (see Figure 5, page 82). We animate wind vector representations over time. This visualization improves researchers understanding of the wind inputs and results in modifications to the respective model algorithms.

We have also applied volume visualization techniques to Regional Oxidant Model outputs. Using a splatter technique originally developed by Westover,[4] we generated a simultaneous visualization of the three atmospheric layers of the model (see

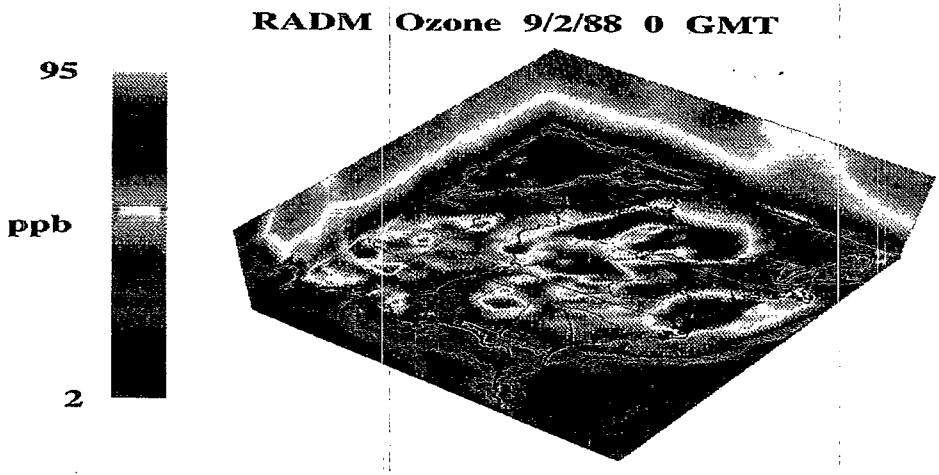


Figure 3. Cutting planes visualization for ozone concentration values of the Regional Acid Deposition Model, which is a nonlinear air quality model used for assessing acid rain impacts. The U.S. EPA Atmospheric Research and Exposure Assessment Laboratory supplied the data.

Figure 6, page 82). The graphical representation displays cloud-like structures for each of the chemical species. This provides a new way for atmospheric researchers to view their data. Interestingly enough, our scientists did not find the volume visualization effort insightful. The researchers had trouble comprehending the interaction between the three layers depicted. They requested that we return to the surface modeling approach to visualize the individual Regional Oxidant Model layers for the respective chemical species.

We developed movie sequences of 15-day episodes of Regional Oxidant Model ozone concentrations, incorporating discrete color mapping display and titling with time and data stamping. We also examined output with standard visualization tool kits and extended these tool kits with customized modules to support environmental researchers visualization requirements.

Total Ozone Mapping Spectrometer

In a cooperative project between the EPA and NASA Langley Research Center, the EPA Visualization Center has visualized data from the Total Ozone Mapping Spectrometer (TOMS) satellite. The

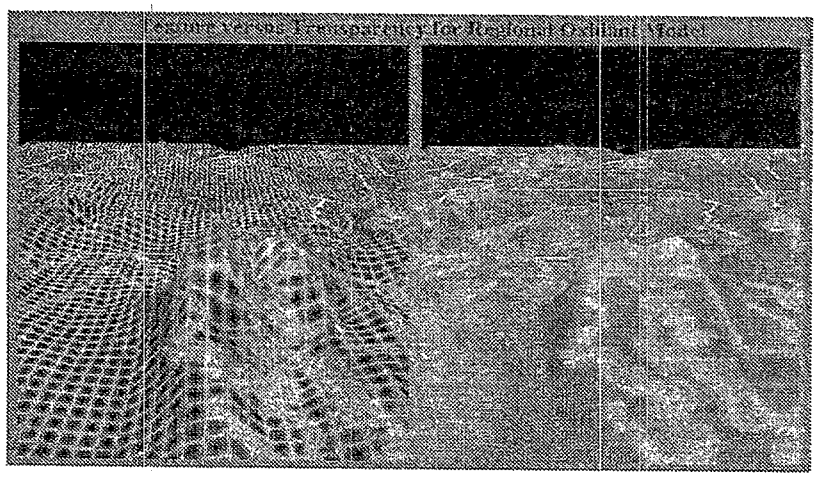


Figure 4. Texture (left side) and transparency (right side) mapping for ozone concentrations of the Regional Oxidant Model, northeastern domain. Researchers use this model to develop control strategies for the Clean Air Act. Data came from the U.S. EPA Atmospheric Research and Exposure Assessment Laboratory.

TOMS instrument measures total atmospheric ozone by analysis of back-scattered solar radiation across several ultraviolet bands. The instrument is aboard a south-north, sun-synchronous, polar-orbiting satellite and performs more than 200,000 measurements each day over the entire globe. The measurements are then processed into 1- degree latitude by 1.25-degree longitude cells (see Figure 7, page 82). The purpose of the EPA portion of the study is to compare the results from the agency's nonlinear models (like the Regional Oxidant Model) with the measured data available from TOMS.

Logistic problems included format conversions and projections to various coordinate systems. The EPA Visualization Center staff worked extensively on writing code to support converting the TOMS data into a 3D visual display. We also encountered issues associated with handling missing data values. Large-scale dropouts in the TOMS data resulted from substation failures during the satellite tracking and data collection process. The computational process for handling the smoothing of the data resulted in potential artifacts, which environmental researchers reviewed and critiqued. Our visualization efforts focused on developing an adaptive filtering process to handle variations in missing data regions, data confidence, and sampling density issues (see Figure 8, page 83).

The large size of the TOMS data set poses other concerns. Interactive viewing of the daily change of global ozone for 365 days requires computationally intensive hardware currently not available on standard desktop visualization workstations. This concern highlights our need to address metacomputing for environmental data sets.

Unsolved Problems

There remain a number of unsolved systems problems with the visualization of environmental data. Our current workstations provide insufficient computational power to interactively step through the time steps of large 3D data sets. Large data sets likewise present storage and access difficulties. The need to support visualization on multiple platforms has not yet been completely addressed. Similarly, we have just begun to address the issues associated with remote access. Remote access issues appear in our collaborations with scientists at remote sites and in our use of the National Environmental Supercomputing Center (NESC) in Bay City, Michigan.[5]

A number of software limitations and challenges also remain. Foremost is the need for a general-purpose data import tool for visualizing multiple data formats. Tool kit data conversion utilities currently available begin to address the problem, but stop short of a complete solution. For example, our scientists often run models on a mainframe or supercomputer and wish to view the model output. Data conversion routines that run on workstations often cannot handle large binary files generated on a supercomputer.

Data collected from atmospheric monitoring sites causes other problems. These sites might be scattered irregularly across the area of interest. The data collected does not exhibit the grid structure (either regular or irregular) expected by standard tools (see Figure 9, page 83). Finally, some data associates values with irregularly shaped areas (such as counties) and requires the functionality found in GIS packages. Until the tool kit data conversion utilities develop sufficient power to handle all the data used by scientists, we will continue writing conversion programs for each new type of data.

As the EPA Visualization Center continues to support agency visualization needs, we will develop new approaches and tools. The installation of our first high-performance computer at the NESC has

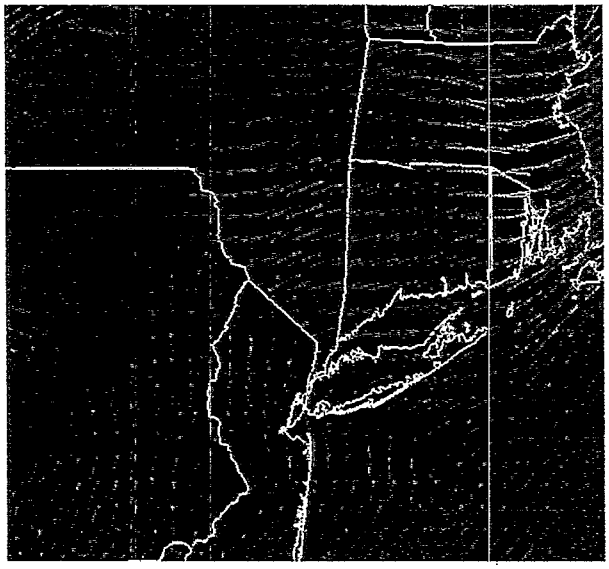


Figure 5. Visualization of wind vector inputs to the Regional Oxidant Model. Meteorologists at the U.S. EPA use visualization to examine inputs to the Regional Oxidant Model before executing it on a high-performance computing platform. Data came from the U.S. EPA Atmospheric Research and Exposure Assessment Laboratory.

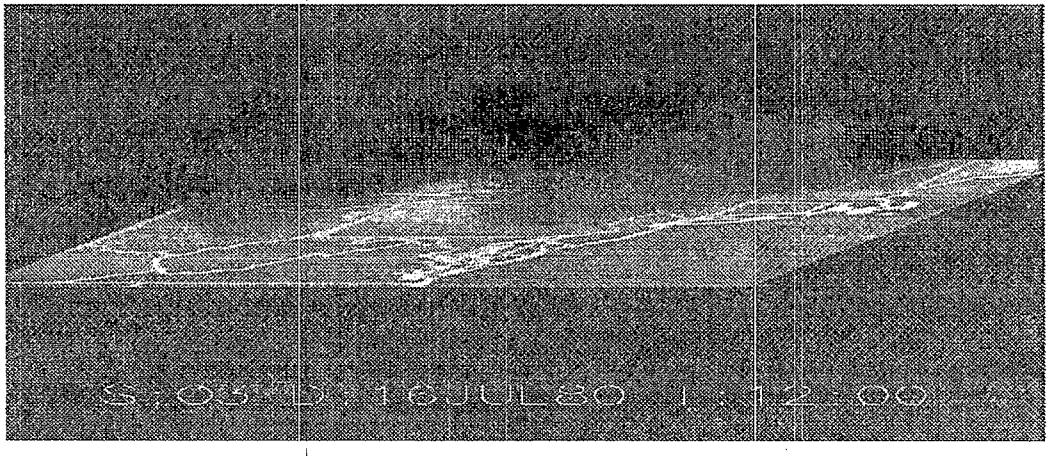
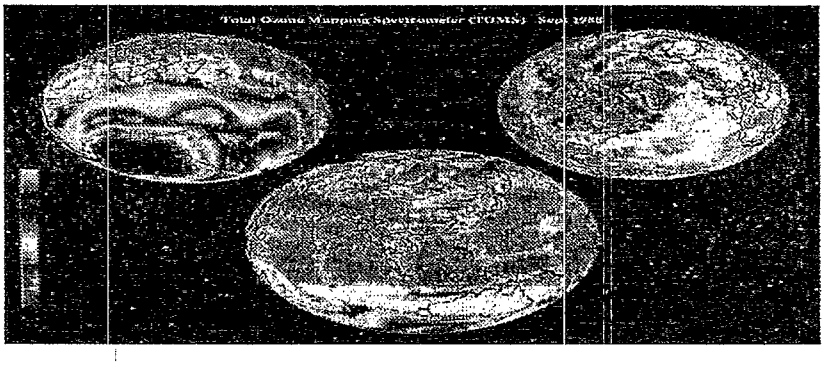


Figure 6. This image shows three layers of the Regional Oxidant Model using a splatter volume-visualization technique. Because of the complexity associated with the simultaneous display of the three layers, environmental researchers requested that we use standard surface renderings for each layer. We obtained the data from the U.S. EPA Atmospheric Research and Exposure Assessment Laboratory.

Figure 7. Visualization of Total Ozone Mapping Spectrometer data. This visualization projects the TOMS data onto a spherical object in three views, helping EPA and NASA Langley researchers examine total global ozone distribution. NASA A Goddard supplied the data



resulted in the need to address high-speed data transfer, storage, and compression requirements for visualizing environmental data at other EPA research sites. As part of our participation in the U.S. Federal High Performance Computing and Communications (HPCC) effort, EPA will develop techniques to transfer high- performance tools to key state, federal, and industrial users with decision making responsibility. These future techniques are indeed grand challenges.

Acknowledgments

We thank the many environmental researchers throughout the EPA who have brought us challenging visualization projects. Special thanks are also in order to our colleagues at the North Carolina Supercomputer Center, University of California at Santa Barbara, NASA Langley Research Center, Numerical Design, Computer Sciences, Sterling Software, and Unisys who have provided support and suggestions on our various projects.

---

[1]

[2] Theresa Rhyne, Mark Bolstad, and Penny Rheingans, Martin Marietta Technical Services, Visualization Support, Research Triangle Park, NC.

[3] Lynne Petterson and Walter M. Shackelford, U.S. Environmental Protection Agency, Research Triangle Park, NC.

[4] L. Westover, Footprint Evaluation for Volume Rendering, *Computer Graphics* (Proc. Siggraph), Vol. 24, No. 4, Aug. 1990, pp. 367-376.

[5] A. Cullati, R. Idaszak, and T. Rhyne, Scientific Visualization Efforts at the U.S. Environmental Protection Agency, *Landscape and Urban Planning*, Special Issue on Data Visualization Techniques in Environmental Management, Vol. 21, No. 4, May 1992, pp. 323-326.

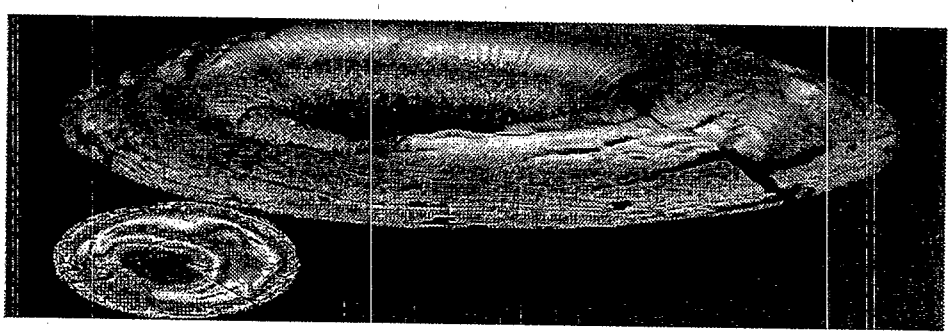


Figure 8. Visualization of Total Ozone Mapping Spectrometer data near the Antarctic region. The image is a polar orthographic projection of data from NASA Goddard, with missing data values shown as square holes.

Figure 9. Abstract representation of ozone concentrations from the Regional Oxidant Model for the northeastern U.S. The geographic map of the domain distorts with the height values of the ozone concentrations. We obtained the data from the U.S. EPA Atmospheric Research and Exposure Assessment Laboratory.

“

The machine does not isolate man from the great problems of nature but plunges him more deeply into them.

”

Antoine de Saint-Exupry (1900-1944)
French aviator, writer

Molecular Modeling on Supercomputers for Risk Evaluation [1]

Research Objective

The U. S. Environmental Protection Agency (EPA) is often faced with the problem of making regulatory decisions about chemicals for which there is a paucity of data on possible health effects. The Office of Toxic Substances receives applications for the registration of approximately 2000 new chemicals a year. Of those, there is acute toxicity data for only about 35% of the chemicals. There is little other animal toxicity data and some data on physical properties. We are learning about the many chemicals produced by secondary processes in the atmosphere for which health data is only now beginning to be collected. Additionally, there are chemicals in the atmosphere produced by combustion, chemicals in drinking water, and chemicals in waste dumps for which the human health data is scarce.

While the necessary data for the evaluation of the potential hazard of these chemicals is being developed, structure activity relationships developed from molecular modeling paradigms provide a useful tool for initial evaluation and provide a rational basis for the prioritization of testing needs. In order to more efficiently use these techniques, it is necessary to understand the molecular basis for the toxicity of chemical classes of environmental importance.

Approach

The actions of a xenobiotic chemical in a biological system can be produced by a variety of specific interactions and combinations of nonspecific processes. Interacting specifically with biomolecular targets like receptors, enzymes, and DNA, the xenobiotic may mimic endogenous substances or interfere with normal functions. The capacity of the agent to penetrate into the cell membrane or organelles where the physico-chemical properties of the agent can cause changes in the local environment to modulate normal processes is a basis for nonspecific effects. In this research, molecular modeling methods used in theoretical chemistry and pharmacology will be applied to this problem. Causal models for the appropriate biomolecular processes are developed in order to provide information on the relationship between chemical structure/properties and toxicity. It also provides an additional basis for the interpretation of experimental results. The structure activity relationships thus developed may be used to extrapolate the chemicals and circumstances for which there is a scarcity of heath data. The advances in computer architecture, the development of new molecular modeling software, and the increased availability of experimental information on the mechanism of chemical toxicity have greatly increased our capacity to develop rational molecular models for the toxic effects of xenobiotics. Supercomputing has and will continue to play an important part in the development and application of these models.

Accomplishment Description and Significance

Acrylates and similar chemicals are an important class of industrial chemicals. We have used molecular modeling methods, along with other techniques, to analyze the basis for the toxicity of chemicals in this class. A short description of that analysis follows. Michael addition was the postulated mechanism for the

interaction of chemicals in this class with DNA and other electron rich biological molecules (See Figure 1, page 88). This interaction does not require metabolism. Other competing pathways are metabolic epoxidation of the double bond and enzymatic de-esterification (See Figure 2, page 88). Epoxidation could provide a pathway for binding to biologically relevant molecules, and also provide a mechanism for elimination. Using *ab initio* quantum mechanical techniques, a three dimensional structure and a charge distribution were obtained for both acrylates and methacrylates. From these charge distributions, the molecular electrostatic potential (MEP) was computed for both of these moieties (See Figure 3, page 89). From the MEP it was observed that a negatively charged species, or the negatively charged part of a biopolymer, could approach more closely to the relevant reactive center of the acrylate moiety than the similar reactive center in a methacrylate. This strongly implies that acrylates are more likely to bind to biopolymers than the corresponding methacrylates. Again, *ab initio* quantum mechanical techniques were used to model Michael addition between these chemicals and small electron donor targets. The results of these calculations indicate that Michael addition to a reasonable mechanism for activity. From these calculations, it is possible to predict that a simple acrylate is likely to be more biologically active than the corresponding methacrylate in systems with minimal metabolism and that the addition of metabolic enzymes decreases the acrylate activity but may increase the methacrylate activity. Both of these predictions were confirmed by genetic bioassays which also provide a way to compare the two different pathways in Figure 2 (page 88). Even with the addition of metabolic enzymes, the acrylate is much more active than the corresponding methacrylate further indicating the importance of the Michael addition pathway. The implication of these studies is that positive data for the activity of a methacrylate can be used to imply activity for the corresponding acrylate, and negative data for an acrylate can be used to imply that the corresponding methacrylate is inactive. However, the reverse of these paradigms is not true. The confirmation of Michael addition as a mechanism for biological activity, using molecular modeling techniques, allows the chemical class to be expanded to include all chemicals that contain an electron withdrawing group adjacent to a double bond. More realistic biomolecular targets have been used to enhance the insight that this model provides.

Polycyclic Aromatic Hydrocarbons (PAHs) are a class of pervasive environmental chemicals produced by the incomplete combustion and pyrolysis of fossil fuels and other organic materials. Molecules within this class show considerable variation in toxicity. Some class members are powerful mutagens and animal carcinogens, while other molecules show no such activity after considerable testing. The activity of many class members falls between these two extremes. While the details of the molecular mechanism of action depend on the specific PAH and the particular test system used, there is always at least one metabolic oxidation step to a reactive epoxide. Experimental scientists in HERL have been studying the biological activity of PAHs and their specific mechanisms of action. Cyclopenta-PAHs (cPAHs) (PAHs that contain a five membered ring in addition to six membered rings) form one of the relevant subclasses. The carbon-carbon bond that closes the five membered ring is often the bond that is metabolically activated to an epoxide. The direction of ring opening for that epoxide indicates the atom that will be bound to the nucleophilic site in DNA (See Figure 4, page 89). In conjunction with the experimental studies underway in HERL, computational methods were used to predict the direction of the epoxide

ring opening for a series of cPAHs. In the initial studies the semi-empirical quantum mechanical method AM1 was used to obtain three dimensional structures of the epoxides and carbocations and *ab initio* quantum mechanical methods were used to obtain energies and charge distributions for those structures. The differences in energy between carbocation pairs resulting from the same epoxide was used to predict the direction of ring opening and the charge distribution used to provide information about the reactivity of the carbocation. It was found that the cPAHs being considered divided into two groups, one for which the energy difference between the possible carbocations was large (>7.5 kcal/mol) and those for which this difference was small (< 4.0 kcal/mol). It was possible to enunciate structural differences between the classes. Available experimental data partially confirmed this result. They suggest that the trends observed by the energy differences between carbocations are correct, but the computational methods may be overestimating this difference.

In later studies, as better computational resources became available, *ab initio* quantum mechanical methods and a semi-empirical method that includes the solvent water (AM1/SM2) were used to improve these results. We found that the improved molecular geometries changed individual carbocation energies significantly but the difference in energy between carbocation pairs was not changed significantly. However, the energy differences between carbocation pairs was lowered significantly by the inclusion of solvent and the results improved. The implication of these results is that, for charged molecules or reactive intermediates at least, a greater impact on the results will be made by methods that include solvent.

In most of the studies where molecular modeling techniques are applied to problems of environmental interest, the environmental agent (that is, the acrylate or the PAH or other molecule) is modeled and its reactivities or interacts with a small surrogate target studied to predict how it will act in a biological system. However, more and more experimental information is becoming available about the actual biomolecular targets for these agents (receptors, enzymes and DNA, for instance). As algorithms and computational facilities improve, it becomes possible to model the interaction of the agent or the ultimate toxicant with its actual target. In order to obtain experience for this eventuality, a specific DNA sequence has been modeled (the sequence was chosen because it is a known target sequence for an environmental agent being studied.) using the molecular mechanics features in the program Discover. Figure 5, page 90, shows the preliminary results of these computer experiments. The importance of including the solvent molecules is clear from these results.

Future Plans

Molecular modeling methods will be applied to additional classes of importance to the Agency. Methods that include solvent in both quantum mechanical and molecular dynamics and mechanics will be employed. Biomolecular targets will be included in the models for activity as the specific targets are identified experimentally.

---

[1] James R. Rabinowitz, Ph.D., CMB/GTD/HERL, U.S. EPA, Research Triangle Park, NC.

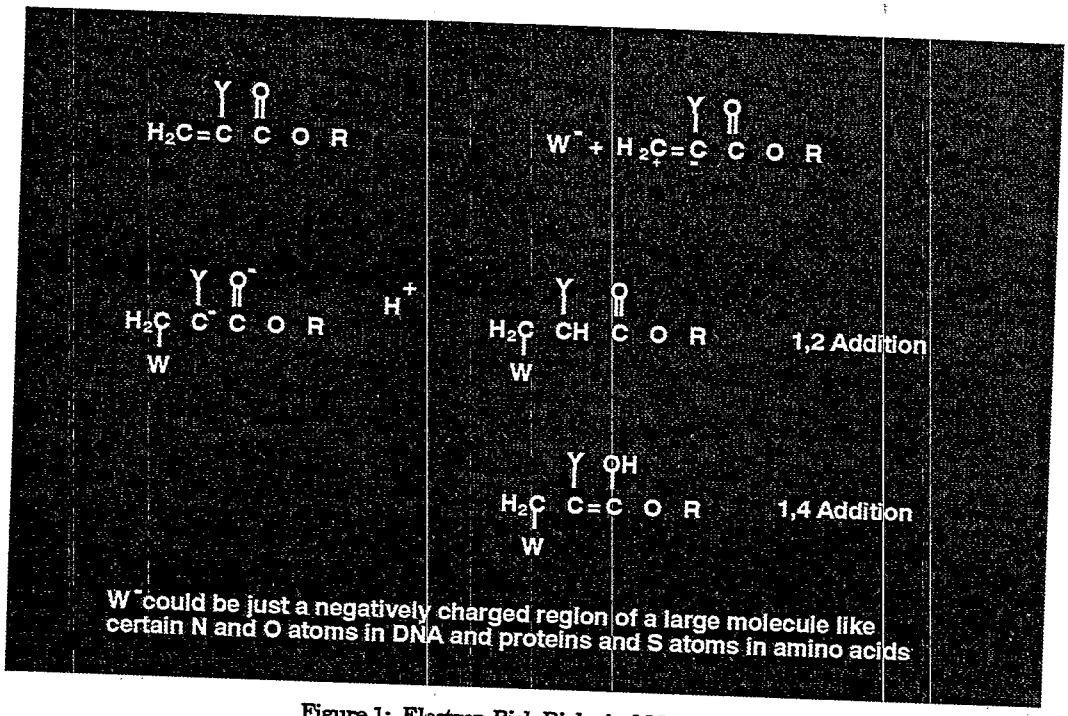


Figure 1: Electron Rich Biological Molecules

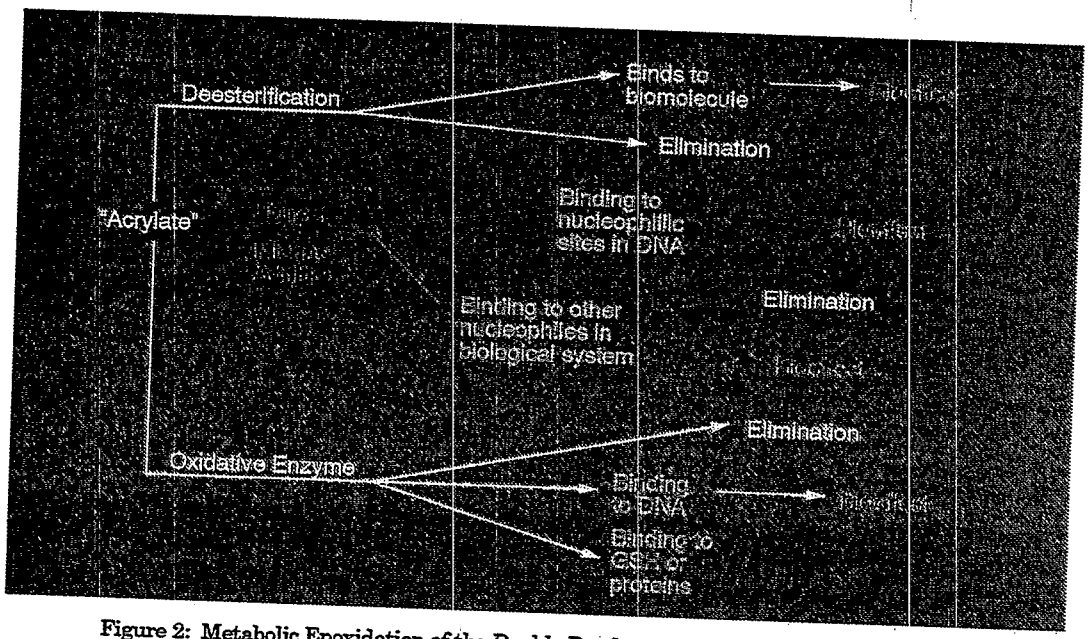


Figure 2: Metabolic Epoxidation of the Double Bond and Enzymatic De-esterification

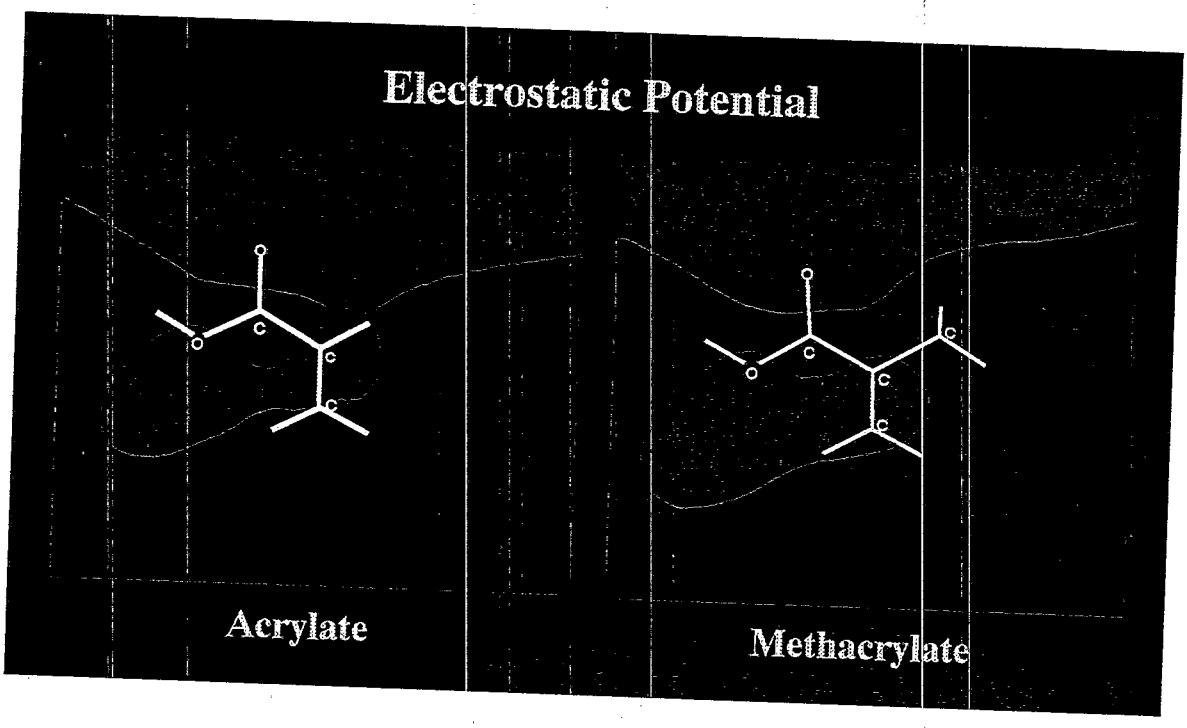


Figure 3: Molecular Electrostatic Potential (MEP) Computation

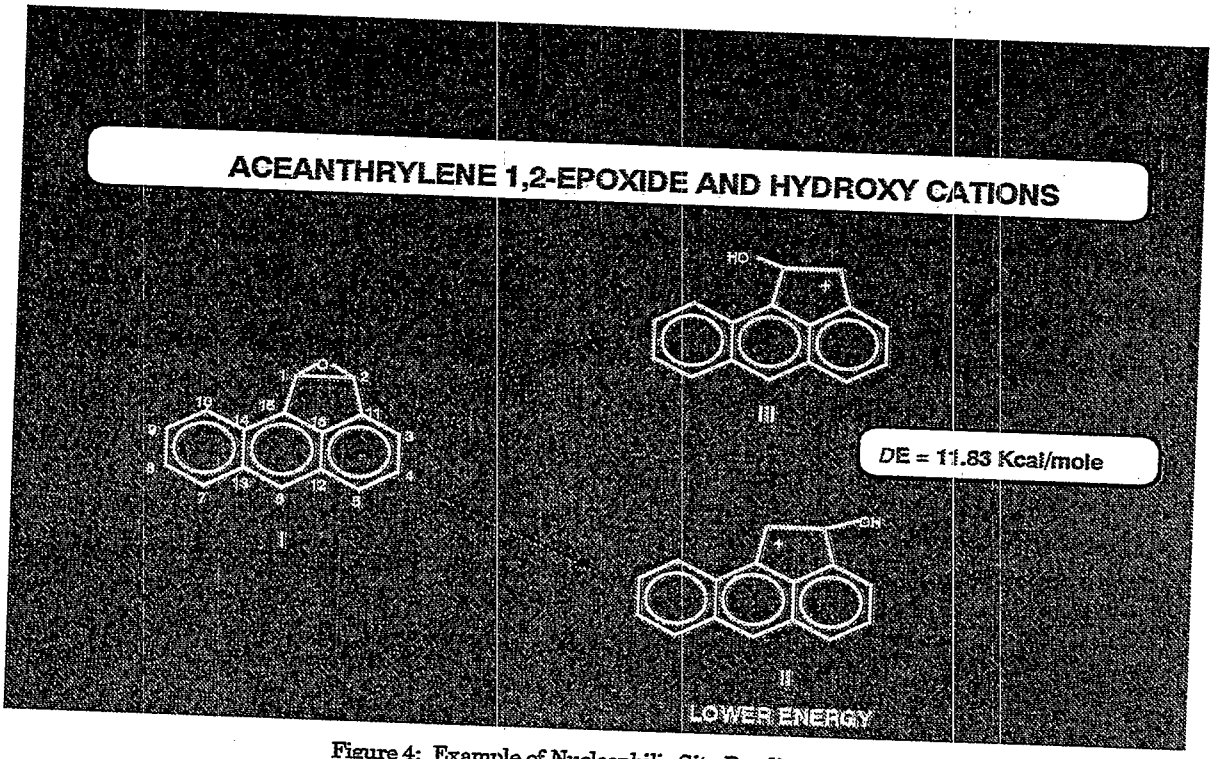


Figure 4: Example of Nucleophilic Site Bonding

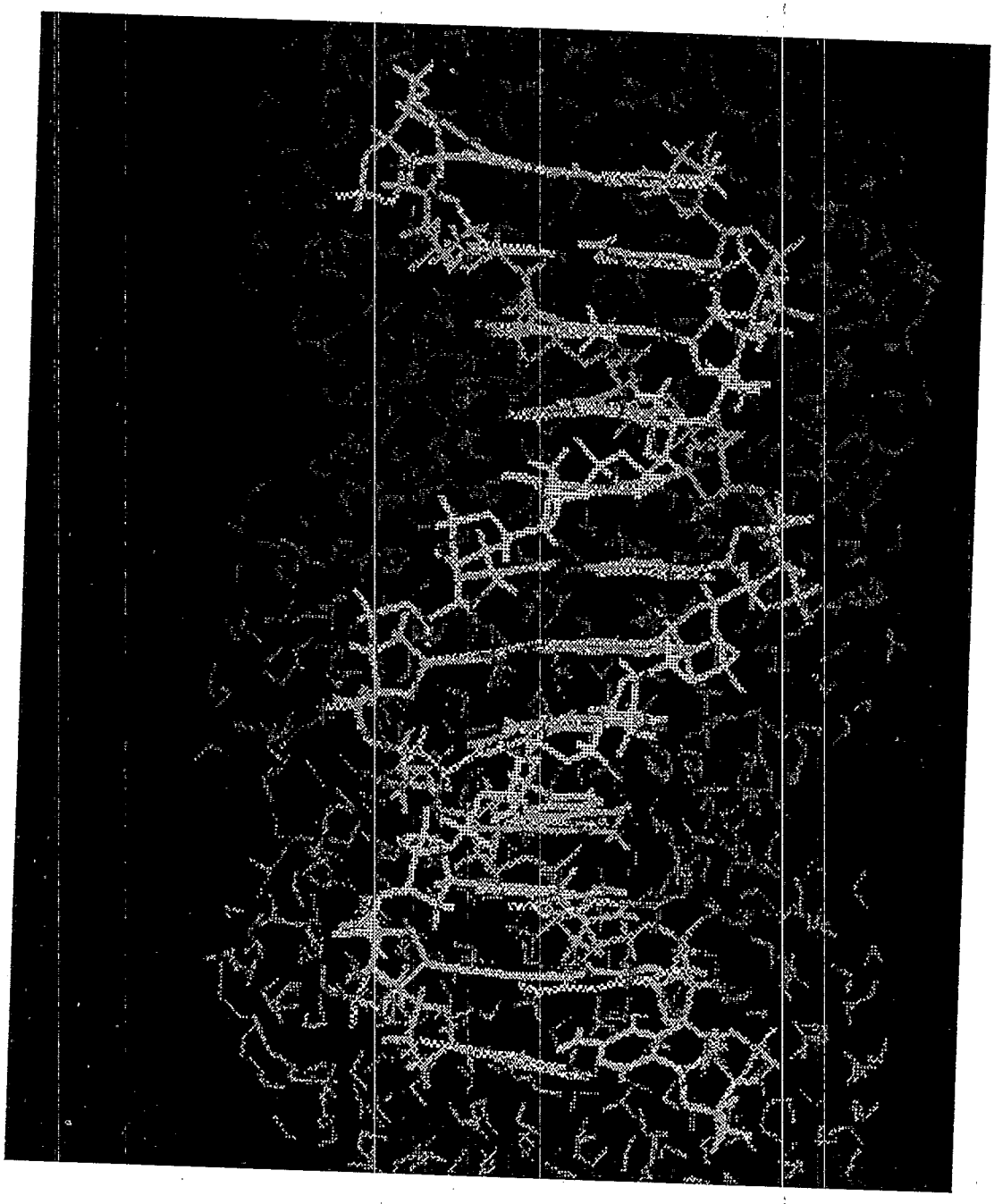


Figure 5: DNA Sequence Model

Regional Ozone Modeling to Support Clean Air Act Mandates [1,2]

Abstract

Attainment of the ozone National Ambient Air Quality Standards (NAAQS) has been a particularly difficult environmental problem to solve. In fact, nearly 100 areas of the United States are still not attaining the NAAQS despite the efforts of the U. S. Environmental Protection Agency (EPA), states, and industry over the last 15 years. In 1990 the Congress reauthorized the Clean Air Act (CAA) in part to mandate programs and milestones for attainment of the ozone NAAQS. In order to support states in meeting these requirements and schedules, the EPA has committed to a regional scale modeling program using the Regional Oxidant Model (ROM). The purpose of the proposed work is to use ROM simulations to support development of credible control strategies to solve the ozone problem for many of the most seriously polluted areas of the Eastern U. S. Because of the computational complexity of models such as ROM, the space and time scales of the applications, and the number of simulations required, it is necessary to run this model in a supercomputing environment.

The key objectives of this project are to 1) assess the relative benefits of alternative emission control strategies for reducing regional ozone concentrations and 2) provide the data bases containing the regional model predictions to states for use in estimating the impacts of controls on future levels of pollutant transport.

Results from completed applications have been used by the EPA for developing policies on the relative effectiveness of alternative emissions reduction scenarios. Analyses for the Ozone Transport Commission have been used to assess the benefits of emissions reductions from mandated CAA control programs.

Background

In November 1990 the United States Congress amended the Clean Air Act (CAA) to provide for attainment and maintenance of the National Ambient Air Quality Standards (NAAQS) to protect public health and welfare against the harmful effects of air pollutants. Major portions of the Clean Air Act Amendments (CAAA) contain provisions oriented toward attainment of the ozone NAAQS through programs that reduce ozone precursor emissions of volatile organic compounds (VOC), nitrogen oxides (NO_x), and carbon monoxide (CO). Section 182 and guidance issued by the U. S. Environmental Protection Agency (EPA)[3] require that urban scale photochemical grid modeling be used by states to demonstrate attainment in areas classified as having moderate (interstate), serious, severe, and extreme ozone problems. Most of these urban modeling efforts involve the Urban Airshed Model (UAM)[4]. In addition, Section 184 of the CAAA establishes an Ozone Transport Commission (OTC) to address the specific regulatory and technical issues associated with modeling and implementing regional control strategies needed to achieve attainment in the Northeast. The formation of the OTC is in recognition of the complexity of the ozone problem in this region where summer meteorological conditions, combined with the spatial distribution and large magnitude of ozone precursor emissions, contribute to long range and interurban transport across political boundaries over several days.

Overview and Objectives

A key component of the modeling by states is the estimation of expected future year urban boundary conditions which reflect changes in upwind emissions between current baseline and future attainment dates required by the CAAA. The EPA has developed a methodology for deriving these boundary conditions from predictions of the Regional Oxidant Model[5,6] (ROM). The EPA is committed to providing the ROM applications that couple with the episodes and emissions scenarios to be simulated with UAM by states. The objectives of this project are to use the capabilities of the National Environmental Supercomputing Center (NESC) to perform the ROM simulations necessary to support the states. Because of the computing requirements of ROM for the spatial and temporal scales needed for these applications, using the NESC is the only available means to achieve results in time to meet the schedules required in the CAAA.

Also, states in the OTC are examining regional controls beyond those in the CAAA that may be needed for attainment in this region. The EPA has completed several ROM analyses for the OTC to examine the relative effectiveness of possible regional strategies. Regional strategies selected by the OTC will be used by individual states as an initial baseline set of controls in their attainment demonstration modeling. This report presents the results of ROM simulations performed for the OTC.

Accomplishments

The OTC modeling analysis has focused on developing an initial regional

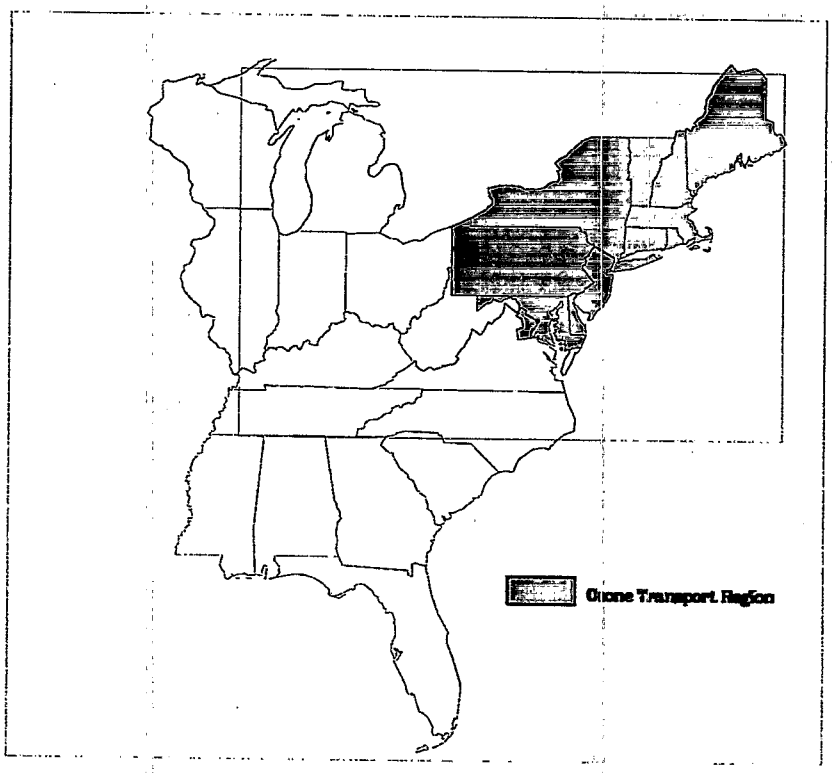


Figure 1: Map of Northeast ROM domain and the Ozone Transport Region.

assessment of the impacts on Northeast ozone levels due to the control programs specified in the CAAA. The results have not only been valuable to the OTC, but also to the EPA since they provide the only integrated assessment of the likely ozone benefits of controls mandated by the CAAA across a region containing over 25 nonattainment areas.

Results from Completed Analyses

The OTC assessment included ROM simulations for emissions scenarios described below. The location of the Transport Region in the ROM Northeast simulation domain are shown in Figure 1, page 92.

Base Case Scenario

The Base Case Scenario was largely created from the 1985 National Acid Precipitation Assessment Program (NAPAP)[7] emissions inventory. In the Transport Region, area sources and highway vehicles account for the bulk of man-made VOC emissions. However, biogenic emissions make up half of the total VOCs. Although they are highest in rural areas, biogenics contribute about 30% to total VOC emissions in many of the Northeast urban areas. An exception is New York City where biogenics are only about 10% of the total. In contrast to VOC, NO_x emissions are dominated by highway vehicles and point sources, with very low natural emissions.

Of importance for designing regional strategies is the spatial distribution of emissions. In the Transport Region, 84% of the man-made VOC and 78% of the NO_x emissions are from sources in nonattainment areas. In addition, 72% of the VOC and 61% of the NO_x in the region are from sources along the Northeast Corridor from Washington, DC to southern Maine. In rural areas, biogenics dominate VOC emissions, but the largest contribution to NO_x comes from electric generation by utilities.

2005 Scenario

The Base Case emissions were projected to 2005 considering the effects of growth, existing controls and the new stationary and mobile source control programs in the 1990 CAAA. The year 2005 was selected since it is near or beyond the dates when most of the areas in the Transport Region are required to be in attainment. Although the focus of this analysis is on the Transport Region, the appropriate CAAA control was also applied to emissions in other nonattainment areas.

Impact on Emissions of CAAA Controls

The result of applying the above control programs and estimates of emissions growth (where applicable) was a net reduction in man-made VOC emissions in the Transport Region of 46% between the Base Case and the 2005 Scenario. When biogenic emissions are factored in, the overall reduction in VOC is only 24%. For NO_x, the reduction in emissions is 36% (recall that natural NO_x emissions, as currently quantified, are relatively insignificant.)

ROM Ozone Predictions

The ozone predictions from ROM for the two scenarios simulated are presented and compared in terms of 1) spatial patterns and the percent change in episode maximum concentrations and 2) changes in daily maximum ozone. Ozone predictions from the lowest layer in ROM are used in the comparisons.

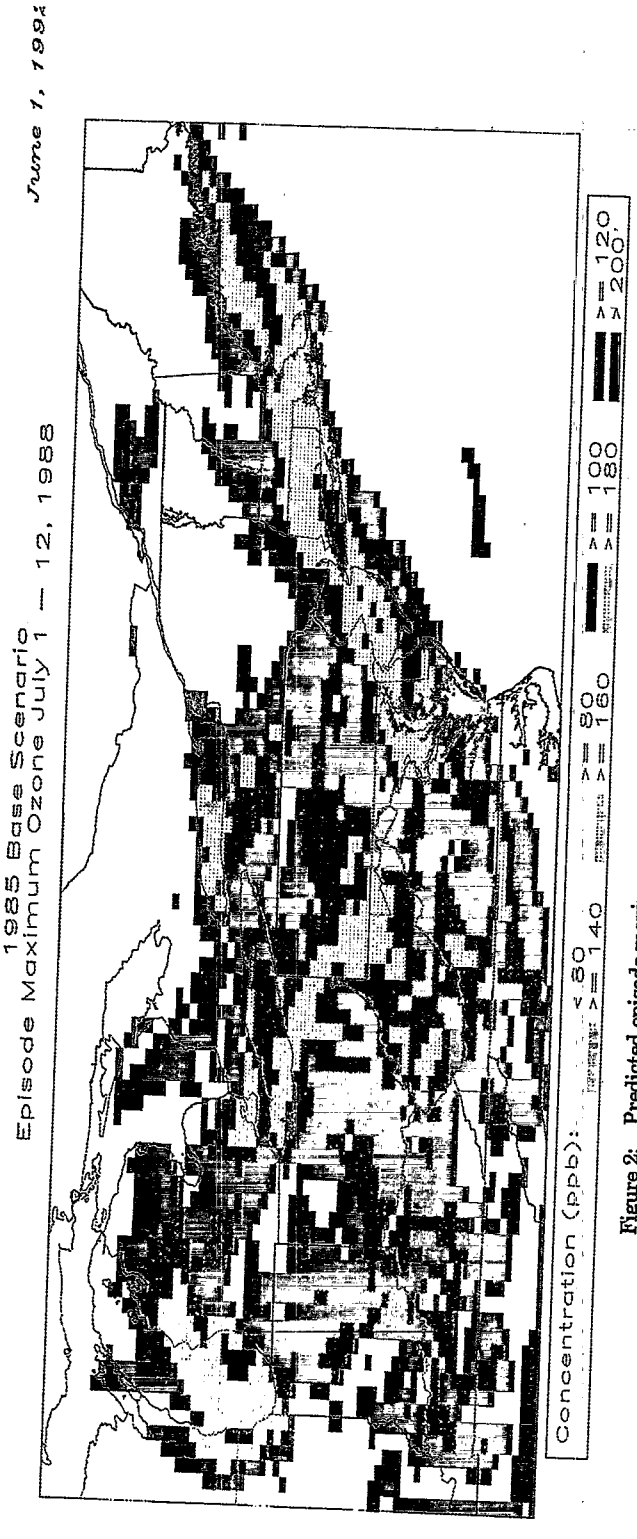


Figure 2: Predicted episode maximum ozone concentrations for the 1985 Base Case scenario.

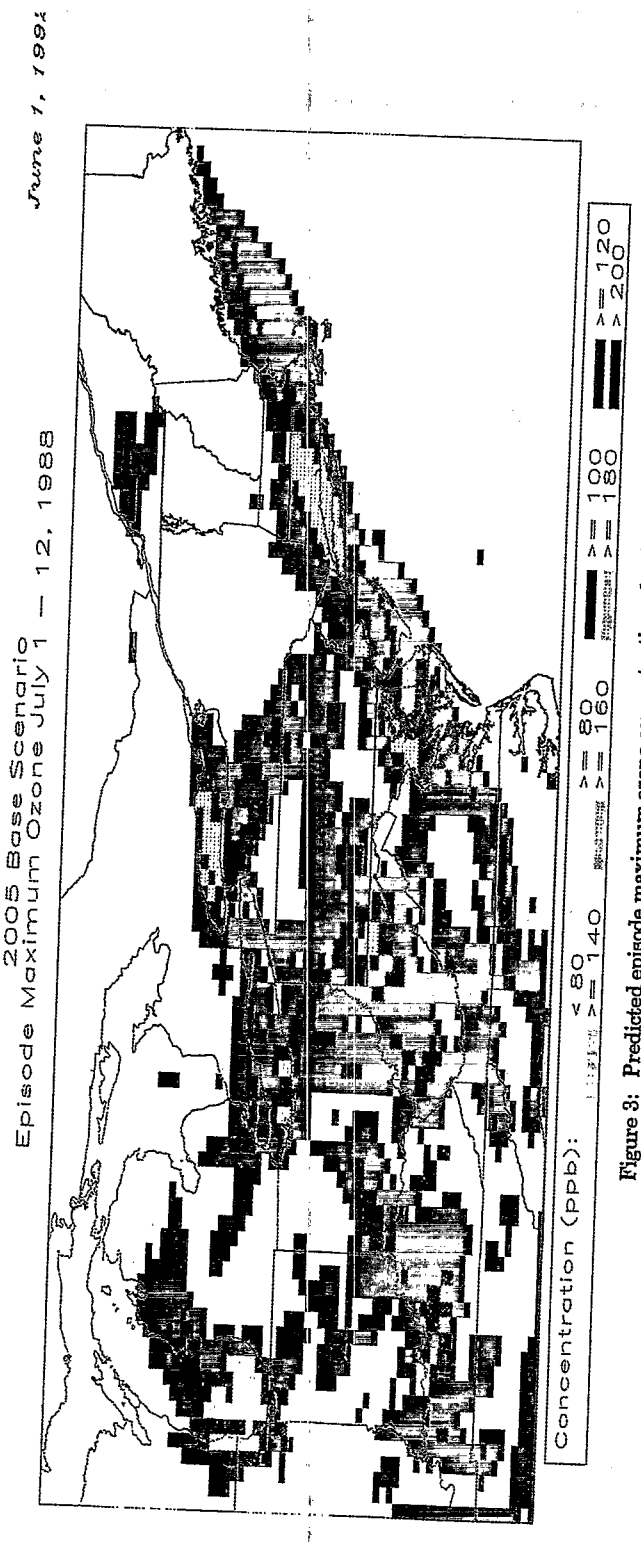


Figure 3: Predicted episode maximum ozone concentrations for the 2005 scenario.

Comparisons for Episode Maximum Ozone

The spatial distribution of episode maximum ozone for the Base Case Scenario is shown in Figure 2, page 94. The episode maximum values represent the highest 1-hour ozone concentrations predicted in each grid all over the entire episode. The figure indicates ozone concentrations above the level of the NAAQS stretching along the entire Northeast Corridor with values exceeding 180 to 200 ppb near Washington, DC/Baltimore and Philadelphia, and from New York City across Connecticut. High ozone levels are also predicted near and downwind of other cities (e.g. Chicago, Cleveland, Pittsburgh, and Toronto) as well as across rural sections of several states (e.g., Virginia, West Virginia, Pennsylvania, and New York).

Predictions for the 2005 Scenario in Figure 3, page 94, indicate that the CAAA controls result in a substantial reduction in peak values as well as the aerial coverage of concentrations above 120 ppb. The number of grids with ozone above 120 ppb in the Transport Region declined by 56%. Peak values along the Northeast Corridor are reduced by 15 to over 20%. Elsewhere in the region, peak ozone levels declined by 10 to 20% across Maryland and Pennsylvania with lesser reductions across upstate New York. Despite these reductions, levels are still predicted to remain above 120 ppb in many of areas of the region.

Comparisons for Daily Maximum Ozone

The 95th percentile of daily maximum ozone concentrations for selected areas in the Transport Region indicates reductions in ozone of approximately 20% from the New York City area northeastward along the Corridor with slightly less reduction (16 to 18%) southwest of New York. In Pittsburgh, daily maximum ozone declined by 13%. These reductions are similar in magnitude to those revealed by the episode maximum concentrations above. The 95th percentile values following the implementation of CAAA control measures decline to near or below the level of the NAAQS in most areas. The exception is New York City and southern New England, which is immediately downwind of New York on most days in this episode.

Summary and Conclusions

Preliminary applications of the ROM have been made to assess the expected benefits of emissions reductions specified in the 1990 CAAA on ozone in the Northeast. This analysis is part of a larger effort to develop a set of regional and/or subregional strategies that, when combined with urban area-specific control programs, will provide for attainment of the ozone NAAQS in the Northeast. The ROM applications include simulations for a Base Case Scenario and a 2005 Scenario that includes the net effects on emissions of growth together with controls in the CAAA. The results indicate that the CAAA controls may provide significant reductions in ozone levels in this region. However, even with these controls, ozone may remain close to or above the NAAQS by the year 2005 in portions of this region under meteorological conditions of the type simulated. Thus, optional regional control programs are being considered to provide the additional emissions reductions which may be needed for attainment. The expected effectiveness of these programs will be assessed as part of upcoming regional and urban model applications.

[1] Norman C. Possiel, National Oceanic and Atmospheric Administration. The author is on assignment to the U.S. EPA, Office of Air Quality Planning and Standards, U.S. EPA, Research Triangle Park, NC.

[2] Richard Wayland, U.S. EPA Project: Regional Modeling for the Office of Air Quality Planning and Standards (ROMO) U.S. EPA, Research Triangle Park, NC.

[3] *Guideline for Regulatory Application of the Urban Airshed Model*, EPA-450/4-91-013, U.S. EPA, Research Triangle Park, NC, 1991, 89 pp.

[4] R. E. Morris, T. C. Myers, and J. L. Haney, *User's Guide for the Urban Airshed Model*, Volume I: User's Manual for UAM (CB-IV), EPA- 450/4-90-007a, U.S. EPA, Research Triangle Park, NC, 1991, 259 pp.

[5] R. G. Lamb, *A Regional Scale (1000 km) Model of Photochemical Air Pollution: Part 1.* Theoretical Formulation, EPA-600/3-83-025, U.S. EPA, Research Triangle Park, NC, 1983, 239 pp.

[6] J. O. Young, M. Aissa, T.L. Boehm, C.J. Coats, J.R. Eichinger, D.J. Grimes, S.W. Hallyburton, W.E. Heilman, D. T. Olerud, S. J. Roselle, A.R. Van Meter, R. A. Wayland, and T. E. Pierce, *Development of the Regional Oxidant Model Version 2.1*, U.S. EPA, Research Triangle Park, NC, 1989, 55 pp.

[7] M. Seager, J. Langstaff, R. Walters, L. Modica, D. Zimmerman, D. Fratt, D. Bulleba, R. Ryan, J. Demmy, W. Tax, D. Sprague, D. Mudgett, and A. Werner, *The 1985 NAPAP Emissions Inventory (Version 2): Development of the Annual Data and Modelers' Tapes*, EPA-600/7-89-012a, U.S. EPA, Research Triangle Park, NC, 1989, 692 pp.

Regional Acid Deposition Model (RADM) Evaluation [1]

EPA Research Objectives

Regional air quality models are needed and used to extrapolate outside current conditions, therefore, these advanced models are developed with parameterizations and physical and chemical mathematical descriptions as close to first principles as possible. The purpose of the evaluation is to test the science incorporated into the advanced models. Evaluation is diagnostic, to explore quality of predictions and develop an appraisal of model strengths and weaknesses. The data used were specially collected for the RADM evaluation as part of the National Acid Precipitation Assessment Program (NAPAP) and a binational effort, the Eulerian Model Evaluation Field Study (EMEFS). The data were collected over a two-year period with special, several-week intensives that used very advanced instruments to collect air concentrations to provide data that would support the most diagnostic testing.

Overview of Project

Early evaluation research concentrated on examining the predictions for the sulfur cycle. Significant improvements to the RADM were accomplished (see references). Current research is investigating the nitrogen cycle, which is much more complex. This investigation focuses on the testing the ability of the model to accurately replicate in time and space the conversion (or oxidation) of nitrogen oxides (NO_x) to their oxidized products, PAN and nitrates (particulate nitrate, NO_3^-, and nitric acid, HNO_3). Measurements taken by aircraft carrying sophisticated instruments to measure air quality in the EMEFS's 1988 aircraft intensive are used for the diagnostic testing.

Background and Approach

The observations were developed from measurements taken during 12 aircraft flights over a 35-day period. The standard 80-km. version of RADM2.6 was used to simulate the 35-day period and a data-probe was "flown" through the model. These "data" from the model are compared to equivalent data from the aircraft measurements. An early sensitivity study on a 3-day period using a smaller grid resolution indicated that grid size may be affecting the rate of conversion of NO_x to PAN and HNO_3. Thus a full set of runs was prepared for the August 25 to September 29, 1988 period that includes the aircraft flights and using a 20-kilometer version of RADM) the High Resolution RADM (HR-RADM), using a subdomain of the 80-km. RADM that encompassed the aircraft flights. The meteorology was interpolated from 80 kilometers to 20 kilometers to maintain consistency in scales that are resolved, and because the more advanced meteorological modeling capability at 20 km., requiring a new convective parameterization designed for scales less than 25 km., was not yet ready. The HR-RADM was run as a one-way nest within the full RADM domain.

Accomplishments Using the NESC's Cray

The initial 80-km. RADM runs for the August 25 to September 29, 1988 period required approximately 48 hours on a single-processor Cray Y-MP. The HR-RADM runs for this same period required approximately 180 hours on a single processor of the NESC Cray Y-MP. It required about five weeks to prepare the emissions. The model runs on the Cray were able to be completed in about two weeks.

Scientific Results and Relevance to EPA Mission

The full set of comparisons showed that reducing the grid size from 80 km. (typical for regional acid deposition models) to 20 km. (typical for regional oxidant models) had little effect on the rate of conversion of NO_x to total-nitrate. The conversion to PAN increased. The manner of the changes in the model predictions indicated that changing grid size does not have the effect one expects. It was very important to be able to test the entire period at the higher grid resolution to avoid the possibility that a single case may not be representative of behavior over the entire period.

This study enhances our understanding of the working of regional model photochemistry for rural ambient concentrations conditions. Proper computation of the photochemistry for rural conditions is important to the ability of models to support exploration and establishment of appropriate emissions controls to reduce and eliminate violations of the ozone health standard. Examination of nitrogen chemistry is important because it is a central part of the oxidation process forming ozone and because rural oxidant production is generally believed to be NO_x-limited.

Future Objectives and Plans

The evaluation will continue with additional sensitivity studies directed at understanding meteorological influences, especially temperature, on the RADM chemistry. Preliminary indications are, that to improve temperature predictions from the mesoscale meteorological model, its vertical resolution will need to be increased from 15 to 25 layers and parameters affecting the surface heat flux, such as a soil moisture, updated. Once the newly adapted meteorological model has been tested, roughly 60 Cray Y-MP hours will be required to regenerate new meteorology for the September 1988 evaluation period, 80 Cray Y-MP hours to generate new 80-km. RADM results, and 300 Cray Y-MP hours to generate new HR-RADM results for the next round of diagnostic testing of the chemistry in RADM.

Publications and Reports

Dennis, R. L., W. R. Barchet, T. L. Clark, S. K. Seilkop, and P. M. Roth, 1990: Evaluation of regional acid deposition models (Part 1), NAPAP SOS/T Report 5. In: *National Acid Precipitation Assessment Program: State of Science and Technology, Volume I*. National Acid Precipitation Assessment Program, 722 Jackson Place NW, Washington, D.C.

Dennis, R. L., J. N. McHenry, W. R. Barchet, F. S. Binkowski, and D. W. Byun, 1993: Correcting RADAM's sulfate underprediction: discovery and correction of model errors and testing the corrections through comparisons against field data. *Atmospheric Environment* **27A**, 975-997.

---

[1] Robin L. Dennis, Talat Odman, Richard D. Cohn, William Hwang, and Daewon Byun, Regional Acid Deposition Model Evaluation, U.S. EPA, Office of Research and Development, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC.

Atmospheric Deposition of Nitrogen to Chesapeake Bay

EPA Research Objectives

Nitrogen is the primary cause of eutrophication in Chesapeake Bay. Nitrogen input from the atmosphere represents a significant source of nitrogen to the Bay (25-35% of the nitrogen loading). Water quality models have incorporated atmospheric nitrogen, but in a very simple manner. One objective of this research is to provide more accurate estimates of the quantity and the pattern of nitrogen loading from the atmosphere to the Chesapeake Bay watershed and the Bay itself. These estimates will be provided as inputs to the water quality models for the watershed (the HSPF model adapted by the Chesapeake Bay Program Office) and the Bay (the 3-D Bay Water Quality model developed by the Army Corp of Engineers). Another objective of this research is to determine the extent of the airshed that is primarily responsible for the atmospheric nitrogen affecting the Bay watershed. The airshed will be larger than the watershed. The overall purpose is to develop an understanding of which controls of NO_x emissions to the atmosphere will have the greatest benefit on reducing the nitrogen loading to coastal estuaries. This work is important to the Chesapeake Bay Program Office's efforts to achieve a 40% reduction in controllable nitrogen loading to the Bay by the year 2000, and to the upcoming 1996 Agency decision on the amount of Phase 2 NO_x controls required by the 1990 Clean Air Act Amendments (CAAA).

Overview of Project

Development of more accurate spatial fields of nitrogen loading estimates involves estimation of annual average nitrogen deposition to coastal areas using the Regional Acid Deposition Model (RADM). These deposition estimates are made for 1985 emissions and representative meteorology. They are also made for 2005 emissions projections representing estimates of changes stemming from growth combined with emissions reductions called for in the 1990 Clean Air Act Amendments with the same meteorology. Development of an understanding of the airshed influencing the Chesapeake Bay watershed involves using RADM as a laboratory of the real world to carry out sensitivity studies that elucidate the contributions of different emissions sources to the Bay watershed. This source-receptor understanding is very difficult and nearly impossible to develop from empirical data and requires the designing of sensitivity studies that will extract that information from a mathematical model.

Background and Approach

Because the RADM is very computationally intensive, it is not feasible, with today's computing power, to simulate an entire year's worth of meteorology to develop annual average estimates of deposition loading. Instead, annual averages are developed from a weighted average of a statistical sample of 30 5-day model runs. The average is representative of meteorology for the 1982 to 1985 period, which has a rainfall pattern very close to a 30-year average. Meteorological events (synoptic progressions of high and low pressure systems) with similar 850-mb wind-flow patterns were grouped or classified by applying cluster analysis to them. This resulted in 19 sampling groups or strata. Meteorological cases were randomly selected from each stratum, based on the number of wind-flow patterns in that stratum and on the number in each of

the other strata. This procedure approximates proportionate sampling. The number of cases, 30, was set after carrying out a sampling-error analysis on wet sulfur deposition and taking into consideration computer resource limitations. These are termed the aggregation cases.

Development of a source-receptor understanding on an annual basis requires an experimental design that will extract this information from sensitivity studies with RADM. Because NO_x emissions contribute to oxidant production and there is a dynamic interplay between the production of ozone and nitric acid, the dominant form by which nitrogen is deposited to the Earth's surface, the modeling of nitrogen must incorporate full photochemistry as is done in RADM. As a first approximation to the source-receptor relations implicit in the model calculations, emissions sources of interest are subtracted from the emissions fields. The 30 aggregation cases are run and the results subtracted from results obtained with unperturbed emissions fields. For this study, the objective was to develop an understanding of the difference in the range of influence of ground-level NO_x emissions (such as automobiles) from upper-level NO_x emissions (such as power plants) with regard to nitrogen deposition.

Accomplishments Using the NESC's Cray

The basic 1985 and 2005 RADM model runs were the same as those produced in the 1990 CAAA Projections project (requiring 240 hours on a single-processor Cray Y-MP). The two sensitivity runs that were part of this study required 240 hours on a single-processor Cray Y-MP.

Scientific Results and Relevance to EPA Mission

The comparison of predicted wet and dry nitrogen deposition from RADM indicated that dry deposition of nitrogen appears to be less than wet. Up to now, many researchers were assuming that dry deposition equaled wet deposition. The model result was confirmed by comparing the model estimates with National Dry Deposition Network data. Thus, the model results are helping to provide more accurate estimates of the magnitude of nitrogen deposition from the atmosphere to the coastal estuaries. The model results have also identified the size of the gradients in the pattern of total nitrogen deposition to the watershed. This information will now be input to the water quality models.

The new, projected reductions in total nitrogen deposition expected to result from the 1990 CAAA are estimated to be 50% larger than earlier estimates used in sensitivity studies with the water quality models. Thus, reductions in NO_x emissions due to the 1990 CAAA could play a more beneficial role for Chesapeake Bay than previously thought. As a result, continued study of atmospheric reductions appears to be important to Chesapeake Bay work.

The sensitivity study on the range of influence of NO_x emissions on nitrogen deposition depending on the height of the emissions (surface or tall-stack) produced a somewhat unexpected result. The range of influence of surface emissions appears to be roughly 70% that of emissions from tall stacks. This is different than the "conventional" wisdom which would "predict" that the range of influence from tall stacks would be much greater. It is possible that conventional wisdom has been influenced by study of the sulfur system where the primary specie, SO_2, plays a significant role in the total deposition. In the nitrogen system, nitrogen deposition is almost entirely due to the secondary specie nitric acid, HNO_3 (ignoring ammonia for the moment). The distance over which NO_x

emissions from the western Pennsylvania region appear to noticeably influence the nitrogen deposition is approximately 400 to 500 kilometers. This means that the airshed affecting Chesapeake Bay is significantly larger than the watershed and is expected to include many sources along the upper Ohio River.

Future Objectives and Plans

Future plans call for from two to three additional sensitivity runs to help better define the airshed affecting the Bay. Plus, the relative contribution from different economic sectors to the nitrogen deposition are important to establish where the effort should be placed in the study of future control options. National trends mask regional differences that are important to regions such as the Chesapeake Bay. The additional sensitivity runs are expected to require 240-360 Cray Y-MP hours and the economic sector study on 1985 and 2005 emissions to require 720 Cray Y-MP hours. The exploration of control options is also planned and is expected to require 240 to 360 Cray Y-MP hours. In addition, the degree of error in the technique used to approximate the nitrogen source-receptor relations needs to be established. Such an experiment would be expected to require the order of 240 Cray Y-MP hours. The Chesapeake Bay Program would also like to develop an estimate of the portion of the nitrogen deposition that is coming from several major urban areas within the watershed. This will require using the High-Resolution RADM to better resolve the urban areas. The base runs plus the urban sensitivities would be expected to require 1,800 Cray Y-MP hours.

Publications and Reports

In the process of being written.

---

[1] Robin L. Dennis, Talat Odman, Richard D. Cohn, William Hwang, and Daewon Byun, Regional Acid Deposition Model Evaluation, U.S. EPA, Office of Research and Development, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC.

“Pollution is nothing but the resources we are not harvesting. We allow them to disperse because we’ve been ignorant of their value.”

R. Buckminster Fuller [1895-1983]
American Architect, Engineer

Projected Effects of the 1990 Clean Air Act on Acidic Deposition [1]

EPA Research Objectives

Developing accurate estimates of the impact of the 1990 Clean Air Act Amendments (CAAA) on acidic deposition and atmospheric sulfate (key to visibility degradation in the Eastern United States) are important to the Agency. The amount of reduction in sulfur deposition to be anticipated by 2005 or 2010 due to implementation of Title IV Phases I and II sets an important baseline for understanding how much mitigation in deposition we are expected to achieve and how much farther we might need to go to provide protection to ecological resources. The reduction in deposition loading in Canada that is likely coming from the United States and vice versa is important to the U.S. Canada Air Quality Accord. These estimates are important to the Canadians for them to project whether they will achieve their goal of wet sulfur deposition being below 20 kg-SO_4^{2-}/ha. As well, the European community is interested in estimates of the long-range transport across the Atlantic of sulfur-related acidic deposition that could be affecting them. The objective is to develop best estimates from evaluated and well-characterized models of changes in acidic deposition loading, visibility impacting pollutants, and oxidants. Also, the cross-program effects from the different Titles of the 1990 CAAA need to be characterized.

Overview of Project

Development of estimates of future deposition involves creation of estimates of future emissions that account for population and economic growth plus the incorporation of emissions controls called for by the 1990 CAAA. The new emissions are input to RADM simulations to estimate the new deposition, assuming the same meteorology as today's. A difficult element for the projection of future emissions is the estimation of power-plant retirements, and the installation of new generating capacity to make up the difference in on-line capacity and projected demand. Of special difficulty is locating or siting potential future plants for the modeling. These projections are generated by experts in the field of emissions estimation and projection.

Background and Approach

Because the RADM is very computationally intensive, it is not feasible, with today's computing power to simulate an entire years worth of meteorology to develop annual average estimates of deposition loading. Instead, annual averages are developed from a weighted average of a statistical sample of thirty 5-day model runs. The average is representative of meteorology for the 1982 to 1985 period, which has a rainfall pattern very close to a 30-year average. Meteorological events (synoptic progressions of high and low pressure systems) with similar 850-mb wind-flow patterns were grouped or classified by applying cluster analysis to them. This resulted in 19 sampling groups or strata. Meteorological cases were randomly selected from each stratum, based on the number of wind-flow patterns in that stratum and on the number in each of the other strata. This procedure approximates proportionate sampling. The number of cases, 30, was set after carrying out a sampling-error analysis on wet sulfur deposition and taking into consideration computer resource limitations. These are termed the aggregation cases.

Two emissions cases were developed for the year 2005. The first emissions

projection case considered Title II (automotive NO_x emissions) and Title IV, Phase 2 (acid rain-related utility emissions of SO_2 and NO_x) emissions reductions. The second emissions projection case considered the addition of Title I (ozone State Implementation Plan requirements on NO_x emissions) to the other two CAAA titles. The projections represented full implementation of the CAAA titles. Thus, emission reductions for SO_2 included all provisions of Phase 2, which are more representative of 2010 emissions of SO_2.

Accomplishments Using the NESC's Cray

The basic 1985 and the Titles-II and IV only 2005 RADM model runs required 240 hours on a single-processor Cray Y-MP. The "sensitivity" run for 2005 that incorporated Titles I, II, and IV required an additional 120 hours on a single-processor Cray Y-MP.

Scientific Results and Relevance to EPA Mission

The reductions of total sulfur deposition due to the Title IV acid rain controls produced by this study were similar to those projected for the National Acid Precipitation Assessment Program's 1990 assessment. Comparison of the two cases indicates that the reduction in nitrogen deposition across the northeast (affecting the Chesapeake Bay watershed in particular) is expected to be a factor of 1.2 larger due to consideration of Title I State Implementation Plan requirements in addition to the controls mandated by Titles II and IV. Thus, it is important to incorporate the impacts of interprogram (inter-pollutant) air quality mandates of the 1990 Clean Air Act on estimates of future air quality and deposition loading. This is consistent with the new emphasis in EPA on considering the full range of multi-media and multi-program effects and taking a more holistic perspective towards pollution control.

Future Objectives and Plans

Future plans call for updating the 2005 projections, in coordination with the Regional Oxidant Model runs by Office of Air Quality Planning and Standards, using the new 1990 emissions inventory as a baseline, more current economic growth projections and the latest mobile source emissions model from EPA-Ann Arbor. Plans are also to assess the potential for differences in sulfur deposition at regions of particular interest across the eastern United States that might occur through the trading of SO_2 emission allocations.

Publications and Reports

In the process of being written.

---

1 Robin L. Dennis, Regional Acid Deposition Model Evaluation, U.S. EPA, Office of Research and Development, Atmospheric Research and Exposure Assessment Laboratory, Research Triangle Park, NC.

The Role of the Zebra Mussel (*Dreissena Plymorpha*) in the Uptake in Food Chain Transfer of Polychlorinated Biphenyls (PCBs) [1]

Abstract

In this research project we are investigating the possible role of the zebra mussel (Dreissena Polymorpha) in the increase of Polychlorinated Biphenyls (PCBs) in Great Lakes fish. A computational approach using high performance computers, such as a Silicon Graphics (SGI) workstation and a Cray supercomputer, to run code and visualize output of the model was used to deal with the volume of data involved in the research of this problem.

Background preparation for the investigation began during Saturday Tutorial sessions held at the National Environmental Supercomputing Center (NESC). This continued during the Summer Research Institute at Saginaw Valley State University by researching literature and contacting scientists doing work on zebra mussels and food chain modeling. A food chain model developed by John Connolly (Connolly, et.al.) was then applied to a simple food chain, in Western Lake Erie, generating preliminary data. The output data was then visualized on a Silicon Graphics scientific workstation using the Explorer visualization software.

Future plans include refinement of the model, locating accurate data for input variables and improvement of the current visualization. Along with continuing the research project, a course in environmental research and computational science will be developed and integrated into the curriculum in Bay City Central High School for the 1994-95 school year.

Research Objectives

Recent studies have shown that PCB levels have increased in Lake Erie fish over the last few years. The purpose of this research project is to explore whether zebra mussels are contributing to this increase. Determination of this role will be of importance to researchers, fisheries and recreationalists who use or examine any fish that have been or will be colonized by the zebra mussel.

Approach

The first step is to research current literature to identify a specific zone and food chain in western Lake Erie into which the zebra mussel has been inserted. Next, identify PCB concentrations in feces, pseudofeces and zebra mussel tissue in the food chain and examine how PCBs may be biomagnified. The next steps are to identify feeding rates at each trophic level, and mechanisms controlling uptake and loss of PCB's.

Based on the above steps we will construct a mathematical model of this food chain, and translate this mathematical model into a high level language such as FORTRAN. This code will run on a Silicon Graphics workstation or if necessary on the Cray Y-MP 8I located at the National Environmental Supercomputing Center (NESC) in Bay City, Michigan. The output of the computational model will be visualized using the Explorer software package on an SGI workstation.

Accomplishments to date

A typical food chain involving the zebra mussel in Lake Erie has been identified. The box diagram on the right hand side of Figure 1 shows a direct feeding relationship between yellow perch and zebra mussel as reported by Fisher. (Fisher, 1993). We used a mathematical model for bioaccumulation of contaminants by John Connolly (Connolly, 1992) to represent this relationship (Figure 2, page 107).

Data on zebra mussel PCB tissue concentrations have been collected from research literature and personal contacts with various researchers in the field. The tissue concentration used for the computational model was taken from Russ Kreiss (Kreiss, 1993). Published data on PCB uptake and loss mechanisms have not yet been identified.

The mass balance equation has been put into FORTRAN code to generate synthetic data of PCB levels over a 90 day time period. The output of this code was transferred into a 2D array data format readable by the Explorer visualization software and visualized using a Silicon Graphics workstation. The visualization in Figure 3, page 107, shows how the PCBs concentrations in prey and the water column relate to PCB concentrations in fish tissue (as represented by varying colors).

The visualization output has been analyzed and possible scenarios have been explored for variables which contribute to the increase of contaminants (e.g., PCB concentrations in the water column and in prey.) Although our visualization indicates that PCB concentrations in the perch are highly dependent on contaminant concentrations in the zebra mussel, at this point, due to lack of substantial data, no reliable conclusions can be drawn.

Future plans

The most pertinent need is to continue literature review and conversations with mentor scientists to obtain more accurate and current data as it becomes available.

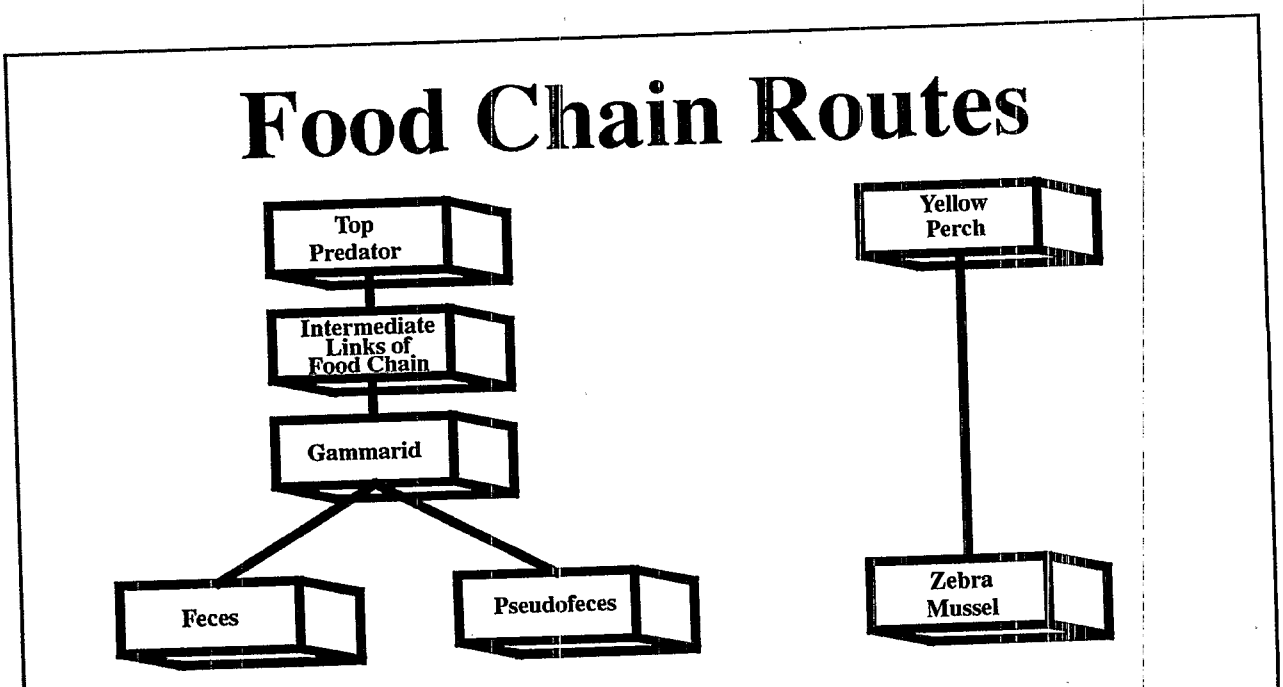


Figure 1. Feeding relationship between zebra mussel and yellow perch.

Mass Balance Equation Representing Uptake and Loss of Contaminants

$$dv/dt = K_u C_d + \propto C v_p - (K + G)v$$

v = concentration of contaminants (PCB's) in the animal (perch)
v_p = concentration of contaminants (PCB's) in the prey (zebra mussel)
K_u = uptake rate from water
C_d = concentration of contaminants (PCB's) in food
$\propto$ = assimilation efficiency of contaminants (PCB's) in food
C = consumption rate of food
K = excretion rate
G = growth rate of the animal

$K_u C_d$ = uptake of contaminants (PCB's) through gills
$\propto C v_p$ = contaminant (PCB) assimilation through feeding
(K + G)v = loss of contaminants (PCB's) through excretion and growth

Figure 2. Mass balance equation for bioaccumulation of contaminants (by J. Conolly)

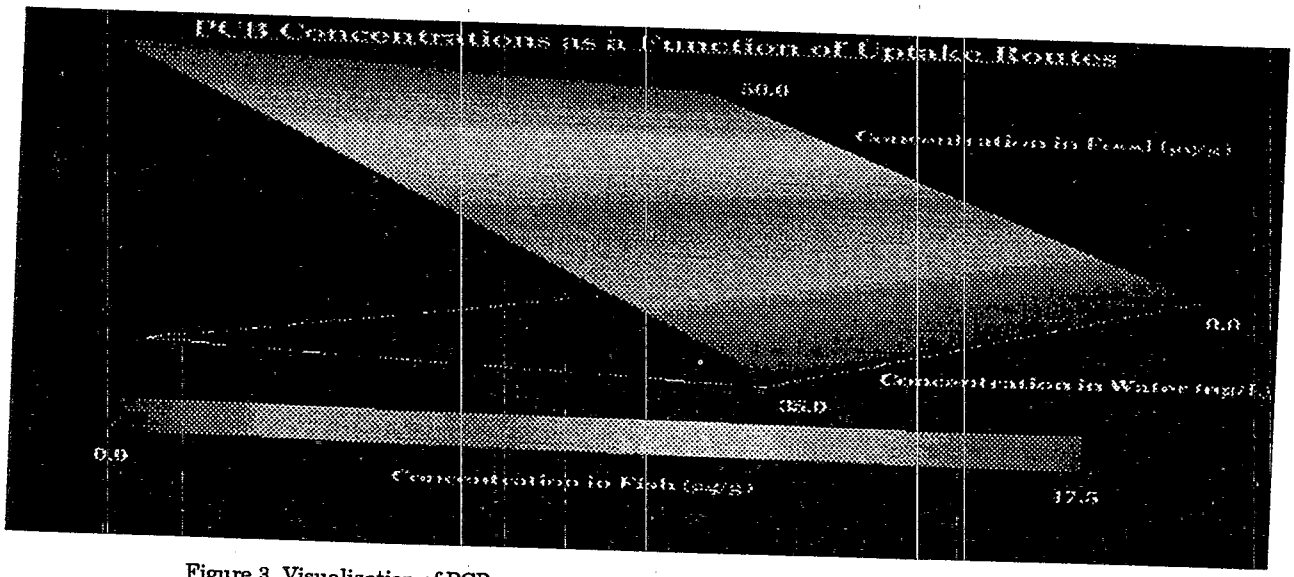


Figure 3. Visualization of PCB concentration in prey, water column and fish tissue.

The mathematical model will be refined to more accurately represent bioaccumulation and gain/loss of PCBs. A graph of the actual numbers which are represented by the model will also be added to our visualization.

A second food chain route (represented on the left hand side of Figure 1 (page 106), will also be identified by determining the intermediate food chain links from a specific species of gammarid to the upper trophic levels. The FORTRAN code will be refined to represent the new model and include place holders for other food sources.

As an extension of the EarthVision program, environmental research and computational science will be incorporated into the existing curriculum at Bay City Central High School using an integrated curriculum approach. Teachers in disciplines such as math, science, computer science and art have agreed to work together to develop units in the courses which will become part of a multidisciplinary environmental-computational-science curriculum. A Silicon Graphics workstation with ongoing access to the Cray is to be installed, and academic activities which will utilize the equipment will be developed.

Literature cited

Connolly, John P., Thomas F. Parkerton, James D. Quadrini, Scott T. Taylor, Andrew J. Thumann, *Development and Application of a Model of PCBs in the Green Bay, Lake Michigan Walleye and Brown Trout and Their Food Webs*, Environmental Engineering & Science Program, Manhattan College, October, 1992.

Fischer, Susan W., *The Role of Zebra Mussels in Contaminant Cycling*, Zebra Mussel Information Clearinghouse, Volume 4, Number 1, January/February 1993.

Kreiss, Russell, Telephone interview on September 9, 1993 regarding measured PCB concentrations in whole zebra mussels taken from Sterling State Park on Lake Erie.

---

1 M. Neal, J. Bisel, N. Mazloom, K. Kukla, E. Gatza, and J. Schroeder, Bay City Central High School EarthVision Team.

Summary of Heavy Metal Distribution Data in the Saginaw Bay and a Computer Model to Interpret these Data [1]

Abstract

The Great Lakes make up a substantial portion of the world's fresh water supply and the Saginaw Bay watershed is one of the largest and most heavily populated and industrialized segments feeding Lake Huron. Contamination of the Saginaw Bay is interesting both locally and nationally. Industrial and municipal waste treatment includes sophisticated removal techniques for most organic contaminants but removal of certain inorganic contaminants, particularly heavy metals, has proven to be difficult.

This project initially focuses on nickel ion species because many industrial pipes are lined with nickel alloys and this metal has recently come under closer scrutiny as a possible health hazard. It is hoped, however, that this research can be applied to a broad spectrum of heavy metal contamination in the Saginaw Bay.

Research Objectives

The project objectives include the collection and visualization of historical heavy metal data for the Saginaw Bay using computer graphics. A mathematical model of how the ions interact with the water, sediments and living organisms will then be developed, visualized and tested by comparing it to the actual historical data.

Approach

Historical data on contamination in the Saginaw Bay will be collected from sources such as the Saginaw Bay Watershed Initiative located on the campus of Saginaw Valley State University and also from STORET. A grid will be placed over a map

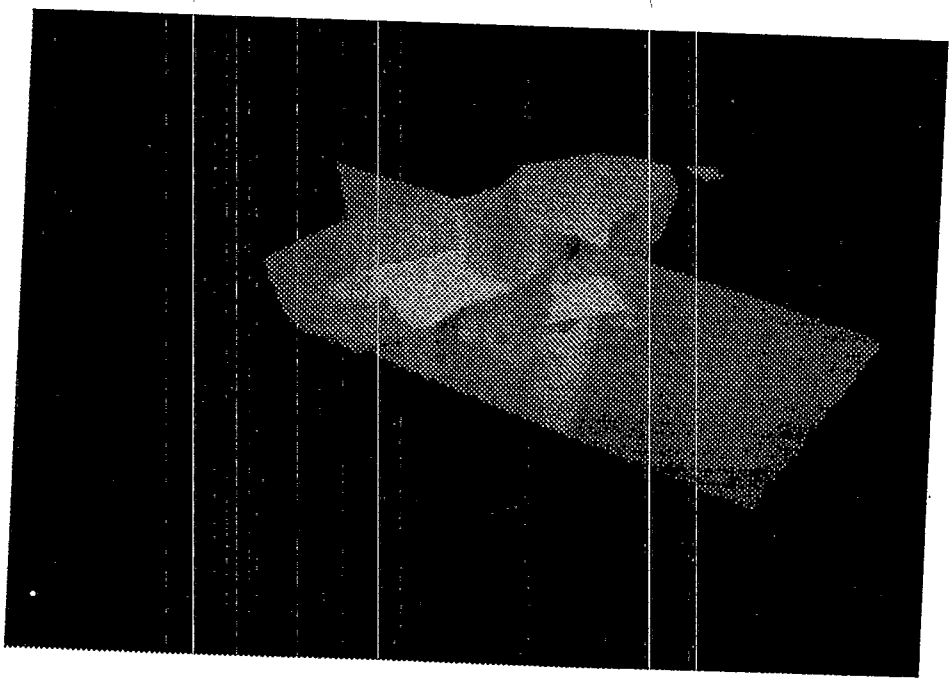


Figure 1: Contamination isosurface and transverse orthogonal slice rendering of historical data (1978) on nickel ion distribution in the Saginaw Bay. Red represents highest concentration. Total depth of slice is 50 cm. Red outer ring is an outline of Saginaw Bay. Bay City is at the bottom of the photo.

of the Bay and each sampling site will be given a set of coordinates. The sampling data will then be entered into a three dimensional array to include concentrations at various depths in the sediments and stored as a data file. Using the visualization software package called Explorer on a Silicon Graphics workstation, the data will be visualized in various ways including, contour mapping, isosurfacing, orthogonal slicing and volume rendering of ion concentrations. Data sets separated in time by several years will be visualized.

For the purpose of modeling, the Bay will be divided into three dimensional 'cells' (3.5km X 3.5km x 50 cm) centered on grid coordinate intersections. Mathematical expressions of the more obvious interactions between the ions and the environment within a 'cell' and each cell's influence on surrounding cells will be written as FORTRAN code. The code will be executed on the workstation or, if necessary, on a CRAY-YMP at NESC in Bay City, Michigan. The results will be visualized as before and compared to the historical data. By using an earlier historical data set as initial conditions and executing the program forward in time to a later historical data set, the model can be tested and modified.

Accomplishments to Date

Some historical data on heavy metal contamination in the Saginaw Bay (J. A. Robbins) have been encoded and visualized as described. These data values were interpolated for sampling sites where none existed for some visualizations.

The mathematical model is still in the early stages of development. In general, Saginaw Bay is a very complex system to model and there is a danger that, due to its shallow depth, any detailed model might be overwhelmed by storm activity. Nevertheless, nickel ion distribution in the Saginaw Bay does seem to be associated with high clay content in the sediment.

Future Plans

The model will be fine tuned and tested against other data sets. Several questions that will be investigated include: Does point source contamination affect distribution differently than multiple site contamination? How do instantaneous emissions compare with long term seepage in affecting heavy metal distribution? What, if any, contributions to heavy metal contamination in the Saginaw Bay come from the main body of Lake Huron itself?

Since this EarthVision project is part of the EPA's Grand Challenge for High School Students, the research will be used to teach students the fundamentals of computational science and the interdisciplinary nature of scientific investigations.

---

1 N. Sefcovic, C. Rohde, B. Weeden C. Stark, D.Sealey, and G. Barker, Saginaw Center for the Arts and Sciences EarthVision Team.

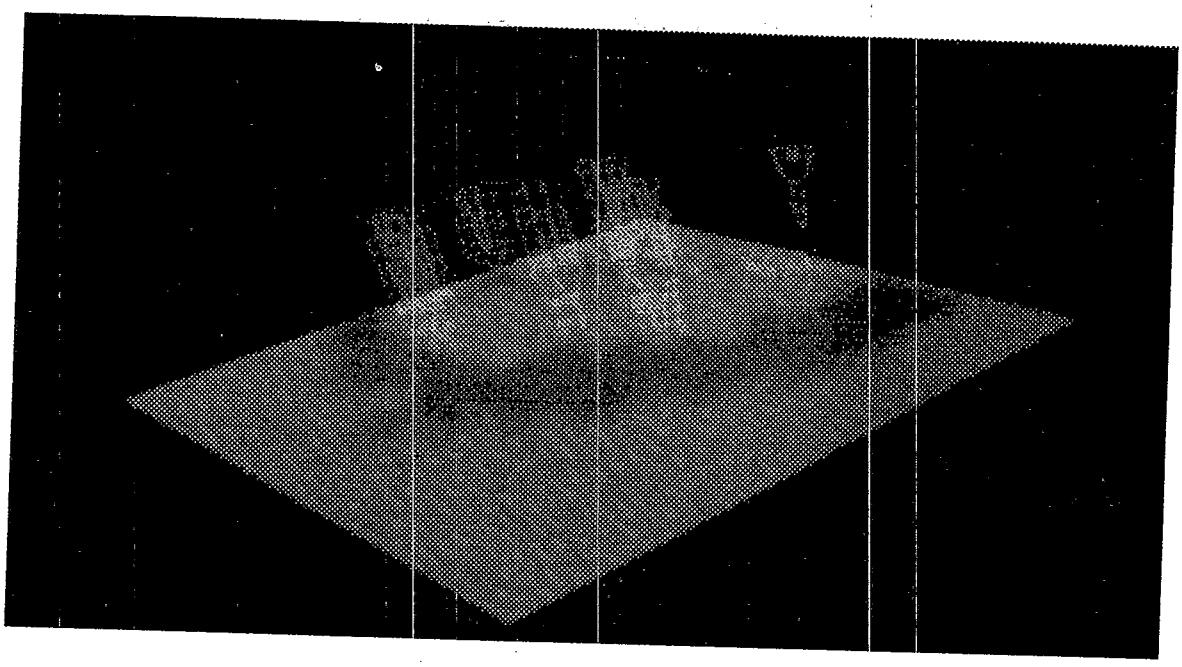


Figure 2. Visualization of historical data (Robbins) on nickel ion distribution in the Saginaw Bay. Cylinders represent core samples with contours at 5cm intervals. Orthogonal slice represents interpolated data (pink is highest concentration). Red outer ring is an outline of Saginaw Bay with Bay City at far left.

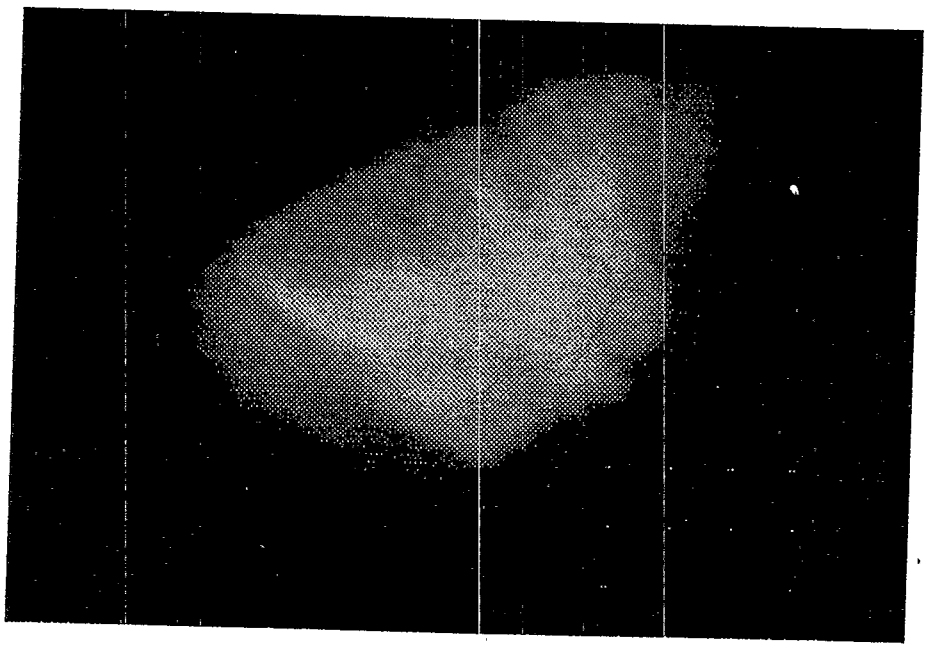


Figure 3. Volume rendering of same data. Purple is highest concentrations (about 60 mg / kg.) Bay City is at lower left.

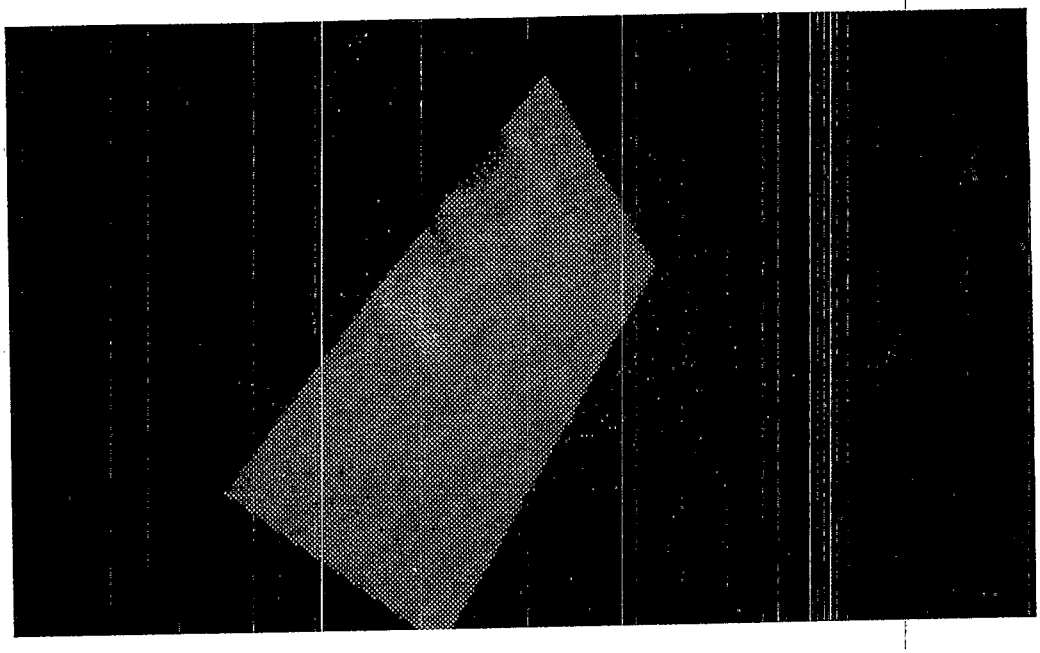


Figure 4. Combination multiple level contours and longitudinal orthogonal slice. Red is highest concentration. Bay City is at lower left.

EPA
United States
Environmental Protection Agency
National Environmental Supercomputing Center
135 Washington Avenue
Bay City, MI 48708-5845

Official Business
Penalty for Private Use
$300