United States
Environmental Protection
Agency

Office of
Research and Development
Washington, DC 20460

EPA/600/3-91/022
February 1991

Surface Waters Monitoring and Research Strategy - Fiscal Year 1991

Environmental Monitoring and Assessment Program

EPA/600/3-91/022
February 1991

ENVIRONMENTAL MONITORING AND ASSESSMENT PROGRAM (EMAP) SURFACE WATERS MONITORING AND RESEARCH STRATEGY FISCAL YEAR 1991

Environmental Research Laboratory
Office of Research and Development
United States Environmental Protection Agency
Corvallis, Oregon 97333

Environmental Monitoring Systems Laboratory
Office of Research and Development
United States Environmental Protection Agency
Las Vegas, Nevada 89153

Environmental Monitoring Systems Laboratory
Office of Research and Development
United States Environmental Protection Agency
Cincinnati, Ohio 45219

February 1991

Printed on Recycled Paper

EMAP-SURFACE WATERS MONITORING AND RESEARCH STRATEGY
FISCAL YEAR 1991

February 1991

By:

Steven G. Paulsen[a], David P. Larsen[b], Philip R. Kaufmann[c], Thomas R. Whittier[d], John R. Baker[e], David V. Peck[e], Joan McGue[e], Robert M. Hughes[d], Dennis McMullen[f], Don Stevens[d], John L. Stoddard[d], James Lazorchak[g], Wesley Kinney[h], Anthony R. Selle[d], Randy Hjort[d]

With contributions from:

W. Scott Overton[i], Jim Pollard[e], Daniel Heggem[h], Gary Collins[g], Marilyn Morrison[d], Colleen Birch Johnson[d], Sandra Thiele[d], Nita Tallent-Halsell[e], Kit Peres[e], Susan J. Christie[d], and Janet Mello[d]

The research described in this report has been funded by the U.S. Environmental Protection Agency. This document has been prepared at the EPA Environmental Research Laboratory in Corvallis, Oregon, through cooperative agreements CR814701 with University of Nevada-Las Vegas, CR815168 with Utah State University, and CR815422 with Oregon State University, and contract numbers 68-C8-0006 with ManTech Environmental Technologies, Inc., and 68-03-3249 with Lockheed Engineering and Sciences Company. The report has been subjected to the Agency's peer and administrative review and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

[a] Environmental Research Center, University of Nevada-Las Vegas, NV.
[b] U.S. EPA Environmental Research Laboratory, Corvallis, OR.
[c] Utah State University, c/o U.S. EPA Environmental Research Laboratory, Corvallis, OR.
[d] ManTech Environmental Technologies, Inc., U.S. EPA Environmental Research Laboratory, Corvallis, OR.
[e] Lockheed Engineering and Sciences Company, Environmental Monitoring Systems Laboratory, Las Vegas, NV.
[f] Technology Applications, Inc., Environmental Monitoring Systems Laboratory, Cincinnati, OH.
[g] U.S. EPA Environmental Monitoring Systems Laboratory, Cincinnati, OH.
[h] U.S. EPA Environmental Monitoring Systems Laboratory, Las Vegas, NV.
[i] Oregon State University, Corvallis, OR.

TABLE OF CONTENTS

Section **Page**

List of Illustrations viii
List of Tables xi
Executive Summary xiii

1.0 INTRODUCTION 1

1.1 Content and Organization of the Research Plan 1
1.2 An Overview of EMAP 2
1.3 Legislative Mandate for Surface Water Protection 3
1.4 Current Efforts to Monitor Surface Waters 4
1.4.1 State 305(b) Water Quality Monitoring 4
1.4.2 National Contaminant Biomonitoring Program 5
1.4.3 National Stream Quality Accounting Network 5
1.4.4 Hydrologic Benchmark Network 5
1.5 Unanswered Questions for Surface Waters 5
1.5.1 Societally Important Surface Water Values 6
1.5.2 Hazards to Aquatic Systems 8
1.6 Specific Objectives of EMAP-Surface Waters 8
1.7 Focus and Purpose of the Research Plan 8

2.0 APPROACH AND RATIONALE 9

2.1 Overview of the EMAP-Surface Waters Approach 9
2.1.1 Design 13
2.1.2 Indicator Approach 17
2.2 Data Quality 20
2.2.1 The Data Quality Hierarchy 24
2.2.2 The Role of DQOs in EMAP 25
2.2.3 DQOs in the Surface Water Component 25
2.3 Reporting 28
2.4 EMAP-Surface Waters Program Limitations 28

3.0 MONITORING NETWORK DESIGN 29

3.1 Populations, Sample Units, Frames, and Other Sampling Concepts 29
3.1.1 Surface Water Populations: Defining Lakes and Streams 29
3.1.2 Subpopulations 30
3.1.3 Sample Units 31
3.1.4 Frame Development 32
3.1.4.1 Lake Frame 32
3.1.4.2 Stream Frame 35
3.2 The Two-stage Sample: Tiers 1 and 2 36
3.2.1 Tier 1 Resources 36
3.2.1.1 Lakes 36
3.2.1.2 Streams 37
3.2.1.3 Tier 1 Association Rules 41
3.2.1.4 Adequacy of the Grid for Tier 1 Sample Selection 41
3.2.2 Tier 2 Sampling 41
3.2.2.1 Tier 2 Site Selection Example 42
3.2.2.2 Mechanics of Tier 2 Site Selection (Design Option 2) 44

3.3 Higher Grid Densities 45
3.4 Summary 50

4.0 INDICATOR DEVELOPMENT AND EVALUATION 55

4.1 Conceptual Framework for Indicators of Condition 55
4.1.1 Response Indicators 56
4.1.2 Exposure and Habitat Indicators 56
4.1.3 Stressor Indicators 59
4.2 Establishing Unimpaired (Nominal) Condition 59
4.3 Response Indicator Selection 60
4.3.1 Lake Indicators 61
4.3.2 Stream Indicators 65
4.4 Exposure and Habitat Indicators 66
4.5 Stressor Indicators 70
4.6 Strategy for Indicator Development and Implementation 71

5.0 POPULATION ESTIMATION AND ANALYTICAL APPROACHES 77

5.1 Estimating Population Characteristics 77
5.1.1 Population Extrapolation from EMAP Sample Data 77
5.1.2 Tier 1 Resource Estimates 79
5.1.3 Tier 2 Estimates and Descriptions 80
5.1.4 Analytical Approaches for Detecting Population Differences, Changes, and Trends 81
5.2 Components of Variance in Regional Population Sampling 85
5.2.1 Variance Model 85
5.2.2 Indicator Variance Magnitudes Estimated from Existing Data 86
5.2.2.1 Alternatives for Comparing Variances 86
5.2.2.2 Coefficients of Variation 91
5.2.2.3 Relative Variances 91
5.3 Design Considerations for Estimating Population Status 92
5.3.1 Effects of Index Variability on Estimates of Population Status 92
5.3.1.1 Theoretical Basis 92
5.3.1.2 Simulated Effects of Index Variability 95
5.3.2 Acceptable Levels of Index Variability for Status Estimates 97
5.3.3 Required Indicator Sampling Intensity for Status Descriptions 99
5.4 Examples of Effects of Indicator Variability on Population Descriptions and Trend Detection 104
5.4.1 Eastern Lake Survey 104
5.4.2 Minnesota Pollution Control Agency 104
5.4.3 Vermont Department of Environmental Conservation 104
5.4.4 Ohio Environmental Protection Agency 104
5.4.5 Descriptions of Status 105
5.4.6 Differences and Changes in Regional Subpopulations 107
5.4.7 Regional Trends 107

6.0 FIELD SAMPLING DESIGN 113

6.1 Index Concept 114
6.1.1 General Design Objectives of Index Sampling 114
6.1.2 Applications of the Index Concept in Synoptic Assessments 114
6.1.2.1 National Stream Survey (NSS) 114
6.1.2.2 National Lake Survey (NLS) 115
6.2 Temporal Considerations for Field Sampling 116

6.3 Index Sampling Design for Lakes and Reservoirs 116
6.3.1 Temporal Variability 116
6.3.2 Summer Lake/Reservoir Index Period: The Proposal 116
6.3.3 Spatial Variability within Lakes and Reservoirs 119
6.3.4 Spatial Sampling Index for Lakes and Reservoirs: The Proposal 120
6.4 Index Sampling Design for Streams and Rivers 120
6.4.1 Temporal Variability 120
6.4.2 Summer Baseflow Stream/River Index Period: The Proposal 121
6.4.3 Spatial Variability within Streams/Rivers 122
6.4.3.1 Upstream-downstream Variations 122
6.4.3.2 Other Stream Habitat Variations 123
6.4.4 Spatial Sampling Index for Streams and Rivers: The Proposal 123
6.5 Conclusions 124

7.0 LOGISTICS APPROACH 125

7.1 Logistics Implementation Components 125
7.2 Logistics Issues 127
7.2.1 Staffing 127
7.2.2 Access 127
7.2.3 Data Confidentiality 128
7.3 Field Operations Scenario: Lakes 128
7.3.1 General Logistics Scenario 128
7.3.2 Daily Activities Scenario 129
7.4 Field Operations Scenario: Streams 132
7.5 Organizational Structure 132

8.0 QUALITY ASSURANCE PROGRAM 135

8.1 Data Quality Requirements 136
8.1.1 Precision and Bias 136
8.1.2 Comparability 139
8.1.3 Completeness 139
8.1.4 Representativeness 139
8.1.5 Tolerable Background Levels 139
8.1.6 Method Detection Limits for chemical Constituents 140
8.2 Organization and Staffing Requirements 140
8.3 Quality Assurance Documentation 140
8.4 Quality Control Guidelines 141
8.4.1 Biological Measurements 144
8.4.2 Chemical Measurements 144
8.4.3 Habitat Quality and Site Characterization Measurements 146
8.4.4 Archival of Samples and Specimens 147
8.5 Data Review, Verification, and Validation 147
8.6 Assessment of Data Quality 148
8.7 Quality Assurance Reporting 149

9.0 INFORMATION MANAGEMENT 151

9.1 Definition of Information Management 151
9.1.1 Overview of EMAP-Surface Waters Information Management 151
9.1.2 Objectives of Surface Waters Information Management 151
9.2 Initial System Concept 152
9.2.1 User Requirements 152
9.2.1.1 Issues 152

9.2.1.2 Data Categories ... 152
9.2.1.3 Identified User Groups ... 152
9.2.1.4 Data User Requirements ... 152
9.2.2 System Management and Functional Requirements ... 153
9.2.3 Data and Information Flow ... 154
9.3 Organization of EMAP-Surface Waters Information Management Program ... 154
9.3.1 Surface Water Information Manager ... 154
9.3.2 Surface Waters Information Center (SWIC) ... 154
9.4 Operational Specifications ... 154
9.4.1 Project Management System ... 157
9.4.1.1 Communications ... 157
9.4.1.2 Sample Tracking Information ... 157
9.4.2 Data and Sample Collection ... 158
9.4.3 Processing and Storage of Indicator Data ... 158
9.4.4 Data Access and Transfer ... 158
9.4.5 Data Analysis and Reporting ... 158
9.4.6 Data Documentation, Access, and Archival ... 158
9.4.7 Geographic Information System (GIS) ... 159
9.4.8 Integration ... 159
9.4.9 Existing Information/Data ... 159
9.5 QA/QC for Information Management Activities ... 159
9.5.1 Field Data Collection ... 159
9.5.2 Data Transfer ... 160
9.5.3 Sample Tracking ... 160
9.5.4 Data Entry/Verification ... 160
9.5.5 Data Validation ... 160
9.5.6 Data Analysis ... 160
9.5.7 Archive and Backup ... 161
9.5.8 Configuration Management (CM) ... 161
9.6 Resources: Overview ... 161
9.7 Implementation Plan ... 161

10.0 COORDINATION ... 163

10.1 Within EMAP ... 163
10.1.1 Tier 3 Activities in Surface Waters ... 163
10.1.2 Across EMAP Elements ... 163
10.2 Within EPA (Office of Water and Regions) ... 164
10.3 Other Federal Agencies ... 165
10.3.1 U.S. Fish and Wildlife Service ... 165
10.3.2 USGS National Water Quality Assessment Program ... 165
10.4 State Agencies ... 166
10.5 Research Organizations ... 166
10.6 Conclusions ... 167

11.0 EXPECTED OUTPUTS AND TIMELINES ... 169

11.1 Reporting ... 169
11.1.1 Data Evaluation Reports ... 169
11.1.2 Internal Program Reports ... 169
11.2 Timelines ... 170

12.0 FISCAL YEAR 1991 FIELD AND ANALYSIS ACTIVITIES 171

12.1 General EMAP-Surface Waters 171
12.2 TIME Project 171
12.3 Data Analysis Activities 172
12.4 Research Needs 172

REFERENCES 173

LIST OF ILLUSTRATIONS

Figure		Page
2-1	U.S. Environmental Protection Agency regions	10
2-2	Aggregations of Omernik's ecoregions	11
2-3	Concept of a four-tiered approach in EMAP	12
2-4	The baseline grid (not randomized) for North America containing about 12,600 points in the conterminous United States	15
2-5	The landscape characterization hexagons are one-sixteenth of the total area and are centered on the sampling points	16
2-6	Enhancement factors for increasing the base grid density	16
2-7	Spatially interpenetrating samples on a 4-year rotating cycle	18
2-8	General approach for identifying indicators	21
2-9	Indicator approach for EMAP-Surface Waters showing candidate indicators and the top-down approach to problem identification and diagnosis of probable cause	21
2-10	Cumulative frequency distribution for Index of Biotic Integrity in a region showing how data might be displayed	22
2-11	Long-term assessment objectives for EMAP-Surface Waters	22
2-12	The three phases of the data quality objective process	22
2-13	Quality objective hierarchy in the data quality objective process	24
3-1	Example of the lake frame	33
3-2	Example of the stream frame, with nodes identifying stream segments that are the stream sample units	34
3-3	Cumulative size distribution for lakes in the Northeast	38
3-4	Map of the spatial distribution of lakes in the Northeast	43
3-5a	Hexagon capture rates for lakes (1-250 ha) in the Northeast	46
3-5b	Areas that group similar numbers of lakes for Tier 2 selection	47
3-5c	The structure of the EMAP grid, illustrating the fourfold decomposition that will be the basis of the interpenetrating sample design	48
3-5d	Selected hexagons containing one or more lakes that will constitute the Tier 2 sample	49
3-6	Acid sensitive regions of interest in the Northeast	51

3-7 Areas of the Northeast requiring threefold augmentation of the grid to supply enough sample sites for the project 53

4-1 The conceptual process for proceeding from measurements to indicators to assessment of condition 57

4-2 The process of proceeding from measurements on fish assemblages to indicators such as Index of Biotic Integrity (IBI) or Index of Well Being (IWB) 58

4-3 The indicator development process, showing the objectives, methods, and evaluation techniques used in each phase 74

5-1 Frequency distributions are plotted for any identifiable population 78

5-2 Example of estimated distribution plots with upper confidence bounds, generated both as numbers (upper plot) and area (lower plot) 82

5-3 An example of comparing population distributions using the chi-square test 84

5-4 Convolution of a standard normal distribution with normally distributed index variability 94

5-5 Simulated effect of index measurement variance on survey results observed from sampling 100 lakes from a hypothetical asymmetrically distributed population with a true mean of 25 and an SD_{unit} of 50 96

5-6 Absolute (a) and relative (b) bias in population percentiles for ratio of index variance to among-sampling unit variance 98

5-7 Nomogram for determining sample allocation between replications within sample unit and adding new sample units in a region based on ratio of cost for replications and new sites and the variability within the index period and among sample units 100

5-8 Cumulative distribution functions (CDFs) describe the status of regional populations of lakes and streams, as this figure illustrates for total phosphorus in the Northeast, based on results of the Eastern Lake Survey 105

5-9 Cumulative distribution functions (CDFs) of the Index of Biotic Integrity for Ohio streams, and chlorophyll-*a* for Minnesota illustrate the status of two indicators having relatively large index variation relative to population variation 106

5-10 Cumulative distribution functions (CDFs) of the Index of Biotic Integrity for Ohio streams illustrating situations in which regional subpopulations are quite similar (Region 3 versus Region 5) and situations in which clear differences among regional subpopulations occur (Region 1 versus Region 5) 108

5-11 Cumulative distribution functions (CDFs) showing the similarity in the Secchi disk transparency status of lakes in Vermont measured during two index periods of the same year, July and August 1985 (upper panel); median values are identified mid-graph. The lower panel shows another way of displaying data when sites are paired 109

5-12 Cumulative distribution functions (CDFs) using the Index of Biotic Integrity for Ohio streams show year-to-year differences in the status of small streams 110

5-13 Cumulative distribution functions (CDFs) taken from the Vermont lakes data set on trophic condition showing comparisons made on paired sites: (a) total phosphorus during March, comparing 1978 and 1979, and (b) Secchi disk transparency during July, comparing 1987 and 1988 111

5-14 Superimposed yearly cumulative distribution functions (CDFs) show the confusing "spaghetti" pattern in multiple overlays of CDFs, based on nine years of data on Secchi disk transparency in Vermont lakes (upper panel). Results obtained from yearly CDFs can be more clearly presented to illustrate trends, or lack of trends, by selecting yearly values representing different quartiles (lower panel) 112

6-1 Seasons of lowest stream flow in the conterminous United States 117

7-1 Flowchart of potential EMAP-Surface Waters daily field activity 131

8-1 Hierarchical model of variance components of importance to EMAP-Surface Waters 138

8-2 Stages in the data acquisition and management process for EMAP-Surface Waters where data quality can be controlled or assessed 143

9-1 Data and information flow for the EMAP-Surface Waters Information Center 155

9-2 Concept of the EMAP-Surface Waters Information Center in relation to the EMAP Information Center 156

LIST OF TABLES

Table		Page
2-1	Sample Fact Sheet and Logic Statement for EMAP-Surface Waters	26
3-1	Estimates of Stream Reach Length Based on Segments Represented on Three 1:100,000-scale Topographic Maps	35
3-2	Summary of the Size Distribution of Lakes in the Northeast (EPA Regions 1 and 2)	39
3-3	Summary of Average Capture Rate of Perennial Stream Segments, Lakes, and Reservoirs (Sample Units per Hexagon) in 40-km^2 Hexagons for the Humid (Eastern) and Arid (Western) Regions of the United States	42
3-4	Estimates of Geographic Regions of Interest in the Northeast Relative to Acidic Deposition and Number of Tier 1 Hexagons, Number of Lakes, and Number of Hexagons with Lakes	52
3-5	Estimates of Tier 1 Coverage of Subpopulations of Interest for Acidic Deposition in the Northeast	52
4-1	Candidate Indicators for Inland Surface Waters	59
4-2	Evaluation of Some Candidate Indicators for Inland Surface Waters by EMAP Selection Criteria	62
4-3	Suggested Physical Habitat Measures for Lakes	67
4-4	Proposed Chemical Measurements Contributing to EMAP Geochemical Habitat Quality Indices	69
4-5	Linkages Between Potential Endpoints, Measurements, Metrics, and Response Indicators for EMAP-Surface Waters	72
4-6	Linkages Between Potential Endpoints, Measurements, and Exposure Indicators for EMAP-Surface Waters	73
5-1	Tier 1 Inventory Estimates and Estimated Variance of Estimates are Provided from the Tier 1 Sample for All Resources and Classes of Resources	80
5-2	Coefficients of Variation (CV) of Surface Water Indicators -- A Number of Examples from Lake and Reservoir Studies	87
5-3	Coefficients of Variation (CV) of Surface Water Indicators -- A Number of Examples from Stream and River Studies	88
5-4	Components of Indicator Variability as a Percentage of Among-Sample Unit Variance (S^2_{unit}) -- Examples from Lake and Reservoir Studies	89
5-5	Components of Indicator Variability as a Percentage of Among-Sample Unit Variance (S^2) -- Examples from Stream and River Studies	90
5-6	Bias of Apparent F(x) as a Function of d^2, Assessed for the Normal Distribution Function with Homogeneous Normal Measurement Error	95

5-7 Rough Categories of Index Measurement Precision (Sum of Analytical, Temporal, Spatial Components) Relative to Regional Variance for Possible EMAP Indicators 101

5-8 Rough Costs of Index Measurement Replication Relative to Costs of Additional Sites (Sample Lakes or Streams) 102

6-1 Time Scales and Sources of Variability in EMAP Indicator Characteristics -- Some Examples 118

7-1 EMAP Logistical Elements for Implementation of Surface Waters Monitoring Programs 125

7-2 Proposed General Field Measurement Variables for Lakes 128

8-1 Criteria for Selection of Appropriate Sampling and Analytical (or Measurement) Methodology 137

8-2 Responsibilities of the EMAP QA Coordinator and Surface Waters QA Officer 141

8-3 Quality Assurance Related Documentation of EMAP-Surface Waters 142

8-4 Quality Control Activities Associated with Biological Measurements 145

8-5 Quality Control Activities Associated with Chemical Measurements 146

11-1 Proposed EMAP-Surface Waters Implementation Schedule 170

EXECUTIVE SUMMARY

EMAP-SURFACE WATERS MONITORING RESEARCH STRATEGY -- FISCAL YEAR 1991

THE ENVIRONMENTAL MONITORING AND ASSESSMENT PROGRAM

The Environmental Monitoring and Assessment Program (EMAP) was initiated in 1988 to provide improved information on the current status of ecological resources in the United States and on long-term trends in their condition. Seven broad resource categories have been defined within EMAP: near-coastal waters, wetlands, inland surface waters, the Great Lakes, forests, arid lands, and agroecosystems. In addition, seven coordination and integration functions have been established to assist the groups working in the seven resource areas and to ensure consistency:

1. Monitoring design
2. Development of indicators of ecological condition
3. Landscape characterization
4. Quality assurance and quality control
5. Field sampling logistics
6. Information management
7. Integration and assessment

This document describes the rationale, objectives, and primary elements of the EMAP-Surface Waters program, which has been designed to assess the condition of U.S. lakes and streams. Separate research plans are being prepared for the other EMAP resource groups, as well as for each of the coordination and integration functions. As an integrated multi-resource program, the success of EMAP, and of EMAP-Surface Waters, will depend on the close cooperation and coordination among these various program components.

OBJECTIVES OF EMAP-SURFACE WATERS

The overall goal of EMAP-Surface Waters is to provide a quantitative assessment of the current status of and long-term trends in surface water condition on regional and national scales. The specific, long-term objectives of EMAP-Surface Waters are:

- Estimate the current extent (location, number, surface area or length) of lakes and streams, on regional and national scales, with known confidence.

- Estimate the current status, changes, and trends in indicators of the condition of the nation's lakes and streams, on regional and national scales, with known confidence.

- Monitor indicators of pollutant exposure and habitat condition within lakes and streams and seek associations between human-induced stressors and ecological condition that identify possible causes of adverse effects.

  Publish annual statistical summaries on the extent and the status of indicators of ecological condition of lakes and streams, and periodic interpretive reports on the status and trends indicators of ecological condition of lakes and streams to the EPA Administrator and the public.

OVERVIEW OF THE EMAP-SURFACE WATERS APPROACH

This research plan presents the rationale, objectives, approach, and plan for establishing the monitoring program that will be required to assess the status and trends in ecological condition of U.S. surface waters. It also includes a description of the EMAP component that will specifically address the response of surface waters to acidic deposition [the TIME project (Temporally Integrated Monitoring of Ecosystems)].

EMAP-Surface Waters is primarily a top-down program being designed to evaluate the health or condition of surface waters with respect to endpoints of concern. Given multiple societal concerns and values, EMAP-Surface Waters proposes to evaluate the condition of lakes and streams with respect to each of these issues:

- Biological integrity
- Trophic condition
- Fishability

The design and indicator strategies described in this report are intended to be adequate for exploring a variety of questions on national and broad regional scales rather than to answer site-specific questions. In this context, EMAP-Surface Waters will provide information with which we can determine (1) what proportions of our lakes, streams, and rivers are degrading or improving, where, and at what rate, (2) what the likely causes are of adverse effects, and (3) whether or not systems are responding as expected to control and mitigation programs. Information derived from the program will help in setting national priorities.

COORDINATION WITH OTHER MONITORING PROGRAMS

EMAP-Surface Waters is being designed and funded by the U.S. Environmental Protection Agency's (EPA) Office of Research and Development (ORD). However, the states, other offices and regions within EPA, and other federal agencies [e.g., U.S. Fish and Wildlife Service (USFWS); U.S. Geological Survey (USGS)] have contributed to its development and will participate in the collection and use of EMAP data. This coordination will avoid duplicative monitoring efforts, facilitate the exchange of data, and increase the expertise available for refining the program design and interpreting the monitoring results. EMAP is not intended as a substitute for other monitoring and research efforts, but instead will provide a framework for integrating existing and new data.

EMAP'S HIERARCHICAL DESIGN

Ultimately, EMAP will involve four tiers, or types of activities related to monitoring and assessing ecological condition (Figure 1):

- **Tier 1:** Landscape characterization, to determine the distribution and extent (numbers and area) of ecological resources in the United States.
- **Tier 2:** Field sampling, to include a suite of biological, chemical, and physical measurement to determine the conditions of a subset of Tier 1 sites.
- **Tier 3:** More intensive sampling to focus on special subpopulations of concern or to obtain more detailed diagnostic information.
- **Tier 4:** Ecological research, to complement the monitoring data collected in Tiers 1, 2, and 3.

This document addresses activities primarily at Tiers 1 and 2. The Tier 3 activity, the TIME project, is discussed briefly throughout this report, as it pertains to modifications of specific applications of the general approach.

Tier 1 Sampling Frame

To achieve its objectives, EMAP-Surface Waters will use standardized sampling methods and an unbiased probability-based sampling design to monitor lakes and streams over broad geographic areas and for multiple decades. The proposed design strategy is based on a permanent national sampling framework consisting of a hexagonal plate containing a triangular grid of approximately 12,600 points placed randomly over the conterminous United States (Figure 2); a similar array is available for Alaska and Hawaii. A defined area (40-km^2 hexagon) around each point will be characterized using existing maps, aerial photography, satellite imagery, and existing databases of landscape information. These 40-km^2 hexagons describe an area sample representing one-sixteenth of the area of the United States and provide the basis for the Tier 1 estimates of surface water extent and distribution.

The characterization carried out on the 40-km^2 hexagons will provide an inventory of all aquatic resources within them. Thus it will provide a one-sixteenth sample of the extent, abundance, and distribution of any aquatic resource that can be identified during the landscape characterization. The complete landscape description at Tier 1 ensures that all ecological resources are included by definition of the monitoring coverage.

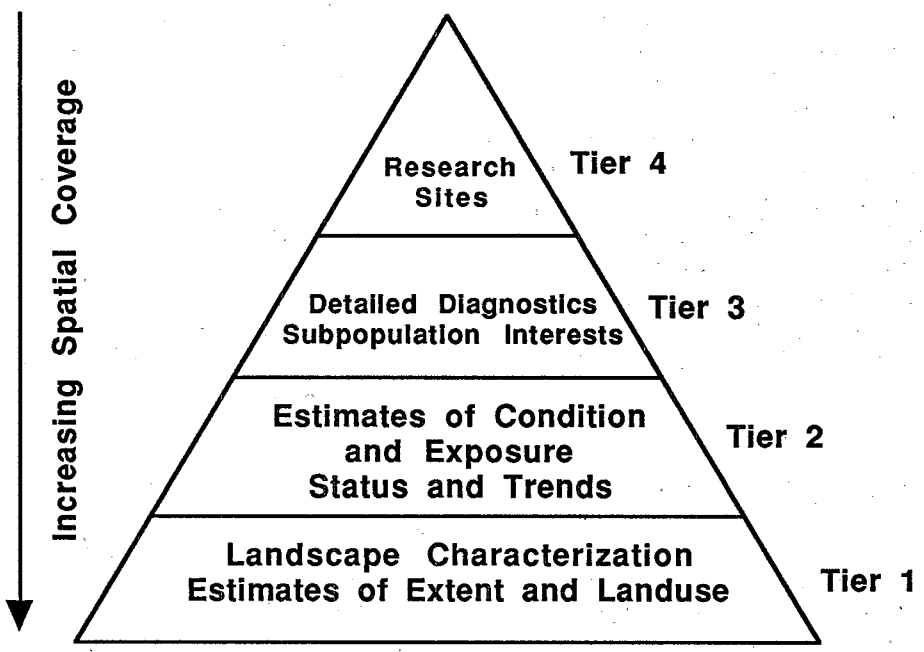


Figure 1. Concept of a four-tiered approach in EMAP. Spatial coverage is maximized in lower tiers; temporal coverage increases at the higher tiers.

Tier 2 Sampling Frame

A subsample of the Tier 1 sample, selected by probability methods and stratified as needed to ensure an adequate sample size for each surface water class and region of interest, will be visited for detailed site characterization of indicators of condition of the resource. Data obtained from these field samples will constitute the information to be used for reporting on regional status and trends in ecological condition and exposure. This second stage of the double sample will constitute Tier 2 of the design.

The Tier 2 sample will be implemented for resources of current concern. If a need to monitor an additional class of resource is identified, it will be necessary to return to the Tier 1 level, identify the extent and location of that resource, and select appropriate sites for field sampling. However, the existence of the landscape characterization provides a ready-made frame for sample selection, so that a field program can be mounted with minimal preparation. This provides the potential for quick response to new or emerging issues of ecological condition.

The Tier 2 sites will be sampled on a four-year cycle, that is, one-fourth of the sites in a region will be visited each year. A small subset of sites will be visited annually during the first decade of the program to estimate year-to-year variation and to identify any important trends early. By the fifth year, all sites will have been sampled and a second cycle will begin, using the same subsets of resource units. The sample sites will be partitioned so that the basic systematic triangular grid (at one-fourth the density) will be retained in each annual subsample, to maintain a nearly uniform spatial distribution of sites each year. Analyses of surface water status will be reported annually, as four-year running averages over the four interpenetrating sample subsets. A few new sites will be selected each year, and others will be dropped, to minimize any bias that might be

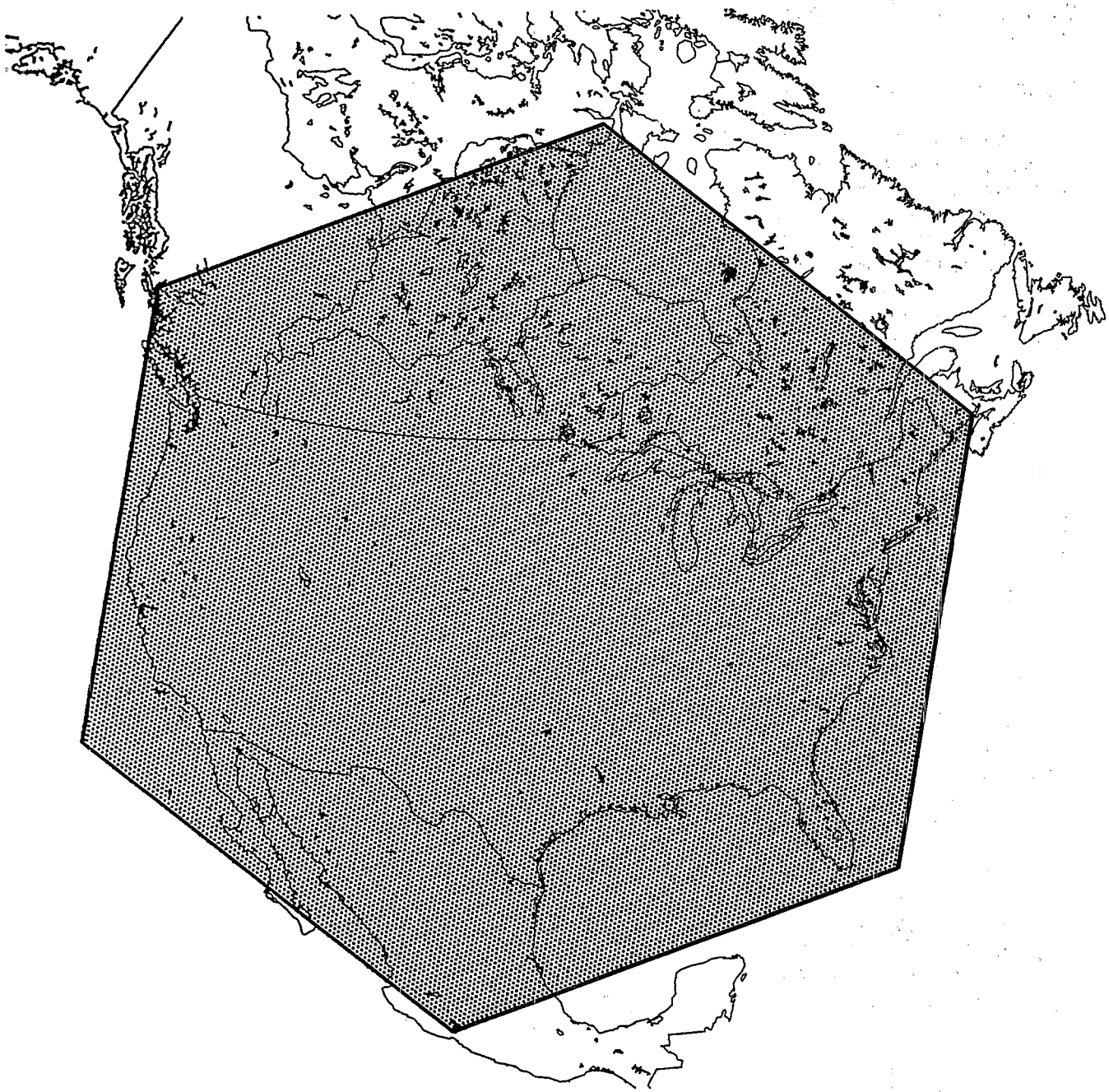


Figure 2. **The baseline grid (not randomized) for North America, containing about 12,600 points in the conterminous United States. Spacing between points is about 27 kilometers.**

introduced by the knowledge that certain sites are EMAP sites, resulting in differential treatment of those sites.

Flexibility of the EMAP Design

An important characteristic of the EMAP design is its flexibility. The results from Tiers 1 and 2 can be summarized according to any subpopulation (e.g., surface water class) or spatial partitioning (e.g., region) of interest. The basic EMAP grid can easily be enhanced for greater sampling density in regions or for lake or stream classes of particular concern. The landscape characterization at Tier 1 provides a ready-made frame for sample selection, so that new or supplemental field programs can be implemented quickly in response to new or emerging issues. The outputs from the EMAP-Surface Waters monitoring network can be analyzed and expressed in a variety of ways, to address a diversity of policy-relevant questions. The EMAP design will provide information on specific indicators measured during a specific index period, as a "snapshot" of the overall condition of a system.

INDICATOR APPROACH

In the EMAP-Surface Waters approach, ecological risk assessment will focus on the ecosystem attributes or endpoints of concern. Two types of endpoints are used in risk assessment:

1. **Assessment endpoints:** Formal expressions of the actual environmental values that are to be protected. They should have unambiguous operational definitions, have social or biological relevance, and be amenable to prediction or measurement (e.g., probability of a >10% reduction in game fish production).

2. **Measurement endpoints:** The quantitative results of the measurements taken to characterize assessment endpoints. Measurement endpoints must correspond to or be predictive of assessment endpoints (e.g., the fraction of streams that exceed the LC_{50} for a toxicant to largemouth bass).

The term "indicator" has been adopted within EMAP to refer to the specific environmental characteristics to be measured or quantified through field sampling, remote sensing, or compiling of existing data. Four types of indicators are being selected and developed for EMAP-Surface Waters:

1. **Response indicator:** A characteristic of the environment measured to provide evidence of the biological condition of a resource at the organism, population, community, or ecosystem level or organization (e.g., trophic state index, fish assemblage).

2. **Exposure indicator:** A characteristic of the environment measured to provide evidence of the occurrence or magnitude of contact with a physical, chemical, or biological stressor (e.g., nutrient concentrations, tissue residues, toxicity tests).

3. **Habitat indicator:** A physical, chemical, or biological attribute measured to characterize the condition necessary to support an organism, population, community, or ecosystem in the absence of pollutants (e.g., availability of snags, substrate of stream bottom, vegetation type and extent).

4. **Stressor indicator:** A characteristic measured to quantify a natural process, an environmental hazard, or a management action that causes changes in exposure and habitat (e.g., land use).

Assessment endpoints by definition must relate to the environmental values identified for each resource category in EMAP. Therefore, the ideal measurement endpoints for EMAP are response indicators. EMAP will collect data on response, exposure, and habitat indicators at its field sampling sites. Data on land use and land cover will be collected as part of the landscape characterization activity within EMAP. Additional stressor indicators will be assembled from other sources.

Selection of assessment endpoints, and the subsequent selection of indicators (Figure 3), requires consideration of public values, policies, and current threats to surface waters, and an understanding of environmental stresses and their impact on the chemical and physical habitat, and hence on the biota. The endpoints, and thus the indicators representing them, must motivate change in policy when poor conditions are found. Figure 4 provides a general overview of EMAP-Surface Waters indicators, relating the selected

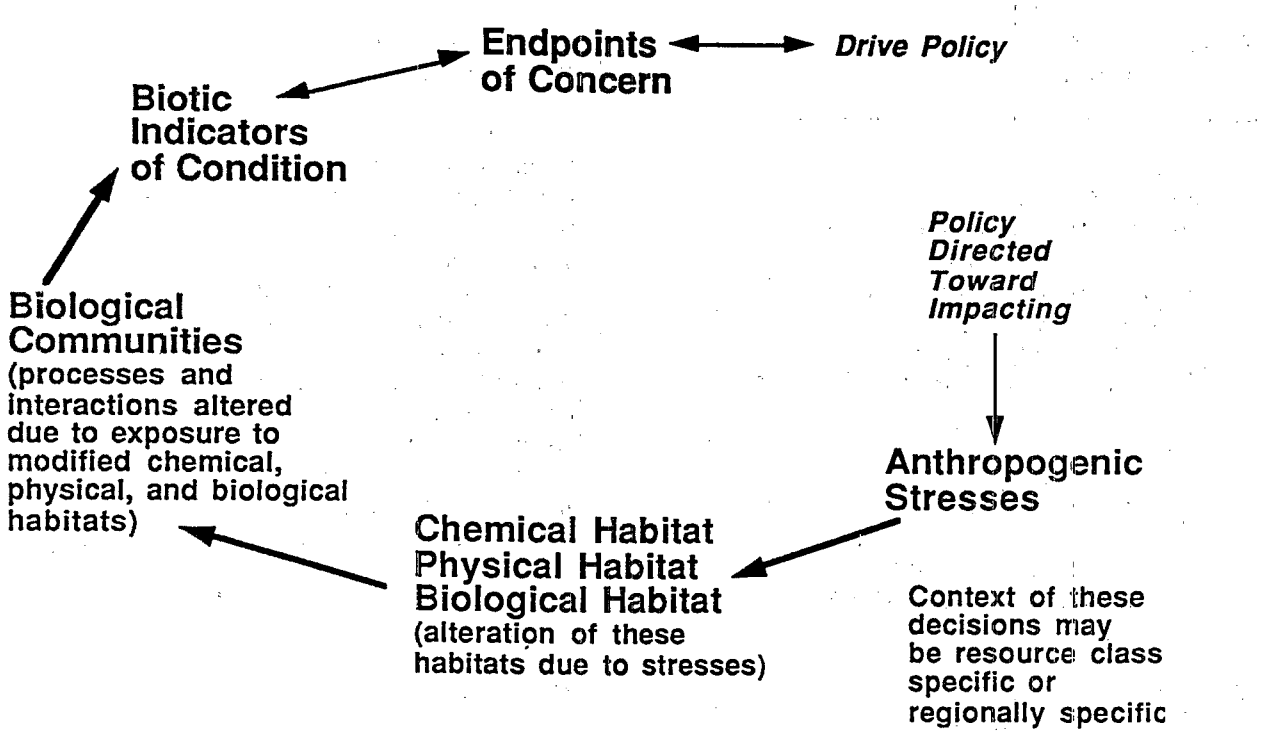


Figure 3. General approach for identifying indicators. It requires (1) consideration of the public values or endpoints (which drive policy) and the potential stresses (at which policy is directed), and (2) knowledge of the interactions between chemical, physical, and biological processes.

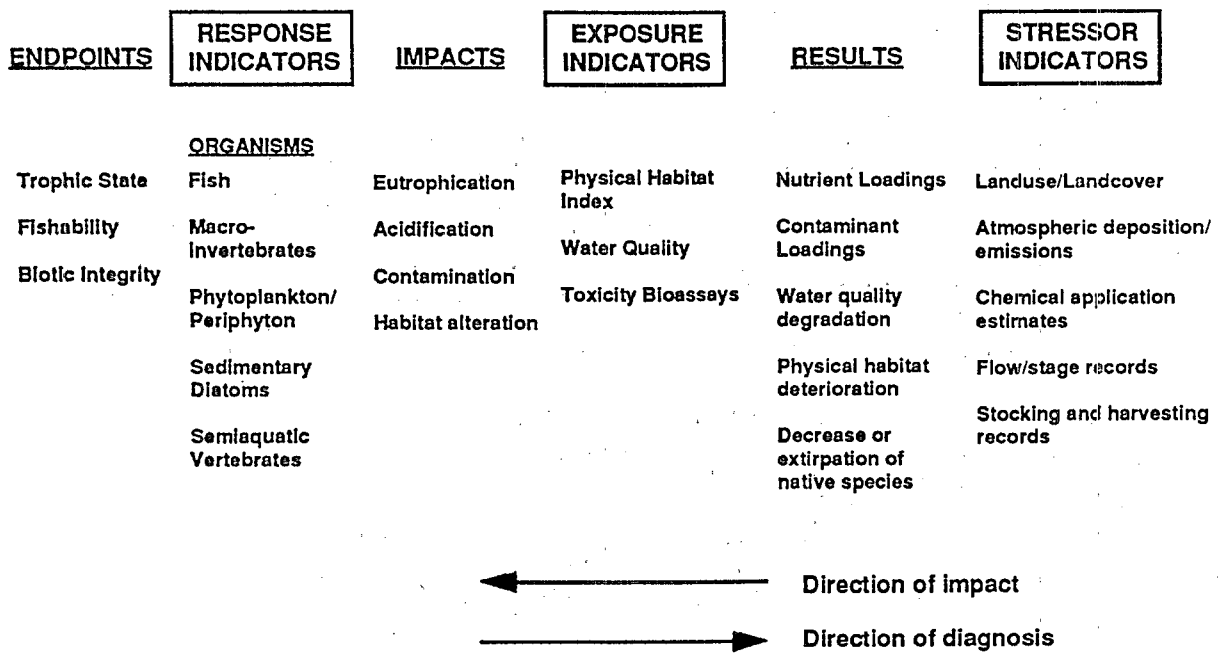


Figure 4. Indicator approach for EMAP-Surface Waters showing candidate indicators and the top-down approach to problem identification and diagnosis of probable cause.

Table 1. Candidate Indicators for Inland Surface Waters

Response
Trophic state index
Sedimentary diatom assemblage
Zooplankton assemblage
Macroinvertebrate assemblage
Fish assemblage
Gross external anomalies
Semiaquatic wildlife assemblage (birds)
Exposure/Habitat
Physical habitat quality
Chemical habitat quality
Sediment toxicity
Chemical contaminants in fish
Chemical contaminants in sediments
Biomarkers
Stressor
Land use and land cover
Human and livestock population density
Atmospheric emission and deposition
Use of chemicals
Pollutant loadings
Flow and channel modification
Lake level and shoreline modification
Introduced species
Stocking and harvesting records

endpoints, response indicators, exposure indicators, and stressors. Table 1 lists indicators being considered for use in EMAP-Surface Waters.

ASSESSING SURFACE WATER CONDITION

Surface Water Populations

To answer questions about the condition of lakes or streams, it is necessary to specify explicitly what constitutes the group under consideration. Target populations can be regionally defined, such as lakes in the Northeast, or defined on the basis of certain attributes, such as lakes larger than 20 ha. One of the general goals of EMAP is to include all ecosystems; therefore at the broadest level, the target population for EMAP-Surface Waters is all lakes and streams in the United States, excluding the Great Lakes, for which a separate EMAP component has been developed.

Within EMAP, subpopulations are the classes of resource types about which we wish to make statements of condition and trends. Tier 1 resources are one kind of subpopulation; they function as strata in the EMAP design. Other kinds of subpopulations are groups of lakes or streams that are not strata, but classes of interest. They may be classes within a particular Tier 1 resource, or classes that include more than one Tier 1 resource. Subpopulations may, but need not be, hierarchical. There is great flexibility to establish a variety of subpopulations on which to focus the estimation of condition and trends. Subpopulations serve two major purposes: (1) to increase the precision about estimates of condition and trends and (2) to target groups on which to report.

Establishing Unimpaired (Nominal) Condition

The assessment of ecosystem condition (or health) requires both (1) the determination of certain criteria considered indicative of a healthy sustainable resource and (2) the absence of known stressors and detectable symptoms of ecosystem stress. The challenge for EMAP-Surface Waters is to conduct such an assessment using the types of information and measurements that can be collected within the constraints of the EMAP design.

EMAP has adopted the terms nominal and subnominal to refer to healthy and unhealthy conditions, respectively. Classification of a lake or stream as nominal or subnominal will rely not on any single indicator, but on the full set of monitored response, exposure, habitat, and stressor indicators. Specific approaches for dealing with apparent inconsistencies in indicator signals, or for formally combining indicators into a joint index of surface water condition, will be explored as part of the EMAP-Surface Waters indicator development process.

The process of selecting and implementing the four types of indicators cannot be separated from the problem of how the data collected and the resultant indicators will be used to make statements about lake and stream conditions. The indicators will show that some of the waterbodies are impaired (subnominal, unhealthy). The task is to establish the criteria that define impaired

systems, that is, to determine that below a certain index or metric score, a lake or stream is considered to be in unacceptable condition, relative to a particular endpoint of concern. Some of the approaches for establishing these criteria include selecting and assessing reference sites, using historical data and/or pristine sites, using ecological models, and basing the criteria on the empirical distribution of indicator values.

EMAP-Surface Waters will interpret ecological condition based on a combination of these approaches, also relying heavily on a series of regional reference sites and knowledge of historic conditions as described through paleolimnological reconstructions, expert opinions, and previous biological data collections.

POPULATION ESTIMATION AND ANALYTICAL APPROACHES

Indicator measurements taken in the field and in the laboratory will be used to index the status and characteristics of each sample waterbody. EMAP sample lakes and streams will be chosen as a probability sample from explicitly defined populations of surface waters, thus information from these sample sites can be used to infer characteristics of the population from which the sites were drawn. Each sample waterbody will have a weighting factor inversely proportional to its sample inclusion probability. For example, if there were a 1 in 10 chance of selecting a particular lake for field sampling, that lake would represent 10 lakes in a statistical sense. These sample site weightings will be used to extrapolate from index conditions in the set of sample sites to index conditions in the population statistically represented by those sites.

Tier 1 Resource Estimates

There are several kinds of Tier 1 estimates, depending on the data on which they are based and the nature of the populations and the selection process. For example, resource inventories are provided by the landscape descriptions and other data on the 40-km^2 characterization hexagons established on the grid points. These Tier 1 estimates will typically be of the areal extent of surveyed resources, by whatever classes have been identified during the Tier 1 classification.

Estimates of total surface area of lakes could be provided by size class and geographic region, as could estimates of the numbers of lakes. Estimating the numbers of any class of lakes or streams is possible, since waterbodies can be unambiguously represented by points.

Tier 2 Resource Estimates

The Tier 2 sample consists of a subset of surface waters on which the suite of indicators will be measured. Measurements of additional variables may also be obtained from maps, remote imagery, and other sources for use in higher resolution classification of surface waters. Tier 2 estimates will be of the areal extent of classes defined by the Tier 2 information, and of the numbers of resource units in those populations. The Tier 2 sample will lend itself to more complete population descriptions in terms of the indicator variables that have been measured on the samples of population units.

Analytical Approaches for Detecting Population Differences, Changes, and Trends

An important feature of EMAP-Surface Waters will be the description of differences among subpopulations and changes in subpopulations over time. For example, it may be important to illustrate or describe differences among lakes influenced by different land use patterns or practices, or to determine whether the condition of different types of lakes differs (e.g., drainage versus seepage lakes, or lakes versus reservoirs). During the early years, it will be important to determine whether there are detectable year-to-year differences within subpopulations of lakes, especially if changes in the magnitude of stressors are occurring or are anticipated. As more data are obtained, it will become feasible to detect trends in a variety of subpopulations.

An assortment of techniques will be used in the examination of EMAP data for change and trend. An extensive range of standard statistical techniques is applicable. Linear model techniques, such as analysis of variance and regression will be used with various assumptions regarding spatial and temporal variance components, statistical independence, explanatory variables, nature of trend, and nature of change. Nonparametric alternatives, based on ranks or signs of differences, will also be used.

FIELD SAMPLING DESIGN

The basic design is to sample all Tier 2 sites once in the late summer period. Water and sediment samples from lakes and reservoirs will be collected in a mid-lake location approximately at the deepest point of the lake. Biological samples will be collected from representative major habitat types within lakes. Physical habitat indices will be used to describe the lake and its riparian zone as a whole. In streams and rivers, the sampling location will be a reach, 30 channel-widths long, located midway between the mapped upstream and downstream confluences defining the stream segment that is the sample unit. Water samples will be taken at a single mid-channel location in the middle of this reach, and biological samples will be systematically sampled in representative habitat types within the reach. Physical habitat characteristics will be assessed over the entire 30-channel-width section and its riparian zone.

INDEX CONCEPT

There are both temporal and spatial aspects of index measurements. The optimal field sampling design depends upon the objectives of monitoring, as well as on the variability of the indicators to be measured. The overall data quality objectives of EMAP require that indicator measurements be relatively stable, but responsive to year-to-year changes in biotic stress and exposure.

In the temporal context, an ideal index period is a time during which the values of response and exposure indicators are relatively stable. It must be a time period when indicator biota are present and measurable. To facilitate diagnosis of probable causes of subnominal indicator values, it is advantageous for the index time period to be a season in which indicator organisms are exposed to maximum environmental stress.

In a spatial context, an index value should provide in a single measurement, or an integration of a number of measurements, information adequate to represent each sampling unit in classification, regional extrapolations, or trend analysis. If the sampling unit is heterogeneous, the index measurement should, through multiple sampling points, sufficiently incorporate the spatial variability of indicator measurements within the sampling unit (lake or stream). Alternatively, the measurements might be made at locations which themselves integrate the ecological stresses impinging on the sampling unit as a whole. For example, pelagic fish and mid-lake water quality might reflect benthic biogeochemical processes, chemical inputs from streams, and littoral spawning, rearing, and feeding areas.

VARIABILITY

Temporal Variability

Temporal variability in the surface water characteristics to be measured by EMAP indicators occurs over a range of time scales. Variability over decades, years, and seasons, and within seasons, may affect the precision and interpretation of status and trends in the condition of lakes and streams. Some sources of variability are cyclical, and might be considered "noise"; others may be trends that EMAP should quantify.

Rivers and streams are generally more persistent than lakes on a geological time scale and their long-term characteristics are probably more stable. However, flowing waters have much smaller hydrologic retention times than lakes, so the amount and character of water in their channels more closely reflects fluctuations in the sources and sinks for water, energy, and chemical substances. As a result, flowing waters often exhibit greater seasonal, within-season, and diel variation than do most lakes, for those characteristics that depend on the amount and flowpath of hydrologic inputs.

Spatial Variability

Spatial variability in lakes and reservoirs occurs in both vertical and horizontal dimensions. Vertical variability in temperate lakes occurs primarily because of temperature stratification. Horizontal variation in lake characteristics occurs because of a wide variety of factors, including lake size, morphometry, and substrate, and the location and magnitude of inlets and outlets. For streams, stream discharge, gradient, channel width, water depth, substrate, and many other physical characteristics change in a predictable fashion along the course of a stream from its headwaters to its mouth. Similarly, geology, soils, vegetation, and the influence of groundwater on streamflows change markedly from stream headwaters to the mouths of rivers.

The basic design, sampling surface waters once in the late summer season, should be adequate for meeting EMAP-Surface Waters goals in most regions, if some systematic spatial replication within sample sites is employed. However, the precision of EMAP indicator measurements will be evaluated through temporal replication on subsets of 20 to 30 samples lakes and streams in each region.

LOGISTICS APPROACH

Implementing the EMAP-Surface Waters program will require comprehensive logistics planning. A logistics plan must be developed before the start of field activities to assure that program goals will be met. Regional logistics plans will be updated annually. The logistics plan will address 15 elements:

1. Overview of logistical activities
2. Staffing
3. Communications
4. Sampling schedule
5. Site access
6. Reconnaissance
7. Waste disposal plan
8. Safety plan
9. Procurement and inventory control
10. Training
11. Field operations
12. Laboratory operations
13. Data management activities
14. Quality assurance
15. Logistics review/recommendations

Field activities will be start in 1991 with a lake pilot program in the Northeast. Additional regions will be phased into the program in each of the following years, as funding permits. The TIME project will also begin in the Northeast in FY91.

ORGANIZATIONAL STRUCTURE

Coordinating the logistics activities of 50 or more surface water teams across the nation will be very difficult. Regionalizing these logistics activities into various centers will be the most effective mechanism for conducting EMAP field operations. The EPA regional offices can provide the initial structure for these logistics centers. The EPA regions have firsthand knowledge of the environmental conditions within their respective regions and will have a major role within EMAP, part of which could be in logistics. The EPA regions represent the Agency's primary contact with the states, and can work with the states at the program level. Securing cooperation from the states for EMAP is essential because of requirements regarding collection permits and access permission. The EPA regions and states also have highly experienced field personnel, and their participation in EMAP field activities would be extremely beneficial. As key personnel directing field team activities year after year, they will provide the program with critical continuity.

The long-term success of EMAP depends on the development of an interagency program with common goals for the monitoring of the ecological condition of the environment. Surface water monitoring alone could involve numerous agencies within the Department of the Interior, the Department of Agriculture, and the Department of Defense. As EMAP evolves into an interagency program, agreements between agencies will have to be established to define responsibilities.

QUALITY ASSURANCE

Quality assurance must be an integral part of any program that intends to produce useful information. Consistent with this precept, and with EPA's policy of ensuring that all environmental data are of known and documented quality, EMAP-Surface Waters will include a comprehensive quality assurance/quality control (QA/QC) program. Major elements of the program will include:

1. Developing and documenting standard operating procedures (e.g., methods manuals).

2. Conducting staff training.

3. Maintaining suitable facilities and equipment.

4. Using QC samples to validate analytical data and the methods used to collect data.

5. Conducting external audits.

6. Performing extensive data verification and validation checks on the database management system.

The development of data quality objectives will provide the framework for balancing the tradeoffs between the quality of data needed to make sound decisions and EMAP constraints and costs.

INFORMATION MANAGEMENT AND REPORTING

To be of maximum use, data must be transformed into useful information as quickly as possible. Therefore the goal of EMAP-Surface Waters is to produce annual statistical summaries of the monitoring results for the preceding year within nine months after data collection has been completed. These reports will provide summaries of response, exposure, and habitat indicators for the regions sampled, but with minimal interpretation.

Interpretive reports will be published for the Congress, interested scientists, and decision makers after each four-year sampling cycle is completed. Special scientific reports and peer-reviewed papers also will be published periodically, to address particular topics of interest.

A computerized information management system (IMS) will be established to ensure that EMAP-Surface Waters data are made available in a timely manner to other EMAP resource groups, to other local, state, and federal organizations, and to academic institutions. The system must be capable of drawing from historical databases already in place for similar environmental monitoring efforts, and it must be compatible across the IMSs for all EMAP resource groups.

RESEARCH NEEDS

The EMAP-Surface Waters program must be accompanied by an aggressive research program. The range of researchable issues is quite long. Key issues can be divided into four categories:

1. Indicators of condition
2. Diagnostic indicators
3. Approaches for defining nominal and subnominal
4. Statistical approaches

Extensive work is required to refine our selection of indicators of condition. Existing indicators need regional refinement and new indicators are required to fill in gaps. Additional work needs to be done to improve our diagnostic capabilities with biological indicators.

Although the decision on what is an acceptable condition for a water body is ultimately a societal one, EMAP-Surface Waters is obligated to provide creative approaches to addressing this issue. The process of gathering scientists, policy analysts, managers, and decision makers together to tackle this problem will offer many challenges.

Many statistical challenges exist in addressing national assessments such as those proposed. The process of combining multiple indicators into a single statement of condition will require an extensive statistical effort. The task of doing trend analysis on population distributions is not simple.

And finally, the effectiveness of regulatory activities must be assessed. Through the EMAP-Surface Waters program, we hope to evaluate changes and trends in response indicators and then associate them with changes or trends in exposure or stressor information. This area must be addressed soon.

SUMMARY

EMAP is not intended to describe all components of an ecosystem or resource type, explain how systems function, address specific stressors to lakes and streams (e.g., acidification; chemical contamination), or replace compliance monitoring. The program will, however, provide information about the condition of specific indicators at a specific index period as a "snapshot" of the overall condition of a system. EMAP-Surface Waters will provide an ongoing monitoring framework within which the condition of lakes and streams can be assessed on a broad scale, so that the magnitude and geographical location of various problems can be assessed and mitigation and research priorities can be set more objectively. It will be possible to add new variables or make regional modifications so that new issues can be more quickly evaluated.

The task at hand is to start a scientifically sound, policy-relevant program that is solid enough to supply useful information, yet flexible enough to incorporate changes as our understanding grows and the program develops. The indicators proposed thus far are a beginning.

"I am always doing what I cannot yet do in order that I might learn how to do it."

Vincent van Gogh

1.0 INTRODUCTION

Vincent van Gogh was certainly not noted for his ecological or statistical contributions. Yet he is considered by many to have been a genius in his field. His thoughts on improving his own skills are insightful. This quotation from van Gogh captures an interesting perspective: often, the only way to really learn how to do something is to start. The situation is certainly far from that extreme with respect to ecological monitoring, but it is a useful perspective to consider.

As you read this research plan, you will find that we cannot currently reach all the objectives we have set. We are painfully aware of this. But we can implement aspects of the program, providing policy-relevant information and filling portions of the existing information gaps. As we implement these portions, we will learn and the program will improve. By implementing aspects of the program, we will more clearly define the key features that are missing and that should be given high priority on a research agenda. If the program does not stimulate or is not accompanied by an aggressive research agenda, its full potential will not be realized.

Over the next few years, we will be building a database from which refinements to the program can grow. The task at hand is to start a scientifically sound, policy-relevant program that is solid enough to supply useful information, yet flexible enough to incorporate changes as the program and our understanding develop.

In this document, we describe our vision of what is needed to evaluate the ecological condition of our nation's surface waters. We ask your assistance in evaluating whether the vision is sound, the approach rational, and the strategy adequate to provide for continual improvement in the program.

1.1 CONTENT AND ORGANIZATION OF THE RESEARCH PLAN

This plan is organized into 12 sections. Section 1 is a general introduction to the Environmental Monitoring and Assessment Program (EMAP) and the Surface Waters component, and their focus and objectives, followed by a summary of the approach for meeting these objectives. Each successive section provides additional levels of detail about specific aspects of the approach such as design, indicators, or information management. Sections of the Research Plan are:

- Section 2, Approach and Rationale: a brief overview of all aspects of EMAP–Surface Waters (EMAP-SW).
- Section 3, Monitoring Network Design: a detailed description of the sampling design, and site selection.
- Section 4, Indicator Development Strategy: a list of the selection of indicators, their intended use, and the process of continual improvement in achieving descriptions of ecological condition for surface waters.
- Section 5, Estimation and Analysis: an overview of the procedures and issues related to data analysis for status, trends, and associations between indicators.
- Section 6, Field Sampling Design: a discussion of index periods and issues pertinent to sampling the lakes and stream reaches within this framework.
- Section 7, Logistics: an outline of the issues related to conducting a field program of this type and issues that require resolution.
- Section 8, Quality Assurance: a summary of the approach and procedures that will be implemented to ensure that the quality of the data collected is high enough to meet program objectives.
- Section 9, Information Management: an explanation of data management issues and a description of the procedures for ensuring that users will be provided with collected data quickly and efficiently.

- Section 10, Coordination: an explanation of the approach for integrating information within EMAP-Surface Waters, with other components of the program, and with programs outside EMAP.

- Section 11, Expected Outputs and Timelines: a description of the reports and summaries expected from EMAP-Surface Waters and an overview of proposed timelines for implementing EMAP-Surface Waters.

- Section 12, Fiscal Year 1991 Field and Analysis Activities: an outline of the specific objectives for the field activities and data analysis activities planned for the coming year.

1.2 AN OVERVIEW OF EMAP

The need to establish baseline environmental conditions against which future changes can be documented with confidence has grown more acute with the increasing complexity, scale, and social importance of environmental problems such as global atmospheric change, acidic deposition, sustainability, and the loss of biological diversity. It is therefore critical for us to develop and implement monitoring programs that can permit quantitative, scientific assessments of the complex effects of pollutants and other stressors on ecosystems.

In 1988, the U.S. Environmental Protection Agency's (EPA) Science Advisory Board recommended implementing a program within the EPA to monitor ecological status and trends and to develop innovative methods for anticipating emerging environmental problems before they reach crisis proportions. More recently, the EPA has been urged to establish a program to assess whether the nation's efforts to protect the environment are producing the expected results in maintaining and improving environmental quality. In response to the need for better assessments of the condition of the nation's ecological resources, the EPA's Office of Research and Development began planning the Environmental Monitoring and Assessment Program (EMAP).

EMAP is a strategic approach designed to meet the growing need to identify and quantify the extent, magnitude, and location of degradation or improvement in environmental condition. When fully implemented, EMAP will answer critical questions for policy- and decision-makers and the public, such as: What is the current extent of our ecological resources? For example, how many hectares are there of estuaries, lakes, forests, deserts, wetlands, and grasslands, and how many kilometers of streams? How are these resources distributed geographically, and is this changing? What percentage of the resources appears to be adversely affected by pollutants or other human-induced environmental stress, and in which regions are the problems most severe or widespread? Which resources are degrading, where, and at what rates? Are adversely affected ecosystems improving in response to control and mitigation programs?

In order to answer these questions, an integrated monitoring network with the following objectives will be implemented within EMAP:

- Estimate current status, extent, changes, and trends in indicators of the condition of the nation's ecological resources on a regional basis with known confidence.

- Monitor indicators of pollutant exposure and habitat condition, and seek associations between human-induced stresses and ecological condition that identify possible causes of adverse effects.

- Publish annual statistical summaries and periodic interpretive reports on status and trends to the EPA Administrator and the public.

Assessments of whether the condition of the nation's ecological resources is improving or degrading require that data be collected over large geographic and temporal scales. For this reason, comparability of data among geographic regions and over extended time periods is critical. Meeting this need simply by aggregating data from many individual, local, and short-term networks has proven ineffective. There is also a need for extrapolating the findings from the sites sampled to the populations of lake and stream resources as a whole. EMAP networks therefore will be statistically designed to allow extrapolation to the entire population, to provide unbiased estimates of status and trends in indicators of ecological condition, with known uncertainty, and to determine associations among indicators of condition, exposure, and stress, all with quantifiable

confidence limits. This will be accomplished to focus on national and regional scales over periods of years to decades. These characteristics distinguish EMAP from other current monitoring efforts.

EMAP is being designed around six primary activities:

- Strategic evaluation, testing, and development of indicators of ecological condition, pollutant exposure, and habitat condition, and protocols for collecting data on these indicators.
- Design and evaluation of a comprehensive and versatile integrated monitoring framework.
- Nationwide characterization of the extent and location of ecological resources.
- Demonstration studies and implementation of integrated sampling designs.
- Development of data handling and quality assurance, as well as spatial analysis and statistical procedures, for efficient analysis and reporting on status and trends data.
- Assessments of the probable causes of environmental conditions and trends.

Seven ecological resource groups form the basic structure of EMAP: agroecosystems, arid lands, forests, Great Lakes, near-coastal systems, inland surface waters, and wetlands. Of course, ecological resources could be split in a variety of ways, but our goal is to ensure that the groups encompass all ecological resources. Six activities cross the entire program and play a key coordination role in its success: monitoring systems design, development and evaluation of indicators of ecological condition, quality assurance, logistics, information management, and integration and assessment. Two components function both as ecological resources and as cross-cutting coordination activities: air/deposition and landscape characterization. This document describes the development of the Surface Waters component of EMAP.

1.3 LEGISLATIVE MANDATE FOR SURFACE WATER PROTECTION

Although 75% of the earth's surface is covered by water, only 0.02% of the volume is contained in rivers and lakes (Ehrlich et al. 1977). In its 1990-91 report, the World Resources Institute (1990) identified the quantity and quality of fresh water, for human and other biological uses, as constituting a growing concern worldwide. In the conterminous United States, inland surface waters comprise less than 2% of the area (Geraghty et al. 1973), a small proportion that belies their importance for our society. For example:

- Inland surface waters provide 79% of the water used daily in the United States (Ehrlich et al. 1977).
- Water use in the 7 southwestern states exceeds runoff 9 out of 10 years (Ehrlich et al. 1977).
- Swimming and fishing rank first and second, respectively, among all outdoor participatory sports (USDI 1989).
- 31.5 million people fished inland waters in 1988; they spent $329.8 million (USDI 1989).

Understanding the importance of surface waters and expressing concern about their extensive degradation, Congress in 1972 established the protection of surface waters as a priority with the passage of the Federal Water Pollution Control Act (P.L. 92-500). The primary objective of P.L. 92-500 was to restore and maintain the physical, chemical, and biological integrity of the nation's waters [Section 101(a)]. An interim goal was to provide for the protection and propagation of fish, shellfish, and wildlife and for recreation in and on the water [Section 101(a)(1)]. Section 105(d)(3) requires the EPA to conduct, on a priority basis, an accelerated effort to develop and apply improved methods of measuring the effects of pollutants on the chemical, physical, and biological integrity of water. Section 304(a)(1) states that the EPA shall develop and publish criteria on the effects of pollutants on biological community diversity, productivity, stability, and eutrophication. Section 305(b) mandates biennial reports that assess the extent to which all waters provide for the protection and propagation of a balanced community of aquatic life.

Since 1972, the Act has been further strengthened. The Water Quality Standards Regulation (U.S. EPA 1983) requires that states designate aquatic life uses consistent with the goals of the Act, establish criteria for standards to protect those uses, and develop an antidegradation policy protective of waters exceeding the criteria. The Water Quality Act of 1987 amends P.L. 92-500 and emphasizes ambient standards and assessments as the driving forces behind further pollution abatement. Section 303(c)(2)(B) allows states to adopt criteria based on biomonitoring (a key EMAP component). Section 304(l)(A) requires listing of waters not expected to attain protection and propagation of balanced biological communities. Section 304(m)(2)(g) requires the EPA to study the effectiveness of applying best available pollution controls for protecting balanced communities. Section 314(a) requires trophic classification of all publicly owned lakes and an assessment of the status and trends of water quality in those lakes. Finally, Section 319 mandates identification of waters that cannot protect balanced aquatic communities without nonpoint source pollution controls.

The Federal Water Pollution Control Act and subsequent amendments in 1977, 1981, and 1987 [collectively called the Clean Water Act (CWA)] not only mandated the protection of surface waters but established, by law, the requirement for several monitoring activities. Monitoring for permit limit compliance is established under section 308. The CWA also provides for short- and long-term monitoring under Title I–Research and related Programs, section 106(e)(1). The EPA is authorized under this section to fund states to monitor, compile, and analyze data on the quality of navigable waters, including biological monitoring and classification according to trophic status. These data are to be annually updated and included in the biennial reports required under section 305(b) of the CWA.

The amendments to the Clean Air Act, passed in the 1990 session of Congress, require that the effects of acidic deposition on surface waters be monitored, in order to assess surface water response to decreases in emissions. Because this mandate is more specific than the general EPA mandate to monitor the condition of surface waters, and because the amendments require an assessment in a relatively short period of time (by January 1, 2003), EMAP will include a component designed specifically to address the question of aquatic response to acidic deposition (the TIME project).

Other legislation requires assessment of environmental risk to aquatic communities as well. Of particular importance are stressors with potential regional impacts that fall under the Federal Lands Policy and Management Act (FLPMA), the National Environmental Policy Act (NEPA), the Resource Conservation and Recovery Act (RCRA), and the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA). In addition, the Endangered Species Act mandates assessments and protection of rare and threatened species and the National Forest Management Act of 1976 requires conservation of animal diversity.

1.4 CURRENT EFFORTS TO MONITOR SURFACE WATERS

Given these legal mandates to protect and restore surface water quality, what efforts currently exist to document the conditions of surface waters? A brief description follows of four current programs that document some aspect of surface water condition: monitoring by the States in support of Section 305(b) requirements (conducted by States and reported through the EPA), the Fish and Wildlife Service National Contaminant Biomonitoring Program (NCBP), the U.S. Geological Survey's National Stream Quality Accounting Network (NASQAN), and the U.S. Geological Survey's Hydrologic Benchmark Network. This section is not intended to be a complete description of these programs or a complete listing of every monitoring program in existence, but rather a brief discussion of four national programs of relevance to EMAP. A new stream monitoring program, the U.S. Geological Survey's National Water Quality Assessment Program (NAWQA), is described in Section 10 as a program with which EMAP needs to interact closely.

1.4.1 State 305(b) Water Quality Monitoring

As described in Section 1.3, the States are currently required to monitor the quality of surface waters in support of section 305(b) of the Clean Water Act and report these findings to the EPA. For its part, the EPA attempts to establish consistency among these efforts by preparing guidelines for the States in reporting their information. Use of consistent measures and approach is promoted and based on the degree

to which a waterbody is in compliance with the State water quality standards established for that waterbody. Water quality standards within the States consist of the objective for water quality expressed as "beneficial use," plus numeric and narrative criteria designed to ensure maintenance of the beneficial use. The degree to which beneficial use is supported is ranked by four categories: fully supporting, fully supporting but threatened, partially supporting, and not supporting. The States have considerable discretion in determining exactly what variables are measured, which methods are used, how many and which waterbodies are assessed, and how decisions about the degree of use support are made. The biennial report submitted by the EPA to Congress (U.S. EPA 1990d) uses the State reports to summarize the total river miles or lake acres in each of the four support categories, the major causes of use impairment, the source of pollution in those waters not fully supporting their uses, and the number of waters affected by toxic pollutants.

1.4.2 National Contaminant Biomonitoring Program

The National Contaminant Biomonitoring Program (NCBP) is maintained by the U.S. Fish and Wildlife Service to document temporal and geographic trends in concentrations of environmental contaminants that may threaten fish and wildlife resources. Originated as part of the multiagency National Pesticide Monitoring Program, the NCBP has been in operation since 1967. The fishery component of the program analyzes tissue from samples collected at over 112 stations at key points in major rivers throughout the United States and in the Great Lakes. Whole fish samples of 2 bottom-feeding species and 1 representative predatory species are analyzed for 22 organic chemical compounds and 7 metals (Schmitt et al. 1985, Lowe et al. 1985). The river stations were selected in conjunction with the USGS NASQAN stations. Collection of samples from all stations takes place during the fall of two successive years.

1.4.3 National Stream Quality Accounting Network

In 1971, the U.S. Geological Survey began the National Stream Quality Accounting Network (NASQAN) with the objectives of (1) accounting for the quantity and quality of water moving within and from the United States, (2) depicting area variability, (3) detecting changes in stream quality, and (4) laying the groundwork for future assessments of changes in stream quality. In 1986, these objectives were modified: (1) assess water quality trends nationally, and to the extent possible, relate those trends that are detected to upstream land and water use, (2) account for the mass transport of selected constituents from the continent, (3) account for the transport of selected constituents to estuaries and/or other significant water bodies (e.g., the Great Lakes) that may be considered critical or otherwise of national significance, and (4) account for the quality of water crossing international boundaries (in a small number of cases). Of the 412 sites on major U.S. rivers, 58% are monitored bimonthly and 42 quarterly. Variables measured are predominantly major anions and cations, selected metals, suspended sediments, dissolved oxygen, and fecal coliform and streptococcal bacteria. Periphyton, phytoplankton, organic carbon, and pesticide measurements were all discontinued between 1979 and 1981. Several descriptions of trends in water quality variables have resulted from this program (Smith et al. 1987).

1.4.4 Hydrologic Benchmark Network

This network, established in 1964, is intended to document natural changes in hydrologic characteristics, provide better understanding of the hydrologic structure of natural basins, and provide a comparative basis for studying the effects of humanity on the hydrologic environment. The 58 stations located across the country are minimally impacted by humans. More than 50 physical and chemical measurements are made at fixed sampling intervals. Samples are collected quarterly at 73% of the stations, bimonthly at 23% of the stations, and monthly at 4% of the stations.

1.5 UNANSWERED QUESTIONS FOR SURFACE WATERS

The mandate to protect and restore the condition of surface waters has been outlined in Sections 1.3 and 1.4, along with current efforts to monitor this progress. What then do we know about the condition of our lakes, streams, and rivers from a national perspective? In evaluating the NASQAN data for 1974-1981, Smith et al. (1987) determined that widespread decreases in fecal bacteria and lead concentrations had occurred, but that concentrations of nitrate, chloride, arsenic, and

cadmium had increased over broad regions. Their general conclusions were that point source controls had been effective, although nonpoint source contaminants were on the rise, and that more detailed work was necessary to evaluate the policy implications of the apparent changes.

Results from the NCBP (Schmitt et al. 1985, Lowe et al. 1985) show no significant trends nationally in the geometric means of metal concentrations in fish tissue. However, small but significant decreases in DDT were observed, PCB residues remained widespread but showed declines, concentrations of dieldrin did not change, and chlordane concentrations were slightly lower and less widespread. The impact of the reported changes on aquatic biota is not well documented. The EPA, in summarizing the 1986 State 305(b) reports, estimated that designated beneficial uses were supported in 74% of the river miles and 73% of the lake acres assessed (U.S. EPA 1987c). In summarizing the 1988 reports (U.S. EPA 1990e), the EPA estimated that 70% of the river miles and 74% of the lake acres assessed were reported to support designated uses. One might legitimately argue that a difference between 74% and 70% is unimportant. The issue, however, is that not only do we not know if the differences are actual and/or meaningful, but we currently have no mechanism for determining this.

In either case, these numbers present a not too gloomy picture of the condition of lakes and streams. They stand in stark contrast, however, to other available information:

- 20% of the native fishes of the western United States are reported to have become extinct or endangered (Miller 1989).
- 32% of the Colorado River basin native fishes are reportedly extinct, endangered, or threatened (Carlson and Muth 1989).
- Since 1910, annual Columbia River salmon and steelhead runs reportedly have declined by 75-85%, or 7-14 million fish (Ebel et al. 1989).
- The Missouri River commercial fish harvest is reported to have been reduced by over 80% since 1945 (Hesse et al. 1989).
- Since 1850, 67% of the fish species from the Illinois River and 44% from the Maumee River have declined or disappeared; 70% to 84% of these species were lost from small or medium-sized streams (Karr et al. 1985).
- Since 1933, the Tennessee River system has lost 20% of its mollusk species (Williams et al. 1989) and 46% of the remaining species are endangered or seriously depleted throughout their ranges (Jenkinson 1981).
- 38 states reported fish consumption closures, restrictions, or advisories in 1985 (Moyer 1986).

We possess a great deal of information about the water quality of our aquatic resources, but we have little consistent understanding of overall ecological conditions or changes in conditions. Our understanding of the reliability or level of uncertainty surrounding the information is minimal. There is some sense that many of the most severe problems have been relieved, but we may in fact have lost ground to the pressures of increased human impact and less obvious nonpoint sources of pollution or habitat alteration. Apparently, we lack a quantitative framework from which we can produce reliable estimates of the condition of lakes and streams in the United States, especially in the areas of nonpoint sources of pollution and habitat alteration. This concern is also reflected in the EPA's recent Science Advisory Board report (SAB 1990), which lists physical habitat alteration and biological diversity among the major ecological risks requiring the attention of the EPA. The National Research Council (1990) reached similar conclusions in their recent review of the USGS NAWQA Program (see Section 10.0): we genuinely need a long-term national assessment of quality on a broad spatial scale in the United States. Brown and Roughgarden (1990) also reflected this concern in their call for a National Institute of Ecology with a U.S. Ecological Survey as a prime component.

1.5.1 Societally Important Surface Water Values

To be effective, the information from a national monitoring program must prompt action when required. This means that the information produced must be related to perceptions of aquatic health and represent issues and values of concern and importance to the public, aquatic scientists, and decision makers. In ecological risk

assessment, these perceptions are called endpoints. Suter (1990) and Hunsaker et al. (1990) defined ecological endpoints as the environmental entity of concern and the descriptor or quality of the entity. Such endpoints represent concepts that are societally important but that tend to be nebulous or abstract and do not lend themselves to direct measurement.

Historically, for surface waters, these endpoints have been expressed as designated uses, which have included habitat for aquatic life, fishing, swimming, navigation, and drinking water supply. Some of these uses depend directly on ecological condition (e.g., habitat for aquatic life, fishing). Others are less directly dependent on ecological condition (e.g., drinking water, navigation), although attainment may be indirectly associated with the consequences of degraded ecological condition. Also, as currently expressed in the Clean Water Act, physical, chemical, and biological integrity embody a societal concern and desire that surface waters be unimpaired and healthy. In particular, increasing attention is drawn to the biological condition of ecosystems--or biological integrity--not explicitly defined by legislation but defined by Karr and Dudley (1981) as "a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitats in the region." This is similar to one current concept of biological diversity defined by Noss (1990) as the diversity of life and its processes which has functional, structural, and compositional attributes at the genetic, species, community, and ecoregion levels of organization. This concept is much broader than simply a rare and endangered species issue.

In the early stages of EMAP-Surface Waters development, our plan was to address the ecological condition of lakes and streams as a single endpoint of concern: biological integrity. After extensive consideration and discussion with others, it became apparent that not all of society's ecological values related to surface waters fall neatly under the biological integrity umbrella. In fact, many aquatic systems are managed in a manner supported by the public as an improvement in ecological condition, but which reduces biotic integrity. One of these practices is the management of lakes and streams to increase game fish, usually by replacing naturally occurring species and stocks with exotic species or hatchery stocks. Lakes are also fertilized to increase fish production. Thus, attaining the societal value of fishability often conflicts with the other societal values of biotic integrity and clear water.

There is also considerable public interest in the trophic condition (often expressed as clarity or productivity) of surface waters, and in particular, lakes, to the extent that assessing the trophic condition of lakes appears as part of the Clean Water Act, Section 314. The Clean Lakes Program evolved to support restoration of lakes undergoing cultural eutrophication, and the North American Lake Management Society (NALMS) formed to deal primarily with management of trophic condition. The desire for lakes with clear water is often in direct conflict with the desire for lakes that support a productive fishery or a diverse fish and wildlife fauna.

Given multiple societal concerns or values, EMAP-Surface Waters proposes to evaluate the condition of lakes and streams with respect to each of these issues:

- Biological integrity
- Trophic condition
- Fishability

There are a few issues to be considered relating to these three endpoints of concern. First, as already mentioned, it is likely that for individual waterbodies our assessments of these endpoints will conflict (e.g., a lake may have an excellent fishery but be in undesirable biotic and trophic condition). In this case, it is important to remember that EMAP will not focus on individual waterbodies but will make its assessments at regional and national scales. However, the regional proportion of lakes and streams with conflicting assessments will be of interest. We must also develop regional expectations for biotic integrity, trophic state, and fishability, because these vary among regions for natural reasons, such as ecoregional and zoogeographic differences. Second, some lakes (and streams) are naturally fishless or eutrophic. This means EMAP-Surface Waters will need (1) to make this determination for each waterbody, (2) to evaluate these waterbodies as separate subpopulations (see Section 3), and (3) to estimate the regional proportion of waterbodies naturally occurring in this condition. Third, fishability is a value-laden endpoint, which complicates its assessment (what is good fishing for one is not necessarily good

fishing for all). In any case, our primary focus will be on biotic integrity.

1.5.2 Hazards to Aquatic Systems

We have classified current hazards to inland surface waters into nine somewhat overlapping categories. Some of these reflect the effects of land use or pollutants, others are related directly to resource management:

1. Eutrophication (anthropogenically induced nutrient enrichment)
2. Chemical Contamination (toxics)
3. Atmospheric pollutants (e.g., acidic deposition)
4. Habitat alteration (physical structure, substrate, riparian vegetation)
5. Flow modification
6. Thermal alteration
7. Introduced species
8. Harvest imbalance (overstocking or overharvesting)
9. Global climate change

One objective of the program is to relate any degraded condition of aquatic resources to these various hazards or combination of hazards, and any improved condition of aquatic resources to reduction in these hazards. However, diagnosing problems to this level of resolution is a very long-term objective.

1.6 SPECIFIC OBJECTIVES OF EMAP-SURFACE WATERS

Objectives for EMAP-Surface Waters parallel the program objectives. More specifically they are to:

- Estimate the current extent (location, number, surface area or length) of lakes and streams, on regional and national scales, with known confidence.

- Estimate the current status, changes, and trends in indicators of the condition of the nation's lakes and streams, on regional and national scales, with known confidence.

- Monitor indicators of pollutant exposure and habitat condition within lakes and streams and seek associations between human-induced stressors and ecological condition that identify possible causes of adverse effects.

- Publish annual statistical summaries on the extent and the status of indicators of ecological condition of lakes and streams, and periodic interpretive reports on the status and trends indicators of ecological condition of lakes and streams to the EPA Administrator and the public.

1.7 FOCUS AND PURPOSE OF THE RESEARCH PLAN

This Research Plan presents the rationale, objectives, approach, and plan for establishing a monitoring program designed to assess the status and trends in ecological condition of the nation's lakes and streams. It also includes descriptions, where appropriate, of the EMAP component (the TIME project) that will specifically address the response of surface waters to acidic deposition. The current document is not intended to be a final plan for establishing these monitoring programs but rather part of a continually improving process. It is our immediate objective to inform potential EMAP data users of the approach proposed for describing and monitoring the condition of lakes and streams, and to elicit input and advice from the scientific, management, and regulatory communities that will improve the program. We recognize from the outset that no one agency or group has the expertise to develop a program of this magnitude independently. We fully acknowledge that we currently lack many of the tools needed to reach the full potential of EMAP. However, we firmly believe that we can begin the process and by so beginning we can make strides toward improvement. The proposed design, we believe, provides sufficient flexibility to allow modifications as we learn from pilot and demonstration programs. The indicators we have proposed are a beginning. They are not complete, but they will provide useful information on the condition of aquatic resources. An aggressive research program for developing improved indicators, as well as other tools, is urgently needed for monitoring in general, and we hope that EMAP can provide some of the stimulus for the development of such a program. In several sections of this research plan, we present the major issues and problems associated with that topic and current options for addressing the issue. Early emphasis has been on developing a lakes program, with a stream and river program to follow. In summary, we welcome all advice and input for this program.

2.0 APPROACH AND RATIONALE

Section 2 is an overview of the EMAP approach and highlights those aspects that make it unique. This section also includes an overview of the technical approach for meeting the goals and objectives previously outlined for EMAP-Surface Waters. The limitations of EMAP are also presented. What is needed to accomplish the EMAP-Surface Water objectives? What is the fundamental concept of EMAP? These subjects are introduced in the overview and explained in detail in the remaining sections of the plan.

2.1 OVERVIEW OF THE EMAP-SURFACE WATERS APPROACH

What sets the EMAP-Surface Waters approach apart from existing monitoring efforts? The fundamental distinctions occur in the areas of design and indicator strategy. This is primarily a top-down program being designed to evaluate health or condition of surface waters with respect to endpoints of concern and then to statistically extrapolate those findings to the ecological resource populations of interest. Our discussion of this difference is not intended to criticize existing programs but rather to acknowledge that they were designed for different objectives and as such cannot effectively be used to meet the broad regional and national scale objectives of EMAP. The strategy for EMAP employs the following features that allow estimation, with known confidence, of indicators of the ecological condition of regional surface water populations:

- Explicit definition of surface water target populations and their sampling units, and identification of an explicit frame for listing or otherwise identifying all the potential sampling units within each target population (Section 3).

- Probability sample site selection from the population frame. A uniform grid and clustered sampling will be used to obtain a randomized, systematic sample of surface waters with a geographical distribution reflecting that of the population (Section 3).

- Representation of ecological conditions in sample lakes and streams using ecological indicators within an index concept (Section 4.

- Uniform sampling and analytical methods for a suite of response, exposure, and stressor indicator measurements (Sections 6 and 8).

- A documented program of rigorous quality control, quality assurance, and quality assessment (Section 8).

These features, coupled with a sound understanding of the ecological function of lakes and streams, comprise an effective program capable of providing increasingly valuable information about the condition of, and our protection for, ecological resources.

EMAP-Surface Waters will estimate the condition of lakes, reservoirs, streams, and rivers on a national scale as well as on relatively broad regional scales. Data from the program will allow estimation of the spatial extent and geographic distribution of various classes of surface waters. Additionally, it will estimate the current status of and future changes or trends in indicators of ecological condition. The design and indicator strategies described in this section are intended to be adequate for exploring a variety of questions on national and broad regional scales rather than to answer specific questions that might be asked. In this context, EMAP-Surface Waters will provide information with which we can determine (1) what proportions of our lakes, streams, and rivers are degrading or improving, where, and at what rate, (2) what the likely causes are of adverse effects, and (3) whether or not systems are responding as expected to control and mitigation programs. Information derived from the program will help in setting national priorities. For example, in order to set research and mitigation priorities for streams and rivers on a national scale, we must be able to determine the relative importance or magnitude of toxic contaminants compared with habitat alteration problems as they affect stream condition. Given this focus, the size of the United States, and the number of types of ecological resources that should be addressed, EMAP-Surface Waters must confine its general activities to regions no smaller than the EPA regions or the aggregate ecoregions of Omernik (Figures 2-1 and 2-2) until cooperative efforts with other federal agencies and States permit a sampling density with greater spatial resolution.

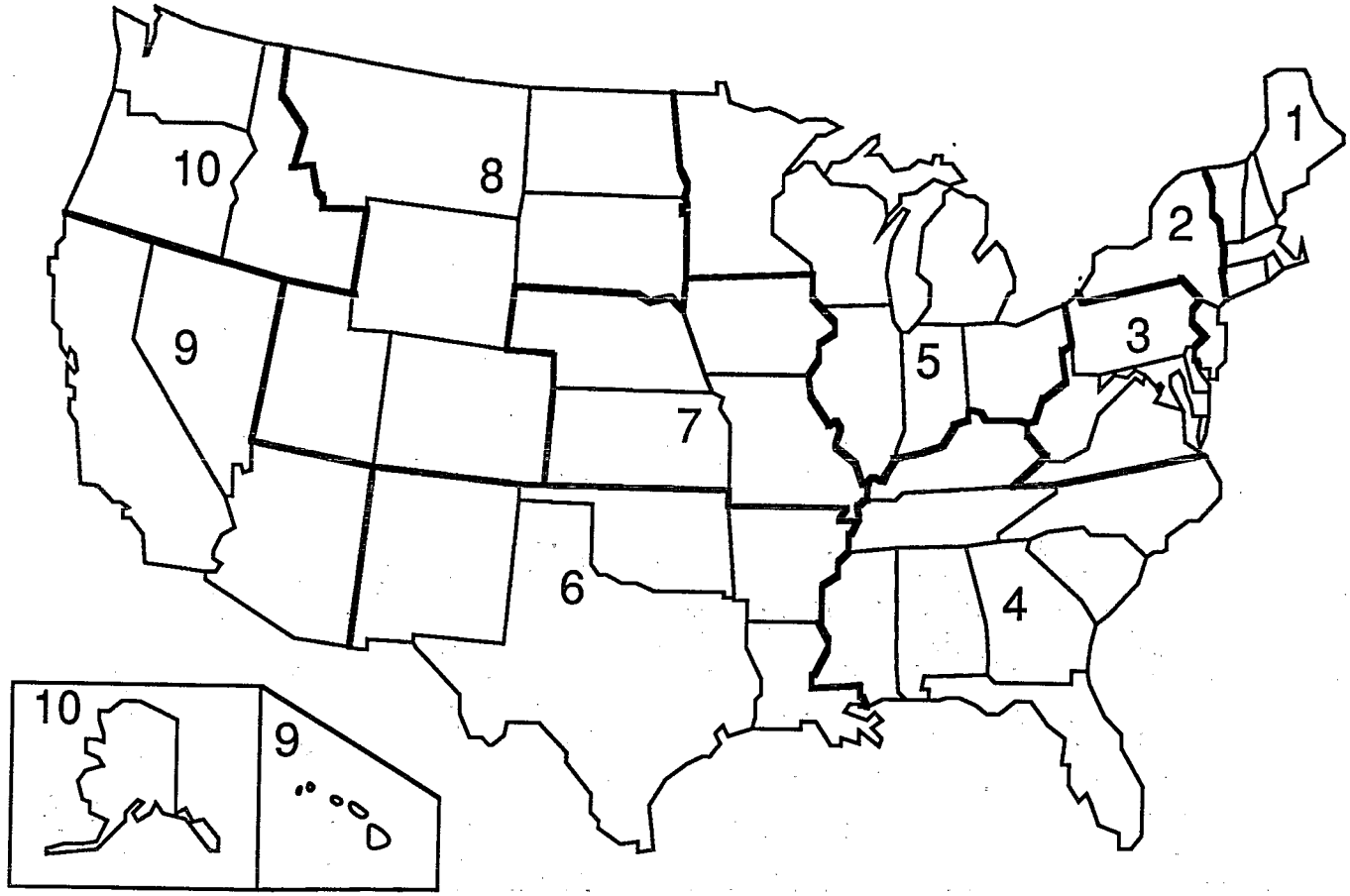


Figure 2-1. U.S. Environmental Protection Agency regions.

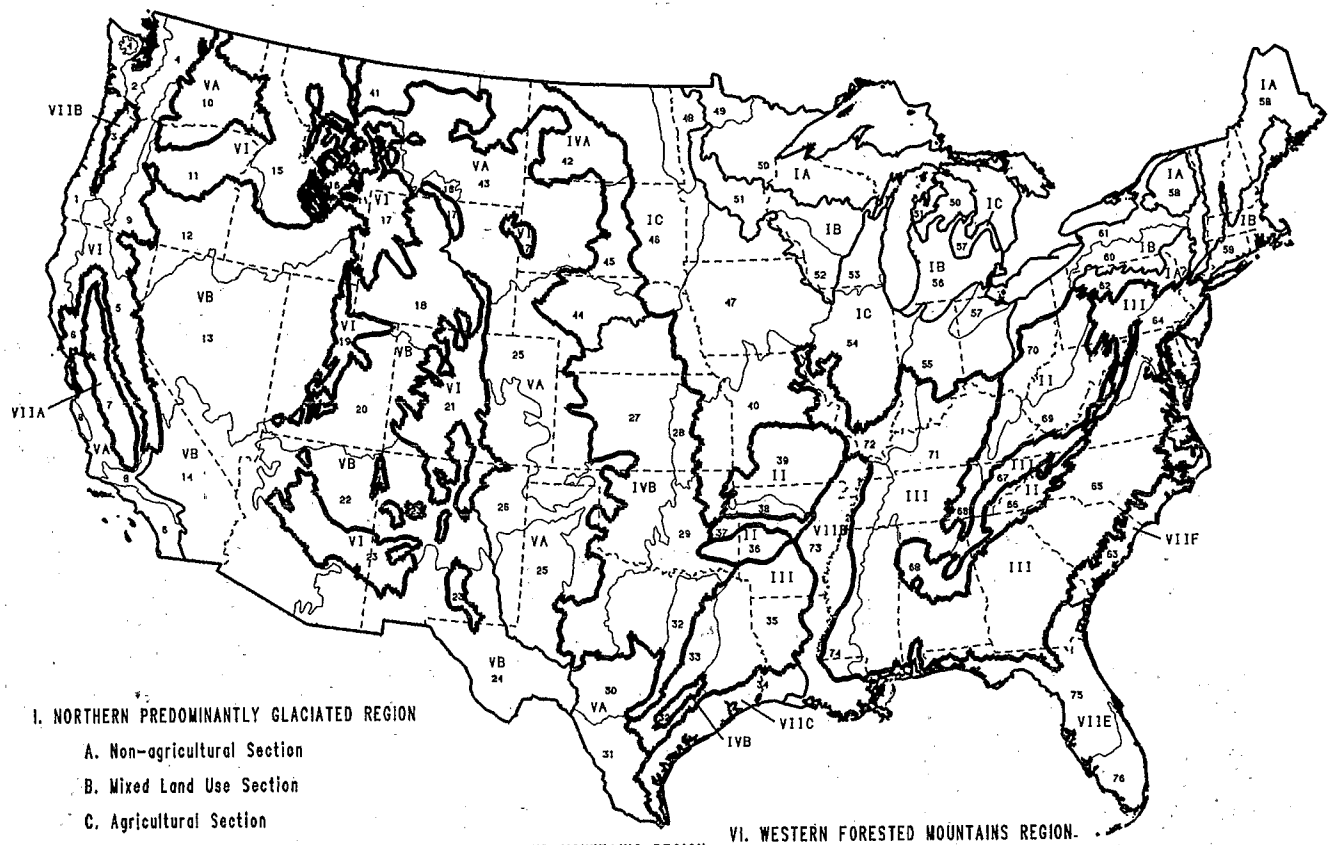


Figure 2-2. Aggregations of Omernik's ecoregions. These regions are the approximate size of regions for which resolution will be possible with base funding in EMAP-Surface Waters. These regions are an effective tool for analyzing the results of water quality and aquatic biological research.

A comprehension of the four-tiered approach to meeting EMAP-Surface Water objectives is essential for reading this document and understanding EMAP as a whole. Figure 2-3 illustrates the basic levels of the four-tiered approach. A tier is a level and type of activity related to monitoring and assessing ecological condition. Most of the discussion presented in this document focuses on Tiers 1 and 2, but two additional tiers are important for long-term planning.

Tier 1 of the four-tiered approach is the broadest level and the one in which estimates of the extent of a resource and its geographical distribution are generated. It involves the characterization of existing resources and their pattern of use on the landscape. For EMAP-Surface Waters, this means that during Tier 1 we will generate the number and surface area of lakes, as well as their geographical distribution. For streams, we will develop estimates of the number of stream reaches and total miles of stream. The initial source of this information will be the USGS 1:100,000-scale Digital Line Graphs (DLGs). The estimates will be periodically updated using new remote imagery (satellite imagery and aerial photography). In some instances, Tier 1 activities may also provide information about the condition of these resources, if the data can be derived from various remote sensors.

Tier 2 activities involve field sampling to determine the conditions of a subset of Tier 1 sites. A suite of biological, chemical, and physical measurements will be obtained from each of these sites and the information aggregated to make statements about the regional condition of that resource type rather than statements about the specific condition of any particular site.

Tier 1 and 2 activities will allow the successful completion of the primary objectives of EMAP and will be the priority in early funding. However, other information needs will require an additional, but complementary, level of effort at Tiers 3 and 4. The need for two types of information will drive

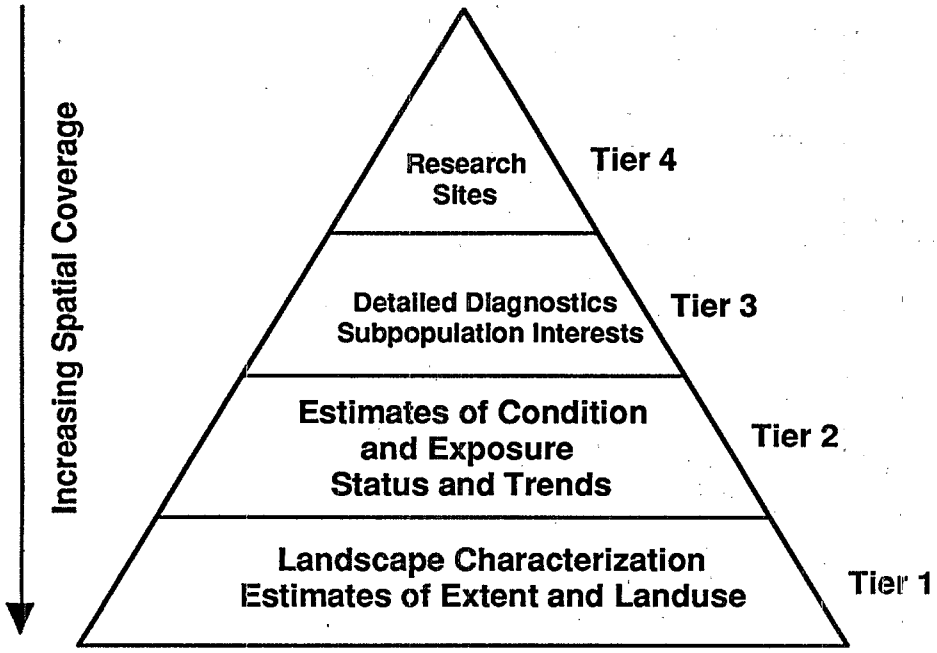


Figure 2-3. Concept of a four-tiered approach in EMAP. Spatial coverage is maximized in lower tiers; temporal coverage increases at the higher tiers.

Tier 3 activities within EMAP: more explicit regional diagnostic information, and information about status, trends, and diagnostics for more specific subpopulations of interest, such as playa lakes or saline lakes.

During Tier 1 and 2 activities, if a resource exhibits poor conditions throughout a region, but the likely cause cannot be determined from the available data, then more detailed work in that area could be implemented to identify the source of the problem. The problems resulting from selenium toxicity, most notably at Kesterson Wildlife Refuge but also widely observed throughout the West, provide an example of the type of situation that might warrant movement to Tier 3. If the EMAP-Surface Waters approach were applied to this example, Tier 2 activities might identify widespread deaths and deformities in waterfowl, but with unknown cause. The Kesterson Wildlife Refuge would be the focus of concern, but given the apparent extent of the problem throughout the western United States, a more detailed study of regions might be undertaken and the link to selenium toxicity identified. This kind of study could be done under a Tier 3 concept within EMAP. More extensive spatial sampling and perhaps intensive temporal sampling might be required. A probability sampling approach might still form the basis of the monitoring effort.

A second use of Tier 3 might be to evaluate concerns that exist nationwide but are restricted to a specific subpopulation not adequately covered using the standard EMAP grid density. It is important, before embarking on this type of Tier 3 activity, to have an understanding of the size of this subpopulation in relation to the resource as a whole. This activity should also be accompanied by funding above the base program so that it does not impact the Tier 1 and 2 efforts. The current acidic deposition scenario is an excellent example of this situation. The initial questions of status and extent (Tiers 1 and 2) for aquatic effects of acidic deposition were addressed during the National Surface Water Survey (NSWS), conducted as part of the National Acid Precipitation Assessment Program (NAPAP). Changes in emissions have now been proposed under the renewal of the Clean Air Act. There is a need for a program that evaluates how effectively these emission reductions improve conditions in affected aquatic systems. High-interest areas, such as the Adirondacks, Southern Blue Ridge, and the Sierras, are too small for the Tier 1 EMAP grid density to evaluate adequately; interpreting any regional changes in aquatic conditions in these areas will require a few well-defined sites that receive more temporally intensive monitoring than that provided by the annual regional visits of Tiers 1 and 2.

Tier 4 is a research component that complements Tier 1, 2, and 3 activities. It is important to think of Tier 4 as an activity rather than a specific place. It is the activity of ecological research that is critical, rather than a particular place. Initially, EMAP will not establish new research sites, but will rely on existing programs to play this role. The National Science Foundation's Long Term Ecological Research (LTER) Program maintains a network of 17 sites across the United States. The Department of Energy has a Research Park network. Additional sites are maintained by other federal agencies and university researchers. These are all sites of active research on the functional and structural aspects of ecosystems. These networks provide (1) important information on system characteristics, which will be useful in developing and testing indicators of ecological condition, (2) detailed process information characterizing aspects of systems not amenable to general monitoring, and (3) a link between spatially extensive and temporally intensive data, to increase the interpretation of EMAP results. Tier 4 should generate input that modifies activities at Tiers 1, 2, 3 and should in turn be influenced to evaluate hypotheses generated from results of Tier 1, 2, and 3 activities. This interaction between Tier 4 and Tier 1, 2, and 3 activities should ensure that activities of the first three tiers continue to be grounded in sound ecological principles. Although EMAP will not initially establish its own Tier 4 sites, we propose that some of the existing research sites be included in the Tier 2 sample. As EMAP proceeds, gaps in the existing Tier 4 networks may be identified and filled.

This document addresses activities primarily at Tiers 1 and 2. The Tier 3 activity, the TIME Project, is discussed briefly throughout this document, as it pertains to modifications or specific applications of the general approach, and more completely in Stoddard (1990).

2.1.1 Design

In trying to describe resources such as lakes or stream reaches for the nation, we could either

census all the systems or sample a subset and extrapolate from that subset to the entire population of interest. Given the number of lakes and stream reaches, the census approach is impractical. Thus an approach for selecting and sampling a subset is needed. We have chosen to select this subset on a "probability" basis; that is, our subset is selected in such a way that the probability of including any particular lake or stream reach can be calculated. Information from the sampled systems can be weighted by the inverse of the inclusion probability to calculate unbiased regional estimates.

Meeting the EMAP-Surface Water objectives outlined in Section 1 requires developing a design that is capable of:

- Estimating, with known uncertainty, the status and health of any regionally defined resource.
- Describing baseline data leading to rigorous detection and description of trends in status and health of regionally defined resources.
- Identifying associations among attributes, both within and among resources, to establish possible causes of impaired condition.
- Quickly responding to new issues and questions.

Important requirements and features of the design include:

- Explicit definition of target populations and their sampling units.
- Explicit definition of a frame for listing or otherwise representing all the potential sampling units within each target population.
- Use of probability samples on well-defined sampling frames to estimate population attributes rigorously through randomization.
- Flexibility to accommodate a variety of resource types and a variety of problems, some of which have not yet been specified.
- Hierarchical structure permitting sampling at a coarser or finer level of resolution than the general grid density, giving flexibility at global, national, regional, or local scales.
- Ability to focus on subpopulations of potentially greater interest, such as specific types of lakes rather than all lakes.
- Ability to quantify statistical uncertainty and sources of statistical variability for populations and subpopulations of interest.

The proposed design strategy is a randomized, systematic sampling of sites selected in association with a uniform spatial grid. This sampling strategy employs a permanent national sampling framework consisting of a hexagonal plate containing a triangular grid of approximately 12,600 points placed randomly over the conterminous United States (Figure 2-4); a similar array is available for Alaska and Hawaii. A defined area (40-km^2 hexagon) around each point will be characterized using existing maps, satellite imagery, aerial photography, and existing databases of landscape information (Figure 2-5). These data will be periodically updated, most likely at 10-year intervals.

These 40-km^2 hexagons describe a sample area one-sixteenth of the area of the United States, and provide the basis for regional estimates of the areal extent of resources and numbers, and their changes over time. The collection of resource units contained within these 40-km^2 hexagons for any explicitly defined subpopulation (e.g., lakes between 10 and 2,000 ha) constitutes a Tier 1 sample. From data collected at Tier 1 (e.g., landscape descriptions), classification and further subpopulation development can be conducted. The sample of resources reflects the true properties of that resource. This follows from the rigorous prescription of the protocol.

The triangular nature of the grid points also permits us to increase or decrease the density of the grid three-, four-, or sevenfold, as well as by multiples of these values (Figure 2-6). This flexibility allows for sampling at densities appropriate for the resource, whether it be common or scarce, and at the scale of resolution desired. The design is flexible enough for sample selection at a resolution appropriate for use by individual States should they desire.

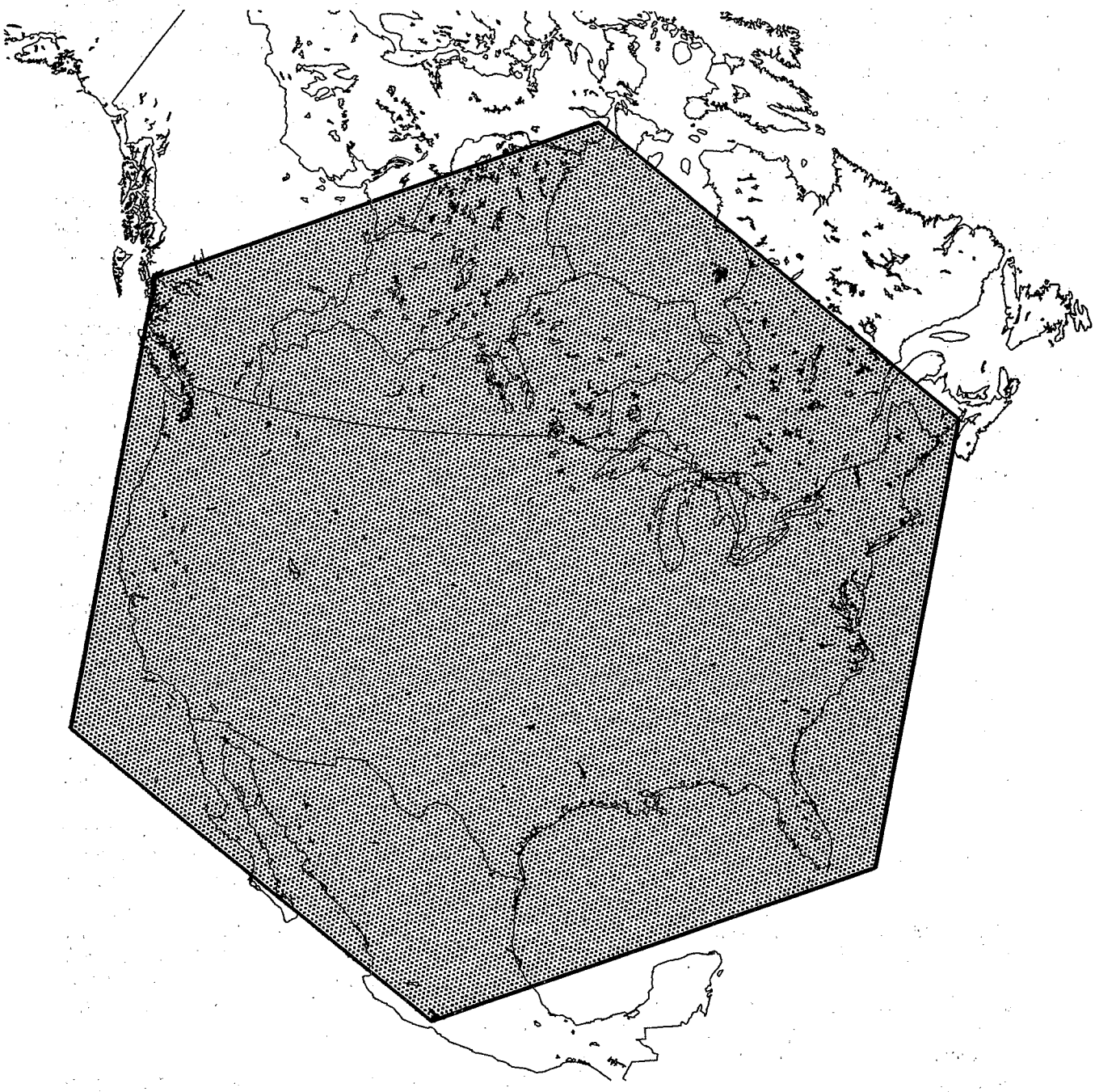


Figure 2-4. The baseline grid (not randomized) for North America containing about 12,600 points in the conterminous United States. Spacing between points is about 27 kilometers.

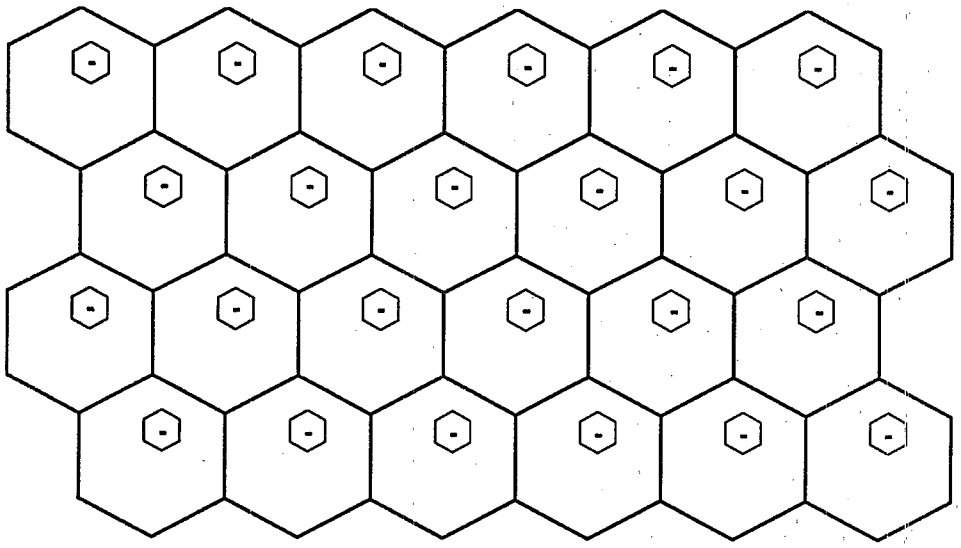


Figure 2-5. The landscape characterization hexagons are one-sixteenth of the total area and are centered on the sampling points. The randomly positioned sampling grid occupies a common but randomly selected position in each of the base tessellation hexagons.

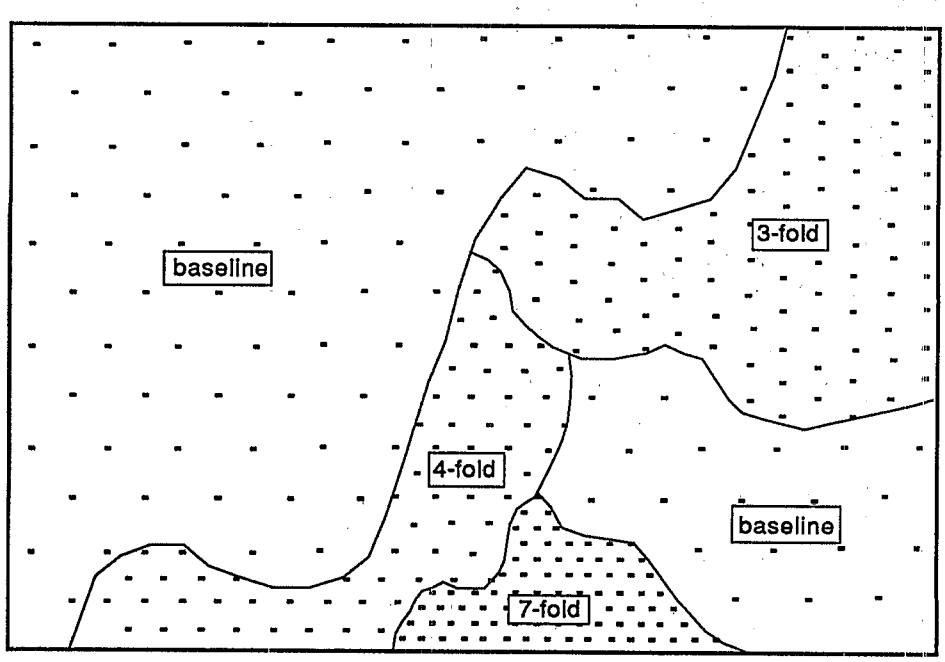


Figure 2-6. Enhancement factors for increasing the base grid density. Enhancement will be made only in the sample grid, not in the base grid.

The characterization carried out on the 40-km^2 hexagons will provide an inventory of all aquatic resources within them. Thus it will provide a one-sixteenth sample of the extent, abundance, and distribution of any aquatic resource that can be identified during the landscape characterization. A subsample of this Tier 1 sample, selected by probability methods, will be visited for detailed site characterization of indicators of condition of the resource. Data obtained from these field samples will constitute the information to be used for reporting on regional status and trends in ecological condition and exposure. This second stage of the double sample will constitute Tier 2 of the design.

The complete landscape description at Tier 1 ensures that no ecological resource is excluded by definition of the monitoring coverage, so that there are no orphan resources. This does not mean that every resource of potential interest is necessarily monitored at the Tier 2 or field level. The Tier 2 sample is implemented for resources of current concern. If a need to monitor an additional class of resource is identified, it will be necessary to return to the Tier 1 level, identify the extent and location of that resource, and select appropriate sites for field sampling. However, the existence of the landscape characterization provides a ready-made frame for sample selection, so that a field program can be mounted with minimal preparation. This provides the potential for quick response to new or emerging issues of ecological condition.

An important aspect of the EMAP design is the temporal and spatial interpenetrating nature of site characterization and field visits. Whereas the sampling grid consists of 12,600 points distributed across the conterminous United States, only one-fourth of these will be considered each year and a subset of those will be sampled. Thus, over a four-year cycle, each grid point will be considered. After four years, the cycle will be repeated, revisiting grid points and sites selected during the first four-year cycle. Figure 2-7 illustrates the temporal and spatial distribution of the interpenetrating approach. During the early years of the program, a subset of sites will be visited each year to evaluate year-to-year variability and to identify important trends that might be detected early. A few new sites will be selected each year to estimate any bias that might be introduced by knowledge that certain sites are EMAP sites, resulting in differential treatment of those sites or surrounding areas (see discussion in Section 6).

2.1.2 Indicator Approach

Traditionally, environmental monitoring programs have focused on individual chemical or biological species at the local scale. The concentration or response of these individual species was assumed to be equivalent to poor or good environmental condition. Individual chemicals and organisms, however, do not exist in isolation, but rather interact with other physical, chemical, and biological factors to produce ecosystem response. In addition, we have learned over the last two decades that continuing, persistent, and cumulative pollution is occurring not only at the local scale but also on regional, continental, and global scales. In some instances, control and regulatory programs at the local scale have aggravated or contributed to problems at the regional scale (e.g., acidic deposition). The complexity of environmental processes and the concept of regional and national scale effects require a new approach to environmental monitoring, both in terms of what we measure and where we measure it.

Instead of focusing on problems that receive the greatest media or public attention, we need to focus on problems that pose the greatest risk to the environment. In *Unfinished Business: A Comparative Assessment of Environmental Problems* (U.S. EPA 1987b), the EPA considered the highest potential risks to ecosystems to be global warming and stratospheric ozone depletion, followed by regional problems of habitat alteration, nonpoint source pollution, and risks from criteria air pollutants. Because environmental problems are becoming increasingly complex (i.e., cumulative effects from multiple pollutants at multiple scales), traditional monitoring approaches and indicators alone are not enough to assess ecological condition. EMAP is being designed with a "top-down" approach in a risk assessment framework. This top-down approach puts the primary focus on the endpoints of concern rather than on the environmental perturbations or stressors.

Risk assessment is the formal framework adopted by EPA to estimate the risks to human health and welfare arising from various hazards and associated mitigation strategies (Yosie 1987, Deisler 1988). Although initially applied by EPA in

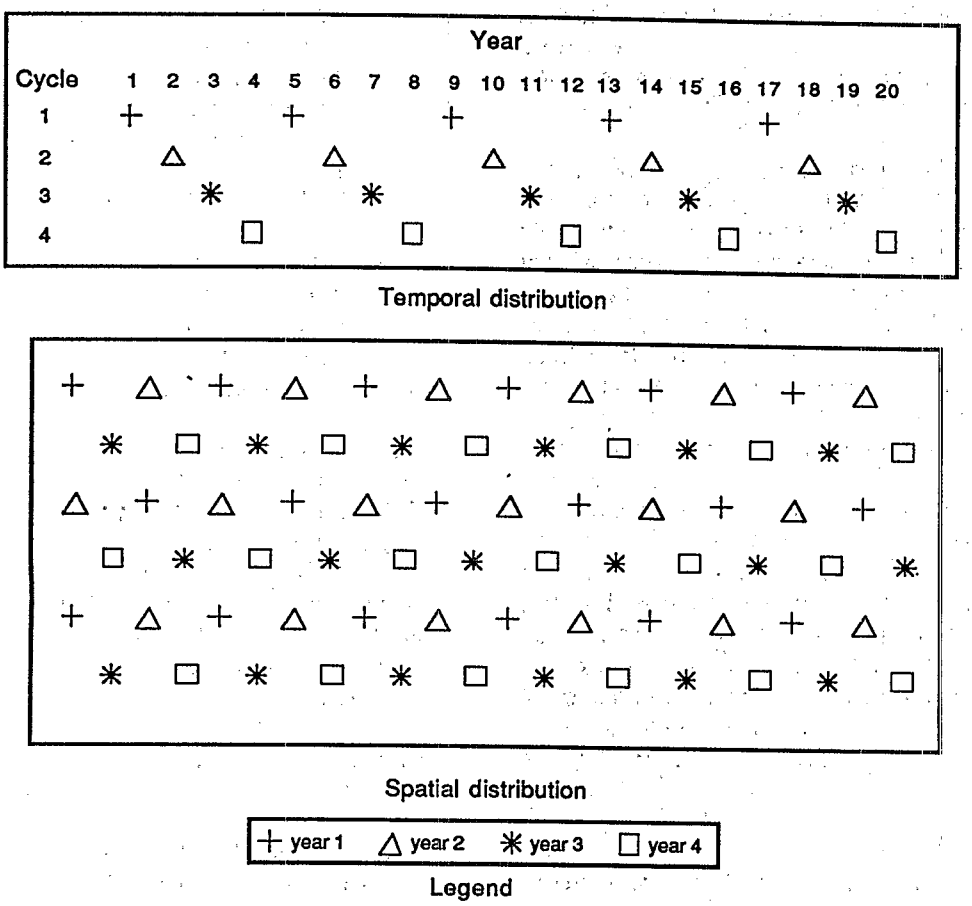


Figure 2-7. Spatially interpenetrating samples on a 4-year rotating cycle.

the human carcinogen arena, the risk assessment approach is now being considered for evaluating ecological risks at various spatial scales (Bascietto et al. 1990, Hunsaker et al. 1990). Ecological risk assessment procedures estimate the effects of anthropogenic activities on ecological resources and allow the significance of those effects to be interpreted with quantified uncertainty estimates. The ideal output of the risk assessment process is an estimated probability that an event of a certain magnitude will occur (e.g., a 90% probability of loss of fish populations in 20% of the lakes in a region; Suter 1990). Key components of risk assessment include (1) selection of endpoints, (2) qualitative and quantitative descriptions of the sources of the hazard (e.g., locations and emission levels for pollutant sources), (3) identification and description of the reference environment within which effects occur or are expected to occur, (4) estimation or measurement of spatiotemporal patterns of exposure, and (5) quantification of the relationship between exposure in the modified environment and effects on biota (Hunsaker and Carpenter 1990). The adaptation of the standard risk assessment paradigm to accommodate ecological risk at different time and space scales has been proposed (Suter 1990, Hunsaker et al 1990, Warren-Hicks et al. 1989) and is currently under active discussion within EPA (Bascietto et al. 1990).

Although the sequence of activities in traditional risk assessment moves from source to exposure assessment and then to effects assessment, an epidemiological or top-down approach to ecological risk assessment is also appropriate (Suter 1990, Fava et al. 1987). In effects-driven risk assessments, we observe an effect or response,

then work from this response toward identification of the hazard or stressors by using exposure information. This top-down or effects-driven assessment is under development for EMAP.

Ecological risk assessments focus on the valued ecosystem attributes or endpoints of concern. Suter (1990a) recognized two types of endpoints that are used in risk assessments. **Assessment endpoints** are formal expressions of the actual environmental value that is to be protected. They should have unambiguous operational definitions, have social or biological relevance, and be amenable to prediction or measurement (e.g., a decrease in biodiversity, probability of a >10% reduction in game fish production, or the proportion of waterfowl killed within a region of pesticide use). The quantitative results of the measurements taken to characterize assessment endpoints are termed **measurement endpoints**. Measurement endpoints must correspond to or be predictive of assessment endpoints. Examples of measurement endpoints include the percentage of trees that exhibit visual symptoms of foliar damage due to oxidants, or the fraction of streams that exceed the LC_{50} for a toxicant to largemouth bass.

Assessment endpoints by definition must relate to the environmental values identified for each resource category in EMAP. Therefore, the measurement endpoints for EMAP are ideally response indicators. Exposure and habitat indicators are not our first choice as measurement endpoints because they do not integrate the effects of multiple pollutants or the effects of unknown hazards. Stressor indicators are farthest removed from the effects of interest and therefore are also poor measurement endpoints. Definitions of these four indicator types are:

- **Response Indicator**: A characteristic of the environment measured to provide evidence of the biological condition of a resource at the organism, population, community, or ecosystem level of organization (e.g, trophic index, fish assemblages).

- **Exposure Indicator**: A characteristic of the environment measured to provide evidence of the occurrence or magnitude of contact with a physical, chemical, or biological stressor (e.g., nutrient concentrations, tissue residues, toxicity tests).

- **Habitat Indicator**: A physical, chemical, or biological attribute measured to characterize the condition necessary to support an organism, population, community, or ecosystem in the absence of pollutants (e.g., availability of snags, substrate of stream bottom, vegetation type and extent).

- **Stressor Indicator**: A characteristic measured to quantify a natural process, an environmental hazard, or a management action that effects changes in exposure and habitat (e.g., land use).

EMAP will collect data on response, exposure, and habitat indicators at its field sampling sites. Data on land use and land cover will be collected as part of the landscape characterization activity within EMAP. Additional stressor indicators will be assembled from other sources. Within EMAP, indicators are identified through the development of conceptual models of ecosystems. These models may be based primarily on how current and anticipated stresses affect ecosystems, or from a perspective of the structural, functional, and recuperative features of "healthy" ecosystems.

Specific indicators for EMAP-Surface Waters are discussed in Section 4; they should be seen as a candidate list reflecting current progress but requiring refinement. The suite of indicators selected must be used in a coordinated manner to provide a comprehensive picture of surface water condition.

Selection and development of indicators for EMAP-Surface Waters will be a long-term process. Some indicators are ready for implementation on a national basis, others have been used effectively in certain regions and require modification for new regions, and still others have been used only in isolated instances and require demonstration of their regional and national applicability. A multiphase process has been identified to guide the current selection of indicators and their future development:

- Phase 1 - identification of issues (environmental values and apparent stressors) and assessment endpoints.

- Phase 2 - development of a set of candidate indicators linked to the identified endpoints and responsive to expected stressors.

- Phase 3 - screening of the candidate indicators based on a set of indicator evaluation criteria, selecting as research indicators those that appear to fulfill key requirements, rejecting those that clearly don't, and holding in a state of evaluation candidate indicators that may, in the near future, advance to research status.

- Phase 4 - quantitative testing and evaluation of the expected performance of research indicators on regional scales, to identify the subset of developmental indicators suitable for regional demonstration projects.

- Phase 5 - implementation of a core set of indicators in a full national EMAP program with annual sampling and statistical summaries, and periodic reevaluation of indicators.

Selection of assessment endpoints, and the subsequent selection of indicators (Figure 2-8), requires consideration of public values, policies, and current threats to surface waters, and an understanding of biological and chemical processes. Figure 2-8 suggests that selection of biotic indicators of condition is driven by our understanding of biological communities, their interactions, and the impact of various stresses on the chemical and physical habitats in which the organisms exist. Selection of indicators is equally influenced by the endpoints or values of concern. These endpoints, and thus the indicators representing them, must be sufficient to motivate change in policy when poor conditions are found. The exposure/habitat and stressor indicators must be selected based on our current understanding of environmental stresses and their impact on the chemical and physical habitat and hence on the biota.

As discussed in Section 1, primary consideration in EMAP-Surface Waters will be given to those ecological characteristics for which we manage surface waters, specifically:

- Biological integrity
- Trophic condition
- Fishability

Figure 2-9 provides a general overview of EMAP-Surface Waters indicators, relating the selected endpoints, response indicators, exposure indicators, and stressors. Figure 2-10 shows how the results for one indicator (Index of Biotic Integrity) might be displayed in a cumulative frequency distribution (CDF). The Y-axis shows the cumulative percentage of the stream miles, and the X-axis shows the IBI scores. In this instance, an IBI of 40 was delineated as the acceptable score. Forty-two percent of the stream miles had a score of 40% or lower. We could transfer the data into a bar chart showing 58% of the population in nominal or acceptable condition and 42% in subnominal condition.

In addition to knowing what proportion of the surface waters are in good or poor (nominal/subnominal) condition, it is important to determine the likely causes of these conditions. In EMAP, this diagnosis will be achieved primarily by using the exposure, habitat, and stressor indicators as defined earlier. In limited situations, the response indicators themselves may shed some light on likely causes of current conditions or trends. Statistical associations between the occurrence of poor conditions as defined by response indicators and values for the exposure, habitat, and stressor indicators will be used to infer the most likely categories of probable cause. Correlative analyses of these types cannot prove causality but will narrow the range of likely explanations for observed regional patterns. Ultimately, we would like to be in a position to display the results for each of the three endpoints with a nominal/subnominal split, and partitioning of the subnominal sites to the probable cause for poor conditions (Figure 2-11).

2.2 DATA QUALITY

An essential complement to the careful design and indicator strategies outlined earlier in this section is consideration of data quality. Production and assurance of quality data should be an integral part of any program that intends to produce useful information. EMAP is committed to producing data useful for its intended purposes. To meet this objective, it is necessary to have a process for identifying the quality of data needed. The following discussion represents a process that has just begun for EMAP-Surface Waters but that should be ongoing during the early stages of implementation.

Data quality objectives (DQOs) are statements of the level of uncertainty a decision maker is willing to accept in results derived from environmental

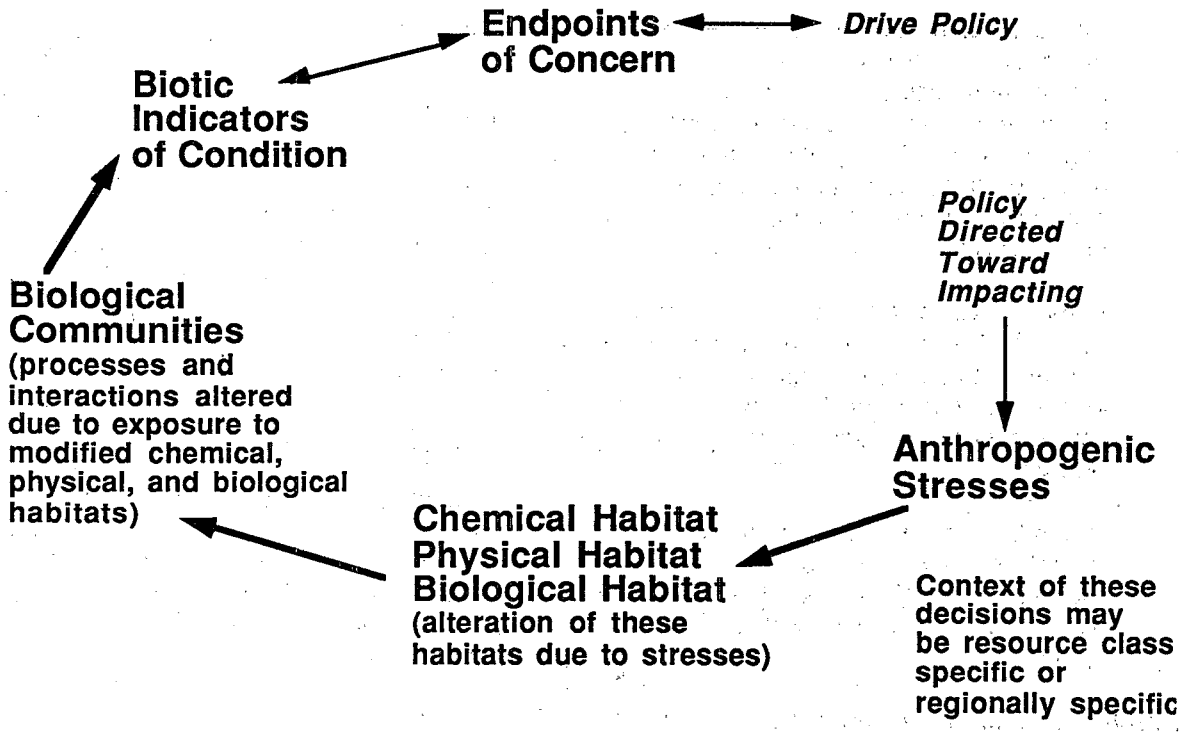


Figure 2-8. General approach for identifying indicators. It requires consideration of the public values or endpoints (which drive policy), the potential stresses (at which policy is generally directed), and the current understanding of interactions between chemical, physical, and biological processes.

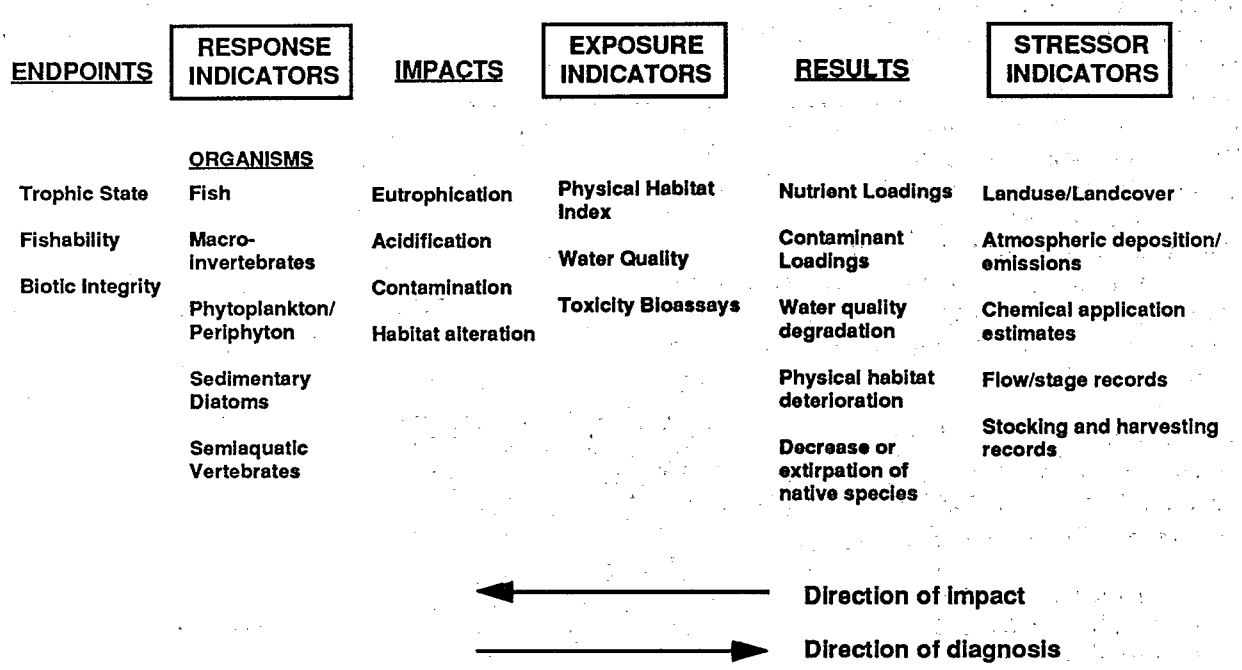


Figure 2-9. Indicator approach for EMAP-Surface Waters showing candidate indicators and the top-down approach to problem identification and diagnosis of probable cause.

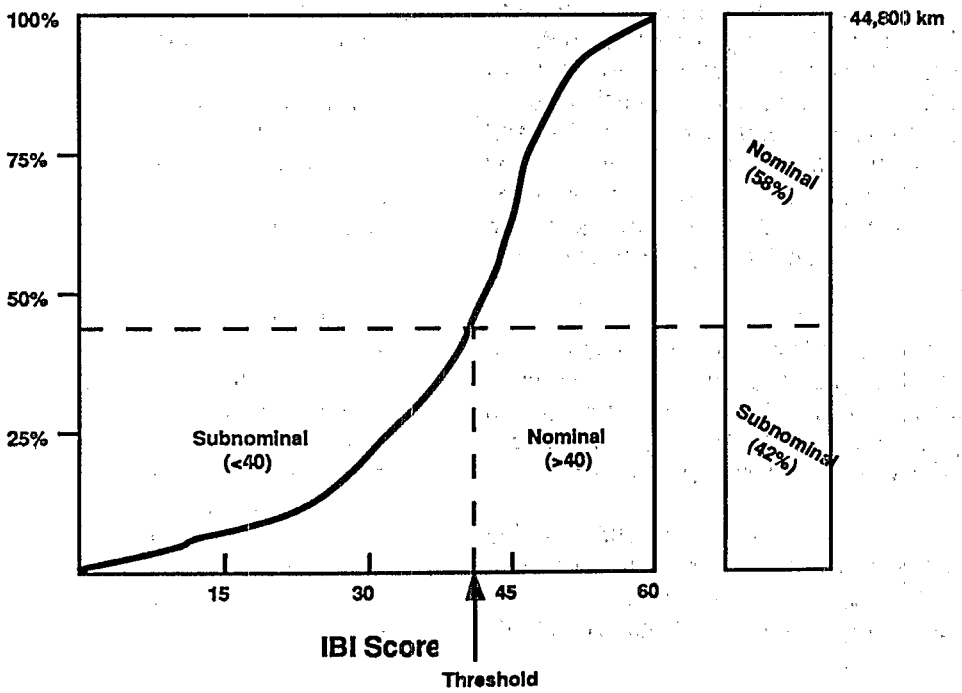


Figure 2-10. Cumulative frequency distribution for Index of Biotic Integrity in a region showing how data might be displayed. With a threshold value of 40 identifying systems in poor or good condition, the data can be transferred to a simple bar chart displaying the proportion of the stream miles in nominal or subnominal condition.

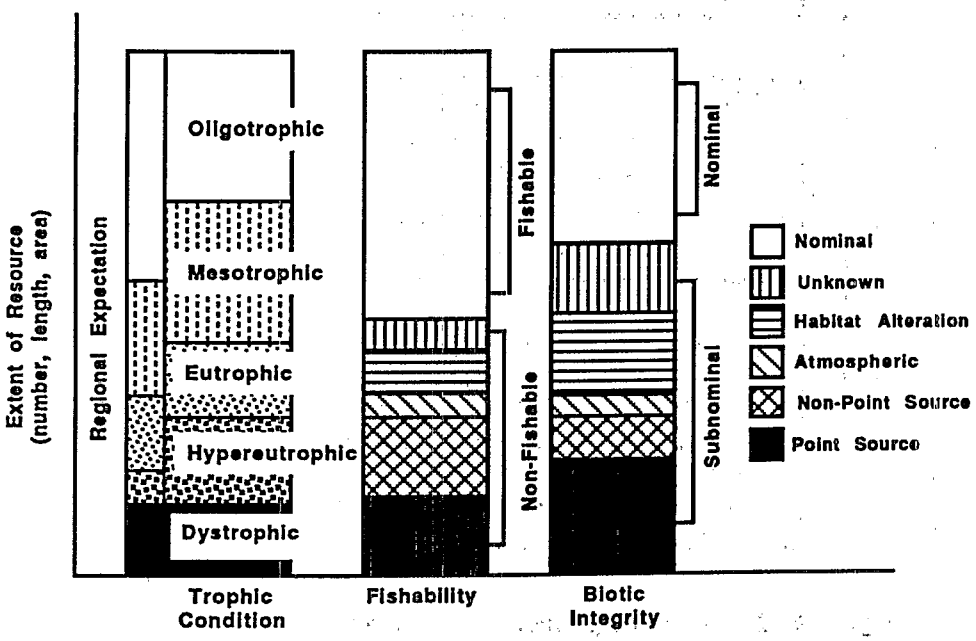


Figure 2-11. Long-term assessment objectives for EMAP-Surface Waters.

data. The DQOs are generated in a multi-stage process (Figure 2-12). Stage I involves defining the major questions or problems of concern. At this stage, the focus is on the decision maker or data user. Information needs are identified, based on an understanding of how the environmental data to be collected will be used. Also, resource and time constraints are identified and the consequences of Type I and Type II errors are examined.

A Type I error results in a false positive (e.g., data indicate adverse ecological effects when no such effects exist), whereas a Type II error results in a false negative (e.g., data indicate no adverse ecological effects when they actually exist).

Stage II involves defining the information needed to answer the questions or make the decisions identified in Stage I. This process includes developing pertinent subordinate questions that may need to be answered in order to fully address the problem. Also, at this stage, the population of interest should be clearly defined. There should also be an identification of specific design constraints, i.e., the desired confidence in the results. Finally, Stage II should examine existing data and confirm the need for new data, in cases where data do not exist to provide the necessary information.

Stage III involves determining a scientific approach to data collection and the data quality requirements for that approach. This process includes considering as many approaches to collecting the necessary data as possible, along with considering the levels of data quality required to meet the constraints specified in Stage I and Stage II. This stage should also serve to identify research activities needed to meet the requirements of Stages I and II.

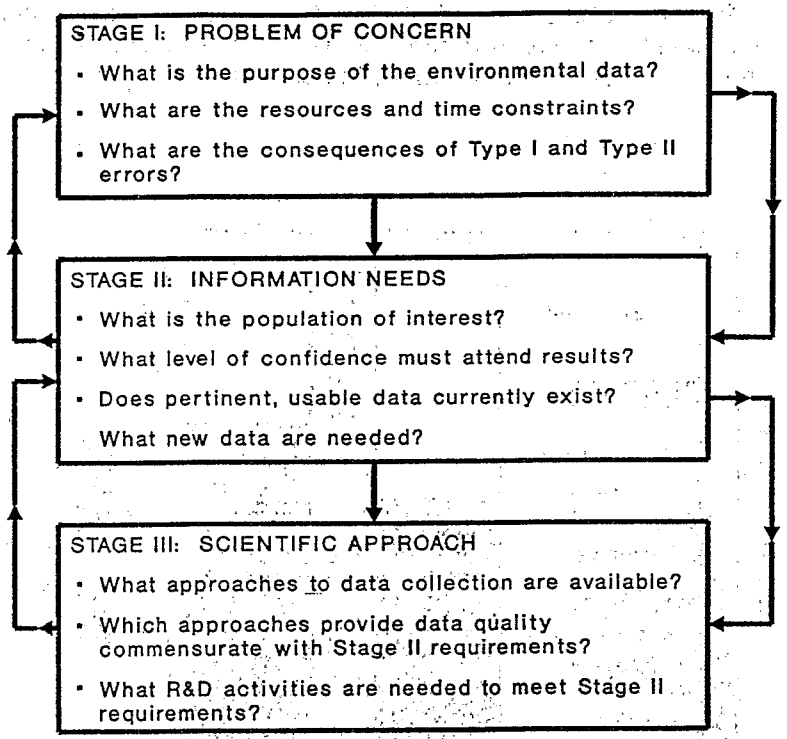


Figure 2-12. The three phases of the data quality objective process.

Throughout the process, there is continuous feedback and communication at all levels involving the decision makers, the scientists involved in identifying specific information needs, and those individuals collecting the data necessary to provide this information. EMAP is committed to the DQO process as a means of optimizing the allocation of limited resources and assuring that data collected in the program provides the information needed to meet program goals.

2.2.1 The Data Quality Hierarchy

Data quality exists at several levels (Figure 2-13). Measurement data quality includes attributes such as precision, accuracy, representativeness, comparability, and completeness associated with the measurement of environmental variables. At the next level, it also includes the uncertainty associated with the methods used to assimilate these measurement data into an assessment (i.e., provide information from the data). For EMAP, this can be perceived as "indicator quality," since the indicators are the tools used to provide information about specific aspects of ecosystem condition. Factors affecting quality at this level comprise not only measurement data quality, but also sampling design and statistical data analysis. At the next level, these indicators are aggregated into an overall assessment of system condition. The uncertainty associated with each indicator must be included in an estimate of the overall uncertainty of this aggregate assessment. This uncertainty will then be compared to ecosystem-level quality objectives to assure that data collection and interpretation activities are consistent with program objectives in terms of the quality of the information provided. Finally, EMAP intends to integrate information across ecosystems in order to make regional-scale assessments of ecological condition. Again, the uncertainty associated with each component of the evaluation (i.e., individual ecosystem

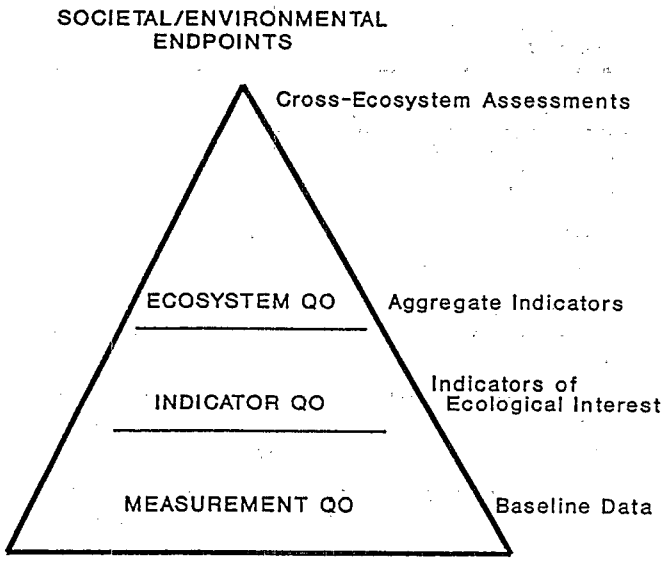


Figure 2-13. Quality objective hierarchy in the data quality objective process.

assessments) must be incorporated into an estimate of overall uncertainty for the assessment and compared to cross-ecosystem quality objectives. The DQO process requires that sources of variability be identified at each level and that all relevant sources be considered in generating estimates of uncertainty at any level of the hierarchy.

2.2.2 The Role of DQOs in EMAP

The EMAP mission provides the Stage I input to initiate the DQO process. The EPA perceived a need within the Agency and in other client groups for information regarding the current extent of various ecological resources (i.e., how much of each resource exists), the current status or condition of those resources, and some indication of trends in extent and condition over time. The need for environmental information has been stated in very qualitative terms at this level.

At this point in the process, the tools necessary to measure extent and define condition within each of these ecosystems are being developed. This, in turn, will allow policy and decision makers to articulate the requirements for data quality in more quantitative terms. The program is now in Stage II, with extensive feedback to Stage I, and the process may require several iterations.

In Stage II, each resource group within EMAP must develop a series of indicators that, in aggregate, allow for an overall assessment of ecosystem condition. Quantitative "logic statements" must be developed describing the data to be collected for each indicator and the way in which the data will be used to provide information on system condition. These statements should include critical values above which the system is in an acceptable or marginally impacted condition (i.e., nominal) and below which the system is significantly impacted (i.e., subnominal). This critical value must be scientifically defensible. Where possible, these statements should also relate each indicator to endpoints of societal interest or concern, so the ramifications of changes in system condition can be understood and appreciated by a variety of client groups.

In addition to developing these logic statements, a series of error constraints must be developed that identify all known sources of error or uncertainty associated with the indicator. These will include measurement error (the difference between sample values and in situ true values). Measurement error can be further divided into analytical error (error associated with the measurement process) and error from other sources such as sample design, collection, handling, storage, and preservation. Sampling error is a function of natural spatial and temporal variability and sampling design. Wherever possible, estimates should be provided for each source of uncertainty. In this way, factors that contribute significantly to the overall variability of the indicator are identified and the effectiveness of various options in resource allocation can be evaluated. For example, if spatial variability is the major factor in the overall uncertainty associated with an indicator, and measurement error is small by comparison, it may be judicious to use a less precise and less costly method of analysis and invest more resources in increasing the sampling density within a region to reduce the overall uncertainty in the data.

Early in the program, individual indicators will be used to make discreet assessments of condition. Tools for making aggregate ecosystem assessments and cross-ecosystem assessments will be developed over time. The DQO process should provide the framework for this development, assuring that assessment tools at all levels provide information of sufficient quality to meet program objectives.

2.2.3 DQOs in the Surface Water Component

In order to assimilate the information necessary to evaluate indicators of potential interest to the Surface Water project, a questionnaire was developed and sent to researchers with expertise in the use of each indicator. Upon completion, these questionnaires will become "fact sheets" that provide the initial input for the development of indicator quality objectives. Table 2-1 is an example of a fact sheet for the sediment toxicity test. For some indicators, data exist that allow quantitative estimates to be developed for each source of uncertainty identified for the indicator. Historical databases for these indicators are currently being evaluated by the surface waters resource group. However, for many indicators of potential interest to the Surface Water project, no data of the appropriate scale or methodology exist from which to generate such estimates. Therefore, data from monitoring activities performed in 1991 will be used to develop quantitative logic statements and error constraints for these indicators.

Table 2-1. Sample Fact Sheet and Logic Statement for EMAP-Surface Waters

EMAP-SURFACE WATERS INITIATIVE INDICATOR FACT SHEET

INDICATOR: Lake Sediment Toxicity

1. What is the problem or environmental concern that this indicator will address?

   Are the bottom sediments of lakes toxic to aquatic life?

2. How does this problem relate to societal interests or concerns?

   If bottom sediments are toxic to aquatic life, they may also be a hazard to humans. These sediments also provide critical habitats for many species necessary to maintain the aquatic food chain. As a result, sediments that are toxic to these species will adversely affect the natural balance of the ecosystem, commercial and recreational fisheries, and the aesthetics of lake environments.

3. What information is needed to address the problem?

   A method of measuring sediment toxicity.

   Subissues:

   a. A definition of what is to be considered toxic (acute versus chronic, degree of response, etc.).
   b. One or more indicator test organisms that, once exposed to sediment, exhibit an adverse response (e.g., mortality, impaired growth, reduced reproduction, etc.) if the sediment is toxic.
   c. A proper laboratory test design to reflect in-field toxicity.

4. What data will be collected in the field? By what method or methods?

   Collection of sediment samples.

   Subissues:

   a. Sampling device (core, grab, etc.)
   b. Number of samples per lake
   c. Depth of sample

5. What laboratory analyses will be performed? By what method or methods?

   Laboratory bioassays.

   Subissues:

   a. Choice of species
   b. Type of exposure (solid-phase, pore water, elutriate)
   c. Acute versus chronic
   d. Chronic endpoints
   e. Test duration

(Continued)

Table 2-1. Sample Fact Sheet and Logic Statement for EMAP-Surface Waters (Continued)

Proposal:

EMAP-Surface Waters will perform 10-day acute solid-phase bioassays with *Hyalella* sp. amphipods and 7-day solid-phase chronic assays with *Ceriodaphnia* and fathead minnow larvae. Overlying, clean water (reconstituted laboratory water) will be used in all tests. The endpoints for *Ceriodaphnia* tests are mortality and reproduction; the endpoints for the fathead minnow test are mortality and growth.

6. How will the field/laboratory data be used to provide the information necessary to address the problem? Be specific.

Data from laboratory bioassays will be analyzed statistically to determine if field-collected sediment samples produce mortality significantly greater than control (control sediment to be defined). Such samples will be categorized as acutely toxic. Samples for which chronic endpoints show responses significantly different from controls will be categorized as chronically toxic. By inference, these sediments should fail to support similarly sensitive species of comparable life-stage in the field and would result in population and community level impacts in these systems.

7. What are the sources of variability for the indicator? Where possible, provide estimates of the variability associated with each source.

Examples of sources of variability:

Measurement - Analytical
Other (sample collection, handling, etc.)

Sampling - Spatial
Temporal

Measurement variability can be due to differences in the health of test organisms, to variability associated with multiple operators, or to variability in control sediment or laboratory water between laboratories or over time. For the *Ceriodaphnia* and fathead minnow tests, single laboratory precision is approximately 30-35%. Amphipod bioassay data will be assessed to determine the variability associated with this method. Nonmeasurement sampling variability is unknown at this time.

Sampling: Representativeness of sediment sample, that is, how well does a sediment sample represent all sediment in a given lake?

Test representativeness, that is, how well does the laboratory bioassay assess the toxicity of the sediment in the field?

DRAFT LOGIC STATEMENT FOR SEDIMENT TOXICITY INDICATOR

What number and percentage of lakes in a given region of the United States have sediments toxic to aquatic life, as demonstrated by effects observed relative to controls in laboratory bioassays?

2.3 REPORTING

EMAP-Surface Waters will produce four types of reports:

1. Annual statistical summary reports
2. Periodic interpretive reports
3. Special interest reports
4. Scientific articles

To be of maximum use, the data must be transformed into useful information as quickly as possible. The objective of EMAP-Surface Waters is to publish summaries of the preceding field season surveys within nine months of data collection. These reports will provide summaries of response, exposure, and habitat indicator data for resources and regions sampled, but will provide minimal interpretation.

Interpretive reports will be published following completion of four-year sampling cycles. These reports will attempt to summarize indicator results for the preceding four years and will integrate information from the suite of response, exposure, habitat, and stressor indicators to determine the regional and national status of lakes and streams. Trends will be evaluated and the likely causes for current conditions and changes in condition will be assessed.

Special issue reports will be used to present data pertinent to specific issues such as changes in surface water chemistry resulting from implementation of the Clean Air Act.

2.4 EMAP-SURFACE WATERS PROGRAM LIMITATIONS

It is important to describe the limitations of EMAP-Surface Waters, as well as what the program will attempt to do. The program is not intended to describe all components of an ecosystem or resource type. It will not describe how systems function. It will, however, provide information about the condition of specific indicators at a specific index period as a "snapshot" of the overall condition of a system.

EMAP has not been optimized to address specific stressors to surface waters such as acidification or chemical contamination. Studies addressing specific stressors are envisioned as the third tier of EMAP. Currently, the only Tier 3 activity planned for EMAP is the TIME project, which will modify the EMAP-Surface Waters design to address the response of surface waters to changes in acidic deposition. EMAP is not intended to be compliance monitoring and will not replace the need for those activities. In general, EMAP is intended to provide a common sampling frame within which to assess the condition of lakes and streams on a broad scale so that the magnitude and geographical location of various problems can be assessed and mitigation and research priorities can be set more objectively. The monitoring program is not intended to be truly anticipatory, but rather to provide an ongoing monitoring framework within which new variables can be added or regional modifications can be made so that the magnitude and extent of new issues can be more quickly evaluated.

3.0 MONITORING NETWORK DESIGN

The broad EMAP objectives and the more specific surface water components of those objectives have been discussed in previous sections. Meeting these objectives requires a design capable of:

- Estimating, with known uncertainty, the status and health of any regionally defined resource.
- Describing baseline data leading to rigorous detection and description of trends in status and health of regionally defined resources.
- Identifying associations among attributes, both within and among resources, to establish possible causes of impaired condition.
- Responding quickly to new issues and questions.

Important requirements and features of the design described in this section include:

- Explicit definition of target populations and their sampling units.
- Explicit definition of a frame for listing or otherwise representing all potential sampling units within each target population.
- Use of probability samples on well- defined sampling frames to rigorously estimate population attributes through randomization and use of probability methods for sample unit selection.
- Flexibility to accommodate a variety of resource types and a variety of problems, some of which have not yet been specified.
- Hierarchical structure that permits sampling at a coarser or finer level of resolution than the general grid density, giving flexibility at global, national, regional, or local scales.
- Ability to focus on subpopulations of potentially greater interest (e.g., specific types of lakes rather than all lakes).

A general overview of the EMAP design was presented in Section 2.0; greater detail about the design can be found in Overton et al. (1990). This section describes how the general EMAP design will be used in the surface water component of EMAP. It includes sections on specification of explicit populations of lakes and streams, sample units, frame development, and association rules by which lake and stream sample units will be selected for field visits.

This section should give the reader a clear picture of how sample units will be selected to meet the design objectives of EMAP. Some decisions have been made regarding pieces of the design puzzle; for other pieces, various options under consideration will be presented, some only briefly. Our initial attention is on lakes, particularly lakes in the Northeast, the area of the country where EMAP-Surface Waters will conduct a pilot study. Examples from this area are used to illustrate some of the decisions, procedures, and unanswered questions.

Consideration of streams is less well developed. The present schedule calls for a stream pilot study in FY92 that will primarily address questions about identifying indicators for arid streams of the west. Thus the directions and decisions regarding streams are more tentative and general.

3.1 POPULATIONS, SAMPLE UNITS, FRAMES, AND OTHER SAMPLING CONCEPTS

3.1.1 Surface Water Populations: Defining Lakes and Streams

To answer questions about the condition of lakes or streams, it is necessary to specify explicitly what constitutes the group under consideration. This target population of lakes or streams constitutes the body about which estimates of condition will be made. Target populations can be regionally defined, such as lakes in the Northeast, or defined on the basis of certain attributes, such as lakes larger than 20 hectares. However, one of the general goals of EMAP is to include all ecosystems, so that there are no "orphan" systems. Therefore, at the broadest level, the target population for EMAP-Surface Waters is all lakes and streams in the United States (Alaska and

Hawaii included), excluding the Great Lakes, for which a separate component has been developed.

What constitutes a lake or stream? Cowardin et al. (1979) developed a classification system for wetlands and deepwater habitats of the United States. Deepwater habitats are "permanently flooded lands lying below the deepwater boundary of wetlands" and wetlands are "lands transitional between terrestrial and aquatic systems where the water table is usually at or near the surface or the land is covered by shallow water." This classification includes definitions of deepwater lacustrine (lake-like) and riverine habitats that serve as a starting point for explicit definitions of lakes and streams.

However, several modifications to the Cowardin definitions are necessary, primarily to distinguish what we generally consider lakes and streams from wetlands. In the Cowardin classification system, a deepwater habitat is > 2 m in depth and has a surface area of at least 8 ha. This is used as an approximate break between lacustrine (lake) and palustrine (marsh, swamp) habitats. Thus, our definition of a lake will correspond with Cowardin's deepwater lacustrine habitat for lacustrine waterbodies > 8 ha. To cover lakes < 8 ha, the combination of deepwater and lack of emergent vegetation will be used to separate lakes from wetland habitats. If the depth is > 2 m, the waterbody will be considered a lake. If the depth is < 2 m, the spatial extent of emergent vegetation will be used; if coverage is < 30% of the waterbody area, the habitat will be classified as a lake. One hectare will form the lower limit of habitats considered as lakes.

An additional element of the distinction between lakes and wetlands is the shifting of lakes to wetlands due to natural filling, anthropogenic filling, or climatic variations. Wetlands may shift to lakes as a function of climatic variation or direct management activities (e.g., impoundments or dredging). Tabulation of the shifts from lakes to wetlands and back will identify the long-term trends in resource shifts. Shifts may be observable as part of the periodic landscape characterization, or as a result of repeated site visits.

For streams, the primary modification of the Cowardin classification will be to include shallow as well as deepwater riverine habitats. As described later, the resource unit of interest will be stream channels. Since stream traces as shown on maps will be used to identify stream units, it will not be possible to determine without a field visit whether the channel represented by the stream trace actually contains water, particularly during the proposed low flow index periods. There are several possible states (ephemeral, intermittent, flowing) for the stream channel; tabulation of the state of stream channels relative to these possibilities will be necessary.

In practical terms, maps will be used to represent the populations of interest and to identify and select sample units (lakes and streams) (see Section 3.1.4 on frame development). The maps chosen will contain most of the water bodies meeting the above definitions, but they will contain errors of misclassification, omission, or location. In order to estimate the extent of lake and stream resources of different types, it will be necessary to estimate the error rates of the maps. Both landscape characterization (including a wetlands inventory component) and site visits will be used to determine if a mapped representation is an accurate depiction of the desired resource type. The probability-based design and site selection procedures permit inferences about map accuracy and estimation of the error in the map population estimates.

3.1.2 Subpopulations

The role of classification has been considered for surface waters. Other resource groups have chosen to classify for a variety of reasons, including:

- Flexibility of design, allowing the design to be modified for specific classes (e.g., sampling each class with a different approach).
- Sampling allocation, to prevent oversampling or undersampling of classes or to increase the precision of estimates about particular classes.
- Logistical and reporting convenience.
- Acknowledgment of natural regional classes among resource types.

For example, EMAP-Forests identifies particular forest types that can be delineated regionally. EMAP-Near Coastal defines biogeographic prov-

inces, and within each province, estuarine types. EMAP-Wetlands defines approximately 16 wetlands classes for which estimates can be made.

Within EMAP, subpopulations are the classes of resource types about which we wish to make statements of condition and trends. Tier 1 resources are one kind of subpopulation; they function as strata in the EMAP design. Other kinds of subpopulations are groups of lakes or streams that are not strata, but are classes of interest. These may be classes within a particular Tier 1 resource, or classes that include more than one Tier 1 resource. Subpopulations may, but need not, be hierarchical. There is great flexibility to establish a variety of subpopulations on which to focus the estimation of condition and trends. Subpopulations serve two major purposes: (1) to increase the precision about estimates of condition and trends, and (2) to target groups on which to report.

For surface waters, a variety of classifications have been discussed; in fact, there are so many ways in which surface waters could be classified, with substantive arguments for each, that we decided to make classification a part of the Tier 1 and Tier 2 evaluations. The flexibility of the design permits examination of candidate subpopulations to determine whether sufficient sample sizes could be obtained to make reasonable estimates of condition. If the examination indicates an insufficient sample size, the grid density can be increased to meet the desired sample size (if costs of the increased sampling can be met). This same flexibility permits combining sites into subpopulations after Tier 2 sampling (post-stratification) to make statements about the condition for different subpopulations.

Some examples illustrate possible surface water subpopulations: size classes of lakes or streams, reservoirs versus lakes; warmwater versus coldwater lakes, small seepage versus drainage lakes, alpine lakes, high-elevation streams, streams in EPA Region III, lakes in the glaciated Upper Midwest, lakes with acid neutralizing capacity (ANC) < 50 μeq/L, and lakes in agricultural or forested land. The flexibility of the EMAP design allows the specification of subpopulations at the Tier 1 level from landscape information or other available data, or at the Tier 2 level after collection and processing of field data.

Initial stages of the program will focus on the most general subpopulations, such as EPA regions (desirable by many as reporting units), and ecological regions or lake or stream regions to describe ecological conditions across relatively homogeneous areas. We are also considering size classes, to prevent oversampling the numerous small lakes and streams (see Section 3.2.2.1).

Subpopulations will also play a large role in the design of the TIME component of EMAP. In this case, subpopulation groups will be created early in the collection of field data, in order to assure that the required density of sites exists in each subpopulation. As described in Section 3.3, regional trend detection for subpopulations of acid sensitive lakes requires that approximately 15 (or more) sites be located in each of the regional subpopulations. It is anticipated that TIME subpopulations will be formed on the basis of sampled information. In particular, chemical characteristics of lakes in the Northeast pilot will be used to describe subpopulations of lakes that can be expected to respond similarly to changes in the levels of acidic deposition that result from emissions reductions.

The potential utility of other kinds of subpopulations will be evaluated during the early phases of EMAP-Surface Waters. Although a large number of potential subpopulations can be envisioned, we will be limited by the overall budget for routine sampling.

3.1.3 Sample Units

Sample units are the individual lakes or stream segments of interest and populations constitute the sum of all possible sample units. Identifying each sample unit uniquely is a necessary step toward site selection, to prevent duplicate selection by a single identification.

Individual lakes and reservoirs that meet the criteria defining lakes will be sample units. Operationally, the sample units will be those that occur on maps chosen to represent the populations of interest. Since maps will be used to select sites, and some of the criteria for defining lakes are not evident from the maps, field characterization will be used to identify waterbodies that are not lakes (do not meet our modified Cowardin criteria). A map error can be estimated by tabulating the numbers of non-lakes and making population inferences.

For streams, sample units will be segments of the stream network defined by the confluences or upstream termini of stream channels appearing on the selection maps. Segments could be defined based on morphometric properties such as channel type, gradient, or change in stream substrate. The logistics of implementing these types of criteria are excessive, however, relative to the possible improvement over using confluence rules. Maps will be used to select sites, and estimates of map error will be made via site visits, reconnaissance, or characterization.

Each sample unit will be uniquely defined by a node. As described later, the USGS Digital Line Graphs (DLGs; 1:100,000 map scale) will be used to create the lake frame. These digital maps conveniently contain label points for each lake, generally located at the center of the lake. Each label point uniquely identifies a lake in this digital database. Thus, for lakes, the label point will be the unique node (e.g., Figure 3-1).

For streams, we are presently planning to identify stream segments by locating nodes at the confluences and upstream termini of stream traces indicated on maps, as well as at the outlets of lakes and reservoirs. Each node would uniquely define a downstream segment (Figure 3-2). Another option would be to locate nodes at the midpoint of a stream segment identified by confluences and termini. This option would have advantages for the few instances in which streams bifurcate. In this case, each node would uniquely specify the segment within which it was embedded. The midpoint alternative would be relatively easy to implement if DLGs are used, but more difficult and time consuming if midpoints are located manually and need to be precise.

These stream segment definitions create some very short and some extremely long segments. A quick estimate based on sampling from three 1:100,000-scale maps is presented in Table 3-1. We tentatively suggest a minimum stream segment length of 500 m and a maximum of 5 km, but we have not evaluated the consequences of such delimiters. These are arbitrarily selected and intended to place bounds on the range of lengths of stream segments treated as sample units. We solicit arguments proposing objective criteria for minimum and maximum stream segment lengths.

3.1.4 Frame Development

In addition to identifying the target population and the sample units, an essential component of a survey design is to specify the frame. The frame consists of the list of sample units that make up a population, or a representation of that list, such as a map. Explicit list frames that identify each potential sample unit are often constructed. The National Lake Survey used list frames for various parts of the country, constructed by digitizing the location of lakes identified on maps (e.g., 1:250,000-scale USGS topographic maps for the Eastern Lake Survey; Linthurst et al. 1988).

Sometimes it is inefficient to construct a list frame because the population is so large. In these cases, an alternative is to create a map that represents the target population. Sites are selected from the map, with well-defined rules. This was the route chosen for the National Stream Survey (Messer et al. 1986). Maps represented the target population of stream segments in selected regions. A systematic rectangular grid of points, superimposed on the map, combined with a site selection rule, was used to select stream segments for characterization and subsampling. In this case, the map represented the target population, as a virtual list frame, without explicitly listing each sample unit.

3.1.4.1 Lake Frame

For lakes, it is feasible to create a list frame from which sample units can be randomly drawn. The USGS 1:100,000 series topographic maps contain an inventory of lakes that includes lakes < 1 ha across the conterminous United States. Digital versions of these maps are available (referred to as digital line graphs, or DLGs); they contain several layers, including one of hydrography (stream and lake traces). The advantages of using DLGs for lake selection include automatic processing with attendant decrease in human errors, knowledge of locational accuracy (USGS specifications), and lake areas included as part of the file.

We developed a method of extracting the lake coverages from this DLG series and creating a point coverage of all lakes represented in the DLG series, each with a unique identifier. This cover-

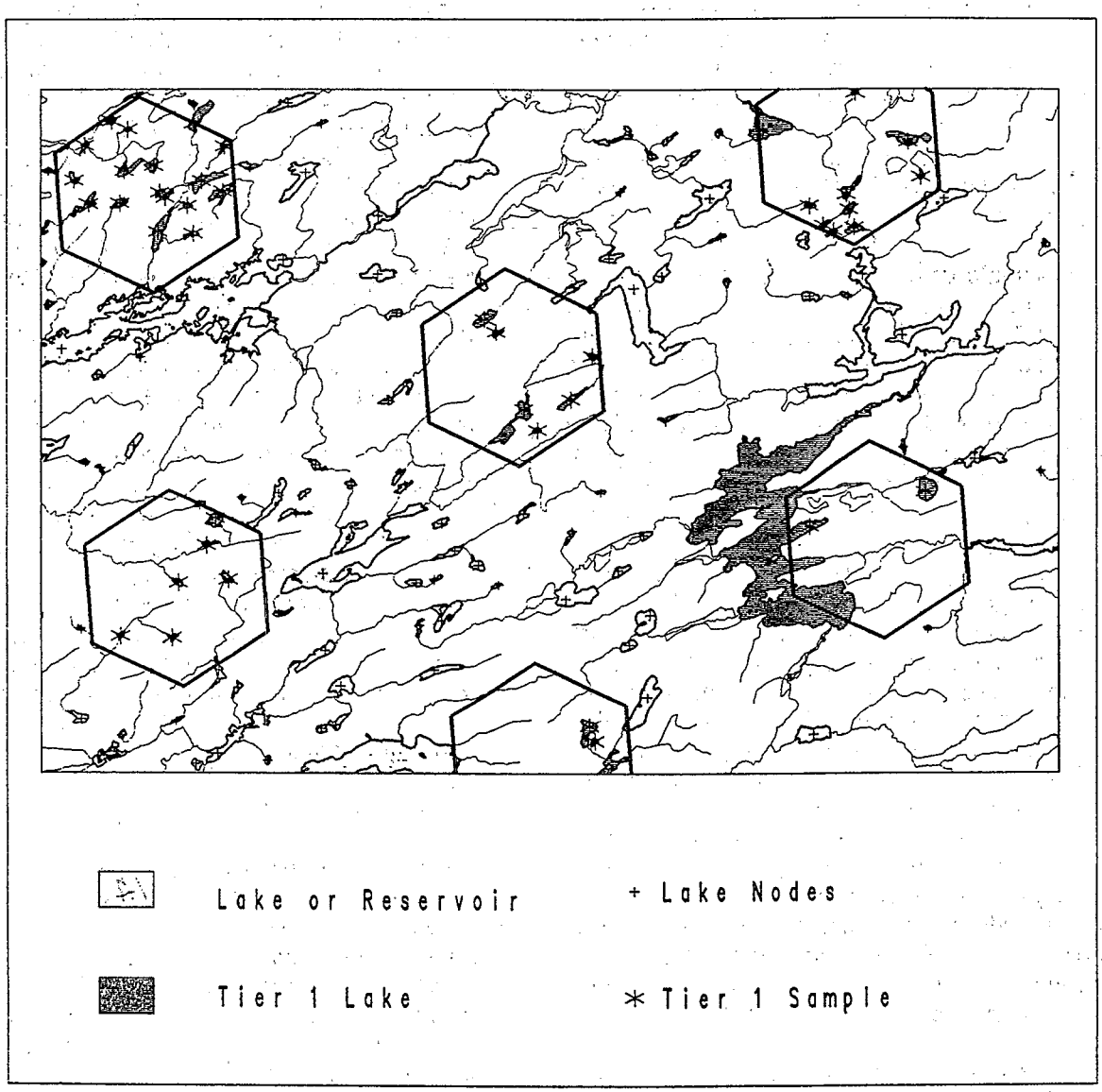


Figure 3-1. Example of the lake frame. Each lake is identified by a unique node. Lake nodes occurring within hexagons identify the Tier 1 sample.

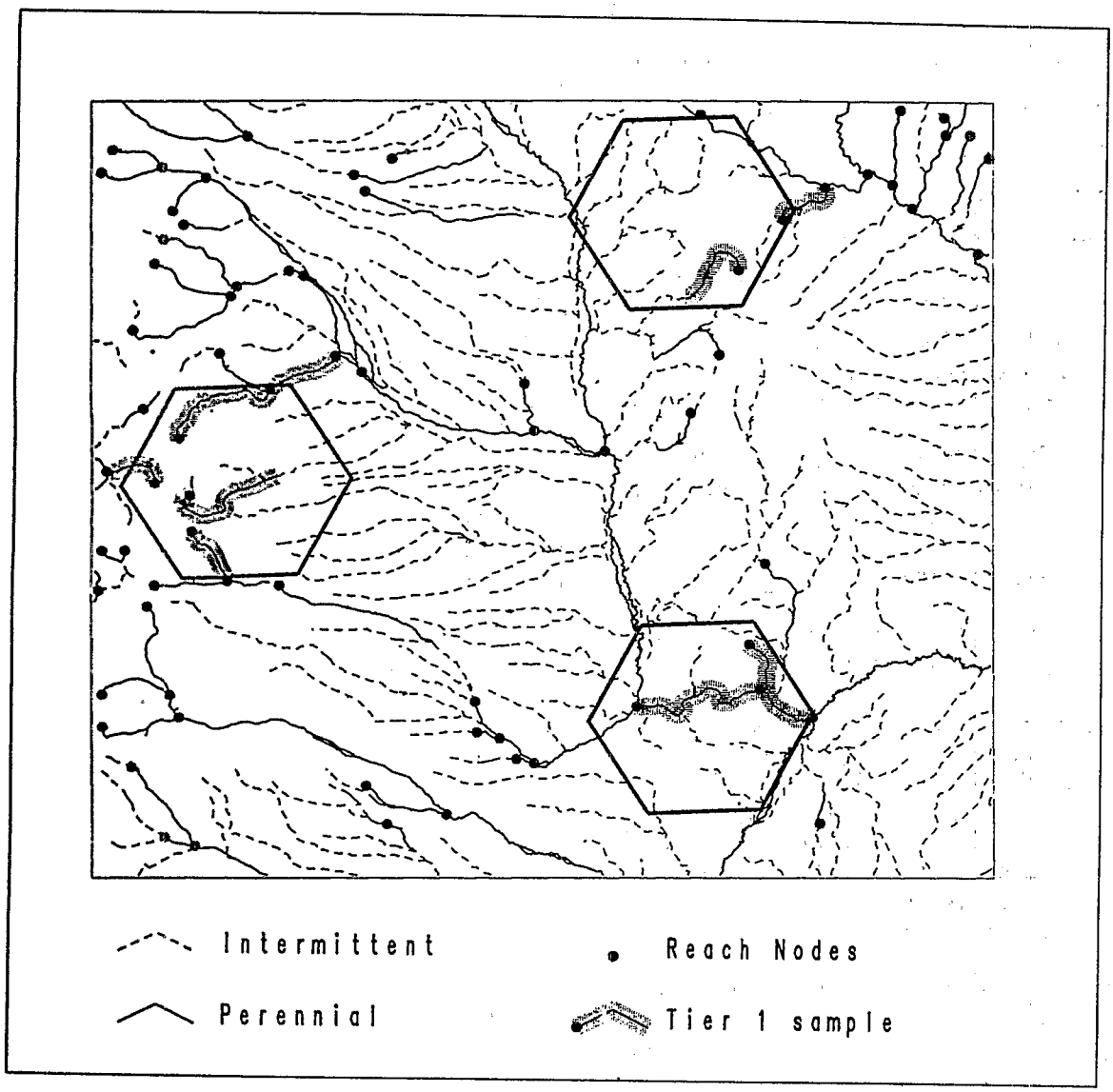


Figure 3-2. Example of the stream frame, with nodes identifying stream segments that are the stream sample units. Tier 1 stream segments, located by their unique nodes, are highlighted. Each node uniquely identifies a downstream segment.

Table 3-1. Estimates of Stream Reach Length Based on Segments Represented on Three 1:100,000-scale Topographic Maps

Map	Segment length (km)				
	Mean	Std. dev.	Max.	Min.	n
Tacoma, WA	2.4	2.2	10.1	0.1	50
Hendersonville, NC	1.2	1.2	7.2	0.1	158
Steamboat Springs, CO	2.5	2.0	9.5	0.5	46

age could be used as a list frame from which to draw the sample lakes; however, to be consistent with the EMAP grid design, we will superimpose the grid on this lake frame and select lakes occurring within the 40-km^2 hexagons using one of the association rules described in Section 3.2.1.3.

3.1.4.2 Stream Frame

We are less definite about the frame materials for streams. We will have to use maps as the frame for selecting sample units, because constructing list frames is too costly and is inconsistent with the use of EMAP's grid design. There are no compelling reasons to use a list frame for streams. Several options for the choice of maps are under consideration. Leading candidates are the USGS 1:100,000 and 1:250,000 topographic series. Site selection should occur from a relatively consistent set of maps across the region of interest.

To our knowledge, the USGS series of topographic maps at various scales is the only relatively consistent set of maps that could be used for selecting streams. Several map scales are available. The 1:24,000 (7 1/2 minute) series are too detailed for EMAP, and the manual labor involved in locating grid points and 40-km^2 hexagons on thousands of 7 1/2 minute maps would be overwhelming. These maps could be used to verify and characterize sites selected from the other maps. Other options are outlined in the following paragraphs.

3.1.4.2.1 USGS 1:100,000 Topographic Series Paper Maps and Digital Line Graphs

The paper maps are intended to cover the entire conterminous United States, but not all are available in topographic version. Some are available as planimetric maps. Stream and lake representation (detail) varies from place to place, but the coverage is more detailed (contains more streams and lakes) than that of the 1:250,000-scale maps. Disadvantages of using the paper maps include manual site selection, variation in hydrographic resolution from place to place (though less severe than on 1:250,000-scale maps), and the possibility of too much detail at the small stream size.

The digital line graph (DLG) version of these maps consists of several layers, including land ownership and political boundaries, transportation, and hydrology. The database is large, requiring massive storage capacity. About 1,800 1:100,000-scale maps cover the conterminous United States. The database contains 43 gigabytes, requiring 241 tapes for storage. The hydrography coverage alone requires five discpaks.

For streams, this digital database might be used, but with some modifications. We will have to address the massive storage requirements of this system.

3.1.4.2.2 U.S. EPA's River Reach File

The River Reach File (RRF) is essentially an enumeration of all lakes and river reaches at a 1:500,000 scale. It is being upgraded to incorporate the USGS 1:100,000 DLGs. Proto-

types have been completed; complete national conversion is underway, funded in stages by EPA regions and the Office of Water. The version based on the DLG is called RF3. Conversion of the DLGs into the River Reach File includes edge matching, so traces from one map to another can be tracked electronically. It may be feasible to conduct Tier 1 site selection automatically. One great advantage of a functional River Reach File is that the file is tied to several EPA databases that characterize the reaches in different ways. We are actively investigating the use of RF3 as a lakes and streams frame.

3.1.4.2.3 1:250,000-scale USGS Topographic Series Maps

These paper maps cover the entire conterminous United States. Stream and lake representation (detail) varies from place to place. Because these are available only as paper maps, their use would require manual site selection. The main disadvantages include manual site selection, variation in hydrographic resolution from place to place, and removal of detail so that this series does not contain as many of the smaller streams as the more detailed series. None of these disadvantages is insurmountable. The 1:250,000 series maps were used quite successfully for site selection for the Eastern Lake Survey (Landers et al. 1988) and National Stream Survey (Messer et al. 1986) conducted as part of the National Acidic Precipitation Assessment Program.

The choice of map representation will depend ultimately upon whether the maps represent the resource of interest adequately, the ease of use of the maps, and the relationship of the map databases to other databases. It would be ideal if the RF3 could be used as an EMAP frame, so evaluation of its use is of high priority.

3.2 THE TWO-STAGE SAMPLE: TIERS 1 AND 2

3.2.1 Tier 1 Resources

In EMAP, the first stage in the two-stage sampling process is called a Tier 1 sample, and a Tier 1 resource is a first-stage stratum. The function of designating a resource as a Tier 1 resource is to insure that an adequate sample size is generated to make population estimates. Thus resources that are of particular interest, or disinterest, or rare, or too spatially limited to get adequate representation may be designated as Tier 1 resources. Other examples of Tier 1 resources include any resource that has the potential for precluding the selection of others of greater or equal interest, or a resource that requires special treatment.

Tier 1 resources are strata on which Tier 2 samples are drawn in EMAP's two-stage sampling process. Initial designation of Tier 1 resources does not preclude future revision as necessary to focus on particular subpopulations inadequately represented during the initial phases.

Lakes and streams will constitute separate Tier 1 resources. For each, several size classes are under consideration as a further distinction of Tier 1 resources. The possible need to distinguish several size classes as Tier 1 resources is driven by the characteristic of the EMAP design that selects sites in proportion to their abundance. In fact, any nonstratified random selection process would select sample units in proportion to their abundance. Since there are substantially greater numbers of small lakes and streams than large, small systems could preclude the selection of larger systems. One way of balancing the selection is to identify size classes as separate Tier 1 resources and allocate samples according to desired sample sizes for each stratum (see Section 3.2.2.1).

Also, the grid design is inefficient for selecting the relatively large, unique systems. It is preferable to list these systems separately and to sample them from a list frame, or to census them. This special interest group will contain lakes larger than 2,000 ha (half the size of the 40-km^2 hexagons) and the largest rivers in the United States. We have not identified cutoff sizes yet; the subject is under discussion.

3.2.1.1 Lakes

The distinction between lakes and reservoirs is a conventional way to stratify; however, lakes and reservoirs will be considered together at the Tier 1 resource level. Reservoirs can be described as a separate subpopulation, if necessary, for finer scale descriptions. There are several reasons for combining lakes and reservoirs at the Tier 1 level:

- In many areas, lakes and reservoirs function similarly, and display similar kinds of

problems, so distinguishing between them provides little additional insight.

- Similar indicators of condition will be used for lakes and reservoirs.

- In some instances natural lakes are artificially controlled, making them a transition between lakes and reservoirs.

- It is often difficult to distinguish between lakes and reservoirs on the basis of mapped information, particularly for the smaller impoundments.

- There tends to be an inverse relationship between the occurrence of reservoirs and natural lakes; in lake-rich regions, relatively few reservoirs occur; in lake-poor regions, a greater density of reservoirs occurs. Site selection based on the grid design should reflect the relative distribution and abundance of lakes and reservoirs. Sampling density can be modified as necessary by increasing or decreasing the density of the basic grid (Overton et al. 1990). However, see Thornton et al. (1990) for discussion of differences between lakes and reservoirs.

In any case, the natural lake/reservoir status of any site selected for field visit will be recorded; data summaries can be created for each group and the condition of each subpopulation can be described. If sufficient interest or justification develops to consider reservoirs as a separate Tier 1 resource, the flexibility of the design will accommodate this.

As noted earlier, lake size is being considered for allocation of lakes into different Tier 1 resources. The largest lakes (e.g., those > 2,000 ha) constitute a separate group, because they are infrequently captured by the grid and increasing the grid density is an inefficient way to select sparsely distributed lakes. They can be identified and listed as part of the Tier 1 sample selection process and will constitute a separate Tier 1 resource.

The remainder of the lakes can be divided into size classes, primarily as a device to assure adequate selection of larger lakes. Each size class will constitute a separate Tier 1 resource. A recently completed summary of the size distribution of lakes in EPA Regions 1 and 2 (Northeast) indicates that of the ~18,000 lakes > 1 ha, 70% are < 10 ha, supporting the need to establish size criteria as Tier 1 resource delineators (Figure 3-3 and Table 3-2).

As already mentioned, lake size is used to identify separate Tier 1 resources to prevent overwhelming the sample selection with small lakes. Stratification by size classes permits the selection of lakes such that inclusion probabilities within strata are equal, a desirable property for minimizing the variance of some kinds of population estimation for that stratum and for simplifying the calculation of population statistics. See Section 5 for a discussion of inclusion probabilities and weights.

It is likely that stratification will produce vastly different inclusion probabilities among strata, so we will face the need to combine sample units with different inclusion probabilities if, as is likely, subpopulations are developed across the size strata. Hence, it may be more desirable to develop other methods to balance the selection of large and small lakes, in which case all lakes will be considered as a single stratum.

3.2.1.2 Streams

In order to represent the broad range of stream sizes, from headwater mountain creeks to rivers as large as the Mississippi or Columbia, it is necessary to consider several possible Tier 1 resources based on size. The discussion begins with a description of the two ends of the size spectrum, then moves on to the intermediate sizes. Five size classes of streams are suggested, from Class I (smallest) through Class V (largest); each class could constitute a separate Tier 1 resource. At the present time, only very general size criteria are proposed. There is no need to make explicit decisions about streams yet, so we present the broad issues, with some discussion about how these issues are being resolved. Part of our activity over the next year will be to refine our definitions of stream size classes and our approaches for considering the broad spectrum of stream sizes.

The characteristic of "size" for streams has no single measure that is satisfactory for all purposes. The closest measure of stream size is probably a measure of discharge, which might be average annual discharge, low flow discharge, or some other expression of the hydrograph. A disadvantage of using stream flow is the general lack of information on stream flow at the Tier 1 level.

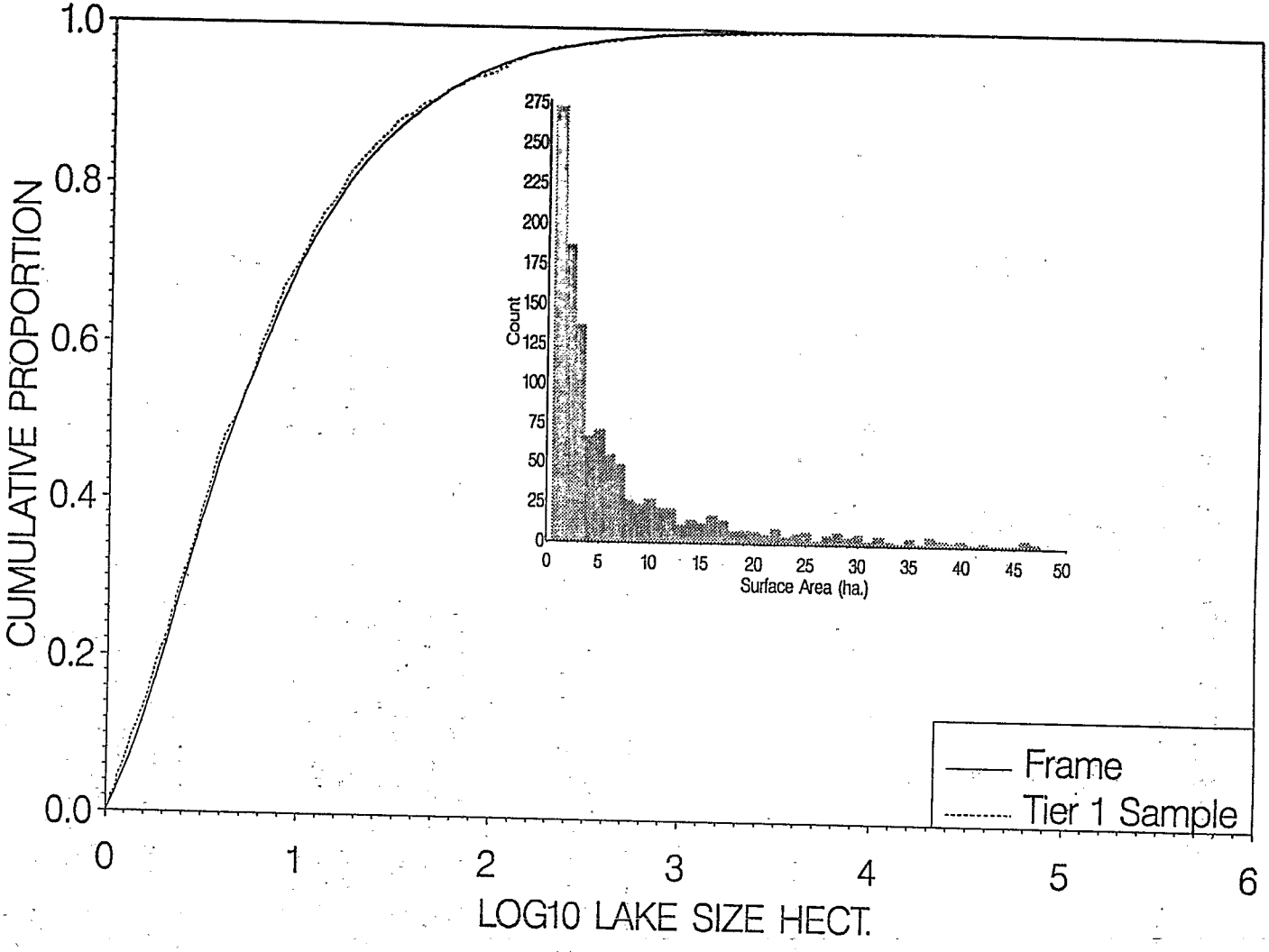


Figure 3-3. Cumulative size distribution for lakes in the Northeast. The insert is a histogram of the size distribution for lakes with areas < 50 ha.

Table 3-2. Summary of the Size Distribution of Lakes in the Northeast (EPA Regions 1 and 2)[a]

Part A

Size class (ha)	Tier 1 Sample		DLG Inventory	
	Number of lakes	Total area (ha)	Number of lakes	Total area (ha)
1-5	662	1,600	10,745	27,200
5-10	219	1,500	3,565	25,300
10-15	94	1,150	1,476	18,000
15-20	58	1,000	917	15,800
20-100	143	6,000	2,665	115,500
100-250	41	5,900	548	82,500

Part B

Size class (ha)	DLG Inventory	
	Number of lakes	Total area (ha)
250-500	227	78,300
500-1,000	112	77,900
1,000-2,000	62	84,600
> 2,000	59	442,300

[a] Part A describes the Tier 1 sample and the DLG inventory for lakes ≤ 250 ha in surface area. Part B describes the DLG inventory for lakes > 250 ha. The DLG database hydrography is described in the text.

Other measures of size or surrogates for size include watershed area (calibrated regionally), stream length, stream order, channel dimensions, or wetted usable area. The last measure pertains to the area of the stream channel that aquatic organisms use as habitat, usually measured during a low flow period.

3.2.1.2.1 Class I

These are the smallest streams of potential interest; they are generally represented on 1:100,000- or 1:250,000-scale maps as first or second order (Strahler) perennial or intermittent streams and have a high potential for being dry or intermittent during the proposed EMAP sampling/index period. They can contain subsurface flow or enduring pools during parts of the year. In much of the arid and semi-arid West, these streams generally do not contain flowing water, yet the stream channels function as an important ecological resource. Therefore for this class, the proposed resource unit is a channel segment rather than a stream segment. Three general types of channels can be described:

- Dry: These channels function primarily to discharge water during episodic storm events; they retain an insufficient amount of surface or subsurface water to support hydrophytic vegetation in the channel or riparian zone; the riparian zone is inhabited by upland vegetation up to the stream bank and sometimes in the stream channel.

- Generally dry: These channels often contain no flowing water, but the channel (including bank) substrate contains enough water to support hydrophytic vegetation in the riparian zone. Particularly in the west, these ribbons of green in an otherwise dry landscape act as corridors linking different landscape types and are sites for locally high biodiversity. They generally do not contain traditional stream communities of

macroinvertebrates and fish, except those with enduring pools. Nonetheless, they are an important aquatic ecosystem.

- Perennial flow: These channels generally contain flowing water and regionally typical stream fish and macroinvertebrate communities, as well as hydrophytic riparian vegetation. They are the conventional streams that we generally discuss.

The above distinctions might not be so important in the humid regions of the country, but for vast areas of the West, where flowing water is limited, and where the flowing water that does exist is often diverted from the stream channels, it is important to focus on the channel itself. One possible indicator of the gain or loss of stream resources in arid regions might be the proportion of systems that change from one of the above classes to another. These shifts are apparent during wet/dry climatic cycles.

The EMAP grid is efficient at selecting these stream segments because their proportion and density is high. Our expectation is that on the average, three to six first and second order segments will be captured per hexagon in humid regions of the country, but on the average, only one segment per two hexagons can be expected in dry regions. The EMAP grid density appears capable of selecting a sufficient number of these segments for EMAP estimation.

3.2.1.2.2 Class V

These are the largest rivers; they stand out by their uniqueness and should receive special attention. Examples are the Columbia, the Mississippi, and the Missouri. The USGS has summarized the largest rivers in the United States based on three criteria of size: discharge at the mouth, watershed area, and length. It might be necessary to treat these systems apart from the other EMAP systems because each one could be tracked and described separately, perhaps one-fourth of them each four years.

3.2.1.2.3 Class IV

We usually call this group of streams rivers, but they are smaller than the Class V rivers. They are sparsely distributed relative to the EMAP grid, thus they are infrequently captured. Creating an adequate sample could pose a problem. It can also be argued that these rivers should be treated as an extensive resource, included with the Class V rivers as a single Tier 1 resource to be sampled by probability methods other than the grid.

One option for identifying rivers that fall into this class is based on an estimation of the stream density necessary to meet a minimum sample size. Lower stream densities would result in sample sizes too low for adequate estimation of condition. An alternative is to enumerate and map this class of rivers based on criteria similar to those used by the USGS to identify the largest rivers in the nation. An inventory could be conducted for each USGS hydrographic basin.

For Class IV and V rivers, a stream's length, from a point well below its headwaters to its mouth, would be considered the resource unit. The point at which a Class V or IV river begins could be located such that the watershed area above it is similar in size to stream watersheds in the region selected from the EMAP grid. Thus the river would not be sampled up to its headwaters, but to some point substantially downstream. A list of these rivers and a map identifying the origin and terminus of each river resource unit can be created.

3.2.1.2.4 Classes II and III

These streams fall between the small Class I streams and the Class IV rivers. They are approximated by third to fifth order on USGS 1:100,000- and 1:250,000-scale maps. It might be reasonable to treat this group of streams as a single class. We expect that an adequate sample size can be obtained in humid regions via the EMAP grid where, on the average, one to two segments per hexagon occur. In dry regions, the capture rate is about 1 segment in 10 hexagons, so it may be necessary to increase the grid density to acquire a large enough sample.

3.2.1.2.5 Other

Artificially created diversions, canals, and other waterways that do not follow the natural drainage pattern of streams could be important resources in some areas where most of the flow of water is through these artificial systems. These areas and systems will be considered on a case-by-case basis as stream implementation proceeds.

3.2.1.3 Tier 1 Association Rules

Association rules guide the selection of sites from the frame. They are required because the systematic grid points of the EMAP design do not fall on lake or stream nodes. Two options have been considered for lake and stream sample selection. One is to choose the nearest lake or stream, via its unique node, to the grid point within each 40-km^2 hexagon. The second is to create a list of lakes or streams that occur within the hexagon and then select one of these for each Tier 1 resource, via a probability sample.

Given that the two options for selecting the Tier 2 sample do not result in appreciable differences in precision, there are several reasons for preferring the list approach over the nearest lake or stream approach. First, the list approach conforms to the standard sampling techniques of double sample and cluster sample, and thus has well known statistical properties. Second, the list approach is easier to implement. Finally, and perhaps most important, the full set of sample units in each hexagon would be carried into the second stage of the two-stage selection process, and so will be available for modification of the sample at a future time.

3.2.1.4 Adequacy of the Grid for Tier 1 Sample Selection

To gain familiarity with the grid design and to evaluate how many lakes and stream segments fall within the 40-km^2 hexagons, we sampled 59 1:100,000-scale USGS topographic maps in different parts of the country. We used a mylar overlay, with the 40-km^2 hexagons outlined at the Tier 1 density, and counted lakes and stream segments that fell within the hexagons. Table 3-3 summarizes the capture rates and the frequencies with which no lakes or streams appeared within the hexagons. For this study we selected only nodes associated with perennial stream traces without distinguishing size classes.

Although the table summarizes only our results for broad geographic regions, we also evaluated the capture rates at the level of Omernik's aggregate ecoregions (Figure 2-2). For example, in the Middle Atlantic Coastal Plain ecoregion (2 maps, 18 hexagons), an average of 10.4 streams were captured per hexagon and all hexagons captured streams. At the other extreme, in the Subhumid Agricultural Plains ecoregion, Northern Section (4 maps, 36 hexagons), no streams were captured (on average) and 97% of the hexagons were without stream segments. In areas where stream density is fairly high, from 2 to 7 stream segments were captured per hexagon, with approximately 20% of the hexagons without stream segments. In areas where stream density is low, from 0 to 0.4 stream segments were captured per hexagon, with more than 75% of the hexagons without segments.

For lakes, regional variation was less severe. On the average, one or two lakes occurred per hexagon, regardless of the region sampled, except for the driest regions. In some cases, more than 10 lakes were present in a hexagon. About 50% of the hexagons had no lakes, but with regional variation. The lowest no capture rate (about 30%) was in the glaciated regions of the Upper Midwest and the Northeast. The highest no capture rates occurred in the xeric regions (60-90%). Reservoirs were sparsely distributed. The no capture rate was about 80-90%.

We concluded that for small systems, the capture rate seems adequate. In some areas where lake or stream density is low, it will be necessary to increase the grid density, if it is deemed important to make subpopulation estimates in these areas. Because of the flexibility of the grid design, changes in density can be accomplished as particular regions are implemented for sampling.

3.2.2 Tier 2 Sampling

Physical, chemical, and biological measurements (Section 4) made at Tier 2 sites, plus landscape descriptions, will complete the information needed to estimate condition and to classify subpopulations. Measurements to be taken at Tier 2 sites are covered in Section 4. This section describes the process by which Tier 2 sites are selected from the candidate Tier 1 sites. Where to sample within Tier 2 sites (index sites) and when to sample (index periods) are addressed in Section 6.

A Tier 2 subsample will be drawn from the collection of sites comprising the Tier 1 sample. The size of the Tier 2 sample will be determined through a combination of the following criteria:

- Importance of estimating resource status in a particular region, relative to the availability of funds to conduct the field work.

Table 3-3. **Summary of Average Capture Rate of Perennial Stream Segments, Lakes, and Reservoirs (Sample Units per Hexagon) in 40-km^2 Hexagons for the Humid (Eastern) and Arid (Western) Regions of the United States[a]**

Region	Perennial stream segments	Lakes	Reservoirs
Humid/Subhumid average (36 maps)	5.2	1.7	0.3
% no capture	19%	48%	85%
Arid/Semi-arid average (23 maps)	0.3	1.1	0.4
% no capture	84%	63%	75%

[a] Data taken from 59 USGS 1:100,000-scale topographic (T) or planimetric (P) maps; nine hexagons per map were sampled.

- Expected variation of lake or stream types within the region.
- Expected variation of indicator values within the region.
- Quantity of the resource in particular regions; if many lakes or streams occur in a region, it may useful to increase the sampling effort to characterize the many possible subpopulations.

An operational guideline is that there should be 50 to 100 Tier 2 sites per subpopulation over the four-year cycle in order to make adequate subpopulation estimates. Larger samples do not increase precision enough to warrant the increased costs (Linthurst et al. 1986).

3.2.2.1 Tier 2 Site Selection Example

EMAP-Surface Waters plans to conduct a pilot survey of lakes in the northeastern United States (covering EPA Regions 1 and 2). The following example illustrates the site selection process as it is being applied to this pilot project.

The design report (Overton et al. 1990) describes two options for selecting Tier 2 sites from the Tier 1 candidate sites. Discussed with reference to lakes, the options are:

1. OPTION 1: Within each 40-km^2 hexagon that contains one or more lakes, select a lake using one of the rules of association; this reduced Tier 1 sample would be subsampled to yield the Tier 2 sample.
2. OPTION 2: Select a subset of the 40-km^2 hexagons, then select a lake from each of the selected 40-km^2 hexagons using one of the association rules.

Option 2 has advantages for lake and stream site selection and will be used to illustrate the Tier 2 site selection process for EPA Regions 1 and 2.

USGS 1:100,000 Digital Line Graphs were used in the lake Tier 1/Tier 2 site selection process. These digital coverages contain almost all lakes represented on USGS 1:100,000-scale topographic maps; lakes < 1 ha are included as part of this series. A map of the spatial distribution (Figure 3-4) and the size distribution by area (Table 3-2 and Figure 3-3) illustrate the nature of this data base. A great majority of the lakes are < 20 ha, creating a situation in which small lakes would dominate any sample unless strata (Tier 1

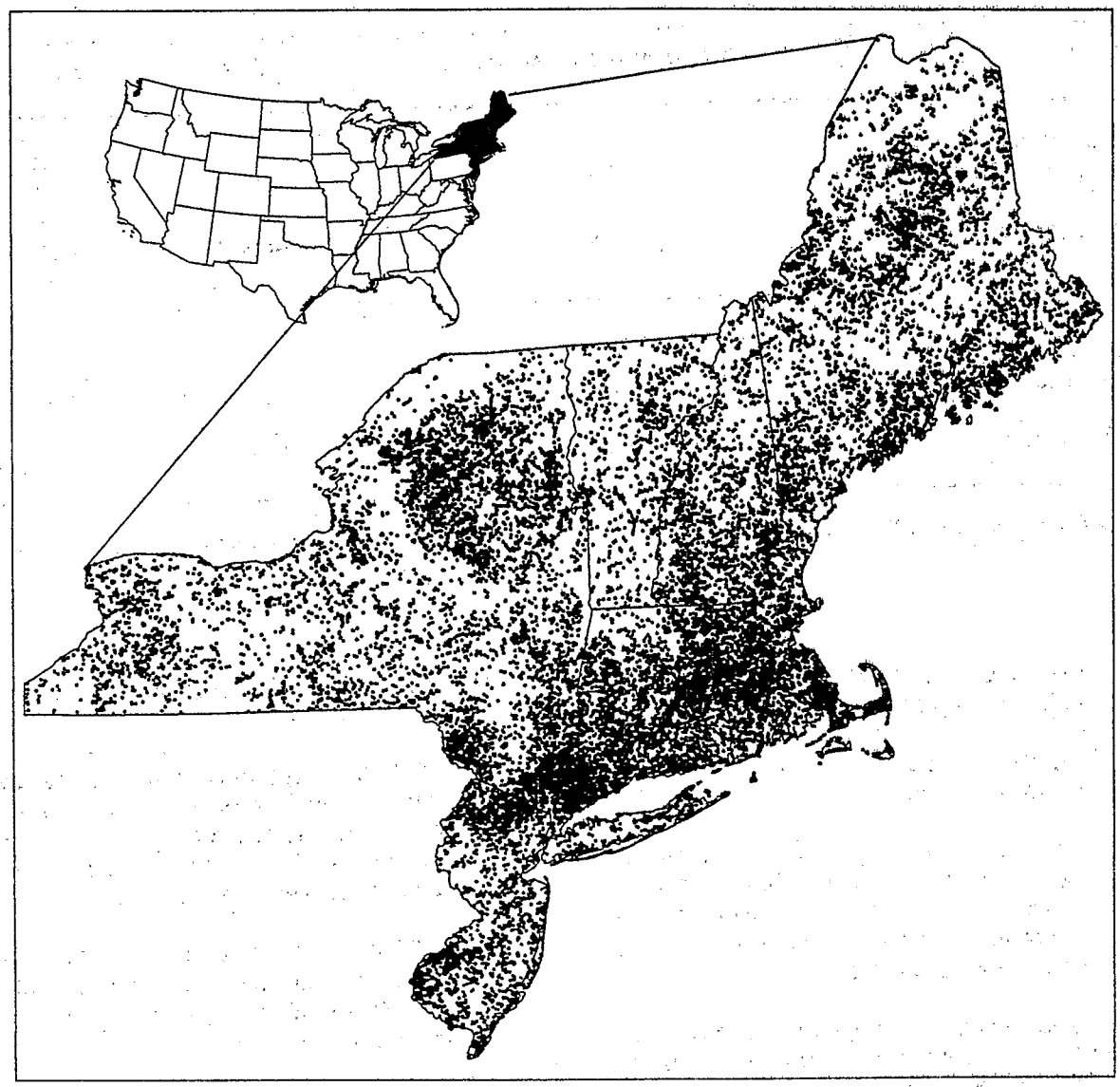


Figure 3-4. Map of the spatial distribution of lakes in the Northeast. Size class represented is 1-250 ha in area.

resource designation) based on size were constructed.

The Northeast accounts for only 4.4% of the conterminous United States. If samples were allocated in proportion to area of a region, ~35 samples per year would be allocated to the Northeast, based on the assumption that we can afford to sample 800 lakes (Tier 2) nationally when EMAP-Surface Waters is fully operational. Because of the greater density of lakes in some regions of the country, it might be appropriate to allocate a disproportionate amount of the sampling effort to these areas, to make refined estimates of condition of subpopulations. For the pilot, we estimate we can afford to sample ~100 lakes; this number will be used for planning.

We have tentatively selected three Tier 1 resource groups based on lake size (area): small (1-20 ha), large (20-2,000 ha), and off grid (> 2,000 ha). We are continuing to discuss the boundary between small and large lakes; it will lie in the range of 10 to 20 ha. This illustration uses the Tier 1 resource "small lakes".

The following process describes one way to select a Tier 2 sample that meets EMAP criteria for population estimation. We are currently evaluating a second process and will use the pilot of FY91 to evaluate the two. We use a Tier 2 sample of 40 each for the small and large lakes, with 20 for special purposes (e.g., lakes > 2,000 ha or reference lakes). Tier 1 sample sizes are: 910 lakes > 1 ha and < 20 ha, and 188 lakes > 20 and < 2,000 ha. There are 49 lakes > 2,000 ha in the Northeast. Over a four-year cycle, 160 Tier 2 samples would be selected for small and large lakes. At the present time, we think it is important to oversample, understanding that we can cut back the sample size in future years without interfering with the estimation process significantly. Nearly all the Tier 1 sample of large lakes would be selected. This could argue for grid augmentation for this size group.

3.2.2.2 Mechanics of Tier 2 Site Selection (Design Option 2)

Options 1 and 2 are both two-stage sampling designs, but Option 2 follows the standard two-stage format more closely than Option 1. Either option can be used with any of the association rules, and Option 2 with the "mini-list random selection" is exactly the standard two-stage design format. Option 2 adds some flexibility in resampling if the initial selection results in nontarget lakes.

The steps in the Tier 1/Tier 2 site selection process are:

- Intersect the complete array of 40-km^2 hexagons with the lake coverage to extract those lakes contained within hexagons in the region of interest. This Tier 1 database contains numbers of lakes per hexagon and lake areas, so identifying small, large, and off-grid lakes per hexagon is straightforward.

- Map the location of each hexagon with its associated number of lakes (including 0) in each Tier 1 resource category, for example, a map of the small (1 to <20 ha) lake captures.

- Select a subset of hexagons with the following constraints:

  - Probability of selection is proportional to the number of lakes contained within each hexagon.

  - Spatial distribution of lakes is adequately preserved.

  - One lake is chosen randomly from each of the selected hexagons.

The result is a set of Tier 2 lakes whose weights (or inclusion probabilities) are approximately equal, and whose spatial distribution represents the spatial distribution of the universe of lakes.

The procedure for meeting these constraints is to delineate areas containing approximately equal total number of lakes. The areas should be spatially compact. The process by which the delineation is accomplished is statistically irrelevant; it can be done subjectively, arbitrarily, or by some objective protocol. There may be some advantage to a subjective grouping using some criteria of similarity (e.g., land use, ecological subregion) other than proximity.

The number of lakes used to define these areas is determined from the desired Tier 2 sample size and a constraint that two hexagons be chosen

within each area, on the average. The number of lakes is calculated as:

$$\frac{2 * (\text{Tier 1 sample size})}{(\text{Tier 2 sample size})}$$

Hexagons are selected within each of the areas such that the probability of hexagon selection is proportional to the number of lakes in the hexagon. Then a lake is selected randomly from each chosen hexagon.

A specific example of these steps is illustrated for small lakes in the sequence of panels of Figure 3-5, with a Tier 1 sample size of 1,217 and a Tier 2 sample size of 80. Figure 3-5a illustrates the hexagon capture rate for this size class. Areas are delineated such that each area contains approximately 30 lakes: 2 * (1,217/80), as in Figure 3-5b. These areas were arbitrarily selected, subject to the constraint that each area is compact and that approximately 30 lakes are contained in each area.

For this exercise, we assumed that each hexagon with more than eight lakes was counted as if eight lakes were present. There has been some discussion about the need to limit the number of sample units per hexagon; the present illustration assumes a particular value. Upon conclusion of the discussions, the actual site selection will follow specific agreed-upon rules.

Hexagons were chosen from each of these areas, generally two hexagons per area with probability proportional to number of lakes per hexagon. This was done by conceptually stringing the hexagons along a line, each hexagon represented by a line segment proportional to the number of lakes within the hexagon. All hexagons within each area are strung together sequentially, then the next area is joined. With a random start, hexagons were selected systematically by a pointer that landed on the line every 15.2 units. This pointer distance is the ratio of the Tier 1:Tier 2 sample size. The result of this process is the selection of hexagons in proportion to the number of lakes they contain; hexagons with a greater number of lakes have a higher probability of selection.

Then one lake is selected randomly per selected hexagon. The result of these steps constitutes the Tier 2 sample (four year). The year 1 sample is obtained by choosing selected lakes in year 1 hexagons; year 1 hexagons are identified as part of the interpenetrating structure of the EMAP grid (e.g., Figure 3-5c). Locations of Tier 2 hexagons are illustrated in Figure 3-5d.

This general process will be used for lakes and streams throughout the country, as the frame materials are developed; the outcome will be the list of Tier 2 lakes and streams selected for field visitation. This Tier 2 subsample meets the EMAP requirements of random selection and spatial representation.

3.3 HIGHER GRID DENSITIES

Increasing the grid density is the preferred method for increasing the sample size as needed for any subpopulation. For lakes and streams it appears that local high-interest areas may require more spatially intense sampling than is provided for by the standard grid density. These areas will be defined as the need arises; they are likely to occur in arid regions of the country, or in areas where a particular subpopulation of high interest is sparsely distributed.

It will be necessary to use an iterative process to select the Tier 2 sample. The first step is to define the desired Tier 2 sample size, then determine whether an adequate number of sites is available from the grid. If not, the grid density will be increased until the desired sample size is achieved. The flexibility of the design makes this feasible.

A specific case of the need for higher grid density sampling occurs in the design of the TIME portion of EMAP-Surface Waters. The main emphasis in TIME is the detection of trends, rather than descriptions of current status; good descriptions of the acid sensitive status of current populations of lakes and streams are already available from the NSWS surveys (Linthurst et al. 1986, Landers et al. 1987, Kaufmann et al. 1988). In order to detect regional trends in a policy-relevant time span (as dictated by the Clean Air Act, this time span is about 10 years), and because of the variability inherent in acid-sensitive lakes and streams, we determined that stratification of sites will be necessary. Stratifying sites into subpopulations narrows the various components of variance that affect our ability to detect trends. Simulations of trends suggest that approximately 15 sites will be needed per subpopulation, in order to meet the requirements for trend

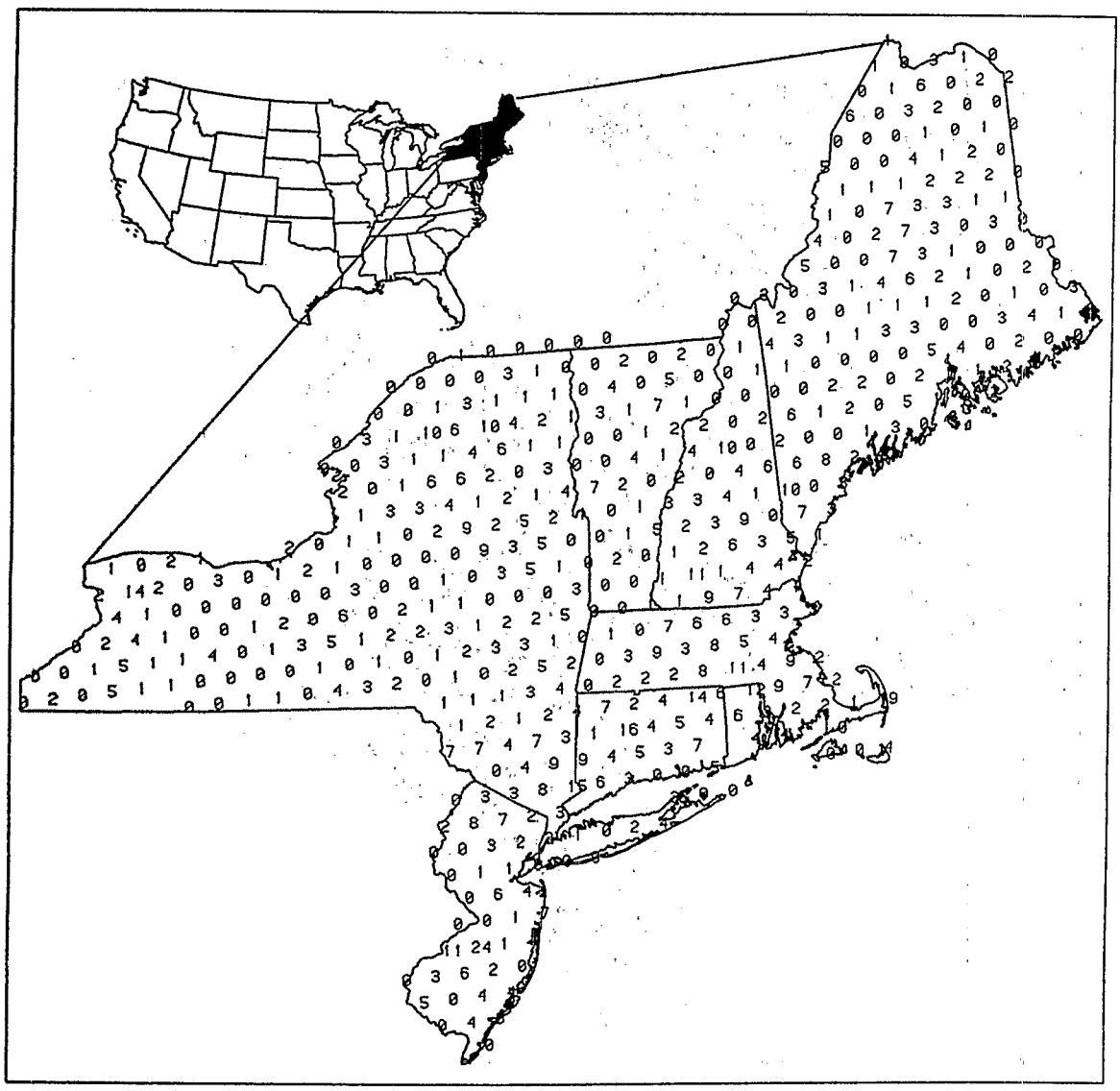


Figure 3-5a. Hexagon capture rates for lakes (1-250 ha) in the Northeast.

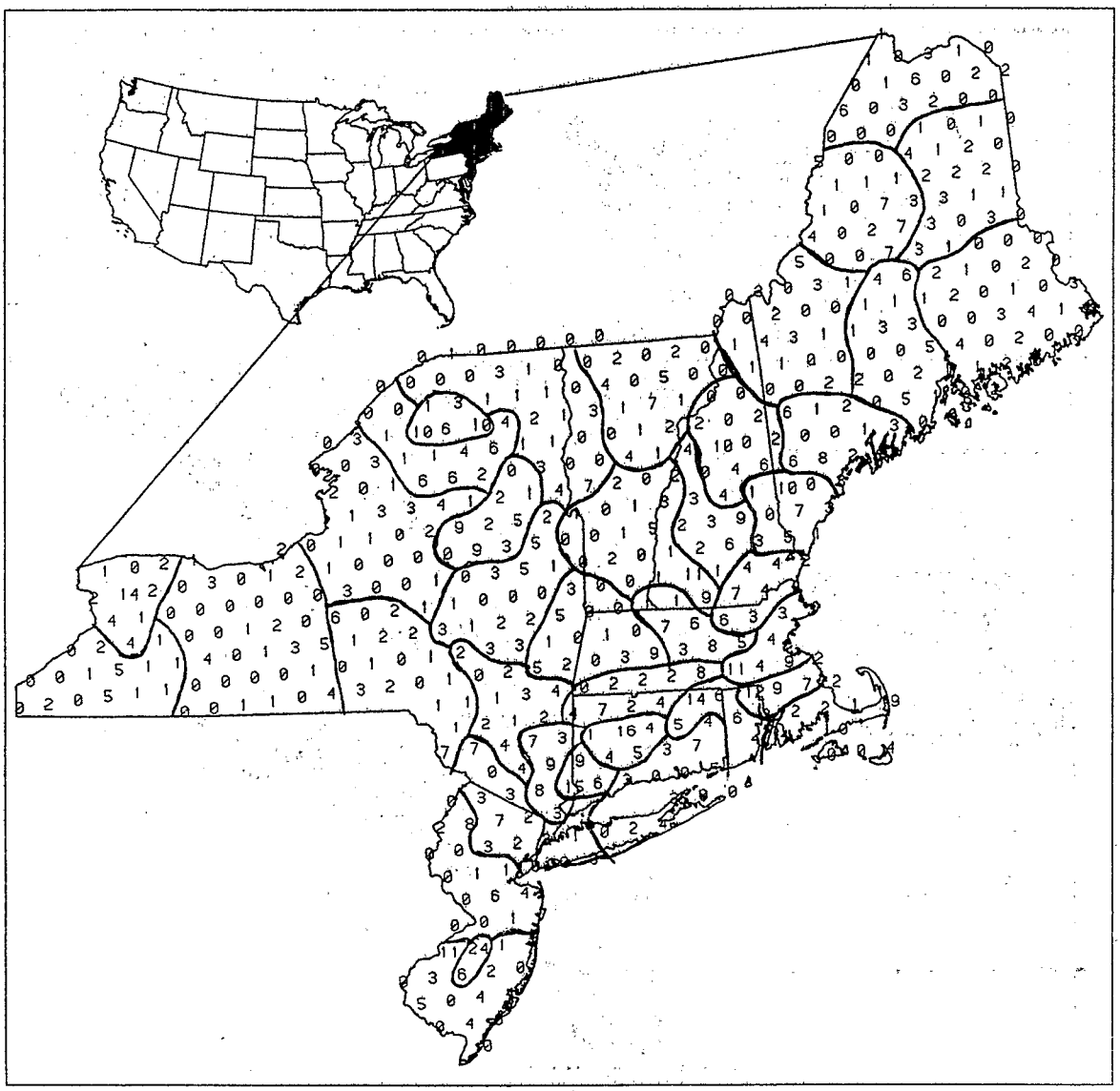


Figure 3-5b. **Areas that group similar numbers of lakes for Tier 2 selection. Number of lakes per area is 2 * Tier 1 sample size/Tier 2 sample size.**

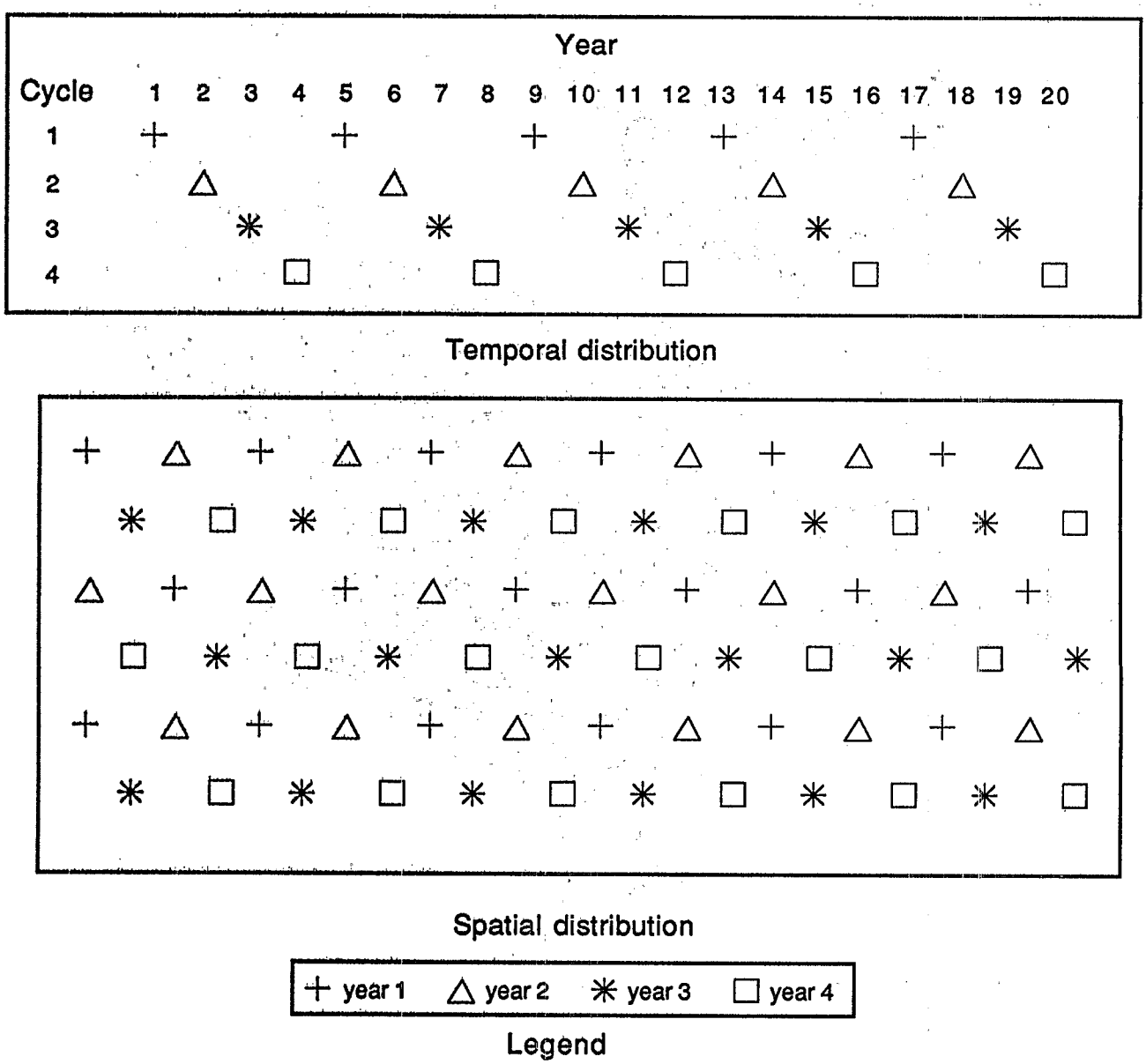


Figure 3-5c. **The structure of the EMAP grid, illustrating the fourfold decomposition that will be the basis of the interpenetrating sample design.**

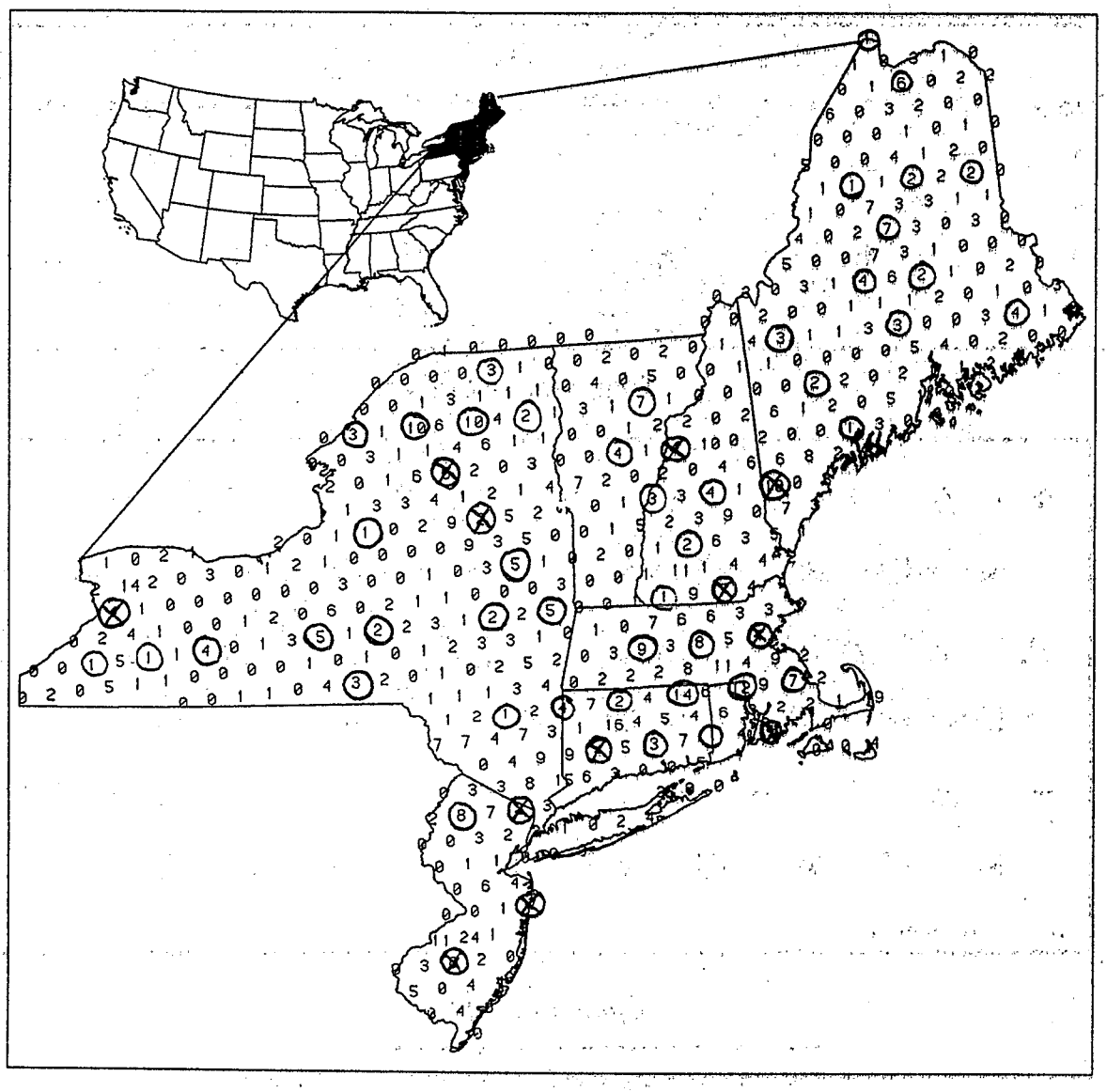


Figure 3-5d. **Selected hexagons containing one or more lakes that will constitute the Tier 2 sample. One lake per hexagon will be selected randomly within the hexagons, except for those marked with an "x," from which two lakes will be sampled.**

magnitude and length of time to detect a trend (see TIME conceptual plan; Stoddard 1990).

Based on information gathered in the Eastern Lake Survey, we estimate that approximately 10 subpopulations will be necessary to adequately describe the types of lakes found in the Northeast, with 15 samples per year required. A total of 150 Tier 2 lakes will therefore be necessary to detect trends in the region.

Figure 3-6 indicates the approximate geographic regions that are sensitive to acidic deposition in the Northeast (Stoddard 1990). Table 3-4 displays the number of Tier 1 hexagons and lakes within these regions. These figures indicate that the total number of lakes is sufficient in Tier 1 to supply the needed coverage for TIME. Samples for TIME, however, need to be evenly spread over the subpopulations of interest, many of which are not clearly mapable at this point.

Estimates of the coverage of subpopulations of interest was generated based on frequency of occurrence of these subpopulations in the Adirondacks and extrapolated to the Northeast. Table 3-5 presents these estimates. The results suggest that the final Tier 2 density will be sufficient for the coastal area only and that Tier 1 density will be acceptable for the subpopulation of drainage lakes with moderate base cations in Vermont and New Hampshire. A threefold augmentation of the grid will be necessary to select the required density of sites in the regions where the subpopulations are most likely to be found (i.e., the Adirondacks, and the high-elevation portions of Vermont and New Hampshire, Figure 3-7).

3.4 SUMMARY

This section describes how the general EMAP design is applicable to the Surface Water component of EMAP. It explains how the populations of interest are specified and how the sample units for lakes and streams are selected. It describes frame development, rules for site selection, and the Tier 1/Tier 2 site selection process. As part of the discussion, several pilot studies are presented to illustrate how the EMAP design will be used for surface waters.

Our conclusion is that in general the grid design will be adequate for selecting much of the surface water resource, but there are some systems (the larger, rare lakes and streams) and some regions where off-grid site selection will be necessary.

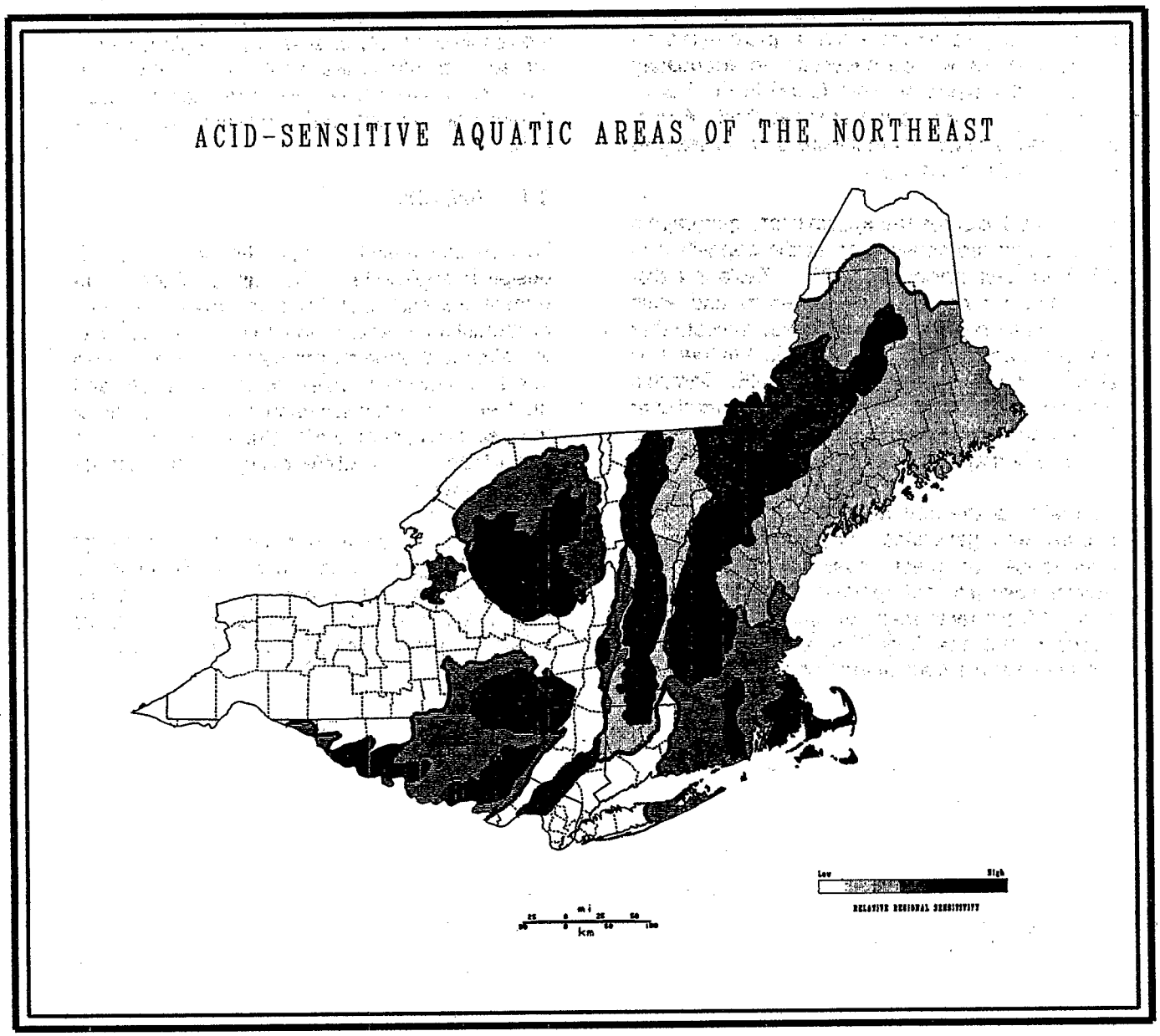


Figure 3-6. Acid sensitive regions of interest in the Northeast.

Table 3-4. Estimates of Geographic Regions of Interest in the Northeast Relative to Acidic Deposition and Number of Tier 1 Hexagons, Number of Lakes, and Number of Hexagons with Lakes

Subregion	Number of Hexagons	Number of Lakes	Number of Hexagons with Lakes
Adirondacks	41	131	36
Southwest Adirondacks	13	38	11
Northeast Coastal Zone	94	204	66
High interest areas in Vermont and New Hampshire	20	73	18
Mid-Northeast	153	509	152

Table 3-5. Estimates of Tier 1 Coverage of Subpopulations of Interest for Acidic Deposition in the Northeast

Subpopulation	Number of Tier 1 Grid Sites
Coastal New England	66
Low base cation, low DOC drainage lakes in the Adirondacks	8
Low base cation, high DOC drainage lakes in the Adirondacks	5
Moderate base cation, low DOC drainage lakes in the Adirondacks	2
Moderate base cation, high DOC drainage lakes in the Adirondacks	3
Mounded, low DOC seepage lakes	2
Mounded, high DOC seepage lakes	2
Low base cation, low DOC drainage lakes in Vermont and New Hampshire	3
Low base cation, high DOC drainage lakes in Vermont and New Hampshire	3
Moderate base cation drainage lakes in Vermont and New Hampshire	38

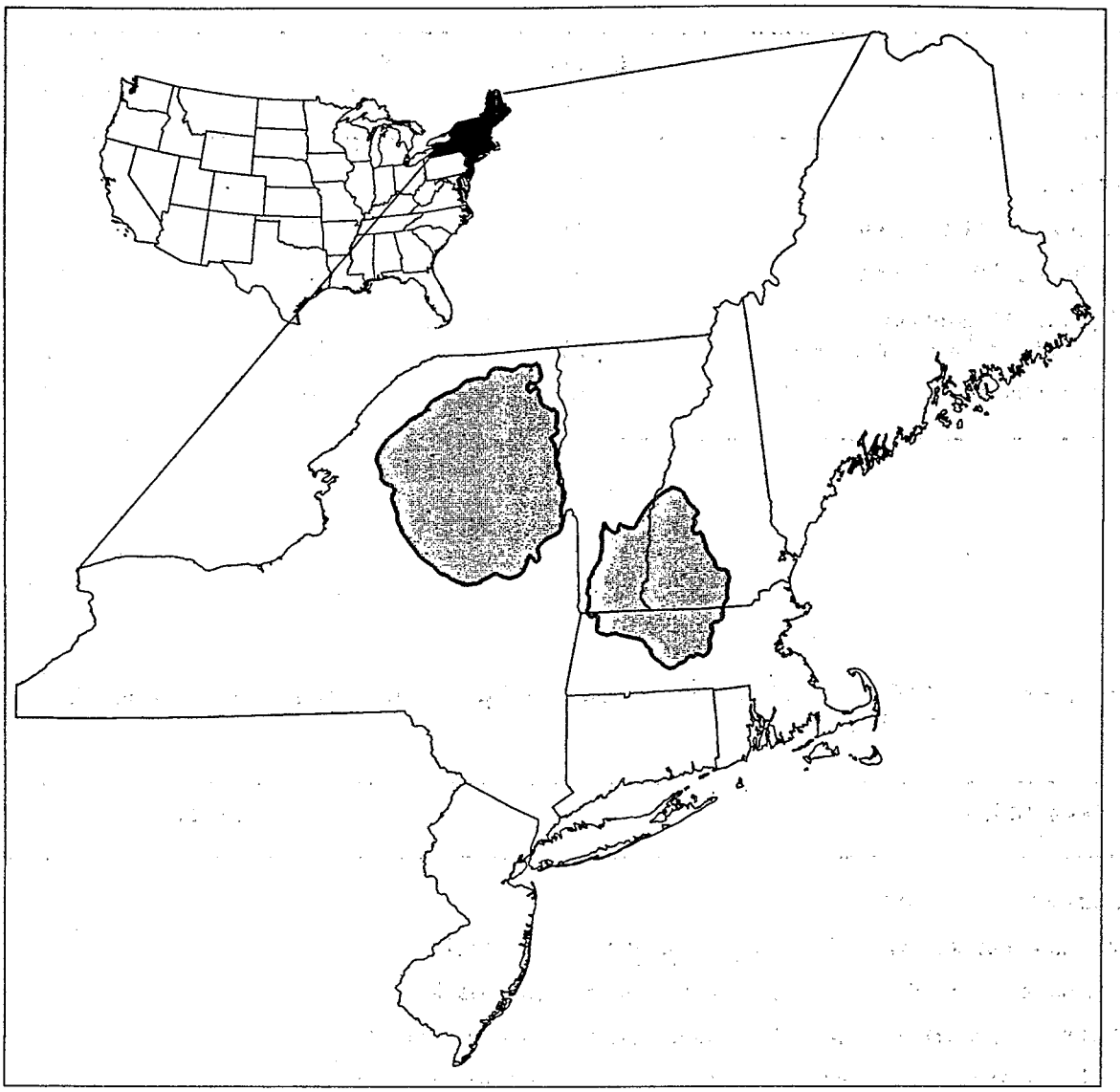


Figure 3-7. **Areas of the Northeast requiring threefold augmentation of the grid to supply enough sample sites for the project.**

4.0 INDICATOR DEVELOPMENT AND EVALUATION

Section 4 addresses the issues of selecting, evaluating, and using a set of measurements to assess the ecological condition of surface waters. This complex process is constrained by the goals of EMAP-Surface Waters (to make regional and national assessments of the ecological condition of lakes and streams over decades) and by the realities of actually performing field sampling at these scales.

This section discusses a wide range of difficult topics, including what organisms and chemical and physical features to target, when and where, within a waterbody, to sample, how to convert the measurements (data) into meaningful assessments of ecological condition and how to relate the measurements to issues of concern to society and science. There are no simple solutions for these issues that will satisfy every need. The material presented here is a beginning, with many aspects needing further substantial effort.

This section presents, for surface waters, an overview of EMAP indicators (4.1), criteria for determining whether a waterbody is impaired (4.2), ecological response indicators chosen for surface waters (4.3), exposure/habitat indicators (4.4), stressor indicators (4.5), and a strategy for the continuing process of selecting, developing, evaluating, and using indicators (4.6). The subsections on response indicators for lakes (4.3.1) and for streams (4.3.2) include synopses of the rationale for their inclusion, and issues related to sampling and making assessments from the data. Before any field work is performed, operations manuals (laboratory and field methods, quality assurance, field implementation) will be prepared.

4.1 CONCEPTUAL FRAMEWORK FOR INDICATORS OF CONDITION

Assessments of surface water condition must relate to the values and problems of present or potential concern to the public and to scientists. These issues are referred to as endpoints of concern, and are reflected in the ways lakes and streams are managed. For EMAP-Surface Waters, these values or endpoints are trophic state, fishability, and biotic integrity, as described in Section 1.5. Thus, our assessments of condition will be made in terms of each of these endpoints, rather than the single entity of surface water condition.

Neither ecological health (as a single entity analogous to human health) nor the status of the endpoints of concern can be assessed directly (Karr 1981, Rapport 1989, Schaeffer et al. 1988). Health assessments are initiated by making a series of measurements. Yet, the raw data from these measurements, by themselves, are not sufficient. These data must be evaluated within the framework of ecosystem theory and known hazards to surfaces waters (Section 1.5), and converted into indicators of condition.

The selection of indicators and data collection methods is constrained by the broad-scale perspective of EMAP. We will use an index sampling approach, which is a sampling regime that targets one or a very few sample times and locations within each waterbody, limiting effort on individual units to maximize regional coverage with available resources. The index sampling concept and the choices of when and where in a waterbody to sample are explained in greater detail in Section 6. Index sampling measurements are intended to be single "snapshots" of conditions to accompany snapshots from other waterbodies within the population of interest. The interest is in the collection of pictures of the population of lakes or stream reaches, not the individual snapshot.

The snapshot characterizations derived from index sampling have been criticized because the data are inadequate for addressing ecological processes, for describing seasonal or diel variability, for understanding individual lakes or streams, and for detecting ecologically important events such as episodic acidification, peak algal densities, or sensitive life stages. It is certainly true that an index sample is not adequate for these purposes, but that is not the objective of the program. Other investigators are addressing these issues in research programs designed specifically for those purposes. Duplication of their efforts with a new program would be counterproductive.

In a program that uses multiple indicators, it is not possible to choose an index time and location optimal for all indicators and all possible stresses. However, it is necessary that the sampling times,

locations, and protocols be adequate for expressing ecological conditions. For example, in lakes, a summer epilimnetic sample may be enough to characterize water chemistry and trophic condition, but would miss low pH associated with snowmelt, or spring phytoplankton blooms in lakes.

As discussed in Section 2.1.2, EMAP has four types of indicators of ecological condition: response, exposure, habitat, and stressor. It is important to understand that these categories provide a conceptual framework, acting as a tool to guide the process of selecting, evaluating, and implementing the actual measurements used to assess ecological condition. These categories are not meant to be rigid boxes in which each measurement type may serve only one function.

4.1.1 Response Indicators

Response indicators are derived from measurements that describe the biotic condition of organisms, populations, communities, or other components or processes of the aquatic ecosystem as they relate to the endpoints of concern (fishability, trophic condition, biotic integrity). Response indicators are the focal points upon which the health of surface water systems will be determined. The measurements should be amenable to quantifying the integrated response of the ecological resources to individual or multiple stressors. Response indicators should clearly relate to aspects of the environment valued by the public and the scientific community, such as overenriched or brown waters, biodiversity, ecosystem sustainability, and wilderness recreational value. Response indicators may also be chosen to identify emerging problems. Exposure and habitat indicators can also be used to define waterbody types and to describe expected conditions.

Figure 4-1 illustrates the conceptual process of condition assessment using response indicators. The rectangles represent types of information. The diamonds indicate processes of information refinement. The raw data from a collection of measurements must be evaluated within the ecological context (waterbody type and size, season, geographic location, etc.) that defines what is expected for similar waterbodies. These evaluations produce scores for metrics (standards of measurement). The metric scores can be further aggregated into indices. This process can range from simply summing unweighted metric scores to including metric scores in complex models. Both metric scores and index scores are used as response indicators in the process of deriving an overall statement of condition. The complexity of information (amount of detail) is reduced at each step of the process.

Figure 4-2 presents a simplified example of this process using fish assemblages. The measurements are the fish species and their abundances. These data are evaluated in their ecological context (e.g. warmwater streams in the Midwest, draining watersheds of < 100 km^2). Among the metrics that could be scored from these data are species richness and proportion of exotic species. These metrics must be calibrated by the number of species and the proportion of exotic species expected in unimpaired warmwater streams of that size in the Midwest. The individual metric scores can be used as indicators of biotic integrity and fishability in the overall assessment of those endpoints. These scores can also be incorporated into an index, such as the Index of Biotic Integrity, which in turn becomes part of the final assessment process. Using both metric scores and index scores preserves more of the detail of the raw data.

4.1.2 Exposure and Habitat Indicators

Exposure and habitat indicators serve a diagnostic function when measured in conjunction with response indicators. They are used to identify the likely causes of impaired conditions as detected by the response indicators. The exposure and habitat indicators are characteristics of the aquatic environment that give evidence of the occurrence and/or magnitude of a response indicator's contact with a physical, chemical, or biological impact.

Historically, we have evaluated primarily water quality (i.e., measures of physical and chemical characteristics of surface waters). These assessments have relied largely on what we have called exposure indicators.

EMAP-Surface Waters plans to use measures such as sediment toxicity, chemical contaminants in fish, and general water chemistry (chemical habitat) as exposure indicators. Habitat indicators are attributes that describe the physical, chemical, and biological conditions of the environment. They are extremely useful for classifying water

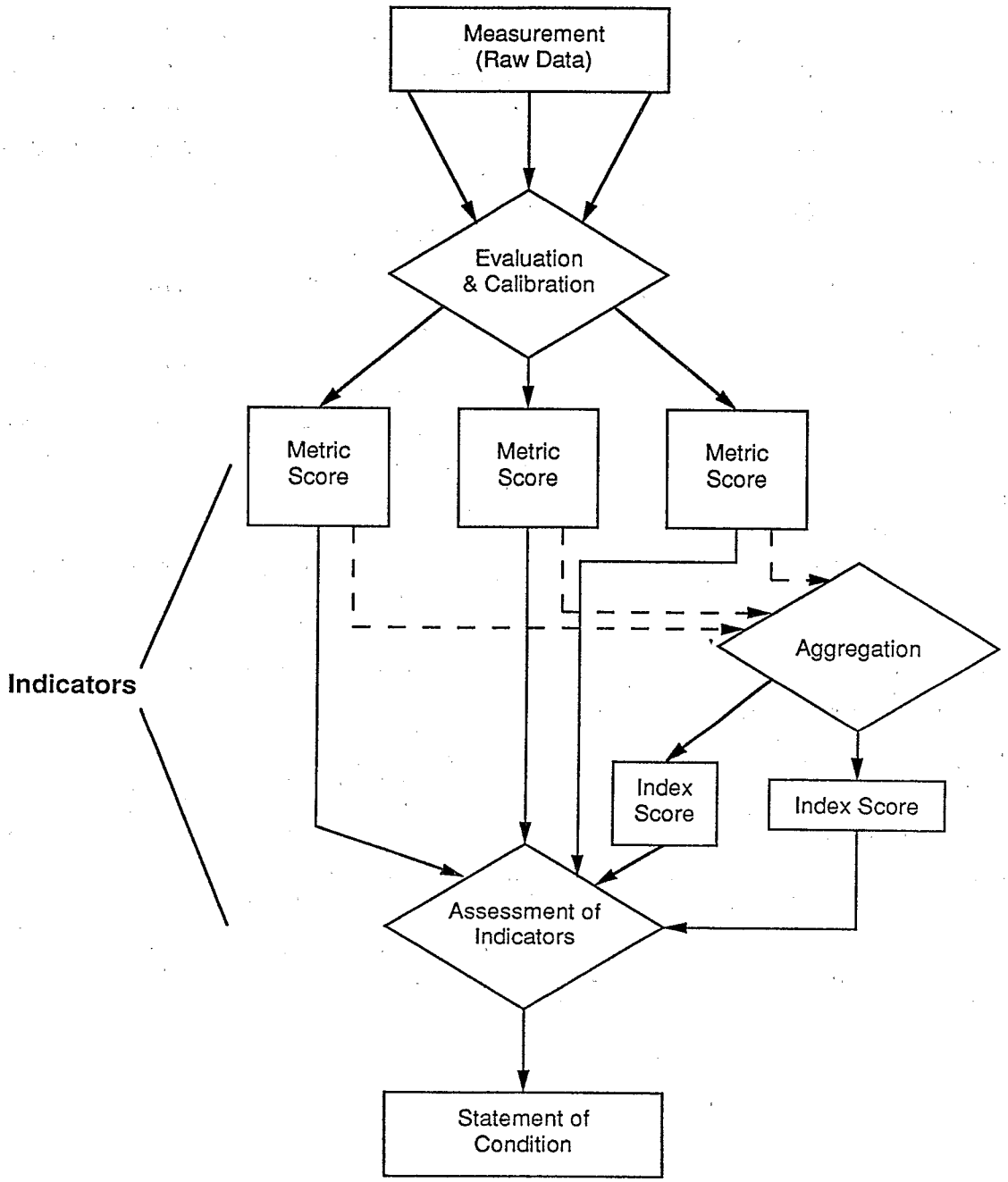


Figure 4-1. The conceptual process for proceeding from measurements to indicators to assessment of condition.

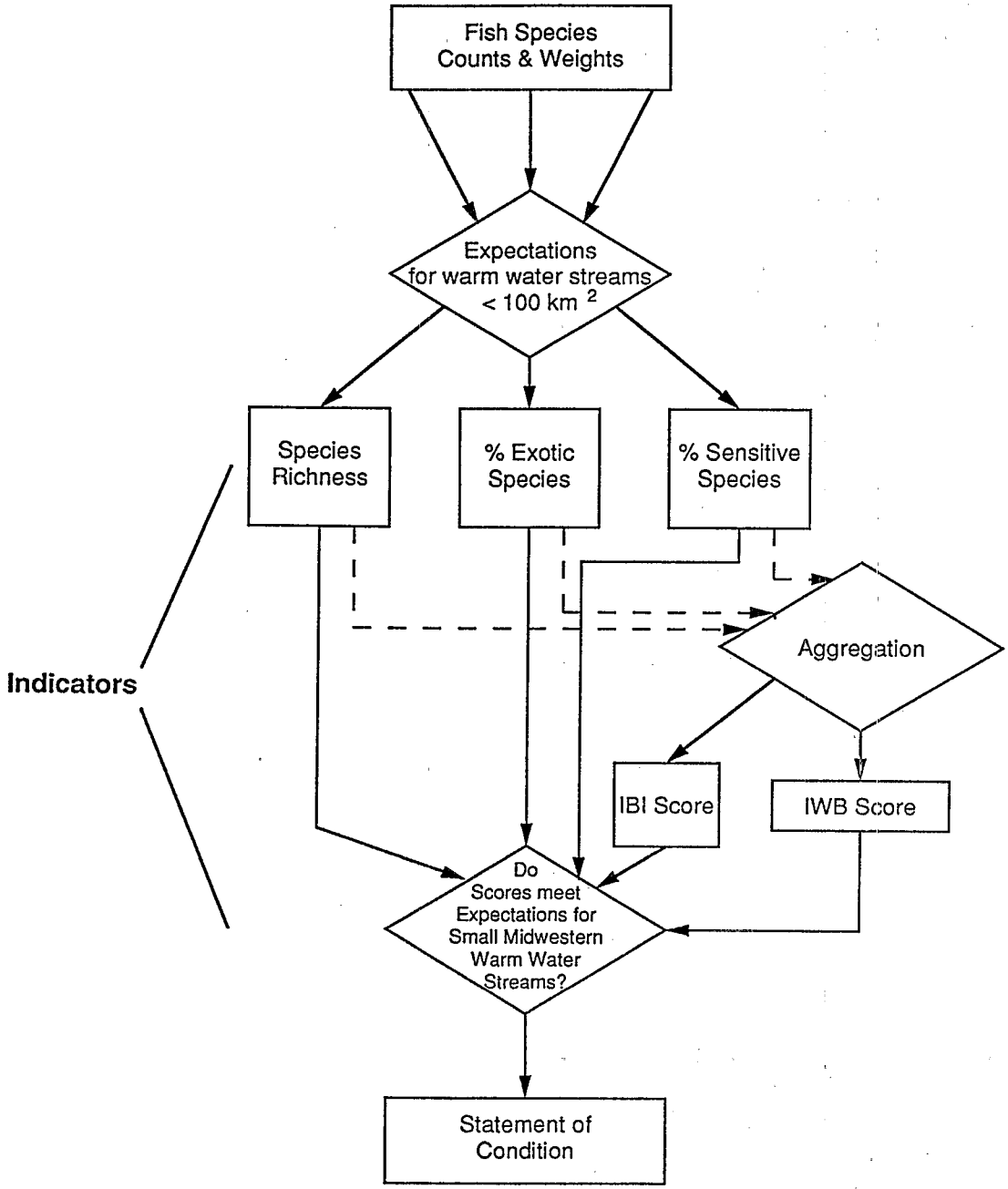


Figure 4-2. The process of proceeding from measurements on fish assemblages to indicators such as Index of Biotic Integrity (IBI) or Index of Well Being (IWB).

body types and for providing natural limits on system behavior. The chemical and physical habitat indicators can be used to model conditions that promote healthy ecological situations, as well as indicate probable causes of poor condition. These indicators are discussed for lakes and streams in Section 4.4.

4.1.3 Stressor Indicators

Stressor indicators describe characteristics, processes, or activities that give rise to exposure (impact) within aquatic ecosystems. They can be thought of as characterizing the sources of exposure. Some examples include land use and land cover, pesticide application rates within the watershed, livestock management activities, human population densities, export of pollutants, flow and channel modification, presence of introduced species, and stocking and harvest records. Stressor indicators may include natural processes, such as climate cycles. However, the primary intent of stressor indicators is to characterize the activities over which humans have control and that can create deleterious exposures. Stressor indicators proposed for EMAP-Surface Waters are summarized in Section 4.5.

The primary role of exposure, habitat, and stressor indicators within EMAP is to identify the probable causes for impaired conditions. Thus, the assessment process for these indicators differs from that envisioned for response indicators. In addition, exposure and habitat indicators are important in defining expected conditions, so that nominal/subnominal conditions can be set. Except for the final assessment of ecological condition, many of the steps in Figure 4-1 may be used to assess these diagnostic indicators. After metric and index scores have been determined, further analyses will be required to relate these indicators to the behavior of the response indicators. Correlation of unacceptable conditions in response indicators with various exposure, habitat, and stressor indicators is the first level of this analysis. Further diagnosis will undoubtedly be required. This is an area with many research opportunities. It is also worth noting that this diagnosis stage will be most fruitful after EMAP has successfully used the response indicators to characterize the condition of lakes and streams.

Table 4-1 presents indicators currently being considered for use in EMAP-Surface Waters.

Table 4-1. Candidate Indicators for Inland Surface Waters

Response
Trophic state index
Sedimentary diatom assemblage
Zooplankton assemblage
Macroinvertebrate assemblage
Fish assemblage
Gross external anomalies
Semiaquatic wildlife assemblage (birds)
Exposure/Habitat
Physical habitat quality
Chemical habitat quality
Sediment toxicity
Chemical contaminants in fish
Chemical contaminants in sediments
Biomarkers
Stressor
Land use and land cover
Human and livestock population density
Atmospheric emission and deposition
Use of chemicals
Pollutant loadings
Flow and channel modification
Lake level and shoreline modification
Introduced species
Stocking and harvesting records

4.2 ESTABLISHING UNIMPAIRED (NOMINAL) CONDITION

The process of selecting and implementing all four types of indicators cannot be separated from the problem of how the data collected and the resultant indicators will be used to make statements about lake and stream conditions. The indicators will show that some of the waterbodies are impaired (subnominal, unhealthy). The task is to establish the criteria that define impaired systems, that is, to determine that below a certain index or metric score, a lake or stream is considered to be in unacceptable condition, relative to a particular endpoint of concern. Some of the approaches for establishing these criteria include selecting and assessing reference sites, using historical data and/or pristine sites, using ecological models, and basing the criteria on the empirical distribution of indicator values.

EMAP-Surface Waters will interpret ecological condition based on a combination of these approaches, but will rely heavily on a series of regional reference sites (Hughes et al. 1986, Hughes 1989) and knowledge of historic conditions as described through paleolimnological reconstructions, expert opinions, and previous biological collections. This approach will allow us to integrate professional judgement, an understanding of historical or pristine conditions, and knowledge of current ecosystem research to select the least disturbed but typical sites of a region. The response indicator values, and the statistical variability found in the reference sites can be used to develop the expected or acceptable conditions for the region. These sites can also be monitored to help detect natural spatial and temporal variation. We expect that the waterbodies selected for Tier 2 sampling (Section 3.2.2) will represent the overall range of conditions for each region and that very few of the best sites will be selected. Purposely selecting additional unimpaired sites will ensure a robust model of unimpaired conditions. It will also be useful to examine the range of stressors impinging on the Tier 2 waterbodies, and to purposely select additional systems to ensure that the model of severely impaired systems is well defined.

Another option for determining whether a system is impaired is to use one or more ecological models. Models may be based on historical data from field collections for a key variable or set of variables, on complex ecosystem studies, or on laboratory toxicity tests. Empirical models that identify relationships, such as maximum fish species richness as a function of lake area or stream size, are an effective way to establish criteria. Habitat classification models should be useful for assessing exposure and habitat indicators. However, key exposure and habitat variables differ considerably across the nation. For example, stream ecosystem research stresses physical habitat in the Pacific Northwest, flow regime and habitat in the Intermountain West, and sediment in the Midwest. The fact that models tend to be stressor or species specific and/or applicable to limited geographic areas will constrain their usefulness at first. However, as EMAP progresses, models should become increasingly important for refining criteria.

Selecting an impaired/unimpaired boundary based on the empirical distribution of indicator data from Tier 2 sampling sites has been suggested from time to time. For example, the median value, or the 25th percentile, or some obvious breakpoint in the distribution of values could be selected as a criterion. There are problems with using this type of purely empirical approach. First, distributions of most indicators are continuous (without clear breakpoints). Second, given the extent and frequency of diffuse pollution, there is a high probability that even the highest quality grid sites may be at least partially impaired.

North America's biomes and ecoregions vary considerably, and they offer only a coarse classification of waterbodies. All indicators do not follow the same regional boundaries, because indicators are controlled by different limiting factors. At a minimum, though, we should expect different indicator scores in different regions because of differences in habitats and species. The combination of reference sites, regional expert opinion, and historical data offers a means of evaluating indicator scores against regional potentials or standards. These then can be combined into national scores.

We do not expect to answer any of the questions quickly. No matter which methods are used to estimate the impaired or unimpaired condition, there will still be considerable natural variability across a region and among lake types and sizes. Determination of the boundaries will therefore have to involve some degree of judgment. The question thus becomes, "With what level of confidence do we want to avoid making a type 1 or type 2 error and which is of greater concern?" In many respects, the definition of acceptable condition is a social not a scientific one. Our intent in EMAP is not to defend any specific value as the one and only value but rather to lead the discussion on how to objectively establish this criterion. EMAP will produce data that can provide the foundation for the assessment and take part in the technical discussion on establishing these boundaries. We believe that we also have some obligation to make an initial decision as to where the boundaries might be. At that point, however, many competing social, economic, scientific, and political issues must be considered.

4.3 RESPONSE INDICATOR SELECTION

The process of selecting indicators has been complex, governed by EMAP's emphasis on ecological response and on broad-scale assess-

ments. It has been guided by consultations with other professionals in aquatic ecology and by examination of available databases and the scientific and management literature. Much of the effort has been directed to response indicators, which address the first of the EMAP objectives. The discussions here are not intended to be a review of existing literature but rather a summary of our plans to use various taxa and indicators.

There is a general feeling that our ability to measure and interpret biological data lags behind our ability to measure and interpret chemical data. In some respects this is true. However, our ability to interpret the meaning of chemical data in terms of its impact on the biota is somewhat limited. If the biota are the concern, it is prudent to take direct measurements, even if our ability to interpret them is in its early stages.

Some indicators are considered "core," in that established sampling and evaluation methodologies exist. Others are developmental indicators, depending on the amount of research or testing of sampling and/or evaluation methods that is needed.

EMAP-Surface Waters has several general criteria for evaluating the usefulness of response indicators. Indicators should be biological and should incorporate elements of ecosystem structure and function. The selected indicators should correlate with changes in other unmonitored biological components. They must have clear connections with the three endpoints of concern, and must be responsive to a broad array of potential stressors. An ideal indicator is applicable in a broad range of surface water types across the nation. The response indicators must be sensitive to varying levels of stress, but not produce excessive false alarms. This includes sensitivity to important episodes that occur outside the sampling period. Table 4-2 presents the indicators being considered for EMAP-Surface Waters and indicators that have been considered but not selected.

Additional criteria apply to all types of indicators. Useful indicators are cost effective, providing considerable information in a limited amount of sampling time. They should be easily implemented and provide reproducible results with low sampling variability.

4.3.1 Lake Indicators

This subsection describes the response indicators chosen for lakes, and discusses some of the variety of concerns that affect their implementation in the Tier 2 sampling (Section 3.2.2). In the first year in each region, part of the pilot program will be dedicated to addressing these concerns.

The proposed response indicators for lakes, discussed in the following paragraphs, are:

- Trophic state
- Fish assemblages
- Macrobenthos assemblages
- Zooplankton assemblages
- Sedimentary diatom assemblages
- Semiaquatic wildlife assemblages

Trophic state indices directly address the trophic state issues of water clarity, noxious algal blooms, and anaerobic conditions in lakes. It will not be possible to completely characterize the trophic state of many lakes with single mid-summer samples. That kind of information requires sampling during spring or fall turnover, as well as during peak algal blooms. However, for EMAP's purposes, an adequate characterization can be achieved with a single mid-lake epilimnetic summer sample.

Trophic state will be assessed from Secchi disk transparency measures and concentrations of chlorophyll-*a*, total phosphorus, and total nitrogen. These data will be converted to one or more indices, such as Carlson's Trophic State Index (Carlson 1977, Shapiro 1979, Walker 1984, and Brezonik 1984). Information on total suspended solids, dissolved organic carbon (DOC), and macrophytes will be used to evaluate nonalgal turbidity and production and eventually to modify current indices.

Fish assemblages will be used to address two of the endpoints of concern: biotic integrity and fishability. The overall assemblage structure (species, abundances, and weight, length, and general condition of individuals) will be used to evaluate biotic integrity; the sports fish portion of the assemblage will be used to address fishability. In most regions and lake types, fish assemblages are relatively stable over the summer index period.

Table 4-2. Evaluation of Some Candidate Indicators for Inland Surface Waters by EMAP Selection Criteria (H=high, M=medium, L=low)

	Critical					Desirable						
Indicator	Correlate Unmeasured[a]	Regional Application[b]	Integrate Effects[c]	Relate Monotonic[d]	Simple Quantification[e]	Important[f]	Responsive[g]	Anticipatory[h]	Standard Method[i]	Low Measure Error[j]	Historical Data Base[k]	Cost Effective[l]
Selected as Research Indicators												
Lake trophic status	L	H	M	H	M	M	H	M	H	H	H	H
Fish MEM	H	M	H	H	H	H	H	M	M	H	M	H
Top carnivore index: fish	H	M	M	H	H	M	H	M	M	H	M	H
Macroinvertebrate assemblage	H	H	M	H	H	H	H	M	M	M	M	M
Sedimentary diatom assemblage	H	M	H	H	H	H	H	H	H	H	M	H
Semiaquatic vertebrates	L	L	H	H	H	H	L	H	H	L	M	H
External pathology: fish	L	H	L	M	H	L	M	M	H	H	L	H
Water column/ sediment toxicity	L	L	L	H	H	L	M	L	H	L	L	L
Chemical contamination: fish	L	L	L	H	H	L	H	L	H	H	M	L
Physical habitat quality	H	H	H	M	H	H	H	H	M	M	L	H

[a] Correlates with changes in processes of other unmeasured ecosystem components.
[b] Is applicable on a regional basis.
[c] Integrates temporal and spatial effects.
[d] Relates unambiguously and monotonically to an environmental value or a relevant exposure or habitat variable.
[e] Can be quantified by synoptic monitoring or by automated monitoring in a cost-effective manner.
[f] Is important to overall ecological structure and function.
[g] Responds to stressors or management strategies.
[h] Is anticipatory by providing an early warning of widespread changes.
[i] Has a standard method of measurement.
[j] Exhibits low measurement error.
[k] Has a functional data base or can be generated from accessible data sources.
[l] Is cost effective (has low cost and high informational value).

(Continued)

Table 4-2. Evaluation of Some Candidate Indicators for Inland Surface Waters by EMAP Selection Criteria (H=high, M=medium, L=low) (Continued)

	Critical					Desirable						
Indicator	Correlate Unmeasured[a]	Regional Application[b]	Integrate Effects[c]	Relate Monotonic[d]	Simple Quantification[e]	Important[f]	Responsive[g]	Anticipatory[h]	Standard Method[i]	Low Measure Error[j]	Historical Data Base[k]	Cost Effective[l]
Routine water chemistry	M	M	L	M	L	H	M	M	H	M	H	M
Biomarkers	L	H	L	M	L	L	H	H	M	M	L	L
Zooplankton	L	M	L	M	L	H	H	M	H	L	L	L
Human-made organics/heavy metals	L	L	L	M	H	L	M	L	H	L	L	L
Considered but not Selected												
Water column bacteria	L	L	L	H	L	L	H	M	H	L	L	L
Phytoplankton	L	M	L	M	L	H	H	H	H	L	L	L
Growth rate	L	H	L	L	L	L	L	M	H	L	L	L
Primary production	L	H	L	L	L	L	L	M	H	L	L	L
Community response	L	H	L	L	L	L	L		H	L	L	L
Nutrient cycling	L	H	L	L	L	L	L	M	L	L	L	L

[a] Correlates with changes in processes of other unmeasured ecosystem components.
[b] Is applicable on a regional basis.
[c] Integrates temporal and spatial effects.
[d] Relates unambiguously and monotonically to an environmental value or a relevant exposure or habitat variable.
[e] Can be quantified by synoptic monitoring or by automated monitoring in a cost-effective manner.
[f] Is important to overall ecological structure and function.
[g] Responds to stressors or management strategies.
[h] Is anticipatory by providing an early warning of widespread changes.
[i] Has a standard method of measurement.
[j] Exhibits low measurement error.
[k] Has a functional data base or can be generated from accessible data sources.
[l] Is cost effective (has low cost and high informational value).

Sampling will generally be done with sonar and overnight sets of gill nets in deep waters, and with traps and seining and/or electrofishing in shallow areas. All these methods have possible restrictions. Seines are not effective in weedy, rocky, or snag-filled areas. Gill nets may not be allowed in some lakes due to fish mortality. Sonar data do not indicate species, traps are difficult to set on dropoffs, and electrofishing requires a boat and trailer. There is a need for research to assess problems of variability and effectiveness with these methods in the various regions and lake types and with various fish assemblages. However, we lack sufficient funds to use all gear in all lakes sampled in this summer's (1991) pilot. We will determine the minimum gear requirements for use in all lakes and test the other types of gear in a subset of lakes.

Currently, there are no indices of biotic integrity for fish assemblages in lakes comparable to Karr's IBI (Karr et al. 1986), which was developed for warmwater streams, or for fishability (fishing quality) in lakes, although the importance of fish in lakes is well documented (Northcote 1988, Tonn 1990). Such indices are research tasks for EMAP-Surface Waters. Biotic integrity will be assessed by metrics such as species richness, percent sensitive species, percent exotic species, age/size structure, rates of external anomalies, and rates and types of species replacement. All these metrics must be evaluated within the context of fish assemblage type (Tonn and Magnuson 1982).

Fishing quality is an endpoint of concern for lakes. Fishability will be assessed by metrics such as presence, relative abundance, and age/size structure of game fish populations, addressing the questions: Are there any large game fish? Are populations reproducing? The presence of external anomalies, as well as consumption criteria violations, address concerns about perceptions of edibility and the health hazards of eating fish. We also plan to use a questionnaire, such as that in Plafkin et al. (1989), to assess fishing quality.

Benthic macroinvertebrate assemblage structure will be used to evaluate biotic integrity. Macroinvertebrates play a significant role in the food web, between primary producers and fish, and are important indicators of biotic integrity in their own right. Macroinvertebrates are a particularly critical class of organisms to monitor in those systems that naturally have very few or no fish. This group has been used historically as a biomonitoring tool. Sampling will probably be done with a combination of several methods, such as sweep and kick nets, handpicking from natural substrates, Ekman or Ponar dredges, and sediment corers. Research is needed on collection method variability to develop protocols to assure consistent, repeatable sampling.

Currently, there are no indices of biotic integrity for macroinvertebrate assemblages in lakes. Considerable research is needed to develop these indices, to modify indices that are in use for streams, and to determine the sensitivity of species to various stressors. Candidate indices and metrics include: taxa richness, number and proportion of Ephemeroptera and Trichoptera taxa, a series of ratios of certain taxonomic groups (e.g., Chironomidae, scrapers, shredders) to the total, and assemblage level indices such as the Biotic Index (Hilsenhof 1988).

Zooplankton assemblages will be evaluated to assess biotic integrity primarily, and fish food abundance to a lesser degree. Zooplankton perform the same role as macroinvertebrates, but in the water column, and they are relatively easy to sample. Our major concern with zooplankton is their species richness and what the proportional abundance of various taxa can tell us about a lake's quality. Zooplankton are sensitive to fish predation and fish introductions, as well as to trophic state and dominant phytoplankton, and they are the dominant animals in most lakes in terms of numbers and biomass.

Because we are interested mostly in rotifers, cladocerans, and copepods, we will take an index sample by vertically towing a fine mesh Wisconsin net at the deepest point of the lake. Multiple tows will be taken at a subset of lakes, and some lakes will be sampled each month during the pilot to assess spatial and temporal variability.

Data interpretation will center on species richness, relative abundance of generally tolerant or intolerant taxa, and feeding guilds. Zooplankton are typically not very speciose, with 4-15 species per tow in most small lakes (Lind 1974, Ruttner 1963).

Sedimentary diatom assemblage data will be used to evaluate the extent and time course of environmental and biological change (biotic integrity). The sensitivity of individual species to

changes in water chemistry can be matched with fossils preserved in deep sediments to track temporal changes in such factors as pH, metals, and nutrient levels. Samples collected from the top of the lake sediments integrate the diatom assemblages over the full year. The entire sediment core preserves the lake's history.

Sample collection, preparation, and analysis methods are fairly well developed for lakes. A single core of bottom sediments provides sufficient material to develop diatom profiles for the entire lake. We will examine diatom assemblages from the top of the core and from 50-100 cm deep in the core. Recent developments in statistical analysis of diatom data make it possible to calibrate top assemblages with current condition, and then to reconstruct historic conditions from deep assemblages. The use of diatom analyses in paleoreconstructions has been growing quickly. Diatoms have been used effectively to evaluate trends in acidification, eutrophication, and general pollution conditions (Agebeti and Dickman 1989, Charles et al. 1990, Lowe 1974, Stoermer et al. 1985).

Semiaquatic wildlife assemblages may be used to further assess biotic integrity. This indicator includes a variety of birds, mammals, amphibians, and reptiles that spend a substantial portion of their lives in or near lakes. Many of these species are important recreationally (for hunting or as "watchable" wildlife), and as a group linking lakes with their watersheds. This indicator is probably the least developed of those being evaluated, but it has considerable potential for future work.

An example of this kind of indicator is the U.S. Fish and Wildlife Service's (USFWS) Breeding Bird Survey (Robbins et al., 1986). The methodology could be modified for lakeside and/or canoe observations. These data could be assessed using a modification of the Brooks et al. (in press) response guild system. However, the optimal index period for birds is in the spring, which does not coincide with the mid-summer sampling period currently planned. Considerable research is needed to develop and assess protocols for surveying this component of lake ecosystems. We propose to conduct this research at a subset of 20-30 lakes this summer (1991).

4.3.2 Stream Indicators

This subsection describes the response indicators chosen for streams and discusses some of the variety of concerns that affect their implementation in the Tier 2 sampling. Many questions and topics discussed here are the same as those for lakes, but there are considerable differences in the details. In the first year in each region, the pilot program will be dedicated to addressing these concerns.

Fish assemblages will be used to address biotic integrity and fishability. As in lakes, fish assemblages are relatively stable during the late summer index period and interannually (Matthews 1986, 1990, Ross et al. 1985, Karr et al. 1986). However, Grossman et al. (1990) found high interannual coefficients of variation among stream fish species, and Matthews (1990) found high spatial variation in dominant stream fishes. The overall assemblage structure (species, abundances, and weight, length, and condition of individuals) will be used to evaluate biotic integrity; the sports fish portion of the assemblage will be used to address fishability. In most regions, streams will be sampled by electrofishing and/or seining. Electrofishing is not effective in streams with very low conductivity, in very turbid or colored systems, and with some species. Seines are not effective in weedy, rocky, or snag-filled areas. There is a need for research to assess the variability and effectiveness of these methods in the various regions and stream types.

The Index of Biotic Integrity (IBI) for stream fish assemblages has proven useful in the Midwest (Karr et al. 1986, Steedman 1988), and has been modified for use in other regions (Miller et al. 1988, Oberdorff and Hughes, in press). A research task for EMAP is to modify and evaluate the IBI in other regions where the expected species diversity is great enough to use the IBI framework. In regions and systems that naturally have few fish species, biotic integrity will be assessed by metrics such as percent sensitive species, percent exotic species, age/size structure, rates of external anomalies, and rates and types of species replacement.

As it is for lakes, fishing quality is an endpoint of concern in some kinds of streams (e.g., trout streams). Until a "Fishing Quality Index" can be developed, fishability will be assessed by metrics such as game fish species present, relative abun-

dance of game fish, age/size structure of game fish populations (Are there any large game fish? Are populations reproducing?), presence of external anomalies (Do the fish look edible?), and consumption criteria violations (Are the fish actually safe to eat?).

Benthic macroinvertebrate assemblage structure will be used to evaluate biotic integrity. Macroinvertebrates play a significant role in the food web, between primary producers and fish, and are important indicators of biotic integrity. The usefulness of macroinvertebrates increases in streams that naturally have very few or no fish. This group has been used extensively as a biomonitoring tool, to assess the effects of point sources of pollution as well as nonpoint sources (Plafkin et al. 1989, Fiske 1988, Metcalfe 1989). Sampling will probably be done with a combination of sweep and kick nets, and handpicking from natural substrates. Research is needed on collection method variability to develop protocols to assure consistent, repeatable sampling.

There are a number of indices of biotic integrity for macroinvertebrate assemblages in streams (Lenat 1988). However, considerable research is needed to further evaluate these indices and to determine species sensitivity to various stressors in many regions. Candidate indices and metrics include: taxa richness, number and proportion of Ephemeroptera/Plecoptera/Trichoptera (EPT) taxa, a series of ratios of certain taxonomic groups (e.g., EPT, Chironomidae) to the total, foraging guilds (percent shredders, grazers, collectors) as suggested by Cummins and Klug (1979), and assemblage level indices such as the Biotic Index (Hilsenhof 1987).

Semiaquatic wildlife assemblages may also be able to augment biotic integrity assessments for streams. For example, Brooks et al. (in press) found bird guilds to be more sensitive indicators than fish or benthos of landscape deterioration in Pennsylvania streams. In the long term, it may be possible to expand the Breeding Bird Survey to include transects that coincide with EMAP-Surface Water sampling units.

4.4 EXPOSURE AND HABITAT INDICATORS

As described in Sections 2.1 and 4.1.2, exposure and habitat indicators are intended to serve a diagnostic function when measured in conjunction with response indicators. That is, when response indicators demonstrate impaired ecological conditions, exposure and habitat indicators will be used to identify the probable (proximal, immediate) causes of that condition. In addition, exposure and habitat indicators are important in defining expected conditions, so that nominal/subnominal conditions can be set. Most water quality monitoring now relies on exposure indicator measurements to define and detect impairment.

As with response indicators, the broad scale of EMAP, along with its objectives, directed the selection of exposure and habitat indicators. For EMAP, it is not necessary (or practical) to completely characterize the water chemistry or substrate of each waterbody. Instead, the choice of methods (and targeted data) is based on the physical and chemical attributes that most directly and extensively affect the response indicators (and the endpoints of concern). For example, pesticides are less likely to occur in detectable amounts in mountain lakes than in agricultural streams. Thus, pesticides may be targeted in the latter but not the former. These indicators (and data) also need to be amenable to the index sampling concept. That is, they should provide robust characterizations of chemical and physical conditions without a need for targeting short periods of maximum concentrations (greatest stress conditions).

This section summarizes the five exposure and habitat indicators for both lake and stream systems. Conceptually, there is little difference between lakes and streams relative to these indicators, but there are some obvious differences in the details of designing, implementing, and analyzing exposure and habitat indicators in each.

Physical habitat quality, or the size, shape, and hydrology of the stream channel or lake basin, characterizes the physical conditions that provide the setting for different natural biological and chemical conditions. Alterations are likely to impact the ecological condition. Thus, physical habitat information is needed to fully interpret the response indicator data (Gorman and Karr 1978, Platts et al. 1983). For example, if the physical habitat score is high but the biological scores are low, then a chemical impact is probably causing the low response indicator score. On the other hand, if the physical habitat is severely degraded, the biological component will be impaired no matter how good the chemical habitat is (Plafkin

et al. 1989). Research is needed to define these relationships.

In general, physical habitat has four components: waterbody size and persistence, and habitat type and complexity (Table 4-3). Size may be a combination of surface area, maximum and average depth, and volume of water. For streams, volume is a function of cross section area and gradient. Persistence is a measure of the stability of the volume of water (e.g., lake level fluctuations, frequency and regularity of stream floods or droughts). Habitat type refers to a combination of water depth and velocity and substrate type. Some habitat types for lakes are profundal (deep), littoral (shallow, along the shore), and vegetated or not. Some stream habitat types are riffles, pools, and backwaters. Substrate type refers to either the bottom material (sand, mud, cobble) or to objects in the water on which organisms live (vegetation, logs, stumps, tree roots). Habitat complexity refers to the number and proportion of the different kinds of water depth and velocity, habitat, and substrate types occurring in the waterbody.

Some of this information will be obtained from maps and existing data (e.g., waterbody size). The remainder will be collected during Tier 2 sampling. Generally, these data will be collected by recording observations made along one or more transects of the lake or stream. A research task for this indicator is to develop a set of measures that adequately characterizes each waterbody, and that provides consistent, quantitative, repeatable data.

Table 4-3: Suggested Physical Habitat Measures for Lakes

Lake Size and Persistence Index

- Lake surface area (1:24,000-scale maps)
- Lake depth at subjective deepest point (field survey)
- Lake level fluctuation
  - Annual percent change in maximum depth (field shoreline survey)
  - Annual percent change in lake area (field shoreline survey)
  - Level control structures
- Lake hydrologic residence time (T_r = [est. volume]/[runoff × topographic watershed area])

Lake Habitat Complexity Index

- Littoral dominance
  - Percent littoral surface area (field sonar survey)
  - Temperature and dissolved oxygen stratification
- Bottom habitat complexity
  - Bathymetric maps
  - Coefficient of variation in lake depth along two cross-lake transects (field sonar survey)
  - Profundal substrate
  - Littoral substrate
  - Macrophytes
- Lake shoreline complexity
  - Shoreline development (1:24,000-scale maps)
  - Shoreline development (field estimate)
  - % artificial shoreline (several types)
  - Shoreline vegetation

Chemical habitat quality is the suite of variables that indicates (1) the general chemical climate surrounding aquatic organisms and (2) possible water column stresses on the biota (Hynes 1974, Warren 1974). Much variability in aquatic species composition and abundance not explained by physical habitat is likely to be associated with natural and anthropogenic differences in water quality. Levels of major ions, pH, temperature, and dissolved oxygen often control the chemical behavior and biotic stresses caused by other chemical constituents (e.g., toxins, nutrients, metals). Quantitative description of the water quality is a necessary part of the diagnosis of proximal causes of impairment, and facilitates an evaluation of the linkages between human activities and biotic responses.

EMAP-Surface Waters will examine four general attributes of water quality: geochemical type (ionic strength), acid-base status (sources of acidification), nutrient status, and redox status (possible toxic conditions) (Table 4-4). There is some overlap among the chemical species collected to assess these categories. The geochemical type variables (e.g, Na, K, Si, Mg, Ca, SO_4, NO_3, Cl) indicate the association between water quality and soil weathering processes, and some anthropogenic disturbances. Acid-base status includes measures of hydrogen ion activity and buffering capacity. The former controls the solubility and toxicity of many chemicals, the latter describes the degree to which the waterbody resists pH changes. Nutrient status addresses the supply of chemical compounds that often limit the growth of algae and macrophytes. The redox status of waters (assessed by dissolved oxygen, pH, and temperature) is a major control on the solubility, mobility and toxicity of many chemicals, including nitrogen and toxic heavy metals. Additional clues to anthropogenic disturbance (e.g., road salting, sewage disposal, irrigation) may be assessed with other combinations of chemical measurements.

For lakes, water temperature (profile), dissolved oxygen (profile), conductance, pH, and Secchi disk transparency will be measured mid-lake. Epilimnetic (0.5 m) and near bottom (for lakes low in dissolved oxygen) water samples will be collected for laboratory analyses, using National Surface Water Survey methods. For streams, grab samples will be taken in the approximate deepest water.

Sediment toxicity tests (Nebeker et al. 1984) will address possible biological exposure to toxic materials that have accumulated in sediments, which potentially affect biotic integrity and fishability. Toxicity and toxic materials act as a kind of endpoint of concern for the public, partially because of human health concerns (which are not directly within EMAP's scope). EMAP is not designed to pinpoint toxic hotspots, but it should be able to characterize, on a regional basis, whether and how much toxics are impacting biotic integrity and fishability.

Toxicity may be demonstrated by any of several adverse responses shown by test organisms (e.g., mortality, impaired growth, reduced reproduction), when compared with controls. In addition, there are several possible methodologic choices, such as which test species to use, the type of exposure (solid phase, pore water, elutriate), acute versus chronic, and test duration. Other research questions relate to the variability of laboratory operation, the representativeness of the collected sediments, and how to define control sediments.

EMAP-Surface Waters will evaluate 10-day acute solid-phase bioassays with *Hyalella* amphipods and 7-day solid-phase chronic assays with *Ceriodaphnia* and *Pimephales* larvae. Overlying, clean water will be used in all tests. The measurement endpoints for *Cerioclaphnia* tests are mortality and reproduction, and for *Pimephales* larvae, mortality and growth. We will be unable to conduct sediment toxicity tests at all pilot sites; however, we plan to do them for a subset of 40-50 sites.

Measurements of chemical contaminants in fish (Schmitt and Brumbaugh 1990, Schmitt et al. 1990) will provide direct assessments of the exposure to chemicals that fish have experienced, and the potential health hazard for humans as well as for other biota.

Numerous research questions still need to be answered for this indicator. Several of these questions involve the risk assessment process, such as which chemicals to target, whether to use whole fish or fillets, which species to use (game species or not), and which health criteria to use (there are large differences in criterion levels among regulatory agencies). The answers to some of the questions will influence other issues, such as what reproductive state to target (some

Table 4-4. Proposed Chemical Measurements Contributing to EMAP Geochemical Habitat Quality Indices

Measurement	Geochemical type	Acid-base	Nutrient	Disturbance	Redox
Temperature	X		X		X
Dissolved oxygen	X		X	X	X
Transparency (Secchi depth)	X		X	X	
Conductance	X			X	
TSS (unfiltered)		X		X	
Turbidity			X	X	
pH (field)	X	X		X	
pH (closed-system)		X			
pH (air-equilibrated)		X			
ANC (Gran)	X	X			
DIC		X			
DOC	X	X	X	X	
Color	X		X		
Na	X	X		X	
K, Mg, Ca	X	X	X		
SO_4	X	X		X	
Cl, NO_3, Si	X	X	X	X	
Total N (unfiltered)	X	X	X	X	
NH_4		X	X	X	
Ttoal P (unfiltered)	X	X	X	X	
Al_{tm} (closed headspace)	X	X			
Al_{im} (closed headspace)	X	X			
Total Al	X	X			
Chlorophyll-*a*			X	X	

chemicals accumulate in the gonads), which chemical analyses to use, and whether to archive samples.

Throughout this activity, we will maintain close coordination with the U.S. Fish and Wildlife Service National Contaminant Biomonitoring program. The proposed analytes are: Aldrin, alpha-BHC, Chlordane, Dacthal, DDD, DDEE, DDT, Dieldrin, Endrin, Heptachlor, Heptachlor epoxide, Hexachlorobenzene, Lindane, Mirex, Nonachlor, Oxychlordane, PCBs, Toxaphene, arsenic, cadmium, copper, lead, mercury, selenium, and zinc.

A set of **biomarkers** taken from fish blood, gill, and liver tissues will be evaluated to diagnose early signs of stress. We are particularly concerned with screening for short-lived chemicals, such as carbonates, pyrethroids, and parathion, that are difficult to detect by other means. These indicators are also being examined for their ability to give early warning signals of ecosystem stress. A third reason for using biomarkers is to provide an organism level indicator to complement the ecosystem (trophic state) and assemblage level response indicators.

We propose to evaluate biomarkers at a subset of lakes because of limited funds and because the sensitivity of biomarkers in moderately impaired and minimally impacted lakes remains questionable. A second research objective is to compare biomarker sensitivity and responses with those of the more typically used assemblage, ecosystem, and exposure indicators.

4.5 STRESSOR INDICATORS

As described in Sections 2.1 and 4.1.3, stressor indicators will be used, along with the exposure and habitat indicators, to further diagnose the probable causes of impaired conditions. Presumably, the stressor indicators will address the human actions causing impairment rather than the proximal causes. For example, pesticide application is a source of stress that shows up as chemical contaminants or toxic sediments (exposure indicators) and impaired biological assemblages (response indicators). This section describes these indicators only from the broadest view, because the highest priority for development and implementation goes to the response indicators first, then to the exposure and habitat indicators, and finally to the stressor indicators.

Few of the data needed to assess the stressor indicators will be obtained from Tier 2 field sampling. Instead, these data will be collected from a variety of existing data sources, including maps and management and regulatory agency reports and databases. A major source of data for the stressor indicators will come from the landscape characterization of the 40-km^2 hexagons made for Tier 1. One issue for the surface water component of EMAP is whether the hexagon characterization is adequate, or whether the complete watersheds for the Tier 2 waterbodies should be characterized.

Land use and land cover characterizations will be the most general stressor indicators. Land use will describe the most prevalent types of human induced stress (Omernik 1977, Smart et al. 1985). For example, agricultural land use tends to be associated with channelization, increased nutrient and sediment loading, increased pesticide levels, and decreased width of riparian woodlands, leading to less stream shading, more erosion, and less in-stream cover. Urban land use is associated with increased toxic materials, nutrient loading, lake shoreline alteration, and channel modifications. Areas with mining activities tend to have increased heavy metals and decreased pH. Land cover integrates human activity (land use) and natural vegetation potentials, such as climate and soil type. Remote sensing and available maps are extremely useful sources of this type of information.

Human and livestock population density will describe the general level or intensity of human-induced activity. That is, greater numbers of people generally mean more kinds of stress. Livestock density is important in areas of low human density, because a concentration of cattle, hogs, or poultry in feedlots or farms usually means greater nutrient loading. Where livestock graze, there is high potential for riparian vegetation removal, bank trampling, erosion, and nutrient loading (Platts 1979).

Pollutant loadings can be assessed by analyzing current pollution discharge permits, both municipal and industrial. These will further describe the kinds and intensity of nutrient, chemical, and thermal stresses.

Flow and channel modifications as well as lake-level and shoreline alterations can be assessed from landscape characterizations, regulatory agency records (e.g., Army Corps of Engineers), and habitat evaluations during Tier 2 sampling. Normally, channelization removes a great deal of the habitat diversity from streams, as does dredging in lakes. Flow modification often stresses aquatic ecosystems by disrupting the natural patterns of flood and low flow, as well as modifying temperature regimes (Karr et al. 1983, Simpson et al. 1982).

Stocking, harvesting and species introduction records provide information about the biological stresses in aquatic ecosystems. Management practices generally intended to enhance fishing quality often degrade the biotic integrity of surface waters. Stocking activities also mean that the waterbody cannot maintain the desired harvest levels. Introduced species, intentional or otherwise, have considerable potential for decreasing biotic integrity (Miller et al. 1989).

In examining all the indicators, we will first interpret the raw data to assess status, detect trends, and correlate probable causes. This information will be useful, but some researchers have found that guilds, indices, and proportionate abundances are less variable and more interpret-

able than presence/absence, absolute abundance, population abundance or biomass, or individual physical and chemical variables (Rahel in press, Plafkin et al. 1989, Karr et al. 1986, Landres 1983, Carlson 1977). Of course indices can camouflage or distort natural phenomena (Jackson et al. 1990, Boyle et al. 1990), especially when they are not grounded in ecological theory or concepts, when they are simple mathematical manipulations, or when an index developed for one region or use is applied without adjustment to a much different region or use. Regardless of the indices used, the raw database will be maintained to facilitate multivariate statistical analysis, data exchange, and later reinterpretations as our ecological insights grow.

Tables 4-5 and 4-6 present potential linkages between endpoints, measurements, and response and exposure indicators, respectively, providing an overview of the indicator approach for EMAP-Surface Waters. We hope ultimately to link the stressor indicators with the ultimate causes of environmental degradation, resource overconsumption, human overpopulation, and an incompletely developed sense of environmental ethics (Ehrlich et al. 1977, Nash 1989).

4.6 STRATEGY FOR INDICATOR DEVELOPMENT AND IMPLEMENTATION

Given the spatial, temporal, and ecological enormity and complexity of EMAP-Surface Waters, it is essential to have a basic strategy that defines a process, and a set of goals and criteria by which to evaluate whether the steps have been successfully completed. Each phase in the general strategy can be expanded to include as much detail as needed to assure achievement of its goals. The strategy described here has been modified from the overall EMAP Indicator Strategy (Knapp et al., 1991).

The indicator development strategy does not propose a linear process (Figure 4-3). Instead, there may be considerable branching off and feeding back to previous steps. Some indicators may progress relatively quickly through to full implementation; others may stall or progress slowly for many years. There are six basic phases:

1. Identify environmental values and apparent stressors.

2. Develop a set of candidate indicators that are linked to endpoints of concern and responsive to expected stressors.

3. Screen candidate indicators to select research indicators with reasonably well-established databases, methods, and responsiveness.

4. Quantitatively evaluate expected regional scale performance of research indicators to identify developmental indicators for regional demonstrations.

5. Demonstrate developmental indicators on a regional scale, using the sampling frame, methods, and data analyses intended for full (core) implementation.

6. Implement core indicators with annual sampling and data analyses and periodic reevaluation of indicators.

The first two phases are meant to generate ideas for endpoints and indicators. These two phases encourage broad-scale, lateral thinking, with a focus on breadth rather than depth of coverage, and may be revisited at any time (Figure 4-3). Essentially, EMAP-Surface Waters has completed the first pass through phases 1 and 2, using a series of workshops, informal literature reviews, and exploratory data analyses.

Phases 3 to 5 are oriented toward critical evaluation and integrative filtering of the candidate indicators down to a defensible, practical set of core indicators. Where phases 1 and 2 are inclusive, these three phases focus on excluding indicators that are currently not feasible within EMAP and on documenting the value of those selected. Each of these phases uses its own detailed research plan, sets of evaluation criteria, and peer reviews of the decisions made.

In practice, EMAP-Surface Waters will need to develop individual phase 3 to 6 strategies for lakes and streams on a regional basis, because (1) the surface waters component is being implemented in several regional steps (i.e., the first year will be a pilot study of lakes in the Northeast, with other regions and streams added sequentially over five years), and (2) there are major differences among waterbody types, sampling

Table 4-5. Linkages Between Potential Endpoints, Measurements, Metrics, and Response Indicators for EMAP-Surface Waters

Societal Concerns (Endpoints)	Measurements	Metrics	Indicators (with scoring criteria)
Fishable Waters (No. or %)			**Fishability Index (0 or 1)**
Gamefish sp. present	Sp. identification	Relative abundance	Absent (0)
Gamefish abundance	No. of individuals	Catch per unit effort (CPUE)	< .002 (1/8H) (0)
Gamefish size	Individual length/weight	% keepers, % trophy	< 50% keepers (0)
Fish appearance	External anomalies	% anomalies	> 5% with anomalies
Fish edibility	Toxic concentrations	Consumption criteria violations	Restricted > once/yr (0)
Fishery sustainable?	Individual weight, length, scales; Stocking records, catch restrictions	Age/size structure % wildfish, % keepers	Juveniles absent (0) Adults stunted (0) < 50% wildfish (0) < 50% keepers (0)
Trophic Condition			
Noxious algal blooms, surface scums	Pigment concentrations, visual, sediment diatom sp. and abundance	% blue-green algae % nuisance sp.	> 10% (0) > 10% (0)
Macrophytes	Macrophytes	% of lake macrophyte dominated	> 25% littoral zone (0)
Low transparency	Secchi depth	Secchi depth	< 2m (0)
Noxious taste/odor	Threshold odor	Threshold odor number (TON)	TON > 10 (0)
Fish kills	Hypolimnion oxygen concentrations	% depth < 3 mg DO/l	> 50% (0)
Trophic state	Pigment concentrations; total phosphorus; Secchi depth; total nitrogen	Trophic state index (TSI)	Trophic State Index (< 30, 30-60, > 60)
Change in trophic state	Sediment diatom sp. and abundance	% oligo-/meso-/eu-/dystrophic sp.; dominant sp.; % epiphytic; % planktonic	Rate of historical trophic state change (Δ TSI/yr)
Biological Integrity[b]			**Biointegrity Index[a]** (1) / (3) / (5)
Dead lakes/streams	No. of individuals[c] and no.	CPUE	< 33% / 33-67% / > 67%
Declines in sp. richness	Sp. identification[c] and no.	No. of species	< 33% / 33-67% / > 67%
Declines in sensitive sp.	Sp. identification[c] and no.	% sensitive	< 33% / 33-67% / > 67%
Increases in tolerant sp.	Sp. identification[c] and no.	% tolerant	> 25 / 10-25 / < 10
Increases in exotic sp.	Sp. identification[c] and no.	% exotic	> 10 / 1-9 / < 1
Reproductive failure/population sustainability	Individual length and weight (fish) No. of individuals (birds)	Age/Size structure CPUE	> 1+ / 1+->1+ (3) / 0+->3+ < 33% / 33-67% / > 67%
Evidence of kills	Questionnaire	Kill frequency	1-5 yr / > 5 yr / no kill
Increased anomalies	External anomalies (fish)	% anomalies	> 5 / 1-5 / < 1
Decreased abundance	No. of individuals[2]	CPUE	< 33% / 33-67% / > 67%
Decreased max individual size	Individual length and weight (fish)	% old growth	< 1 / 1-9 / > 10
Historical dislocation	Species and abundances (diatoms)		< 25 / 25-75 / > 75

[a] The Index of Biotic Integrity (IBI) and the Biotic Condition Index (BCI) will be used for stream fish and macroinvertebrate assemblages; criteria for both are based largely on regional reference site values.

[b] Determined from regional reference site values.

[c] Includes sedimentary diatoms, zooplankton, macroinvertebrates, fish, and birds for lakes; macroinvertebrates, fish, and birds for streams.

Table 4-6. Linkages Between Potential Endpoints, Measurements, and Exposure Indicators for EMAP-Surface Waters

Societal Concerns (Endpoints)	Measurements	Indicators
Are fish safe to eat?	Muscle toxics (metals, organics) concentrations	Violation of consumption criteria > 1/yr
Are waters becoming more eutrophic?	P, N, Ca, Cl, TSS, Secchi, temperature, pigments Lake sedimentary diatom sp. and abundance	Trophic State Index (TSI) % eutrophic, mesotrophic, oligotrophic
Are waters becoming more saline?	Cl, TDS Lake sedimentary diatom sp. and abundance	Total chloride, total dissolved solids % halophilic, halophobic
Is water acidity changing?	pH, DIC, ANC, Al, SO_4, NO_3, Cl, DOC, TN, Na, K, Mg, Ca, NH_4 Lake sedimentary diatom sp. and abundance	ANC, anions, cations % acidophilic, acidophobic
Are waters warming?	% canopy (streams) Lake sedimentary diatom sp. and abundance	% canopy (< 25, 25-50, 51-75, > 75) % stenothermal, metathermal, eurythermal
Are waters becoming more turbid?	Secchi depth, turbidity, TSS	Secchi depth, turbidity, total suspended solids
Are sediments toxic to aquatic life?	Hyalella, Ceriodaphnia, and Pimephales mortality and reproduction or growth	Toxicity significantly greater than controls
What is the critical level of habitat alteration for eliminating habitat sensitive fish and wildlife?	Lakes: % littoral dominance, temperature and DO profiles, lake area, maximum depth, level fluctuation, substrate and vegetation cover, depth variation, shoreline development, land use	Lake Habitat Quality Index (LHQI) Stream Habitat Quality Index (SHQI) (% regional reference site value < 25, 25-50, 51-75, > 75)
To what degree is biological impairment due to natural habitat conditions?	Streams: widths, depths, substrate and vegetation cover, embeddedness, flow, channel alteration/complexity, pool/riffle/run/bend, bank stability, riparian vegetation, immediate land use	

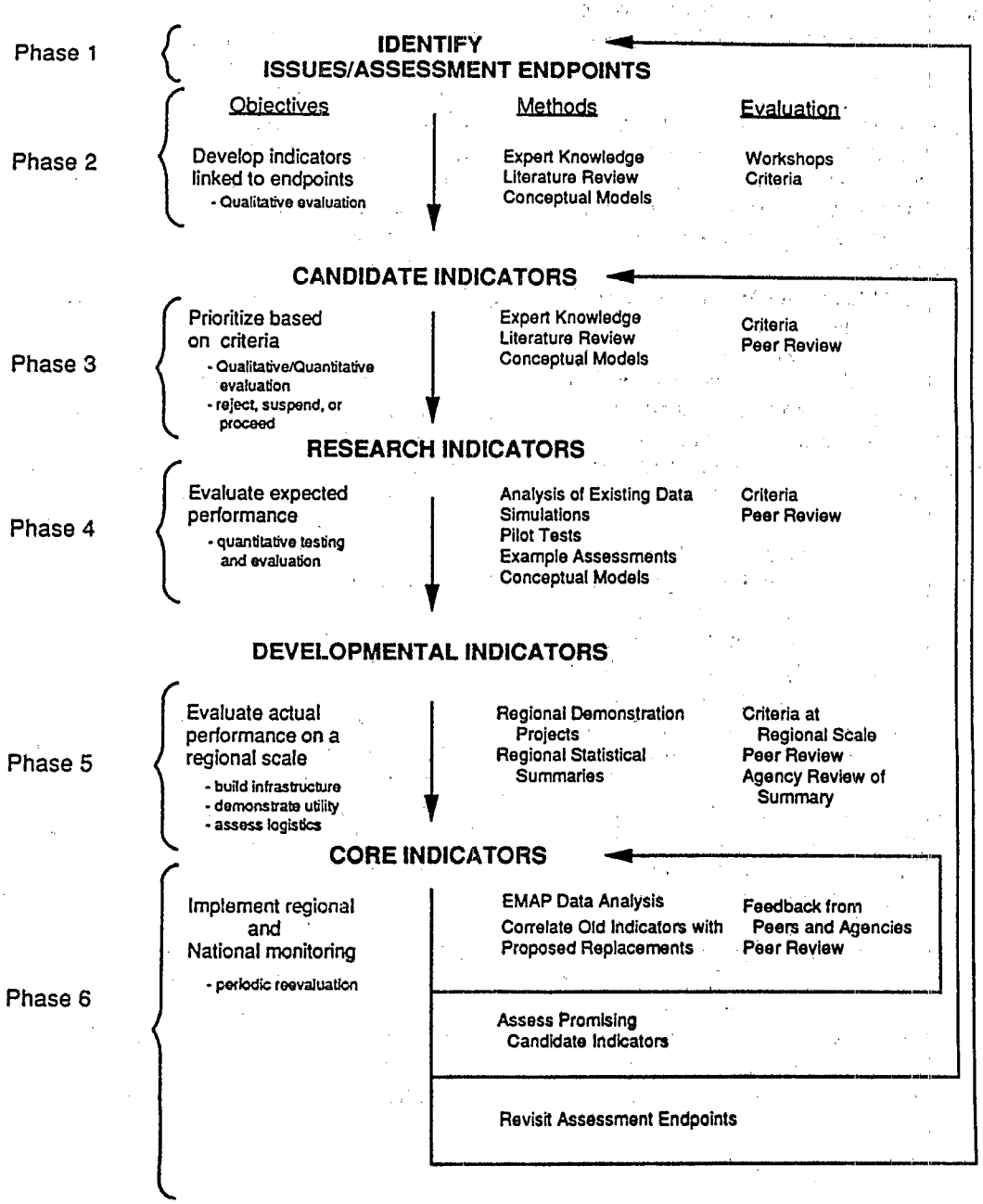


Figure 4-3. The indicator development process, showing the objectives, methods, and evaluation techniques used in each phase.

methods for lakes and for streams, current level of development of indicators in various regions, and regional importance of the endpoints of concern. An overview of a generalized plan for phases 3 to 6 is presented in the following paragraphs.

Phase 3 (select research indicators) relies on literature reviews and expert knowledge from professional scientists and water resource managers who have experience with a region's lakes or streams. Study of available species distributions and water quality, sampling methodologies, and interpretive techniques is emphasized. The results of this phase should be (1) a list of research indicators for the region and waterbody type, along with reasons for their selection, and (2) a proposed set of tasks for the next phase. EMAP-Surface Waters has completed the first pass through this phase for lakes in the Northeast.

Phase 4 (select developmental indicators) increases the level of detail of the information collected for phase 3, concentrating on questions of spatial and temporal variability, data interpretability, trend and association objectives, and proposed field and laboratory methods. During this phase we will acquire and analyze existing databases, simulate the minimum detectable trends (given the expected variability of the data), and determine preferred index periods and locations. A mock assessment should be developed to evaluate the utility of the complete suite of indicators. In some cases there should be a pilot test of indicators, especially in the first years of EMAP.

Phase 5 (identify core indicators) uses a regional demonstration project to address issues posed by the selection criteria and identified by the earlier phases. This phase should demonstrate whether the sampling and analysis protocols are adequate for generating regionally interpretable information.

In EMAP-Surface Waters, phases 4 and 5 are not necessarily sequential. All indicators selected have been field tested to some degree somewhere. Remaining issues include variability and regional applicability (evaluated by demonstration projects), as well as cost, redundancy, and methodological questions best evaluated in pilot studies. This summer (1991), we propose to conduct a demonstration on about 60 lakes and a pilot on a subset (20-30) of these lakes.

Phase 6 is the full implementation of the core indicators at regional, and eventually national, scales. This phase must balance continuity of methods, to maximize trend detection, with a set of procedures for continually improving the indicators. EMAP-Surface Waters must be open to other sources of information, along with the development and implementation of core indicators. Historical data and information from local experts should continually be sought and used to help interpret impairment and change.

We must make an effort to include citizen monitoring of more waters, or of some waters more frequently. For example, Secchi depth and benthic macroinvertebrate assemblages (U.S. EPA 1990c, Cummins and Wilzbach 1985) offer easily applied and reliable measures of condition. Most importantly, we must develop formal agreements or informal mechanisms to incorporate data from other federal agencies (e.g., The USGS National Water Quality Assessment Program, the USFWS Breeding Bird Survey, and the National Contaminant Biomonitoring Program) and from State museums and water quality and fish and wildlife agencies.

5.0 POPULATION ESTIMATION AND ANALYTICAL APPROACHES

Indicator measurements taken in the field and in the laboratory will be used to index the status and characteristics of each sample waterbody (Section 4.0). Because EMAP sample lakes and streams will be chosen as a probability sample from explicitly defined populations of surface waters, information from these sample sites can be used to infer characteristics of the population from which the sites were drawn. Each sample waterbody will have a weighting factor inversely proportional to its sample inclusion probability. For example, if there were a 1 in 10 chance of selecting a particular lake for field sampling, that lake would represent 10 lakes in a statistical sense. These sample site weightings will be used to extrapolate from index conditions in the set of sample sites to index conditions in the population statistically represented by those sites.

This section describes the type of data analyses planned and some of the approaches we will employ for quantitatively estimating population status, changes, and trends and evaluating uncertainty in those estimates. Section 5.1 describes the basis for population description and extrapolation, in which the main source of uncertainty derives from the selection of sample sites to represent population characteristics. Analytical approaches for detecting population differences and trends are outlined in the last part of Section 5.1. An additional source of uncertainty in population status and trend analysis results from variability in our indicator measurements. We organize the variability of indicator measurements into a conceptual model of variance (5.2.1) and quantify the major components of variability in a number of example indicators (5.2.2). We then evaluate the effect of indicator variability on the robustness of population status descriptions (5.3). Finally, we provide examples of population status, change, difference, and trend analyses (5.4), illustrating the effects of indicator variability and some design and analytical options for accommodating the degree of variability we may encounter in EMAP-Surface Waters.

5.1 ESTIMATING POPULATION CHARACTERISTICS

5.1.1 Population Extrapolation from EMAP Sample Data

Data collected by the sampling effort will be analyzed to estimate, with known confidence, the current status, extent, changes, and trends in the condition of the nation's surface waters. In conjunction with data from other monitoring sources, the data will be examined to uncover correlative associations between indicators of stress and indicators of condition. Status and extent will be portrayed through the use of descriptive statistics and visual displays of spatial pattern. Descriptors of status will include the usual descriptive statistics as well as estimates of the proportion of the population in various categories. The kinds of analytic output to be generated will be dictated by the specific objectives of the various aspects of EMAP, with some constraint imposed by the design. Estimates of such proportions are facilitated by the cumulative distribution function (CDF), and the CDF will be a prime focus of the analyses.

Cumulative distribution functions, such as the one shown for lake pH in Figure 5-1, represent the collective variation of an attribute over a set of units such as lakes. The CDF shows the estimated proportion of lakes in the population having a value of the variable (in this case pH) less than any specified value. The confidence bound is a 95% one-sided confidence bound on the estimated number, scaled by the estimated total number of lakes. Examination of curves such as Figure 5-1, constructed from response indicator survey data, will provide estimates of the number, area, or length of aquatic resources at or below certain condition values for that indicator, and will constitute some of the primary analytical outputs of the EMAP program.

It is necessary that the properties of estimates and analyses be adequate for the EMAP objectives and that any constraint be acceptable in this respect. The algorithms of estimation and analysis are specific for the design and oriented to the nature of the needed output. Descriptive statistics (for example, mean, median, standard deviation) and the CDF will be generated using the Horvitz-

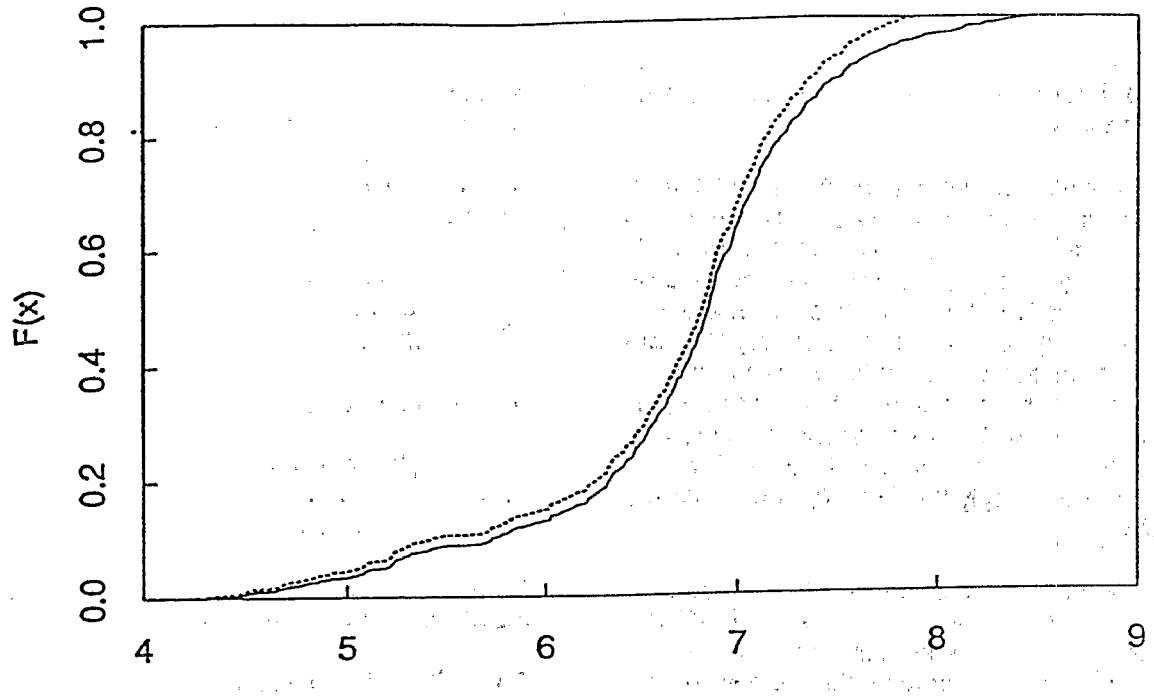


Figure 5-1. Frequency distributions are plotted for any identifiable population. This plot is for the variable pH in Region 1 of the Eastern Lake Survey. The confidence bound has been suppressed above the figure border.

Thompson (HT) estimation formulae (Horvitz and Thompson, 1952). These formulae permit estimation from any probability-based sampling design, and require only specification of the inclusion probabilities. The inclusion probabilities are quantifiable for any probability sample, and the strict requirement that EMAP designs be probability designs has provided the assurance that these formulae can be used.

Inclusion probabilities for EMAP sample units are of two kinds, first and second order. First order inclusion probabilities are simply the probabilities with which the individual sampling units are included in the sample. They must be known for each selected unit. They are generated and archived at the time of sample selection and will be designated by the symbol π_i, referring to the ith sampling unit. Second order, or pairwise, inclusion probabilities are the probabilities with which two specific sampling units are included in the sample. They are designated as π_{ij}, with obvious extension of the notation, referring to the probability of simultaneously including units i and j. Certain design features are required to determine π_{ij}, for example, stratum or cluster identification, and sample size. That information will also be retained as part of the individual data record, so that storage and use of the data are kept uncomplicated.

Estimation formulae are simplified by use of weights rather than inclusion probabilities, where

$$w_i = \frac{1}{\pi_i} \quad \text{and} \quad w_{ij} = \frac{1}{\pi_{ij}}$$

In practice, it is also appropriate to store the weights rather than the inclusion probabilities as part of the data set. The HT formulae are then expressed as:

$$\hat{T} = \sum_{i \epsilon S} w_i y_i \qquad (1)$$

$$\hat{V}[\hat{T}] = \sum_{i \epsilon S} y_i^2 w_i (w_i - 1) + \sum_{i \epsilon S} \sum_{\substack{j \epsilon i \\ j \neq i}} y_i y_j (w_i w_j - w_{ij}), \qquad (2)$$

where y is any attribute, and $\hat{T}y$ is the total of that attribute over any specific identified population. The summation is restricted to the specific set of units, S, in the sample or any subset of the sample defined by a specific population. Estimates

over subpopulations are thus provided by subsetting the sample.

In some instances, interest will focus on totals; in other cases, there will be more interest in means or percentiles. An estimate of the mean is obtained by dividing $\hat{T}$ by the N, number of units in the population. If N is not known, an estimate is provided by setting y=1 and applying (1). Estimators of proportions are obtained using indicator functions. For example, to estimate the proportion of the population with attribute A, set $y_i = 1$, if sample unit i has attribute A, and set $y_i = 0$ otherwise, and apply (1). Then $\hat{T}y/N$ provides the desired estimator.

Basic population parameters will be estimated by some form of (1), and so will be strict HT estimators. The variance estimator (2) is unbiased if all pairwise inclusion probabilities are positive and known exactly. However, in many cases, design features will require approximations in calculating joint inclusion probabilities. The approximation developed by Overton (1987) will be used in most instances. An heuristic derivation of the joint inclusion probability for units i and j can be given as follows. The Tier 2 sample will be selected using a method similar to Madow's (1949) method, in which the sample units are given an order (which may be random, but need not be). Let

$$w_T = \sum_{i=1}^{N} w_i \qquad w_{Tk} = \sum_{i=1}^{k} w_i$$

A random number, r, between 0 and w_T/n is drawn, and unit k is selected if $w_{T,k-1} \le r + w_T/n \le w_{Tk}$, for i = 0,1,...,n-1. For this selection method, $\pi_i = nw_i/w_T$. The probability of drawing unit j given unit i has been drawn should be similar to the probability of drawing unit j in n-1 drawn with the total weight reduced by the weight of unit i, that is, the conditional probability of j given i should be close to $(n-1)w_j/(w_T - w_i)$. Multiplying this expression by π_i gives $\pi_i \pi_j (n-1)/(n-\pi_i)$ as an approximate expression for the joint inclusion probability. The required symmetry in i and j can be achieved by replacing the π_i in the denominator with $(\pi_i+\pi_j)/2$, which is equivalent to subtracting the average weight of units i and j from the total weight in the expression for the conditional probability. Simulation studies (Overton and Stehman 1987; Stehman and Overton 1987a,b) have demonstrated that this approximation and the variance estimator (2) perform well in sampling circumstances such as EMAP.

A more convenient computational form of the approximation is given by

$$\pi_{ij} = \frac{2(n-1)\pi_i \pi_j}{(2n-\pi_i - \pi_j)}, \qquad (3)$$

or, in terms of weights,

$$w_{ij} = \frac{(2nw_iw_j - w_i - w_j)}{[2(n-1)]}. \qquad (4)$$

Sample stratification and cluster sampling require modifications to inclusion probability specification.

5.1.2 Tier 1 Resource Estimates

There are several kinds of Tier 1 estimates, depending on the data on which they are based and the nature of the populations and the selection process. For example, resource inventories are provided by the landscape descriptions and other data on the 40-km^2 characterization hexagons established on the grid points. These Tier 1 estimates will typically be of the areal extent of surveyed resources, by whatever classes have been identified at the Tier 1 classification. For example, estimates of total surface area of lakes could be provided by size class and geographic region, as could estimates of the numbers of lakes. The landscape description units represent 1/16th of the area per grid point, so that w=16, and y is either the area of the lakes or the number of lakes in the size class in the 40-km^2 hexagon.

Let the areal extent, a_r, of a particular class of lakes be identified in the 40-km^2 hexagon. Then, with i now indexing the grid points,

$$\hat{A}_r = 16 \sum_{s} a_{ri}$$

estimates the total area of this class representing any spatial or other subpopulation, and where the summation is taken over the set of grid points where the class occurs. Then the variance of this estimate is estimated by (2) with $w_i = 16$ and w_{ij} from (4) for all i and j.

Estimation of numbers of any class of lakes is possible, since lakes can be unambiguously

represented by points. Point representation is necessary in order to unambiguously count the number of such units in the 40-km^2 hexagons. Given such a count, n_{fi} over the n grid points,

$$\hat{N}_f = 16 \sum_{s} n_{fi}$$

estimates the total number of units in the defined class, with variance again estimated by (2) with $w_i = 16$ and w_{ij} from (4).

Estimates representing resource inventory will then be summarized in a table, similar to the one presented in Table 5-1.

5.1.3 Tier 2 Estimates and Descriptions

The Tier 2 sample consists of a subset of surface waters on which the suite of indicators will be measured. Measurements of additional variables may also be obtained from maps, remote imagery, and other sources for use in higher resolution classification of surface waters. Tier 2 estimates will be of the areal extent of classes defined on the Tier 2 information, and of the numbers of resource units in those populations. A table similar to Table 5-1 can be generated for Tier 2 estimates. In addition, the Tier 2 sample lends itself to more complete population description in terms of the indicator variables that have been measured on the samples of population units.

Estimating formulae for Tier 2 will follow the general forms (1) and (2), but the inclusion probabilities are somewhat more complex, representing the sampling process at both tiers. When possible, selection at Tier 2 will be made in a manner that cancels out the variable inclusion probabilities that were generated at Tier 1, but additional variable probability will be generated by the use of stratification at Tier 2 selection. There will be some restriction on the ability to remove Tier 1 variable probability, and the general formulae will be used throughout.

In addition to inventory estimates, a general characterization of these populations will be by the estimated CDF for the variable, y, of interest. Indicator variables will play such a role. In this, it is necessary to consider only a particular range of the variable y as defining a subpopulation, and estimate the number of lakes, for example, in the population that are in this subpopulation. This is repeated for all values of y in the sample, and the estimated CDF plotted as in Figure 5-1. Such distributions are generated for a variety of y-variables, such as numbers (frequency distributions), areas (areal distributions), lengths, or other attributes. For example, it may be useful to generate the distribution of stream miles on the variable, y, indicating the miles of stream that have a specific attribute.

Table 5-1. Tier 1 Inventory Estimates and Estimated Variance of Estimates are Provided from the Tier 1 Sample for All Resources and Classes of Resources

	Estimated area by class					Estimated number of units				
Resource	1	2	3	....	k	1	2	3	....	k
A										
B										
C										
...										

This can be repeated for any geographic subdivision or other spatial partitioning of the United States, or for any population defined in any manner.

The page of descriptive statistics that was prescribed (Overton 1985) for the National Lake Survey (NLS), and also used for the National Stream Survey (NSS), will also serve EMAP descriptions (see Figure 5-2). Quantiles are interpolated from the fitted distribution function. In the NLS, confidence bounds were defined for the number of units in the class in question, and presented in scaled form.

Extension to EMAP demands greater versatility, and we recommend that the distributions be generated both as numbers and as proportions, and that the confidence bounds be generated accordingly. Then a single page of description might report only one kind of distribution, but in both forms. To keep the two forms distinct, we recommend that the scale of the distribution of numbers be labeled as numbers, but that the plots be of fixed dimension to facilitate comparison. Confidence bounds on the distribution of numbers should continue to be one-sided, as in the NLS (Overton 1985), and ascending and descending analyses should be provided, depending on the variable. The ascending analyses provide upper bounds on the numbers of resource units having values of the variable below a particular value, and descending analyses provide upper bounds on the numbers of units having values above a particular value. Confidence bounds on proportions of numbers of units will usually be appropriately provided by exact binomial bounds, and these are to be two-sided, with descending analyses unnecessary. Binomial bounds can be used only when the variable probabilities have been eliminated by the selection process at Tier 2; otherwise, it may be necessary to use ratio variances for these confidence bounds. Combined strata will involve means of binomials, and will also not be suitable to exact binomial bounds.

Distributions of area or length will follow the direct estimation methodology (Overton 1985) but proportions of areas and lengths must be treated as ratios, and so will require special treatment in generating confidence bounds.

5.1.4 Analytical Approaches for Detecting Population Differences, Changes and Trends

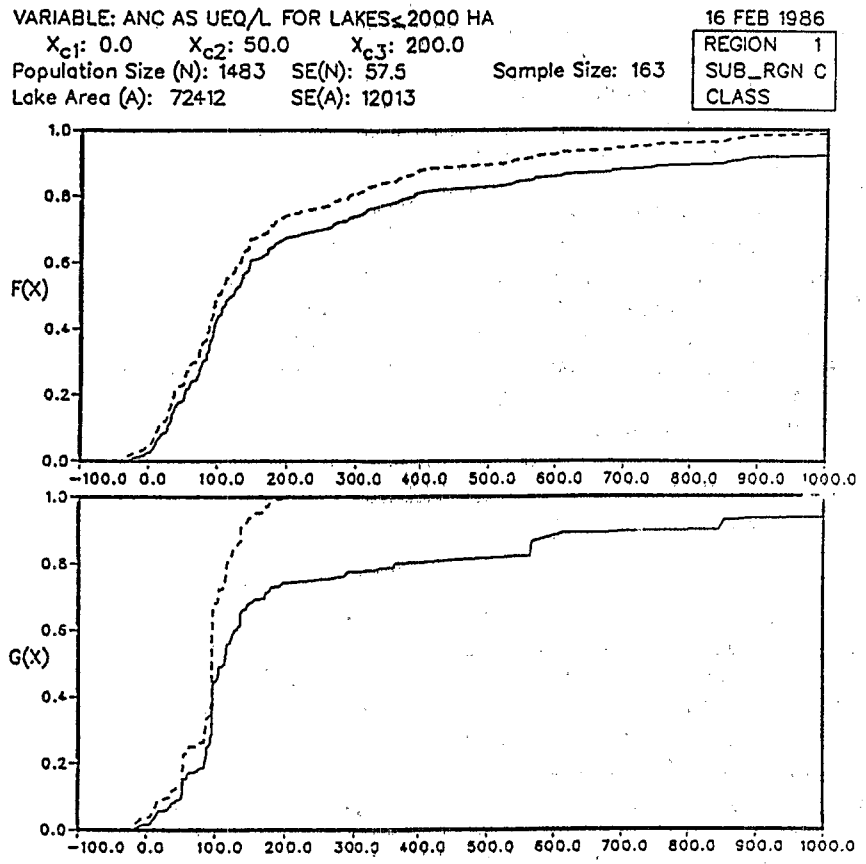
An important feature of EMAP-Surface Waters will be the description of differences among subpopulations and changes in subpopulations over time. For example, it may be important to illustrate or describe differences among lakes influenced by different land use patterns or practices, or to determine whether the condition of different types of lakes differs (e.g., drainage versus seepage lakes, or lakes versus reservoirs). Also, during the early years, it will be important to determine whether there are detectable year-to-year differences within subpopulations of lakes, especially if changes in the magnitude of stressors are occurring or are anticipated. As more data are obtained, it will become feasible to detect trends in a variety of subpopulations.

An assortment of techniques will be used in the examination of EMAP data for change and trend. An extensive range of standard statistical techniques is applicable. Linear model techniques, such as analysis of variance and regression will be used with various assumptions regarding spatial and temporal variance components, statistical independence, explanatory variables, nature of trend, and nature of change. Nonparametric alternatives, based on ranks or signs of differences, will also be used (e.g., Loftis et al. 1989).

Factors that influence the ability to detect differences, changes, and trends include:

- Inherent variation
- Sample size
- Level of significance
- Level of power
- Size of the differences to be detected
- Temporal and spatial autocorrelation

Inherent variation refers to the natural variation that occurs within and among systems, as well as the variation introduced during the measurement process, as described in Section 5.2. Sample sizes of 50 to 100 are expected to be adequate for most subpopulation descriptions. Relatively little reduction in uncertainty occurs with increased sample sizes above 50 to 100 (Linthurst et al. 1986). The level of significance (α) refers to the probability of correctly detecting a change when a change is present. In most studies, α is set at the 0.05 level, implying that if a change has occurred, we would expect to detect it 95% of the time. It may be appropriate to relax this constraint somewhat for EMAP purposes.



Min: −31.00 Q_1: 52.05 Q_2: 94.20 Q_3: 146.20 Q_4: 387.74 Max: 4046.60

Median: 119.88
Mean: 323.69
Std. Dev.: 557.05

PROPORTIONS AND NUMBERS BELOW THE VALUE OF X_c

P_{c1}: 0.024	N_{c1}: 35	N_{cu1}: 63
P_{c2}: 0.177	N_{c2}: 262	N_{cu2}: 334
P_{c3}: 0.676	N_{c3}: 1002	N_{cu3}: 1100

Min: −31.00 Q_1: 84.14 Q_2: 95.86 Q_3: 128.15 Q_4: 363.31 Max: 4046.60

Median: 112.94

PROPORTIONS AND AREA BELOW THE VALUE OF X_c

G_{c1}: 0.016	A_{c1}: 1152	A_{cu1}: 2664
G_{c2}: 0.090	A_{c2}: 6501	A_{cu2}: 9294
G_{c3}: 0.744	A_{c3}: 53894	A_{cu3}: 72587

Figure 5-2. Example of estimated distribution plots with upper confidence bounds, generated both as numbers (upper plot) and area (lower plot). Examples are given of descriptive statistics used in the National Lake Survey that could also be used for EMAP: F(x) and G(x) distributions of ANC for the target population of lakes (≤2,000 ha) samples in Region 1 (Northeast), Eastern Lake Survey - Phase I. Source: Linthurst et al. (1986).

An important consideration in the selection of techniques for change and trend detection will be the sensitivity of the tests, or, in statistical terms, the power of the tests. Statistical power is the probability that a test of an hypothesis will be rejected given that the hypothesis is false. Generally, power increases with increasing departure from the hypothesis, that is, a large change is more likely to be detected than a small change. A statistical test of a particular hypothesis can be characterized by specifying the size of the test (probability of rejecting a true hypothesis) and its power function (probability of rejecting a false hypothesis as a function of degree of departure from the hypothesis). Where choices among alternative methods of change and trend detection exist, the more powerful test will be preferred.

Power will also be used to evaluate the adequacy of the design, and of the sample sizes in various subpopulations. The evaluation will be in terms of describing the probability of detecting a change or trend of a given magnitude. If the magnitude of change that is detectable with, say 80% probability, is unacceptably large for some subpopulation, then the sample size for that subpopulation may have to be increased. Provisions for doing so are incorporated in the general EMAP design.

One routine method for determining detectable differences among subpopulations is a comparison of mean values with a Z-statistic. A general equation that can be used to estimate MDC (minimum detectable change) is:

$$MDC = (Z_{\alpha/2}+Z_{\beta}) * s * [2(1-r)/n]^{1/2} * [1 + \rho(n-1)]^{1/2}$$

where: $Z_{\alpha/2}$ and Z_{β} are tabulated "Z" values for various levels of significance (α) and power (β)

s = sample standard deviation

r = temporal autocorrelation

n = sample size

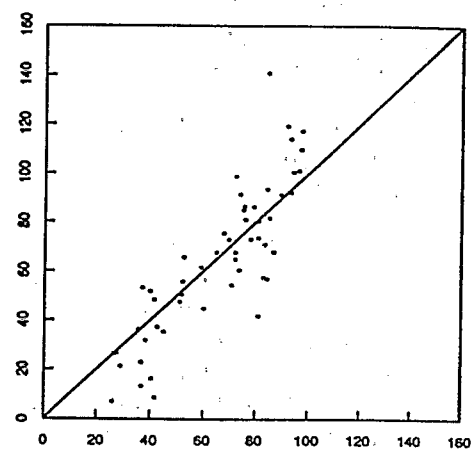
ρ = spatial autocorrelation

for a two-sided test for differences; if a one-sided test is sufficient, Z_{α} can be used.

However, EMAP will also use techniques that do not focus solely on change in mean value or central tendency. Some changes might be of great interest but would not have large effects on mean values, such as changes in variance. For example, one sign that an ecosystem is stressed is an increase in variability. Since the CDF will be the primary tool for summarizing and presenting population descriptions, analyses of change and trend will utilize the CDF, or some property thereof. An example is the examination of the quantiles for differences. The test is carried out by using the CDF of the pooled subpopulations to define classification criteria (e.g., quintiles of the combined population) and doing a chi-squared test of homogeneity. Figure 5-3 illustrates this approach. Such a test is sensitive to changes in location as well as changes in scale or shape of the distribution. The test is applied in the conventional manner if both samples have approximately uniform inclusion probability. Extension is needed if there is appreciable variation in inclusion probability in either sample.

Comparison of subpopulations also provides a general technique for multivariate comparison. A population can be classified on several variables or attributes, and the CDF of the classes compared by the chi-squared test prescribed above. In the case of repeat observations of the same population units, a more powerful test can be used. Again, a procedure based on a chi-squared test and sensitive to changes in location, scale, and shape can be used. The test was used by Overton (1989) to examine repeat measurements of lake chemistry for change. The first step in the procedure is a sign test of the differences of the two observations. Then the combined distribution is used to determine median and quintile values. Chi-squared tests of counts from each sample in the class intervals defined by combined population percentiles are used to test for changes in scale and shape. An important aspect of EMAP's interpenetrating design is that it will achieve its full potential power to detect change only after repeat visits to all sites, that is, after two complete cycles.

Repeat visits to a site permit a paired analysis that essentially eliminates the population variation component. Thus, ability to detect change will increase greatly in years 5 through 8 of the sampling. Similarly, the power to detect a persistent trend will continue to increase as more years of data become available.



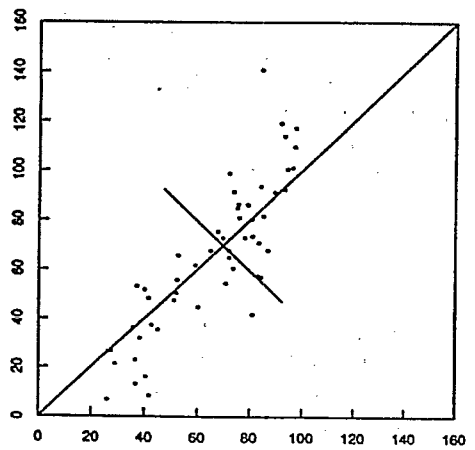
OLD ANC CLUSTER II

FALL 1986 VS. FALL 1984

THE TOTAL NUMBER OF POINTS PLOTTED: 53
NUMBER OF POINTS ON DIAGONAL LINE: 0

CHI-SQUARE = 0.01886793
1 DF CHI-SQUARE CRITICAL VALUE = 3.84

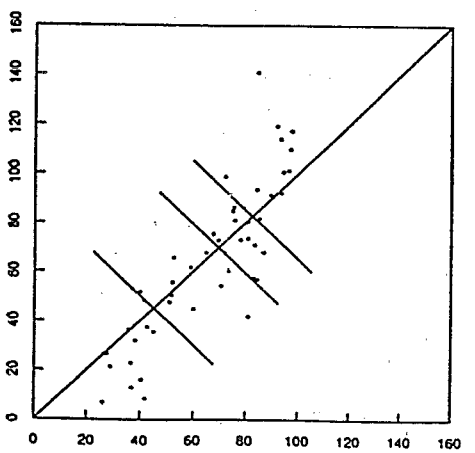
26	
	27



ADDITIONAL POINTS WITH SUM EQUAL TO MEDIAN: 1

CHI-SQUARE FOR INDEPENDENCE = 4.923077
1 DF CHI-SQUARE CRITICAL VALUE = 3.84

9	17
17	9



ADDITIONAL POINTS WITH SUM EQUAL TO A QUARTILE: 0

CHI-SQUARE FOR INDEPENDENCE = 10.15385
3 DF CHI-SQUARE CRITICAL VALUE = 7.81473

3	6	6	11
10	7	7	2

Figure 5-3. An example of comparing population distributions using the chi-square test. ANC values are for samples taken in fall of 1984 and 1986 from lakes during Eastern Lake Survey - Phase I and Phase II. The expectation is for points to fall on the 1:1 line if no difference exists. Specific portions of the distributions can be evaluated by dividing the analysis by the median, quartiles, or other percentiles.

The development of techniques for trend detection for regional subpopulations is at a relatively primitive state and much will be done over the next several years. Loftis et al. (1989) evaluated several techniques for trend detection in individual waterbodies. They recommend two nonparametric tests to use with seasonal sampling: an analysis of covariance on ranks and a seasonal Kendall test. They briefly addressed extending these tests to detecting trends in regional mean values, but that issue remains unresolved.

Trend detection methods that are specific to the interpenetrating design structure of EMAP are being developed. In particular, methods that account for temporal correlation and its effect on the power of tests of trend are needed.

5.2 COMPONENTS OF VARIANCE IN REGIONAL POPULATION SAMPLING

5.2.1 Variance Model

For any characteristic of interest to be sampled from a surface water population, uncertainty in population estimates may result from a number of sources, ranging from sample selection to analytical error. A useful approach that provides the information necessary for optimizing effort at various levels of sampling is to consider the magnitudes of variation in the environmental characteristic over a range of spatial and temporal scales. In doing this, it is important to distinguish between the uncertainty in estimating population statistics (e.g., CDFs, medians, quartiles) and the actual magnitude of variation in the environmental characteristic itself.

For example, consider an indicator variable such as fish species richness in a population of lakes within the Northeast region. We are interested in describing the regional variation and trends in the true species richness within these lakes. If we had perfect knowledge of the true species richness in each lake within the Northeast, and species richness did not vary within the lakes, we could describe with virtual certainty (by a census) the distribution of species richness values in the regional population of lakes. If, rather than taking a census of the lakes, we look only at a subset (a "sample") of the total number of lakes, there is uncertainty in our estimates of the true population statistics introduced by the sampling process. Section 5.1 describes the procedures for quantifying the uncertainty in estimating statistical parameters of the true population distributions (CDF's, medians, quartiles, etc.) that results from selecting a sample of lakes, rather than taking a census of the population. Those confidence bounds do not explicitly incorporate uncertainty in measuring the indicator values in each individual sample lake.

Unlike the idealized situation in the previous paragraph, species richness in each of our sample lakes might actually vary over annual, seasonal, and short-term time scales. We can circumvent this problem somewhat by targeting an index time and position in each lake, but even these index values of species richness may vary within this temporal and spatial window and may vary cyclically from year to year. To complicate the matter, our measurements of species richness may vary even if true species richness in each lake remains invariant. Our measurements are not perfect because of variations in fish sampling efficiency, gear selectivity, variations in the location and behavior of fish at different times, and uncertainties in taxonomic identification.

We may describe variation in an index measurement (e.g., fish species richness or Secchi disk transparency measured at a defined window of space and time) as being derived from a number of sources:

- True differences in index values among surface water sample units (lakes or stream segments) across a region: e.g., fish species richness differences in lakes or streams across the Northeast.

- Among-year temporal variation in each surface waterbody in the absence of trends: e.g., normal fluctuations in fish species richness between summer index periods of different years occurring within single lakes or stream segments.

- Index sampling variation within each waterbody, including:

  - a within-season temporal component: e.g., variation in fish species richness within a late summer index time period, and

  - a spatial component: e.g., spatial variation in species richness at a given time but in different potential index locations

within a mapped stream segment or lake identified as a sample unit. A subtle but important distinction is that this is not necessarily the spatial variation of the index value as measured in all possible locations in the lake or stream segment, but only those locations that might be sampled within the index definition. For some variables, this might include the whole lake, for others only positions within the profundal zone, for example.

- Measurement variation (error): Here we are defining measurement error as the variation in the estimation of an indicator (e.g., species richness) that might be observed if the measurement were replicated at a single time and place (the "single place" may be much larger than a single point). Measurement error derives from variation in sample collection technique, handling, analysis, or species identification.

Application of the index concept eliminates the need to include variability at times and places outside of the index sampling window or location. Consequently, seasonal variation outside the index time period and spatial variation outside the defined index location are not included in our variance modelling. Those variances, however, have great bearing on the selection of the indicator measurements and the chosen definition of the index time and location. The spatial and temporal variation not included in the model is important in subsequently limiting the interpretation of population statistics describing index conditions in a region. A clear example of this is found in the NSS (see subsection 6.1.2.1), in which the index sampling window specifically excluded episodic variability in stream chemistry associated with storm flows. Interpretation of the biological relevance of population distributions of index chemistry benefitted by subsequent consideration of the episodic characteristics of streams as they relate to measured index chemistry (Kaufmann et al. 1988, 1989, Eshleman 1988, Baker et al. 1990).

Based on the components described, a generalized hierarchical model of variance in index values within a population of surface waters may be proposed. For an indicator variable, y:

$$y = \mu + \epsilon_{unit} + \epsilon_{annual} + \epsilon_{index\text{-}temporal} + \epsilon_{index\text{-}spatial} + \epsilon_{measurement}$$

where each of the component deviations (errors) are normally distributed with expected means of zero and respective variances of σ^2_{unit}, σ^2_{yr}, $\sigma^2_{in\text{-}t}$, $\sigma^2_{in\text{-}sp}$, and σ^2_{m}. Raw measurement values may have to be transformed (e.g., log transformation) to satisfy these assumptions. The model relating the total variance σ^2_{tot} to the component variances is then additive:

$$\sigma^2_{tot} = \sigma^2_{unit} + \sigma^2_{yr} + \sigma^2_{in\text{-}t} + \sigma^2_{in\text{-}sp} + \sigma^2_{m}$$

As discussed earlier (5.1), descriptions of the condition of populations of lakes and streams will rely primarily on cumulative distribution functions and statistics derived from them (means, medians, quantiles, and proportions). Each year these descriptors will summarize the status of various regional subpopulations. Differences among and changes within regional subpopulations will also be identified, and as EMAP accumulates data over time, trend detection will play an increasing role.

The various components of variation identified on the previous page influence the certainty with which the descriptions can be made. An important goal of EMAP is to quantify the uncertainty with which condition, differences, changes, and trends can be estimated. The various components of variation take on differential importance depending upon the types of descriptors to be made. We have begun to evaluate the role of these components in an attempt to describe how well we expect various indicators to perform and to develop explicit performance criteria for the indicators (data quality objectives).

5.2.2 Indicator Variance Magnitudes Estimated from Existing Data

5.2.2.1 Alternatives for Comparing Variances

Information on the magnitudes of variance indicator variables of likely interest to EMAP-Surface Waters is summarized in Tables 5-2 to 5-5. We will continue to acquire data of this type on candidate indicators. Values of S^2 in the tables are sample estimates of population values of σ^2. The tabulated information is from a variety of regional sampling efforts as cited in the tables.

Table 5-2. Coefficients of Variation (CV) of Surface Water Indicators -- A Number of Examples from Lake and Reservoir Studies

Variable[a]	Study[b]	CV_m	CV_{in-t}	CV_{in-sp}	CV_{yr}	CV_{unit}	Mean
Secchi	Vermont	9.2%	12%	na	12%	43%	4.9 m
Secchi	Minnesota	na	29%	na	17%	66%	1.7 m
Secchi	ELS-II	na	16%	inc[e]		47%	3.7m
Chl a	Vermont	30%	26%	na	27%	72%	4.3 µg/L
Chl a	Minnesota	na	49%	na	44%	87%	17.5 µg/L
Chl a	Knowlton	12%	45%	na	20%	76%	33.0 µg/L
TP	Vermont	13%	23%	na	26%	62%	24.8 µg/L
TP	Minnesota	na	41%	na	25%	81%	62 µg/L
TP	Knowlton	10%	27%	na	13%	103%	63 µg/L
TP	ELS-II		21%	inc[e]		259%	8.7 µg/L
SO_4	ELS-II	2.3[c]	9.0%	inc[e]	12%	42%	124 µeq/L
SOBC	ELS-II	3.3[c]	3.3%	inc[e]	19%	72%	374 µeq/L
ANC	ELS-II	14[c]	12%	inc[e]	26%	144%	61 µeq/L
pH (fall)	ELS-II	0.10[c]	0.06	inc[e]	0.13	0.76	5.9
DI-pH (core)	PIRLA-II	0.04[c]	time-integ.[d]	0.10	na	0.76[f]	5.9[f]
DI-pH (surficial core)	PIRLA-II	0.04[c]	time-integ.[d]	0.21	na	0.76[f]	5.9[f]

[a] Secchi = Secchi disk depth, Chl a = chlorophyll-a, TP = total phosphorus, SO_4 = sulfate, SOBC = sum of base cations, DI-pH = diatom-inferred pH; for pH variables, we report SD rather than CV; 100 SD/grand mean; for pH variables, we report SD rather than CV.

[b] Vermont lakes [n=17 (TP); n=63 (Chl); n=63 (Secchi)] data from Eric Smeltzer (pers. comm.); Minnesota lakes [n=69 (TP); n=57 (Chl); n=69 (Secchi)] data from Steve Heiskary (pers. comm.); ELS-II data (n=145) from Herlihy et al. (1991); Knowlton data are as reported by Knowlton et al. (1984) for eastern lakes and reservoirs (n=186); PIRLA-II data for ELS-II Adirondack lakes from Dixit et al. (1990).

[c] These measurement CVs were independently calculated and were not included in the nested analysis of variance.

[d] Diatom cores integrate conditions over the whole year, effectively eliminating within-year temporal variability.

[e] Spatial variance component CV^2_{in-s} for mid-lake sampling is included within the temporal component CV^2_{in-t}.

[f] For diatom-inferred pH, we have calculated these means on the basis of the Adirondack lakes population from ELS-II.

Table 5-3. Coefficients of Variation (CV)[a] of Surface Water Indicators -- A Number of Examples from Stream and River Studies

Variable[b]	Study[c]	CV_m	CV_{in-t}	CV_{in-sp}	CV_{yr}	CV_{unit}	Mean
IBI-stream	Ohio	na	12%	na	10%	37%	32 units
IBI-river	Ohio	na	15	na	13%	28%	30 units
Sp.-rich stream	Ohio	na	16%	na	17%	50%	14 species
Sp.-rich river	Ohio	na	19%	na	24%	27%	15 species
Turb.	NSS	9.7%[d]	41%	46%	na	127	3.8 NTU
T-dP	NSS	4.2%[d]	54%	89%	na	232	59 μg/L
NO_3	NSS	3.5%[d]	13%	14%	na	196	66 μeq/L
SO_4	NSS	1.0%[d]	3.5%	4.8%	na	119%	282 μeq/L
Cond.	NSS	0.56%[d]	3.5%	6.4%	na	86%	115 μS/cm
ANC	NSS	1.2%[d]	1.5%	2.3%	na	95%	830 μeq/L
pH	NSS	0.035[d]	0.43	0.69	na	0.93	6.6 pH units

a 100 SD/grand mean; for pH variables, we report SD rather than CV.

b IBI = Index of Biotic Integrity (Karr et al. 1986) in wadable streams and rivers sampled by boat; Sp.-rich = total number of fish species in wadable stream and large river sampled sites; Turb. = turbidity; T-dP = total dissolved phosphorus; NO_3 and SO_4 are filtered samples; Cond. = specific conductance; ANC = acid neutralizing capacity by Gran titration; pH = pH determined in closed-headspace sample.

c Ohio = Ohio Department of Environment stream and river data provided by Chris Yoder and Ed Rankin (pers. comm.); NSS = National Stream Survey data on 500 Mid-Atlantic and Southeast streams [Kaufmann et al. 1988 (in press)].

d These measurement CVs were independently calculated from NSS measurement variances reported by Cougan et al. (1988) and population means in column 8--a procedure not within nested analysis of overall variance.

Table 5-4. Components of Indicator Variability as a Percentage of Among-Sample Unit Variance (S^2_{unit}) -- Examples from Lake and Reservoir Studies

Variable[a]	Study[b]	S^2_{m}/S^2_{unit}	$S^2_{in\text{-}t}/S^2_{unit}$	$S^2_{in\text{-}s}/S^2_{unit}$	S^2_{yr}/S^2_{unit}	SD_{unit}
Secchi	Vermont	4.6%	7.2%	na	7.5%	(2.1 m)
Secchi	Minnesota	na	19%	na	7%	(1.1 m)
Secchi	ELS-II	na	12%	inc[e]	na	(1.8 m)
Ln-Chl a	Vermont	17%	13%	na	14	~3 μg/L
Ln-Chla	Minnesota	na	32%	na	25	~15 μg/L
Log_{10}-Chl a	Knowlton	0.14%	1.8%	na	0.37	~25 μg/L
Ln-TP	Vermont	4.4%	14%	na	18	~15 μg/L
Ln-TP	Minnesota	na	26%	na	9.5%	~50 μg/L
Log_{10}-TP	Knowlton	0.058%	0.42%	na	0.098%	~65 μg/L
TP	ELS-II	na	0.67%	inc[e]	na	(23 μg/L)
SO_4	ELS-II	0.30[c]	4.5%	inc[e]	8.2%	(53 μeq/L)
SOBC	ELS-II	0.22[c]	0.21	inc[e]	7.2%	(268 μeq/L)
ANC	ELS-II	0.91[c]	0.73	inc[e]	3.4	(87 μeq/L)
pH (fall)	ELS-II	1.7[c]	5.3	inc[e]	13	(0.77 pH unit)
DI-pH (core)	PIRLA-II	0.3[c]	time-integ.[d]	1.7	na	(0.77 pH unit)[f]
DI-pH (surficial core)	PIRLA-II	0.3[c]	time-integ.[d]	7.5	na	(0.77 pH unit)[f]

[a] Secchi = Secchi disk depth, Chl a = chlorophyll-a, TP = total phosphorus, SO_4 = sulfate, SOBC = sum of base cations, DI-pH = diatom-inferred pH.

[b] Vermont lakes [n=17 (TP); n=63 (Chl); n=63 (Secchi)] data from Eric Smeltzer (pers. comm.); Minnesota lakes [n=69 (TP); n=57 (Chl); n=69 (Secchi)] data from Steve Heiskary (pers. comm.); ELS-II data (n=145) from Herlihy et al. (in press); Knowlton data are as reported by Knowlton et al. (1984) for eastern lakes and reservoirs (n=186); PIRLA-II data for ELS-II Adirondack lakes from Dixit et al. (1990).

[c] These measurement variances were independently calculated and were not included in the nested analysis of variance.

[d] Diatom cores integrate conditions over the whole year, effectively eliminating within-year temporal variability.

[e] Spatial variance component $S^2_{in\text{-}s}$ for mid-lake sampling is included within the temporal component $S^2_{in\text{-}t}$.

[f] For diatom-inferred pH, we have calculated these SD's on the basis of the Adirondack lakes population from ELS-II.

Table 5-5. Components of Indicator Variability[a] as a Percentage of Among-Sample Unit Variance (S^2) -- Examples from Stream and River Studies

Variable[b]	Study[c]	$S^2_{m/S^2_{unit}}$	$S^2_{in\text{-}t/S^2_{unit}}$	$S^2_{in\text{-}s/S^2_{unit}}$	$S^2_{yr/S^2_{unit}}$	SD_{unit}
IBI-stream	Ohio	na	11	na	7.4	(12 unit)
IBI-river	Ohio	na	29	na	20	(8.3 unit)
Sp.-rich stream	Ohio	na	11	na	12	(7.1 species)
Sp.-rich river	Ohio	na	50	na	75	(4.1 species)
Turbidity	NSS	0.59[d]	11	13	na	(4.87 NTU)
T-dP	NSS	0.03[d]	5.4	15	na	(138 μg/L)
NO_3	NSS	0.03[d]	0.44	0.52	na	(130 μeq/L)
SO_4	NSS	0.0075[d]	0.086	0.16	na	(337 μeq/L)
Conductance	NSS	0.0042[d]	0.16	0.55	na	(99 μS/cm)
ANC	NSS	0.017[d]	0.025	0.060	na	(790 μeq/L)
pH	NSS	0.0014[d]	0.26	0.54	na	(0.93 unit)

[a] 100 SD/grand mean; for pH variables, we report SD rather than CV.

[b] IBI = Index of Biotic Integrity (Karr et al. 1986) in wadable streams and rivers sampled by boat; Sp.-rich = total number of fish species in wadable stream and large river sampled sites; Turb. = turbidity; T-dP = total dissolved phosphorus; NO_3 and SO_4 are filtered samples; Cond. = specific conductance; ANC = acid neutralizing capacity by Gran titration; pH = pH determined in closed-headspace sample.

[c] Ohio = Ohio Department of Environment stream and river data provided by Chris Yoder and Ed Rankin (pers. comm.); NSS = National Stream Survey data on 500 Mid-Atlantic and Southeast streams [Kaufmann et al. 1988 (in press)].

[d] These measurement variance ratios were calculated using independently derived NSS measurement variances reported by Cougan et al. (1988) divided by S^2_{unit}--a procedure not within the nested analysis of overall variance.

Some of the data are from probability sampling [e.g., National Surface Water Survey (NSWS)], others are not. The purpose of these tables is to provide a somewhat representative range of indicator precision and regional heterogeneity upon which we can base preliminary decisions regarding the necessary temporal and spatial frequency of sampling efforts. In order to express variance information on comparable grounds, we provide the information first in terms of coefficients of variation (100 X SD/grand mean) for indicator data in lakes (Table 5-2) and streams (Table 5-3). Using coefficients of variation, the magnitudes of the variance components can be compared with the general level of the indicator values in the sample or population of surface waters. The second pair of tables (5-4 and 5-5) describes the same information, but in terms of relative variances (the variance components are expressed as percentages of the among-sample unit variance). The among-sample unit variance, S^2_{unit}, not shown in the tables, would have a value of 100% in every row. Rather than providing S^2_{unit}, we provide its square root, SD_{unit}, expressed in the appropriate indicator measurement units, to facilitate visualizing the magnitudes of variance. The component variances were calculated using a nested analysis of variance, so the component variances add up to the total population variance (i.e., sample unit variance does not include that amount due to index temporal and spatial differences).

5.2.2.2 Coefficients of Variation

Coefficients of variation (CV) among lakes and streams in a region tend to vary from 28% to 232% of the mean, with biological variables tending to be less variable than chemical variables (Tables 5-2 and 5-3). This finding may simply reflect the generally smaller size (and presumably more homogeneous characteristics) of regions examined in the biological studies. It may also reflect the natural constraints on species pools, or the intentional constraints placed on the range of indices such as IBI.

With one exception, CV's for among-year variability range between 10% and 26%. CV's for spatial variation within the index location were available only for NSS and ELS-II data. For NSS data, $CV_{in\text{-}sp}$ varied from 2% to 5% for the higher precision chemical variables at the bottom of the table, and from 14% to 46% for "noisier" variables such as nitrate, phosphorus, and turbidity. Spatial variability of mid-lake profundal samples in the ELS-II is included in $CV_{in\text{-}t}$, which is a measure of temporal variability within the index time period and ranges from 3.3% for the sum of base cations (SOBC) to 21% for total phosphorus. Quick examination of the phosphorus mean shows that levels were extremely low in the ELS-II lakes--elevating the CV's for that variable. Within-index period temporal CV's were mostly between 2% and 10% for the precise chemical variables and from 12% to 20% for most of the biological variables. The highest CV's for index temporal variability were chlorophyll-*a* and variables such as phosphorus and stream turbidity, which are biologically active or responsive to stream flow.

Measurement CV's were below well below 3% for all the precise group of chemical variables, with the exception of ANC in ELS-II. The CV for this variable is inflated by the fact that ANC can take negative values, and also that it does not vary widely within the ELS-II region. Coefficients of measurement variation for other chemical variables are mostly under 13%. Measurement CV's for chlorophyll-*a* ranged from 12% to 30%, and probably account for a moderate fraction of the high value (49%) observed within season in the Minnesota data, which in our data set did not enable calculation of measurement variability.

5.2.2.3 Relative Variances

Variance components expressed as a percentage of among-sample unit variance in a region (Tables 5-4 and 5-5) are useful because they allow calculation of the degree of distortion and uncertainty that these components cause in the descriptions of population means and distributions (see Section 5.3). To sum the percent relative variances across rows in the relative variance tables, set S^2_{unit} (the variance across sample units in the region) at 100% and add it to the other variance components. The total variance is then always a percentage greater than 100% that expresses the combined amount of index measurement variance in excess of true differences among sample lakes and streams across the region. For example, a variance sum of 120% means that index measurement variances equal to 20% of S^2_{unit} contribute to the total observed variance observed in the regional sample.

Among-year variability is less than 10% of among-sample unit variance for most indicators other than phosphorus, chlorophyll-*a*, and the

stream biological measures, with the exception of small stream IBI (column 6 of Tables 5-4 and 5-5). The river fish measures (20% to 75%) and Minnesota chlorophyll (25%) have the greatest between-year variance (20% to 75%) among the studies we examined. It seems likely that a large proportion of what is interpreted here as biological variability may simply result from differences in the location of mobile fish populations and the capture efficiency of sampling gear. Alternately, the high variance of river fish measurement components relative to among-river differences might also reflect a narrow range of meaningful differences in "condition" across the regional set of large Ohio rivers sampled.

Spatial variance within the profundal zone of northeastern lakes sampled by the ELS-II was included as part of the within-fall index period variance, and for all the variables presented in Table 5-4, those combined variances were less than about 5% of the variation among lakes. Diatom-inferred pH from surficial cores was more variable over the lake bottom (7.5%) than was pH for deeper cores (1.7%). Among the example stream studies, information on the spatial component of variance in streams was available only for physico-chemical data from the National Stream Survey (Table 5-5). These upstream/downstream differences were surprisingly small, amounting to less than 1% of the among-stream variance except for phosphorus (15%) and turbidity (13%). It is likely that the within-sample unit spatial component is a large part of the variance for biological measurements (e.g., fish, macroinvertebrates) in streams, rivers, and lakes). Stratification or targeting of specific physical habitat types will probably be necessary to reduce the effect of large spatial variability. This problem is discussed in Sections 6.3 and 6.4 and, with physical habitat characterization, is a major area of productive pilot research that is likely to result in major gains in the precision of biological indices.

Temporal variability of most chemical variables within the index sampling period is less than 5% of among-lake or stream variability across the regions examined in Tables 5-4 and 5-5. The chemical constituents that are biologically active and sensitive to stream discharge, such as phosphorus and turbidity, appear in most cases to be more variable (5% to 26%). Chlorophyll variability differs substantially among regions, a result that may reflect our failure to consistently isolate observations from an appropriate index sampling period in our analysis of some of the data sets. Secchi disk transparency and small stream fish measurements (species richness and IBI) have moderately low within-index temporal variability in some regions (7% to 11%). Variability of large river fish species richness and IBI (50% and 29%) was much higher than that for smaller streams, and probably reflects spatial habitat sampling variations, capture efficiency, and fish population mobility rather than true variation in condition of the fish community within the summer index time period.

With one exception, the measurement component of variability is less than 2% of the among-unit variance for all of the chemical variables shown in Tables 5-4 and 5-5. In most cases it is extremely small, less than 0.1%. The exception is total phosphorus in Vermont lakes (4.4%), which was somewhat higher in relative terms because the concentrations involved were quite low (population mean 25 μg/L) and the among-lake variability was also low (SD = 15 μg/L). Measurement variability for Secchi disk transparency (field duplicate measurements) in Vermont lakes (4.6%) was approximately the same as that for total phosphorus. The measurement variability of chlorophyll analyses in Vermont lakes (17%) was higher than that reported by Knowlton et al. (1984) (Table 5-5), probably reflecting the relatively low concentrations and among-lake variation in the Vermont lakes.

5.3 DESIGN CONSIDERATIONS FOR ESTIMATING POPULATION STATUS

5.3.1 Effects of Index Variability on Estimates of Population Status

5.3.1.1 Theoretical Basis

Section 5.1 describes the algorithms for extrapolating population means and their standard errors from index values at surface water sample sites. These extrapolation procedures do not separately account for variance in the index values within individual surface water sampling units. Rather, they estimate the total sampling variance in the population. If large imprecision exists in the estimation of the index values for each sample unit, we may be misled into believing that there is a great deal of regional heterogeneity among surface waters, when in fact a large portion of the population variance is due to index measurement imprecision rather than true differences among

sample sites across a region. In this context, we consider index measurement error to be in the broadest sense--to include all components of variability besides that due to the regional variation among sample units themselves. Define for this discussion:

$$\sigma^2_{index} = + \sigma^2_{yr} + \sigma^2_{in\text{-}t} + \sigma^2_{in\text{-}sp} + \sigma^2_{m}$$

Total population sampling variance then becomes:

$$\sigma^2_{tot} = \sigma^2_{unit} + \sigma^2_{index},$$

and the variance of the population mean is:

$$\sigma^2_{mean} = \frac{\sigma^2_{tot}}{n_{unit}} = \frac{\sigma^2_{unit} + (\sigma^2_{index}/n_{rep})}{n_{unit}}$$

where n_{unit} is the number of sample units and n_{rep} is the number of temporal (or spatial) index measurement replicates. It is clear from this expression that if determining the population mean is the sole objective, greater precision can always be gained by increasing the number of sites, rather than the number of replicates, no matter what the magnitude of index variability. Cost differences come into play to determine the most economical allocation of sampling effort between sample units and replications (see Section 5.3.3). Furthermore, without some replication, even if it is not systematic (i.e., across all sample sites), we will know what the population variance is, but we will not know whether it is due to the imprecision of index determinations on each site or to true differences among sites within a region.

In many cases in EMAP, the population mean values of indicators will not be of greatest interest in describing the status of a resource. EMAP will often rely more heavily on the shape of the regional population distributions and other statistics, such as percentiles and estimates of the percentage of surface waters with indicator values above or below nominal or subnominal values. Overton (1989) assessed the effects of index measurement variation (σ^2_{index}) on estimated distribution functions describing status, as provided by EPA's National Surface Water Survey (Linthurst et al. 1986, Landers et al. 1987, Kaufmann et al. 1988, in press). As does the EMAP design, the surveys employed probability sampling from well-defined populations of waterbodies, selected from well-defined and explicit frames. Rigorous inference is permitted by these probability designs. Ordinarily, in sample surveys, the effect of index measurement variation that has an expectation of zero is not critical, as such error contributes no bias to the usual estimates of population means or totals, and the contribution of such error to the variance of such estimates is ordinarily low. However, the distributions of the biological or chemical variables in surveys such as the NSWS and EMAP can be biased by substantial variance in the measured indicator values for each sample waterbody.

In the context of sample surveys, cumulative distribution functions (CDF), such as the one shown for lake pH in Figure 5-1, represent the collective variation of an attribute over a set of units such as lakes. Examination of such CDF curves, constructed from response indicator survey data, will provide estimates of the number, area, or length of aquatic resources at or below certain condition values for that indicator, and will constitute some of the primary analytical outputs of the EMAP program. Index measurement variance increases the collective variation in the data, leading to bias in estimates of the number, area, or length of waterbodies having values of response indicators above or below specified criteria values.

The case of a normal population with normally distributed index measurement variation is well known and serves to identify the nature of the bias problem and establish an intuitive base for assessing its magnitude. The following discussion and development within this section borrows heavily from Overton (1989).

The curves in Figure 5-4 illustrate the general pattern of distortion of the true population distribution (CDF Curve E), as successively greater magnitudes of index measurement variance are added (convoluted) to the distribution curves D, C, B, and A. The convolution of the normal population distribution with index measurement error distributions of progressively greater variance leads to a family of normal distributions with successively greater total variance (Overton, 1989).

Overton (1989) examined proportional error models and several asymmetric distributions and concluded that their behaviors are in essence similar to that of near-symmetric distributions with homogeneous normal measurement error. The models for more asymmetric distributions still

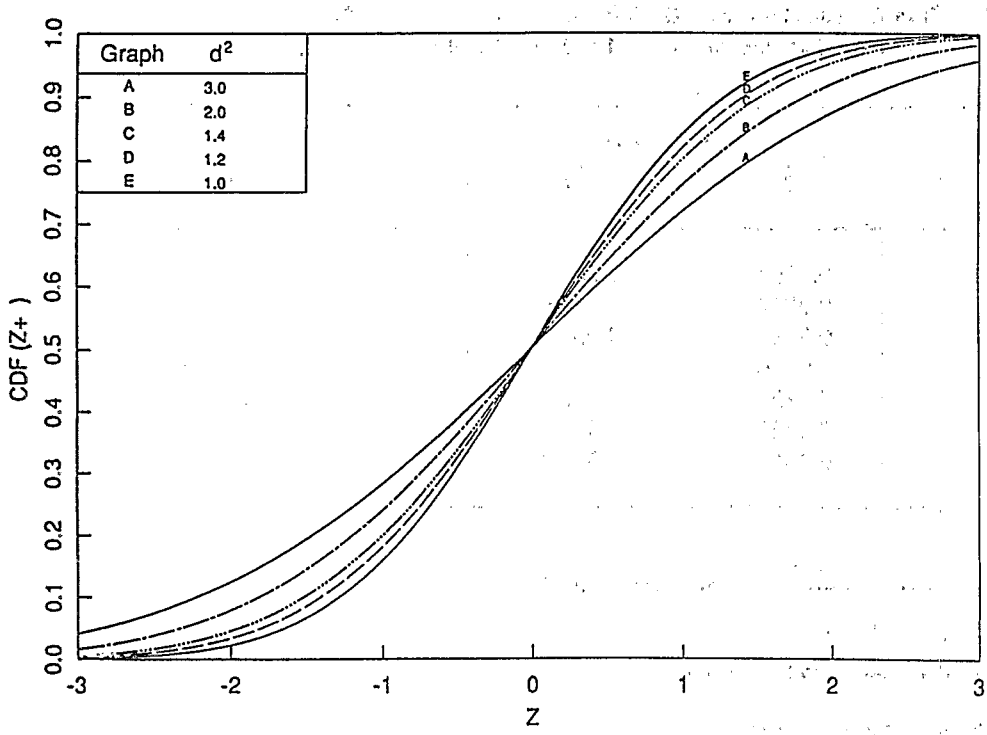


Figure 5-4. Convolution of a Standard Normal Distribution with normally distributed index variability, where $d^2 = (\sigma^2_{unit}+\sigma^2_{index})/\sigma^2_{unit}$ (figure from Overton, 1989).

show maximal bias in the neighborhood of ±1 SD, but maximal bias increases as asymmetry increases.

The similarity in the behavior of the different distribution and error models examined by Overton (1989) allows an assessment of acceptable index measurement variance to be made in terms of the normal population with homogeneous error. Table 5-6 shows bias of CDFs resulting from index measurement variance over a range that we might encounter in the selection of indicators and the design of a field sampling program. Maximum bias occurs in the neighborhood of ±1 SD from the population mean, and relative bias continues to increase toward the lower tail of the CDF. For larger index measurement variances, maximum bias will be farther from the population mean.

When index measurement variance is half the magnitude of the population variance, the estimated population distribution is biased considerably (relative biases of +0.31 and +1.25 times the true percentages at Z values of ±1 SD and ±2 SD, respectively). For values of index measurement variance less than 50% of the variance among sample units, maximum bias for symmetric populations is still close to ±1 SD. Consideration of bias at different CDF values of F(x) also identifies the bias at the Z = ±1 SD values to be the most important. In the lower end of the distributions shown in Figure 5-4, the population fractions that have index values below -2 SD are so small that they may be of little concern, and the bias may therefore be of little importance in general interpretations. Toward the mean, bias decreases to zero; so that assessing bias at Z = ±1 SD insures that it is negligible near the mean. At +1 SD, bias again increases to its maximal value, but relative bias is small, so it is clearly the lower end of the CDF that dominates our concern. Thus Overton (1989) emphasized a focus on the bias at -1 SD as the most crucial, not only representing the worst case, but also the values of greatest interest, and the values at which the maximal bias has the greatest importance.

Table 5-6. Bias of Apparent F(x) as a Function of d^2, Assessed for the Normal Distribution Function with Homogeneous Normal Measurement Error[a]

d^2	At ±1 SD		At ±2 SD	
	Bias	Rel. Bias	Bias	Rel. Bias
1.001	0.0001	0.0008	0.000	0.0024
1.01[b]	0.0012	0.0076	0.0005	0.0238
1.05	0.0059	0.0372	0.0027	0.1200
1.10[c]	0.0115	0.0726	0.0055	0.2424
1.20	0.0220	0.1387	0.0112	0.4921
1.50	0.0485	0.3054	0.0285	1.2521
2.00	0.0811	0.5111	0.0559	2.457

a $d^2 = \dfrac{\sigma^2_{unit} + \sigma^2_{index}}{\sigma^2_{unit}}$; from Overton (1989).

b Design standard suggested by Overton (1989).

c allowable limit suggested by Overton (1989).

For normally distributed populations with normally distributed measurement errors, Overton (1989) concluded:

- Bias is maximal between 1 SD and 2 SD from the population mean; for small relative magnitudes of index measurement variance, maximum bias occurs close to ±1 SD.
- Bias is zero at the mean; for small relative index measurement variance, bias is negligible near the mean and in the tails.
- Relative bias is large only for very small (highly negative) values of Z (where Z is a multiple of the population standard deviation).

5.3.1.2 Simulated Effects of Index Variability

In addition to this theoretical analysis, we also used simulations to assess the effect of a broad range in magnitudes of index variance (Figure 5-5). These simulations represent a sampling of 100 lakes from a hypothetical asymmetrically distributed population with a true mean of 25 and an SD_{unit} of 50 (among-lake variance of 2500). We used Monte Carlo techniques and a proportional error model that would be appropriate for many variables with asymmetrical distributions and multiplicative error structures that are of interest in EMAP. The hypothetical distributions were randomly sampled to yield the mean observed sample distributions (solid lines). The crosses depict the range within which 90% of the simulated sample distributions fell.

The six observed distributions (a through f in Figure 5-5) all have the same among-lake mean and variance, with the identical "true" frequency distribution represented in each figure by a broken line. Index measurement variance increases from 1% of among-lake variance in Figure 5-5a to 100% of among-lake variance in Figure 5-5f. To clarify, in case f, half the observed total variance is due to index measurement variability and the other half is from differences among lakes across the region. When index measurement error is only 1% of the magnitude of among-lake variance (Figure 5-5a), the observed

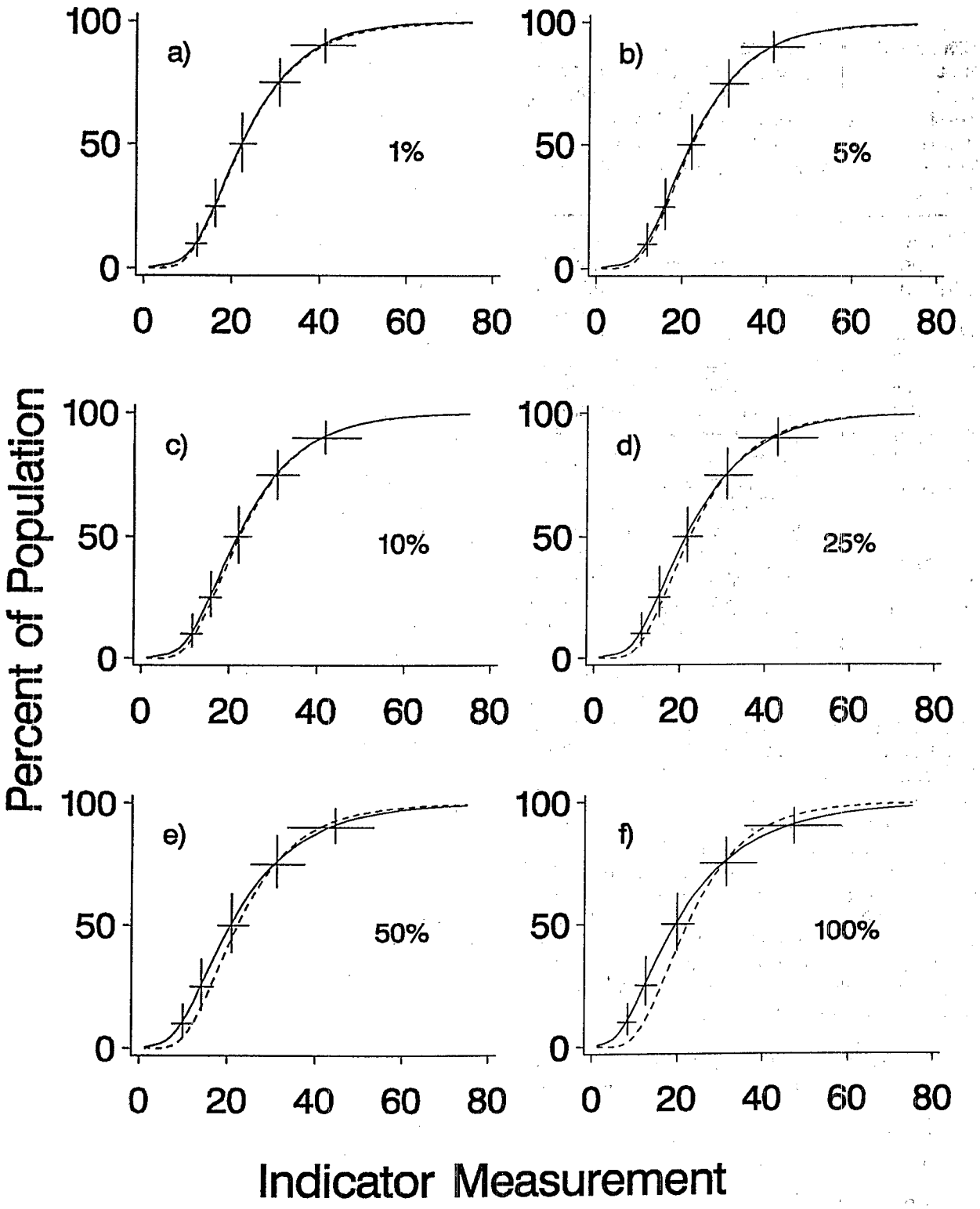


Figure 5-5. Simulated effect of index measurement variance on survey results observed from sampling 100 lakes from a hypothetical asymmetrically distributed population with a true mean of 25 and an SD_{unit} of 50. The true distribution is shown by the dotted line. Crosses show 90% range of simulated sample distribution. Index measurement variance increases from 1% to 100% of among-lake variance in (a) through (f).

distribution is almost identical to the true distribution. A positive bias in lower percentiles becomes progressively greater as index measurement variance becomes larger. This bias results in an overestimation of the number of lakes below some named value of the variable along the X-axis.

Estimates for the percentiles Q_5, Q_{10}, Q_{25}, Q_{50}, Q_{75}, and Q_{90} have insignificant absolute bias when index measurement variance is 10%, but acquire progressively greater positive biases as index measurement variance increases (Figure 5-6a). These absolute biases in percentiles do not exceed 10% of the total population number until index measurement variance rivals that of among-lake variance. However, Figure 5-6b shows that when index measurement variance exceeds 25% of among-lake variance, errors in the 5th percentile can exceed 50% (i.e., saying that 5% of the population is < X, when the true percentile is really 2.5%). With the same indicator measurement variance (25%), bias in estimating the population median is still quite small (5%) in relative terms (we say that 50% of the lakes are below some value, when the true value is only 47.5%). Once index measurement is 50% or 100% of among-lake variance, biases are quite large (Figures 5-6a and 5-6b). In fact the true population percentages lie entirely outside the 90% range of the simulated sample distributions for values of the indicator in the region of the 10th and 25th percentiles (Figures 5-5e and 5-5f).

Figures 5-6a and 5-6b show that the absolute and relative magnitudes of bias in population estimates can be quite divergent, and they have different implications for interpretation. It is perhaps necessary to clarify by illustration the meanings of absolute and relative bias. Imagine that we have sampled 100 lakes in a regional population, with single observations, using a very "noisy" indicator that has an index measurement variance 50% of that derived from true differences in lake condition across the region. From this survey, we state that 5% of the lakes are "impaired" and 25% are "threatened." Figure 5-5e shows our observed frequency distribution, our estimate of the true population distribution. Figure 5-6a shows that in fact the estimate of 5% impaired is overestimated by about 3.5%; really only 1.5% of the population is impaired. In a relative sense (Figure 5-6b), this is an overestimate of about 70% (i.e., 3.5%/5.0%).

5.3.2 Acceptable Levels of Index Variability for Status Estimates

Overton (1989) suggested that for the normal distribution, a reasonable upper bound of index measurement error yielding acceptable bias over most of the estimated population distribution is 10% of σ^2_{unit}. For this level of index variability, the maximum bias is approximately 0.115, and critical relative bias is approximately 0.07 times the estimated population proportion (Table 5-6). At values of index measurement variance < 1% of sample unit variance, maximum bias (0.001) and relative bias (<0.01) are very small, and Overton (1989) suggested this degree of relative precision in index measurement might be viewed as a design standard.

It should be noted that acceptable bias is ultimately a function of the precision required in any given population condition statement. Information from even very imprecise indicators may be very useful. As a hypothetical example, suppose that the Z value of -2.0 in Figure 5-4 is "subnominal". The true population CDF (Curve E) shows that 2% of the population is subnominal. An indicator without impressive precision (e.g., index measurement variance = 20% of among-lake variance) estimates that about 3% of the population is subnominal. A very "noisy" indicator, with measurement variance equal to the among-lake variance, overestimates the population percentage of subnominal waters by over 3 times (7%), but still allows the policy-relevant statement that < 10% of surface waters are subnominal. Some caution is in order, however, for if the populations and their variances are extremely asymmetrical, even the mildly precise indicator (indicator measurement variance = 20%) could overestimate the subnominal proportion by about 4 times (12%), not allowing even the rough assessment that < 10% of the population is subnominal. In the asymmetrical distribution, index measurement variance must be < 10% to yield the correct rough conclusion that < 10% of the population is subnominal.

In light of the illustrations in the previous paragraph, the index measurement variance criteria values of 1% (design standard) and 10% (upper bound of acceptability) recommended by Overton (1989) appear to be reasonably robust guidelines for design of indicators and field sampling efforts, particularly if analysis of trends

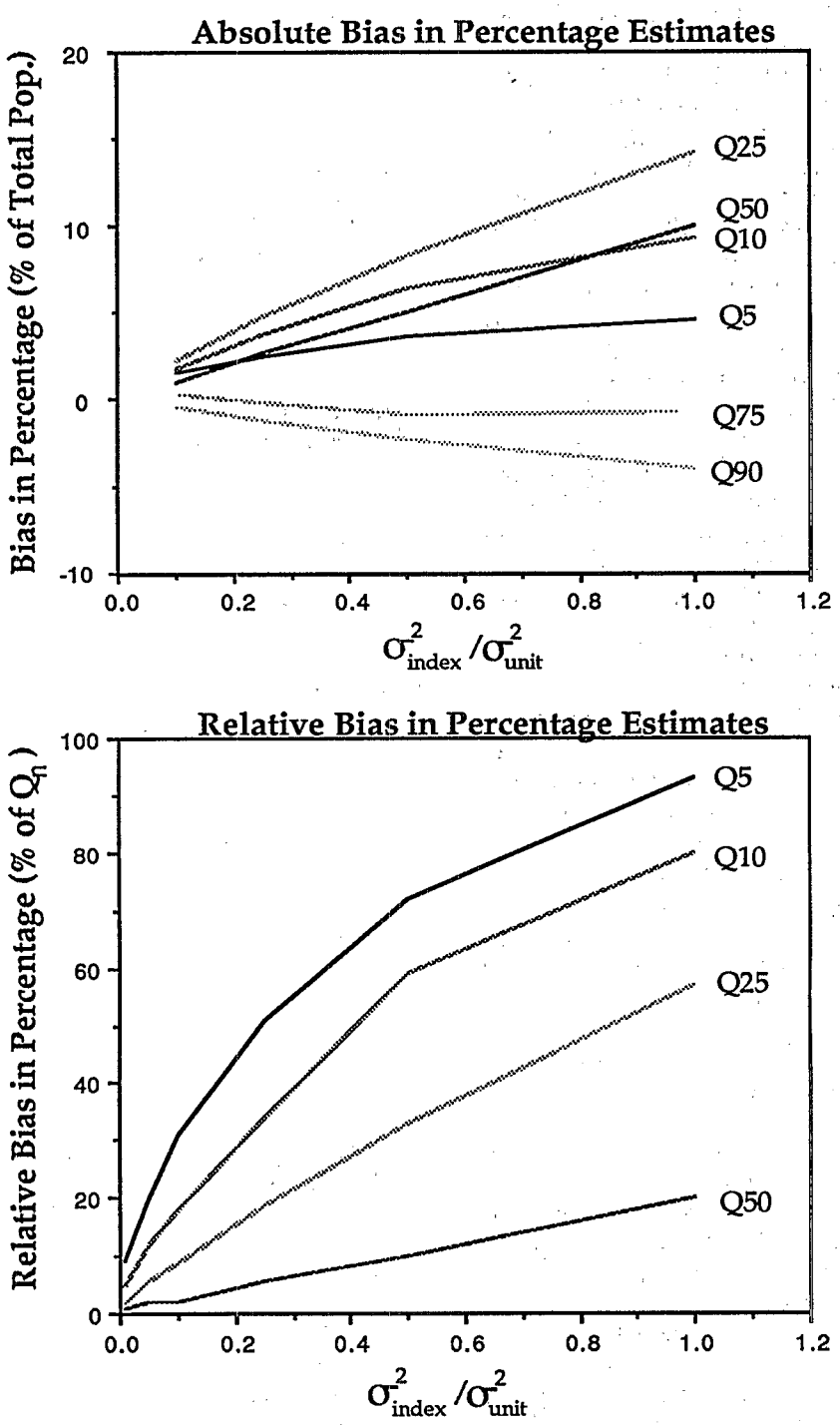


Figure 5-6. Absolute (a) and relative (b) bias in population percentiles for ratio of index variance to among-sampling unit variance.

is desired. For population status estimates alone, however, these guidelines may be rather stringent, particularly if population variances are approximately normal in distribution. Considerably noisier indicators will probably be useful in these applications, depending both upon the degree of symmetry in their error distributions and on the precision and resolution of population descriptions desired. The differences in biases regarding statements near the tails of CDF's based on indicators with different variance distributions (e.g., normal versus asymmetrical) underscore the importance of obtaining information regarding indicator variability during EMAP pilot studies. Such studies may also be a necessary component of ongoing EMAP.

In general, it appears likely that if index measurement variance is greater than about 50% of variance among sample units (e.g., lakes), and we believe that the regional population of lakes spans a meaningful range of the indicator variable, then the index measurement may be too imprecise to answer with confidence some questions posed concerning the proportions of the population within stated ranges of ecological condition defined by the indicator values. In such cases, we must either refrain from making high-resolution statements, seek a more precise indicator, or achieve greater precision in the indicator by better targetting, by replicating samples from each waterbody in space or time, by taking more spatially extensive samples within the waterbodies, or by deconvolution of population distribution estimates after the fact.

5.3.3 Required Indicator Sampling Intensity for Status Descriptions

An examination of the relative magnitudes of variance components in regional data sets can be used to assess optimal allocation of sampling effort among sample units (sites), within-index period replicate measurements, and duplicate measurements to reduce the biases in estimated population distributions. Sampling intensity must be adequate for robust calculations of index values that do not, by their uncertainty, distort population distribution functions or preclude timely detection of relevant trends. The strategy for evaluating the allocation of sampling effort between replicates and additional sample sites across a region is similar, whether the goal is minimizing the error in the estimated population mean or minimizing bias and uncertainty in descriptions of the population distribution. Both cases involve a comparison of the magnitude of variation for individual site index values with variation among all sites across a region of interest.

It is clear from the expression for σ^2_{mean} (page 5-17) that, if the cost of index sample replications is equal to that for adding sample units, the degree of precision in estimating the mean gained for each incremental increase in total sample number is always greater for increases in n_{unit} than for n_{rep}. This is true even if index variation (σ^2_{index}) is greater than the regional variation among sample units such as lakes (σ^2_{unit}). The advantage of replicating index measurements within sample sites comes rapidly into play if the cost of such replicates is substantially lower than that for additional sites and if σ^2_{index} is large relative to σ^2_{unit}.

The optimum allocation of sampling effort between numbers of measurement replications (n_{rep}) and sample sites (n_{unit}) for minimizing the standard error of the population mean, adapted from Snedecor and Cochran (1980, p. 453) is:

$$\left(\frac{n_{rep}}{n_{unit}}\right) = \sqrt{\left(\frac{c_{unit}}{c_{rep}}\right) * \left(\frac{\sigma^2_{index}}{\sigma^2_{unit}}\right)}$$

We calculated optimum sample allocations for a reasonable range of cost and variance ratios that we might expect in measuring response and exposure indicators in EMAP-Surface Waters (Figure 5-7). These calculations show that, for cases in which costs of index measurement replication are roughly equal to the cost of adding new sites, there is no advantage in replication. This conclusion holds no matter what the variance ratio of sample units to index measurement. We caution that this conclusion pertains to minimizing error in estimation of the population mean. Replication of index measurements at sites (though not necessarily all sites) is certainly useful in revealing the portion of the variance observed over a region that is simply due to index measurement error, rather than to true differences among sample units (e.g., among lakes).

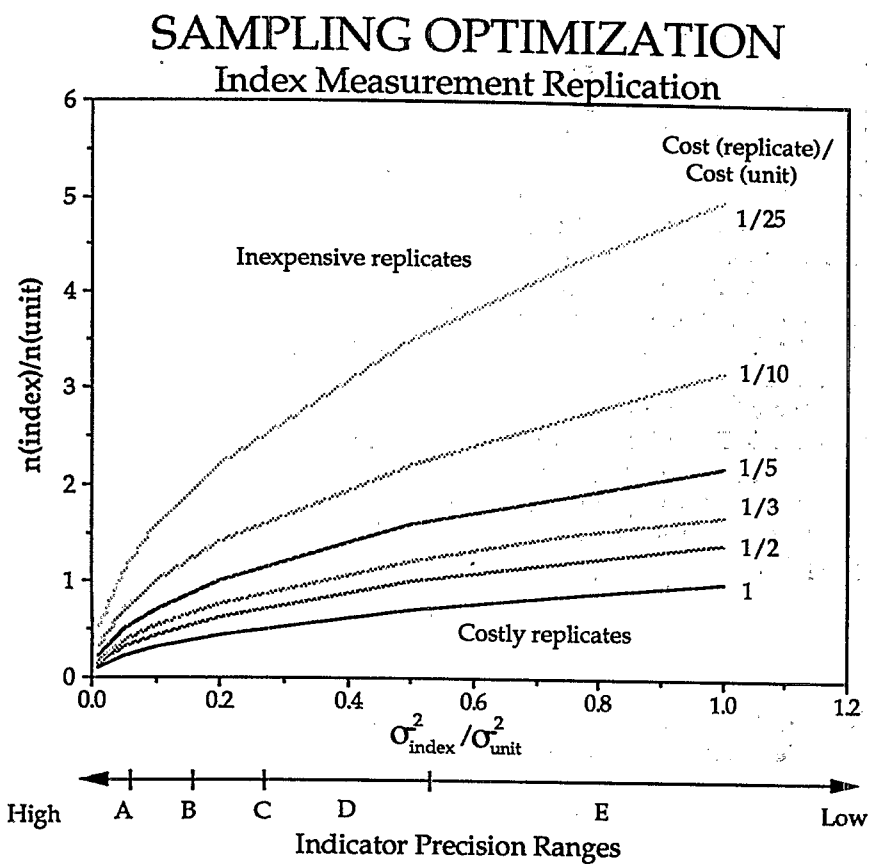


Figure 5-7. Nomogram for determining sample allocation between replications within sample unit and adding new sample units in a region based on ratio of cost for replications and new sites and the variability within the index period and among sample units.

Replication of index measurements within individual water bodies becomes more advantageous than adding sample sites in reducing the error of the estimated population mean when the ratio of replication cost to new site cost is greater than the ratio of index measurement variance to among-site variance $[(c_{rep}/c_{unit}) > (\sigma^2_{Index}/\sigma^2_{unit})]$, as illustrated in Figure 5-7. For example, if index measurement variance is one-fifth that among sample sites (x axis value of 0.2 in the figure), systematic replication is not justified unless the cost of each replicate index measurement is less than one-fifth the cost of adding an addition sample unit (third curve up from the bottom in the figure). In most cases, replicate costs (e.g., an additional visit during the index period) will probably cost more than one-fifth that for a new site, unless the replicates are spatial or analytical replicates on the same lake or stream during a single site visit. If index measurement variance is unknown (the case for most EMAP variables) or is expected to be large, replication on at least a subsample of lakes (or streams) is necessary to determine the portion of the measured regional variation in surface water condition that derives simply from measurement variation. This information might allow a recalculation of regional means, variances and frequency distributions that would be observed if replicated index measurements at each waterbody had been averaged (this recalculation and correction procedure is termed "deconvolution").

To apply Figure 5-7 to decisions about field sampling design in EMAP, we need information

about the ranges of indicator variability within sample units and across regions and the relative cost of replication compared with additional samples. This information is in Tables 5-7 and 5-8.

Table 5-7 (based on Tables 5-4 and 5-5) shows rough categories of relative precision in indicator measurements that correspond with ranges A through E in Figure 5-7. At the far left of the x-axis (range A), index measurement variance is almost insignificant relative to the range of variation among different surface waters across a region. At the other extreme, range E extends to conditions where the magnitude of index measurement variance rivals or equals that observed in sample units across a region. If the range of an indicator variable measured across a region describes a meaningful difference in surface water condition, then indicators with variance characteristics in the A range can be considered "very precise" indicators, and those in the E range can be considered "noisy." These indicator precision categories must be interpreted in light of the meaningfulness of the range of surface water condition across the population used as a basis for variance comparisons. For example, in cases where a population of surface waters is homogeneous, index measurement variation (though perhaps small in absolute terms) may be large compared with the range of indicator measurements observed among different waters in the region. In such cases, the fact that the indicator is noisy relative to the range of variation in condition in the population of surface waters simply means that condition within individual

Table 5-7. Rough Categories of Index Measurement Precision (Sum of Analytical, Temporal, Spatial Components) Relative to Regional Variance for Possible EMAP Indicators[a]

Group[b]		
A Very high precision	0-5%	NSS: SO_4, Cond., ANC, pH, NO_3 ELS-II: ANC PIRLA-II: DI-pH (deep core)
B High precision	5-15%	ELS-II: SOBC, SO_4 PIRLA-II: DI-pH (surficial cores)
C Medium precision	15-25%	Ohio (streams): fish species richness, IBI ELS-II: pH, Secchi depth NSS: T-dP, turbidity Minnesota, Vermont: Secchi depth
D Low precision	25-50%	Ohio (large rivers): IBI Vermont: Chl a, TP Minnesota: TP
E Noisy	>50%	Ohio (large rivers): fish species richness Minnesota: Chl a

[a] Table based on summing columns 3 to 6 in Tables 5-4 and 5-5; information on all index measurement variance compartments was not available for many variables (e.g., inter-annual variance for NSS and within-reach variance for Ohio IBI). Data sources are cited in Tables 5-4 and 5-5.

[b] See also Figure 5-7.

Table 5-8. **Rough Costs of Index Measurement Replication Relative to Costs of Additional Sites (Sample Lakes or Streams)**[a]

Category	$\frac{C_{unit}}{C_{replicate}}$	Type of Measurement (Examples)
Inexpensive	10-25	Duplicates of routine physical, chemical variables (Secchi depth, Chl a, D.O., SO_4)
		Spatial replicates of routine physical, chemical variables -- same visit, same waterbody, but different location (Secchi depth, Chl a, D.O., SO_4)
Moderate	5-10	Spatial replicates of complex analysis on same visit, same waterbody, but different location (paleo cores, macroinvertebrate and fish assemblage, habitat measurements)
		Temporal replication of simple measurement or collection of easily obtained sampled (Volunteer measures of Secchi depth, or recovery of sediment trap by returning field crew)
Costly	1-2	Temporal replicates (repeat visits) (Secchi depth, Chl a, D.O., SO_4, IBI, physical habitat variables, chemical habitat variables)
Very costly	1	Temporal replicates of very costly analyses

[a] Relative costs are crudely estimated assuming a total cost of $5,000 per site/visit, $300 per site cost of full suite analytical duplicate of routine chemistry variables, somewhat higher for paleoecological analysis of sediment cores, and macroinvertebrate, and fish assemblage sampling/analysis.

surface waters is nearly as variable as the condition of different surface waters across a region. The indicator may be noisy in the homogeneous surface water population but adequate to discern differences in surface water condition in a more heterogeneous region.

Many of the chemical habitat variables and the diatom-inferred pH measurements from deep cores are in the "very high precision" categories (A), in which index measurement variance is less than 5% of sample unit variation. We caution that inter-annual variability is not available for the NSS chemical variables, but if ELS-II data are any indication, that component might add another 3% to 8% to the index measurement variability and would move those variables down to group B. The "high precision" group (B), has index measurement variance between 5% and 15% of among-sample unit variance. Group B includes diatom-inferred pH from surficial cores, and chemical variables that are not very active biologically, such as the sum of base cations (SOBC) and lakewater SO_4. The "medium precision" group (C) includes many biologically active and flow sensitive variables such as phosphorus, turbidity, and Secchi disk transparency. The medium precision group also includes the determination of fish species richness and IBI in small streams. Chlorophyll-*a*, phosphorus in some regions, and large river IBI determinations are of relatively lower precision

(Group D), with index measurement variances between 25% and 50% of among-site variance in the data sets we examined. Among the indicator measurements we examined, only Minnesota lake phosphorus and fish species richness in large rivers had combined index measurement variance components summing to more than 50% of among-site differences.

The relative cost categories in Table 5-8 are very crudely estimated, and assume that the cost of adding a new site is $3,000 to $5,000. Observe that most of the precise indicators, such as the chemical variables (group A in Table 5-7), are also the least expensive to replicate, as long as the replication takes place during the same sample site visit. For most of these, even spatial replication during the same sample visit is probably not justified, except in very large lakes and streams (but see discussion about littoral and backwater habitats in 6.3.3 and 6.4.3). Spatial replication appears to be advantageous in cases where the indicator is of moderate to low precision (groups C, D, E), AND the major source of indicator measurement variation either is analytical (or methodological) or is due to spatial heterogeneity within the sample lake or stream. Likely examples of this case are IBI or fish species richness in lakes and streams.

The general conclusion of the analysis discussed over the last several paragraphs is that inexpensive, easily obtained duplicates or replicates are desirable at all sites for high variance (noisy) indicator measurements. Replicates on a subsample of sites are advantageous in cases where the cost of replication is high relative to the cost of adding new sites in a region. Optimum allocation of sampling activities among samples over a region and spatial, temporal, and analytical replicates within single lakes or stream segments depends on accurate information concerning variances and costs at various levels of sampling. Pilot research activities will include efforts to quantify the relative costs and relative magnitudes of variance components.

Compared to procedures for optimally allocating sampling effort for obtaining precision in estimated population means, the procedures for obtaining adequate precision and accuracy in estimated distributions of indicator values in surface waters are much more difficult to evaluate. Once evaluated, however, they are in some ways more straightforward to describe. Section 5.3.1 developed the theoretical basis for evaluating the effect of index measurement variance on population distributions observed in regional surveys.

If the variance of an indicator is excessive, it can be cut in half by replicating the sample and using the average of the two measurements as the index value for each waterbody. The effect of variance of this refined indicator is then evaluated in the manner described in Sections 5.3.1 and 5.3.2, but the indicator measurement variance is reduced because the index value for each sample site is now the average of replicate measurements.

The variance tables (5-4 and 5-5) show that combined index measurement variances for most of the chemical exposure and diatom-inference indicators examined are less than 10% of among-site variance (many are less than 5%), probably precise enough to justify single measurements for describing status. The more precise of the fish assemblage indicators and the flow-responsive and biologically active chemical species have combined index measurement variances of about 10% to 30%. At the high variance end of this range, it would be advantageous to replicate spatial sampling within waterbodies to reduce measurement error. This would reduce bias in estimation of true population distributions. For some indicators, such as fish assemblage, variance might more effectively be reduced by increasing the spatial extent of fish collection within each lake or stream reach. When combined index measurement variance exceeds about 30% of the among-site variance in a region, as occurs with fish species richness and IBI in large rivers, index measurement variance must be reduced if we need to make fine resolution statements about the condition of the regional population. By fine resolution, we mean, for example, statements about the fifth or tenth percentiles rather than simply the median. Wherever possible, it will be more efficient and economical to increase spatial replication within sample waterbodies, rather than to increase the temporal frequency of sampling within the index season. However, knowledge of the component sources of indicator variability is necessary (from literature or pilot studies). This information will allow evaluation of where in the sampling hierarchy (analytical duplicates, spatial replicates, temporal replicates, etc.) increases in sample sizes are likely to result in greater indicator precision.

5.4 EXAMPLES OF EFFECTS OF INDICATOR VARIABILITY ON POPULATION DESCRIPTIONS AND TREND DETECTION

The previous sections describe approaches that can be used to characterize regional populations, differences among subpopulations, and changes and trends within subpopulations. Earlier sections also cover the importance of the relative magnitude of index variation versus population variation in making these estimates. In this section, we illustrate some of these descriptions with examples selected from data sets similar to the kinds of data we expect to obtain from EMAP, that is, data sets covering reasonably large areas and including repeat visits over a period of several years. These data sets were among those used in Section 5.2 to quantify and illustrate the relative magnitudes of components of variation. A brief description follows of the data sets from which the illustrations are drawn. Only one case, the Eastern Lake Survey, involved a probability design from which estimates of regional condition can be derived. In the other studies, lakes or streams were selected for other specific purposes.

5.4.1 Eastern Lake Survey

As part of the National Acid Precipitation Assessment Program, a variety of chemical attributes were measured in lakes > 4 ha in the northeastern United States. Lake populations were identified, target populations were specified, and probability samples were drawn for field visitation. The bulk of the sample lakes were visited once during fall overturn in 1984; a subset of lakes was visited three times during the subsequent year, during spring, summer, and fall, to estimate the magnitude and influence of seasonal variability on estimates derived during a single index period. For our illustrations, we use total phosphorus sampled in these lakes. Data are from Linthurst et al. (1986) and Herlihy et al. (in press).

5.4.2 Minnesota Pollution Control Agency

The Minnesota Pollution Control Agency conducts routine lake monitoring throughout the state to estimate trophic condition. Total phosphorus, chlorophyll-*a*, and Secchi disk transparency are measured in selected lakes several times from ice-out in spring to late summer. Over 1,000 lakes have been monitored in the past 10 years, with consistent methods. Only a small subset of the lakes was visited repeatedly in different years. These data can be retrieved from the EPA's STORET system. We selected a subset of lakes that were sampled repeatedly among years from the database of over 1,000 lakes. These data were obtained in cooperation with Steve Heiskary and Bruce Wilson of the Minnesota Pollution Control Agency.

5.4.3 Vermont Department of Environmental Conservation

The Vermont Department of Environmental Conservation samples lakes throughout Vermont, targetting trophic condition and visiting lakes several times during the year. Lakes are sampled during spring overturn for total phosphorus and during the open water seasons for Secchi disk transparency and chlorophyll-*a*. A citizen's monitoring program gathers additional Secchi disk readings at numerous lakes to supplement the Department's monitoring program. Many of the lakes were sampled repeatedly for seven to nine years for total phosphorus and Secchi disk transparency; thus this database is useful for illustrating the kinds of trend detection techniques EMAP will be using initially. Eric Smeltzer of the Vermont Department of Environmental Conservation made the data available to us.

5.4.4 Ohio Environmental Protection Agency

The Ohio Environmental Protection Agency conducts routine stream sampling for a variety of physical, chemical, and biological attributes. Most of the monitoring program is devoted to identifying and characterizing impacts and impairments associated with point source discharges, but some sampling has been devoted to characterizing minimally impacted systems. In the illustrations below, we have drawn from the Agency's extensive fish assemblage data collection efforts. As part of their routine fish collection program, streams are visited two to three times during the summer and sampled by electrofishing methods over 100 to 300 meter sections of the streams.

During the past 10 to 15 years, several thousand sites have been sampled, so there is an extensive database from which to draw illustrative material. Relatively few of the sites have been visited over a period of years, so the primary purpose of these data is to illustrate differences among regional subpopulations. The Agency aggregates the raw fish collections records into indices that charac-

terize each site each time the site is sampled. Our illustrations use one of the indices, the Index of Biotic Integrity, modified from the one developed by Karr et al. (1986). Chris Yoder has made various parts of this database available for our exploratory use.

5.4.5 Descriptions of Status

As discussed earlier, one of the initial outputs of EMAP-Surface Waters will be descriptions of the status of regional populations and subpopulations of indicators of interest. Cumulative distribution functions will form the basis of these descriptions, with summaries of various population parameters included. An example of this type of summary is given by the regional description of total phosphorus in lakes in the Northeast derived from the Eastern Lake Survey (Figure 5-8). In this example, confidence bounds on population status can be established because the survey was based on probability methods. As in the surface water surveys, only the upper bound is shown. A summary attached to the figure gives selected population parameters. During the initial phases of EMAP, we expect much of the output to be this type of descriptive summary of various populations and subpopulations.

Part of the Eastern Lake Survey program documented the minor contribution to variance that was due to index variability relative to population

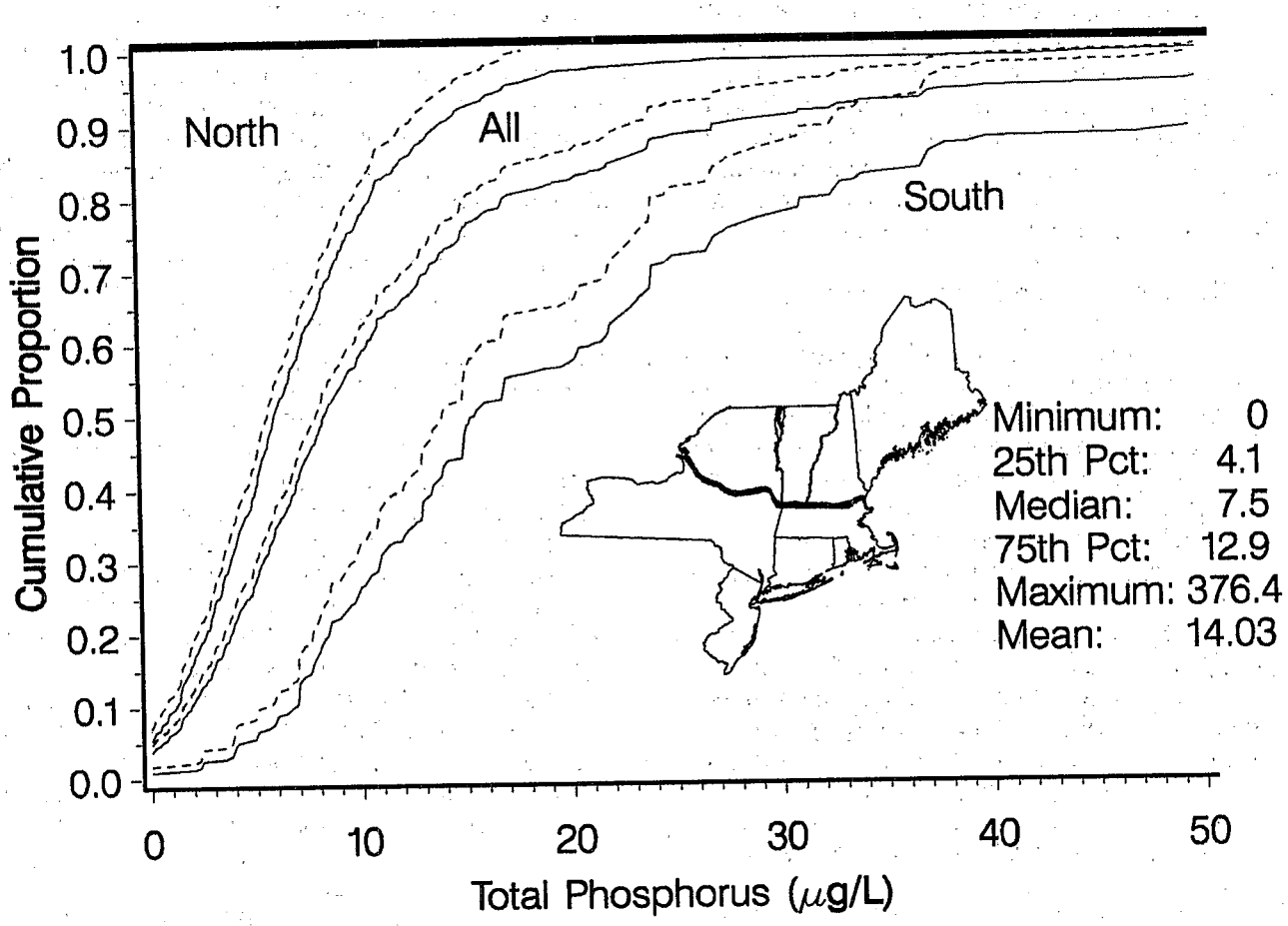


Figure 5-8. Cumulative distribution functions (CDFs) describe the status of regional populations of lakes and streams, as this figure illustrates for total phosphorus in the Northeast, based on results of the Eastern Lake Survey. The population of lakes in the Northeast is split into two parts to show regional subpopulations with clearly differing status. Upper 95% confidence bounds on each distribution are included.

variation, so distortion of the CDF was not an issue. For some of the EMAP variables, distortion will be of importance. Figure 5-9 is an example drawn from the Maine database, showing a CDF derived from data collected in August 1985. Chlorophyll-a is one of the variables for which we expect index variation to be an important component of total variation and therefore to contribute some distortion to the display. The illustration shows the CDF derived from the data collected and the estimated true CDf with the distortion removed by deconvolution. Observe that, like the simulations shown in Section 5.3, distortion is greatest toward the ends of the distribution. A similar example is given for the Index of Biotic Integrity in Ohio, a variable that is less "noisy" than chlorophyll-a is likely to be.

5.4.6 Differences and Changes in Regional Subpopulations

During its initial phases, EMAP-Surface Waters will identify possible subpopulations of interest for targetting further specialized investigations. It will be useful to determine whether characteristics of the proposed subpopulations differ from each other. As more data are collected, there will be increasing interest in evaluating the significance of temporal changes within subpopulations. Most statistical methods test the significance of differences in selected population parameters such as means or medians; EMAP-Surface Waters will evaluate differences over the entire distributions (CDFs). Differences and changes might be more readily detected by looking at the entire distribution rather than some specific part.

Comparisons will rely on chi-square tests on proportions of the distributions that fall in different quartiles or quintiles. The expectation is that, under the scenario of no difference, proportions expected in each category will be equal; if differences should be present, they could be detected by shifts in the proportions within each quartile or quintile. The general procedure for paired tests is shown in Figure 5-3; for unpaired tests, individual subpopulations are scored relative to the average CDF of the combined subpopulations. The following examples summarize the chi-square tests, indicating the probability that the calculated chi-square score would be exceeded by chance. High probability values indicate that real differences among subpopulations are highly likely.

The initial set of illustrations covers differences among subpopulations. We expect large differences among various candidate subpopulations, and these will be evident graphically. As we begin to target particular kinds of subpopulations at a finer scale, the role of chi-square tests will become more important. As an example of large differences in subpopulations, the Eastern Lake Survey total phosphorus data for the Northeast can be split into two distinct regional subpopulations, illustrated by clear differences between the CDFs in Figure 5-8. The northern subpopulation of lakes occurs primarily in forested landscapes in which soils are generally nutrient poor; the southern subpopulation occurs in land with substantial agricultural development and urbanization.

Another example, taken from the Ohio Environmental Protection Agency database, compares two ecoregional subpopulations, Regions 1 and 5 in the Huron Erie Lake Plain and the Eastern Corn Belt Plains (Figure 5-10). Again, observe the substantial differences. On the other hand, other cases will produce subpopulations that are not as distinguishable (Figure 5-10).

Temporal differences in populations or subpopulations are likely to be more difficult to distinguish or detect than major regional differences. Our illustrations include two types of time intervals: within-season comparisons and between-year comparisons. Since index sampling will be the basis for much of the descriptions, it is of interest to determine whether differences among potential index periods occur. For some of the databases, we can compare different possible index periods. For example, the Vermont trophic measurements taken in July and August can be compared as different index periods (Figure 5-11). Extending this type of analysis among years produces comparisons of data collected during the same index period but in different years. We can illustrate two types of comparisons EMAP will conduct. The first will be during the early years when relatively few sites will be revisited and comparisons will be based on independent samples (Figure 5-12). The second will be after the first four-year cycle, when comparisons can be made with repeat visits (Figure 5-13).

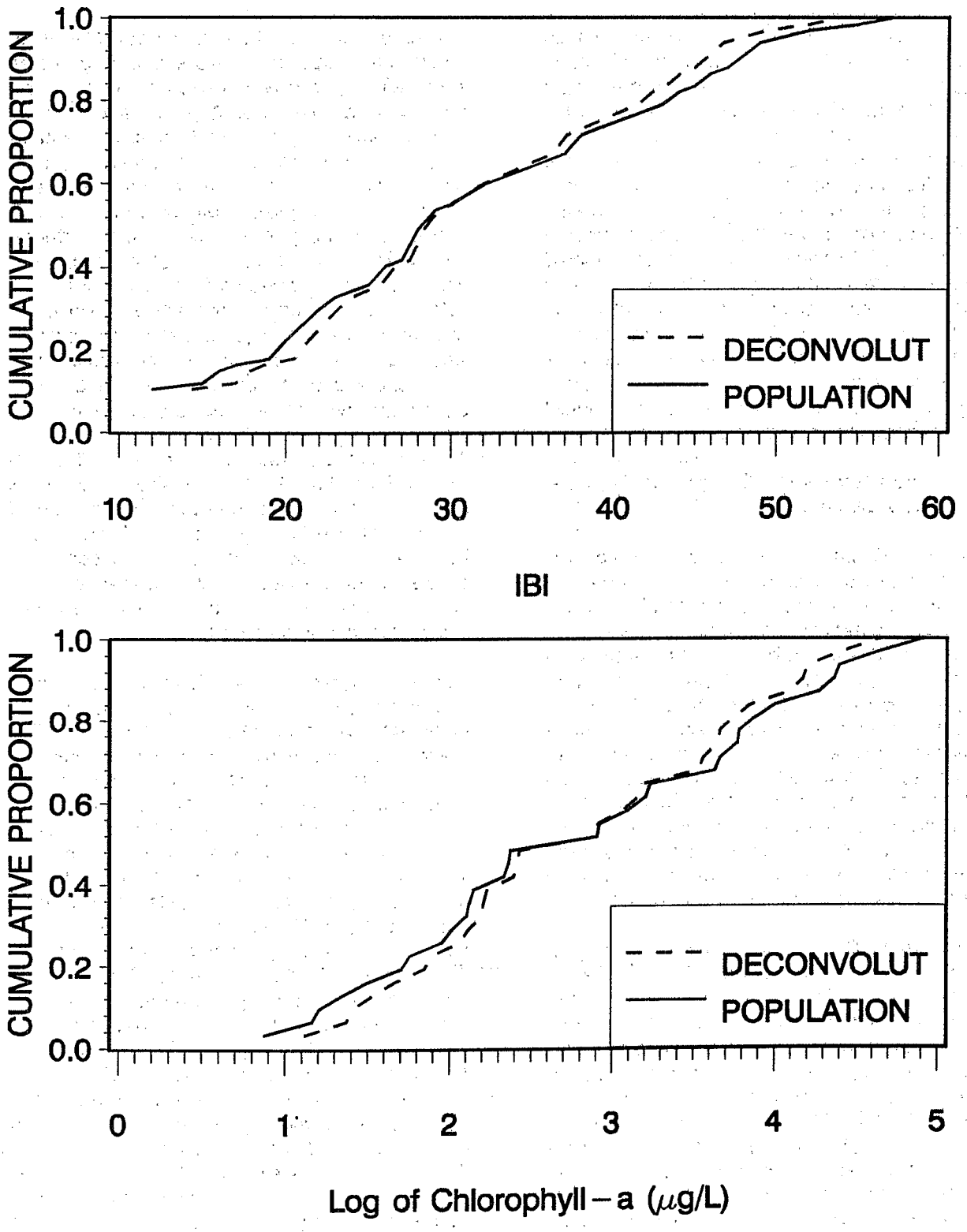
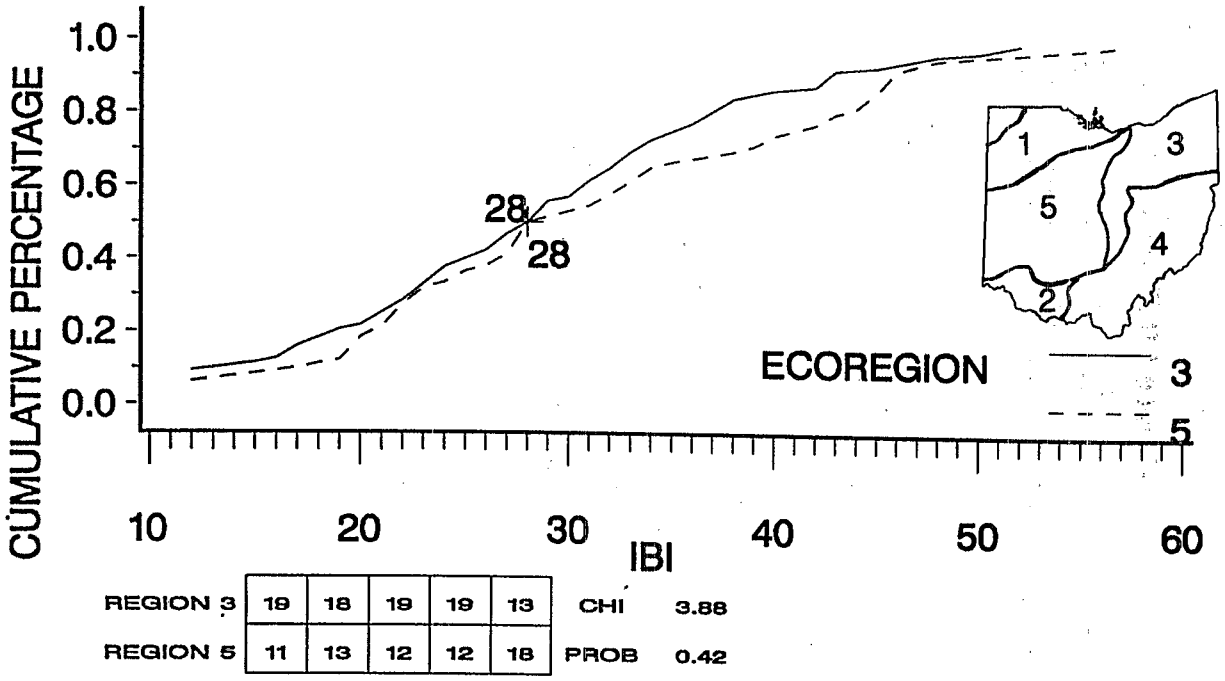
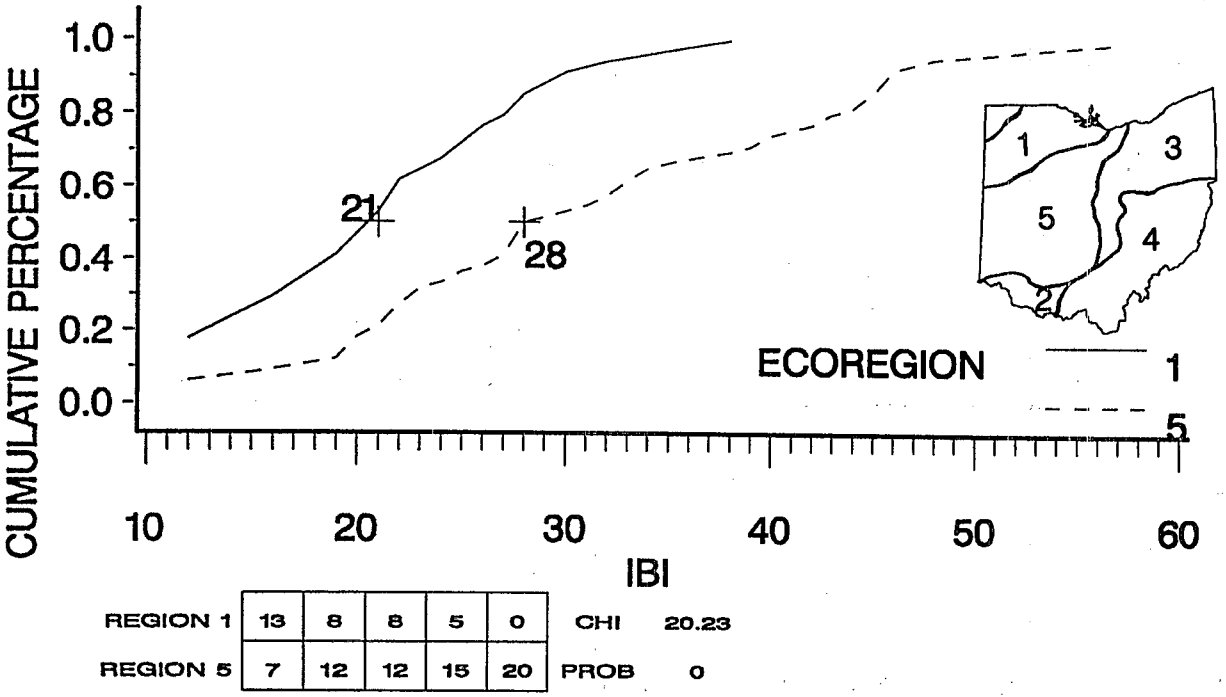


Figure 5-9. Cumulative distribution functions (CDFs) of the Index of Biotic Integrity for Ohio streams, and chlorophyll-a for Minnesota illustrate the status of two indicators having relatively large index variation relative to population variation. The distorted CDFs can be deconvoluted to present a truer picture of the CDF of the regional population.



REGION 3	19	18	19	19	13	CHI	3.88
REGION 5	11	13	12	12	18	PROB	0.42



REGION 1	13	8	8	5	0	CHI	20.23
REGION 5	7	12	12	15	20	PROB	0

Figure 5-10. **Cumulative distribution functions (CDFs) of the Index of Biotic Integrity for Ohio streams illustrating situations in which regional subpopulations are quite similar (Region 3 versus Region 5) and situations in which clear differences among regional subpopulations occur (Region 1 versus Region 5). Chi-square summaries indicate differences in subpopulations if the probability is low (e.g., 0.10).**

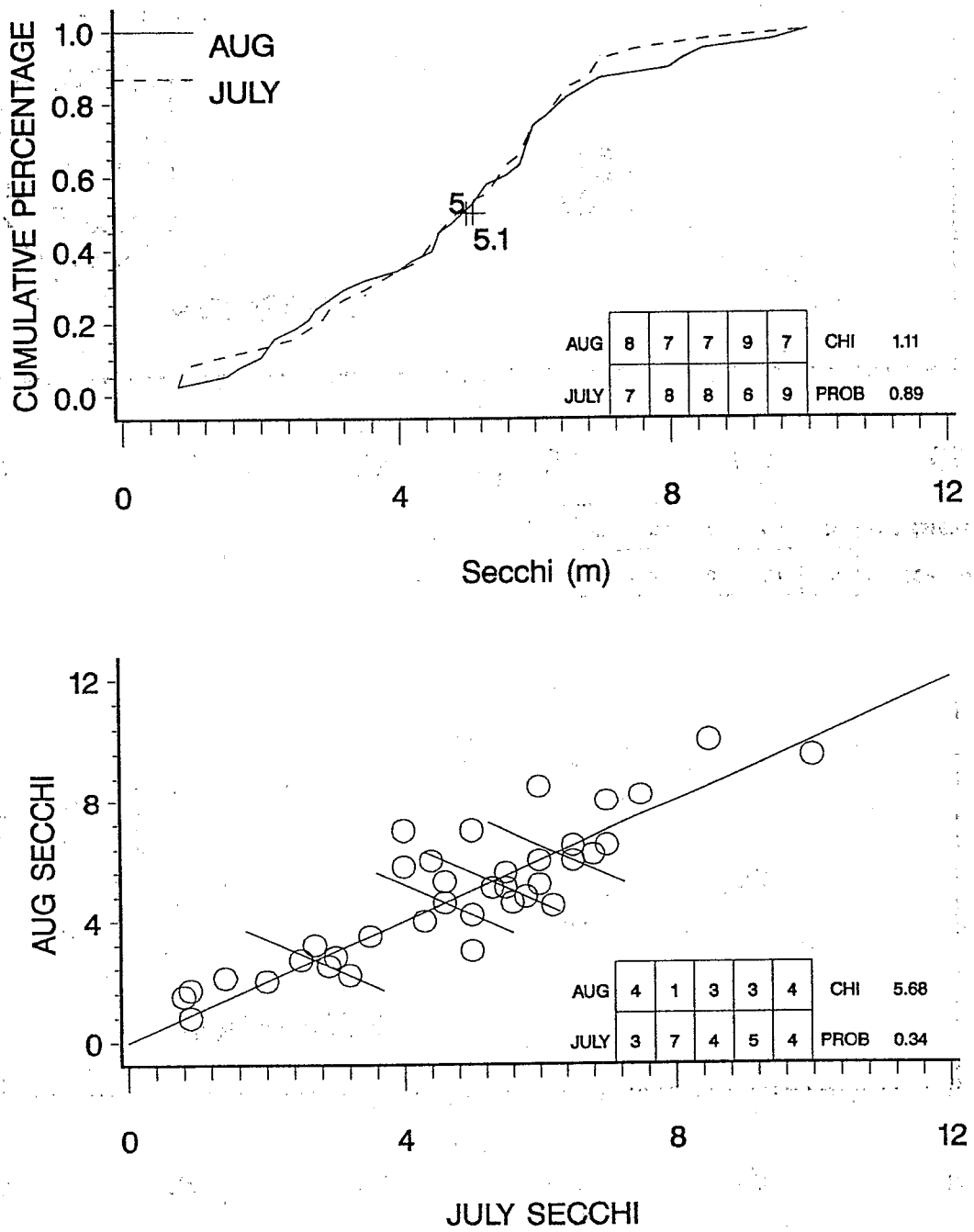
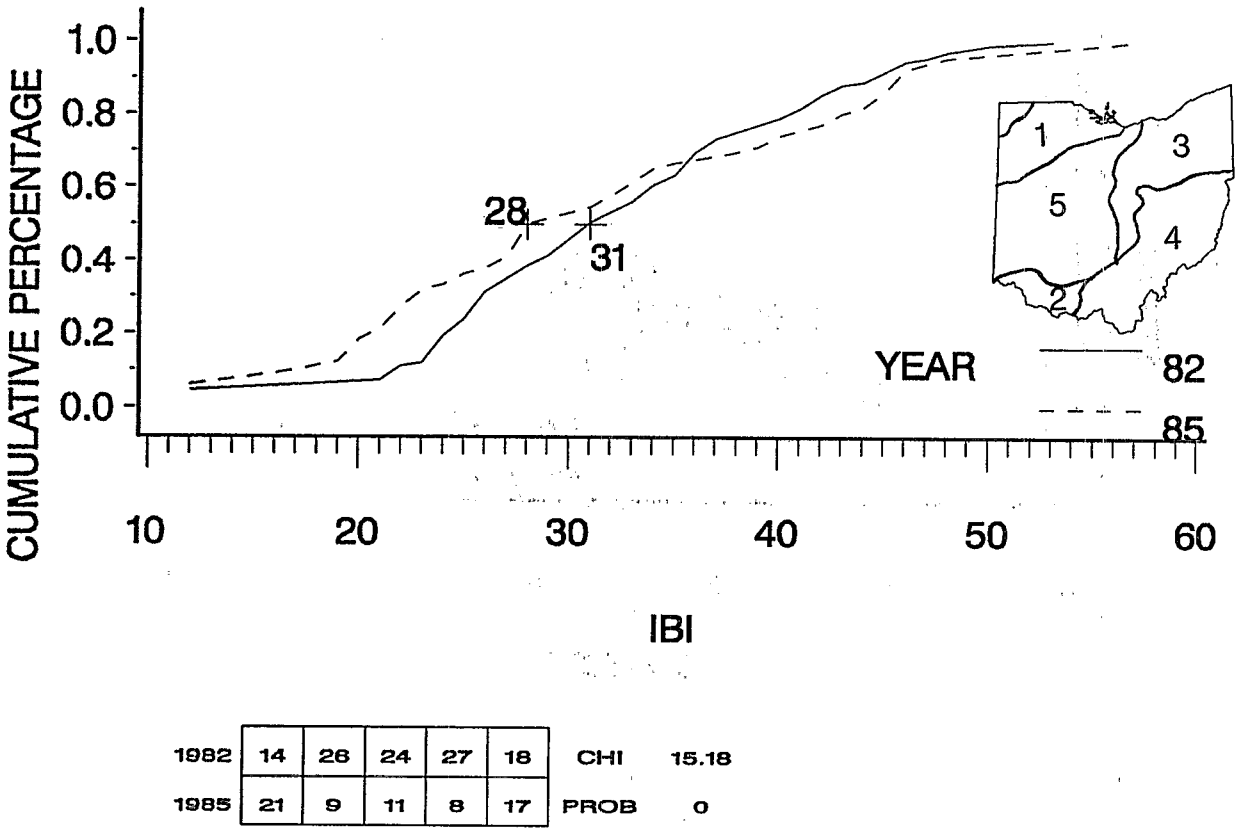


Figure 5-11. **Cumulative distribution functions (CDFs) showing the similarity in the Secchi disk transparency status of lakes in Vermont measured during two index periods of the same year, July and August 1985 (upper panel); median values are identified mid-graph. The lower panel shows another way of displaying data when sites are paired; the data set contains several measurements made during each month; for the comparisons, one value was selected randomly for each month for each lake. Chi-square tests indicate no difference in these two index periods, in both cases.**



1982	14	26	24	27	18	CHI	15.18
1985	21	9	11	8	17	PROB	0

Figure 5-12. **Cumulative distribution functions (CDFs) using the Index of Biotic Integrity for Ohio streams show year-to-year differences in the status of small streams. In this case, sample sizes are quite large, so relatively small differences can be detected among streams with independent samples. In this illustration, the difference in the median value was 3 units (28 for 1985 and 31 for 1982).**

5.4.7 Regional Trends

Although we have not worked out all the methods for trend detection, we can use some of the data to illustrate the nature of the distributions over time. Vermont's database consists of data for many lakes that were sampled each year for up to nine years, from 1981 to 1988. When seven separate CDFs for Secchi disk transparency are overlaid, they display a "spaghetti" pattern (Figure 5-14, upper panel). Other ways of displaying these data are more useful for illustrating trends (or lack thereof), such as trends in particular quartiles (Figure 5-14, lower panel), or in the proportion of lakes that exceeds (or is less than) a particular critical value, such as an impairment criterion.

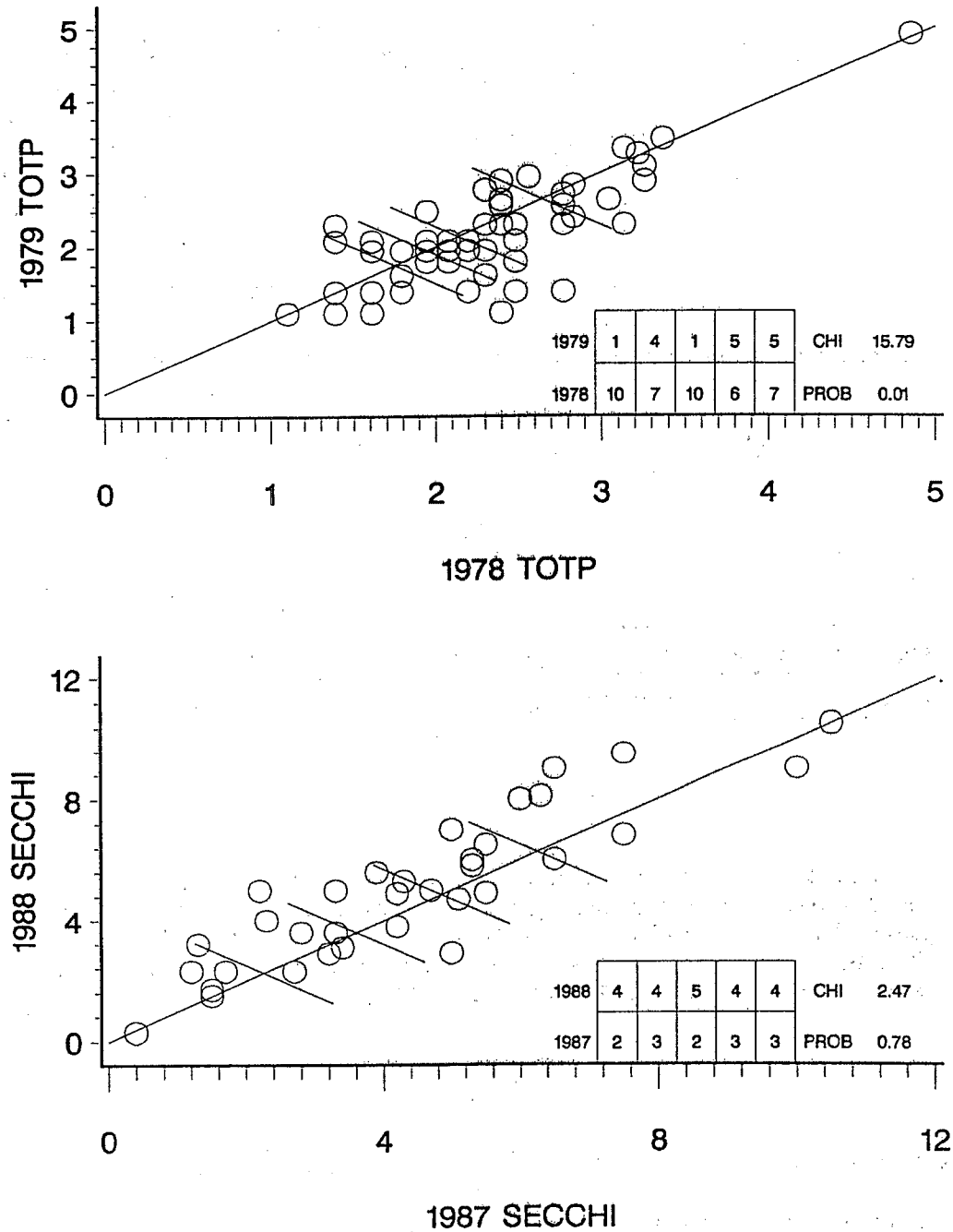


Figure 5-13. Cumulative distribution functions (CDFs) taken from the Vermont lakes data set on trophic condition showing comparisons made on paired sites: (a) total phosphorus during March, comparing 1978 and 1979, and (b) Secchi disk transparency during July, comparing 1987 and 1988. In one case, the test indicates high probability of differences between the two years (total phosphorus), but in the other, differences are not detectable.

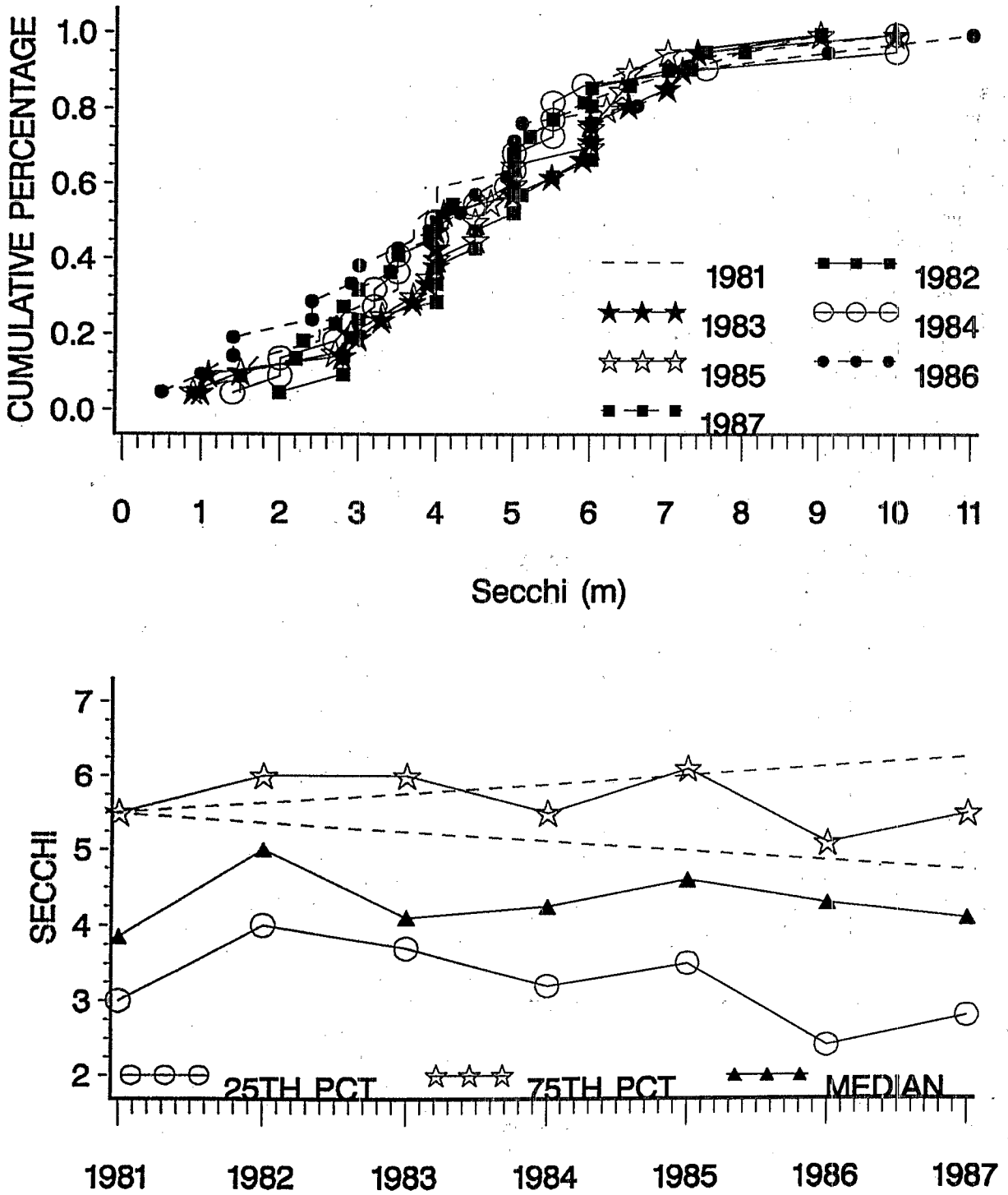


Figure 5-14. Superimposed yearly cumulative distribution functions (CDFs) show the confusing "spaghetti" pattern in multiple overlays of CDFs, based on nine years of data on Secchi disk transparency in Vermont lakes (upper panel). Results obtained from yearly CDFs can be more clearly presented to illustrate trends, or lack of trends, by selecting yearly values representing different quartiles (lower panel). The dashed lines represent a 2% per year change in Secchi disk transparency, for comparison with the trend graphs.

6.0 FIELD SAMPLING DESIGN

This section describes the design approach for gathering field information and for obtaining chemical and biological samples from surface water sites selected for field visits in Tier II of EMAP. We explain where and when field measurements will be made on these lake and stream sample units. The surface water sample units to be visited are lakes, reservoirs, and stream segments within several size classes as described in Section 3.0. The actual measurements to be made and the chemical/biological samples to be collected are described in Section 4.0.

We should view the data quality objectives (DQOs) for the field sampling design in the context of (1) the DQOs for the spatial design (site selection) and (2) the choice of EMAP indicators. For total quality assurance (QA), the DQO of Tier 1 sample site selection (Section 3.0) is to provide a set of field sampling sites representative of the populations of surface waters within the regions of interest. The DQOs of indicator design are to provide practical measures that accurately and precisely measure the ecological condition, exposures, and stresses at these sampling sites. The field sampling design described in this section defines, within each sample unit (lake or stream segment), the specific places and times at which these indicator measurements will be made. The overall DQOs of field sampling design are to provide adequate representation of the sites in order to classify the condition of the sites and make robust estimates of surface water status and trends on national and regional levels.

Biological, chemical, and physical characteristics of surface waters vary both temporally and spatially. The degree to which sampling must incorporate this variability depends upon how robust the indicator measurements are and what information is required of them. The field sampling design is <u>not</u> intended to provide adequate sampling frequency and coverage to describe this variability at each individual site, unless changes in this variability are a major ecosystem response to environmental stress. Ideally, indicator measurements might be made at single well-chosen points in space and time within each lake and stream. However, temporal or spatial variability in the indicator measurements at individual sites may prevent robust site condition classification and thus regional status descriptions. In such cases, several courses may be taken:

- Take more samples and express the index value for each water body as the mean of these replicate measurements (e.g., sample more frequently if temporal variability is high; sample more locations within the lake or stream if spatial heterogeneity is high).

- Target the sampling to a more specific time and location. Temporal sampling may be targeted more to the time of greatest biological stress, or it may be targeted to a more stable time period that can be associated with stressful conditions by calibration. Spatial sampling within heterogeneous lakes or streams may be stratified to collect separate sets of information from different habitat types (e.g., littoral versus profundal, riffle versus pool). Similarly, sampling may be aimed specifically at one habitat type of greatest interest, or one that is most influenced by the condition of surrounding habitats.

- Deconvolute, or correct bias in population distributions calculated from the sample data, based on quantification of the magnitude of index measurement variance in a subset of sample sites. Deconvolution can be used to compensate for index measurement error in population distributions, but no analogous procedure can be used to extract trends from excessively "noisy" index data.

Our basic design is to sample all surface waters once in the late summer period. Water and sediment samples from lakes and reservoirs will be collected in a mid-lake location judged to approximate the deepest point of the lake. Biological samples will be collected from representative major habitat types within lakes. Physical habitat indices will be used to describe the lake and its riparian zone as a whole (Section 4.0). In streams and rivers, the sampling location will be a reach, 30 channel-widths long, located midway between the mapped upstream and downstream confluences defining the stream segment that is the sample unit. Water samples will be taken at a single, mid-channel location in the middle of this

reach, and biological samples will be systematically sampled in representative habitat types within the reach. Physical habitat characteristics will be assessed over the entire 30-channel-width section and its riparian zone.

6.1 INDEX CONCEPT

6.1.1 General Design Objectives of Index Sampling

Many broad-scale surveys and trend monitoring networks have employed indices to characterize sample units, limiting sampling effort on individual units in order to maximize regional coverage with available personnel and financial resources. The index values obtained in such studies are intended as single values of each exposure or response indicator variable that are robust enough to represent each sample lake and stream for the purposes of classifying its condition, tracking trends, and extrapolating these sample data to regional and national scale descriptions. There are both temporal and spatial aspects of index measurement, and the optimal field sampling design depends upon the objectives of monitoring, as well as the variability of the indicators to be measured.

The overall DQOs of EMAP require that indicator measurements be relatively stable, but responsive to year-to-year changes in biotic stress and exposure. The relationships among stressor, exposure, and response indicators have been discussed in Sections 2.0 and 4.0, as has the interpretation of ecological condition based upon response indicator measurements. As they apply to the field sampling design, EMAP's overall DQOs require that the indicator measurements be taken at locations and times that balance often opposing requirements for indicator stability, spatial representativeness, and relevance to biotic stresses.

In the temporal context, an ideal index period is a time during which the values of response and exposure indicators are relatively stable. It must be a time period when indicator biota are present and measurable. To facilitate diagnosis of probable causes of subnominal indicator values, it is advantageous for the index time period to also be a season in which indicator organisms are exposed to maximum environmental stress. In some cases the periods of maximal environmental stress to indicator organisms may be very short-lived or episodic. In these cases it is advantageous to choose an index period for which conditions may be related in a predictable way to conditions during the most stressful times, so that exposure might be inferred by calibration from other temporally intensive sampling.

In the spatial context, an index value should provide in a single measurement, or an integration of a number of measurements, information adequate to represent each sampling unit in classification, regional extrapolations, or trend analysis. If the sampling unit is heterogeneous, the index measurement should, through multiple sampling points, sufficiently incorporate the spatial variability of indicator measurements within the sampling unit (lake or stream). Alternatively, the measurements might be made at locations which themselves integrate the ecological stresses impinging on the sampling unit as a whole. For example, pelagic fish and mid-lake water quality might reflect benthic biogeochemical processes, chemical inputs from streams, and littoral spawning, rearing, and feeding areas.

6.1.2 Applications of the Index Concept in Synoptic Assessments

Several examples are useful for illustrating the application of the temporal and spatial aspects of the index concept.

6.1.2.1 National Stream Survey (NSS)

The objective of the EPA's National Stream Survey (NSS) (Kaufmann et al. 1988, in press) and other acid rain related stream surveys (Knapp et al. 1988) was to index conditions of biological exposure to low pH, high aluminum, and other chemical conditions related to stream acidification associated with acidic deposition. Streams were sampled during "baseflow" (not storm flows) during the spring season after snowmelt but before leafout. Chemistry during this index time period was shown to be quite stable in southeastern streams (Messer et al. 1986, 1988), and populations could be adequately described based on single spring baseflow measurements in sample streams. In the mid-Atlantic region, where spring season chemical variability was greater (Kaufmann et al. 1988, Baker et al. in press), stream sampling sites were characterized by averaging two spring baseflow measurements.

At each sampling visit, the NSS measured characteristics at the upstream and downstream ends of each stream segment (sample unit). Water chemistry samples were taken at mid-depth and mid-stream, ignoring spatial heterogeneity that might be present in relatively small volumes of water (relative to seasonal runoff) outside the main channel flow. On the average, the distance between measurement locations (i.e., the distance between confluences on 1:250,000-scale maps) was 3 km. In the low ANC streams of greatest interest in the NSS, the median downstream chemical change was +5 μeq/L ANC and +0.06 pH units per km of stream length. To incorporate these characteristic downstream trends, chemical index values measured at each end of the sample stream segments were used to estimate, by interpolation, the chemical characteristics of the whole segment.

Chemical conditions more stable than spring baseflow might have been observed during the summer baseflow period in NSS streams. The intent of the NSS index, however, was not only to minimize within-season and episodic variability, but also to maximize the probability of sampling chemical conditions potentially limiting for aquatic organisms. For a spring index stream sampling period to be biologically relevant, sensitive life stages of aquatic biota must also be present during the sampling period. Spring is the season when acid-sensitive swim-up fry of many sport fish are present in streams within these regions, although fry of some trout (*Salmo gairdneri*, *S. trutta*, *Salvelinus fontinalis*) populations may also be present at other times of the year. Prolonged periods of low pH and ANC are most likely to be observed in the winter and spring seasons in streams within the Northeast, Mid-Atlantic, and Southeast (Colquhoun et al. 1981, Murdoch 1986, Witt and Barker 1986, Olem 1986, Nix et al. 1986, Ford et al. 1986, Kaufmann et al. 1988).

The NSS effectively indexed the chronic exposure of aquatic biota to acid stresses. However, the most stressful times of exposure for fish fry are probably not during spring baseflows, but rather during episodic pH depressions that occur during spring storm flows. It is not practical to sample storm episodes on a synoptic scale (Messer et al. 1987), but episodic conditions can be inferred to some extent from spring baseflow index chemistries using hydrologic models that also take into account the chemistry of precipitation (e.g., Eshleman 1988, Gerritsen et al. 1989).

In contrast to biologically relevant chemistry (in the context of acidic deposition), general biotic condition in response to acid stresses need not necessarily be measured during the actual time of maximum stress. The presence and abundance of nonanadromous fish species have in most cases been measured during the traditional summer field season. Although fish presence and absence must be interpreted in relation to other natural and anthropogenic stresses, this biological information can be compared with spring season pH and then related to modelled episodic conditions (e.g., Baker et al. in press, Morgan and Janicki in press).

6.1.2.2 National Lake Survey (NLS)

The EPA's Eastern and Western Lake Surveys (Linthurst et al. 1986, Landers et al. 1987) employed a single mid-lake sample taken at a depth of 1.5 m during fall turnover to represent the nonlittoral lake water volume. The relatively well-mixed fall turnover period was targetted to minimize spatial (particularly vertical) variability. Although sampling might also have been undertaken during spring mixing following ice-out, the fall period had the advantages of (1) relatively stable water chemistry over a sampling window of several weeks, (2) ease of access, and (3) safety. Unlike the strategy used in the NSS, the National Lake Survey (NLS) did not directly target the season of maximal acid stress. As in streams, spring pH minima in lakes often coincide with the presence of acid-sensitive developmental stages of many fish species. However, the robust chemical classifications of lakes based on fall sampling can be related by calibration to the more biologically relevant acid stresses that occur during snowmelt (Eshleman 1988), or in the spring season (Baker et al. in press, Herlihy et al. in prep.). In addition, fish species presence and abundance measured in traditional mid-summer surveys have effectively been related to fall index chemistry. As for streams, the influence of acid stresses must be evaluated within the context of other natural and anthropogenic stresses such as fishing harvest and physical habitat quality changes.

6.2 TEMPORAL CONSIDERATIONS FOR FIELD SAMPLING

Sections 2.0 and 4.0 discuss the rationale for measuring EMAP-Surface Water response and exposure indicators in the late summer season. An EMAP design goal is to target a time period of high stress to aquatic biota. This design goal facilitates the diagnosis of probable cause of ecological impairment. The index period must also be a time when robust measurements of ecological condition can be made.

For a number of biotic stresses, late summer is generally a time of high stress due to many coincident factors. In most of the United States, streamflows are at a minimum during the late summer and fall (Figure 6-1). The major exceptions are the Rocky Mountain region, much of Alaska, and parts of the Great Plains, where almost all precipitation is locked up in the form of snow until the spring thaw. Lowest lake and reservoir levels tend, with some lag, to correspond with the season of lowest stream runoff. Furthermore, the highest stream and lake water temperatures tend to coincide with these seasons of lowest runoff in much of the United States. Evapotranspiration and irrigation withdrawals from surface waters are also at a high during the late summer, as are the pollutant inputs from irrigation returns and municipal waste disposal. Low flows and lake levels minimize the dilution of point and nonpoint source pollutant inputs, and high temperatures increase metabolic oxygen demands, while at the same time decreasing the solubility of dissolved oxygen. High water temperatures and resulting rapid metabolic rates increase the exposure of aquatic organisms to toxins and increase the toxic effects of many substances. The potential for algal blooms and hypolimnetic deoxygenation increases during the summer period of thermal stratification in eutrophic lakes and reservoirs. These numerous stress factors converge to make late summer to early fall the period of maximal stress to aquatic biota in many surface waters.

Late summer/early fall stress is common and widespread in surface waters, but this season is by no means the only period of important biotic aquatic stresses in the United States. Examples of important exceptions are in the areas of surface water acidification, lake eutrophication, and stream physical habitat alteration.

Spring is the season in which biotic stresses due to surface water acidification from deposition are typically greatest. Surface waters in areas most affected by acidic deposition are generally most dilute during spring and provide less ANC than during summer and fall. Winter may be the time of critical stress in eutrophic lakes that freeze, because biochemical degradation of organic matter may result in low dissolved oxygen concentrations under the ice that are limiting to fish. Winter is also the critical season of stress in many small high-elevation lakes that freeze to the bottom; in fact, most of these lakes do not support naturally reproducing fish populations. Removal of riparian vegetation from small northern and subarctic streams can result in fish habitat loss due to the formation of anchor ice as heat loss by radiation is increased in these streams, making winter the season of maximal stress for many fish. Similarly, channel and riparian zone activities that reduce physical complexity in streams may reduce slackwater habitat during winter and spring high flows, making these high-flow seasons the periods of maximal stress for fish. For example, removal or depletion of large woody debris in Pacific coastal streams has resulted in reduction of winter cover from high water velocities for anadromous and resident salmonid fishes.

6.3 INDEX SAMPLING DESIGN FOR LAKES AND RESERVOIRS

6.3.1 Temporal Variability

Temporal variability in the lake and reservoir characteristics to be measured by EMAP indicators occurs over a range of time scales. Variability over decades, years, seasons, and within seasons may affect the precision and interpretation of status and trends in condition of lakes and reservoirs (Sections 5.3 and 5.4). Table 6-1 lists examples of time scales of variability and some common sources of variability. 6-1. Some of these sources are cyclical, and might be considered "noise"; others may be trends that EMAP will want to quantify.

6.3.2 Summer Lake/Reservoir Index Period: The Proposal

Based upon the information discussed in Sections 6.2, 4.0, and 5.0, we propose a mid- to late summer index period (generally from July to mid-September) for sampling lakes and reservoirs in most of the United States.

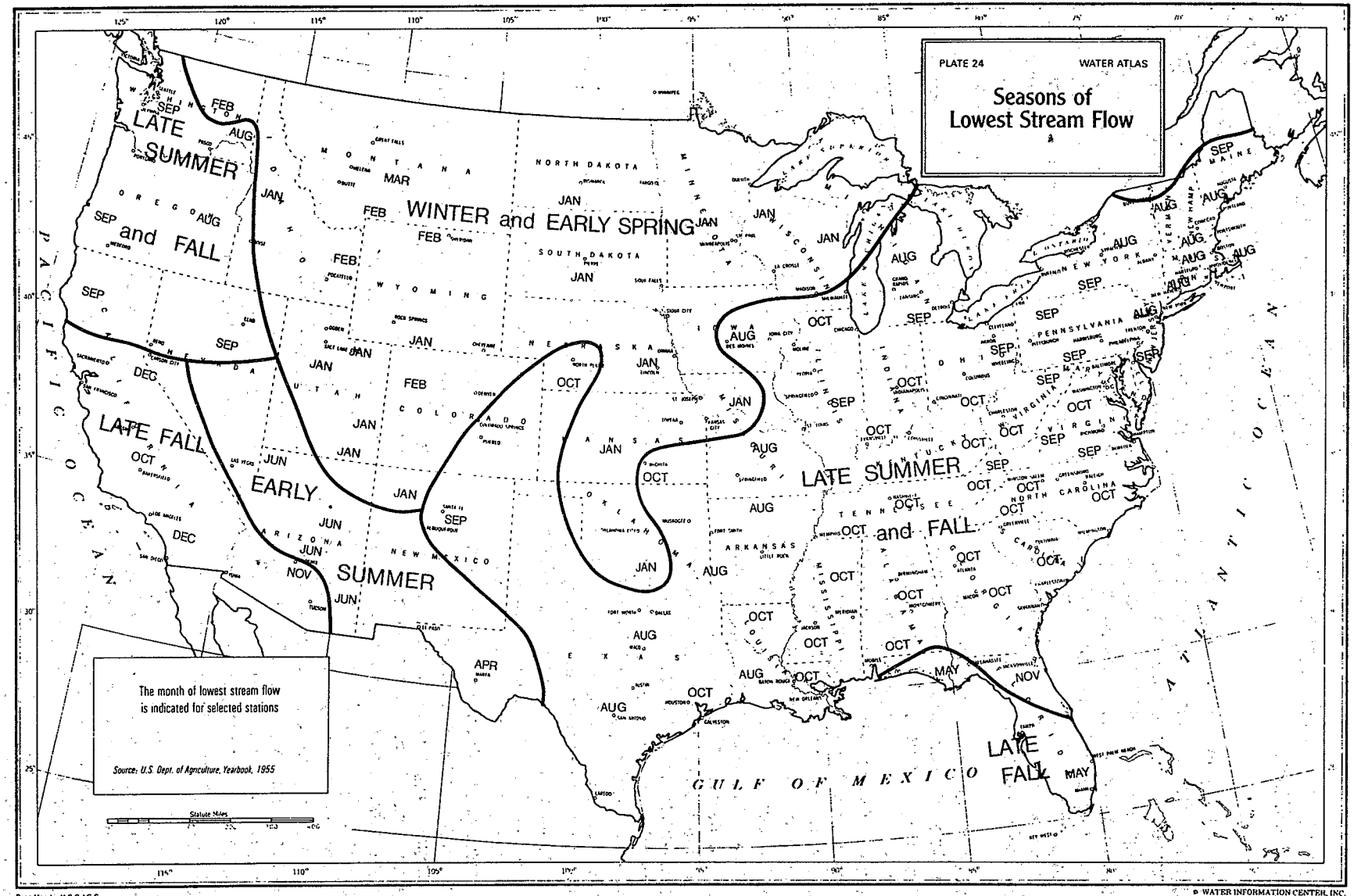


Figure 6-1. Seasons of lowest stream flow in the conterminous United States (Plate 24 from Geraghty et al. 1973).

Table 6-1. Time Scales and Sources of Variability in EMAP Indicator Characteristics -- Some Examples

Time Scale	Source of Variability
Centuries and longer	• Tectonic activity that changes physiography • Continental glaciation • Climatic change • Changes in species pool resulting from migration, extinction, evolution • Residual effects of "resetting" by fire, volcanism
Years to decades	• Cyclic climatic variation (e.g., ENSO) • Land use changes • Residual effects of "resetting" by fire, floods, mass movements, volcanism
Within/Among seasons	• Hydrologic inputs, floods, evapotranspiration • Lake/Reservoir stratification • Seasonal cycles of plant production and die-off, with associated chemical uptake and release • Competitive, synergistic interactions among phytoplankton, macrophytes, and other groups • Grazing, secondary consumption • Life history, developmental stages, macroinvertebrate emergence, fish migration
Diel	• Wind, wave action, temperature, sunlight, seiches • Plant production, uptake, release • Mobility of organisms (active and passive)

We plan pilot studies for each of the major regions before or at the time of full EMAP implementation in those regions. Available information on indicator variability suggests that, for most of our proposed indicator measurements, one visit to each site within the summer index period will probably be adequate to describe regional status. Seasonal sampling and within-season replication in a subset of approximately 20 to 30 of the first year's sites within each region will be aimed at testing the relevance, robustness, and practicality of the proposed index period and sampling frequency. The questions that will be re-evaluated with regard to our choice of index sampling time are as follows:

- Relevance
  - Is the season a period of stressful conditions for biota?
  - Are sensitive indicator species present in sufficient number to be collected effectively?
- Robustness: Is temporal variability sufficiently small to yield acceptable accuracy and precision in regional extrapolations of surface water condition?
- Practicality
  - Is the index sampling period long enough to allow sampling crews to complete planned site visits?
  - Is index period weather suitable for safe sampling?
  - Can field crews find, reach, and gain access to the sampling sites?

- Are enough willing, qualified field technicians available during the index time period?

Our answer to all the above questions is a qualified "yes" for single site visits during the summer index period for lakes and reservoirs in the Northeast, where the first EMAP regional pilot will be undertaken. Pilot activities will test the effectiveness and feasibility of that choice. Except for the arid West, where a spring period may be advantageous for flowing waters, the summer period is our likely choice in the other regions of the country.

6.3.3 Spatial Variability within Lakes and Reservoirs

Spatial variability in lakes and reservoirs can be in both vertical and horizontal dimensions. Vertical variability in temperate lakes occurs primarily because of temperature stratification. Summer stratification occurs with warming of surface layers of water, whereas less marked and less stable winter stratification occurs when water at temperatures < 4°C lies above warmer water (this is usually under ice). Strong summer stratification severely limits vertical mixing, so epilimnetic and hypolimnetic processes become largely separated. Shallow and deep waters take on different physical, chemical, and biological characteristics that reflect the different light levels, atmospheric gas exchange, hydrologic inputs, and biological activities in upper and lower layers. Phosphorus and heavy metal concentrations in hypolimnetic waters may greatly exceed those in the epilimnion if organic decomposition consumes oxygen and allows release of these substances from bottom sediments under reducing conditions. Stratification is usually most pronounced in lakes or reservoirs that are deep in relation to their surface area or length dimensions. Additionally, lakes experiencing only gentle winds due to their climate or topographic setting tend to become more strongly stratified than exposed lakes in windy regions. The depth of thermal stratification can be affected by turbidity. Turbid waters may stratify more intensely than clear waters because solar energy is more rapidly attenuated in the surface waters.

Phytoplankton species and abundance may vary markedly at different depths, even in waters that are not thermally stratified. These differences are the result of differential growth rates, sinking rates, and mobility at different water depths and temperatures and among different planktonic species. Chlorophyll concentrations reflect the differences in algal biomass.

Horizontal variation in lake characteristics measured by EMAP indicators occurs because of a wide variety of factors, including lake size, morphometry, and substrate, and the location and magnitude of inlets and outlets. The spatial variability of EMAP indicator measurements directly associated with water column characteristics (e.g., chlorophyll, temperature, water chemistry) may tend to increase with lake size, particularly if shallow depths and complex morphometry impede lateral mixing. For example, water clarity and phytoplankton production rates often vary markedly at different locations in large lakes or reservoirs with numerous, relatively isolated arms that receive different hydrologic and chemical inputs. Similarly, depth-integrated samples taken at near-shore locations in lakes with gradually sloping bottoms will show a greater influence of littoral epilimnetic processes than those which remain deep very close to shore.

In contrast to measurements of chemistry or phytoplankton in the water column, horizontal variability in measurements of fish or benthic macroinvertebrate assemblages are not as strongly dependent upon the degree of lateral mixing. Rather, they are more strongly dependent upon the spatial heterogeneity of substrate, the water depth, and the distribution of macrophytes and other elements that provide habitat structure and cover. Fish, being highly mobile, may move from profundal to littoral habitats to find food and cover at different times of the day and year. The mobility of fish will often be interpreted as temporal variability if the spatial scale of sampling is not large in comparison with the range of short-term migrations. Different developmental stages of the same fish species may be found in very different types of habitat within the same lake because the stages have different requirements and preferences for water depth, temperature, chemistry, food, and cover.

6.3.4 Spatial Sampling Index for Lakes and Reservoirs: The Proposal

Spatial aspects of our proposal for measuring EMAP indicators differ according to the specific type of indicator measurement:

- Fish assemblage: Fish samples will be collected from profundal and major littoral habitats (beach, rock, rubble, emergent macrophytes, submergent macrophytes). Samples will remain separate during pilot studies. Pilot research will be used to determine how to incorporate fish assemblage data from different habitat types into metrics that index whole-lake biointegrity.

- Benthic macroinvertebrate assemblage: Macroinvertebrates will be collected from profundal and major littoral habitats as will fish. Replicate benthic samples (n=5) from the same habitat type will be analyzed separately for a subset of 20 to 30 lakes in each region in order to evaluate sampling variation. Replicate samples in the remainder may be combined before counting and identification. As for fish assemblage data, pilot research will be used to determine how to incorporate macroinvertebrate assemblage data from different habitat types into whole-lake biointegrity indicators.

- Sedimentary diatom assemblage: Sediment cores and surficial sediments for analysis of algal remains will be collected from profundal habitats only. Replicate sediment cores (n=3) will be analyzed separately in a subset of 20 to 30 lakes in each region to assess spatial variability. For the remaining lakes, a single mid-lake core will represent the lake.

- Trophic state: Water samples will be collected at one mid-lake location. Secchi disk transparency will be replicated at two to four additional sites where profundal sediments are sampled.

- Physical habitat: Measurements as outlined in Section 4.0 are designed to describe the entire lake, including profundal, littoral and riparian areas. Additional characterization activities will describe topographic .drainage basin of lake from maps, aerial photographs, and other studies.

- Chemical habitat: Single water samples will be collected mid-lake at 1.5 m depth. Profiles of dissolved oxygen, specific conductance, pH, and temperature will be measured at the same location. We will assess index variability in a subset of 20 to 30 lakes in each region. We will examine spatial variability by taking multiple profiles in different parts of the lake. We will assess temporal variability in these lakes by taking mid-lake samples at two different times in the index period. We will treat large lakes (surface area > 2,000 ha) differently, obtaining samples from a number of subjectively chosen representative profundal locations (e.g., separate arms of a large reservoir).

6.4 INDEX SAMPLING DESIGN FOR STREAMS AND RIVERS

6.4.1 Temporal Variability

Streams and rivers, like lakes and reservoirs, exhibit temporal variability in the characteristics to be measured by EMAP indicators. Over the longest time scales (years, decades, centuries, and longer), the sources of this variability are much the same as those presented as examples for lakes. Rivers and streams are generally more persistent than lakes on a geological time scale and their long-term characteristics are probably more stable. However, these flowing waters have much smaller hydrologic retention times than lakes, so the amount and character of water in their channels more closely reflects fluctuations in the sources and sinks for water, energy, and chemical substances. As a result, flowing waters often tend to exhibit greater seasonal, within-season, and diel variation than do most lakes, for those characteristics that depend on the amount and flowpath of hydrologic inputs. Important exceptions to this pattern are shallow, ephemeral lakes, which, like other ephemeral waters, exhibit extremes in physical, chemical, and biological variability.

On time scales shorter than a year, characteristics of flowing water ecosystems measured by EMAP indicators respond to seasonal patterns in precipitation, runoff, water temperature, watershed

activities, riparian cover, riparian organic inputs, periphyton productivity, fish migration, and macroinvertebrate life histories. Virtually all characteristics of streams and rivers change to some extent with discharge. Short-term changes in discharge may cause marked changes in stream water chemistry. The chemical changes and increases in water velocity accompanying spates may cause increases in the downstream movement of macroinvertebrates (drift), and may cause fish to move downstream or into lower velocity backwaters or tributaries. Similarly, increases and decreases in discharge can cause marked changes in the extent and quality of physical habitat in a stream or river channel. Channel and bank scour and deposition from extreme flows can cause catastrophic change in macroinvertebrate, fish, and periphyton populations, and may greatly change the morphometry and flow characteristics of the channel when discharges subside. These periodic "resetting" mechanisms are necessary for the long-term maintenance of many ecosystems, despite their immediate detrimental effects. In regional aggregations, effects associated with the natural or altered disturbance regime become "noise," and should, in fact be sampled and characterized as part of the natural variability in regional surface water population descriptions.

6.4.2 Summer Baseflow Stream/River Index Period: The Proposal

Based upon available information discussed in Sections 4.0 and 5.0, we propose a mid- to late summer index period for sampling streams and rivers in most of the United States. This period is generally from July to mid-September. As for lakes, we plan pilot studies in each of the major regions before or at the time of full EMAP implementation in those regions. Available information on indicator variability suggests that, for most of our proposed indicator measurements, one visit to each site during baseflow within the summer index period will probably be adequate to describe regional status. Seasonal sampling and within-season replication in a subset of approximately 20 to 30 of the first year's sites within each region will be aimed at testing the relevance, robustness, and practicality of the proposed index period and sampling frequency. The following questions, which will be re-evaluated with regard to our choice of index sampling time, are almost the same as those for lakes and reservoirs:

- Relevance
  - Is the season a period of stressful conditions for biota?
  - Are sensitive indicator species present in sufficient number to be collected effectively?
  - Are important flowing waters dry during the summer and better sampled during a wetter season?
- Robustness
  - Is temporal variability small enough to yield acceptable accuracy and precision in regional extrapolations of surface water condition?
  - Are stream and river flows and physical habitat quality and extent stable enough to provide a reasonable physical template for biota within the summer index period?
- Practicality
  - Is the index sampling period long enough to allow sampling crews to complete planned site visits?
  - Is index period weather and stream flow suitable for safe sampling?
  - Can field crews find, reach, and gain access to the surface water sampling sites?
  - Are enough willing, qualified field technicians available during the index time period?

As for lakes, our answer to all the above questions is again a qualified "yes" for single site visits during the summer index period for streams and rivers in the Northeast, where the first EMAP regional pilot will be undertaken on lakes. Pilot activities in each region will test the effectiveness and feasibility of the chosen index season. For example, spring sampling will be evaluated for streams in the arid West and for coastal streams with anadromous fish populations. Pilot studies will be undertaken in each region to re-evaluate

the summer period as our likely choice in the Northeast and other regions of the country.

Within the summer index season, stream chemistry may change considerably in some regions. This effect might be minimized by choosing relatively stable low flows of summer, but the profound effects of summer plant and microbial growth cycles will have a large effect on the low flows of summer.

6.4.3 Spatial Variability within Streams/Rivers

6.4.3.1 Upstream-downstream Variations

Stream discharge, gradient, channel width, water depth, substrate, and many other physical characteristics change in a predictable fashion along the course of a stream from its headwaters to its mouth (Strahler 1957, Leopold et al. 1964). Similarly, geology, soils, vegetation, and the influence of groundwater on streamflows change markedly from stream headwaters to the mouths of rivers. Recognizing these physical and chemical patterns, Vannote et al. (1980) proposed the Stream Continuum Hypothesis, stating that the structure and function of biological communities in streams should reflect these upstream-downstream patterns of variation. The reach segment design proposed for sampling small and mid-sized streams in EMAP was chosen for its efficiency in capturing broad-scale regional variations in stream characteristics. It is not the ideal design for describing downstream changes within individual streams. However, examination of longitudinal (downstream) variation in streams is incorporated into the EMAP stream sampling design through stream size stratification (see Section 3.2). Though it is unlikely that EMAP will be able to provide complete, detailed descriptions of the downstream trends in any but the largest river systems, major patterns of upstream-downstream variation will be incorporated into the estimates of regional condition. In addition, specific subsets of stream size can be compared to detect those upstream-downstream trends that are generally consistent among river systems in a given region.

The population sampling units for streams and rivers in EMAP are segments between mapped confluences. A similar design used in the NSS showed fairly consistent qualitative patterns of change in physical and chemical characteristics between upstream and downstream sampling points an average of 3 km apart (see subsection 5.2.2 and Kaufmann et al., 1988). The NSS sample units were segments (reaches) depicted on 1:250,000-scale USGS topographic maps. Though differences between upstream and downstream sample sites were often large, they were small in relation to differences among streams. The imprecision resulting from sampling only one point on each segment would cause only a small distortion of regional population descriptions of stream chemistry.

Despite the fact that NSS results showed that upstream-downstream variation on stream segments was relatively insignificant compared to regional variation (Table 5-5 and subsection 5.2.2.3), interpreters of NSS data found it necessary to use both upstream and downstream chemical information in an assessment of the acid-base status of streams. The NSS experience has bearing on the design for EMAP. In contrast to the EMAP target resource, the NSS targeted a narrower range of stream size (only streams with watershed areas < 155 km^2, but large enough to appear on 1:250,000-scale maps). In addition, the NSS showed that, outside of streams acidic because of mine drainage or natural organic acids, acidic conditions in streams were overwhelmingly concentrated at the upstream portions of the stream segments examined. The biggest disadvantage of sampling only downstream ends of NSS segments would have been the complete avoidance of most of the acidic stream resource--rather than simply poor precision in approximating stream segment characteristics by one point measurement. The NSS needed upstream as well as downstream measurements on sample stream segments in order to extend the scope of inference to streams at the small end of its targeted range. A further benefit, of course, was that relatively precise estimates of the extent of acidic stream length could be made. Because EMAP targets a broader size range of streams than NSS and employs stream size stratification in its design, there is less of a disadvantage in sampling one point on each stream segment--at least for chemistry. Furthermore, the average length and the variation in length of sample stream segments in EMAP could be reduced, compared with NSS, if EMAP used finer resolution 1:100,000-scale maps as a basis for its sampling frame.

In contrast with streamwater chemistry, stream channel morphology and therefore biological assemblages may show cyclical spatial patterns

between mid-segment and confluence areas at the ends of segments. Confluences are often hot spots of biological activity because of tributary inputs of sediment and organic debris. We will need to investigate these effects of within-segment position in EMAP pilot research activities.

6.4.3.2 Other Stream Habitat Variations

We anticipate that a major source of variation in EMAP fish, macroinvertebrate, and periphyton assemblage measurements will be associated with physical habitat differences within each sample stream or river sample segment (reach). Channel features such as riffles, pools, and backwaters provide a wide variety of water velocities, depths, and substrates favoring different biotic assemblages. These habitats also differ in their ability to physically and biochemically retain dissolved and particulate organic matter and other nutrients. Riparian interactions differ among channel habitats and along the course of stream channels.

Stream ecologists frequently report that macroinvertebrate and periphyton assemblages differ more between contrasting habitats (e.g., riffles versus pools) than among similar habitats in streams across often large regions. For this reason it will be necessary for EMAP's field design to stratify sampling within individual stream sample units, or sampling could specifically target one habitat type. Riffle habitats are often sampled in synoptic macroinvertebrate surveys because they often contain numerous and diverse assemblages of species, and because they can be sampled relatively easily using consistent methods from stream to stream.

Single habitat sampling would be an acceptable design for EMAP if the program were interested mainly in ecological changes associated with exposure to changes in water chemistry, temperature, and other characteristics that do not necessarily change the relative abundance of habitat types. To the contrary, EMAP plans to monitor ecosystem changes caused by changes in physical habitat. For this reason, initial field activities will separately sample major physical habitat types within each stream sample site. This pilot research will investigate whether assemblages within particular habitat types may perform the function of overall condition indicators in reflecting changes that occur in the relative abundance of other physical habitat types or the condition of their biota.

6.4.4 Spatial Sampling Index for Streams and Rivers: The Proposal

We propose to represent stream and river segments by sampling a mid-segment location. Logistical field trials and pilot sampling efforts will be undertaken to determine whether, for practical reasons, measurements should instead be taken at the downstream end of the sample stream segment (i.e., just upstream of the confluence that defines the downstream end of the mapped segment). It is often very difficult to find and then access mid-segment positions on small streams with confidence.

Chemical habitat measurements and their associated water samples on each sample stream segment will be collected at a single well-mixed location in mid-channel.

For physical habitat measurements and biological sampling, a field sampling reach 30 channel-widths long will be defined midway between the upstream and downstream ends of each sample stream segment. For example, a field sampling reach approximately 100 m long would be identified on a typical second-order stream channel segment 3 km long that averages about 3 m wide. Similarly, the field sampling reach within a 5-km mapped segment of a major river 50 m wide would be 1,500 km long. Adjusting the field sampling reach length to the size of the stream allows sampling of a segment long enough to incorporate the morphometric and biological variability within the mapped stream segment.

Physical habitat measurements will evaluate stream channel morphometry, substrate, hydraulic retention, habitat complexity, and riparian interactions over the full length of the field sampling reach. Biological samples will be collected separately from major habitat types within the same portion of the stream segment. For three rough ranges of stream size, the types of habitats to be sampled are as follows:

- Small (wadable) streams
  - Main channel riffles
  - Main channel pools
  - Peripheral backwaters
  - Side channels

- Large streams (nonwadable midstream)
  - Main channel riffle accessed by wading from bank
  - Main channel pool accessed by wading from bank
  - Peripheral backwaters
  - Side channels
- Rivers
  - Natural cutbank, or periphery of pool
  - Natural bank on depositional side of bend, or periphery of riffle
  - Bank revetments
  - Navigable sloughs, backwaters, and side channels
  - Wadable backwaters

Sample collections for fish assemblages within major habitat types will remain separate during pilot studies. For macroinvertebrates and periphyton, sampling will be replicated (n=5) within each habitat type. These replicate benthic samples may be combined before analysis for most streams, but will be analyzed separately to evaluate heterogeneity within habitat type for a subset of 20 to 30 streams within each region. Pilot research will be used to determine how to incorporate fish, macroinvertebrate, and periphyton assemblage data from these different habitat types into metrics that index whole-stream condition, diversity, and biointegrity.

6.5 CONCLUSIONS

The field sampling design described in Sections 6.3 and 6.4 defines the specific places and times at which EMAP indicator measurements will be made within each sample unit (lake, reservoir, stream segment, river segment). In earlier sections of this report, we define the index concept (4.0 and 6.1) and qualitatively and quantitatively explore the accuracy, precision, and bias of individual site measurements in a number of synoptic applications of this concept (5.2.2). Using statistical theory and examining simulations in the context of EMAP goals, we estimate the effect of indicator measurement on the expected precision and utility of estimates of population status (5.3 to 5.5).

Our basic design for estimating population status is to sample a statistical subset of surface waters (Section 3.0) once in the late summer season. Preliminary data discussed in Section 6.2 suggest that for most EMAP indicators this sampling frequency should be adequate to meet the goals of EMAP in most regions, if some systematic spatial replication within sample sites is employed. However, the precision of EMAP indicator measurements will be evaluated through temporal replication on subsets of 20 to 30 sample lakes and streams in each region.

Physical habitat measurements and resulting indices of physical habitat quality will be designed to be representative of each sample lake and stream and its associated riparian zone.

Water and sediment samples from lakes and reservoirs will be collected from a mid-lake location judged to approximate the deepest point of the lake. In rivers and streams, water samples will be taken at a single well-mixed mid-channel location in the middle of each mapped stream segment. Pilot study spatial replications within sample units will re-evaluate the suitability of these mid-lake and mid-reach index locations. Fish, macroinvertebrates, and periphyton will be collected separately from major habitats within each lake or stream. Macroinvertebrate and periphyton sampling will also be replicated within habitat types. These benthic replicates will be analyzed separately for a subset of sample sites, and the remainder may be combined before identification and counting. Pilot research will be used to re-assess the adequacy of temporal and spatial sampling. Pilot research will also be used to incorporate surficial sediment diatom information, and habitat-specific information on fish, macroinvertebrate, and periphyton assemblages into metrics that index whole-stream or whole-lake condition, diversity, and biointegrity. The challenge is to develop biological metrics that quantitatively index the changes in communities resulting from changes in habitat availability, as well as other exposures and stresses.

7.0 LOGISTICS APPROACH

This section describes the logistics approach for implementing EMAP-Surface Waters. It includes a summary of requirements for the logistics plans (Section 7.1), a discussion of the major logistics issues (Section 7.2), a field operations scenario for lakes and streams (Sections 7.3 and 7.4), and a proposed organization structure (Section 7.5).

7.1 LOGISTICS IMPLEMENTATION COMPONENTS

Implementing the EMAP Surface Waters program will require detailed, comprehensive logistics planning. Logistics considerations include coordination and oversight of all implementation support activities (access permission, procurement, etc.) and the actual data collection activities. A logistics plan must be developed before the start of field activities to assure that program goals will be met. Regional logistics plans will be updated annually. The logistics plan will address all elements given in Table 7-1 as specified by U.S. EPA (1990).

ELEMENT 1. Overview of Logistics Activities: Summarize the types of activities required to complete the project. Maintain a timeline or Gantt chart showing all critical path milestones (e.g., project design, indicator selection, site selection, access permission, reconnaissance, procurement, methods selection, development of standard operating procedures, and resolution of specific quality assurance [QA] issues). Show required deliverable products such as plans, manuals, and reports. Provide logistics budget summaries.

ELEMENT 2. Staffing and Personnel Requirements: Describe the number of personnel and the organizational structure necessary to accomplish project objectives. Define who is responsible for staffing and interagency and teaming mechanisms. Consider work schedules to determine whether extra positions should be created or whether existing personnel should work overtime. Create a contingency plan for replacing staff members when necessary. Identify key personnel and provide plans for retaining them.

ELEMENT 3. Communications: Address communications among field crews, laboratory crews, and supervisory personnel, and between EMAP participants and any local organizations who should be informed of EMAP field activities. Include plans for tracking samples, data, crews, and equipment and supplies. Discuss how field crews should interact with the public or with the media. Explain how approved changes in standard operating procedures will be documented and communicated for implementation.

Table 7-1. EMAP Logistical Elements for Implementation of Surface Waters Monitoring Programs

1.	Overview of logistical activities
2.	Staffing
3.	Communications
4.	Sampling schedule
5.	Site access
6.	Reconnaissance
7.	Waste disposal plan
8.	Safety plan
9.	Procurement and inventory control
10.	Training
11.	Field operations
12.	Laboratory operations
13.	Data management activities
14.	Quality assurance
15.	Logistics review/recommendations

ELEMENT 4. Sampling Schedule: Based on project, indicator, and statistical design or other program requirements, devise an efficient schedule for field activities. Consider geographical sampling windows within geographical areas and other factors such as climate and site access constraints.

ELEMENT 5. Site Access: Address issues related to gaining access to sampling sites, including scientific collection permits, if required. Develop a list of local contacts to discern property ownership, jurisdiction, and the best site access methods. Address plans to obtain appropriate access permission and applicable collection permits. Consider how to coordinate activities in the same area of more than one resource task group. Discuss ways to arrange long-term access

rights, track changes in ownership of private sites and management of public sites, notify owners and managers before revisiting the sites for future monitoring, and provide contingency plans in case of future failure to obtain re-access permission.

ELEMENT 6. Reconnaissance: Define criteria for selecting base operation sites (considering personnel and technical support requirements), determining geographical location with respect to sampling sites, and defining time constraints imposed by sampling design or climate. Sampling sites identified as having potentially difficult physical or legal access should be visited during field reconnaissance. Additional resources needed for sampling should be identified if the access problem is due to physical conditions. If the access problem is legal, one last attempt should be made to obtain permission to sample.

ELEMENT 7. Waste Disposal Plan: Explain how chemical and biological wastes will be stored, transported, and disposed of safely and legally. Address what permits will be needed for storage, transport, and disposal of wastes.

ELEMENT 8. Safety Plan: Discuss how emergency situations will be evaluated and handled. Determine emergency contact personnel and what emergency services will be available in the field. Explain what procedures will be used to initiate search and rescue operations. List the training or other preventive measures required to conduct field operations safely. Indicate how this field safety plan will be developed in conjunction with laboratory, processing, and materials handling safety plans.

ELEMENT 9. Procurement and Inventory Control: Identify equipment, supply, inventory control and resupply, and services requirements of the field program and the processes by which they will be acquired and maintained. Determine where back-up equipment will be stored and how sites will be resupplied. Consider shipping regulations, especially for chemical and biological materials. Determine what analytical or other services will be needed and the best mechanisms for acquiring them. A procurement schedule should be provided for all items.

ELEMENT 10. Training Program: Describe who will prepare, review, and revise the field training and operations manual and the procedures for field measurements, sampling, sample handling, shipment, data recording, quality control, safety, waste disposal, and communications. Outline a schedule for the completion of these items. Describe training needs and identify who will conduct and review training. Address how personnel will be evaluated to ensure competency.

ELEMENT 11. Field Operations: Indicate the organizations that will perform each of the daily field activities. Describe how and when the daily field activities will be performed. Discuss and schedule the major events within field operations (i.e., mobilization, demobilization, and phase changes in sampling activities). Consider contingencies such as back-up personnel in the event of sickness. Require real-time evaluation to identify and resolve problems.

ELEMENT 12. Laboratory Operations: Indicate what organizations will be responsible for each type of sample preparation or analysis and for formulating each laboratory operations manual. If EPA conducts the activities directly, provide a development plan for providing appropriate laboratory facilities.

ELEMENT 13. Information Management: Describe any data management activities that might be affected directly by field operations. Establish guidelines for the timely and responsive transferral of information from field personnel to data managers. Indicate the groups that will be responsible for preparing and reviewing field data forms; provide a schedule for the completion of these forms. Develop a schedule for completion of the information management plan by the information management group.

ELEMENT 14. Quality Assurance: Describe who will provide input to the QA plan on field sampling, sample handling and preparation, sample shipment, sample disposition, and data management. A schedule for completing the QA plan should be provided to the logistics team and included in the logistics plan. QA activities should be coordinated with other resource groups using similar methods. This effort should identify common methods and standards when possible.

ELEMENT 15. Logistics Review and Recommendations: For each year of study within each resource group, summarize logistics activities. Discuss how personnel will be debriefed to identify and resolve problems. Discuss pilot

studies and associated methods evaluation experiments; present logistics data summaries within the full-scale project.

Field activities will start in 1991 with a lake pilot program in the Northeast. Additional regions will be phased into the program in each of the following years. We plan to begin a stream pilot in the West in 1992. The program will be fully implemented across the nation by 1995 for lakes and by 1996 for streams. A proposed regional implementation schedule for EMAP-Surface Waters is given in Section 11.

7.2 LOGISTICS ISSUES

The complexity of this program poses a number of logistics issues; overlooking or ignoring apparently minor issues or details will eventually jeopardize the success of the program. These issues will be addressed fully in each of the regional logistics plans before the implementation of field activities. This can be accomplished only through long-range planning and coordination. A brief discussion of the major issues (staffing, access, and data confidentiality) is provided in the following sections.

7.2.1 Staffing

Due to the kind of field data needed for indicator evaluation (Section 4.0), field personnel will require a high degree of expertise. They must have knowledge of fish and macroinvertebrate taxonomy, field sampling methods, and sample handling. Finding personnel with expertise in these areas will be a major challenge.

Field personnel will be required to undergo an intensive training program. Training assures that protocols are understood and strictly adhered to, and provides consistency across the program. Training is expensive, and these costs can be reduced in the long-term only by retaining staff. Therefore, retaining key personnel during nonfield seasons is critical to program continuity and cost effectiveness. We may be able to achieve this by employing field personnel for other EMAP tasks (e.g., assessment, integration, reporting), but it will place additional requirements on recruitment of personnel.

Various state resource agencies, EPA regional offices, other federal agencies, and universities have large pools of experienced personnel. Long-term agreements with these agencies and institutions to provide key personnel during the field season may also be a solution to this issue. To accomplish this, EMAP will have to demonstrate its utility to the States and other agencies by providing additional data and information addressing their problems. A concerted effort to inform these agencies of the goals and objectives of EMAP, and getting these agencies involved in the early planning phases of the programs are initial steps to be taken.

7.2.2 Access

Obtaining access information and permission to visit sampling sites is a difficult and time-consuming task. If land is publicly owned, approval must be obtained from the appropriate authority and permission may be conditional (e.g., upon the use of nonmotorized transport). Contingency plans for these conditions will have to be developed. If land is owned privately, each landowner will have to be contacted, and written access permission will have to be obtained. Collection permits will have to be obtained from the states regardless of land ownership. These state collection permits are usually very specific relative to site location, numbers and kinds of specimens, and identification of field personnel. As a result, these permits are difficult and time-consuming to obtain.

Gaining access permission and knowledge of access routes will require reconnaissance. The amount of time devoted to sampling a lake or stream site is dependent upon the physical access conditions. Some sites may be accessible only by foot. Each lake or stream identified as having difficult access should be visited prior to sampling to determine how sampling crews and gear can be transported to the sampling site and how samples can be transported adequately.

Although it sounds like an overwhelming task, access permission and all that it implies can be accomplished. The logistics team for EMAP-Surface Waters was responsible for carrying out the National Surface Water Survey (NSWS) of the National Acid Precipitation Assessment Program (NAPAP) and is an experienced and accomplished team.

7.2.3 Data Confidentiality

Data confidentiality is an issue of particular concern to EMAP. Many landowners may be reluctant to allow access to lakes or streams from their property because they fear regulatory and enforcement actions. Access is not a design constraint, and any denials by landowners could affect population estimates (see Section 3.0). Cooperating agencies within the Department of Agriculture often conduct field programs under an agreement of confidentiality with landowners. EMAP data may have to be aggregated in such a way that individuals cannot be identified to assure landowners and cooperating agencies that site-specific data will not be used against their interests. Agreeing to withhold certain information, however, is in direct conflict with the Freedom of Information Act and EPA's policy on data confidentiality. This issue will have to be resolved soon. The EPA Office of General Counsel is currently being consulted in this matter.

7.3 FIELD OPERATION SCENARIO: LAKES

We present the following field operation scenario to demonstrate that the proposed field activities are logistically feasible within the allotted time frame. This scenario is only one of many that could be developed at this time. It is strictly hypothetical and does not necessarily include all proposed indicator parameters or the order in which activities would take place when the program is actually implemented. Indicators are being evaluated and developed for pilot programs; the actual protocols will be solidified later, as the program evolves.

7.3.1 General Logistics Scenario

The scenario described here and the associated support activities are based on experience gained during the NSWS lake chemistry surveys (Morris et al., 1986; Bonoff and Groeger, 1987) and the Upper Peninsula of Michigan lake acidity fish survey (Baker et al., 1990). The scenario is based on the parameters listed in Table 7-2 and on the following 11 assumptions.

1. The index period will be July and August: a sampling window of about nine weeks.

2. Approximately 800 Tier II sites will be sampled per year, and it will take 4 years to sample all Tier II sites.

3. Site selection is completely random and does not consider site access.

4. The majority of lakes sampled will be smaller than 20 ha (see Section 3.0) and will be accessible by road. Assumptions concerning accessibility are based on the Eastern Lakes Survey (ELS).

5. Distance between sites will be one-fourth the density of the Tier II sites, or about 150 miles.

6. Small, motorized boats will be the primary sampling platform.

7. Fishing gear will be set out overnight; two days will be required at each site.

8. Four-wheel drive vehicles will be used for site access and each sampling team will need a second vehicle for logistics support.

9. Field mobile laboratories will not be used and there will be a minimum of sample preparation in the field.

10. Samples requiring immediate laboratory analysis will be shipped to the appropriate laboratory by overnight courier the day after collection.

11. A field crew will consist of five people: four for field sampling activities and one for logistics support.

Table 7-2. Proposed General Field Measurement Variables for Lakes

• Fish assemblage and gross pathology
• Macroinvertebrate assemblage
• Diatom assemblage from sediment cores
• Macrophyte cover
• Habitat quality
• In situ water quality
• Chlorophyll-*a*/Suspended solids
• Nutrients
• Sediment Chemistry/Toxicity

Based on these assumptions, a field crew of four people will be required to sample a lake over a two-day period. Transit time between sites will require most of one day; with one day off, a field crew could sample two lakes per week. Larger lakes (> 20 ha) and lakes with difficult access will require additional time and/or staff. A total of 40 to 50 field crews will be required to sample all 800 lakes within the index period. Allowance for downtime due to weather and other factors will have to be considered in determining the actual number of field crews. Each field crew will be responsible for sampling 15 to 20 sites in an area approximately the size of Oregon. To organize and coordinate the activities of 40 to 50 field crews, 5 to 10 regional logistics centers will be developed across the nation. These regional logistics centers (see Section 7.4) could possibly be integrated to support the field activities of the other EMAP resource groups (Wetlands, Near Coastal, Forests, Agroecosystems, and Arid Lands).

7.3.2 Daily Activities Scenario

The five-member field crew will be housed at a motel within 50 miles of the sampling site. Additional space will be obtained at the motel or some other location to store equipment and supplies and to calibrate instruments. Crews will begin each morning's activities by calibrating the instruments and checking to see that all necessary equipment and supplies are loaded into the vehicle. The four-member sampling crew will then depart for the lake. The fifth crew member (base site crew member) will remain at the motel base site; this individual will provide logistics support (picking up equipment and supplies, shipping and tracking samples, transmitting data, contacting landowners, etc.) and will provide a communications link with program management. These base site activities could not be completed by one of the sampling crew members because of the demanding schedule, as discussed below. The base site crew member, on a limited basis, could also be a backup sampler if someone were sick or if additional personnel were needed to sample the lake. Figure 7-1 and the following discussion summarize daily activities for the sampling crew.

On day one, the field crew will first verify that they are at the appropriate lake. Lake verification will be based on landscape features and topographic maps, and by the Global Positioning System (GPS) once this system is fully implemented. After verifying the identification of the lake, a calibration check will be performed on all instruments (e.g., the in situ water quality analyzer) to be used that day. The sampling crew will then split into a two-member boat crew and a two-member shore crew. The fishing gear (nets and/or traps) need to be deployed as early as possible, and the shore crew will start to prepare the gear and other equipment immediately. At this time, the boat crew will conduct a habitat evaluation, which will include sonar transects for depth and structure and in situ water quality profiles that will determine thermal stratification and anoxic condition. This information will be used to determine where to deploy the fishing gear. The boat crew will then pick up the fishing gear from the shore crew and deploy it. This activity will be followed by the collection of profundal benthic invertebrates (dredge sample) and sediment diatoms (core sample) at multiple locations. During this time, the shore crew will conduct a shoreline habitat evaluation and collect and sort littoral benthic invertebrates. After collecting the profundal benthic samples, the boat crew will join the shore crew and help process and preserve samples. During all of these operations, data will be entered into field data loggers. These systems will have programmed QA checks for parameter ranges and completeness. Backup hardcopy forms and notes will be filled out to assure data are not lost in the event of data logger failure. The field crew will check the data loggers, forms, and sample labels to assure completeness before returning to the motel base site.

At the motel, the base site crew member will debrief the sampling crews and check the data loggers, forms, and sample labels, and the condition of the samples. The preserved benthic invertebrate and sediment diatom samples will then be prepared for shipment. During this period, the sampling crew will prepare equipment and supplies for the next day. The base site crew member will then download the data from the field data loggers onto a portable computer.

On day two, instruments will be calibrated at the motel and a calibration check will be performed at the lake as described for day one. On the lake, the boat crew will first take various in situ measurements (temperature, dissolved oxygen, pH, conductivity, Secchi depth, etc.) and collect water and sediment samples at the deepest point in the lake. These samples will be transferred to the shore crew to be processed and preserved. After

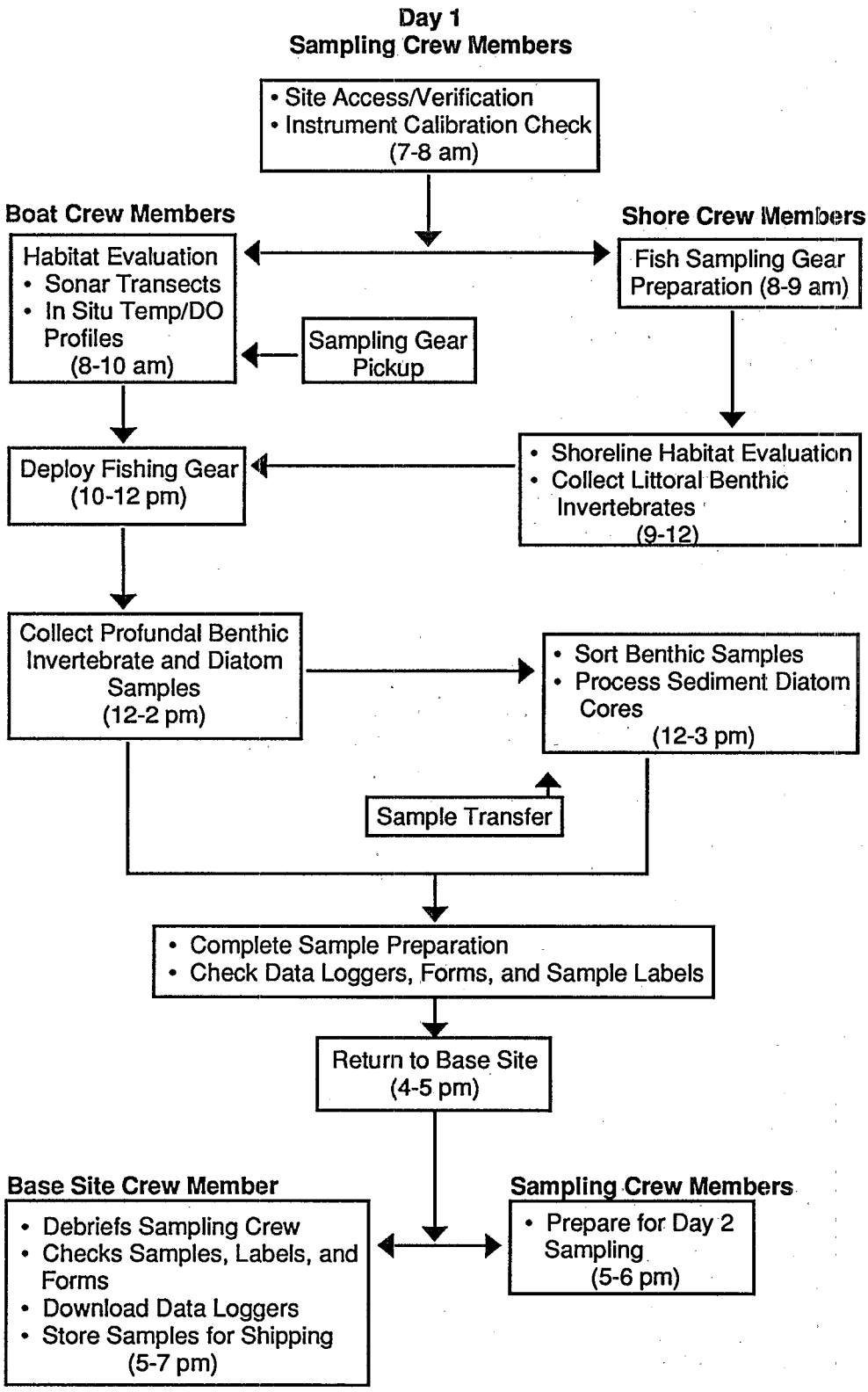


Figure 7-1. Flowchart of potential EMAP-Surface Waters daily field activity (page 1 of 2).

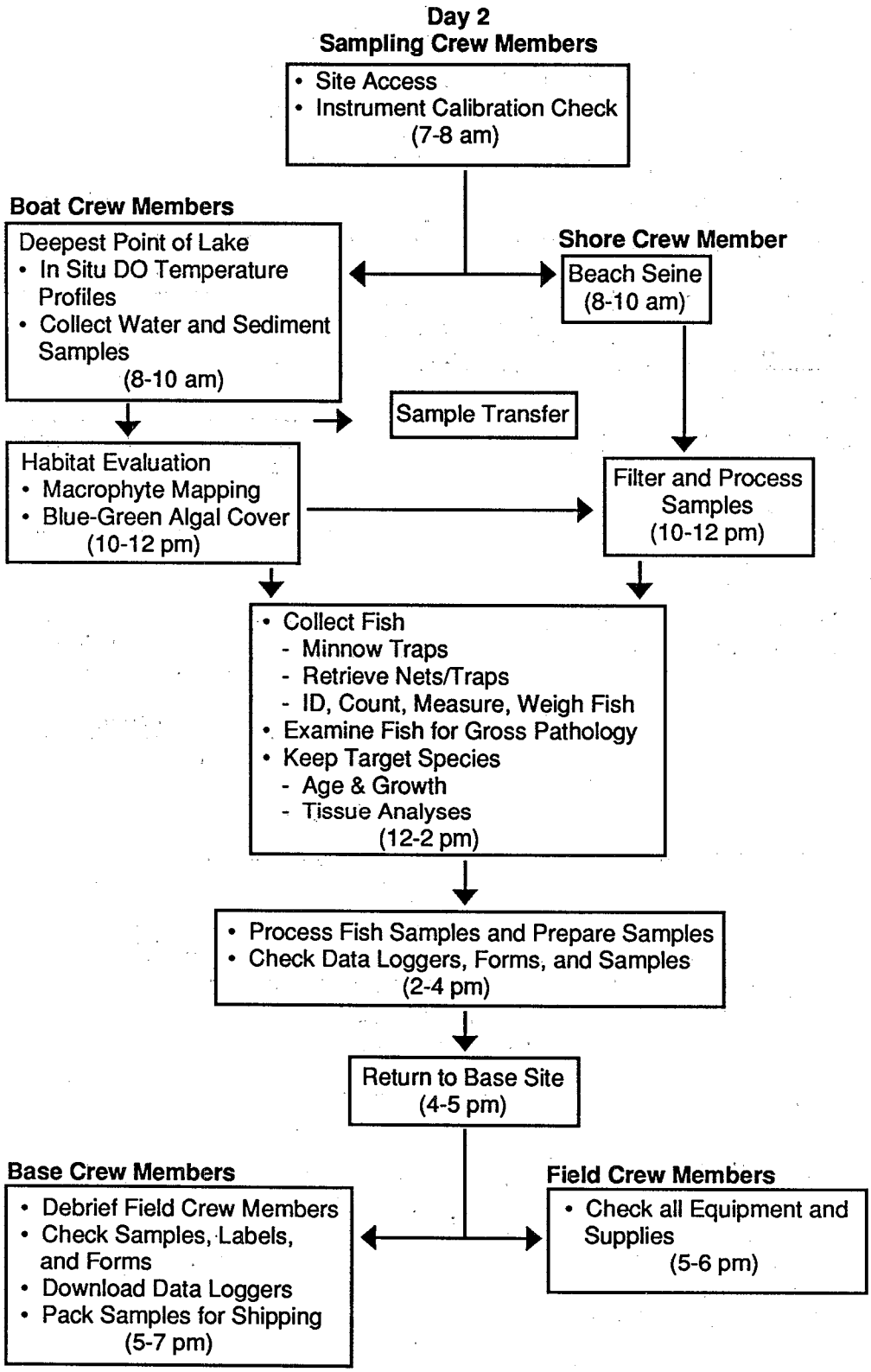


Figure 7-1. FLowchart of potential EMAP-Surface Waters daily field activity (page 2 of 2).

completing these activities, the boat and shore crews will rejoin to retrieve the fish nets and traps set out the previous day and to seine along the appropriate beaches. Members of the sampling crew will also snorkel along the shore of the lake to determine if fish are present and to identify additional fish species that may not have been caught by the sampling gear. Collected fish will be examined for gross pathology. Target specimens will be held and processed on shore for age, growth, and tissue analyses. Before leaving the lake, the field crew will check the data loggers, forms, and sample labels for completeness as described for day one.

At the motel, the base site crew member will debrief the sampling crews and again check the data loggers, forms, and labels, and the condition of the samples. The base site crew member will then download the field data loggers and pack the samples for shipping, as described for day one.

On day three, the field crew will move to a new motel base site close to the next site to be sampled. The base site member will transmit data electronically via a phone modem to the task group information center. After completing this activity, the base site member will take the samples to the overnight courier and pick up additional supplies and equipment. The specific location of the courier will have been predetermined during reconnaissance and verified through daily communications with program management. At the new base site, the sampling crew will go through a complete maintenance check of all the equipment.

7.4 FIELD OPERATIONS SCENARIO: STREAMS

Planning efforts for streams have not been conducted at the same level as lakes because of a later implementation schedule; therefore, we do not present a detailed field operations scenario here. A general scenario is presented, based on experience from the NSWS stream surveys (Knapp et al., 1987; Hagley et al., 1989) and on the following six assumptions.

1. Most of the parameters given in Table 7-2, except Chlorophyll-a, will be measured in streams.
2. Assumptions 1, 2, 3, 4, 7, 8, and 10 presented in Section 7.3.1 are valid for streams.
3. The majority of sites will not be immediately adjacent to a road but will be within a few hundred meters of a road. Assumptions concerning accessibility are based on the NSWS National Stream Survey (NSS).
4. The majority of sites will be first, second, or possibly third order streams. A different sampling frame will be used for fourth and fifth order streams.
5. Only one day will be required to sample a site.
6. A field crew will consist of four to five people.

Based on these assumptions, a field crew will sample three sites a week, with a transit time of one day between sites. A total of 25 to 30 field crews will be required to sample all 800 streams. Moving base sites will be used, as described for lakes. Again, 5 to 10 regional logistics centers will have to be developed across the nation to organize and coordinate the activities of the field crews. These centers will be identical to those developed for lakes and possibly to those of the other EMAP resource groups.

7.5 ORGANIZATIONAL STRUCTURE

Coordinating the logistics activities (staffing, training, deployment, tracking, etc.) of 50 or more Surface Water teams across the nation will be very difficult. As mentioned in sections 7.3 and 7.4, regionalizing these logistics activities into various centers will be the most effective mechanism for conducting EMAP field operations. The EPA regional offices can provide the initial structure for these logistics centers. The EPA regions have firsthand knowledge of the environmental conditions within their respective regions and will have a major role within EMAP, part of which could be in logistics. The EPA regions represent the Agency's primary contact with the states, and are working with the states at the program level. Securing cooperation from the states for EMAP is essential because of requirements regarding collection permits and access permission. The EPA regions can provide the crucial link in obtaining state cooperation for EMAP. The EPA regions

and states also have highly experienced field personnel, and their participation in EMAP field operations would be extremely beneficial. As key personnel directing field team activities year after year, they will provide the program with the continuity critical to EMAP field activities.

Surface Waters is only one of seven EMAP resource groups that will be conducting annual field monitoring surveys. Other resource groups are Wetlands, Great Lakes, Near Coastal, Agroecosystem, Forest, and Arid Lands. Although the various resource groups will not necessarily be sampling at the same location or at the same time, logistics efforts should be integrated as much as possible to reduce costs. Shared regional logistics centers with permanent warehouse facilities and staff will aid in this integration.

The long-term success of EMAP is dependent on the development of an interagency program with common goals for the monitoring of the ecological condition of the environment. Surface water monitoring alone could involve numerous agencies within the Department of the Interior (e.g., U.S. Geological Survey, U.S. Fish and Wildlife Service, National Park Service, U.S. Bureau of Reclamation), the Department of Agriculture (e.g., U.S. Forest Service), and the Department of Defense (e.g., U.S. Army Corps of Engineers). As EMAP evolves into an interagency program, agreements between agencies will have to be established to define responsibilities. As with the EPA regions and states, these agencies have highly experienced field personnel, and it is anticipated that personnel from these agencies will participate in both field activities and the regional logistic centers.

The Boise Interagency Fire Center, which may serve as a model for future EMAP regional logistic centers, is the national logistical support center responsible for coordinating and dispatching the closest suitable personnel, equipment, and aircraft for wild fires that exceed the capabilities of local and regional resources of land management agencies. This center is the result of an interagency program, with agreements between the U.S. Bureau of Land Management, the U.S. Bureau of Indian Affairs, the U.S. Forest Service, the National Park Service, the National Weather Service, and the U.S. Fish and Wildlife Service. The U.S. Bureau of Land Management manages the land, facilities, and changes as a key facet of the center and is host to the other five agencies. The objectives of the Boise Interagency Fire Center are:

- Interagency programs and services are developed through coordination and cooperation.
- Effective use of interagency programs and services by cooperating agencies.
- Equitable cost sharing of interagency programs and services.

Similar arrangements need to be considered for interagency EMAP logistics centers.

8.0 QUALITY ASSURANCE PROGRAM

A comprehensive quality assurance (QA) program will be implemented to ensure that all required data collected within EMAP can be used to achieve the objectives of the program. Quality assurance programs are mandated by the EPA for all environmental data acquisition activities it sponsors or in which it participates (Stanley and Verner 1985). The EMAP Quality Assurance Program Plan (Einhaus et al. 1990) documents the overall policies, organization objectives, and functional responsibilities designed to achieve data quality goals for EMAP activities. The QA program for EMAP-Surface Waters will be designed to ensure that the type, amount, and quality of data collected will be in accordance with the data quality objectives (Section 2.0) established for the program. The design of the overall QA program for the Surface Waters component must consider two major constraints: a broad range in scales of implementation (spatial and temporal) of data collection activities, and the participation of a potentially large number of individuals and organizations in data collection activities.

The major focus of data collection (and thus QA) will be the regional-scale sampling efforts conducted annually to monitor and assess status and trends in ecological condition. In addition, special studies on smaller areas (regional, subregional, or local) will be conducted to address specific questions (e.g., acidification or eutrophication), to investigate possible cause and effect relationships between ecological components (or processes) and changes in ecosystem condition, or to develop and evaluate appropriate ecological indicators. Each of these levels of implementation will require QA programs that differ in focus and intensity. The QA programs should be consistent with each other, however, and with EMAP's overall approach to QA. Such consistency will increase the level of confidence with which data user compare data from different though related programs within EMAP-Surface Waters and use data from different EMAP components.

We expect the participation of federal and state agencies, contractors, private consultants, production-mode analytical laboratories, and scientists from universities or other research institutions. Over the long lifetime of the program, the list of participants will undoubtedly change. Groups will have different levels of expertise in the principles and practices of QA. Existing monitoring programs considered for integration into the EMAP framework may have QA requirements that are initially incompatible with those established for the Surface Waters component, or EMAP as a whole. Differences in sampling and analytical methodology, whether among participating groups, among regions, or as a result of new technologies over the life of EMAP, must be monitored and assessed in order to quantify and minimize their impact on the interpretation of the observed status and trends in ecological condition.

The overall QA program implemented for EMAP-Surface Waters will be based on a philosophy of guidance, assistance, and commitment to improvement. It will not be based on the "enforcement" approach required for regulatory or compliance monitoring programs. In developing information of known quality, consistent with data quality objectives, emphasis will be on consistency in implementation, quality control (QC), prompt corrective action, and continuous improvement. This approach will identify and correct problems as soon as possible, to minimize their impact on data quality. The QA staff within the Surface Waters resource group will receive training in carrying out their responsibilities. Such training should include, but is not limited to, improving quality and productivity, preparing and reviewing QA-related plans and reports, and conducting and reporting audit activities. The QA staff will provide guidance, training, technical support, and tools (e.g., performance audit materials, QA documentation) to all participants to implement QA programs that are consistent with the data quality requirements of the Surface Waters component, and EMAP as a whole. Although the emphasis of the QA program is to provide guidance and support, the management and QA staff must also have the means to deal with instances of poor performance, when necessary, to avoid compromising data quality.

The following subsections outline the general approaches, conceptual rationale, and guidelines proposed for designing and implementing the overall QA program for EMAP-Surface Waters.

Section 8.1 presents aspects of data quality requirements for individual variables, metrics, and indicators. Section 8.2 outlines the organizational structure and staffing requirements of the QA program. Then Section 8.3 describes the documentation that will be required as part of the QA program. Section 8.4 presents the guidelines and proposed approaches for incorporating QC procedures into the data acquisition process. Section 8.5 explains the proposed procedures and guidance for the review, verification, and validation of data. Section 8.6 details the assessment of data quality. Finally, Section 8.7 summarizes reporting activities and products related to the QA program.

8.1 DATA QUALITY REQUIREMENTS

In all data collection activities, data quality requirements will be specified in five areas: precision, bias, comparability, completeness, and representativeness (Stanley and Verner 1985, Smith et al. 1988). In addition (when appropriate), minimum tolerable background levels of chemical constituents will be established. These levels represent the maximum concentrations of constituents that can be contributed by the sample collection and measurement process. Method detection limits will also be specified, monitored, and assessed (Section 8.1.6).

Ideally, data quality requirements will be based on the overall data quality objectives (Section 2). In some cases, the requirements will be qualitative, in others, quantitative. Data quality requirements will also be reviewed periodically throughout the program, and revised as necessary in response to improved capability, additional knowledge, or technological or resource limitations.

Data quality requirements, constrained by sampling, measurement, or logistical considerations, will determine the appropriate methodology. Criteria for initial selection of appropriate chemical and biological methodologies are presented in Table 8-1. These criteria were chosen primarily by means of the approach advocated by Hunt and Wilson (1986), with the addition of criteria relevant to the design of a data management system (Section 9).

8.1.1 Precision and Bias

Precision and bias are estimates of random and systematic error in a measurement process (Kirchner 1983, Hunt and Wilson 1986). Collectively, they provide an estimate of the total error or uncertainty associated with an individual measurement, or set of measurements. In theory, random and systematic errors can be determined at any point in a collection and measurement process. Figure 8-1 summarizes error components of interest for EMAP-Surface Waters (described in Section 4.0). Estimates of the various error components will be determined primarily by replicate sampling; such sampling can be modified to address and control major sources of variability. The statistical design and sampling plan should act to minimize systematic errors in all components except measurement error (σ^2_{meas}). Systematic errors in these components will be minimized by using documented methodologies and standardized procedures, and evaluated using samples of known composition that can be subjected to the entire collection and measurement process. Variance components of the collection and measurement process (e.g, among analytical laboratories or among individuals identifying biological specimens), although of less interest to data interpretation, should be estimated periodically so that QA efforts can be allocated to control major sources of error.

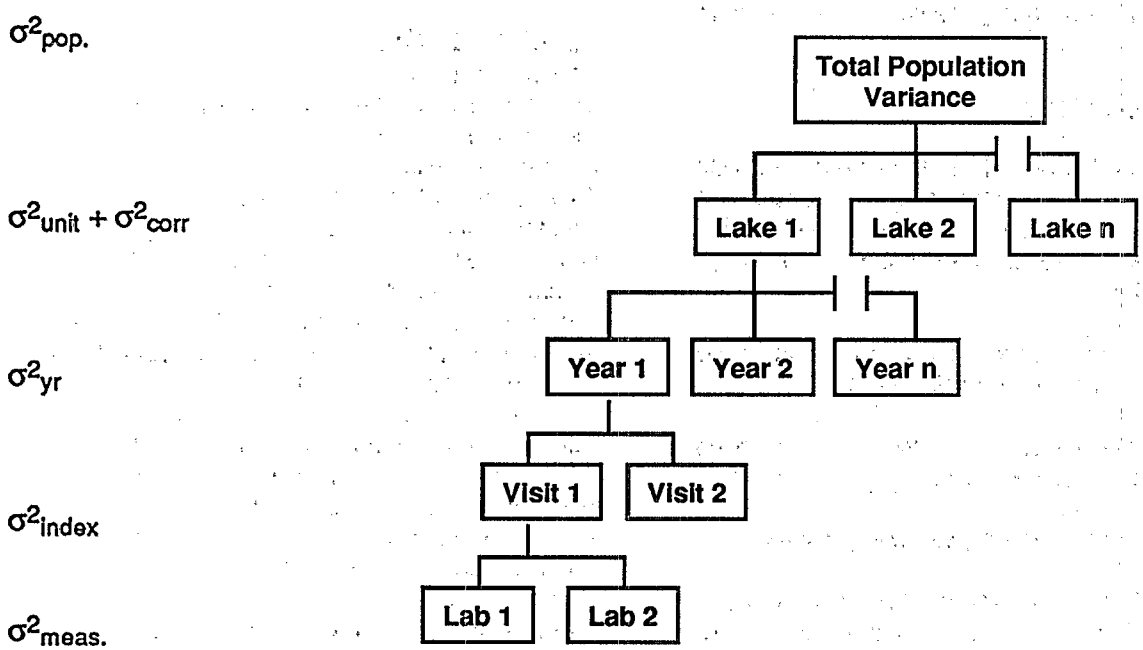
The precision and bias requirements will be used to define criteria for monitoring collection and measurement activities and for maintaining them in a state of statistical control (i.e., the distribution of individual measurements have a stable and predictable distribution over time; Taylor 1987). Estimates of precision and bias are also necessary to evaluate the other three data quality indicators (comparability, completeness, and representativeness).

In general, data from one or more measurements of variables will be combined (and possibly transformed or categorized) into metrics; one or more metrics will be incorporated into an indicator, and one or more indicators will be used to provide an estimate of the ecological health of a population of lakes or streams. An example of the relationship between measurements, metrics, and indicators, and endpoints appears in Figure 4-1. Random and systematic errors associated

Table 8-1. Criteria for Selection of Appropriate Sampling and Analytical (or Measurement) Methodology[a]

Criteria	Comments
Range of values of interest	May need to be tailored for different regions or different projects.
Lowest value of interest (values below this will probably have uncertainties > 100%); or, in case of trace-level analyses, the required limit of detection.	Chemical: Used to determine appropriate limit of detection. Biological: May help define the minimum effort required in obtaining data (e.g., at least 100 organisms in a benthic sample are needed for determining species composition and relative abundances).
Maximum tolerable measurement error (random and systematic)	Defined on basis of sample collection and measurement only). Additional error components of interest can be defined in terms of short-term or long-term (e.g., within-day or within-batch versus among-day, among-batch, or among-group). Should be expressed in both absolute and relative terms: Absolute: M = lowest value of interest Relative: M = some specified percentage of true or most probable value "Knot value" = Absolute/Relative; represents value at which absolute error equals relative error.
Standard reference on which method is based	Describe any required modifications.
Required frequency of sampling or collection	Requirements based on sampling plan; used to estimate variance components of interest.
Collection, analytical, or measurement constraints	Site selection criteria, special equipment requirements, use of hazardous reagents, etc.
Sampling handling considerations or measurement conditions	Appropriate holding times (e.g., APHA 1989) Operational: May be based on data reporting time requirement. Maximum: If greater than operational, point at which sample is no longer considered representative of conditions at time of collection. Containers/preservation techniques
Reporting requirements	Type of variable (numeric, coded, character, categorical, etc.). Reporting Units (mg/L, µeq/L, number of individuals, etc.). Number of significant figures desired, and maximum number of decimal places. In addition to being used to select method of appropriate sensitivity, information is needed to design data base, also for collection forms (hard copy or electronic).
Data reporting time	Time period between collection and incorporation of validated data into data base for use in analysis and reporting.

[a] Based on Hunt and Wilson (1986).



σ^2_{total} (Total Variance) =

- $\sigma^2_{pop.}$ = Sampling Variance of Population (Extrapolation of sample to population)
  - Function of sample size, sample variance
- σ^2_{unit} = Sample Variance
  - Due to differences in index values among sites within a region
- σ^2_{corr} = Variance due to Correlations Among Sites
  - Spatial or temporal patterns of sample sites within a region
- σ^2_{yr} = Variance in index value at a site among index periods (years)
- σ^2_{index} = Variance within an index period
  - Spatial (e.g., among potential sampling locations during a single visit)
  - Temporal (different visits within an index period)
- $\sigma^2_{meas.}$ = Variance in measurement within a single index sample
  - Sample collection, processing, and analysis
  - "Worst-case" estimate would include different teams, different laboratories, different dates of analysis

Figure 8-1. Hierarchical model of variance components of importance to EMAP-Surface Waters.

with these various "composite" values will have to be evaluated regarding potential impact on data quality.

8.1.2 Comparability

Comparability is defined as "the confidence with which one data set can be compared to another" (Stanley and Verner 1985; Smith et al. 1988). For EMAP-Surface Waters, there are three levels of comparability:

1. Short-term: within a single sampling period (EMAP's basic reporting unit).

2. Intermediate-term: over a complete sampling cycle (4-5 yr).

3. Long-term: among sampling cycles.

Data from these assessments are needed to compare and interpret surface water data only. Comparability can be maximized through the use of standardized methodologies and similar QA programs; it can be evaluated quantitatively through estimates of precision and bias, and through carefully designed method comparison studies.

We also need to be cognizant of data comparability external to EMAP-Surface Waters. EMAP-Surface Waters data should be comparable with:

- Data collected by other EMAP ecosystems groups (e.g., agroecosystems, forests, and wetlands).

- Other environmental data sets (e.g., deposition data from NADP or other networks), and data from existing monitoring programs being incorporated or integrated into EMAP-Surface Waters.

- Data that will be collected in the future, whether as part of EMAP itself, or other monitoring efforts that may develop in the future.

In these cases, comparability will need to be evaluated with respect to the available QA and QC data. The degree of comparability required will depend on the intended use of the data (trend detection, associative analyses, etc.).

8.1.3 Completeness

For EMAP-Surface Waters, the requirements for completeness will be based on the amount of data required to make conclusions pertinent to the program (or project-specific) objectives. For example, in order to make a conclusive statement (with an acceptable level of confidence about the condition of a subpopulation of lakes (or streams) with respect to biotic integrity, valid data for indicators related to biotic integrity from a minimum number of sampled lakes (or streams) will be required. This minimum number may differ for other indicators (such as trophic state). Completeness may also need to be evaluated at the level of individual samples or measurements. For example, if an expected taxonomic group is missing from a sample, any metrics or indicators that included this taxon would be suspect, and the sampling site could not be utilized in estimating ecological condition with respect to biological integrity. In the case of chemical samples, valid data may be available for 99% of the samples collected in a region, but if the missing or invalid 1% of the data all come from a group of sampling sites of interest (e.g., seepage lakes), assessments of condition may be compromised.

8.1.4 Representativeness

Representativeness is defined as "the degree to which the data accurately and precisely represent a characteristic of a population parameter, a variation of a property, a process characteristic, or an operational condition" (Stanley and Verner 1985; Smith et al. 1988). Representativeness can be affected by problems in any or all of the other indicators of data quality, as well as by issues such as the location of a sampling site, the time of sampling, and the statistical selection of sampling sites. More specific to a QA program is the representativeness of samples or procedures used to control and assess data quality as compared to the range of conditions being sampled.

8.1.5 Tolerable Background Levels

Background is operationally defined as the amount of contamination due to collection, handling, processing, and measurement. It is most relevant to those chemical constituents present in the environment in very low concentrations. Requirements for tolerable background limits will be based on the lowest concentration of interest that is required to assure that repre-

sentativeness or completeness requirements are not compromised. Careful adherence to sample collection, handling, and processing protocols will minimize background levels. Blank samples of various types will be used to provide estimates of background levels.

8.1.6 Method Detection Limits for Chemical Constituents

For chemical constituents, the method detection limit (MDL) will be defined, monitored, evaluated, and reported. The MDL is important in evaluating data quality, providing the data user with an indication of the lowest concentrations that were determined with confidence. The MDL also assists in determining the capabilities of analytical laboratories with respect to the analytical requirements for EMAP-Surface Waters.

For EMAP-Surface Waters, the MDL is defined as the lowest concentration of a constituent that an analytical process can distinguish from zero concentration, with a single measurement, with 99% confidence. Based on the MDL for each constituent, acceptance criteria will be established. Single measurements of a low-concentration standard will be compared against the acceptance criteria as an indication the MDL is being achieved for a particular batch of samples. Repeated measurements of the low-concentration standard will be used to evaluate the achieved MDL (including among-batch variability) against the required MDL value. The achieved MDL will be estimated as a one-sided 99% confidence interval for the set of low-concentration standard measurements. The achieved MDL values will be reported with all data quality summaries associated with EMAP-Surface Waters.

8.2 ORGANIZATION AND STAFFING REQUIREMENTS

The EMAP QA coordinator, located at EMSL-Cincinnati, has overall responsibility for implementing consistent and adequate QA programs within EMAP as a whole. The QA officer is responsible for designing and implementing the QA program for the Surface Waters component. Table 8-2 summarizes the responsibilities of the QA coordinator and the Surface Waters QA officer. The QA officer will be assisted by one or more coordinators in implementing the large-scale annual sampling operations. Special projects of regional scale will have a designated QA representative to oversee the QA program for the project. We anticipate that each localized, intensive monitoring program will have a designated on-site QA representative, who will be supported by the QA officer and regional QA representatives.

8.3 QUALITY ASSURANCE DOCUMENTATION

Before implementation of field sampling operations, a number of different documents will be prepared (or existing documents utilized) as part of the QA program (Table 8-3).

Primary guidance for implementing the QA program will be provided by the EMAP Quality Assurance Program Plan (Einhaus et al. 1990). The policies, organization, objectives, and functional activities that pertain specifically to the QA program for the Surface Waters component will be detailed in a QA project plan. The Surface Waters QA project plan will be used as guidance in preparing QA project plans for special studies, regional or local in focus. In general, the QA project plan for any specific data collection activity will detail the QC and data quality assessment activities (summarized in the following subsections) that will be used to ensure achievement of required data quality. Additional QA documentation appropriate for specific data collection activities may include guidance documents or QA program plans for other federal or state agencies and facilities, provided they meet or exceed the requirements set forth in the EMAP QA program plan and the Surface Waters QA project plan.

Documentation pertaining to the EMAP-Surface Waters QA program will be reviewed periodically, and revised as necessary to reflect changes based on previous performance, or other modifications to either the QA program or to EMAP in general. Changes in various aspects of the QA program should also be incorporated into revision of standard operating procedures related to sample collection and measurement.

The general QA project plan for EMAP-Surface Waters activities will be reviewed by representatives of the major participants (e.g., EPA laboratories) before submission for formal approval. The existing QA program plan for EMAP (Einhaus et al. 1990) specifies signature approval by the Surface Waters QA officer, the Surface Waters technical director, the associate

Table 8-2. Responsibilities of the EMAP QA Coordinator and Surface Waters QA Officer

QA Coordinator:

- Prepare and revise EMAP QA program plan, and assist in implementation of the overall QA program.
- Provide guidance to EMAP resource groups in developing data quality objectives (DQOs) and auditing programs, and for preparing QA project plans and standard operating procedures.
- Review and approve EMAP QA documents and reports, including QA project plans, DQO documents, and resource group QA annual reports and workplans.
- Review QA programs implemented by individual resource groups.
- Prepare an overall QA annual report and workplan for EMAP.
- Coordinate EMAP QA-related activities with EPA Agency QA personnel.

Surface Waters QA Officer:

- Prepare the QA project plan and QA annual report and workplan for the Surface Waters resource group.
- Coordinate QA-related activities within the Surface Waters resource group.
- Facilitate the development of DQOs and methods selection for the Surface Waters resource group.
- Prepare or coordinate the preparation of data quality assessment reports for inclusion into EMAP Surface Water reports and data bases.
- Conduct audits of data acquisition activities within the Surface Waters resource group.
- Assisting in the preparation of QA project plans for special research projects conducted or sponsored by the Surface Waters resource group.

director for Inland Aquatic Systems, and the EMAP QA coordinator.

8.4 QUALITY CONTROL GUIDELINES

Quality control is applicable to all stages of a data acquisition process, from design through sampling and analysis, data management, and interpretative reporting. For the Surface Waters component, this process will be similar to that illustrated in Figure 8-2. Each stage in the process represents a point at which quality control measures can be implemented (if necessary or desirable to monitor those aspects that are most subject to error or inconsistency).

Those stages conducted after the commencement of field operations also represent points where assessments of data quality (Section 8.7) can be made. In some cases, such assessments are necessary for monitoring sources of error to optimally allocate control measures among points in the process where they are most needed.

General activities that will maximize the success of a QC program include:

- Documentation of procedures related to design, sampling, measurement, information management, data analysis, reporting, and QA (Section 8.3).

Table 8-3. Quality Assurance Related Documentation of EMAP-Surface Waters

- EMAP Quality Assurance Program Plan

  Describes philosophy and QA policies of EMAP and provides guidance for designing and implementing QA programs within EMAP

- Surface Waters Quality Assurance Project Plans

  Detail the QC and assessment activities that will be used in the QA program for Surface Waters

- Field Operations Manuals

  Standard operating procedures for sample collection, handling, and processing, collection of field data, and data management activities (including QA and QC procedures). Also describe other logistical procedures (e.g., sample shipping, waste disposal, communications, safety) conducted in the field.

- Analytical Methods Manuals

  Standard operating procedures for sample analysis (including QA/QC procedures).

- Quality Assurance Project Plans from EMAP Support Groups

  - Landscape characterization
  - Information management

- Other QA program plans and appropriate QA project plans for other participating groups (e.g., agencies, laboratories, principal investigators).

- Standardized training programs to ensure a minimal level of competency in all aspects of the project.

- A program of calibration and preventative maintenance for all sampling and analytical equipment and instrumentation.

- Periodic site visits by knowledgeable members of the QA or management staffs to ensure that sampling and measurement activities are being conducted appropriately, to recommend corrective actions as necessary, and to assist on-site personnel with addressing QA-related issues.

Where appropriate, collection and measurement processes will be monitored through the use of frequent QC checks using samples of known composition, or through replicate measurements. Control charts will be maintained whenever possible. Use of these tools allows for rapid identification and resolution of problems related to sample collection or measurement, and provides documentation that the process is being maintained in a state of statistical control. Specific examples of QC activities are provided in the following subsections.

It is important to recognize that the utility of QC measurements will be constrained, especially in the field, by the relatively brief index period each year (about 2.5 months) and by the turnaround time between collection of samples and subsequent analysis (especially for complex analyses such as organic compounds or fish tissue analyses). Resampling a site because of subsequent identification of a problem will not be feasible in most cases (although sample reanalysis will be, depending on the amount of sample remaining and any holding time considerations). It thus becomes critical to minimize the potential of

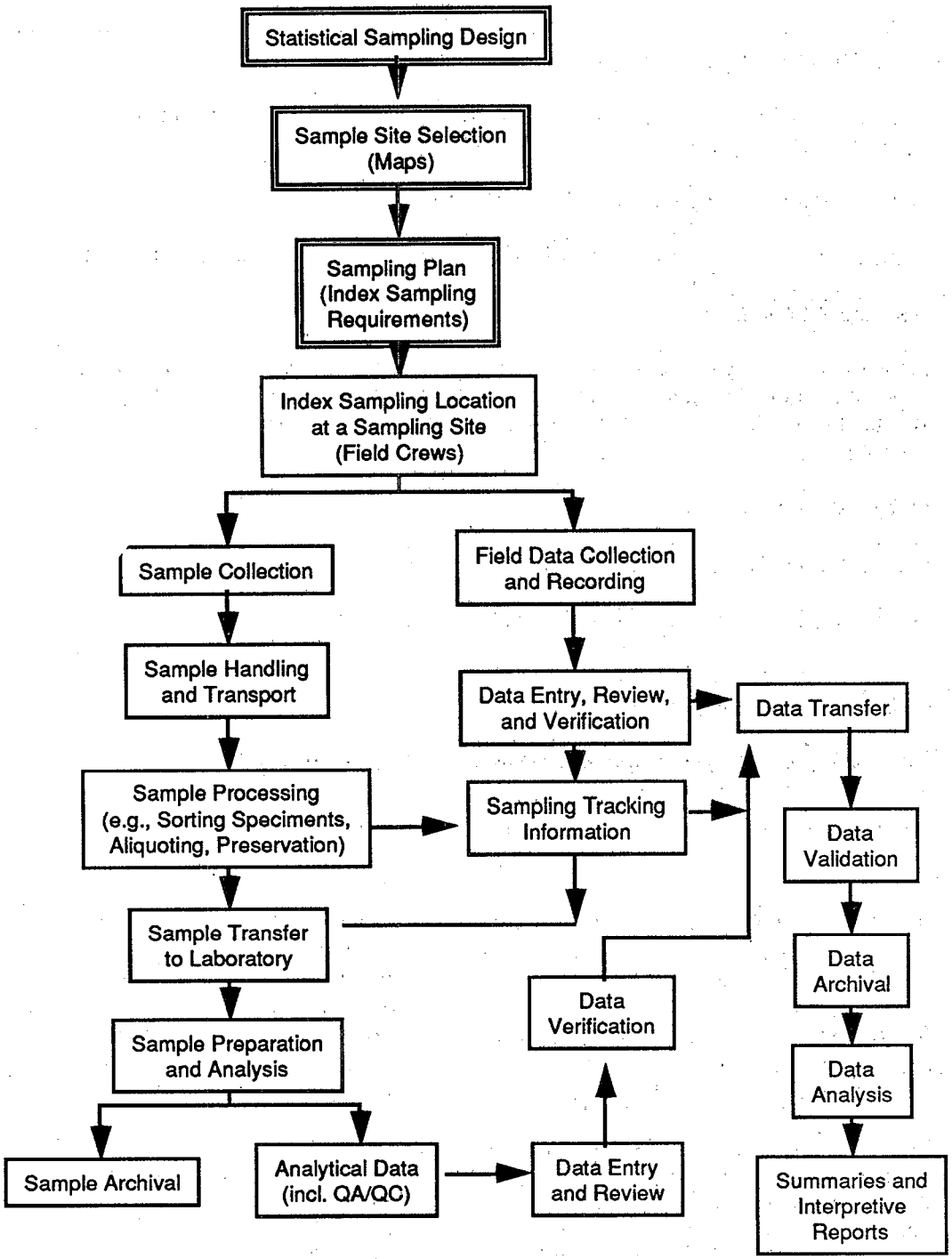


Figure 8-2. Stages in the data acquisition and management process for EMAP-Surface Waters where data quality can be controlled or assessed. Double-lined boxes represent stages conducted before implementation of field sampling operations.

sampling-related errors through careful attention to standardized procedures.

8.4.1 Biological Measurements

Quality control activities applicable to the collection, identification, and associated attributes of fish, benthos, diatoms and other biological samples are summarized in Table 8-4. For collection of fish samples, the activities proposed are based on procedures previously utilized in lakes for the Biologically Relevant Chemistry Project (Cusimano et al. 1990, Baker et al. 1990) and other projects, and for streams, the Episodic Response Project (Peck et al. 1988) and the EPA Rapid Bioassessment protocols (Plafkin et al. 1989), and in guidelines presently being developed for use in Canadian monitoring programs (Hart 1990). For benthos, many of the activities utilized in streams (Peck et al. 1988, Plafkin et al. 1989) are identical (or easily modified) for lake benthic sampling. For sediment core sampling (for diatom assemblages) QC procedures have been developed and evaluated as part of the Paleoecological Investigation of Recent Lake Acidification (PIRLA) project (Charles and Smol in press). For toxicity tests, QC procedures have been developed as part of standardized test methods (e.g., EPA or ASTM); these can be modified as necessary to meet the needs of the EMAP-Surface Waters component.

When possible, some type of control criteria will be established to ensure an adequate sampling effort has been conducted at each site to collect a representative index sample. In cases where this is not feasible, some type of replicate sampling, repeated measurement, or additional effort sampling program will be conducted at a subset of sites to provide an estimate of sampling efficiency or precision. Such estimates can subsequently be used to develop control criteria as the program continues. Repeated sampling and measurement strategies will be designed in conjunction with the sample replication scenarios presented in Section 4.0. Repeated or independent checks on sample processing and taxonomic identifications will be conducted on a subset of samples collected. Reference collections of biological specimens will be developed and maintained by participating groups during the course of the program. Such collections will be eventually be archived at a permanent collection facility (e.g., a museum) for future use (Section 8.4.4).

Several issues pertaining to QC of biological measurements within EMAP-Surface Waters remain, including developing efficient techniques to identify, assess, and monitor systematic errors associated with sample collection, especially fishes in lakes. As presented in Section 5.4, different types of sampling gear are required in lakes to obtain a representative sample of the fish assemblage. Additional sampling programs to estimate and control biases in collection will be costly and complex to conduct for all the different types of gear proposed. We must also assure that sampling and associated QA activities do not result in a reduction in population sizes more severe than would result from any potential environmental stressor. Testing and reporting conditions for bioassay tests will need to be precisely specified, in terms of test organisms, holding conditions, reference conditions, and reporting procedures, to ensure consistency in performance among all participating laboratories. Changes in taxonomic relationships and nomenclature will need to be tracked as new studies are published that revise taxonomic relationships or nomenclature, or report on new species.

8.4.2 Chemical Measurements

Quality control activities for chemical measurements are well documented (e.g., Hunt and Wilson 1986, Taylor 1988). Table 8-5 summarizes these activities. Quality control procedures for dilute surface water samples have been used for large-scale studies during the EPA's Aquatic Effects Research Program (e.g., Silverstein et al. 1988 for lakes, Cougan et al. 1988 for streams), and refined for smaller scale operations for the Episodic Response Project (Peck et al. 1988). Specialized collection and handling procedures may be required for certain types of water samples (e.g., those being analyzed for organic constituents) to minimize contamination and prevent changes in composition between collection and analysis. Use of field blank samples will monitor possible contamination during collection and analysis.

In the laboratory, appropriate types of control samples, and control charts, will be used to monitor and evaluate statistical control of the analytical process. For inorganic analyses, at least one check standard, at a concentration near the middle of the calibration range, will be analyzed periodically with routine samples.

Table 8-4. Quality Control Activities Associated with Biological Measurements[a]

Fish	Benthos	Sediment Cores (Diatom assemblages)	Bioassays
Established criteria for sampling effort; number of individuals, number of taxa, or sampling effort (bias, comparability)	Proper preservation, labeling of samples in field (completeness)	Collect replicate cores (precision)	Establish minimum number of test organisms per trial (comparability)
Repeated sampling by different crew (precision, comparability)	Establish minimum number of individuals required (precision, bias)	Section cores at lake shore (bias, comparability)	Establish minimum number of replicate determinations required for a valid test (completeness, precision)
Proper preservation and labeling of samples in field (completeness)	Independent check on sorting, counting, and identification (precision, completeness)	Proper preservation, labeling of core samples (completeness)	Establish restart criteria based on mortality in control test, or evidence of container effect (completeness, comparability)
Independent checks on counts, weights, and identification; calibration checks of balances and measuring devices (completeness, precision, bias)	Use of standard taxonomic keys and nomenclature (bias, comparability)	Independent checks on counting and identification (precision, completeness)	Establish performance criteria for test results, e.g., minimum number of young required for valid reproductive endpoint (precision, bias, completeness, comparability)
Use of standard taxonomic keys and nomenclature (bias, comparability)	Collection of voucher specimens, curation of reference collections (bias, comparability)	Preparation of photographic plates of different taxa, maintenance of reference collection (bias, comparability)	Use of control charts to monitor replicate precision, response to reference toxicants (precision, bias)
Collection of voucher specimens, curation of reference collections (bias, comparability)		Use of standard taxonomic keys and nomenclature (bias, comparability)	
Established precision levels for length-weight regression equations (precision, bias)		Monitor performance of instrumentation used for dating core layers using blanks and internal standards (precision, bias)	
Replicate age determinations from scales or other structures (precision)			

[a] Attributes of data quality that are addressed for each activity are listed in parentheses.

Additional standards may be necessary to determine detection limits for analytes present in low concentrations. For organic analyses, internal standards may not be available; matrix spikes or duplicate analyses on a subset of routine samples will be required to monitor random and systematic errors.

When possible, standard reference materials (SRM) or certified reference materials (CRM) will be used periodically as nonblind samples to assist laboratories in maintaining statistical control. Such materials will be of most use for chemical analyses of sediment samples, for organic analyses, and for analyses of compounds in fish tissue. Such materials are currently being used

Table 8-5. Quality Control Activities Associated with Chemical Measurements

Field
• Calibration checks of instrumentation using independent standards of similar composition to sample; use of control charts to monitor performance (precision, bias, comparability).
• Preventative maintenance program for all equipment and instrumentation (completeness).
• Standardized procedures for collecting, handling, and processing samples; procedures for minimizing the potential of contamination during collection, handling, or processing (comparability).
• Proper preservation and labeling of samples (completeness).
• On-site review of recorded data and other information (comparability).
• Periodic use of field blank samples and audit material to check for effects of collection and processing (background levels, bias).
Laboratory
• Standardized procedures for preparation, calibration, and analysis (comparability).
• Preventative maintenance program for analytical instrumentation (completeness).
• Routine use of control samples (blanks, check standards, matrix spikes, etc.) and control charts to monitor statistical control of analytical process (precision, bias, method detection limit).
• Periodic use of reference material (standard reference materials, certified reference materials) or other sample of known composition as internal standards to check for systematic errors in analysis (precision, bias).
• Review of analytical data (and associated QC sample data) immediately after analysis and before entry into data base (completeness).

by the Near Coastal component of EMAP, and this program should provide information related to feasibility, cost, and preliminary performance data that can be used in the QA program for Surface Waters. It would be advantageous to subject such reference samples to the entire collection and measurement process, rather than just to the analytical phase. This would assist in monitoring potential errors (random and systematic) associated with sample collection and field processing. The feasibility of implementing such an approach as a QC tool will be investigated as part of the QA program for EMAP-Surface Waters.

8.4.3 Habitat Quality and Site Characterization Measurements

Some site characterization data will be collected by the EMAP Landscape Characterization resource group, using remote sensing techniques. The Landscape Characterization group will document QC activities associated with the landscape characterization measurements in a separate QA plan. For measurements collected directly as part of the EMAP-Surface Waters effort, the most critical QC activities, once standardized methods are implemented, will be the development and use of standardized codes and categories. For

measurements collected from maps, an independent check of the measurements, conducted periodically by a second person or group, would serve to detect and correct errors on a timely basis. For data collected during a site visit, proper calibration of instruments (e.g., calibration of an electronic depth finder against a calibrated sounding line), and repeated measurements by a second person on a subset of sites will be used as the primary means of minimizing errors. Such repeated measurements also will provide estimates of the magnitude of measurement errors. It will be desirable to implement methods to monitor for systematic errors in collecting these data (whether they result from a particular method or from different crews using the method).

8.4.4 Archival of Samples and Specimens

Archival activities will involve samples for chemical analyses, the curation of biological specimens (fish, invertebrates), slides of diatoms from sediment cores, and validated data bases. For chemical samples, samples of water and sediment will generally be archived during a particular index period, in case some type of reanalysis is warranted. Such samples (or a subset) may be preserved and archived for longer periods to permit future analyses of constituents other than those initially determined. An example might include more detailed analyses of samples when the results of bioassay experiments indicate possible toxicity. Such long-term storage may be feasible for some inorganic constituents, but may not be feasible for organic compounds. Samples of fish tissue may be placed in a permanent specimen bank, such as that established by the National Institute of Standards and Technology (NIST) for possible analyses in the future.

Voucher specimens of fish and invertebrates will be collected as part of the routine QC program (Table 8-4). Periodically, such specimens will be placed into a permanent collection. Possible options for curation include the establishment of a specimen banking and curation system specifically for EMAP, or to make arrangements with regional facilities (e.g., national museums, university museums, or state biological survey agencies) to incorporate specimens collected as part of EMAP into permanent collections. The location of archived specimens would be reported in the appropriate summary or interpretative reports.

The existing EMAP QA program plan instructs all participants in the program to retain raw data and associated documentation (e.g., notebooks, instrument output records) for seven years, with permission for disposal required from the program director. Validated data bases associated with EMAP-Surface Waters are anticipated to be archived permanently as part of the overall EMAP information management program.

8.5 DATA REVIEW, VERIFICATION, AND VALIDATION

This aspect of the QA program overlaps with the QA program that will be established for the information management program (Section 9). Quality assurance for Surface Waters information management can be considered operationally as two separate elements. One element, data review, verification, and validation, discussed here, involves ensuring the accuracy of all information that is ultimately entered into or accompanies a data base. The second component, which is also addressed in Section 9.0, involves maintaining the security and integrity of validated data bases once when have been archived and made available for use in data analysis, reporting, or distribution to users outside the Surface Waters component, or outside EMAP. Of primary concern here is the prevention of deletions, alterations, or irretrievability of information stored in data bases.

The general approach in minimizing data-related errors before archival will be to emphasize the review of information at the point of collection or measurement as soon as possible after the sample or measurement has been obtained. Where feasible, data recording will be done electronically, with standardized recording forms being used as backups. In the field, data logging devices (hand held computers that display screens similar to manual field forms) are being considered for use, and are currently being tested by other EMAP task groups. Use of data logging devices would reduce the time required for data entry. These devices can automatically check for erroneous data as they are entered (e.g., range checks on numeric data, misspellings, or invalid codes). Similar types of devices may also be utilized for laboratory measurements.

The review process will be automated to the extent possible, but not to the exclusion of a manual review by qualified and knowledgeable

persons. In order to complete the review, verification, and validation process as quickly as possible, a substantial investment in resources and personnel required for data entry and data review and verification will be required.

Data review will initially involve a check of raw data (e.g., what is recorded on data forms) before entry into an electronic data base. In the field, forms should be reviewed by a second person before leaving a site. Another review should occur immediately after data entry by comparing the entered data to the raw data. Data from QC samples or measurements, and review of control charts, can be used to determine if reanalysis (or remeasurement) is required before data are entered.

The data verification process, which is largely automated, checks that entered values are correct and, when possible, internally consistent within a sample or set of measurements. Examples of verification procedures include range checks, checks for duplicate entries, frequency checks of coded variables to identify inappropriate codes, and format checks to ensure data have been entered in the correct format. Data from QC check samples, blanks, or replicate samples, or redundant measurements for critical parameters (e.g., field versus laboratory pH), can be used to determine if there are problems with sample collection or measurement. Internal consistency checks include ion balance and conductivity calculations for inorganic chemical constituents. For biological samples and measurements, internal consistency checks include summing species proportions in samples to ensure they do not total to more than 100%, checks for missing taxa, evidence of "container effects" in bioassay experiments, and the taxonomic accuracy of species identifications.

Once the entered data are verified as accurate, they are validated by examination against regional expectations to identify and explain outlier samples or sites. Validation may involve comparison with historical data, or the use of association and multivariate analyses.

8.6 ASSESSMENT OF DATA QUALITY

Data quality assessment for EMAP-Surface waters will occur at three levels: (1) within a region (i.e., within a year's sampling), (2) among regions (i.e., within a four-year sampling cycle), and (3) among sampling cycles. Qualitative assessments will include documenting methods, using a sampling design that ensures unbiased and representative samples, and visiting sites to ensure consistency among participating groups. Quantitative assessment will attempt to estimate errors associated with sample collection and measurement that are important in interpreting indicators or in optimizing the QA program through time (adjusting the effort and intensity of QC to areas where it is needed the most).

The primary means of assessing error and uncertainty will be through carefully designed performance evaluation studies. These studies will test null hypotheses related to data quality requirements for random and systematic errors. The design will be based on consideration of Type I and Type II errors and will attempt to provide estimates of (1) total measurement error (Figure 8-1) for use in data interpretation activities, and (2) important components of variance within measurement error that can be used to determine which steps in the collection and measurement process require more, or less, QC emphasis. The sample sizes and frequency of measurement will be optimized to provide the necessary answers in the required reporting period.

Performance evaluation studies could be conducted using performance audit samples for chemical analyses, reference samples for biological measurements, and "round-robin" studies using natural samples (either chemical or biological). For water column chemistry, appropriate performance audit materials are available for most inorganic constituents. Depending on the constituent, appropriate materials for organic analyses may or may not be currently available. Materials for sediment chemistry and fish tissue chemistry may be limited in their availability and appropriateness.

As mentioned previously, an important issue in the program is the impact that using different methodologies, or modifying or changing methodologies over time, will have on data interpretation, particularly in the detection of trends in ecological condition. The QA program for EMAP-Surface Waters will provide standard guidelines for implementing a new methodology in a region or at a specific site. Performance evaluation studies will provide some information on methods comparability, but comparability studies should be a more intensive effort designed to test specific

hypotheses related to the comparability with a previous methodology. Such a comparability study must be conducted, evaluated, and approved before new or modified methods can be implemented.

8.7 QUALITY ASSURANCE REPORTING

In addition to the documentation described in Section 8.3, other types of reports will be produced periodically as part of the QA program, including (1) summary reports of site visits, (2) performance evaluation (or method comparisons) summaries, and (3) assessments of data quality. Summary reports of site visits will serve to identify and track issues and subsequent corrective actions, and provide information to update other QA documentation. The results of performance evaluation studies will be reviewed and returned to participants within a short time (within 30 days) after submission. Evaluation summaries of QA-related data and other appropriate information concerning data quality will be prepared and included in the appropriate EMAP-Surface Waters reports (Section 12).

In addition, the EMAP QA program plan requires the Surface Waters resource group to prepare a QA annual report and workplan (Table 3-2) each fiscal year. This document will be submitted to the EMAP QA coordinator. The QA annual report and workplan summarizes:

- The status of all data acquisition activities
- The status of DQO development
- QA activities (e.g., QA project plans and standard operating procedures)
- A summary of current data quality, including auditing activities
- QA resources utilized and recommendations for modifications in the QA program
- Planned QA-related activities for the upcoming fiscal year

The EMAP QA coordinator will combine the QA annual reports and workplans from each resource group into an overall QA annual report and workplan for EMAP.

9.0 INFORMATION MANAGEMENT

As do all research programs, EMAP-Surface Waters depends upon data: data generated both within and external to the program. Integration of the data is the key to determining the current status, extent, changes, and trends in the condition of our nation's lakes and streams on regional and national scales. Proper data management will be a critical factor in the success of the EMAP-Surface Waters Program. Section 9 describes the rationale, approach, and objectives, and the plan for ensuring the effective management of this essential resource.

9.1 DEFINITION OF INFORMATION MANAGEMENT

Information is data that has been manipulated so that it is useful. In other words, information must have value, or it is still data.

INFORMATION = f(DATA, PROCESSING)

Information management is the efficient management of data and information through all stages and processes. The information Management Program must provide for the effective collection, processing, storage, cataloging, retrieval, dissemination and reporting of all pertinent data and information. All these activities must be considered as part of any comprehensive information management effort.

An information management system (IMS) must support all these data and information activities through the combination of data bases and computer programs that perform the various functions and operations on the data.

9.1.1 Overview of EMAP-Surface Waters Information Management

EMAP-Surface Waters will collect a large volume of data. As the sampling program is expanded, the quantity of data collected will increase exponentially. The ability of EMAP-Surface Waters to manage and disseminate the large amounts of information collected will have a major influence on the success of the program.

An established goal of EMAP-Surface Waters is to publish statistical summaries within nine months after completion of data collection. The timely transformation of data into information will therefore be a key to success. A computerized IMS will be required to ensure that EMAP-Surface Waters data are made available in a timely manner.

The IMS must be made available at various levels to other EMAP resource groups, other local, state, and federal organizations, and to academic institutions. The system must also be capable of drawing from historical data bases already in place for similar environmental monitoring efforts, and it must be compatible across all EMAP Resource Group information management systems.

Although EMAP-Surface Waters information management encompasses a variety of activities, all are directed toward a single fundamental objective: timely and cost-effective access to all data generated or utilized by the program. The Information Management Program must provide for the effective collection, processing, storage, cataloging, retrieval, and dissemination of all pertinent data. All these activities must be considered as part of any comprehensive information management efforts.

9.1.2 Objectives of Surface Waters Information Management

The overall objectives of EMAP-Surface Waters information Management are to:

- Design an IMS that is responsive to user requirements within and outside EMAP-Surface Waters.
- Provide state-of-the-art information management technology within the guidelines of the Office of Information and Resource Management (OIRM).
- Facilitate the wide use of EMAP-Surface Waters data.
- Produce statistical summary reports within nine months after completion of data collection.
- Ensure integration with present and historical data.

- Develop a flexible IMS that can adapt to the programs changing needs.

9.2 INITIAL SYSTEM CONCEPT

9.2.1 User Requirements

9.2.1.1 Issues

An important issue for the EMAP-Surface Waters Information Program Plan is adequate security, to ensure the integrity of the information. We must assure that EMAP-Surface Waters data bases are protected from damage caused by mismanagement, computer viruses, unauthorized access, or hardware/software failures.

The confidentiality of data is also a very important issue. EMAP-Surface Waters may access and integrate data from other agencies that require controlled access and confidentiality agreements. EMAP-Surface Waters, EPA, and inter-agency policies on data confidentiality must be addressed in the maintenance and distribution of data. These issues are discussed at length by Franson (1990).

9.2.1.2 Data Categories

Completed data files will be grouped into the categories in the following list. Access to the data files will vary according the identified user group (subsection 9.2.1.3).

- Level A (Raw) - Data direct from field or analytical laboratory collection.
- Level B (Verified) - Raw data files that have completed the verification analysis defined in quality assurance (QA) plans.
- Level C (Validated) - Verified data files that have completed the validation analysis.
- Level D (Enhanced) - Validated data files in which missing values have been filled in using established statistical procedures.

9.2.1.3 Identified User Groups

The information generated by EMAP-Surface Waters will be used by the following groups:

- Decision makers at all levels of government who set environmental policy.
- Resource managers and regulators who require an objective basis for allocating resources and prioritizing actions, especially those focusing on the protection and enhancement of ecological resources.
- Researchers working on the development and understanding of ecological indicators and processes.
- Those interested in evaluating the effectiveness of the nation's environmental policies for protecting and enhancing ecological resources (e.g., the EPA Administrator and Senior Management staff, Congress, and the public).

9.2.1.4 Data User Requirements

Users and their requirements can be separated into six general categories.

1. EMAP-Surface Waters Core Research Group: Includes those individuals and groups charged with designing, implementing, and interpreting the data from the field sampling programs.

   Requirements: This group needs to have access to a comprehensive data set including:

   - Project management information
   - Sample tracking
   - Shipment tracking
   - Raw data files
   - Quality assurance/Quality control (QA/QC) reports
   - Field logs
   - Logistics
   - Summary reports
   - Maps
   - Verified and validated data sets

   This group also requires access to the data on a time basis as close to real time as possible. Raw data used by this group will not have been through QA. This group needs access to all data described in the other categories.

2. EMAP-Surface Waters team: Includes individuals and groups who are involved in the EMAP-Surface Waters effort but not in the day-to-day field operations. The team includes outside participating agencies, logistical support personnel, EMAP-Surface Waters

QA/QC personnel, program reviewers, and EPA Headquarters personnel.

Requirements: This group requires access to summary information regarding logistics and project management as well as validated data. They require access only to those files that have been validated and verified. They do not require real-time access, nor do they need to have access to a comprehensive data set.

3. EMAP Program: Includes all researchers directly involved in the design, implementation, and analyses of the national EMAP program. This group includes members of other task groups, members of the Synthesis and Integration Team, and personnel in other agencies directly involved in EMAP.

   Requirements: This group requires final summaries regarding logistics and project management. They require access only to those files that have been validated and verified. They do not require real-time access, nor do they need to have access to a comprehensive data set. They need data in a context that can be integrated with data from other disciplines. Document summaries with interpretation and graphic outputs will be most useful.

4. Other federal and state agencies involved in environmental monitoring programs: This group includes the EPA regions, other EPA offices, state agencies, universities, and other scientific research organizations.

   Requirements: This group requires access only to validated and verified data. They do not require any data concerning logistics and project management. They need summarized characterization data for each station sampled. They also need access to an index of available data (both EMAP and historical). They require access only to those files that have been validated and verified. They need data in a form that can be integrated with data from other disciplines. Document summaries with interpretation and graphic outputs will be most useful.

5. Legislators and Environmental Managers.

   Requirements: This group needs summarized and interpreted data. They do not require any data regarding logistics or project management. This group of users will be best served by published reports, maps, and an on-line summary system.

6. The General Public.

   Requirements: This group needs summarized and interpreted data. They do not require any data concerning logistics or project management. This group of users requires published reports, maps, and an on-line summary system.

9.2.2 System Management and Functional Requirements

Information management has different functions within three levels of EMAP-Surface Waters operation: Resource Group projects, the Resource Group Program; and overall EMAP. To support the needs of these levels, the EMAP-Surface Waters IMS must provide the following products and services.

- Resource Group Projects
  - Field data collection (data loggers, bar code readers)
  - Sample tracking
  - Analytical laboratory data collection
  - QA/QC analysis and reporting
  - Data transfer from other agencies
  - configuration management
- Resource Group Program
  - Access to existing data bases
  - Data integration and analysis
  - Data base management
  - Presentation and reporting
- Overall EMAP
  - Data base transfer/access
  - Data base catalog
  - Archival/Backup of data
  - Integration of data from multiple resource groups
  - Training and support

At each of the levels above, programs will be designed, developed, and tested to meet the needs identified. Standardization of components

will facilitate QA activities. Resource group standards will be developed by the EMAP Information Center (EIC) with input and review from the Information Management Committee (IMC). Where appropriate, standards will be implemented as defined.

9.2.3 Data and Information Flow

Figure 9-1 shows the overall flow of information and data through the Information Management System.

9.3 ORGANIZATION OF EMAP-SURFACE WATERS INFORMATION MANAGEMENT PROGRAM

The following structure has been developed for the EMAP-Surface Waters Information Management Program to respond to user and functional requirements.

9.3.1 Surface Water Information Manager

The EMAP-Surface Waters Information Manager is responsible for planning, coordinating, and facilitating information management activities within the group. The Information Manager reports to the Surface Water Technical Director, and is responsible for ensuring that data and information collected by EMAP-Surface Waters are properly captured, stored, and transmitted within EMAP-Surface Waters and EMAP in general. The Information Manager works closely with the Technical Director and technical staff to assure that information management adequately supports EMAP-Surface Waters in achieving its scientific objectives. The Information Manager accomplishes these tasks by establishing and directing a Surface Waters Information Center (SWIC), described in Section 9.3.2.

The Information Manager, as an executive member of the Information Management Committee (IMC), serves as the liaison between EMAP-Surface Waters and the Information Management Director (IMD), representing the needs of EMAP-Surface Waters scientists and managers.

9.3.2 Surface Waters Information Center (SWIC)

The Surface Waters Information Center (SWIC), managed by the Surface Water Information Manager, is the focal point for planning, coordinating, and implementing information management within EMAP-Surface Waters. The SWIC is one of the key elements to successful information management. The SWIC is composed of personnel (e.g., Information Manager, Data Base Administrator, programmers, data clerks) and hardware/software resources that support the EMAP-Surface Waters Information Management Program in the planning, implementation, and operation of the information management activities.

Figure 9-2 shows the overall structure of the Surface Water Information Center.

The following items describe some of the key personnel and their roles in supporting the SWIC:

- Information Manager (see Section 9.3.1)
- Data Base Administrator: This person is responsible for assembling, documenting, and administering the EMAP-Surface Waters data bases.
- Programmer: The programmer provides the programming support required for the information management activities. He or she will develop systems necessary for data collection, tracking, and dissemination (data entry screens, reports, sample tracking, analysis, etc.).
- Data Clerk/Librarian: The Data Clerk/Librarian is the person who helps develop the data bases and the EMAP-Surface Waters data catalog and data dictionary.
- Technical Support for Computer Services: The technical support person will in charge of the installation and maintenance of all hardware, software, and communications. This will not require a full-time person.

9.4 OPERATIONAL SPECIFICATIONS

The IMS must have the flexibility needed to handle the array of data types resulting from sample collection and processing. It must also support a variety of analysis, presentation, and reporting activities. For the 1991 Demonstration Project, the Statistical Analysis System (SAS) will be used for data management. SAS will also be used for most statistical analyses. SAS has been selected as the data management system

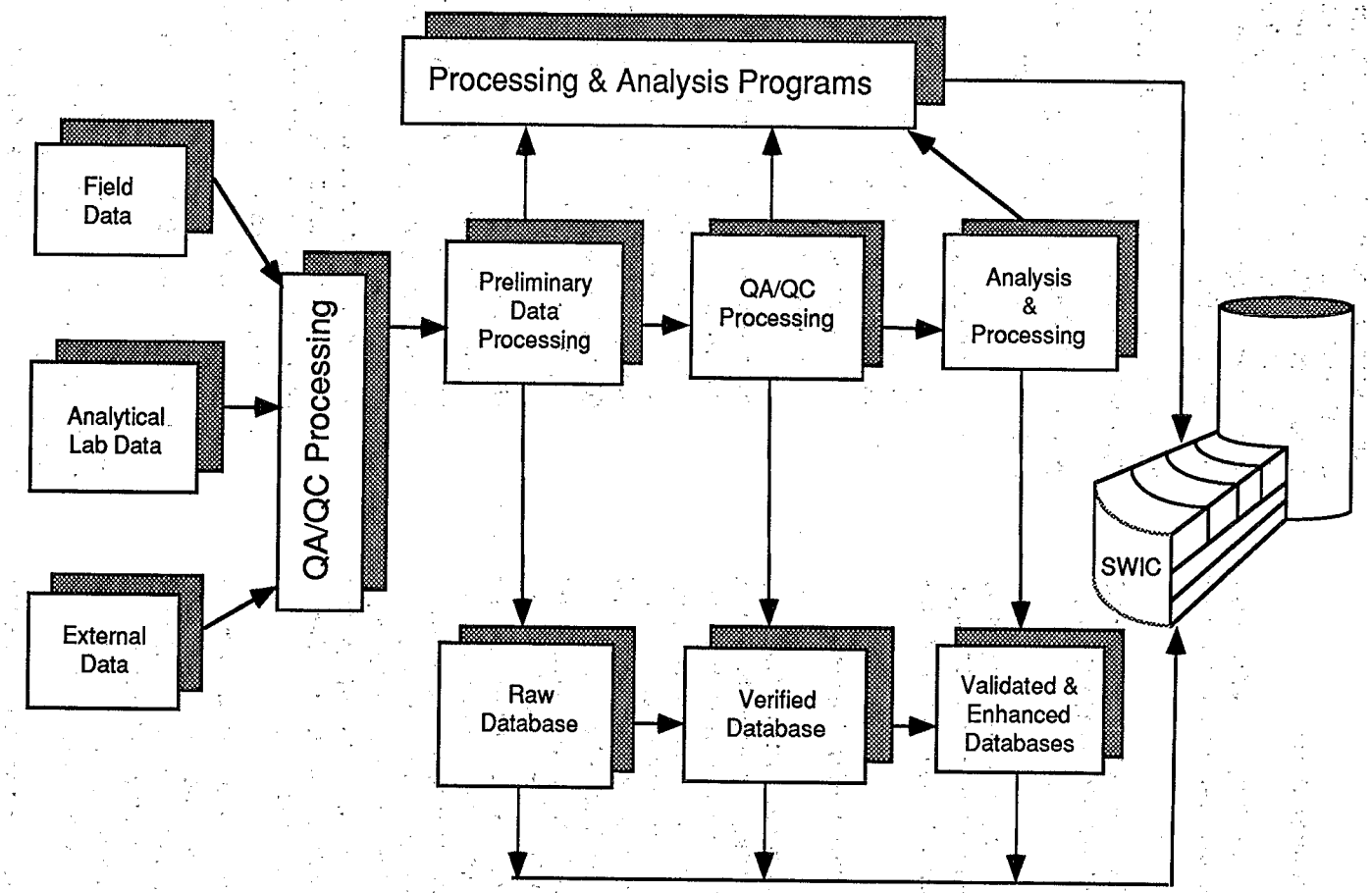


Figure 9-1. Data and information flow for the EMAP Surface Waters Information Center.

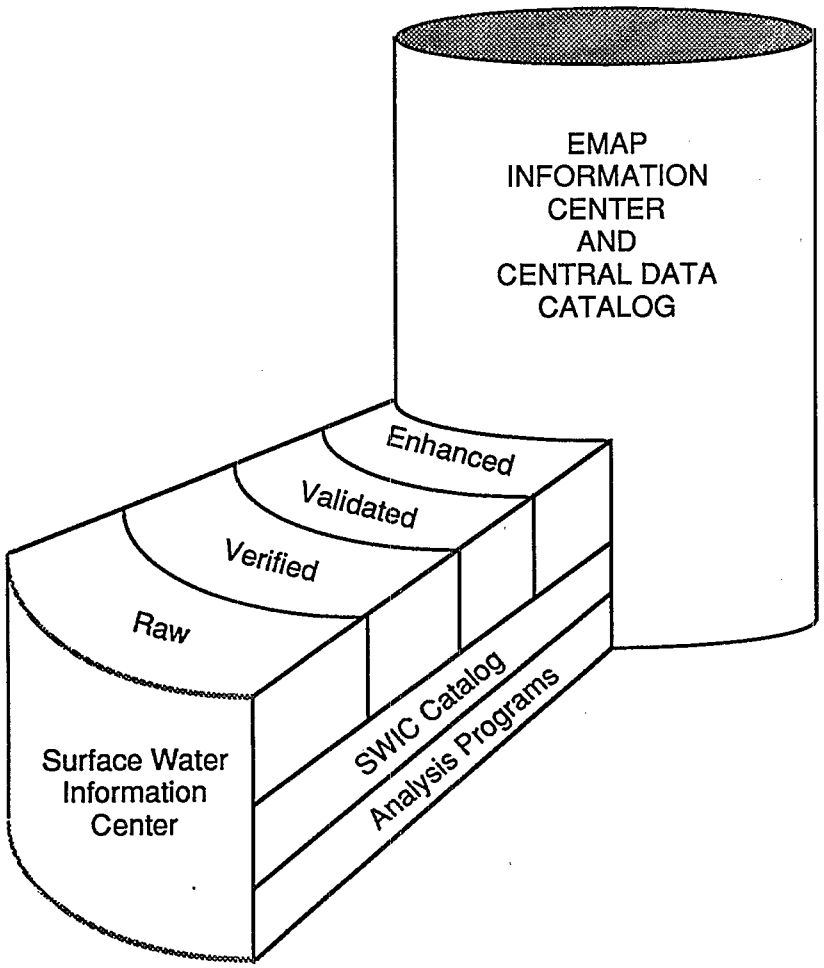


Figure 9-2. Concept of the EMAP-Surface Water Information Center in relation to the EMAP Information Center.

because no relational data base system is available to EPA through current contracting mechanisms. When a relational data management system becomes available to EPA through the Office of Information Resource Management (OIRM), the EMAP-Surface Waters data management system will be converted to the selected relational data base system, allowing users more flexible and efficient access to EMAP-Surface Waters data. After the conversion to a relational data base system, SAS will continue to be used as a principal data analysis tool.

To meet the requirements of the user groups, the EMAP-Surface Waters IMS will consist of the following major components for operations: project management; data and sample collection; processing and storage of indicator data; data access and transfer; data analysis and reporting; data documentation, access, and archival; geographical information system (GIS); and integration of data.

9.4.1 Project Management System

The project management component of the IMS requires frequent and accurate status reports about field collection and laboratory processing activities. The project management component of the IMS is used for this purpose. This component has two major elements: (1) a communications system for rapidly transferring information between field crews, processing laboratories, and the SWIC, and (2) a sample tracking system for monitoring the status of sampling events, sample shipments, and status of analyses on a real-time basis. These two elements of the Project Management Information System are discussed below.

9.4.1.1 Communications

Field crews and processing laboratories will use standard formats to submit data to the SWIC in established time frames. Reliable communications software and programs will facilitate this information exchange. For example, programs will be developed to automatically log remote computers into the central processing center and to perform file transfers into predetermined directories. Initial processing of the data will be accomplished automatically. When processing is complete, the SWIC will be notified and requested to acknowledge that it is aware the data are ready for additional processing.

Field crews will access data through the SWIC communications link. These data will include logistical information, such as the locations of equipment, personnel, and motels, and site-specific information about sampling locations.

9.4.1.2 Sample Tracking Information

The sample tracking system will track samples from initial collection through completion of all analyses and/or processing. To accomplish this, each sampling event and sample type will be assigned a unique identification number. These numbers will be entered into the IMS before data collection. Sample numbers will be bar coded to facilitate data entry by the field crews (if funding permits). Information to be entered for each sample in the sample tracking system available for retrieval and review includes:

- Sampling site name.
- The time the sample was collected, including date, hour, and duration of sampling effort.
- Type of sample (e.g., grab samples to be processed for benthic species composition and biomass, fish tissue sample to be processed for contaminant concentrations).
- Identification of the individual/team that collected the sample.
- A list of the analyses and processing activities planned for that sample, and the status of those analyses and activities (e.g., collection completed, analyses completed).
- Directions to files containing raw data generated for each sample.
- Directions to textual files containing descriptive information about the sampling event (e.g., field team comments).

When samples are transferred from field crews to analytical laboratories, a record of the exchange will be entered into the sample tracking system, both upon release and upon receipt of the materials. The identity and disposition of any sample can then be established by checking the sample status in the IMS. The status of all analyses and results will be available through the sample tracking system. When all processing for

a sample is complete, a "flag" will be set, and users will be notified automatically when logging onto the VAX computer. The GIS system will be linked to the sample tracking system to display the status of sampling activities.

9.4.2 Data and Sample Collection

Field data will be entered directly into a portable field data recorder. These data include direct field measurements, as well as site information (e.g., latitude, longitude, state, sample ID). The data will be automatically verified as described in the QA section of this chapter. Field data will be submitted to the SWIC in established time frames, and in approved formats.

Sample numbers will be assigned to each sample collected to ensure that all samples in the plan are properly identified and tracked. These numbers will be assigned by the field sampling coordinator and entered into the IMS before the sampling event. All sample labels will be produced before the sampling event.

9.4.3 Processing and Storage of Indicator Data

Quality assurance will be performed on all data received by the SWIC, using the procedures described in Sections 8.0 and 9.5, then the data will be converted into SAS data sets. The data sets will be stored in data libraries by indicator type. Following initial data processing, the required data analyses will be performed and summary data bases produced. The IMS will maintain data and relevant analytical results in both raw and summarized form. This will eliminate costly, redundant analyses.

9.4.4 Data Access and Transfer

The EPA VAX network will be the primary means of access to data in the IMS. All data base design work and documentation, including the code libraries, data dictionary, standard operating procedures for data handling, and GIS standards and base coverages will be available over this network. Users who do not have access to the VAX network will be provided with direct dial access to the SWIC, as appropriate. Ultimately, the data, reports, and findings of EMAP-Surface Waters will be important to many other groups within EMAP, to the scientific community, and to the general public. Access authorization will be established under the direction of the EMAP-Surface Waters Technical Director.

Historical data sets or data collected by other organizations that may be important, though not likely to be used regularly, will be documented and processed; QA will be conducted. Access through the IMS to confidential data, or data for which the quality is suspect or cannot be determined, will be limited.

All data made available for general use will be in read-only format, allowing users to access the data without compromising the integrity of the data base. Requests for copies of or access to data in the IMS will be submitted to a qualified person. A schedule will be developed for providing access to these data. The release schedule will depend on the availability of personnel to process the data, as well as the urgency of the request.

9.4.5 Data Analysis and Reporting

Analysis will be done only on data that have passed QA scrutiny. Qualified scientists will develop programs for the analyses. The IMS does not perform assessments of environmental health, but assists the Resource Group in performing data analyses and data presentations. Data exchange interfaces will be developed between the IMS, GIS, and other tools for data analyses. All results will be displayable shortly after they are incorporated into the IMS.

9.4.6 Data Documentation, Access, and Archival

All data received by the SWIC will be converted into SAS data sets in the VAX system. Complete documentation of all data bases stored in the SWIC is of paramount importance. The Data Set Index (DSI) will be the principal data information source and will include a catalogue listing all available data, modes of access, and quality of the data. A data library contained within the DSI will provide users with important information about the contents of each data set. A Central Data Dictionary will document information on standards that have been developed for data sets generated by EMAP and external data sets. The standards will include field names, formats, documentation, acceptable ranges, and codes. A data base backup system will be developed for

complete and rapid recovery of all information in the event of a data base failure.

9.4.7 Geographic Information System (GIS)

A major requirement of SWIC capabilities will be to create maps and perform geographically based analyses; therefore, the data generated for EMAP-Surface Waters will be geo-referenced. Spatial analyses will be accomplished using ARC-INFO, a GIS that is used throughout EPA. ARC-INFO is also used by most of the federal and state agencies participating in the EMAP program. ARC-INFO is a powerful tool which includes extensive analytical capabilities and interfaces with a number of other major software products, including SAS and ERDAS (a common software tool for processing data collected by satellites). ARC-INFO is not user friendly; therefore, user friendly interfaces for routine data analysis and display will be developed by GIS team analysts. EMAP-Surface Waters data analysts will work with other data management groups within EMAP and other agencies to develop standards and coverages for GIS applications. Standards will be developed for data accuracy, naming conventions, and documenting and archiving completed maps.

9.4.8 Integration

All data integrated into the IMS will be converted, if necessary, to EMAP-Surface Waters formats and standards.

9.4.9 Existing Information/Data

Existing data bases constitute an important source of information. They will be used to develop indicators, design the sampling program, and interpret the results of EMAP-Surface Waters surveys. Pertinent data bases will be included in the Data Set Index. Data bases used frequently will be converted to SAS data sets.

9.5 QA/QC FOR INFORMATION MANAGEMENT ACTIVITIES

Another important QA issue associated with a monitoring program the size of EMAP-Surface Waters is ensuring and maintaining the integrity of the large number of values that eventually will be entered into the data management system (NRC 1990, Risser and Treworty 1986, Packard et al. 1989). The usefulness of this system is in part determined by the quality of the information it contains. The scientific importance of this project requires the utmost confidence in the validity of the final data base. QA for information management includes those measures that ensure that information entered into the system is not corrupted.

Data become corrupted in two general ways: (1) incorrect information and (2) missing, incomplete, or nonretrievable information (Kanciruk et al. 1986). Examples of errors of the first type include typographical errors, incorrect species identification, and inaccurate instrument calibration. Although all errors cannot be completely eliminated through data management protocols, the potential for including incorrect information in the data base can be reduced. The second type of error is the omission of important information relating to legitimate data values (Kanciruk et al. 1986). Such pieces of information, called "data qualifiers" assist in the correct interpretation of data values. Qualifying information will be anticipated and a structured system for recording and retrieval of this information will be developed. The system will be designed with flexibility to allow the inclusion of unanticipated qualifiers.

QA and information management personnel will work in conjunction with other EMAP-Surface Waters personnel to prevent the corruption of data and to ensure the integrity of the data through all phases of activity (collection through dissemination and reporting). The features described in Section 9.5.1 through 9.5.8 will provide the foundation for the management and QA of all data through these phases. Automating as much of the standardized QA/QC process as possible will be the ultimate goal.

9.5.1 Field Data Collection

A systematic numbering system will be developed for unique identification of individual samples, sampling events, stations, shipments, equipment, and diskettes. Whenever possible, sample containers, equipment, and diskettes will be prelabelled to eliminate confusion in the field. The pre-labeling will reduce the number of incorrect or poorly affixed labels. If funding permits, bar code readers will be used to facilitate accurate sample number entry and identification in the field.

Standard operating procedures (SOPs) will document the use of the field computer systems (data

loggers) and the data entry procedures. Data loggers are hand held computers that display screens similar to manual field forms; they will automatically check for errors in data being entered (range checks on numeric data, biological species code verification). Explicit contingency plans using field forms will also be made, in the event that the field system fails. If field forms must be used, qualified persons will conduct manual verification before the data are entered into the system.

At periodic intervals, field crews will back up data collected in the field. Procedures for these field back-ups will be defined and crews trained in their use. Data will not be purged from the field computers until the information has been transferred and backed up.

9.5.2 Data Transfer

To minimize the errors associated with entry and transcription of data from one medium to another, data will be captured electronically when possible.

The communications protocols used to transfer data electronically will have mechanisms by which the completeness and accuracy of the transfer can be checked. If a file cannot be verified upon receipt of the data (i.e., incorrect number of bytes), a new file transfer will be requested.

9.5.3 Sample Tracking

The tracking and handling of samples collected in the field will be critical for QA of the resulting data. Sample tracking systems that provide accurate information on the status and location of the samples will be developed.

9.5.4 Data Entry/Verification

As stated previously, data will be captured electronically in order to minimize the errors associated with entry and transcription of data from one medium to another. However, at certain points of data acquisition, it will be necessary to enter these data manually (i.e., results of data analysis). When manual entry is required, the data entry process will include an accuracy check in which data will be entered twice or a 100% visual verification will be performed. The approved method will depend on a cost/benefit analysis. In many instances, the use of bar code labels should eliminate the need for manual entry of routine information.

The following examples show the types of checks that will be performed if manual entry is required:

- Range checks on numeric data: Field and laboratory numeric data and taxonomic species codes must be checked for acceptability using established criteria. Data that fall outside the acceptable range will be flagged and reviewed by QA personnel.

- Certain fields will contain coded information. When codes are used, they will be compared with the codes established by the scientific personnel and information management to assure compliance.

- Data that have been flagged will be reported. QA personnel will review this report and release data that have passed the QA check for addition to the data base. All errors must be corrected before the data can be entered.

9.5.5 Data Validation

All discrepancies identified by the computer will be documented in hard copy. These discrepancy logs will be saved as part of the EMAP-Surface Waters archive. Data will be flagged as described in Section 9.5.4. All identified discrepancies should be brought to the attention of a qualified QA person, who will determine the appropriate corrective action to be taken. Data will not be transferred to the data base until all discrepancies have been resolved by this person. A record of all additions will be kept in hard copy.

9.5.6 Data Analysis

Laboratories will be required to collect data and transmit it to a main repository. These data will include sample tracking information, results, and QA/QC information. Each set of data transmitted or received will include a full set of QA/QC data as specified by the EMAP-Surface Waters QA/QC plan.

Standard formats will be developed for data submitted electronically by the laboratories. A specific format will be developed for each file type within each discipline. Files that do not meet required specifications will be deleted and data will be requested in the proper format.

9.5.7 Archive and Backup

Data generated, processed, and incorporated by EMAP-Surface Waters will be stored and archived on redundant systems. This will ensure that if one system is destroyed or incapacitated, information management personnel will be able to reconstruct the data bases. Procedures will be developed to archive the data, monitor the process, and recover the data if necessary.

Several backup copies of the data (all levels) and of the programs used for processing the data will be maintained. Backups of the entire system will be maintained off site, using designated system backup procedures.

9.5.8 Configuration Management (CM)

Methods must be developed for keeping accurate records of the who, what, when, and why of any modification to the IMS. Whenever possible, a software configuration development and management package should be used, for example, a Computer Aided Software Engineering (CASE) tool.

Configuration management (CM) is the process of controlling the system's organization, construction, and maintenance. Although CM and QA are separate but complementary disciplines, both are required to ensure system integrity.

Five major sets of activities are involved in CM:

- Configuration identification - definition and identification of items subject to configuration control.
- Configuration control - evaluation, coordination, and approval or disapproval of proposed changes to controlled items.
- Configuration status accounting - recording and monitoring of changes to controlled items.
- Data management - maintenance of official correspondence records, CM records, and controlled documentation.
- Configuration auditing - verification that controlled items are what their documentation states they are and that they meet their assigned requirements.

A CM program appropriate for EMAP-Surface Waters will be developed and documented.

9.6 RESOURCES: OVERVIEW

The availability of information management resources is vital to the success of EMAP-Surface Waters. These resources are made available to EMAP-Surface Waters through authorized access, existing capabilities, or the procurement and acquisition of new resources. Existing resources and funding for new resources is limited; therefore, all steps must be taken to assure cost-effective use of all resources. EMAP-Surface Water's approach is to make use of existing and planned resources to the extent possible, and to acquire additional resources only where needed. Existing resources include current EPA staff, existing national and site contracts, EPA standard hardware and software, institutional resources (e.g., universities), other agencies' resources, and other contractor resources.

EMAP-Surface Waters will work with EPA's Office of Information Resources Management (OIRM) and the Office of Administration and Resources Management (OARM) to assure that the EMAP-Surface Waters Information Management Program is properly planned and coordinated with the Agency's IRM Program and that EMAP-Surface Waters resource requirements are conveyed through the appropriate channels. EMAP-Surface Waters will make extensive use of the automated data processing (ADP) resources available at the Office of Research and Development Laboratories. The Information Manager will work with the ADP coordinators and laboratory directors to plan and obtain approval for the utilization of these resources by EMAP-Surface Waters.

EMAP-Surface Waters Information Management will identify an ADP coordinator at each location within the program. Each ADP coordinator is an existing local resource whose responsibilities include assisting the Information Manager in development of the EMAP-Surface Water's ADP resources and ADP plans. These individuals are key figures in assuring that SWIC is fully supported by the ADP capabilities.

9.7 IMPLEMENTATION PLAN

The EMAP-Surface Waters IMS will be implemented in several phases over a four-year period. The SWIC will become operational when a full-

time Information Manager assumes control, basic information management systems are in place, and initial data collection activities have begun.

In FY91, Information Management will expand areas of development and operation for the large-scale pilot. Activities will focus on staffing and start-up of the SWIC. An Information Management Program Plan, and a System and Data Base Design Document will be generated. Initial prototypes of the various systems and data bases identified in the System and Data Base Design Document will be completed before the 1991 field season.

In FY92, Information Management will complete development of its core systems. External data will be integrated with EMAP-Surface Waters data. Prototype systems for electronic data access, the Data Set Index, and the Central Data Catalogue system will be designed and developed.

10.0 COORDINATION

This section focuses on coordination, among components within EMAP and between EMAP-Surface Waters and other state and federal programs.

In developing a complex program such as EMAP, a wide range of issues must be acknowledged and addressed throughout the early phases. The coordination of activities and analyses must occur at multiple levels both within the program and across other programs. For example, within the Surface Water component of EMAP, there is a need to integrate activities in the lake and stream components with one another and to integrate the Tier 1 and 2 activities with those that are occurring or might occur at Tiers 3 and 4 (see Section 2 for a description of the four-tier concept). Equally important is the need to integrate information from the various ecological resource categories within EMAP. Outside EMAP itself is a host of programs that, although they cannot adequately address the objectives of EMAP by themselves, can complement the information from EMAP to provide a clearer view of current status and trends in indicators of condition of surface waters and diagnosis of conditions.

10.1 WITHIN EMAP

10.1.1 Tier 3 Activities in Surface Waters

The primary focus of EMAP-Surface Waters during early phases must be on providing estimates of the extent and status of current conditions at broad regional scales. By doing so, we should be able to more effectively identify priority regions where significant proportions of the resource are in poor condition, and thus in need of attention, or where particularly sensitive subpopulations are at risk. We have described this as Tier 3 level activity. As described in Section 2.0, Tier 3 activities may be of several types: (1) more detailed diagnostic work to determine cause-effect linkages and remediation strategies or (2) follow-on monitoring to evaluate the effectiveness of a specific legislative or regulatory activity by monitoring early warning or particularly sensitive subpopulations. Depending on the exact nature of these activities, they should complement the design of EMAP-Surface Waters and use indicator methodologies that provide comparable and consistent results.

As explained throughout this plan, a Tier 3 activity is being designed within EMAP-Surface Waters to evaluate the effectiveness of the expected revisions to the Clean Air Act. The Temporally Integrated Monitoring of Ecosystems (TIME) project is intended to provide a Tier 3 activity that will elucidate, through monitoring, the changes and trends in chemical conditions in acid-sensitive surface waters of the United States that result from changes in acidic deposition. Specific goals are to:

- Monitor current status and trends in chemical indicators of acidification in acid-sensitive regions of the United States.
- Relate changes in acidic deposition to changes in surface water conditions.
- Assess the effectiveness of the Clean Air Act emissions reductions in improving the acid/base status of surface waters.

The TIME project must be able to make regional estimates of the changes in acid-base status, detect regional trends in a policy-relevant time span, and relate regional changes in surface water condition to regional changes in acidic deposition. The project design is national in scope but regional in implementation (Stoddard 1990). The design, implementation, and analysis of the TIME project is being conducted to maximize the use of existing resources and coordinate all aspects (e.g., quality assurance (QA) and data base management).

As new special projects emerge over the lifetime of EMAP-Surface Waters, we will make similar efforts to enhance the comparability and utility of the information generated.

10.1.2 Across EMAP Elements

An important aspect of EMAP is the inclusion of all ecological resources within the program. From the national perspective, this provides the opportunity to evaluate the relationships between conditions and problems in each resource category, their impacts on one another, and the

potential to more effectively evaluate comprehensive ecological resource management strategies. The achievement of these potentials requires extensive coordination in the selection of indicators, methodologies, and design.

Within EMAP, the focus to date has been the design and evaluation of programs to identify status, trends, and probable cause of conditions within each ecological resource group. Discussions on enhancing the integrative aspect of the program have been gradually increasing and will be the focus of the Integration and Assessment Team. Coordination between ecological resource groups has been facilitated by the coordination teams within EMAP (i.e., statistics and design, indicator development, integration and assessment, information management, QA, and logistics). The activities sponsored by these teams currently provide the framework to ensure our future ability to more fully integrate the information from each aspect of the EMAP.

Workshops have been held by the indicator development coordinator to facilitate discussion between the ecological resource groups. Effective communications will ensure that all groups take into account, as they develop their programs, the advantages of exchanging information with other groups. For example, measurement of nitrate in headwater streams by the Surface Waters group could help identify nutrient leakage for the Forest and Agroecosystem groups.

Coordination is also taking place as the specific resource categories (e.g., wetlands from lakes) are clearly defined and all ecological resources are considered. For example, riparian habitats might fall under wetlands, surface waters, or individual terrestrial groups (e.g., Arid Lands, Forests, Agroecosystems). The Arid Lands, Wetlands, and Surface Waters teams are now planning a joint pilot to evaluate how riparian habitat in the southwest can best be evaluated within the framework of EMAP. Similar coordination will be required between EMAP-Surface Waters, Wetlands, Forests, and Agroecosystems to address riparian habitats in other regions of the country.

The Integration and Assessment team, in conjunction with the Statistics and Design team, and the ecological resource groups will begin efforts during FY91 to evaluate design alternatives that might maximize the ability to integrate information from various ecological resource groups. These efforts can be coupled with the design and pilot activities of each resource group to evaluate long-term options for the program.

10.2 WITHIN EPA (OFFICE OF WATER AND REGIONS)

As described in Section 1.0, the U.S. EPA is responsible for providing the states with guidance in fulfilling their reporting requirements under Section 305(b) of the Clean Water Act. The EPA then integrates the information provided by the states and produces a report for the Administrator and Congress. The difficulty faced by the Office of Water is in reporting national status and trends. The flexibility allowed the states in the choice of methods, designation of beneficial uses and criteria, and selection of monitoring sites for evaluation makes the reporting of national status and trends an all but impossible task. The host of compliance-oriented issues for which the Office of Water and the regions are also responsible has made their charge quite extensive and given them little time to address the issue of national and regional status and trends. We have begun working with the Office of Water and the EPA regional offices to ensure that we effectively provide the national status and trends information they require. We have been working with the biocriteria development activities within the Office of Water to ensure comparable, or at least complementary, approaches between EMAP-Surface Waters and the guidance given the states.

One current limitation concerns spatial coverage. As presently envisioned and budgeted, EMAP-Surface Waters will be able to provide resolution only at a scale of 7 to 10 regions over the conterminous United States, rather than resolution at a state level. We recognize that the program will increase in value to the individual states, as well as to the Office of Water, if state-level resolution can be achieved. We have begun discussions to determine the various options available for achieving state-level resolution.

We hope that, in addition to providing information potentially useful to the Office of Water and to the regions, EPA regional offices will act as the coordination links to the states. The regional offices are familiar with the individual states within their boundaries, understand the state organizational structures, and know the individuals within various state agencies through whom we will need to work. The regions possess much of the expertise

needed to coordinate and implement the program. The regions also provide the nationwide infrastructure necessary to the success of a long-term effort such as EMAP.

In short, not only do we view the Office of Water and the U.S. EPA regional offices as clients for EMAP information, but we view them as important to the implementation and long-term success of the program.

10.3 OTHER FEDERAL AGENCIES

Almost every federal agency that has a mandate or jurisdiction over some natural resource has an interest in surface water resources. These agencies include, but are not limited to, the Agriculture Research Service, Bureau of Indian Affairs, Bureau of Land Management, Bureau of Reclamation, U.S. Fish and Wildlife Service (USFWS), National Science Foundation, National Park Service, U.S. Army Corps of Engineers, U.S. Forest Service, and U.S. Geological Survey (USGS). We firmly believe that EMAP should be a multi-agency program. The concept of the program is important, rather than its location within any particular agency. If aspects of the program objectives are being met through activities in other agencies, then there is no need to duplicate those aspects within EMAP.

Although we will need to coordinate with all the federal agencies mentioned above, we directed our early efforts toward the ones that have active surface water monitoring programs, but no planned interaction with other EMAP resource groups. This led to early interaction with USFWS and USGS. Programs maintained by each of these agencies were discussed in Section 1.0. In Sections 10.3.1 and 10.3.2, we describe potential scenarios for interaction. Discussions have been taking place between EMAP-Surface Waters and both these agencies, but no firm commitments have been achieved.

10.3.1 U.S. Fish and Wildlife Service

As described in Section 1.0, the USFWS maintains the National Contaminant Biomonitoring Program (NCBP). We have discussed with the USFWS their participation in EMAP-Surface Waters in a variety of contexts. We would like to see the concept of NCBP expanded to include the full range of lakes, reservoirs, streams, and rivers that will be sampled under EMAP-Surface Waters. This would not replace the current NCBP, but would function in addition to it. As a major federal agency responsible for biological resources, the USFWS is currently addressing monitoring both on and off the National Wildlife Refuges. Their concept appears to parallel that of EMAP in proposing to use information on ambient biotic assemblages, bioassay or toxicity tests, tissue residue data, and physical habitat data. As USFWS proceeds with these plans, we would like to work with them to ensure that their efforts and ours complement one another. The USFWS has extensive experience and expertise in dealing with aquatic biota from which EMAP-Surface Waters could benefit. The rigorous design and spatial coverage proposed by EMAP-Surface Waters could provide a beneficial context within which to evaluate the results of the National Wildlife Refuge monitoring efforts and the NCBP information.

Cooperation and coordination of USFWS activities and EMAP activities would benefit more than just EMAP-Surface Waters. The USFWS maintains the National Wetlands Inventory, which has begun discussions and interaction on a cooperative program with EMAP-Wetlands. Wildlife populations have been viewed as potentially valuable indicators of ecological resource condition for all of the terrestrial EMAP groups. We have proposed pursuing joint programs in most EMAP resource groups so that new efforts truly demonstrate added value and complement existing activities to the maximum extent possible.

10.3.2 USGS National Water Quality Assessment Program

In addition to maintaining the NASQAN and Benchmark networks, the USGS recently began development of the National Water Quality Assessment (NAWQA) program (Hirsch et al. 1988), which has been reviewed by the National Research Council (1990). The goals of this program are to:

- Provide a nationally consistent description of current water quality conditions for a large part of the nation's water resources.
- define long-term trends (or lack of trends) in water quality.
- Identify, describe, and explain, as possible, the major factors that affect observed water quality conditions and trends.

In many respects, EMAP-Surface Waters and NAWQA sound similar in scope and intent. Both are intended to address aquatic resources on regional and national scales over long periods of time. But each program has a different focus and will provide different information. An obvious difference is that NAWQA will focus on streams and rivers and EMAP-Surface Waters will focus on lakes, reservoirs, streams, and rivers. NAWQA is focusing on the component of aquatic resources that accounts for approximately 60% of the surface water withdrawal use and supplies roughly 60% of the human population served by surface water supplies. EMAP-Surface Waters has selected a national probability design so that the sample information can be extrapolated with known confidence to the entire surface water resource. NAWQA's goals are to describe water quality and to understand, within a hydrologic basin, those factors that influence it. EMAP-Surface Waters will focus on describing ecological condition, aspects of the chemical and physical habitat, and land use patterns that may affect ecological condition.

We believe that each of these programs has its own strengths as well as weaknesses. USGS has a mandate to monitor water quality, as does EPA. USGS has an excellent reputation in the quality of science it conducts, its experience in addressing water quality and hydrology issues, and its ability to commit to long-term programs. The Office of Research and Development within EPA is gaining a reputation for developing creative and innovative approaches to address aquatic resource issues and for designing programs that provide information needed to answer policy related questions. We firmly believe that the public and Congress will be best served if EMAP and NAWQA can be designed and conducted in a manner that will maximize the complementary nature of their information. We have intentionally begun the EMAP-Surface Waters activities with lakes rather than streams to provide more time for EPA and USGS to develop creative ways to ensure that NAWQA and EMAP-Surface Waters do not duplicate efforts but work in a complementary fashion. The two groups have discussed these issues over the past 18 months and will continue doing so until resolution.

10.4 STATE AGENCIES

Interaction with state agencies is important to the success of EMAP. We believe that the states may provide the most effective mechanism for implementing the field aspects of the program and a means of institutionalizing the activity. However, this cannot be done if it simply adds yet another project that the states are expected to carry out but does not give them the resources or support needed to do the job. Many states have the required expertise, or access to it. EMAP will need to coordinate interaction with the EPA regions who already have extensive interaction with state personnel. Participation in the program will certainly be of more interest to the states if we can intensify the activity to provide state-level resolution of information. We need to find creative approaches for satisfying the objectives of the program while supplying information of value to the states. These interactions have just begun in EPA Regions 1 and 2, the northeastern United States, and the FY91 activities will serve as a pilot for how we conduct interaction with the states in the future.

10.5 RESEARCH ORGANIZATIONS

As described in Section 2.0, research organizations such as universities and state and federal research centers are important groups with which EMAP-Surface Waters needs to interact. The research that will allow EMAP to reach its potential will come from these groups. Interaction with these groups during the developmental stages of EMAP will provide a mechanism for ensuring that the program is grounded in sound scientific principles. Existing field sites can provide locations for testing and developing indicators of ecological condition and for understanding the relationships between our indicators of response and indicators of exposure, habitat and stress. Existing sites, such as the Long-Term Ecological Research (LTER) sites funded by the National Science Foundation (Franklin et al. 1990, Magnuson 1990, Swanson and Sparks 1990), not only provide excellent sites for testing, developing, and understanding indicators, but also, if included as part of Tier 2 sampling, may provide information extremely useful for interpreting and understanding the changes or trends in indicator values that are observed. EMAP, on the other hand, can provide useful information that will allow research organizations to evaluate the context of their findings in terms of the portions of the resource their sites might represent and perhaps allow them to extrapolate their findings to a broader area. Once again, this is a situation in which programs certainly have different objectives but

can complement one another in significant ways if an early effort is made to communicate. The interaction at this level should by no means be limited to the LTER program. Important ecological research is occurring at many other locations and settings, and a mechanism must be developed over the next year to take advantage of the potential benefits which can be derived from interaction.

10.6 CONCLUSIONS

The types of interaction and coordination needed to create a successful program that best serves the public interest have been presented. During the next year, EMAP-Surface Waters will pursue these coordination efforts and seek to establish mechanisms to facilitate this coordination. We envision the need for a scientific advisory panel to ensure the sound footing of the program and an interagency coordination panel to facilitate the interaction and coordination needed among participating state and federal agencies.

11.0 EXPECTED OUTPUTS AND TIMELINES

Section 11.0 describes the types of products and outputs expected from EMAP-Surface Waters, along with expected timelines for field activities. The actual timing of these activities depends upon the available funding, but we present a proposed implementation schedule for discussion purposes.

11.1 REPORTING

11.1.1 Data Evaluation Reports

A key element of EMAP-Surface Waters is the production of information from the collected data in a timely manner and the dissemination of that information to users. Four types of reports are currently envisioned:

- Annual statistical summary
- Periodic interpretive reports
- Special interest reports
- Scientific articles

To be of maximum use, data must be transformed into useful information as quickly as possible. The objective of EMAP-Surface Waters is to publish summaries of the preceding field season's surveys within nine months of data collection. These reports will provide summaries of response, exposure, and habitat indicators for resources and regions sampled, but will provide minimal interpretation. Rigorous statistical evaluation of trends and associations between indicators will not be made. These reports are intended simply to get the information out to interested parties as quickly as possible. An example of the Annual Statistical Summary is presented in Paulsen et al. (1990).

Interpretive reports will be published following completion of four-year sampling cycles. These reports will attempt to summarize indicator results for the preceding four years and integrate information from the suite of response, exposure, habitat, and stressor indicators to determine regional and national status of lakes and streams. Trends will be evaluated and the likely causes for current conditions and changes in condition assessed. The later will be accomplished by analysis of the associations between response indicators and the exposure/habitat indicators. This diagnosis of probable cause will be further evaluated using information available for various stressor indicators. A detailed example of an interpretive report and the likely approach for analyses will be developed during the upcoming year.

Special issue reports will be used to present data pertaining to specific issues, such as changes in surface water chemistry resulting from implementation of the Clean Air Act or other special issue projects within EMAP-Surface Waters. Special interest analyses of particular indicators may also be published in this type of report. General reports for a broad spectrum of the public will be produced as part of this series.

Investigators within EMAP, and from programs outside EMAP, will be actively encouraged to produce analyses of the EMAP-Surface Waters results for publication in the peer review literature. One of the best methods of evaluating the quality of data produced in EMAP-Surface Waters will be its use in both policy-oriented and science-oriented publications.

11.1.2 Internal Program Reports

A series of reports and plans will be required before the program can commence field activity. These include, but are not limited to:

- Implementation plans
- Methods manuals
- Quality assurance plans
- Information management plans
- Field operations manuals
- Quality assessment reports

Each of these plans should be considered a "living" document. This implies that they will be updated before field implementation each year, as a result of the evaluation of the previous year's activities. The Implementation Plan will contain the objectives and outline of the field activity for that season. This will include a listing and location of all sites to be visited along with indicators to be collected. Additional details for all aspects of that season's activities, such as quality assurance, information management, or field operations, can then be found in updated program plans and manuals. An assessment of the quality of data collected will be required following each field season.

11.2 TIMELINES

Table 11-1 presents the timelines proposed for the implementation of EMAP-Surface Waters. Pilots are activities designed to test various aspects of the system, such as indicator evaluation or sampling and logistic constraints. Site selection for pilot activities need not be constrained to the EMAP-Surface Waters sampling grid. Demonstrations are activities in which parts of the system continue to be tested. A key aspect of the demonstration is the use of the sampling grid to provide policy-relevant information for some subset of the indicators on a population of aquatic resources of interest. The demonstration may incorporate additional sampling to evaluate indicator performance or examine components of variability for indicators of interest.

The timeline proposed in Table 11-1 represents an extremely optimistic time frame. It assumes adequate and timely funding for acquiring and training the staff needed for program planning, implementation and data analysis.

Table 11-1. Proposed EMAP-Surface Waters Implementation Schedule[a]

FY91	FY92	FY93	FY94	FY95	FY96
Lake pilot/ demo; Regions 1, 2	Lake demo; Regions 1,2	Lake demo; Regions 1,2	Lakes; Regions 1,2	Lakes; Regions 1,2	Lakes; Regions 1,2
	Lake pilot/ demo; Regions 3,4,5[b]	Lake demo; Regions 3,4,5	Lake demo; Regions 3,4,5	Lakes; Regions 3,4,5	Lakes; Regions 3,4,5
		Lake pilot/ demo; Regions 6,7,8	Lake demo; Regions 6,7,8	Lake demo; Regions 6,7,8	Lakes; Regions 6,7,8
			Lake pilot/ demo; Regions 9,10	Lake demo; Regions 9,10	Lake demo; Regions 9,10
	Stream pilot; Regions 8,9	Stream demo; Regions 8,9,10	Stream demo; Regions 8,9,10	Streams; Regions 8,9,10	Streams; Regions 8,9,10
		Stream pilot Regions 6,7	Stream demo; Regions 6,7	Stream demo; Regions 6,7	Streams; Regions 6,7
			Stream pilot; Regions 3,4	Stream demo; Regions 3,4	Stream demo; Regions 3,4
				Stream pilot; Regions 1,2,5	Stream demo; Regions 1,2,5

[a] Regions refers to the EPA regions.

[b] The lake pilot and demonstration in 1992 are scheduled for EPA Regions 3, 4, and 5. If funding constraints limit the scope of activities, Region 5 will be the highest priority. This will ensure that lake pilots and demonstrations are completed in the northern tier of lake states in succeeding years.

12.0 FISCAL YEAR 1991 FIELD AND ANALYSIS ACTIVITIES

"It is common sense to take a method and try it. If it fails, admit it frankly and try another. But above all, try something."

Franklin D. Roosevelt

As a first step in accomplishing the objectives of EMAP-Surface Waters, a field pilot will be implemented in the lakes of the Northeast, EPA Regions 1 and 2, in 1991. The goal of this activity is to demonstrate in the Northeast Region the utility and value of the EMAP-Surface Waters sampling design and indicator approach, and at the same time, to collect the information necessary to improve and develop a technically sound and cost-effective program. Specific objectives include the following:

- Evaluate the adequacy of the proposed Tier 2 sample allocation scheme and collect information to evaluate alternative designs.
- Evaluate the reliability, variability and sensitivity of the proposed response indicators over a range of environmental conditions.
- Refine field sampling methodologies.
- Identify and resolve unexpected logistical problems associated with sampling lakes in the Northeast.
- Develop and evaluate data handling, quality assurance, and statistical procedures for efficient analysis and reporting.
- Produce, on a pilot basis, regional estimates of the condition of lakes in the Northeast based on a limited suite of response, exposure, and habitat indicators.

The Temporally Integrated Monitoring of Ecosystems (TIME) project will also begin in the Northeast in FY91. Specific objectives associated with this activity include:

- Use field information to refine the subpopulations of concern relative to acidic deposition in the Northeast.
- Evaluate the appropriateness of the proposed grid augmentation for addressing issues related to the Clean Air Act.
- Refine the selection of regional long-term monitoring sites to represent subpopulations of concern.
- Produce regional estimates of lakewater quality with respect to acidic deposition.

12.1 GENERAL EMAP-SURFACE WATERS

In order to meet the objects just outlined, the EMAP-Surface Waters pilot will take place in EPA Regions 1 and 2, the northeastern United States. Approximately 60 lakes will be sampled during this effort. Of these sites, approximately 40 will be selected by procedures outlined in Section 3.0. Ten percent of these will be revisited in order to evaluate index period variability. Indicators outlined in Section 5.0 will be collected using the late summer index period.

An additional 20 sites will be sampled to evaluate indicator performance. Ten highly perturbed sites, 10 minimally perturbed sites and 10 randomly selected sites will be chosen and sampled using the protocols outlined for the regional demonstration activity. A pilot implementation to be conducted in spring 1991 will detail the proposed activities.

12.2 TIME PROJECT

The TIME project will be implemented in the Northeast in 1991 as well. This activity will consist of continued monitoring of the approximately 30 LTM sites described in Stoddard (1990), and an additional 85 sites in the subpopulations of interest. The regional sampling in 1991 will help us to (1) confirm the level of augmentation of the standard grid that will be needed for underrepresented areas and (2) begin to generate new estimates of regional condition that will be used to help evaluate the Clean Air Act.

12.3 DATA ANALYSIS ACTIVITIES

Many of the data analyses that were begun in 1990 will continue during 1991. The primary effort will be to develop an approach for examining the associations between response indicators and exposure/habitat indicators. The statistical approaches selected will aid us in resolving outstanding design issues such as (1) whether to use the hexagon or watershed characterization and (2) the density needed to achieve the desired level of uncertainty in population estimates. Continued evaluation of existing data will help to refine aspects of the design such as the level of change and trend detectable with the variability that exists in our indicators. We will pursue increased interaction with members of the academic community, the states, and other federal agencies to identify data sets and analysis approaches that might better allow us to reach our objectives.

12.4 RESEARCH NEEDS

As stated at the outset, EMAP must be accompanied by an aggressive research program. The range of researchable issues is quite long. Key issues can be divided into four categories: (1) indicators of condition, (2) diagnostic indicators, (3) approaches for defining nominal/subnominal, and (4) statistical approaches.

Extensive work is required to refine our selection and choice of indicators of condition. Existing indicators need regional refinement and new indicators are required to fill the gaps. Additional work needs to be done to improve our diagnostic capabilities with biological indicators.

The field of diagnostic indicators of exposure and habitat offers many challenges and potential benefits. Biomarkers appear to be a potentially valuable tool in evaluating both condition and exposure.

Although the decision on what is an acceptable condition is ultimately a societal one, we have an obligation to provide creative approaches to addressing this issue. The process of gathering scientists, policy analysts, managers, and decision makers together to tackle this problem will offer many challenges.

In addressing national assessments such as those proposed, many statistical challenges exist. The process of developing indicators from many metrics is a difficult task but has been accomplished successfully. However, the process of combining multiple indicators into a single statement of condition will require extensive statistical as well as ecological insight and effort. Trend analysis at a single site is somewhat straightforward but the task of doing trend analysis on population distributions is not. This effort will need attention early in the program. In assessing the effectiveness of regulatory activities, we would like to evaluate changes and trends in response indicators and then associate them with changes or trends in exposure or stressor information. Regional analysis of associations among indicators and associations of trends in indicators is an area that must be addressed soon.

REFERENCES

Agbeti, M., and M. Dickman. 1989. Use of lake fossil diatom assemblages to determine historical changes in trophic status. *Can. J. Fish. Aquat. Sci.* 46:1013-1021.

Baker, J.P., D.P. Bernard, S.W. Christensen, M.J. Sale, and others. 1990a. Biological Effects of Changes in Surface Water Acid-base Chemistry, SOS/T Report 13 in Acidic Deposition: State of Science and Technology. National Acid Precipitation Assessment Program, 722 Jackson Place, Washington, D.C.

Baker, J.P., S.A. Gherini, S.W. Christensen, C.T. Driscoll, J. Gallagher, R.K. Munson, R.M. Newton, K.H. Reckhow, and C.L. Schofield. 1990b. *Adirondack Lakes Survey: An Interpretative Analysis of Fish Communities and Water Chemistry, 1984-1987.* Adirondack Lakes Survey Corporation, Ray Brook, NY.

Baker, L.A., J.M. Eilers, R.B. Cook, P.R. Kaufmann, and A.T. Herlihy. In press. Current water chemistry and biogeochemical processes in case study regions: intercomparisons and synthesis. Chapter 18 in D.F. Charles, ed. *Acidic Deposition and Aquatic Ecosystems: Regional Case Studies.* Springer-Verlag, New York.

Baker, L.A., P.R. Kaufmann, A.T. Herlihy, and J.M. Eilers. 1990. Current Status of Surface Water Acid-Base Chemistry, SOS/T Report 9, in Acidic Deposition: State of Science and Technology, National Acid Precipitation Assessment Program, 722 Jackson Place, Washington, D.C. 1990.

Battarbee, R.W. 1986. Diatom analysis. Pages 527-570 in B.E. Berglund, ed. *Handbook of Holocene Paleoecology and Paleohydrology.* John Wiley and Sons, New York.

Bascietto, J., D. Hinckley, J. Plafkin, and M. Slimak. 1990. Ecotoxicity and ecological risk assessment: Regulatory applications at EPA. *Environ. Sci. Technol.* 24(1):10-15.

Bonoff, M.B. and A.W. Groeger. 1987. National Surface Water Survey, Western Lakes Survey - Phase I, Field Operations Report. EPA 600/8-87/018. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Las Vegas, Nevada. 32 pp.

Boyle, T.P., G.M. Smillie, J.C. Anderson, and D.R. Beeson. 1990. A sensitivity analysis of nine diversity and seven similarity indices. *J. Water Pollut. Control Fed.* 62:749-762.

Brezonik, P.L. 1984. Trophic state indices: rationale for multivariate approaches. Pages 441-445 in *Lake and Reservoir Management.* EPA/440/5-84/001. U.S. Environmental Protection Agency, Washington, D.C.

Brooks, R.P., D.E. Arnold, E.D. Bellis, C.S. Keener, and M.J. Croonquist. In Press. A methodology for biological monitoring of cumulative impacts on wetland, stream and riparian components of watersheds. Pages (in press). J.A. Kusler, and G. Brooks, eds. Proc. Internat. Symp.: Wetlands and river corridor management. Assoc. Wetland Managers, Inc., Berne, N.Y.

Brown, J.H., and J. Roughgarden. 1990. Ecology for a Changing Earth. *Bull. Ecol. Soc. Amer.* 71:173-188.

Carlson, R.E. 1977. A trophic state index for lakes. *Limnol. Oceanogr.* 22:361-369.

Carlson, C.A., and R.T. Muth. 1989. The Colorado River: Lifeline of the American Southwest. Pages 220-239 in D.P. Dodge, ed. Proceedings of the International Large Rivers Symposium. *Can. Spec. Publ. Fish. Aquat. Sci.* 106.

Charles, D.F. and J.P Smol. In press. The PIRLA II Project: Regional assessment of lake acidification trends. *Verh. Internat. Verein. Limnol.*

Charles, D.F., M.W. Binford, E.T. Furlong, F.A. Hites, M.J. Mitchell, S.A. Norton, F. Oldfield, M.J. Paterson, J.P. Smol, A.J. Uutala, J.R. White, D.R. Whitehead, and R.J. Wise. 1990. Paleoecological investigation of recent lake acidification in the Adirondack Mountains, N.Y. *J. Paleolimnol.* 3:195-241.

Church, M.R., K.W. Thornton, P.W. Shaffer, D.L. Stevens, B.P. Rochelle, G.R. Holdren, M.G. Johnson, J.J. Lee, R.S. Turner, D.L. Cassell, D.A. Lammers, W.G. Campbell, C.I. Liff, C.C. Brandt, L.H. Liegel, G.D. Bishop, D.C. Mortenson, S.M. Pierson, and D.D. Schmoyer. 1989. Future effects of long-term sulfur deposition on surface water chemistry in the Northeast and Southern Blue Ridge Province. EPA-600/3-89/061. Washington, DC: U.S. Environmental Protection Agency.

Colquhoun, J.R., J. Symula, M. Pfeiffer, and J. Feurer. 1981. Preliminary Report of Stream Sampling for Acidification Studies. Technical Report 812. New York State Department of Environmental Conservation, Albany, N.Y.

Cougan, K.A., D.W. Sutton, D.V. Peck, V.J. Miller, and J.E. Pollard. 1988. National Stream Survey -Phase I: Quality Assurance Report. EPA 600/4-88/018. U.S. Environmental Protection Agency, Las Vegas, Nevada. 205 pp.

Courtemanch, D.L., S.P. Davies, and E.B. Laverty. 1989. Incorporation of biological information in water quality planning. *Environ. Manage.* 13:35-41.

Courtemanch, D.L., and S.P. Davies. 1988. Implementation of biological standards and criteria in Maine's water classification law. Pages 4-9 in T.P. Simon, L.L. Holst, and L.J. Shepard, eds. *Proceedings of the First National Workshop on Biological Criteria.* EPA/905/9-89/003. U.S. Environmental Protection Agency, Chicago, IL.

Cowardin, L.M., V. Carter, F.C. Golet, and E.T. LaRoe. 1979. *Classification of Wetlands and Deepwater Habitats of the United States.* FWS/OBS-79/31. USDI Fish and Wildlife Service. 103 pp.

Cummins, K.W., and M.J. Klug. 1979. Feeding ecology of stream invertebrates. *Ann. Rev. Ecol. Syst.* 10:147-172.

Cummins, K.W., and M.S. Wilzbach. 1985. *Field Procedures for Analysis of Functional Feeding Groups of Stream Macroinvertebrates.* Contribution 1611. Appalachian Environmental Laboratory, University of Maryland, College Park MD.

Cusimano, R.F., J.P. Baker, W.J. Warren-Hicks, V. Lesser, W.W. Taylor, M.C. Fabrizio, D.B. Hayes, and B.P. Baldigo. 1990. *Fish Communities in Lakes in Subregion 2B (Upper Peninsula of Michigan) in Relation to Lake Acidity.* EPA/600/3-89/021. U.S. Environmental Protection Agency, Washington, D.C.

Deisler, P.F., Jr. 1988. The risk management-risk assessment interface. *Environ. Sci. Technol.* 220:15-19.

Dixit, S.S., J.P. Smol, D.S. Anderson, and R.B. Davis. 1990. Utility of scaled chrysophytes for inferring lakewater pH in northern New England lakes. *J. Paleolimnol.* 3:269-286.

Ebel, W.J., C.D. Becker, J.W. Mullan, and H.L. Raymond. 1989. The Columbia River - toward a holistic understanding. Pages 205-219 in D.P. Dodge, ed. Proceedings of the International Large River Symposium. *Can. Spec. Publ. Fish. Aquat. Sci.* 106.

Ecology Committee. 1980. Goals and criteria for design of a biological monitoring system. Science Advisory Board. U.S. Environmental Protection Agency, Washington, D.C. 53 pp.

Einhaus, R.L., D.M. McMullen, R.L. Graves, and P.H. Friedman. 1990. *Environmental Monitoring and Assessment Program Quality Assurance Program Plan.* Revision 0. U.S. Environmental Protection Agency, Cincinnati, OH. 123 pp.

Ehrlich, P.R., A.H. Ehrlich, and J.P. Holdren. 1977. *Ecoscience: Population, Resources, Environment.* W.H. Freeman, Sanfrancisco, CA. 1051 pp.

Eshleman, K.N. 1988. Predicting regional episodic acidification of surface waters using empirical techniques. *Water Resour. Res.* 24:1118-1126.

Fava, J.A., W.J. Adams, R.J. Larson, G.W. Dickson, K.L. Dickson, and W.E. Bishop. 1987. *Research Priorities in Environmental Risk Assessment.* Society of Environmental Toxicology and Chemistry. Washington, D.C.

Federal Register 48 (217):51400-51413.

Fjerdingstad, E. 1965. Some remarks on a new saprobic system. Pages 232-235 in *Biological Problems in Water Pollution.* Public Health Service Publication No. 999-wp-25. Third Seminar. U.S. Public Health Service, Cincinnatti, OH.

Fiske, S. 1988. The use of biosurvey data in the regulation of permitted nonpoint discharges in Vermont streams. Pages 67-74 in T.P. Simon, L.L. Holst, and L.J. Shepard, eds. *Proceedings of the First National Workshop on Biological Criteria.* EPA/905/9-89/003. U.S. Environmental Protection Agency, Chicago, IL.

Ford, J., A.D. Newell, P.L. Brezonik, C.T. Driscoll, J.S. Kahl, and K.E. Webster. 1989. In preparation. *Trends in Surface Water Chemistry for Low Alkalinity Lakes of the Northeastern United States: Results from the EPA long-term Monitoring Project.* U.S. Environmental Protection Agency, Corvallis, OR.

Franson, S. 1990. *EMAP Data Confidentiality: Issues and Suggestions.* Internal Report. U.S. Environmental Protection Agency, Las Vegas, Nevada.

Geraghty, J.J., D.W. Miller, F. Van Der Leeden, and F.L. Troise. 1973. *Water Atlas of the United States.* Water Information Center, Inc., Port Washington, NY. 122 plates.

Gerritsen, J., J. Dietz, H.T. Wilson, Jr., and A.J. Janicki. 1989. *Prediction of Episodic Acidification in Maryland Coastal Plain Streams.* Maryland Dept. of Natural Resources. Annapolis, MD.

Gorman, O.T., and J.R. Karr. 1978. Habitat structure and stream fish communities. *Ecology* 59:507-515.

Grossman, G.D., J.F. Dowd, and M. Crawford. 1990. Assemblage stability in stream fishes: a review. *Environ. Manage.* 14:661-671.

Hart, D.R. 1990. *QA Guidelines for Biology in Aquatic Environmental Protection.* Prepared for Environment Canada, Conservation and Protection, Inland Waters Directorate, Burlington, Ontario.

Hagley, C.A., C.L. Mayer, and R. Hoenicke. 1989. *National Stream Survey - Phase I: Field Operations Report.* EPA 600/4-88/023. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Las Vegas, NV. 34 pp.

Herlihy, A.T., D.H. Landers, R.F. Cusimano, W.S. Overton, P.J. Wigington, Jr., A.K. Pollack, and T.E. Mitchell-Hall. 1991. *Temporal Variability in Lakewater Chemistry in the Northeastern United States: Results of Phase II of the Eastern Lakes Survey.* EPA/ . U.S. Environmental Protection Agency, Washington, D.C.

Hesse, L.W., J.C. Schmulbach, J.M. Carr, K.D. Keenlyne, D.G. Unkenholz, J.W. Robinson, and G.E. Mestl. 1989. Missouri River fishery resources in relation to past, present, and future stresses. Pages 352-371 in D.P. Dodge, ed. Proceedings of the International Large River Symposium. *Can. Spec. Publ. Fish. Aquat. Sci.* 106.

Hilsenhoff, W.L. 1987. An improved biotic index of organic stream pollution. *Gr. Lakes Entomol.* 20:31-39.

Hilsenhoff, W.L. 1988. Rapid field assessment of organic pollution with a family level biotic index. *J. North Am. Benthol. Soc.* 7:65-68.

Hirsch, R.M., W.M. Alley, and W.G. Wilber. 1988. *Concepts For a National Water-Quality Assessment Program.* U.S. Geological Survey Circular 1021. 42 pp.

Horvitz, D.G., and D.J Thompson. 1952. A generalization of sampling without replacement from a finite universe. *J. Am. Statist. Assoc.* 47:663-685.

Hughes, R.M. 1989. Ecoregional biological criteria. Pages 147-151 in *Water Quality Standards for the 21st Century.* U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Hughes, R.M., and S.G. Paulsen. 1990. Indicator strategy for inland surface waters. Pages 4.1-4.20 in C.T. Hunsaker and D.E. Carpenter, eds. *Environmental Monitoring and Assessment Program Ecological Indicators.* EPA/600/3-90/060. U.S. Environmental Protection Agency, Research Triangle Park, NC.

Hughes, R.M., D.P. Larsen, and J.M. Omernik. 1986. Regional reference sites: a method for assessing stream potentials. *Environ. Manage.* 10:629-635.

Hughes, R.M., C.M. Rohm, T.R. Whittier, and D.P. Larsen. 1990. A regional framework for establishing recovery criteria. *Environ. Manage.* 14:673-684.

Hughes, R.M., S.A. Thiele, D. McMullen, J. Lazorchek, S.G. Paulsen, J.P. Baker, D. Peck, J. Pollard, D.P. Larsen, S. Hedthe, T.R. Whittier, P.R. Kaufmann, and J.R. Karr. 1990. Appendix B: indicator fact sheets for inland surface waters. Pages B.1-B.32 in C.T. Hunsaker and D.E. Carpenter, eds. *Ecological Indicator Report for the Environmental Monitoring and Assessment Program.* U.S. Environmental Protection Agency, Research Triangle Park, NC.

Hunsaker, C.T., and D.E. Carpenter, eds. *Ecological Indicator Report for the Environmental Monitoring and Assessment Program.* U.S. Environmental Protection Agency, Research Triangle Park, NC.

Hunsaker, C.T., R.L. Graham, G.W. Suter II, R.V. O'Neill, L.W. Barnthouse, and R.H. Gardner. 1990. Assessing ecological risk on a regional scale. *Environ. Manage.* 14(3):325-332.

Hunt, D.T.E., and A.L. Wilson. 1986. *The Chemical Analysis of Water.* Second Edition. Royal Society of Chemistry, London, England. 683 pp.

Hynes, H.B.N. 1974. *The Biolgoy of Polluted Waters.* University of Toronto Press, Toronto, Ontario.

Jackson, D.A., H.H. Harvey, and K.M. Somers. 1990. Ratios in aquatic sciences: statistical shortcomings with mean depth and the morphoedaphic index. *Can. J. Fish. Aquat. Sci.* 47:1788-1795.

Jenkinson, J.J. 1981. Endangered or threatened aquatic mollusks of the Tennessee River system. *Bull. Am. Malacolog. Union.* 1981:43-45.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. *Fisheries* 6:21-27.

Karr, J.R., and D.R. Dudley. 1981. Ecological perspective on water quality goals. *Environ. Manage.* 5:55-68.

Karr, J.R., L.A. Toth, and G.D. Garman. 1983. *Habitat Preservation for Midwest Stream Fishes: Principles and Guidelines.* EPA/600/3-83/006. U.S. Environmental Protection Agency, Corvallis, OR.

Karr, J.R., L.A. Toth, and D.R. Dudley. 1985. Fish communities of midwestern rivers: A history of degradation. *BioScience* 35:90-95.

Karr, J.R., D.D. Fausch, P.L. Angermeier, P.R. Yant, and I.J. Schlosser. 1986. Assessing biological integrity in running waters: a method and its rationale. Special Publication 5. Illinois Natural History Survey, Champaign, IL. 28 pp.

Kaufmann, P.R., A.T. Herlihy, J.W. Elwood, M.E. Mitch, W.S. Overton, M.J. Sale, J.J. Messer, K.A. Cougan, D.V. Peck, K.H. Reckhow, A.J. Kinney, S.J. Christie, D.D. Brown, C.A. Hagley, and H.I. Jager. 1988. *Chemical Characteristics of Streams in the Mid-Atlantic and Southeastern United States. Volume I: Population Descriptions and Physico-Chemical Relationships.* EPA/600/3-88/021a. U.S. Environmental Protection Agency, Washington, D.C. 397 pp.

Kaufmann, P.R., A.T. Herlihy, M.E. Mitch, and W.S. Overton. In press. Stream chemistry in the eastern United States: 1. Synoptic survey design, acid-base status and regional patterns. *Water Resour. Res.*

Kirchner, C.J. 1983. Quality control in water analyses. *Environ. Sci. Technol.* 17:174A-181A.

Knapp, C.M., C.L. Mayer, D.V. Peck, J.R. Baker, and G.J. Filbin. 1987. *National Surface Water Survey, National Stream Survey (Phase I - Pilot Survey) Field Operations Report.* EPA 600/8-87/019. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Las Vegas, NV. 110 pp.

Knapp, C.M., W.P. Saunders, Jr., D.G. Heimbuch, H.S. Greening, and G.J. Filbin. 1988. *Maryland Synoptic Stream Chemistry Survey: Estimating the Number of Distribution of Streams Affected by or at Risk from Acidification.* Prepared for the Maryland Department of Natural Resources, Power Plant Research Program. Report No. AD-88-2.

Knapp, C.M., D.R. Marmorek, J.P. Baker, K.W. Thornton, J.M. Klopatek, and D.P. Charles. 1991. *The Indicator Development Strategy for the Environmental Monitoring and Assessment Program.* U.S. EPA Environmental Research Laboratory, Corvallis, OR.

Knowlton, M.F., M.V. Hoyer, and J.R. Jones. 1984. Sources of variability in phosphorus and chlorophyll and their effects on use of lakes survey data. *Water Resour. Bull.* 20:397-407.

Landers, D.H., J.M. Eilers, D.F. Brakke, W.S. Overton, P.E. Kellar, M.E. Silverstein, R.D. Schonbrod, R.E. Crowe, R.A. Linthurst, J.M. Omernik, S.A. Teague, and E.P. Meier. 1987. *Characteristics of Lakes in the Western United States. Volume I. Populations Descriptions and Physico-chemical Relationships.* EPA/600/3-86/054a. U.S. Environmental Protection Agency, Washington, D.C. 176 pp.

Landres, P.B. 1983. Use of the guild concept in environmental impact assessment. *Environ. Manage.* 7:393-398.

Lee, D.S., C.R. Gilbert, C.H. Hocutt, R.E. Jenkins, D.E. McAllister and J.R. Stauffer, Jr. 1980. *Atlas of North American Freshwater Fishes.* North Carolina State Museum of Natural History, Raleigh, NC. 854 pp.

Lenat, D.R. 1988. Water quality assessment of streams using a qualitative collection method for benthic macroinvertebrates. *J. North Am. Benthol. Soc.* 7:222-233.

Leopold, L.B., G.M. Wolman, and J.P. Miller. 1964. *Fluvial Processes in Geomorphology.* W.H. Freeman and Co., San Francisco, CA. 522 pp.

Lind, O.T. 1974. *Handbook of Common Methods in Limnology.* Mosby. Saint Louis, MO. 154 pp.

Linthurst, R.A., D.H. Landers, J.M. Eilers, D.F. Brakke, W.S. Overton, E.P. Meier, and R.E. Crowe. 1986. *Characteristics of Lakes in the Eastern United States. Volume I: Population Descriptions and Physico-chemical Relationships.* EPA-600/4-86/007a. U.S. Environmental Protection Agency, Washington, D.C.

Loftis, J.C., R.C. Ward, R.D. Phillips, and C.H. Taylor. 1989. *An Evaluation of Trend Detection Techniques for Use in Water Quality Monitoring Programs.* EPA/600/3-89/037. U.S. Environmental Protection Agency, Corvallis, OR.

Lowe, R.L. 1974. *Environmental Requirements and Pollution Tolerance of Freshwater Diatoms.* EPA/670/4-74/005. U.S. Environmental Protection Agency, Cincinnati, OH. 333 pp.

Lowe, T.P., T.W. May, W.G. Brumbaugh, and D.A. Kane. 1985. *National Contaminant Biomonitoring Program: Concentrations of Seven Elements in Freshwater Fish, 1978-1981.*

Madow, W.G. 1949. On the theory of systematic sampling II. *Ann. Math. Stat.* 20:333-354.

Matthews, W.J. 1986. Fish faunal structure in an Ozark stream: stability, persistence, and a catastrophic flood. *Copeia* 1986:388-397.

Matthews, W.J. 1990. Spatial and temporal variation in fishes of riffle habitats: a comparison of analytical approaches for the Roanoke River. *Am. Midl. Nat.* 124:31-45.

Messer, J.J., C.W. Ariss, J.R. Baker, S.K. Drousé, K.N. Eshleman, P.R. Kaufmann, R.A. Linthurst, J.M. Omernik, W.S. Overton, M.J. Sale, R.D. Schonbrod, S.M. Stambaugh, and J.R. Tuschall, Jr. 1986. *National Stream Survey Phase I - Pilot Survey.* EPA/600/4-86/026. U.S. Environmental Protection Agency, Washington, D.C. 179 pp.

Messer, J.J., C.W. Arris, J.R. Baker, S.K. Drousé, K.N. Eshleman, A.J. Kinney, W.S. Overton, M.J. Sale, and R.D. Schonbrod. 1988. Stream chemistry in the Southern Blue Ridge: feasibility of a regional synoptic sampling approach. *Water Resour. Bull.* 24:821-829.

Metcalf, J.L. 1989. Biological water quality assessments of running waters based on macroinvertebrate communities: history and present status in Europe. *Environ. Pollut.* 60:101-139.

Miller, D.L., P.M. Leonard, R.M. Hughes, J.R. Karr, P.B. Moyle, L.H. Schrader, B.A. Thompson, R.A. Daniels, K.D. Fausch, G.A. Fitzhugh, J.R. Gammon, D.B. Halliwell, P.L. Angermeier, and D.J. Orth. 1988. Regional applications of an index of biotic integrity for use in water resource management. *Fisheries* 13:12-20.

Miller, R.R., J.D. Williams, and J.E. Williams. 1989. Extinctions of North American fishes during the past century. *Fisheries* 14:22-38.

Morgan, R.P., and A.J. Janicki. In press. *Evaluation of Relationships between Sensitivity to Acidification, Watershed Characteristics, and Fish Distributions in Western Maryland.* Maryland Department of Natural Resources, Annapolis, MD.

Morris, F.A., D.V. Peck, M.B. Bonoff, K.J. Cabble, and S.L. Pierett. 1986. *National Surface Water Survey, Eastern Lake Survey (Phase I - Synpotic Chemistry) Field Operations Report.* EPA 600/4-86/010. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Las Vegas, NV. 46 pp.

Moyer, S. 1986. *Aquatic Contaminants: A Threat to the Sport Fishing Industry.* Sport Fishing Institute, Washington, D.C. 165 pp.

Murdoch, P.S. 1986. *Chemical Input-Output Budgets and Stream Chemistry Dynamics during a Two-year Period in Biscuit Brook, Catskill Mountains, New York.* U.S. Geological Survey Water Resources Investigations Report 86 (Draft). Prepared in cooperation with U.S. Environmental Protection Agency.

Nash, R.R. 1989. *The Rights of Nature: A History of Environmental Ethics.* University of Wisconsin Press, Madison, WI. 290 pp.

National Research Council. 1990. *A Review of the USGS National Water Quality Assessment Pilot Program.* National Academy Press, Washington, D.C.

Nebeker, A.V., M.A. Cairns, J.H. Gakstatter, K.W. Malueg, G.S. Schuytema, and D.F. Krawczyk. 1984. Biological methods for determining toxicity of contaminated freshwater sediments to invertebrates. *Environ. Toxicol. Chem.* 3:617-630.

Nelson, M.K., C.G. Ingersoll, and F.J. Dwyer. 1989. *New Standard Guide for Conducting Solid-phase Toxicity Tests with Freshwater Invertebrates.* Draft #5. ASTM Committee E47 on Biological Effects and Environmental Fate, Subcommittee E47.03 on Sediment Toxicity. American Society for Testing and Materials, Philadelphia, PA.

Nix, J.F., K.W. Thornton, D.E. Ford, and J.T. Malcolm. 1986. *Storm Event Sampling of Southwestern Arkansas Streams Susceptible to Acid Precipitation.* U.S. Environmental Protection Agency Cooperative Agreement 811863-01-0 (Draft).

Northcote, T.G. 1988. Fish in the structure and function of freshwater ecosystems: A "top-down" view. *Can. J. Aquat. Sci.* 45:361-379.

Noss, R.F. 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conserv. Biol.* 4:355-365.

Oberdorff, T., and R.M. Hughes. In press. Modification and application of the Index of Biotic Integrity to the Seine River, France. *Hydrobiologia*

Ohio EPA. 1988. *Biological Criteria for the Protection of Aquatic Life.* Division of Water Quality Monitoring and Assessment, Columbus, OH. 384 pp.

Olem, H. 1986. *Episodic Changes in Stream Water Quality in Five Watersheds in the Southern Blue Ridge Province.* Interim Report. Chattanooga, TN.

Omernik, J.M. 1977. *Nonpoint Source Stream Nutrient Level Relationships: A Nationwide Study.* EPA/600/3077/105. U.S. EPA Environmental Research Laboratory, Corvallis, OR.

Overton, W.S. 1985. *Analysis Plan for the Eastern Lake Survey.* Working draft, March 1985. Technical Report 113. Department of Statistics, Oregon State University, Corvallis, OR.

Overton, W.S. 1987. *Phase II Analysis Plan, National Lake Survey.* Working draft, April 1987. Technical Report 115. Department of Statistics, Oregon State University, Corvallis, OR.

Overton, W.S. 1989. *Effects of Measurement and Other Extraneous Errors on Estimated Distribution Functions in the National Surface Water Surveys.* Department of Statistics Technical Report 129. Oregon State University, Corvallis, OR.

Overton, W.S., and S.V. Stehman. 1987. *An Empirical Investigation of Sampling and Other Errors in the National Stream Survey: Analysis of a Replicated Sample of Streams.* Technical Report 119. Department of Statistics, Oregon State University, Corvallis, OR.

Overton, W.S., D. White, and D.L. Stevens. 1990. *Design Report for EMAP. Environmental Monitoring and Assessment Program Part I.* Draft. Department of Statistics, Oregon State University, Corvallis, OR.

Peck, D.V., J.M. Engels, K.M. Howe, and J.E. Pollard. 1988. *Aquatic Effects Research Program: Episodic Response Project Integrated Quality Assurance Plan.* Revision 1. EPA 600/X-88/274. U.S. Environmental Protection Agency, Las Vegas, NV.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. *Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish.* EPA 444/4-89/001. U.S. Environmental Protection Agency, Washington, D.C. 172 pp.

Platts, W.S. 1979. Livestock grazing and riparian/stream ecosystems - an overview. Pages 39-45 in O.B. Cope, ed. *Proceedings of the Forum - Grazing and Riparian/Stream Ecosystems.* Trout Unlimited, Denver, CO.

Platts, W.S., W.F. Megahan, and G.W. Minshall. 1983. *Methods for Evaluating Stream, Riparian, and Biotic Conditions.* General Technical Report INT-138. U.S. Department of Agriculture, U.S. Forest Service, Ogden, UT.

Rahel, F.J. In press. The hierarchical nature of community persistence: a problem of scale. *Am. Nat.*

Rapport, D.J. 1989. What constitutes ecosystem health? *Perspect. Biol. Med.* 33:120-132.

Robbins, C.S., D. Bystrak, and P.H. Geissler. 1986. *The Breeding Bird Survey: Its First Fifteen Years, 1965-1979.* Resource Publication 157. United States Department of the Interior, Fish and Wildlife Service, Washington, D.C. 196 pp.

Ross, S.T., W.J. Matthews, and A.A. Echelle. 1985. Persistence of stream fish assemblages: effects of environmental change. *Am. Nat.* 126:24-40.

Ruttner, F. 1963. *Fundamentals of Limnology.* Translated by D.G. Frey and F.E.J. Frey. University of Toronto Press, Toronto, Ontario. 295 pp.

Schaeffer, D.J., E.E. Herricks, and H.W. Kerster. 1988. Ecosystem health: 1. Measuring ecosystem health. *Environ. Manage.* 12:445-455.

Schindler, D.W. 1987. Detecting ecosystem responses to anthropogenic stress. *Can. J. Fish. Aquat. Sci.* 44:6-25.

Schmitt, C.J., and W.G. Brumbaugh. 1990. National Contaminant Biomonitoring Program: concentrations of arsenic, cadmium, copper, lead, mercury, selenium, and zinc in U.S. freshwater fish, 1976-1984. *Arch. Environ. Contam. Toxicol.* 19:731-747.

Schmitt, C.J., J.L. Zajicek, and M.A. Ribick. 1985. National pesticide monitoring program: Residues of organochlorine chemicals in freshwater fish, 1980-81. *Arch. Environ. Contam. Toxicol.* 14:225-260.

Schmitt, C.J., J.L. Zajicek, and P.H. Peterman. 1990. National Contaminant Biomonitoring Program: residues of organochlorine chemicals in U.S. freshwater fish, 1976-1984. *Arch. Environ. Contam. Toxicol.* 19:748-782.

Shapiro, J. 1979. The current status of lake trophic indices: a review. Pages 53-99 in T.E. Maloney, ed. *Lake and Reservoir Classification Systems.* EPA/600/3-79/074. U.S. Environmental Protection Agency, Washington, D.C.

Silverstein, M.E., M.L. Faber, S.K. Drousé, and T.E. Mitchell-Hall. 1987. *Western Lake Survey-Phase I: Quality Assurance Report.* EPA/600/4-87/037. U.S. Environmental Protection Agency, Las Vegas, NV.

Simpson, P.W., J.R. Newman, M.A. Keirn, R.M. Matter, and P.A. Guthrie. 1982. *Manual of Stream Channelization Impacts on Fish and Wildlife.* FWS/OBS-82/24. U.S. Fish and Wildlife Service, Washington, D.C.

Smart, M.M., J.R. Jones, and J.L. Sebaugh. 1985. Stream-watershed relations in the Missouri Ozark Plateau Province. *J. Environ. Qual.* 14:77-82.

Smith, F., S. Kulkarni, L.E. Myers, and M.J. Messner. 1988. Evaluating and presenting quality assurance sampling data. Pages 157-168 in L.H. Keith, ed. *Principles of Environmental Sampling.* American Chemical Society, Washington, D.C.

Smith, R.A., R.B. Alexander, M. G. Wolman. 1987. Water-quality trends in the nation's rivers. *Science* 235:1607-1615.

Snedecor, G.W., and W.G. Cochran. 1967. *Statistical Methods.* Sixth Edition. The Iowa State University Press, Ames, IA. 507 pp.

Stanley, T.W., and S.S Verner. 1985. The U.S. Environmental Protection Agency's quality assurance program. Pages 12-19 in J.K. Taylor and T.W. Stanley, eds. *Quality Assurance for Environmental Measurements.* ASTM STP 867. American Society for Testing and Materials, Philadelphia, PA.

Steedman, R.J. 1988. Modification and assessment of an index of biotic integrity to quantify stream quality in southern Ontario. *Can. J. Fish. Aquat. Sci.* 45:492-501.

Stehman, S.V., and W.S. Overton. 1987a. Estimating the variance of the Horvitz-Thompson estimator in variable probability, systematic samples. Proceedings of the Section on Survey Research Methods of the American Statistical Association.

Stehman, S.V., and W.S. Overton. 1987b. *An Empirical Investigation of the Variance Estimation Methodology Prescribed for the National Stream Survey: Simulated Sampling from Stream Data Sets.* Technical Report 118. Department of Statistics, Oregon State University, Corvallis, OR.

Stevens, D.L., W.S. Overton, and A. Herstrom. 1987. Spatial representation of National Lake Survey data. Manuscript.

Stoddard, J. 1990. *Environmental Monitoring and Assessment Program: Plan for Converting the NAPAP Aquatic Effects Long-Term Monitoring (LTM) Project to the Temporally Integrated Monitoring of Ecosystems (TIME) Project.* Internal Report. U.S. EPA Environmental Research Laboratory, Corvallis, OR.

Stoermer, E.F., J.A. Wolin, C.L. Schelske, and D.J. Conley. 1985. An assessment of ecological changes during the recent history of Lake Ontario based on siliceous algal microfossils preserved in the sediments. *J. Phycol.* 21:257-276.

Strahler, A.N. 1957. Quantitative analysis of watershed geomorphology. *Trans. Am. Geophys. Union* 38:913-920.

Study Group on Environmental Monitoring. 1977. *Environmental Monitoring.* National Research Council, National Academy of Sciences. Washington, D.C. 173 pp.

Suter, G.W. 1990. Endpoints for regional ecological risk assessments. *Environ. Manage.* 14:9-23.

Taylor, J.K. 1988. *Quality Assurance of Chemical Measurements.* Lewis Publishers, Chelsea, MI. 328 pp.

Thornton, K.W., B.L. Kimmel, and F.E. Payne. 1990. *Reservoir Limnology. Ecological Perspectives.* John Wiley and Sons, New York. 256 pp.

Tonn, W.M. 1990. Climate change and fish communities: a conceptual framework. *Trans. Am. Fish. Soc.* 119:337-352.

Tonn, W.M., and J.J. Magnuson. 1982. Patterns in the species composition and richness of fish assemblages in northern Wisconsin lakes. *Ecology* 63:1149-1166.

Uutala, A.J. 1990. Chaoborus (Diptera:Chaoboridae) mandibles - paleolimnological indicators of the historical status of fish populations in acid-sensitive lakes. *J. Paleolimnol.* 4:in press.

U.S. Department of Interior. 1989. News release. July 17. U.S. Department of the Interior, Washington, D.C.

U.S. EPA. 1984a. *Monitoring Strategy.* Office of Water, Washington, D.C. 132 pp.

_____. 1984b. *Waterbody Surveys and Assessments for Conducting Use Attainability Analyses. Volume III: Lake Systems.* Technical Support Manual. Office of Water Regulations and Standards, Washington, D.C.

_____. 1987a. *Surface Water Monitoring: A Framework for Change.* Office of Water and Office of Policy, Planning, and Evaluation, Washington, D.C. 57 pp.

_____. 1987b. *Unfinished Business: A Comparative Assessment of Environmental Problems.* Office of Policy Analysis and Office of Policy, Planning, and Evaluation, Washington, D.C. 100 pp.

_____. 1987c. *National Water Quality Inventory.* 1986 report to Congress. EPA/440/4-87/008. Office of Water, Washington, D.C. 184 pp.

_____. 1988. *Report of the National Workshop on Instream Biological Monitoring and Criteria.* Office of Water Regulations and Standards, Washington, D.C. 34 pp.

_____. 1989a. *Nonpoint Sources Agenda for the Future.* Office of Water, Washington, D.C. 31 pp.

_____. 1989b. *Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms.* EPA/600/4-89/001. U.S. EPA Environmental Monitoring Systems Laboratory, Cincinnati, OH.

_____, Region IV. 1989. *Region IV Workshop on Biomonitoring and Biocriteria.* Water Quality Management Branch and Ecological Support Branch, Atlanta, GA. 26 pp.

_____. 1990a. *Environmental Monitoring and Assessment Program: Guidelines for Preparing Logistics Plans.* EPA/600/x-90/161. U.S. Environmental Protection Agency, Environmental Monitoring Systems Laboratory, Las Vegas, NV. 31 pp.

_____. 1990b. *Environmental Monitoring and Assessment Program: Quality Assurance Program Plan.* Revision 1. U.S. Environmental Protection Agency, Cincinnati, Ohio.

_____. 1990c. *Volunteer Water Monitoring: A Guide for State Managers.* EPA/440/4-90/010. Office of Water, Washington, D.C. 78 pp.

_____. 1990d. *National Water Quality Inventory.* 1988 Report to Congress. EPA/440/4-90/003. Office of Water, Washington, D.C. 187 pp.

_____. 1990e. *Reducing Risk: Setting Priorities and Strategies for Environmental Protection.* SAB-EC-90-021. U.S. Environmental Protection Agency, Science Advisory Board, Washington, D.C.

U.S. General Accounting Office (GAO). 1981. *Better Monitoring Techniques are Needed to Assess the Quality of Rivers and Streams.* Vol. I. CED-81-30. Washington, D.C.

_____. 1988. *Protecting Human Health and the Environment through Improved Management.* GAO/RCED-88-101. Washington, D.C. 246 pp.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing. 1980. The river continuum concept. *Can. J. Fish. Aquat. Sci.* 37:130-137.

Walker, I.R. 1987. Chironomidae (Diptera) in paleoecology. *Quat. Sci. Rev.* 6:29-40.

Walker, W.W., Jr. 1984. Trophic state indices in reservoirs. Pages 435-440 in *Lake and Reservoir Management.* EPA/440/5-84/001. U.S. Environmental Protection Agency, Washington, D.C.

Warren, C.E. 1974. *Biology and Water Pollution Control.* W.B. Saunders, Philadelphia, PA.

Warren, C.E., W.J. Liss, and M.J. Wevers. 1987. *Artificial Streams for Ecosystem Toxicity Studies.* EA-5079. Electric Power Research Institute, Palo Alto, CA.

Warren-Hicks, W., B.R. Parkhurst, and S.S. Baker, Jr., eds. 1989. *Ecological Assessment of Hazardous Waste Sites: A Field and Laboratory Reference Document.* EPA/600/3-89/013. U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C. 251 pp plus appendix.

Williams, J.E., J.E. Johnson, D.A. Hendrickson, S. Contreras-Balderas, J.D. Williams, M. Navarro-Mendoza, D.A. McAllister, and J.E. Deacon. 1989. Fishes of North America endangered, threatened, or of special concern: 1989. *Fisheries* 14:2-20.

Witt, E.C., and J.L. Barker. 1986. *Stream Chemistry Response during Episodes of Acidic Rainfall and Snowmelt Runoff on Laurel Hill, Somerset County, Pennsylvania, November 1983 to July 1985.* U.S. Geological Survey Open File Report 86 (Draft). Prepared in cooperation with U.S. Environmental Protection Agency.

World Resources Institute. 1990. *World Resources 1990-91: A Report by the World Resources Institute.* Oxford University Press, New York.

Yosie, R.F. 1987. EPA's risk assessment culture. *Environ. Sci. Technol.* 21(6):526-531.

☆U.S. GOVERNMENT PRINTING OFFICE: 1991 -548-187/25642

United States
Environmental Protection
Agency

Center for Environmental Research
Information
Cincinnati OH 45268-1072

BULK RATE
POSTAGE & FEES PAID
EPA
PERMIT No. G-35

Official Business
Penalty for Private Use, $300

Please make all necessary changes on the above label, detach or copy, and return to the address in the upper left-hand corner.

If you do not wish to receive these reports CHECK HERE ☐; detach, or copy this cover, and return to the address in the upper left-hand corner.

EPA/600/3-91/022