United States
Environmental Protection
Agency

Office of
Research and Development
Washington, DC 20460

EPA/620/R-93/003
February 1993

Surface Waters 1991 Pilot Report

Environmental Monitoring and
Assessment Program

EMAP-Surface Waters 1991 Pilot Report

EDITORS:

David P. Larsen
U.S. EPA Environmental Research Laboratory
200 SW 35th Street
Corvallis, OR 97333

Susan J. Christie
ManTech Environmental Technology, Inc.
U.S. EPA Environmental Research Laboratory
200 SW 35th Street
Corvallis, OR 97333

AUTHORS AND CONTRIBUTORS:

Chapter 1: D.P. Larsen[1] and S.G. Paulsen[2]

Chapter 2: R.M. Hughes[3], C. Burch Johnson[3], S.S. Dixit[4], A.T. Herlihy[5], P.R. Kaufmann[5], W.L. Kinney[6], D.P. Larsen[1], P.A. Lewis[7], D.M. McMullen[8], A.K. Moors[9], R.J. O'Connor[9], S.G. Paulsen[2], R.S. Stemberger[10], S.A. Thiele[3], T.R. Whittier[3], and D.L. Kugler[3]

Chapter 3: D.P. Larsen[1], K.W. Thornton[11], N.S. Urquhart[5], S.G. Paulsen[2]

Chapter 4: D.P. Larsen[1] and N.S. Urquhart[5]

Chapter 5: S.G. Paulsen[2], N.S. Urquhart[5], D.P. Larsen[1]

Chapter 6: J.R. Baker[12] and N. Tallent-Halsell[12]

---

1 U.S. EPA Environmental Research Laboratory, 200 SW 35th Street, Corvallis, OR 97333.

2 University of Las Vegas, c/o U.S. EPA Environmental Research Laboratory, 200 SW 35th Street, Corvallis, OR 97333.

3 ManTech Environmental Technology, Inc., c/o U.S. EPA Environmental Research Laboratory, 200 SW 35th Street, Corvallis, OR 97333.

4 Department of Biology, Queen's University Kingston, Ontario K7L 3N6, Canada.

5 Oregon State University, c/o U.S. EPA Environmental Research Laboratory, 200 SW 35th Street, Corvallis, OR 97333.

6 U.S. EPA Environmental Monitoring Systems Laboratory, Las Vegas, NV 89193-3478.

7 U.S. EPA Environmental Monitoring Systems Laboratory, 3411 Church Street, Newtown, OH 45244.

8 Technology Applications, Inc., c/o U S. EPA Environmental Monitoring Systems Laboratory, 3411 Church Street, Newton, OH 45244.

9 Department of Wildlife, University of Maine, 238 Nutting Hall, Orono, ME 04469-0125.

10 Biology Department, Room 101, Gilman Hall, Dartmouth College, Hanover, NH 03755.

11 FTN Associates, 3 Innwood Circle, Suite 220, Little Rock AR 72211.

12 Lockheed Engineering and Sciences Co., 980 Kelly Johnson Drive, Las Vegas, NV 89119.

EPA/620/R-93/003
February 1993

EMAP-Surface Waters 1991 Pilot Report

David P. Larsen and Susan J. Christie, Editors

February 1993

ACKNOWLEDGEMENTS

In any undertaking with a scope as broad as that of EMAP-Surface Waters, numerous people contribute in important ways not reflected by authorship on documents such as this. We hesitate attempting to list all those who contributed, because we are sure to miss some. Instead, we will identify the organizations whose staff generated important material included in this report:

- EMAP-Surface Waters and associated laboratory staff in Corvallis, Las Vegas, and Cincinnati, including EPA and on-site contractor personnel (ManTech Environmental Technology, Inc., Lockheed Engineering and Sciences Co., and Technology Applications, Inc.), and guest workers from Oregon State University and University of Nevada at Las Vegas.

- Environmental Services Division of EPA Regions I and II.

- Personnel on cooperative agreements with Queens University, Dartmouth College, and University of Maine.

- Members of the lake sampling crews of miscellaneous origin

- Members of the peer review panel.

We especially appreciate the members of our sampling crews for their diligent efforts in obtaining data of outstanding quality. We also thank members of our peer review panel for their initial evaluation of the results contained in this report. The following people provided official technical reviews of this report: Steve Bartell, Charles Goldman, Steve Hedtke and Stephen Lozano, Thomas LaPoint, Eugene Stoermer, Deborah Coffey, and Don King. Many others provided informal but important review comments.

Financial support from U.S. EPA to the following institutions led to much of the work reported here:

- Oregon State University (CR818606 and CR816721)
- Queens University (CR818707)
- Dartmouth College (CR819689-01-0)
- University of Maine (CR819659 and CR818179)
- ManTech Environmental Technology, Inc. (68-C8-0006)
- Lockheed Engineering and Sciences Co. (68-C0-0049)
- Aquatech (1B0517NTSA)

The research described in this document has been funded by the U.S. Environmental Protection Agency. This document has been subjected to the Agency's peer and administrative review and approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

ANC	—	acid neutralizing capacity
ANSP	—	Academy of Natural Sciences at Philadelphia
APHA	—	American Public Health Association
AVHRR	—	advanced very high resolution radiometry (from satellite)
BI	—	biotic index
CCA	—	canonical correspondence analysis
CDF	—	cumulative distribution function
Chl-a	—	chlorophyll-*a*
CV	—	coefficient of variation
DCA	—	detrended correspondence analysis
DIC	—	dissolved inorganic carbon
DLG	—	digital line graphs
DO	—	dissolved oxygen
DOC	—	dissolved organic carbon
EMAP	—	Environmental Monitoring and Assessment Program
EMAP-SW	—	EMAP-Surface Waters
EPA	—	Environmental Protection Agency
FIA	—	Forest Inventory Analysis, performed by U.S. Forest Service
FWS	-	Fish and Wildlife Service
GLM	—	General Linear Model (SAS procedure)
GPS	—	Global Positioning System
ha	—	hectare
km	—	kilometer
m	—	meter
mL	—	milliliter
NALMS	—	North American Lake Management Society
NASS	—	National Agricultural Statistics Survey
NES	—	National Eutrophication Survey
OEPA	—	Ohio Environmental Protection Agency
PCA	—	principal components analysis
QA/QC	—	quality assurance/quality control
REDOX	—	reduction-oxidation
RF3	—	U.S. EPA River Reach File, Version 3
SAS	—	Statistical Analysis System

SD	—	Secchi Disk (transparency)
SE	—	standard error
TIME	—	Temporally Integrated Monitoring of Ecosystems
TP	—	total phosphorus
TSI	—	Trophic State Index
USDA	—	U.S. Department of Agriculture
USGS	—	U.S. Geological Survey

SYMBOLS

mg/L	—	milligrams per liter
°C	—	degrees Celsius
μeq/L	—	microequivalents per liter
μg/L	—	micro-grams per liter
σ^2_{index}	—	index variance
σ^2_{lake}	—	population variance
$\sigma^2_{lake \star year}$	—	lake-year interaction effects
σ^2_{meas}	—	measurement variance
σ^2_{res}	—	residual variance
σ^2_{year}	—	year variance

CHAPTER 1
INTRODUCTION

1.1 OVERVIEW OF THE ROLE OF PILOT STUDIES

In 1989, the U.S. Environmental Protection Agency (EPA) initiated the Environmental Monitoring and Assessment Program (EMAP) to provide improved information on the current status of, and long-term trends in, the condition of the nation's major ecological resources, including inland surface waters. Before full-scale implementation of EMAP-Surface Waters (EMAP-SW), we anticipate conducting a series of lake monitoring pilot studies, followed by demonstration studies, to show whether or not EMAP can produce the kinds of results intended.

It is often difficult to say exactly what constitutes a pilot study, except to describe it as the effort expended to answer specific questions. Pilot studies can be specific field studies, desktop analyses, or both. They can be conducted on a regional scale, using the EMAP probability-based design as the basis for lake selection, or they can occur at specially selected sites, depending on the nature of the questions to be addressed or the lake populations of interest. Pilot studies sometimes address questions that can be answered without conducting field studies; for example, they can evaluate the results of other studies or include workshops that seek to answer particular questions.

The demonstration studies that follow the pilot studies are designed to produce regional-scale estimates similar to the estimates expected from EMAP. In some cases, demonstration studies can answer relevant questions just as well as pilot studies; in fact, a regional-scale implementation study may be the only way to answer some questions.

The FY 1991 pilot studies included elements of all three types of preliminary study (regional surveys, special field studies, and desktop analyses). The fundamental role of a pilot study is to focus on critical questions, the answers to which are necessary for effective implementation of EMAP monitoring. We began by asking, "What critical pieces prevent us from implementing regional or national monitoring?" or "Why can't we presently initiate an EMAP-type monitoring program?" or "What will be the consequences of proceeding with monitoring without an answer to this set of questions?" The *EMAP-Surface Waters Northeast Lakes Pilot Implementation Plan* posed a set of questions to be addressed in a series of pilot studies (Pollard and Peres, 1991). The questions, recast here to set the framework for this pilot report, focus on lake monitoring but they apply equally well to stream monitoring.

A. One basic set of questions focuses on what we will monitor and includes the following:

   1. What indicators will we use as part of the basic monitoring program? What is a sufficient number of indicators on which to base full-scale monitoring? In this report, we limit the discussion primarily to selection of those indicators that will be used to address the primary objectives of EMAP, the response indicators (see Chapter 2).

   2. Where, when, and how will we measure the indicators? This question has two parts. The first is the design-based, probability selection of the lakes and streams to be sampled: How will we select the lakes for field visitation? (Survey Design Evaluation; see Chapter 3.) The second is: Once the lakes have been selected, how and when will we sample them? (Index Characterization; see Chapter 2.)

   3. How will we use the indicators to make statements about condition, associations, and probable cause of impaired and unimpaired conditions? (See Chapter 5.)

B. Indicator variability at various spatial and temporal scales affects our ability to characterize status (condition) and detect trends in ecological condition. These topics are covered in Chapter 4 of the report:

   1. What are the important components of variance needing characterization?

   2. How does the magnitude of these components affect our ability to describe status and to detect trends?

   3. What choices do we have to reduce variance (e.g., through methodological improvements), to minimize its effects (e.g., through efficient design modifications), or to *remove its effects (e.g., through mathematical manipulation)?*

C. Implementing a regional-scale monitoring program such as EMAP requires extensive logistical planning and testing. These issues are addressed in Chapter 6. Basic questions include:

   1. Can we assemble and deploy sampling teams to obtain lake and stream samples within the selected temporal index window?

   2. Can we train different teams well enough to keep team-to-team differences minor and not compromise program objectives?

   3. Can we coordinate the work of multiple field crews well enough for smooth, consistent sample collection and field recording? Can we ship and track samples to appropriate analytical facilities or data processors in a timely manner and without loss of samples?

1.2 FY 1991 PILOT STUDIES

During the summer of 1991, EMAP-SW conducted its first series of field pilot studies, with lakes as the resource of interest and the northeastern United States (EPA Regions I and II) as the area of interest. A basic criterion for the pilot studies was that we would use the EMAP grid design and probability methods to select a set of lakes for sampling to answer the questions outlined in

Section 1.1, and that we would select other lakes only if we were unable to answer the questions by sampling the probability lakes. For some indicators, a regional-scale pilot study would serve to answer some questions. For other indicators, however, this was not so. In particular, it was evident that we were not prepared to obtain index samples of the fish assemblage, the riparian birds, the benthos assemblage, and the physical habitat within reasonable budget constraints. Therefore, we selected a special set of lakes for use in developing methods to index lakes for these attributes.

We also faced the need to begin data collection to answer some very specific questions raised by reauthorization of the Clean Air Act, which mandates significant reduction of sulfate emissions (10 million metric tons) and nitrate emissions (2 million metric tons) over the next 10 years (Public Law 101-549, Title IV, Section 401b). The Clean Air Act also mandates an assessment of the effect of this reduction on aquatic ecosystems: could we detect lake and stream responses to this magnitude of reduction in sulfate emissions after 10 years? To some extent, the National Surface Water Survey (Linthurst et al., 1986; Landers et al., 1987; Kaufmann et al., 1988), including the Long-Term Monitoring Project (Newell and Hjort, 1991), were pilot studies that evaluated the ability of design-based probability surveys to collect and interpret limnological information addressing critical policy questions. We felt confident that a regional-scale survey could be implemented to address the mandates of the Clean Air Act. In addition, through regional-scale pilot studies, we could begin data collection for some indicators (e.g., indicators of the trophic condition of lakes and lake chemical characteristics) to answer questions about variance and logistics not answerable any other way.

Therefore, the field-scale pilot studies had two parts. One part was designed to develop indexing protocols for indicators with which we were not confident we could index lakes. We selected a set of lakes for this part of the pilot activity and called them *indicator lakes*. Chapter 2 focuses on questions about indicators. The second part of the pilot activity focused on logistics and variance questions about those indicators for which we were confident we could obtain index samples (trophic condition and chemical characteristics). In this part, we also began data collection on indicators in order to address the Clean Air Act issues. We called the set of lakes selected for this purpose the *probability lakes*, because we used the EMAP grid design and probability methods to select lakes for field visitation.

In addition, a significant amount of thought has gone into refining the conceptual framework on which EMAP is founded. We further explored the concepts of *ecological values*, *response indica-*

tors, *index samples*, *description of status*, and *nominal-subnominal*. A major part of the 1991 pilot addressed questions related to indicator development, particularly for the societal value of biological integrity. Furthermore, we have been identifying and evaluating the influence of important components of variation on our ability to describe status and detect trends in condition. These refinements will guide us in selecting available databases for a preview of EMAP capabilities.

1.3 SCOPE

This report is organized as follows:

- Chapter 2 focuses primarily on the process of selecting response indicators for EMAP monitoring, drawing from numerous discussions on our conceptual framework for indicator selection and from results of the 1991 pilot surveys.

- Chapter 3 describes how we selected the probability lakes by using the general EMAP design guidelines and what we learned from that exercise. It covers questions related to the Tier 1 and Tier 2 selection process, nontarget and noninterest lakes, and the setting of inclusion probabilities (or expansion factors).

- Chapter 4 discusses the set of questions that address how well we think the proposed design will describe status and detect trends in condition. We describe and illustrate important components of variance with one set of indicators describing trophic condition. We use available data, along with data derived from the EMAP pilot, to show the framework for estimating the influence of variance. The results of this part of the pilot offer guidance on the types of databases most useful for describing variance for the other indicators.

- Chapter 5 briefly summarizes how annual statistical reports might be structured using data collected on some indicators from the probability lakes.

- Chapter 6 covers the logistics component of the pilot study.

This report does *not* cover the following:

- Regional-scale monitoring conducted to answer questions about the response of lakes to sulfate emissions reductions mandated by the Clean Air Act [the Temporally Integrated Monitoring of Ecosystems (TIME) component of EMAP].

- Habitat, exposure, and stressor indicators, except as covered briefly in Chapter 2. We are developing our diagnostic strategy.

- The process for identifying particular lake subpopulations, and those subpopulations at a national scale, on which EMAP-SW is likely to target reporting.

- Information management.

A sound quality assurance/quality control (QA/QC) program is fundamental to any monitoring program with the potential scope and duration of EMAP. A QA/QC plan has been developed that describes how the multitude of EPA and EMAP requirements will be addressed during the pilot, demonstration, and implementation phases of EMAP-SW (e.g., Peck, 1992). In addition, periodic reports (e.g., University of Maine, 1991) will focus on the QA/QC results of particular aspects of EMAP-SW activities. This pilot report covers some aspects of QA/QC, including how to obtain a representative, repeatable sample of various biological assemblages in lakes (Chapter 2); how to obtain a representative sample of lakes on which to make the measurements (Chapter 3); how to evaluate variance components and their influence on our ability to estimate status and detect trends (Chapter 4). These chapters lay the foundation for how we currently address, and how we expect to address, several fundamental aspects of assuring that data collected during EMAP-SW monitoring are of known quality. The future reports will address these topics in more detail.

Chapter 2 discusses an issue of interest, and occasionally of concern, to many people. The issue is how to set a criterion that can be used to evaluate whether the condition measured and expressed with indicators is acceptable or unacceptable for a particular lake or stream type within a given geographic location. Hunsaker and Carpenter (1990) and Messer et al. (1991) outline the issue. Messer et al. state: "Operational criteria must be developed for each response indicator to identify the transition from acceptable or desirable (nominal) to unacceptable or undesirable (subnominal) condition...Criteria could be based on attainable conditions under 'best management practices' as observed at regional reference sites..., or on theoretical grounds or management targets." Clearly, this is not strictly a technical issue. We emphasize that our objective in approaching this issue is not to make this decision about thresholds independent of others, but rather to contribute to the process of deciding how to resolve these issues from a technical perspective. We are not advocating any particular management decision or action or societal decision. We do believe, however, that it is important to develop as sound a scientific basis as possible for such decisions. We developed our discussions in Chapter 2 from this perspective.

CHAPTER 2
DEVELOPMENT OF LAKE CONDITION INDICATORS FOR EMAP—1991 PILOT

2.1 INTRODUCTION

The primary objective of EMAP is to estimate, on a regional scale and with known confidence, the status of, and changes and trends in, indicators of the condition of our nation's ecological resources. Given this objective, it is critical that we clearly describe the EMAP-Surface Waters (EMAP-SW) strategy for developing, evaluating, and selecting indicators of the condition of our lakes and streams. Chapter 2 focuses on issues pertaining to indicator development and the results of the 1991 pilot relative to these issues.

The selected indicators will be the foundation for the information presented to decision makers, the scientific community, and the public about the condition of our nation's lakes and streams. An effective indicator will inform decision makers and environmental managers, will be relevant to one of a host of policies that impact these resources, and will provide information upon which decision makers and managers might be willing to act (Hilborn, 1992; Ward, 1989). The indicators ultimately selected for EMAP-SW must be ecologically relevant and scientifically credible and must relate clearly to important biologically oriented characteristics of lakes and streams that are valued by the public.

What then is an indicator? There has been some confusion about this term. Many ecologists equate it with "indicator species," but we do not mean indicator species. ***We define an indicator as an ecological measurement, metric, or index that quantifies physical, chemical, or biological condition, habitat, or stress.*** Measurements conducted on a biological assemblage (e.g., fish, diatoms) must be converted into numerical metric or index scores that can be presented as a distribution of characteristics found within the population of lakes and streams of interest. These quantifiable forms, not the assemblage itself, constitute the indicators.

EMAP has identified four kinds of indicators: response, exposure, habitat, and stressor (Hunsaker et al., 1990; Messer, 1990; Paulsen et al., 1991). These categories are meant to be functional ones that describe the intended use of the indicator, not mutually exclusive pigeon holes for each measurement.

- *Response* indicators demonstrate the biological condition of the resource, at the organism, population, assemblage, or community level of organization. For surface waters, the preferable organization levels are the assemblage and community, and these should quantify the integrated response of the system to single, multiple, and cumulative stressors and relate to the ecological values of the waterbody held by society. The EPA now requires states to monitor response indicators in their biological criteria programs (U.S. EPA, 1990).

- *Exposure* indicators show the occurrence or magnitude of a response indicator's contact with a physical, chemical, or biological stressor. In surface waters, we focus on stressors that are usually of human origin. These indicators are most useful for diagnosing the probable causes of unacceptable condition and historically were the focus of monitoring by state and federal water monitoring agencies.

- *Habitat* indicators are defined as physical attributes necessary to support an organism, population, or community in the absence of pollutants (Hunsaker et al., 1990; Messer, 1990). However, this definition initially omitted chemical and biological habitat components. For surface waters, we consider habitat indicators as characterizing the typically natural chemical, physical, and biological conditions that support biological assemblages in the absence of anthropogenic stressors. They are subjects of study by many basic ecologists and by the U.S. Fish and Wildlife Service's HSI (Habitat Suitability Index; U.S. FWS, 1981) and IFIM (Instream Flow Incremental Methodology; Bovee, 1982) groups. At various scales of resolution, habitat indicators are also useful for classifying major lake and stream types (Brussock et al., 1985; Frissell et al., 1986; Lewis, 1983) and ecological or biogeographic regions (Bailey, 1976; Omernik, 1987).

- *Stressor* indicators were defined by Messer as characterizing natural processes, environmental hazards, or management actions that change exposure and habitat. We propose that stressor indicators are those that quantify management actions, inactions, and policies that ultimately change exposure.

All four types of indicators play a critical role in EMAP. However, because the major objective is to describe the status of and trends in indicators of condition, our primary effort, and the main focus of this chapter, is on the response indicators, as they will be used to describe the condition of our lakes and streams. We have chosen to focus on biological measurements as the foundation for the response indicators because we believe that they most effectively assess the cumulative effects of the many physical, chemical, and biological stressors to which we expose our aquatic resources. Before a response indicator can be effectively used within EMAP, we must (1) link it to one or more of the selected societal values of concern, (2) describe how a lake or stream will be effectively indexed for that indicator, (3) determine whether an indicator can be used to distinguish acceptable from unacceptable conditions, (4) determine whether natural variability and variability within the measurement process will allow us to adequately describe meaningful status or trends, and (5) identify a process for selecting among the number of candidate indicators currently available for use.

2.1.1 Societal Values and Their Response Indicators

In any monitoring program with a mandate as broad as that of EMAP, it is necessary to narrow the focus from the broad goals and objectives to the specific attributes that will form the basis of field monitoring and be suitable for making statements about the condition of resources. As a first step in narrowing the focus, we attempted to identify the fundamental reasons for society's concern about the condition of lakes from a *biological* perspective, which differs from a water quality perspective, such as drinkability, or an industrial use. The selection of these societal values drives the selection of appropriate indicators, and effective indicators must be clearly related to growing societal and scientific concerns over the condition of aquatic resources. After numerous discussions with managers of aquatic resources, individuals in decision making roles, and members of the scientific community, we felt that a reasonable place to start was with biological integrity, trophic condition, and fishability as the initial societal values. This set of values is not necessarily final or complete, but it represents a starting point. In addition, the Federal Water Pollution Control Act and its various amendments specifically mandate that these three attributes be protected, maintained or restored, and reported on periodically.

The societal value of greatest concern to EMAP-SW in the indicator pilot was biological integrity, because we knew least about how to select indicators for it. *Biological integrity can be defined as the ability to support and maintain a balanced, integrated, adaptive community with a biological diversity, composition, and functional organization comparable to those of natural lakes and streams of the region* (Frey, 1977; Karr and Dudley, 1981) *and includes various levels of biological, taxonomic, and ecological organization* (Noss, 1990). Waters in which composition, structure, and function have not been adversely impaired by human activities have biological integrity (Karr et al., 1986). Karr et al. (1986) also defined a system as healthy "when its inherent potential is realized...and minimal external support for management is needed." This value or ethic differs considerably from values oriented toward human use or pollution that are traditionally assessed in water quality and fisheries programs, in which production of a particular species of game fish is the goal (e.g., Doudoroff and Warren, 1957), and may contradict them (Callicott, 1991; Hughes and Noss, 1992; Pister, 1987). We focused our selection on candidate lake indicators that would aid us in assessing the various structural and functional aspects of biological integrity and in determining whether biological integrity was at an acceptable or unacceptable level.

We define trophic condition as the abundance or production of algae and macrophytes. Trophic condition is the focus of an entire section (314) of the Federal Water Pollution Control Act; the federal government and local lake associations spend millions of dollars annually to control the growth of algae and aquatic macrophytes. Trophic state involves both aesthetic (water clarity) and fundamental ecological (production of plant biomass) components. It is, therefore, a key aspect in determining both a lake's relative desirability to the public, its production of fish, and its ecological character or classification by limnologists (e.g., eutrophic or oligotrophic). The trophic state of a lake is largely a function of its geographic location and morphology, requiring us to consider both in assessments of whether or not the lake's condition is acceptable. Trophic condition is nested somewhat within the concept of biological integrity; we separate it here because it remains a primary concern for many lake managers. Like biological integrity, trophic condition must be interpreted relative to some expected condition for the region and waterbody class.

Fishability is defined as the catchability and edibility of fish by humans and wildlife. Fish represent a major human use of an aquatic ecosystem product, fishing is a multimillion dollar recreational industry, and fishing quality and fish edibility are major concerns of millions of anglers. Therefore, protecting fish is the goal of many water quality agencies, and fish drive their water quality standards. In addition, states have established agencies whose mission is to maintain and enhance fisheries. Piscivorous mammals (mink, otter) and birds (eagle, osprey, heron, egret, tern, loon, merganser), which are among our most desirable wildlife, are even more dependent on fish. When their fisheries are depleted or contaminated by toxics, they develop abnormalities, are extirpated locally, or become endangered regionally. Fishability may or may not be nested within the concept of biological integrity, in that we have many systems with catchable stocks of contaminant-free fish that have been managed at the expense of biological integrity. These systems might be considered in good condition with respect to fishability, but in poor condition from the perspective of biological integrity.

Before the pilot study, we concentrated on defining the values and attributes of concern and selecting candidate indicators that the scientific literature suggests are ready for use in EMAP or could be made ready with minimal effort. These candidates were the first to be evaluated. We will continue to invest in the research community for the development of new indicators of condition, but it was important to begin with those that appeared most ready to be evaluated.

The chief response indicators for trophic condition are chlorophyll-*a* and macrophyte coverage, with total phosphorus, total nitrogen, and Secchi disk transparency as important exposure and habitat indicators. Although macrophytes are poor indicators of water column trophic state, they do indicate lakewide trophic state, especially in shallow lakes. Chlorophyll-*a*, total phosphorus, total nitrogen, and Secchi disk transparency are relatively standard and well-accepted measures of trophic condition, thus our primary concern in the 1991 pilot was to evaluate survey logistics and the variances associated with the lake population and indexing process (see Chapters 4–6). Indexing methods for macrophytes were reserved for a later pilot.

In evaluating fishability, our major interests are in the catchability and edibility of fish; that is, are they present and free of contaminants (Appendix 2A). Our 1991 pilot focused on the catchability aspect, in other words, how to go about indexing a lake for its fish assemblage. In future years, edibility will be based on concentrations of chemicals in fish tissue.

Developing indicators of biological integrity poses a much greater challenge than either trophic condition or fishability, because biological integrity involves assessing several aspects of the entire community (see Section 2.4.2). We focused indicator development on biological assemblages. Using a combination of workshops, literature reviews, and extended discussions among EMAP-SW staff, members of various state and federal agencies, and the broader academic community, we reduced a long list of candidates down to those evaluated in the 1991 pilot: fish, zooplankton, sedimentary diatoms, macrobenthos, and littoral/riparian breeding birds.

- Fish were chosen primarily for their societal value, their relationship to the fishability value, and their role as a top consumer in lakes, and because they represent a long-lived assemblage in lakes and therefore reflect the effects of a variety of stressors.

- Sedimentary diatoms are one indicator of primary production in lakes, have served as an effective diagnostic indicator, and can be used to interpret earlier conditions in lakes.

- Zooplankton and benthic macroinvertebrates represent the primary consumer level in lakes and are intermediate lived organisms; they differ in their general habitat requirements.

- Riparian/littoral birds indicate the condition of a lake's riparian and littoral zone, are highly valued by society, and can serve as a top consumer in the absence of fish.

Our current position is that fish, zooplankton, and sediment diatoms form the probationary core assemblages to be used for demonstration monitoring. We will continue to examine additional pilot information to be derived from benthic macroinvertebrate and bird assemblages. The questions we sought to answer about these candidate biological indicators during the 1991 pilot and the results are examined in the remainder of Chapter 2. We focused primarily on (1) determining the sampling effort appropriate for each assemblage (how to obtain an index sample for each, Section 2.2), (2) developing candidate metrics representative of each assemblage (Section

2.4), and (3) evaluating the responsiveness of assemblages and metrics to a set of lake types and stressor gradients (Section 2.4).

2.1.2 Habitat, Exposure, and Stressor Indicators

Although this chapter focuses primarily on the development of response indicators, physical, chemical, and biological habitat play an important role in classifying lakes and in diagnosing probable cause of impairment or unacceptable condition. Therefore, habitat, exposure, and stressor indicators are also briefly discussed in this chapter. The primary objectives of the 1991 pilot were to quantify the necessary sampling (indexing) effort (Section 2.2) and to determine the appropriateness of the habitat, exposure, and stressor metrics via their association with biological responses (Section 2.4).

2.1.3 Indexing

Indexing is the consistent manner in which a waterbody is sampled, both spatially and temporally, and the way in which its condition is numerically represented. Indexing is fundamental to all ecological sampling, whether the objective is a site-specific assessment of the effects of a particular perturbation, or a regional assessment of a population of lakes and streams. Indexing is necessary in EMAP, as in most efforts, because we cannot measure all attributes in all parts of all waterbodies at all times. We must select an appropriate time period for sampling, the sample must adequately represent the waterbody's character, and the lakes and streams selected for sampling must be representative of the population of waters from which they are drawn. We can meet the last requirement through spatially balanced, probabilistic selection of lakes and streams. Selection of the attributes sampled requires indicator development and evaluation (Section 2.4). Selection of a sampling period and sampling locations involves several indexing considerations.

In the issue of indexing, the concepts of quality assurance (QA) also emerge. The role of QA is to insure and document the repeatability and accuracy of a measurement process. Typically, we think of this from the perspective of laboratory analysis, that is, the accuracy and precision of an analytical number given the techniques employed in the analysis. In the context of EMAP, we broaden this concept to incorporate the repeatability of the entire protocol, including the sampling location, the sampling period, and the actual measurement of the indicator, whether by field or laboratory analysis, identification, or enumeration.

Acceptable lake indexing methods had already been determined for some attributes and assemblages before the 1991 pilot. These (chlorophyll-*a*, water quality, sedimentary diatoms, zooplankton) were implemented on the probability lakes to answer questions about the logistics of conducting regional surveys (Chapter 6) and to begin collecting data on population and index variance (Chapter 4). This process will require several more years to evaluate important variance components in different regions of the country (Chapters 4 and 5). For other indicators, standard methods of obtaining an acceptable index sample were not available in the literature or in methodologies of the monitoring community. These indicators (macrobenthos, fish, riparian birds, physical habitat structure) required focused pilot studies to develop efficient indexing protocols appropriate for a single visit by a small crew. Section 2.2 describes the requirements for an appropriate index sample, explains the indexing questions we wished to answer in the pilot, and provides the results and current status of index sampling for each assemblage.

2.1.4 Indicator Evaluation

To ensure consistency among the various EMAP resource groups and comparable evaluation among candidate indicators within a group, EMAP has developed a set of indicator evaluation criteria (Olsen, 1992). We describe how these criteria, and a subset pertinent to evaluation of biological integrity indicators, were used to select and evaluate candidate assemblages (Sections 2.4.1 and 2.4.2). In Sections 2.4.4 and 2.4.5, we explain the process for developing biological integrity metrics from assemblage measurements and a likely process for culling candidate metrics and assemblages. Indicator selection is a particularly critical process because of the role indicators play in linking field measurements, waterbody condition, and ecological values. It is made difficult by the differing sensitivities, discriminating power, and potential for contradictory conclusions among any set of assemblages or metrics. We do not intend to select indicators based on one summer's data, but rather to evaluate the data and present a process for eventually choosing core indicators.

2.1.5 Determining Nominal/Subnominal Condition

Describing the range of conditions found and the trends in these conditions is EMAP's primary objective. However, we suggest that the information will be useful to a wider range of decision makers if we determine the proportion of the waterbodies that are in subnominal (unacceptable) and nominal (acceptable) condition. This is actually nothing new. As a society, we engage in this process routinely when we establish chemical criteria and standards or biological criteria and standards. We are not suggesting that EMAP do this alone, or that there is a single number either nationally or regionally that should be used. The numbers may differ by ecological regions or waterbody classes. We are suggesting, however, that EMAP enter into this activity in concert with state and federal agencies as well as with the scientific community, and develop rational methods for establishing these decision points.

Establishing acceptable or unacceptable condition requires some sort of benchmark or reference condition against which the sample waterbodies can be compared. Several approaches have been suggested for estimating reference condition: reference sites, pristine sites, historical data, paleoecological data, ecological models, empirical distributions, experimental results, and consensus of experts. Each has advantages and disadvantages, which we describe in Section 2.5; we applied a reference lake approach as a first step in assessing response indicators in this pilot (Sections 2.3 and 2.4).

2.1.6 Probability Lakes and Indicator Lakes

The 1991 pilot study consisted of lakes selected using two different protocols, depending on the state of readiness for a particular indicator. For indicators with suitable indexing protocols, we used the proposed probability design. For fish, birds, benthos, and habitat structure, we handpicked lakes from a range of types, sizes, and apparent watershed impact. At these 19 "indicator lakes," we developed plot sampling protocols. We also evaluated all assemblages together at these lakes to develop and assess metrics. With a few exceptions, this chapter focuses on the indicator development conducted at these handpicked indicator lakes. The selection process and the characteristics of these lakes are described in Section 2.3.

2.1.7 Chapter Layout

This introduction has briefly touched upon the issues necessary to move ahead with the selection of indicators to meet the first objective of EMAP. The results and progress of the 1991 pilot for each of these topics are presented in the remainder of Chapter 2, as follows:

- 2.2 Evaluation of Indexing
- 2.3 Selection of Indicator Lakes and their Physical and Chemical Characteristics
- 2.4 Indicator Development and Evaluation
- 2.5 Distinguishing Nominal from Subnominal Condition
- 2.6. Summary and Conclusions

2.2 EVALUATION OF INDEXING

In this section, we offer an overview of the considerations involved in index sampling for surveys of this nature, drawn primarily from Stevens (pers. comm.). Following the overview is a description of how we addressed several key indexing issues relevant to lake surveys.

2.2.1 The General Indexing Concept

Selection of a sampling period and sampling sites involves 10 basic aspects (Table 2-1; Stevens, pers. comm.). The first two involve precision or variance considerations, two are largely logistical concerns, and the remainder are ecological concerns.

Table 2-1. Characteristics of an Ideal Index Sample[a]

1.	Minimizes natural spatial and temporal gradients and periodicities.
2.	Allows a single or a composite sample.
3.	Has wide (regional) applicability.
4.	Provides sampling ease and efficiency.
5.	Distills essential aspects of the ecosystem or assemblage.
6.	Establishes relative condition of the system (is sensitive to stressors).
7.	Is responsive to stressors that occur at different seasons or in different places.
8.	Can be sampled during period of maximum anthropogenic perturbation.
9.	Has biological relevance (information rich).
10.	Is accurate.

[a] From Stevens, pers. comm.

2.2.1.1 Precision or Variance Considerations

(1) Indexing should minimize natural spatial and temporal gradients and periodicities within the lake or stream. In other words, sampling should occur in places and at times that are relatively

invariant, though not necessarily constant. If the system is changing rapidly for phenological or seasonal reasons (case 1) or if the habitat structure undergoes cyclical changes, as in the riffles and pools of streams (case 2), it will be difficult to distinguish anthropogenic effects from natural noise. In case 1, hydrographs and climatographs are examined for gradients during potential index periods; in case 2, littoral/profundal or riffle/pool samples are kept separate until data analysis.

(2) Relative natural variation in the index period and in macrohabitats must be considered to decide whether a single or composite sample is representative enough. If a single sample is likely to differ greatly from another, as is expected for some attributes along the littoral zone of many lakes or in many stream reaches, aggregates of multiple samples are advisable. If not, a single sample is adequate. Similarly, if the index period occurs in a highly variable and unpredictable season, as in the spring and fall in some regions, multiple sampling visits are required. However, EMAP is oriented to single-visit sampling. As a result, we must measure the variation that occurs during an index period and determine its influence on our ability to describe the status of and detect trends in indicators of the condition of lakes. Addressing this issue will be a basic part of the initial years of routine monitoring. We will revisit a subset of lakes (and streams when we begin monitoring streams) within the index period to estimate index variance relative to other variance components. See Chapter 4 for a detailed discussion of the variance components and their relative importance.

2.2.1.2 Logistical Aspects of Indexing

(3) The selected methods should be applicable across a multistate region, both to reduce logistical headaches and to increase the regional applicability of the data collected. We evaluate regional applicability by considering the physical and chemical character of the waterbodies and the distribution and abundance of the indicator assemblages. That is, are the major waterbody types, macrohabitats, and assemblages regionally comparable, or do the salinities, lake level fluctuations, lake morphologies, or obstructions require regional modifications in sampling methods?

(4) Clearly the ease and efficiency of collecting a sample must be considered. In evaluating sampling ease, we are concerned about the availability, work hours, health, safety, and comfort of field crews. In the Northeast, the sampling window for crews was June to September. Chapter 6 details the logistical considerations of regional surveys.

2.2.1.3 Ecological Considerations

(5) An index sample should distill the essential aspects of the system; that is, it should provide important data about the condition of a waterbody, if not critical information about the fundamental nature of the lake or stream. This aspect is evaluated by the ability of the various indicators, at various times and places, to assess biological integrity or trophic state, diagnose unacceptable levels of integrity or trophic state, or indicate different lake or stream types.

(6) An index sample must establish the relative condition of the waterbody. For example, it must produce data that allow discrimination among lakes or streams of varying quality. This aspect is initially evaluated by an indicator's responsiveness along known disturbance gradients and during periods of greatest, or substantial, disturbance.

(7) Ideally, an index sample is responsive to stressors that occur in places and at times distant from the time and place of sampling. A substantial perturbation of the system must be detectable even if it occurs in another season or in a portion of the lake or stream distant from the index site. This aspect of the sample is evaluated by the life cycle, lifespan, and mobility of the assemblage.

(8) An index sample is best collected with periods and places of maximum anthropogenic perturbation in mind. Atmospheric deposition and agricultural chemical runoff are most stressful in early spring when snow and ice meltwaters enter lakes and erosive rains occur, but late summer is usually when eutrophication is of greatest concern because of in-lake processes, warming, and human use. Places of maximum perturbation in lakes are the littoral and profundal zones and the hypolimnion of stratified lakes. Littoral areas are most frequently the sites of anthropogenic changes in physical habitat structure (snag removal, dock construction, beach development, trash deposits) and water quality deterioration (septic field leaching, lawn runoff, acidification, stream inlets). The profundal zone collects sediments, and in stratified lakes these can deoxygenate the hypolimnion as they decay.

(9) We wish to sample biota at biologically relevant times and places, that is, when and where we can most likely obtain an information-rich sample. We evaluate this aspect by examining published life history characteristics of the assemblages sampled.

(10) Indexing must provide an accurate picture of the lake or stream population. Indicators that produce excessive numbers of false positives or false negatives eventually will result in unneces-

sary expenditures or unnecessary loss of ecological values. Overall condition accuracy may be evaluated through the degree of agreement of multiple indicators and responses to known levels of disturbance; taxonomic accuracy of voucher specimens is another component of accuracy.

The foregoing criteria serve to guide in the selection and evaluation of lake indexing.

2.2.2 Index Period, Index Location, and Sampling Gear

Before the lake pilot, we evaluated index period adequacy and sampling locations. Preferred index periods for each attribute were tabulated (Table 2-2) and evaluated in combination. We estimated that EMAP lake surveys would require a 2-month index period to sample approximately 60–80 lakes in this region. This estimate was based on each of 4–5 crews sampling 2 lakes per week for 8 weeks. Late fall through early spring was rejected for safety and accessibility reasons. Late spring and early fall were rejected because of the substantial temperature changes and consequent biological changes that lakes experience over two months during those seasons. June through early September was chosen as the index period. Birds were sampled by separate crews in June because this is the peak breeding season, when birds are least mobile and easiest to identify. July and August are months of substantial thermal, nutrient, and recreational stress on lakes; also, fish, benthos, macrophytes, and zooplankton assemblages are relatively stable.

Index sites fell into two major groups, midlake and littoral (Table 2-3). Based on previous experience and the literature (Herlihy et al., 1990; Smol and Glew, 1992; Tessier and Horwitz, 1991), we chose the midlake area at or near the greatest depth for water quality, chlorophyll-*a*, sedimentary diatoms, sediment toxicity, and zooplankton.

The middle of the lake was selected for water quality and chlorophyll sampling because it best represents the pelagic water volume of the whole lake, is well mixed in the surface layer, and is least influenced by littoral perturbations. Herlihy et al. (1990) found insignificant differences in water chemistry collected from 3 deep water sites at 41 randomly selected lakes in the northeastern United States. The midlake location offers the greatest rain or focusing of diatoms; the sediments there include pelagic and benthic diatoms and are least disturbed by wind, currents, and macroinvertebrates. Sediment samples from such locations are relatively invariant; standard deviations in inferred pH were measured for diatom assemblages of 10 surface sediment samples and 3 sediment cores from the same lake. The standard deviations were 0.21 and 0.10 pH unit, respectively (Charles et al., 1991). For many of the same reasons, a midlake site yields the most

Table 2-2. **Preferred Index Periods for Various Lake Attributes and Assemblages**

Attribute	Spring	Summer	Fall	Winter
Water Quality	●	●	●	
Habitat Structure		●		
Chlorophyll-*a*		●		
Sedimentary Diatoms	●	●	●	
Zooplankton	●	●		
Benthos		●		
Fish	●	●		
Birds		●		
Sediment Toxicity		●		

Table 2-3. **Index Locations for Various Lake Attributes and Assemblages**[a]

Attribute	Midlake	Littoral	Comment
Water Quality	1		0.5 m below surface
Chlorophyll-*a*	1		0.5 m below surface
Sedimentary Diatoms	1		1 40-cm core
Zooplankton	1		vertical tow from bottom
Sediment Toxicity	3		grabs
Benthos	3	3	Stratified subjective
Fish	1–10	1–10	Stratified random, # depends on lake size
Birds		20–24	Randomized systematic, # depends on lake size
Habitat Structure		10	Randomized systematic

[a] Numbers indicate number of samples taken. Single samples are taken at midlake, at or near maximum depth. Multiple samples are taken according to protocols developed during the pilot study and described in the text.

representative sample of sediment for sediment toxicity tests. The maximum depth provides the longest vertical zooplankton tow, any potential hypolimnetic zooplankton assemblage, and the most varied zooplankton pelagic habitats. Tessier and Horwitz (1991) reported only 5% within-lake variation at the index site for zooplankton in the same lakes that Herlihy et al. (1990) studied.

For benthos, birds, habitat structure, and fish, our general concern was to determine if we could obtain an information-rich, discriminating, and repeatable sample with a minimum amount of effort, conforming to our need to spend no more than two days on each lake visit; a single day would be preferable. Our concerns included choosing the location and number of sampling stations, the number of samples per station, and the type of gear or combination of gear types. Addressing these questions was a major thrust of our 1991 pilot survey.

2.2.2.1 Benthos

Our indexing objective for benthic macroinvertebrates was to compare the relative effort required to assess lake condition consistently in 2 to 3 hours by sampling sublittoral, profundal, and littoral habitats and to evaluate the relative information return in terms of numbers of species and individuals collected. Benthos were sampled in two habitat types by petite PONAR grab (Appendix 2B): the deepest area and the sublittoral zone (within which we selected two to three sites). Coarse littoral substrates (cobble, gravel, sand, macrophytes, snags) were sampled at three to four locations by sweep net or handpicking. Sampling times and necessary sampler expertise were recorded.

2.2.2.2 Fish

After numerous discussions with lake ichthyologists and fishery biologists in universities and state and federal agencies, we found no standard protocols for obtaining a lake index sample for the fish assemblage. We planned to collect all the common and most rare species and to estimate the proportional abundance of the abundant and common species, which meant sampling the variety of fish macrohabitats within each lake with habitat-appropriate gear. Our primary means of determining whether methods were adequate was to sample each habitat with the chosen gear to estimate marginal return with increasing sampling effort. Our objective for fish sampling was to determine the combination of gear types and the amount of each required to collect 90% of the species caught by intensive sampling. We also were concerned about whether to use a sampling design that resulted in a probability sample stratified by habitat, or one in which we subjectively chose locations based on expert knowledge about where in lakes fish were likely to be found.

Consequently, we compared catches from both stratified random and subjectively selected sites within a lake.

Ultimately, we needed to develop an indexing methodology that would reduce subjectivity in site selection among different crews to minimize crew variance, optimize sampling effort, ensure that major habitat types were sampled, and produce data that could be related to the proportions of those macrohabitats. These were not simple tasks because of the patchy distribution of fish species, differential mobility of fish species and size classes, variable presence and extent of macrohabitats, our inability to view the lake as a fish does, and variable gear efficiency among different habitats and lakes.

Our pilot study evaluated various gear types to determine which was most cost effective relative to labor and catch at the lake scale and what combinations of gear would be best. In all, seven gear types were evaluated (Appendix 2B): gill net, trap net, minnow trap, eel pot, beach and short seines, and boat electrofishers. The small indicator lakes were sampled for two days with gill and trap nets at four midlake and four littoral locations, chosen using systematic random and subjective methods, in those lakes with features that concentrate fish. Large lakes were sampled for three days with the same gear types at eight locations stratified randomly, and occasionally subjectively located. Electrofishing effort was evaluated at two to four transects on small lakes and at four to seven transects on large lakes.

2.2.2.3 Birds

To our knowledge, riparian and littoral birds had not been quantitatively sampled in New England lakes. Thus our pilot objective for birds was to determine if indexing methods developed for terrestrial and stream bird surveys (Brooks et al., 1991; Robbins et al., 1986) could be adapted to lakes. We were especially concerned with the amount of time required for sampling and the index period variance. We recorded the time it took to collect the data and sampled bird assemblages in each lake twice with an interval of approximately two weeks. An ornithologist indexed birds at a set of 20–24 evenly spaced sites along the shoreline (Appendix 2B).

2.2.2.4 Physical Habitat Structure

Physical habitat structure, except for macrophyte density, has rarely been quantitatively studied in lakes. Our methodological objectives for physical habitat in the pilot were to (1) quantify the amount of time required to obtain the sample from lakes of various sizes, and (2) determine if the variables sampled and the sampling methodology were appropriate. We assessed the appropriateness of the habitat structure data by comparing them with an annotated lake map and an overall assessment of lake character to determine if important lake characteristics were missed by the 10 samples; we also examined the data for associations with biological responses. We adapted variables developed for streams (Plafkin et al., 1989; Platts et al., 1983) to lakes (Appendix 2C) and measured them at 10 evenly spaced shoreline sites.

2.2.3 Indexing Results

2.2.3.1 Zooplankton

One of our objectives for zooplankton was to begin determining the magnitude of indexing variation relative to other variance components. (See Chapter 4 for details on the structure for estimating variance components.) Developing a more complete picture of variance components will require revisits across years as well as within years. The zooplankton estimates summarized in Table 2-4 give us preliminary insight into the relative magnitude of index variance. Variability across lakes is clearly higher than index variance; however, index variance for the body size ratio is of some concern, because it represents about one-half of among-lake variance.

Table 2-4. Zooplankton Index and Lake Variance Estimated for Species Richness and Body Size Ratio, Two Candidate Metrics for the Zooplankton Assemblage

ANALYSIS OF VARIANCE

Response Metric	Index Variance	Lake Variance
Species richness	12.3	46.5
Body size ratio	6.0	14.1

2.2.3.2 Benthos

Our preliminary decision about benthos indexing was to sample the oxygenated profundal zone (sublittoral zone), although not all the samples from all the habitats were processed in time for the information to be included in this report. A cursory analysis revealed that the taxa richness of the profundal deep sites was low, and therefore assemblages collected there contained very little information compared to what we could obtain from the other habitats. Forbes (1925) described a similar situation for midwestern lakes. The high microhabitat heterogeneity seen in the littoral zone was of concern because the sampling effort required to obtain an adequate index sample was likely to be great. As explained in Section 2.2.1, high microhabitat variability requires either greater habitat stratification and more samples or a sizeable number of systematic random samples. Stratification requires considerable experience sampling macroinvertebrates and both choices result in greater field time than the desired 3-hour limit. However, the sublittoral zone sites in the pilot produced a wide range of taxa (2–43) and individuals (7–82) among lakes, yet sites in a lake were physically and biologically similar. In four lakes, triplicate sublittoral samples yielded variances of 0.2 to 4.2 for a biological integrity index, with mean scores of 19 to 26. Future benthos sampling, therefore, is expected to focus on the sublittoral zone, but a final decision must await analysis of the littoral samples and the remaining sublittoral and profundal samples. Pilot work in 1992 was designed to evaluate the effects of number of samples, sampling device (petite PONAR, standard Ekman, K-B corer), and mesh size (589, 417, and 246 μm) on species richness and number of individuals per sample.

2.2.3.4 Fish

The most effective single gear type was a boat electrofisher, in terms of proportion of individuals and species caught (Figure 2-1), although it had serious logistical limits. Knight et al. (1991) also found electrofishing was less selective and captured the most fish and species, and the majority of total species, compared with gill and trap nets, although passive gear added species missed by electrofishers. However, there is some danger of injuring large trout if electrofishers are improperly operated (Reynolds and Kolz, 1988; Sharber and Carothers, 1988). Although we did not kill or mark any fish electrofishing, neither did we examine them for internal injury. Electrofishing is clearly desirable because of its high sampling efficiency and because it harms relatively few fish compared with the mortality for gill and trap nets (usually 100% and occasionally 96%) and the occasional mortality of nontarget species (30 turtles, 1 cormorant, 2 muskrats, 1 beaver) in trap nets. However, the logistical problems of generator weight and night sampling required further pilot research before implementing electrofishing of all lakes, especially those inaccessible by road. Therefore, a separate study comparing various lightweight electrofishers was conducted in a subset of lakes in 1992. Because we could not use boat electrofishers in the 1992 pilot, we examined combinations of the other gear types to see which one would produce an adequate sample of the fish assemblage.

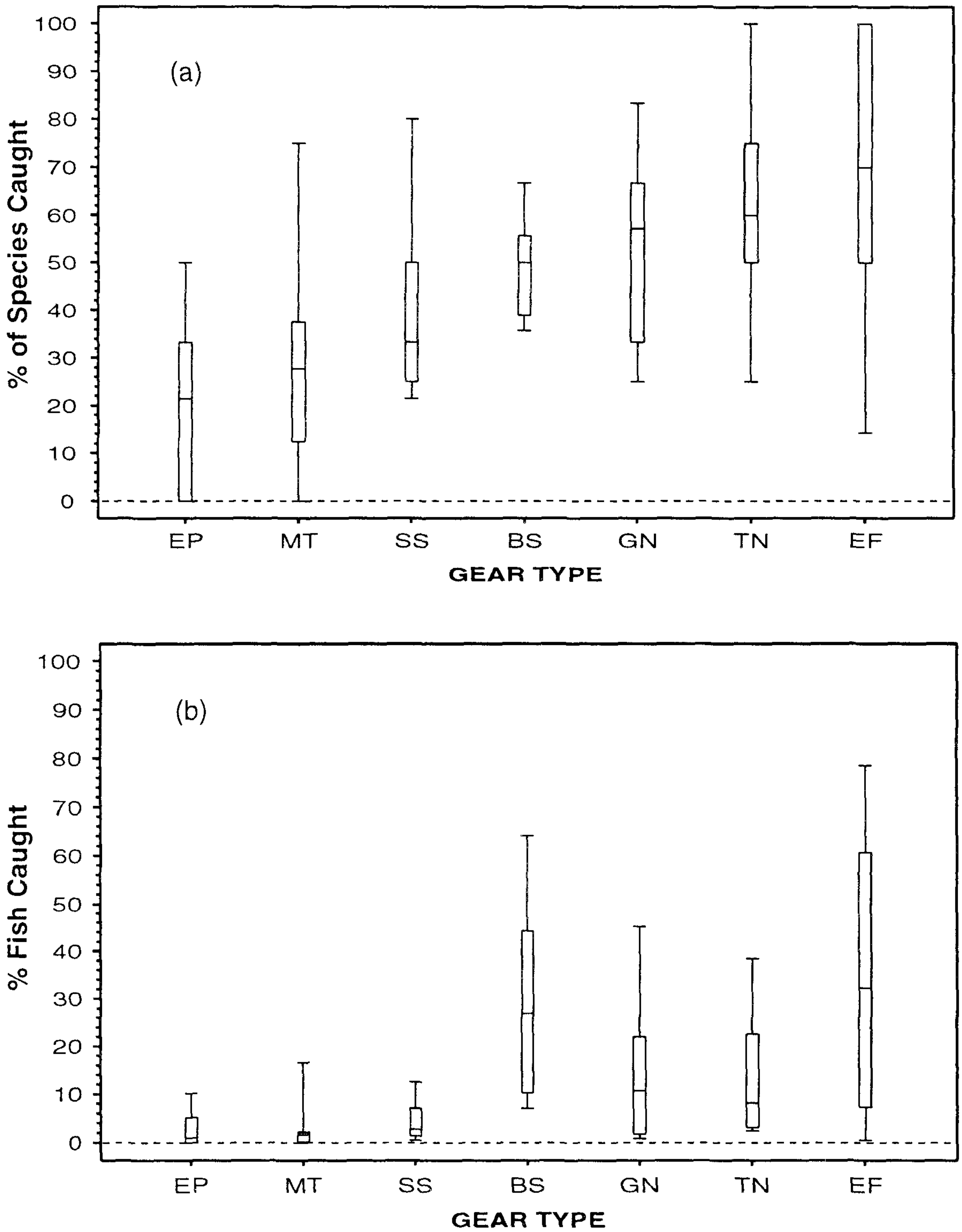


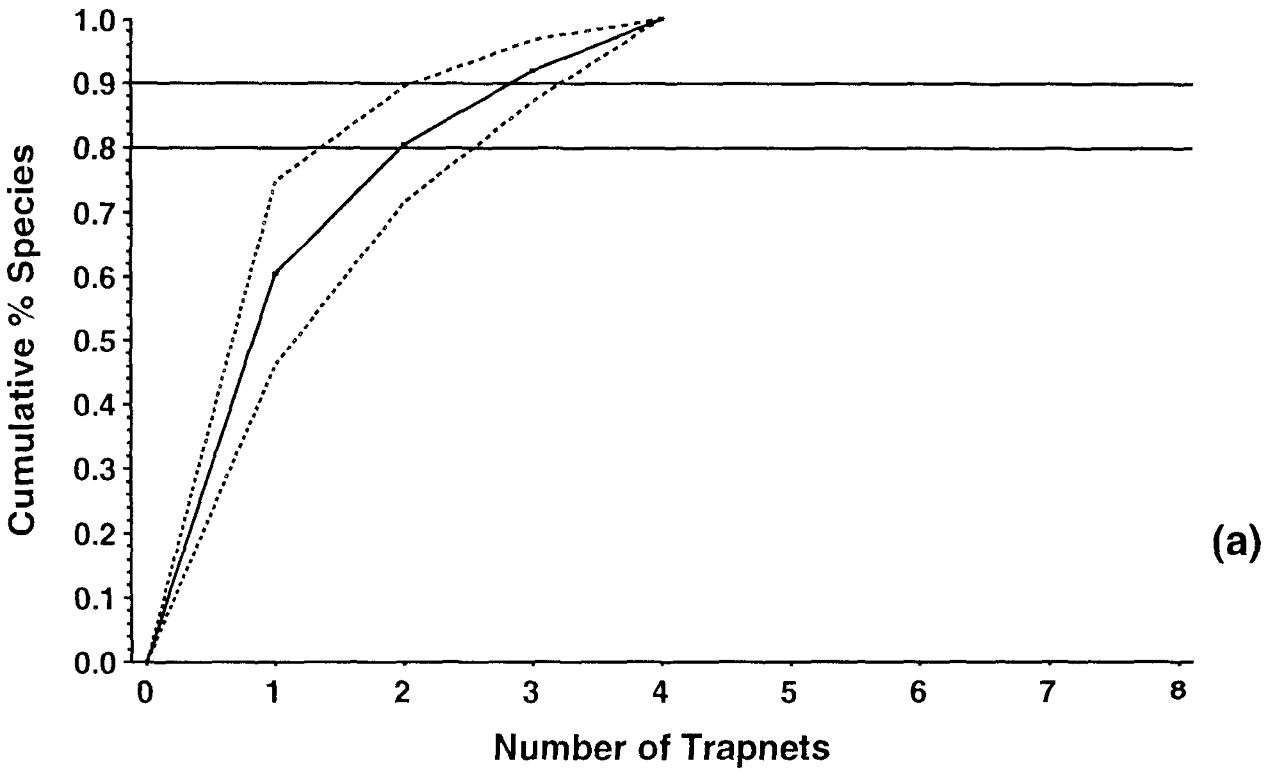
Figure 2-1. **Capture efficiency of fishing gear by (a) % species caught and (b) % individuals caught, both by all gear in combination, for 19 lakes. Box plots show medians, quartiles, and ranges. EP = eel pot, MT = minnow trap, SS = minnow seine, BS = beach seine, GN = gill net, TN = trap net, EF = boat electrofisher.**

Gill nets are the most appropriate gear type for sampling the deep water habitats; trap nets effectively sample the shallow littoral habitats. Initially, EMAP limited the sampling time to two days per lake for small lakes and three days for large lakes. This was not usually enough time for a crew to set and tend any more than four trap nets and four gill nets per day, given other sampling duties and travel time. Based on this four nets per day limit in the pilot, our data revealed that 90% of the species caught in the trap nets could have been caught in three traps in small lakes (7–28 ha) and in six traps in large lakes (100–564 ha) (Figure 2-2); similar results were obtained for gill nets. This does *not* mean that we will always capture 90% of the species present in a lake. The combination of gear types caught the same species as electrofishers in 11 of the 19 lakes; without electrofishing, 1 species would have been missed in 4 lakes, 2 species would have been missed in 3 lakes, and 5 species would have been missed in 1 lake. Consequently, in the 1992 pilot, gill and trap nets were emphasized, with seining and minnow traps in appropriate habitats.

At only two lakes were more species caught using subjective placement than with stratified random sampling (Table 2-5). At six lakes, more individuals were caught at subjectively selected sites. However, stratified random net settings caught more species at four lakes and more individuals at five lakes. We concluded that the occasionally greater catch does not warrant the variance introduced by subjective gear placement by crews with different expertise. More important, stratified random placement allows us to make lake-wide estimates of species and prevents biasing the catch by placing gear in concentrating sites that do not represent a major portion of the lake.

Based on the results of the 1991 pilot, we have developed tentative protocols for sampling fish in lakes. We will use gill nets chiefly in pelagic macrohabitats. If the lake is unstratified with > 4 mg/L dissolved oxygen at the bottom of the deep site, all gill nets will be set on the bottom. The first gill net will be placed near the deepest site and the remainder will be placed midway between it and the shore in randomly selected directions. If the lake is stratified with > 4 mg/L profundal dissolved oxygen, gill nets will be set as just described, but at the bottom and at the thermocline. On lakes > 100 ha, one gill net will be placed at mid-epilimnion and one in the littoral zone, in addition to the profundal sets. If a lake lacks a well-oxygenated hypolimnion or metalimnion, the gill nets will all be placed in the epilimnion and littoral zones.

The trap nets will be placed at systematically selected sites in the littoral zone stratified by major habitat type with the following variables: human influenced or natural, cover or open, cover type,



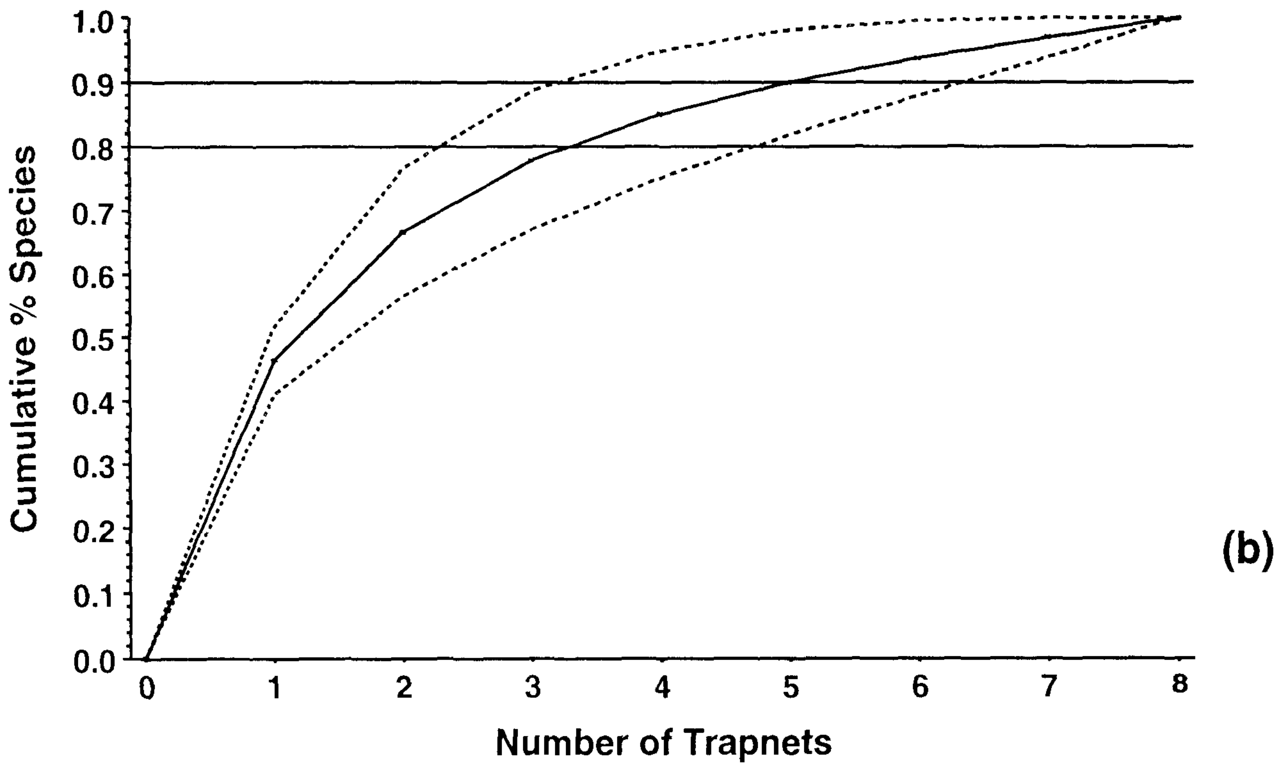


Figure 2-2. Species/effort curve for fish caught by trap nets in (a) small and (b) large lakes. Dashed lines are 90% confidence intervals.

Table 2-5. **Maximum Fish Catches per Net through Use of Stratified Random and Subjectively Placed Trap Nets**

Lake[a]	Random		Subjective	
	Individuals	Species	Individuals	Species
BL	74	3	4	2
TI	18	3	9	3
MA	6	4	0	0
KE	19	4	150	5
TO	21	4	25	6
GR	6	4	9	4
WE	7	5	8	5
DE	23	5	9	4
HA	87	6	9	5
NU	2	2	5	2
NE	5	2	116	2

[a] Two-letter code = the first two letters in a lake's name.

fine or coarse substrate. Some crew discretion will be allowed to ensure that the nets will fish adequately and not be destroyed; that is, they will not be set on extreme drop-offs, snags, boat landings, etc. However, the nets will not be placed at sites expected to be highly productive, such as inlets, outlets, points, shoals, bay mouths, or areas of cover, unless these occur at the systematically selected sites. Fish caught by the systematically placed gear can be extrapolated to whole lake values by multiplying the net catch times the macrohabitat volume and then summing those volumes. Fish caught by beach seining will also be calibrated by beach volume or area. In addition, crews will be expected to spend one additional hour (1) sampling habitats that would otherwise not be sampled (such as the highly productive ones just listed) or (2) using non-standard gear or methods (such as backpack electrofishers, angling, enclosure nets, dip nets, or baited nets), but the catch will be recorded separately.

2.2.3.5 Birds

The circular point count method for estimating bird species and proportionate abundances was easily adaptable to lakes. Index variance between the two visits of the bird crews was relatively low, compared with variability among lakes (Table 2-6). Ratios of index variance to among-lake variance for birds also tended to be lower than ratios for zooplankton (Table 2-4), but the lake populations were not identical. Apparently, a single visit during the breeding season provides an appropriate index of the number of individuals and species per lake, as long as sampling occurs within 4 hours after sunrise and avoids high winds and rain. However, variance components must be evaluated further on a larger set of probability lakes.

Table 2-6. Index and Lake Variance Estimated for 12 Candidate Richness and Tolerance Metrics of the Bird Assemblage

Analysis of Variance		
Response Metric	***Index Variance***	**Lake Variance**
Number tolerant species	1.3	5.3
Number intolerant species	1.4	3.5
Number tolerant individuals	242.7	1735.2
Number intolerant individuals	17.4	100.9
Intolerant species/Tolerant species	0.1	0.3
Intolerant individuals/Tolerant individuals	0.3	2.2
% tolerant species	.001	.004
% intolerant species	.001	.003
% tolerant individuals	.002	.02
% intolerant individuals	.002	.01
Number species	16.1	47.0
Number individuals	1412.6	4801.7

2.2.3.6 Habitat Structure

The 10 systematic sampling sites allowed us to quantify the spatial extent of habitat measurements or visual observations made at the lakes. When we compared results of this systematic physical habitat survey with those from annotated lake maps and focused observations by field biologists, we found no habitat structure concerns missed by the systematic approach. In fact, the process of stopping at 10 sites and carefully investigating them produced information on stressors, habitat conditions, and wildlife that would have been missed by a more cursory shoreline cruise, especially on the large lakes. With 10 observation stops per lake, the physical habitat survey should detect, on average, any habitat structure that comprises at least 10% of the lake shoreline. Consequently, the spatial extent of habitat features must differ by at least 10 percentage points among lakes to be distinguished by the shoreline survey.

The habitat structure protocol was modified slightly for use in 1992. To provide additional information for interpreting bird assemblages, vegetation type will be classified as coniferous, deciduous, or mixed in future surveys. To aid in placing fish sampling gear, a section was added to the field form to classify each of the 10 sites by fish macrohabitat type (natural/human, cover/open/mixed, cover type, substrate type). We found that each observation stop required < 5 minutes to make and record observations; on small lakes, sites were 5-10 minutes apart, and on large lakes with complex shorelines, these times were doubled. We concluded that habitat structure could be assessed in 1.5–3.5 hours on lakes of 7–560 ha. This method of indexing physical habitat structure, therefore, seems appropriate to apply to grid lakes as long as the staff are adequately trained on a diverse set of lakes.

2.2.3.7 Indexing Summary

In summary, we determined that existing indexing protocols were adequate for several lake attributes, including water chemistry, chlorophyll-*a*, sediment diatoms, and zooplankton (Appendices 2B and 2D). For these attributes, the pilot survey focused on an evaluation of the logistical requirements for conducting a regional-scale survey, on measuring the status of lakes in the Northeast relative to these attributes (population descriptions), and on initial estimates of the replicability of these protocols by repeat visits during the index period.

For other attributes (fish, macrobenthos, riparian birds, and physical habitat structure), the pilot survey targeted development of protocols for indexing different types of northeastern lakes. For birds, the circular point count method developed for terrestrial systems could be used along the lake shoreline with little modification. The systematic design for collecting physical habitat structure data along a lake shoreline is an objective way of obtaining quantitative information about a lake's shoreline physical structure. For both these attributes, future research will focus on quantifying the repeatability of making measurements within the index period and among different sampling crews.

For fish, an indexing protocol that stratifies a lake's fish habitat, then samples each habitat type by random placement of the appropriate habitat selective gear is an objective way of obtaining a lake-wide estimate of the fish assemblage. Intensive sampling of each habitat type allowed us to determine the amount of each gear type to deploy to reach our goal of sampling 90% of the species obtained in the intensive sampling. The pilot survey revealed that boat electrofishing was the most efficient method of obtaining fish assemblage data, but that the gear was too cumbersome to transport to lakes without landings. Future surveys will use the fish indexing protocols developed during the 1991 survey to evaluate repeatability of fish sampling during the index period. Also, resources will be devoted to evaluating electrofishing units that can be transported into remote lakes.

Preliminary evaluation of macrobenthos protocols suggests that the sublittoral habitat is the most favorable for indexing a lake's macroinvertebrate assemblages. A protocol that systematically samples this habitat type is expected to emerge as an effective way of indexing benthos in lakes. The 1992 pilot survey evaluated the effectiveness of several samplers and mesh sizes; evaluating replicability must await the development of an operational protocol.

2.3 SELECTION OF INDICATOR LAKES AND THEIR PHYSICAL AND CHEMICAL CHARACTERISTICS

The selection of indicator lakes served three primary functions: (1) developing site indexing protocols for assemblages lacking them, (2) developing metrics and evaluating their sensitivity, and (3) evaluating the sensitivity of the assemblages. For all three purposes, we sought to represent a variety of lake types and catchment disturbance types and intensities. This section explains the selection process, and describes indicator lake physical and chemical habitat characteristics.

2.3.1 Rationale for Selecting the Indicator Lakes

We selected 19 lakes on which to develop and evaluate the use of diatoms, zooplankton, benthos, fish, bird, and habitat structure indicators. We chose lakes that reflect the range of lake temperature and size throughout the northeastern United States (warm, cold, large, small), as well as the range of disturbance types and intensities. The 19 lakes, represented in this chapter by two-letter codes that equal the first two letters of their names, thus represent a broad spectrum of lakes likely to be encountered in EMAP surveys of the Northeast. We assumed that if questions about sampling gear, methods, effort, and index locations were answered for this set of lakes, the protocols developed would be appropriate for most lakes throughout the region, including lakes of intermediate size and temperature. We chose disturbance types for each lake class that appeared to be fairly distinct and common to the Northeast. Also, by picking lakes along particular disturbance gradients, we were able to compare the effects on candidate assemblages and metrics of suspected stressors and known stressor gradients. We emphasized temperature and size because literature and data analyses revealed these as major factors differentiating biological assemblages in New England lakes (Barbour and Brown, 1974; Dixit et al., 1992; Schmidt, 1985; Tessier and Horwitz, 1991; Underhill, 1985). We selected lakes from Rhode Island to northern Maine to maximize latitudinal, ecoregional, and biogeographic differences (Figure 2-3).

In selecting the 19 lakes, our concern was not in choosing lakes whose biota had already been sampled; instead, we simply wished to locate lakes with different types and amounts of human activity (agriculture, silviculture, residential development, fish stocking). Rawson (1939) considered human disturbance as important as geology, topography, and climate in determining the biological character of lakes. We also selected lakes with boat landings to minimize initial access problems and sought to maximize temperature and size differences, recognizing that the presence of boat landings indicates at least some minimal level of disturbance. Candidate lakes and disturbance gradients were determined by consultation with state water quality and fishery biologists and by limited field reconnaissance. Lake selection was made more problematic by the lack of time before sampling to calibrate the subjective perceptions of different state biologists in different ecoregions, creating different assumptions of what were highly and minimally disturbed lakes of a class and stressor type.

More complete and comparable catchment data were obtained for each lake during and after the index period, allowing a more precise depiction of disturbance levels, as summarized in Table 2-7. USGS land use/land cover data, revealed that the catchments of 2 of the 19 lakes (MA, BL) were < 50% forest (Table 2-7). This database does not distinguish cut from uncut forest, although the road density in forested catchments suggests historical cuts. Three lakes had urban lands (MA, KE, WE) and four had agricultural lands (BL, TI, FR, JO) in their catchments. Shoreline disturbances and road and human population densities were also used as indicators of over-

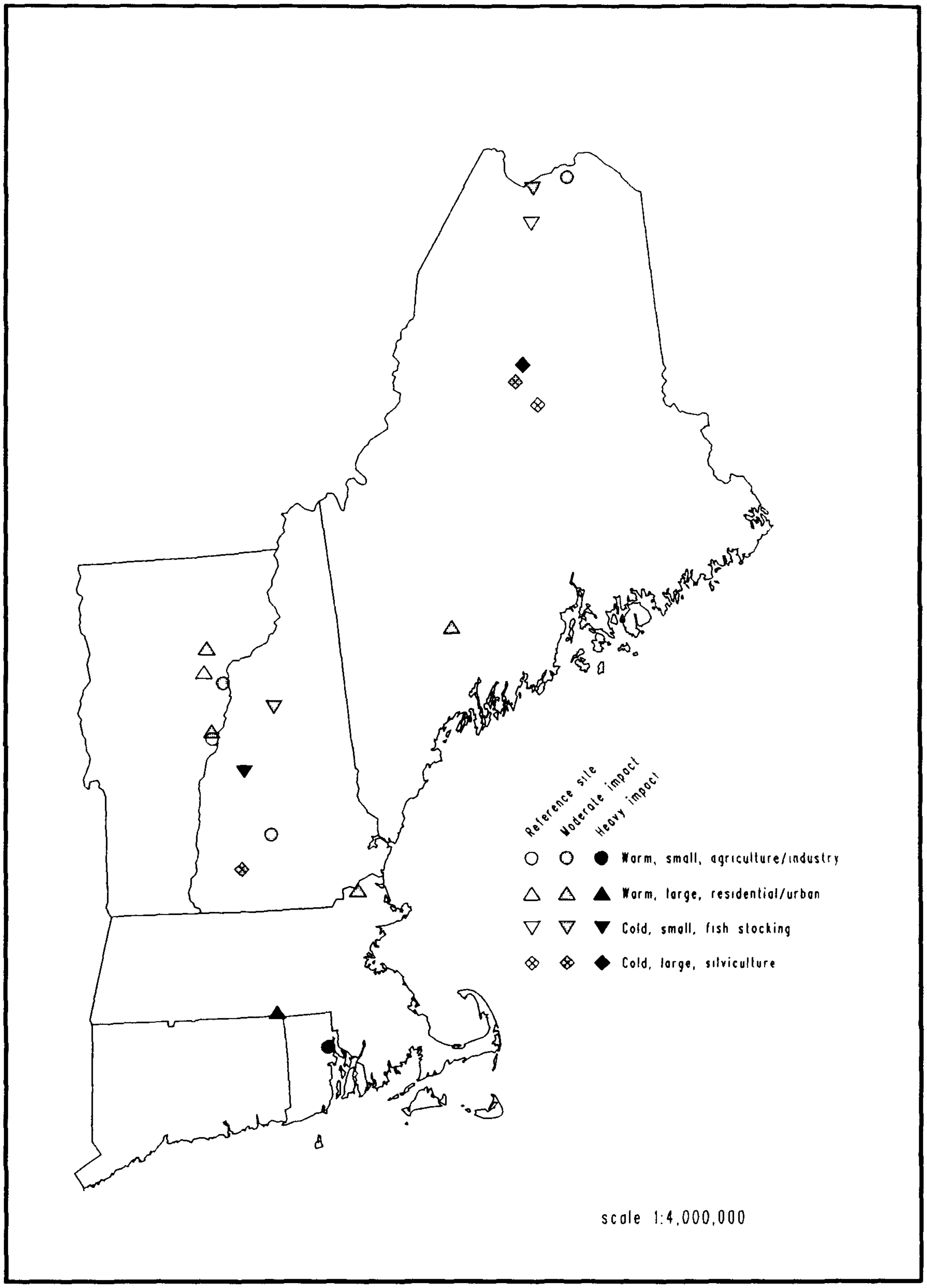


Figure 2-3. Locations of the 19 indicator lakes in New England.

Table 2-7. Catchment Conditions of 19 Indicator Lakes

Lake	Shoreline[a] Disturbance Rank[c]	% Forest	% Agriculture	% Urban	Urban Agricultural Rank[c]	Road Density (m/ha)	Road Density Rank[c]	NPDES[b] and Population Density Rank[c]	Mean Rank[c]
Small Warm (mostly agricultural impact)									
<u>AB</u>[d]	9	92.8	0.0	0.0	1	11.5	9	13	8
BL	6	29.4	63.7	0.0	18	10.6	7	6	10
FR	11	71.6	25.0	0.0	16	25.9	15	16	15
TI	14	52.3	42.1	1.1	17	19.7	14	15	16
MA	16	0.0	0.0	91.0	19	158.4	19	19	19
Large Warm (mostly residential impact)									
<u>GR</u>[d]	15	88.2	0.0	0.0	1	7.5	5	10	7
TO	13	79.2	3.9	1.6	11	19.1	13	7	12
FA	17	87.3	9.2	0.0	12	12.1	10	9	13
JO	19	80.5	13.7	0.1	13	14.7	12	11	14
KE	12	70.2	5.5	10.9	15	35.0	17	18	17
WE	17	59.4	0.2	17.4	14	36.3	18	17	18
Small Cold (mostly fish stocking impact)									
<u>UP</u>[d]	1	91.9	0.0	0.0	1	4.1	2	4	1
HU	5	92.5	0.0	0.0	1	7.1	4	5	4
RU	2	89.8	0.0	0.0	1	26.2[e]	16[e]	14	9
TE	8	93.3	4.1	0.0	10	14.0	11	12	11
Large Cold (mostly silvicultural impact)									
<u>DE</u>[d]	3	87.9	0.0	0.0	1	0.5	1	3	1
HA	3	86.5	0.0	0.0	1	5.6	3	2	3
NE	10	81.8	0.0	0.0	1	8.7	6	1	5
NU	7	77.0	0.0	0.0	1	11.5	8	8	6

[a] Shoreline disturbance was estimated as part of the physical habitat surveys of each lake.
[b] The EPA's National Pollution Discharge Elimination System database contains information about point source discharges to aquatic ecosystems.
[c] Larger numbers reflect greater disturbance.
[d] Underlined lakes are minimally disturbed reference lakes.
[e] U.S. Forest Service campground.

all disturbance. The disturbance rankings used for evaluating metrics is the mean rank tabulated in Table 2-7. High altitude aerial photographs and AVHRR (advanced very high resolution radiometry, from satellite) with 1-km^2 resolution, will be examined in 1992 and 1993. The photographs are appropriate for assessing the land use and land cover of small catchments, but AVHRR data are appropriate only for regional or large catchment assessments.

The lakes underlined in each group in Table 2-7, and throughout the text, represent minimally disturbed reference lakes for each of the four lake classes. The disturbance gradient within each lake class also offered a mechanism for developing and examining preliminary estimates of metric and assemblage sensitivity. Section 2.5 describes the role of these minimally disturbed lakes as reference lakes. The reference condition derived from a set of reference lakes is one way to identify lakes in acceptable or unacceptable condition. Section 2.5 elaborates on this concept.

2.3.2 Description of Indicator Lakes

The physical and chemical habitats of the 19 lakes are described in Tables 2-8 and 2-9 and Figure 2-4. Sampling and analysis methods are tabulated in Appendices 2B, 2C, and 2D. The 21°C, 5-mg/L dissolved oxygen (DO) habitat is considered the limit for salmonids (Ellis, 1937; Scott and Crossman, 1973). All eight coldwater lakes had habitats at < 21°C and > 5 mg/L DO, although RU had a pH level < 6. Four of the warmwater lakes (FA, JO, BL, FR) had some waters at these temperatures and oxygen concentrations, but only BL lacked a substantial DO deficit. Among the coldwater lakes, only HA and TE had seasalt-corrected chloride levels > 30 μeq/L, suggesting substantial catchment perturbation. The reverse was true among the warmwater lakes, where all but TO and AB had excessive chloride. MA's chlorophyll and TP concentrations indicated it was eutrophic. RU and NU showed a greater chlorophyll/phosphorus ratio (response) than the others. In humid temperate regions, chloride concentration is a useful chemical indicator of land use intensity, and unusually high chlorophyll concentrations for a region are often associated with reduced biological integrity as described by Moyle (1956).

Physical habitat structure also differed among the lakes for natural and anthropogenic reasons. Shoreline development, a measure of shape complexity (ratio of shoreline length to the circumference of a circle with the same area as the lake), was > 1.9 for all the large lakes but KE, and five lakes (UP, AB, BL, TO, GR) had littoral areas > 49% of total area. Six lakes (WE, FA, JO, TO, AB, TI) supported mean macrophyte densities that covered more than 29% of the nearshore area.

Table 2-8. Selected Water Quality Characteristics of 19 Indicator Lakes

Lake Name	≥ 4 mg/L DO[a] ≤ 20°C	Chloride[b] (μeq/L)	Mean Secchi Depth (m)	Total Phosphorus (μg/L)	Chlorophyll -*a* (μg/L)
Small Warm					
AB[c]	–	24	Bottom (3.5)	8.8	4.2
BL	+[d]	78	Bottom (4.3)	11.0	3.6
FR	+[d]	195	2.0	14.0	11.3
TI	–	296	1.9	14.0	6.3
MA	–	1578	1.3	30.0	21.4
Large Warm					
GR[c]	–	35	2.8	7.8	6.5
TO	–	28	5.3	4.5	3.2
FA	+[d]	139	7.5	5.0	2.1
JO	+[d]	110	4.1	7.9	4.0
KE	–	671	2.5	13.0	6.5
WE	–	366	3.9	8.6	6.1
Small Cold					
UP[c]	+	7	2.7	11.0	6.8
HU	+	8	4.6	7.5	5.5
RU	+	16	13.3	0.9	2.0
TE	+	228	6.8	1.8	1.3
Large Cold					
DE[c]	+	13	5.8	3.0	3.1
HA	+	39	4.1	6.8	4.9
NE	+	8	2.9	7.0	3.3
NU	+	20	7.6	0.9	2.2

[a] + indicates ≥ 4 mg/L DO and ≤ 20°C; – indicates < 4 mg/L DO and > 20°C.
[b] Sea salt corrected.
[c] Underlined lakes are minimally disturbed reference lakes.
[d] A small lense of water ≥ 4 mg/L DO and ≤ 20°C existed in this lake near the thermocline or near springs.

Table 2-9. Habitat Structure of 19 Indicator Lakes

Lake	Lake Surface Area (ha)	Maximum Depth (m)	Shoreline[a] Development	% Area < 3 m Deep	% Littoral Macrophyte Coverage	Natural[b] Fish Cover/Site	% Natural Riparian Forest	Riparian[b] Disturbance/ Site
Small Warm								
AB[c]	18	3.5	1.78	100	42	2.9	70	0.7
BL	23	4.3	1.82	66	12	3.7	40	0.3
FR	17	11.4	1.14	36	12	2.2	70	0.8
TI	19	15.5	1.73	47	30	1.9	0	0.8
MA	28	4.0	1.23	39	10	1.3	40	1.6
Large Warm								
GR[c]	168	10.5	2.16	50	18	2.7	40	1.2
TO	312	12.5	3.07	56	32	3.7	40	0.9
FA	187	14.2	2.21	30	50	2.3	30	1.5
JO	160	27.3	2.56	39	40	1.7	0	2.5
KE	100	16.0	1.72	18	15	2.1	50	1.0
WE	537	13.2	3.35	39	36	1.6	40	3.0
Small Cold								
UP[c]	7	4.5	1.28	54	3	3.6	90	0.0
HU	26	12.0	1.64	43	5	3.1	40	0.1
RU	16	23.5	1.21	38	17	3.9	90	0.2
TE	19	14.9	1.49	17	18	2.5	70	0.5
Large Cold								
DE[c]	130	41.3	1.96	0.17	10	1.7	90	0.3
HA	453	37.1	2.07	0.28	12	3.5	60	0.0
NE	564	15.5	1.96	0.28	0	1.6	20	0.4
NU	289	22.3	2.84	0.17	3	2.7	80	0.5

[a] Shoreline development is the ratio of the lake perimeter to the perimeter of a circle with the same area as the lake (e.g., circular lake ratio = 1.0).
[b] Fish cover and riparian disturbance express the number of different cover types or disturbances at a site averaged across all 10 habitat monitoring sites in a lake.
[c] Underlined lakes are minimally disturbed reference lakes.

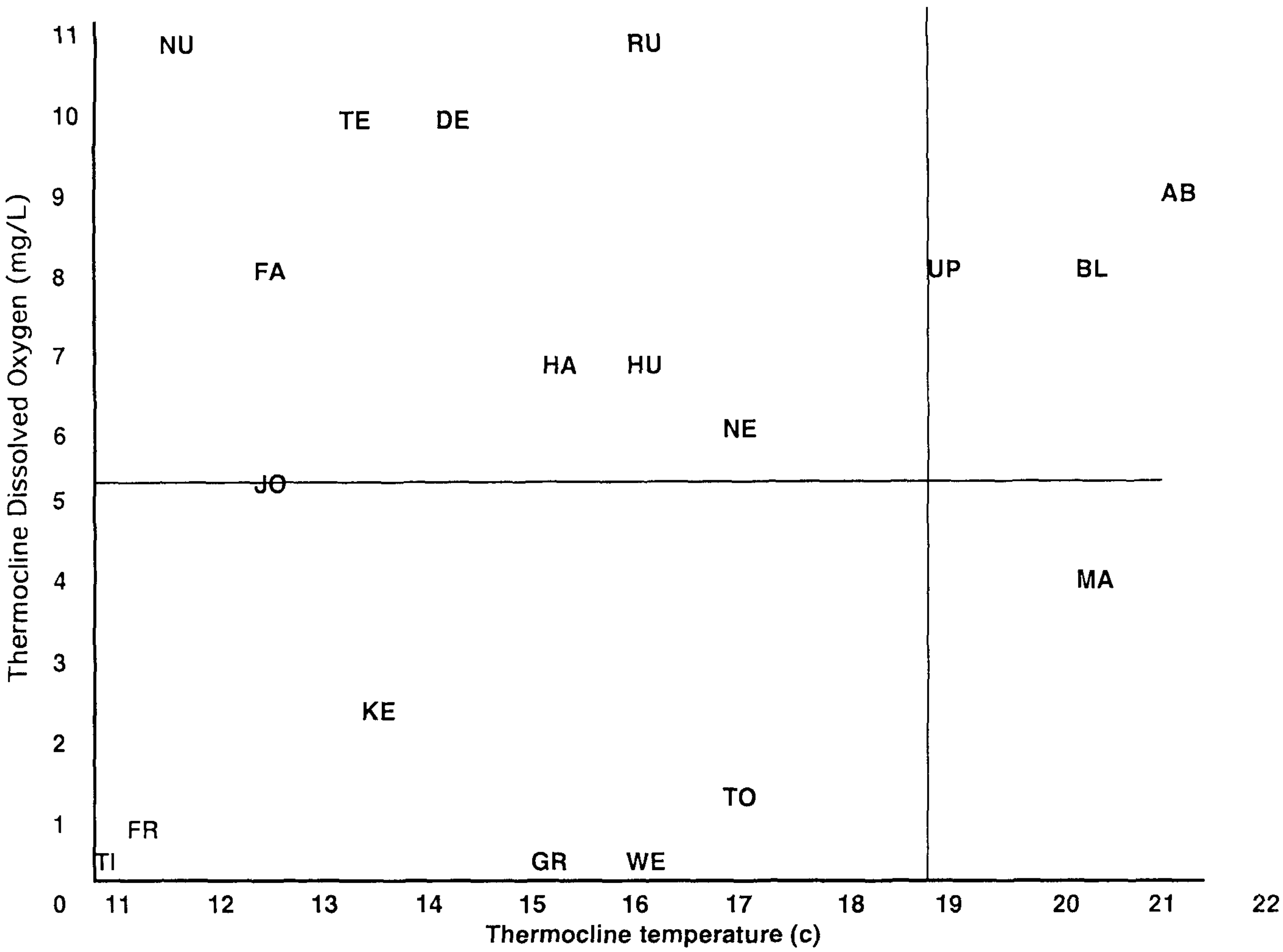


Figure 2-4. Thermocline temperatures and dissolved oxygen concentrations of 19 indicator lakes. The upper left quadrant represents the temperature and oxygen concentrations needed to support a cold water fish population.

Natural forest canopy was dominant along nine lake shorelines (UP, RU, TE, DE, HA, NU, KE, FR, AB); the average number of disturbances/sampling site was > 1 at WE, MA, GR, FA, and JO; and all four lake classes showed fish cover gradients. Clearly, the littoral area and macrophyte density were not always directly related, as one might predict. Increased shoreline disturbance is generally associated with lower biological integrity (see Section 2.4), as suggested by Plafkin et al. (1989).

Sediment toxicity (< 80% of control value) was not found in the four indicator lakes tested (Figure 2-5). In fact, all the indicator lake samples showed greater growth rates than the controls (Figure 2-5b), probably because of more nutritious sediments or the reduced metal solubility of the aerobic, high pH, overlying test water. This result suggests that toxics were unlikely causes of the perturbations observed in these lakes, or that the standard tests used were not sensitive enough. A more nutritious control sediment and modified REDOX (reduction-oxidation) conditions are also warranted.

2.4 INDICATOR DEVELOPMENT AND EVALUATION

2.4.1 Selection of Candidate Assemblages

Clearly we cannot sample all aspects of lakes or even all biological attributes of lakes. Instead, we must implement a process for selecting a set of indicators that will adequately convey the condition of lakes relative to the biological integrity value as efficiently as possible. We chose assemblages, rather than indicator species or community processes, chiefly because we perceived them to be more directly linked to the ecological value of biological integrity. Assemblage information is usually of greater concern to the public and to decision makers than the other options. Also, we understand assemblage ecology and taxonomy better than community processes, and aquatic assemblages appear to have greater diagnostic power, sensitivity, and responsiveness to perturbation than do processes (Schindler, 1987; Ford, 1989; Fausch et al., 1990) or indicator species (Karr, 1987; Landres et al., 1988). Our conceptual model of a lake ecosystem (Figure 2-6) incorporates a basic view of lake macrohabitats (riparian, littoral, pelagic, profundal), as often depicted in limnology and ecology textbooks (Cole, 1975; Smith, 1977; Ruttner, 1963).

Each macrohabitat in the model has assemblages representing the four major trophic levels (producers, primary consumers, secondary consumers, decomposers). Although we are not currently evaluating decomposition in lakes, we are evaluating sediment metabolism in streams. If

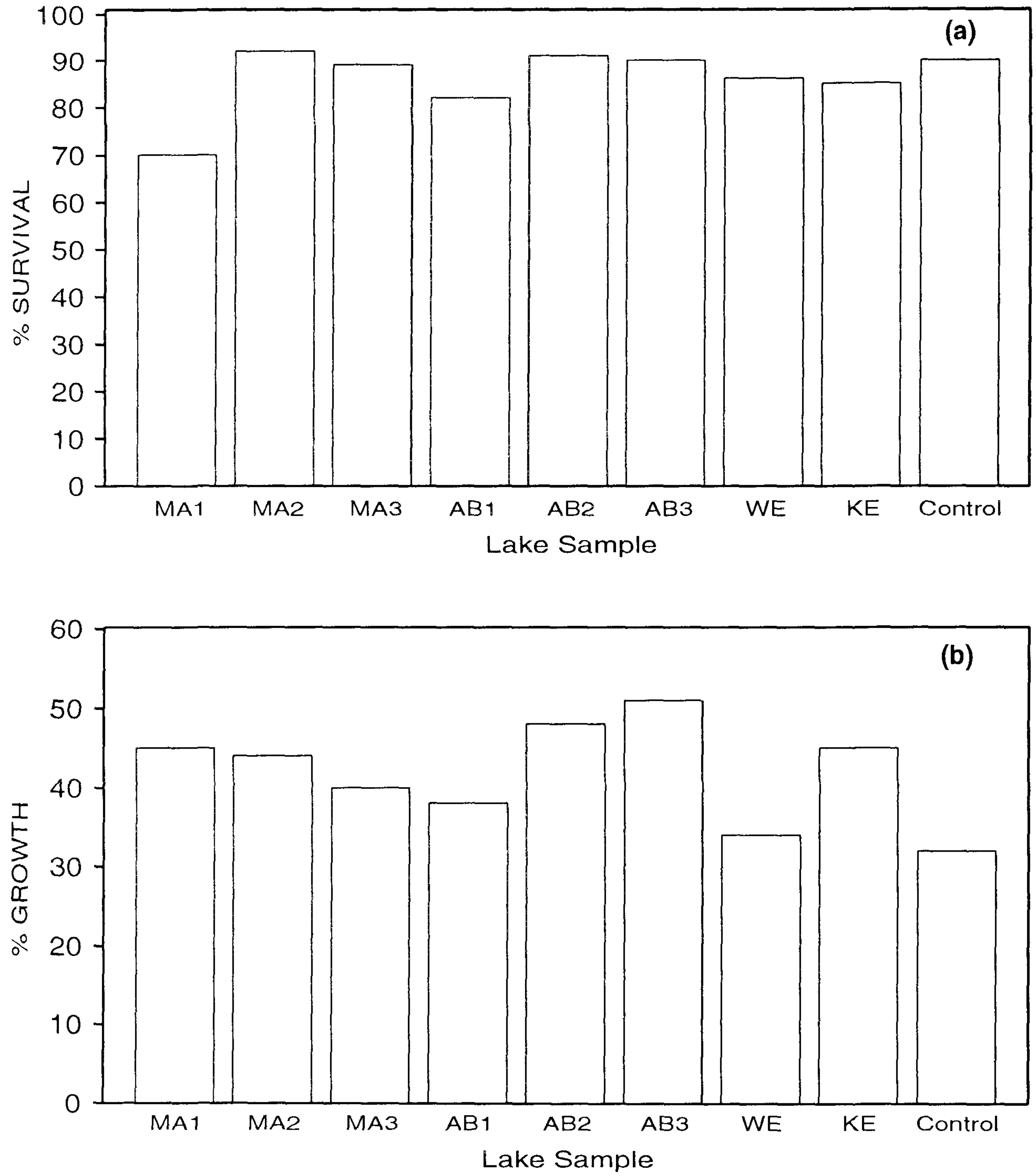


Figure 2-5. **Sediment toxicity of 8 samples from 4 indicator lakes: (a) percent survival, and (b) percent growth.**

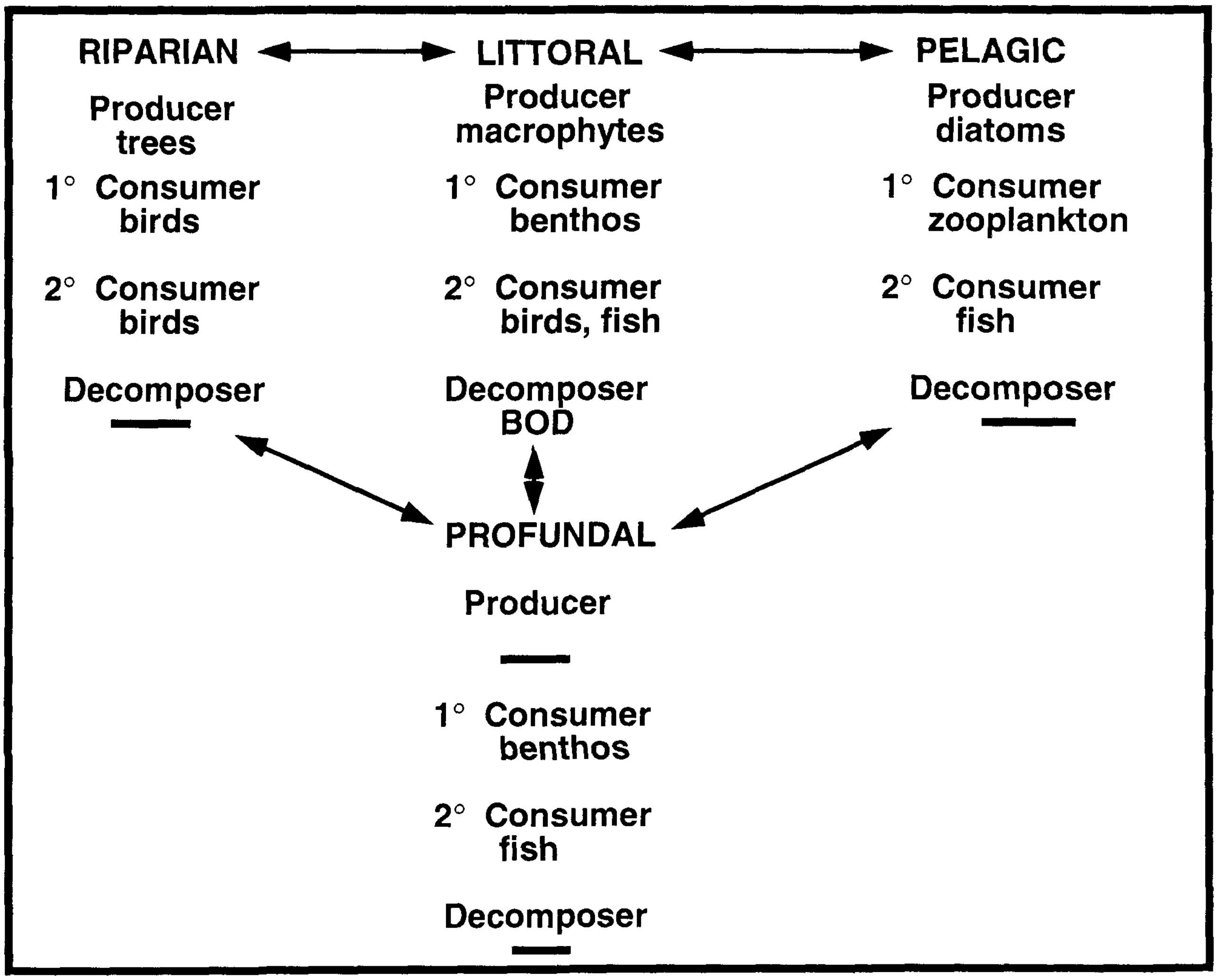


Figure 2-6. Lake conceptual model showing macrohabitats and trophic levels. Only candidate indicators are listed for each trophic level. Note: profundal aquatic insects may serve as prey for littoral and pelagic fishes, and they are an important food of riparian birds when they emerge as adults.

that research proves fruitful, we can use sediment metabolism in lake pilot studies with minor modifications of current methods.

Our initial efforts at assemblage selection included listing possible assemblages and lake attributes, evaluating them through workshops (Halsell, 1990; Whittier, pers. comm.), and conducting literature reviews. We decided *not* to use the following in the pilot: (1) lake periphyton and phytoplankton (except as they are represented by chlorophyll-*a* measurements and sedimentary diatom assemblages), because they were considered redundant with, and more variable than, sedimentary diatoms, (2) lake amphibian assemblages, because they tend to be species depauperate in many lakes, (3) macrophyte assemblages (except as areal cover in trophic state estimates), because they too were believed to be species depauperate in lakes, (4) protozoa assemblages, because they were assumed to be redundant with diatoms and microzooplankton, but less important to the lake, and (5) biomarkers, water column bacteria, fish growth rates, and primary production, because they are not assemblage-level indicators and are highly variable. The candidate assemblages (riparian/littoral birds, fish, benthic macroinvertebrates, pelagic zooplankton, sedimentary diatoms) from which we will derive indicators, and the reasons for their selection, are discussed in the following paragraphs.

Sedimentary diatoms reflect pelagic algal assemblages as well as benthic and littoral assemblages. Although they comprise only a part of the algae in a lake, sedimentary diatoms give us a way to sample the lake's evolution and its condition over recent years. This is impossible with a single phytoplankton sample. Sedimentary diatoms thus have the potential to provide information about the presettlement and preindustrial condition of lakes. Also, the high species diversity of lentic diatoms, and the narrow habitat requirements of many, facilitate their use in diagnosing acidification (Charles et al., 1990), eutrophication (Hall and Smol, 1992), salinization (Lowe, 1974), and general pollution (Lowe, 1974; Stoermer et al., 1985), as well for evaluating trends (Charles et al., 1990).

Zooplankton are the dominant pelagic consumer of most lakes in terms of numbers and biomass. They are sensitive to fish predation (Brooks and Dodson, 1965), acidification (Sprules, 1975; Tessier and Horwitz, 1991), toxics (Keller and Yan, 1991), nutrient enrichment (Siegfried et al., 1989), and thermal condition (Patalas, 1990). Often they are the primary prey of fish.

Like zooplankton, benthos are a trophic link between primary producers and fish. They also are important prey for birds. As the name implies, benthos are the dominant macrofauna of lake bottoms. They have long been used in biomonitoring and they are sensitive to acidification (Lien et al., 1992), nutrient enrichment (Brinkhurst, 1974), toxics (Chapman and Brinkhurst, 1984), and organic enrichment (Hilsenhoff, 1987).

Fish represent the end of the aquatic food chain in many lakes and they are important sources of food and recreation for humans. Unlike the above assemblages, reduced fish diversity has been reported at the stock, species, assemblage, and faunal levels (Hughes and Noss, 1992). Fish are sensitive to migration barriers (Ebel et al., 1989), exploitation (Smith, 1968), toxics (Gilbertson, 1992), acidification (Lien et al., 1992), habitat loss (Miller et al., 1989), and introduced species (Miller et al., 1989).

Traditional aquatic biologists have questioned our inclusion of birds, more than any other assemblage, as a candidate indicator. However, lakes and streams cannot be separated from their catchments (Frey, 1977; Hynes, 1975; Likens and Bormann, 1974; Rawson, 1939), and birds provide a response indicator linking catchment and surface water condition. The riparian zone is an integral part of our conceptual model of a lake ecosystem. Just as we identify habitat specific indicators of other parts of the lake, it is necessary to select indicators of the condition of the riparian zone. Birds are our candidate assemblage for a riparian indicator. Birds are occasionally the top consumer in the littoral and pelagic zones and they are important primary and secondary consumers in the riparian zone. It is useful to monitor birds in the riparian zone because that macrohabitat is often the first and most disturbed lake habitat and the conduit for much water from the catchment to the lake. Furthermore, the riparian zone is the major buffer between a lake and its land use stressors. This is because riparian zones are often wetlands and always active interfaces between terrestrial and purely aquatic systems; consequently, riparian areas themselves demonstrate stress and shape the physical, chemical, and biological character of lakes and streams (Naiman et al., 1992; Gregory et al., 1991). Birds often respond to disturbances before vegetation itself (Redford, 1992), or before the aquatic biota (Brooks et al., 1991; Carson, 1962), and they are sensitive to changes in land use, such as agricultural practices (O'Connor and Shrubb, 1986) and wetland (riparian) damage (Sharrock, 1974). The British Trust for Ornithology has been surveying riparian birds since 1974 (Spellerberg, 1991). Birds provide an enormously popular recreational resource (Payne and DeGraaf, 1975) and the bird fauna continues to decline rapidly. For example, more than 70% of the migrant species in the eastern United States have declined since 1978 and the number of individuals is half what it was in the 1960s (Ackerman, 1992; Senner, 1986; Terborgh, 1989). Birds are sensitive to habitat stressors (Brooks et al., 1991;

Moors and O'Connor, 1992), thermal and recreational stressors (Moors and O'Connor, 1992), and toxic stressors (Carson, 1962) in aquatic ecosystems.

2.4.2 Indicator Evaluation Process

Throughout the evaluation process, we examine how well each candidate meets the 22 criteria listed in Table 2-10. Although many of the criteria are appropriate for selecting both assemblages and metrics, our examples are mostly of assemblages because we are between the assemblage and metric phases of indicator development (Figure 2-7). When we select metrics instead of assemblages, we can substitute "metric" for "assemblage" in many of the following paragraphs. The 22 criteria differ in their importance for selecting assemblages; the most critical ones are so indicated in Table 2-10.

Table 2-10. Response Indicator Selection Criteria (modified from Olsen, 1992)[a]

Richness of Biological Information	Comprehensiveness
1. Knowledge of species' trophic and habitat guilds[b]	16. Applicable to temporary systems
2. Incorporates multiple macrohabitats	17. Suite of indicators of varying size and guilds
3. Includes multiple organismic levels	
4. Spatially and temporally integrative	Societal Relevance
5. High species richness	
6. Easily identifiable species	18. High societal value[b]
	19. Linked with ecological value
Impact Assessment Ability	
	Cost and Logistics
7. Broadly sensitive and responsive[b]	
8. Anticipatory	20. Inexpensively implemented[b]
9. Retrospective	21. Ease of sampling and analysis
10. Diagnostic	22. Rapidly available data
11. High signal/noise[b]	
12. High index period stability	
13. Low measurement error	
14. High among-lake/within-lake variance[b]	
15. Low among-year variance[b]	

[a] These criteria guide the choices of assemblages and metrics for routine monitoring.
[b] Most critical criteria.

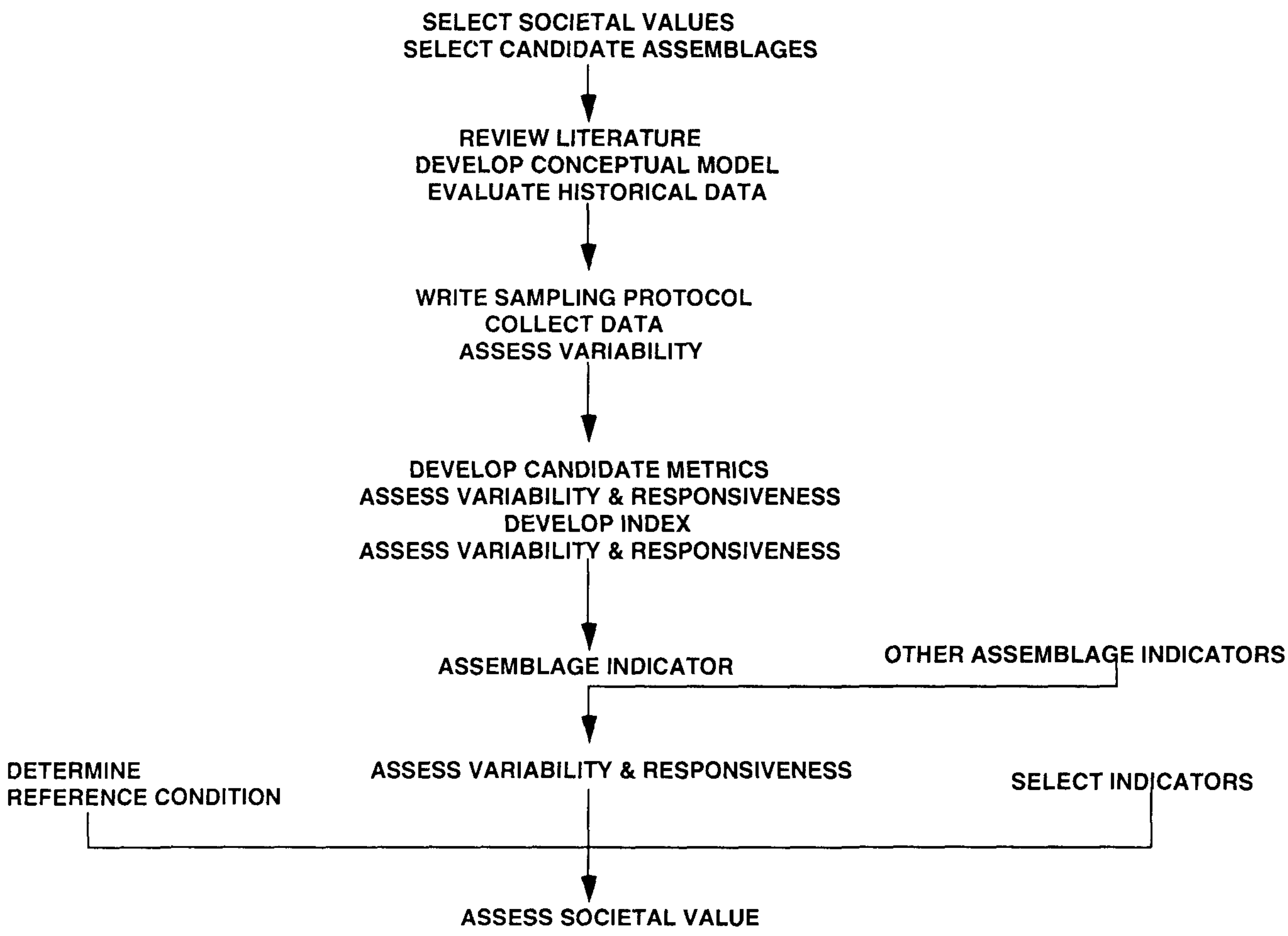


Figure 2-7. The indicator development process showing the five major phases. Each phase involves reevaluation of previous steps.

The first set of criteria is oriented towards the biological information available from the index sample: (1) An assemblage should provide information that allows us to deduce biological, physical, and chemical habitat quality through knowledge of species habitat, trophic, reproductive, and general tolerance guilds. Birds and fish offer more of such guild information than the other assemblages, but sedimentary diatoms are most useful for inferring changes in water quality. (2) It is desirable for an assemblage to represent a number of macrohabitats of the lake (e.g., pelagic, profundal, littoral, and riparian zones). Fish and diatoms rate highest for this criterion. (3) An assemblage should contain information about a number of organismic levels (individual, population, assemblage, community). Such assemblages facilitate analyses at multiple levels of biological organization. Fish provide the most information about multiple organismic levels. (4) An ideal assemblage integrates the spatial and temporal variability in habitat and stressors. Birds, fish, and sedimentary diatoms all seem highly integrative. (5) Assemblages should have high species richness, thereby having the potential to provide greater information about habitat changes. Diatoms are the most speciose of our candidate lake indicators. (6) Assemblages should be relatively easy for trained individuals to identify. Bird species are easiest to identify and most can be identified in the field by trained biologists.

Several criteria are concerned with our ability to assess an impact by means of an assemblage: (7) Assemblages must be responsive and sensitive to a wide range of direct and indirect biological, physical, and chemical stressors on biological integrity, such as land use type and intensity, stocking and harvesting, introduced species, increased turbidity and temperature, decreased snags and macrophytes, and increased concentrations of nutrients and toxics. Fish assemblages appear most responsive to the broad array of stressors (Hughes and Noss, 1992; Karr et al., 1986; Miller et al., 1989; Nehlsen et al., 1991; Williams et al., 1989), but other assemblages are often much more sensitive to individual stressors. Typically, large-bodied species and local endemics are more susceptible to ecological, or actual, extinction (Redford, 1992; Miller et al., 1991; Williams et al., 1989) and their loss can greatly distort some assemblages without outwardly affecting the ecosystem's appearance (Goulding et al., 1988; Paine, 1966; Power et al., 1985; Redford, 1992). (8) The assemblage should be anticipatory; that is, it should suggest early warning of future, more substantial changes. An ideal assemblage is (9) retrospective, indicating past conditions in the population of lakes, and also (10) diagnostic, providing information about probable stressors. Sedimentary diatoms are the most anticipatory, retrospective, and diagnostic indicator, at least to water quality changes. (11) The assemblage must have a high signal/noise ratio so that natural variability does not hinder detection of an impact. The spatial and temporal integration by sedimentary diatoms gives them extremely low noise. A good assemblage (12) is fairly stable during the index period and (13) has low measurement error, including all aspects of field sampling and laboratory analyses. Although birds and fish may be more stable, they are sampled with greater variability, so diatoms reflect the lowest measurement error. (14) For status estimates, an assemblage should exhibit a high ratio of lake population variance/extraneous variance (consisting of index, interaction, and year components). (15) For trend detection, the year component of variance (operating on all lakes together) should be small. These variance components and their effects on status and trends estimation are described in Chapter 4. Diatoms, zooplankton, fish, and birds are believed to have the highest among-lake/within-lake variance; diatoms, zooplankton, benthos, and fish are estimated to have low among-year variance.

It is also useful to examine the comprehensiveness of the suite of assemblages when making selections. For example, small lakes are likely to be temporary, becoming wetlands in some years and lakes in others, or even dry in some years. Thus, an assemblage must be (16) applicable to temporary systems. Birds and sedimentary diatoms are most applicable to ephemeral lakes if EMAP chooses to assess the integrity of such systems. (17) It is wise to limit redundancy yet ensure that a diversity of organism sizes, life spans, taxonomic/functional groups, and macrohabitats are sampled. This suggests monitoring small-sized/short-lived individuals (diatoms or zooplankton), medium-sized/longer lived organisms (benthos), and large-sized/long-lived individuals (fish or birds). Key taxonomic/functional guilds include producers (diatoms or macrophytes), invertebrate consumers (zooplankton or benthos), vertebrate consumers (fish or birds), and decomposers (fungi or bacteria). Important lake macrohabitats are riparian (birds), littoral (birds, fish, benthos, diatoms, microbes, or macrophytes), profundal (benthos, fish), and pelagic (fish, zooplankton, or diatoms).

Five criteria assess societal relevance, cost, and logistics: (18) A satisfactory assemblage has high societal value, that is, people care about these organisms for their own sakes. Vertebrates rank higher than plants and invertebrates for this criterion. (19) All response indicators for biological integrity must be linked directly with some ecological value. In other words, a response indicator of biological integrity should incorporate some aspect of species composition and richness, guild structure, and life history characteristics. Such metrics are best developed for fish and birds. (20) An assemblage must be fairly inexpensive to implement, since sampling and analysis costs limit the number of sites monitored when budgets are fixed. Fish are currently the most expensive lake assemblage; the others appear to have comparable costs. (21) An index sample should be relatively easy to obtain from the field and, if necessary, field samples should be relatively easy to process in the laboratory. (22) Data need to be available for analyses three or four months after the field season. Time-consuming identifications and chemical analyses hinder our ability to plan for the following summer and to meet the EMAP goal of publishing annual statistical summaries nine months after the end of the field season. Riparian bird data are the most easily obtained and most rapidly available.

Two selection criteria or considerations lead to a dilemma. Typically we know more about the guild memberships of species in species-poor assemblages (fish) than of those in species-rich assemblages (zooplankton), but both species richness and guild knowledge are desirable. Presumably this dichotomy results from there being fewer species to know and study in the former group, as well as from disproportionate societal concern and research funding.

2.4.3 Comparison among Assemblages

Although metrics, not assemblages, ultimately will yield the indicators, it is necessary to choose at the assemblage level because assemblages are the functional sampling units. Cost savings come from dropping assemblages, rather than metrics, which usually are only different ways of summarizing assemblage composition data. It will be necessary to choose a smaller suite of core assemblages for national implementation because EMAP lacks the resources to monitor all five assemblages at all lakes annually. The core assemblages could differ from one region to another.

We have reached the following conclusions regarding culling the five assemblages to yield the tentative core assemblages for evaluating biological integrity in lakes:

- Fish, zooplankton, and sedimentary diatoms are tentative core assemblages for at least one cycle of probability lake monitoring.

- The primary reasons for choosing fish are their societal relevance, their representation of many guilds and levels of biological organization, their sensitivity to multiple stressors, their ability to integrate those stressors through long time spans and across multiple habitats, and their role in evaluating the fishability value, as well as the biological integrity value.

- Sediment diatoms were selected because they represent the primary producer component of lake ecosystems, they have low measurement error, they represent several chemical habitat guilds and occupy all lake habitats, they anticipate chemical changes, they provide a historical perspective on lake condition, and they have proven useful for diagnosing many kinds of lake condition. Furthermore, they are easily collected and provide an added perspective to evaluation of lake trophic condition as well as to the biological integrity value.

- Zooplankton are the most abundant primary consumers in lake ecosystems, are prey for aquatic vertebrates, and are relatively easy to collect.

We will continue to evaluate macrobenthos and birds, at a minimum, in northeastern lakes. Macrobenthos, like zooplankton, are primary consumers and prey for vertebrates, but they occupy different macrohabitats and are likely to respond to different stressors. Indexing protocols (sampling a lake for macroinvertebrates) have not yet been fully developed. Our future pilot research on this indicator assemblage will address indexing protocols, variance components, metric responsiveness compared with zooplankton, and redundancy with other indicator assemblages.

There are strong arguments for including the riparian zone in routine lake monitoring and for using birds as a core assemblage. They are represented by many guilds, they are spatially and temporally integrative, and they have very high societal value. Birds are sensitive to physical, chemical, and biological changes in the lake and riparian zone, and they are much easier to sample than vegetation or other animals. In fishless lakes, they often represent the primary vertebrate consumer, although, since most birds are riparian rather than aquatic, they would not be used to assess the same habitat as fish. Bird research will focus on variance components and relative sensitivity to disturbances.

Although we have made tentative decisions about the core and additional assemblages, we will continue to evaluate evidence to support or refute their inclusion. Pilot studies in contrasting areas of the country will be necessary to weigh the regional strengths and weaknesses of various indicators and determine whether the probationary assemblages should be modified. For example, biological integrity measurements may be meaningless in some reservoirs. Bird surveys may be redundant with fish surveys in many lakes, but essential in alpine, prairie pothole, and desert lakes that are naturally fishless and therefore lack another vertebrate consumer. Other indicators, fish tissue contamination, for example, may be sampled initially for baseline information and then sampled infrequently thereafter. Although we are unlikely to monitor all indicator assemblages nationwide, primarily because of cost, a single core collection of indicators that work well for the northeastern United States may be less applicable in other regions.

The seven critical criteria listed in Table 2-10 for selecting assemblages will be essential in this process. We have already assessed the candidate assemblages' societal values and their existing guild information. We cannot accurately estimate their relative implementation costs until after finalizing the field and laboratory protocols for all the candidate assemblages. Meaningful variance estimates require several years of sampling with the established protocols. Similarly, satisfactory estimates of assemblage responsiveness and signal/noise ratio require a greater sample size than the 19 indicator lakes. However, the 1991 pilot is useful as an example of the process of using assemblage sensitivity and signal clarity in assemblage selection. Thus, we sought comparative assemblage information to answer the following questions:

- What is the sensitivity of an entire assemblage to the selected disturbance gradients? This criterion was evaluated by examining the response of the aggregate set of candidate metrics across different types and intensities of catchment disturbance. A particular concern was the total percent of occasions that the assemblage performed as predicted.

- How clear is an assemblage's signal? We evaluated this criterion by comparing the aggregate metric responses for all four minimally disturbed lakes to the responses for the 15 lakes with more highly disturbed catchments.

- Are reference lakes useful for assessing the condition of an aggregate of assemblages? We measured this criterion by comparing the total number of hits (metric score > 1.5X the minimum in a lake class) for each lake class.

As already stated, mere collections of assemblages do not adequately express lake condition; we need a way to extract information from collections of species and their proportionate abundances. The next section describes this process and the role of the surveys of the indicator lakes.

2.4.4 Selection of Candidate Metrics

In order to extract useful information from the assemblage collections that describes the condition of lakes, we have adopted our primary approach from Fausch et al. (1990). In this approach, we select a series of metrics to represent various aspects of an assemblage under consideration. The process involves converting assemblage richness and composition measurements to a set of intrinsically important metrics (e.g., the number of native species), as well as to metrics known or likely to be responsive to the expected array of anthropogenic disturbances to lakes. Fausch et al. (1990) summarize a series of changes that occur when aquatic ecosystems are subjected to various kinds of anthropogenic stress. This list (as modified in Table 2-11) served to guide the selection of many candidate metrics. We based our selection of others on the literature or on evaluated historical databases available in our laboratories. We also used results of the pilot survey to identify metrics apparently responsive to the lake types and gradients we chose—an essential step for assemblages such as zooplankton that are not typically used to assess disturbance. For each assemblage, we developed a conceptual model of how candidate metrics might respond to major stressors (see Table 2-11 and Figure 2-8 for examples) as a way of synthesizing the set of candidate metrics.

Table 2-11. Typical Effects of Environmental Degradation on Biological Assemblages, and Candidate Fish Metrics

A. Typical effects of environmental degradation on biological assemblages (from Fausch et al., 1990; Hughes and Noss, 1992; Margalef 1963).

- The number of native species, and of those in specialized taxa or guilds, declines.
- The percentage of exotic or introduced species or stocks increases.
- The number of generally intolerant or sensitive species declines.
- The percentage of the assemblage comprising generally tolerant or insensitive species increases.
- The percentage of trophic and habitat specialists declines.
- The percentage of trophic and habitat generalists increases.
- The abundance of the total number of individuals declines.
- The incidence of disease and anomalies increases.
- The percentage of large, mature, or old-growth individuals declines.
- Reproduction of generally sensitive species decreases.
- The number of size- and age-classes decreases.
- Spatial or temporal fluctuations are more pronounced.

B. Candidate fish metrics. Metrics may be based on number or percent of individuals or species.

Assemblage Composition • Ordination score • Species richness • Family richness	Habitat Guilds • Pelagic species • Benthic species • Littoral species
Trophic Guilds • Omnivores • Invertivores • Planktivores • Piscivores	Reproductive and Life History Characteristics • Nonguarding lithophils • Young of year species richness • Juvenile species richness • Native adult game species • Trophy fish • Introduced species
Tolerance Guilds • Intolerant species • Sensitive species • Tolerant species	

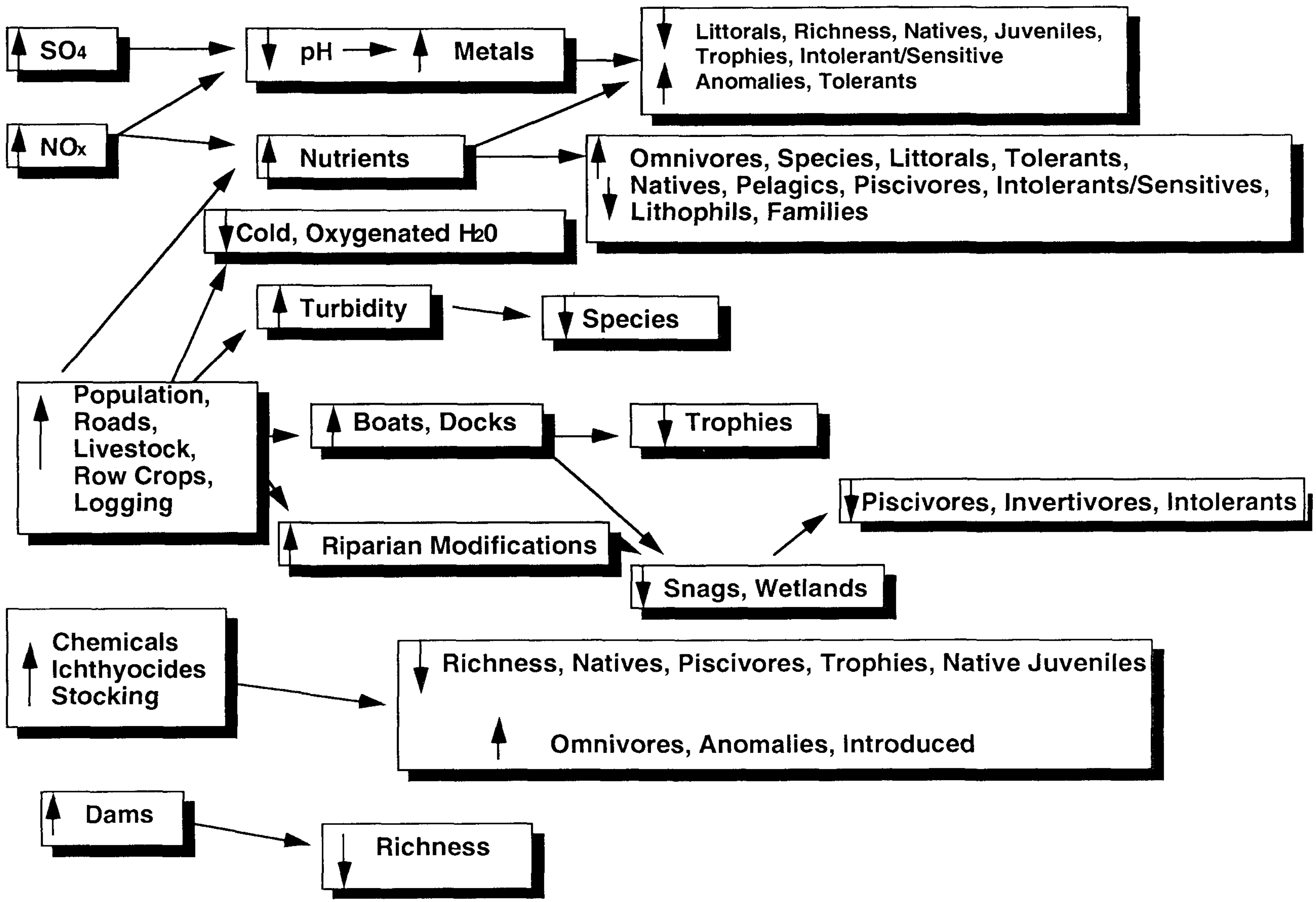


Figure 2-8. An example of a conceptual linkage diagram for fish and biological integrity. Major hazards or stressors applied at the catchment level are at the left, exposure/habitat indicators measured at lakes are in the middle boxes, and candidate response metrics are in the rightmost boxes.

Our goal is to derive a single indicator (from a series of carefully selected assemblages and metrics) to serve as the indicator of biological integrity. However, for the short term, it is more reasonable to choose a series of indicators, with each one representing a lake's condition relative to a particular assemblage. The two options are not mutually exclusive. Thus, a major challenge is to develop and aggregate the set of candidate metrics into an efficient set that extracts enough information from each assemblage to adequately describe the biological integrity of lakes. This process involves evaluating the metrics against the same screening criteria identified in Table 2-10 and evaluating the performance of the metrics in pilot field surveys.

The general process of metric development and evaluation is described in Section 2.4.2, using assemblages as examples. In the 1991 indicator pilot, we focused on two critical criteria for metrics: sensitivity or responsiveness to the disturbance gradients chosen and signal/noise ratio.

Future research on metrics should include lakes experiencing various levels of eutrophication, toxics, acidification, level alterations, and shoreline and riparian modifications. In all these studies, a major concern is in evaluating a metric's sensitivity to disturbance and the clarity of its signal. Such information will aid in the development and selection of a useful set of metrics for assessing condition and diagnosing probable stressors, although it is not conclusive. If a metric does respond as expected, we can be more confident that it will respond to proven stressor gradients. Lack of metric response is one reason for rejecting a candidate metric, unless it is sensitive to other stressors. A large part of a metric's sensitivity and noisiness may be a function of its score interpretation. By rating ranges of metric scores as 5, 3, or 1, Karr et al. (1986), for example, increased the signal clarity of 12 fish assemblage metrics. As with assemblages, further study of metric responsiveness, signal/noise ratio, and the other criteria, particularly costs and variance components, will continue on larger numbers of probability and handpicked lakes in future years.

Some researchers have suggested using only a multimetric index approach; others have encouraged the use of multivariate analysis alone. We plan to use both. It is too early in the indicator development process to choose either path, and each has value. Multivariate analysis is a useful tool for evaluating species or guild structure through use of a single figure and for classifying lake or stream types. Trophic, tolerance, and habitat guilds have proven useful for analyzing assemblages of species whose autecology is fairly well developed (e.g., diatoms, fish, and birds), and species richness and life history characteristics are of intrinsic interest. Each approach also has drawbacks. Multivariate analysis involves considerable subjectivity in its interpretation and axes cannot be interpreted without being associated with other indicators. Multimetric indices require subjective species assignments and considerable calibration; also, metrics are poorly developed for some assemblages. Therefore, we plan to evaluate both multivariate analysis and multimetric indices in the search for indicators that are sensitive to anthropogenic disturbances yet relatively unaffected by natural fluctuations.

Our process for selecting and evaluating candidate metrics for each assemblage initially allowed considerable latitude to the scientists responsible for developing each indicator. General guidance, as described in the previous paragraphs, defined the goals of the process, but the details were left to individual investigators. This process, from which we could all learn, allowed metrics to be developed and defined in diverse ways, within general constraints. The time taken to make data available for each assemblage varied considerably, thus the time for data analysis did, as well. As a result, a variety of metric types and analytical procedures were used, reflecting the investigators' interpretation of the goals as well as their choices of appropriate analytical tools for the task.

As individual investigators developed candidate metrics, they shared the information to assure that individuals didn't deviate too far from the goals, and to propose possible metrics based on those that served other assemblages well. This section describes the metrics proposed for each assemblage and illustrates their sensitivity to the different lake types and to the catchment disturbance gradients.

In the future, we will develop a logically consistent set of metrics across all assemblages, as well as assemblage-specific metrics, and assess various ways of metric selection and evaluation. We expect to do this after a more thorough evaluation of assemblage theory, existing databases, and additional years of EMAP data.

In evaluating the sensitivity of various metrics, the expected condition was a function of the desired direction of change. If low metric scores were considered desirable (e.g., % tolerant species), they were considered unacceptable if they were $\geq$ 2X the minimum for a lake class; marginal scores were those that were 1.5–2X the lowest score. If high scores were desirable (e.g., number of native species), they were rated as unacceptable if they were $\leq$0.5X the highest score in a lake class; marginal ratings were given metrics 0.5–0.75X the high score. These score ratings were considered appropriate by participants in an indicator workshop (Thornton, pers. comm.), but we recognize that they are arbitrary and that metric responses may be nonlinear.

Choosing any arbitrary metric score implicitly suggests that we know what a desirable condition is, regardless of disturbance. This is probably incorrect, but it removes the circularity of evaluating the condition of reference lakes using metrics scored from them. In addition, using arbitrary scores to assess effects reflects our incomplete knowledge about all impacts and the actual or true disturbances in the lake/watershed. At the outset, we also recognized that the sample size was small, hence we could not meaningfully describe the variance associated with these descriptions of the expected condition. However, we use the results from the indicator lakes to illustrate these methods of identifying an expected condition and examining metric sensitivity.

Metric scores were evaluated against each of the four stressor (land and resource use) gradients and lake classes we chose (Section 2.3). The most and least disturbed lakes in each class were examined for substantially different scores, and moderately disturbed lakes were examined for intermediate scores. Making comparisons by lake class avoids developing expectations for considerably different sizes and types of lakes.

2.4.5 Results and Discussion of Indicator Evaluation on Indicator Lakes

This subsection first discusses the results of evaluating each metric and lake class, beginning with diatoms and ending with birds, then compares these results by assemblage.

2.4.5.1 Metric Results and Discussion

In this subsection, each metric is evaluated by lake class, and as before, the four reference lakes are underlined to distinguish them from the others.

Sedimentary Diatoms. Six response metrics (changes in position on a detrended correspondence analysis (DCA) plot, taxa richness, total phosphorus, pH, chloride, and a disturbance index) were evaluated for diatoms. The DCA and richness metrics evaluate the fundamental structure of the assemblage, the chemically oriented metrics capture the habitat tolerance and specialist aspects, and the disturbance index integrates all the metrics. Canonical correspondence analysis (CCA) and weighted averaging regression, based on water quality measurements and diatoms from surface sediments of 66 lakes (19 indicator lakes and 47 probability lakes), were used to calibrate the chemical requirements of 245 diatom taxa. Taxa representing the *range in total phosphorus (TP), chloride, and pH were used to develop diatom-inferred chemical* values. The r^2 between inferred and measured values was 0.77 for TP, 0.66 for chloride, and 0.86

for pH. After further calibrations, the EMAP database and the one generated from other northeastern lakes will be combined to provide more accurate predictions.

Changes in chemical habitat were suggested by several metrics. Diatom-inferred TP indicates an increase for four small warm lakes (AB, MA, TI, and BL), relative to FR, and two large warm lakes (KE and JO), relative to WE, but no cold lakes (Figure 2-9); thus 3 reference lakes and only 5 of 15 disturbed lakes were distinguished. Diatom-inferred chloride change suggests marked increases in BL, TI, MA, and FR relative to AB (small warm lakes), in KE, TO, GR, FA, and JO relative to WE (large warm lakes), in UP and TE relative to RU (small cold lakes), and in DE and HA relative to NU (large cold lakes). Chloride change thus distinguished only 1 reference lake and 10 disturbed lakes. In general, TP and chloride concentrations have increased most in those lakes with high catchment disturbance (Table 2-7; Figure 2-10) and high measured TP and chloride values (Table 2-8). Only one of the indicator lakes (RU) revealed notable acidification, although pH has declined slightly in several of the low TP lakes (Figure 2-11). In general, the pH of lakes with high catchment disturbance has increased, similar to the findings of Cumming et al. (1992) in the Adirondacks. Only 5 disturbed lakes appeared to experience substantial enrichment, despite considerable levels of disturbance in the catchments of 13 lakes. This suggests that (1) the effects of catchment disturbances were not as great as expected, (2) the metrics were not sufficiently sensitive to the disturbances, (3) the diatom assemblages were not affected by the disturbances, or (4) the core was not long enough to provide a presettlement layer, possibly because nutrient enrichment stimulates macrophyte growth and greater sediment accumulation. However, Tables 2-7 to 2-9 indicate considerable catchment disturbance in several of the lakes, and diatoms were affected by nutrient enrichment in 6 lakes and by chloride increases in 13 lakes. Dixit et al. (1992) summarize the sensitivity of diatoms to the nutrient enrichment typically resulting from such disturbances; disturbances that yield more phosphorus and chloride affect diatom assemblages. Because our results clearly indicated that the catchments were disturbed, we assumed that the cores were too short; core length was increased for the 1992 pilot.

A detrended correspondence analysis (DCA) of historical and present diatom flora from the indicator lakes (Figure 2-12) suggests several have experienced anthropogenic or natural changes; these changes in DCA values indicate intrinsic change in assemblage composition. All the small warm lakes (AB, BL, TI, FR, MA) changed considerably between historical and present times, but only BL changed significantly relative to AB. UP, RU, and TE changed more than HU (small cold lakes); HA and NU experienced markedly more change than DE (large cold lakes); and TO FA, JO, KE, and WE changed more than GR (large warm lakes). These changes may have resulted from previously explained responses to chemical changes. As demonstrated in Figures 2-9, 2-10,

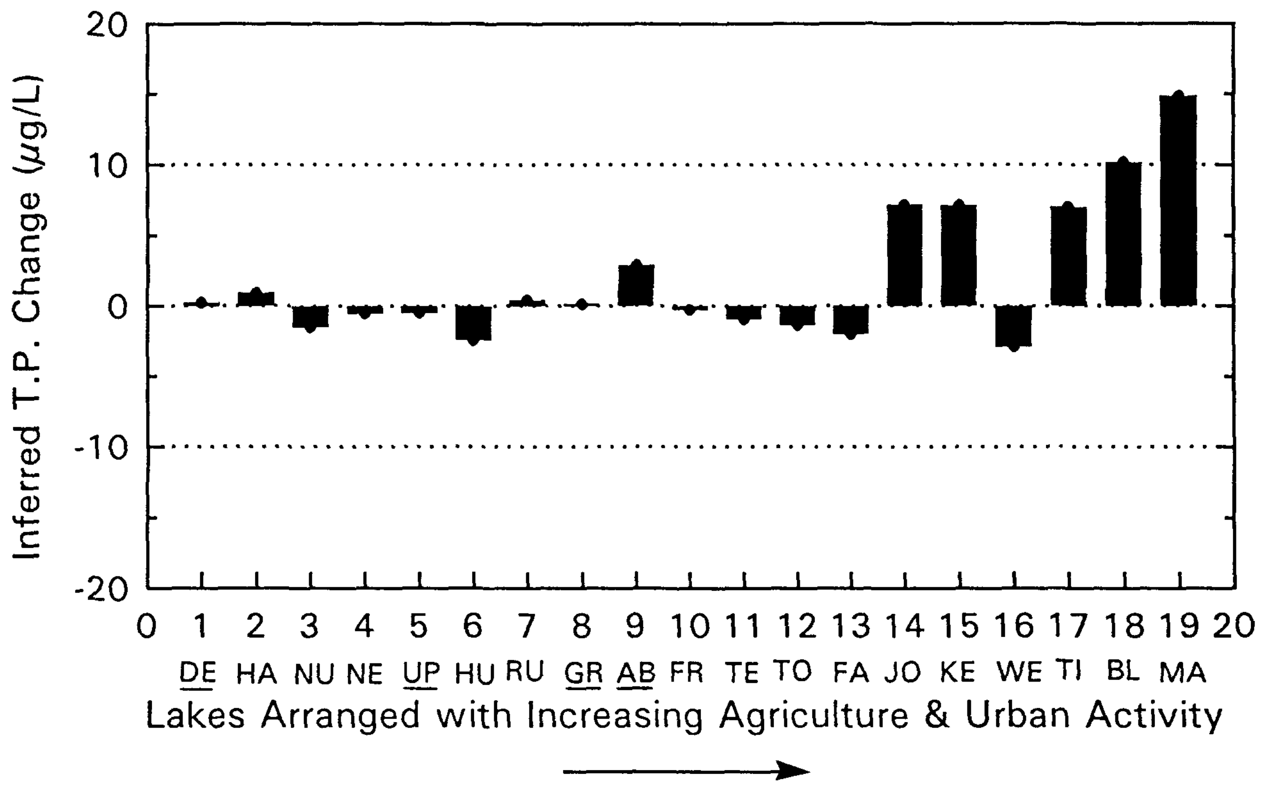


Figure 2-9. Diatom-inferred total phosphorus (TP) change in indicator lakes. Underlined lakes are reference lakes. Agricultural and urban activities were determined from a land use database.

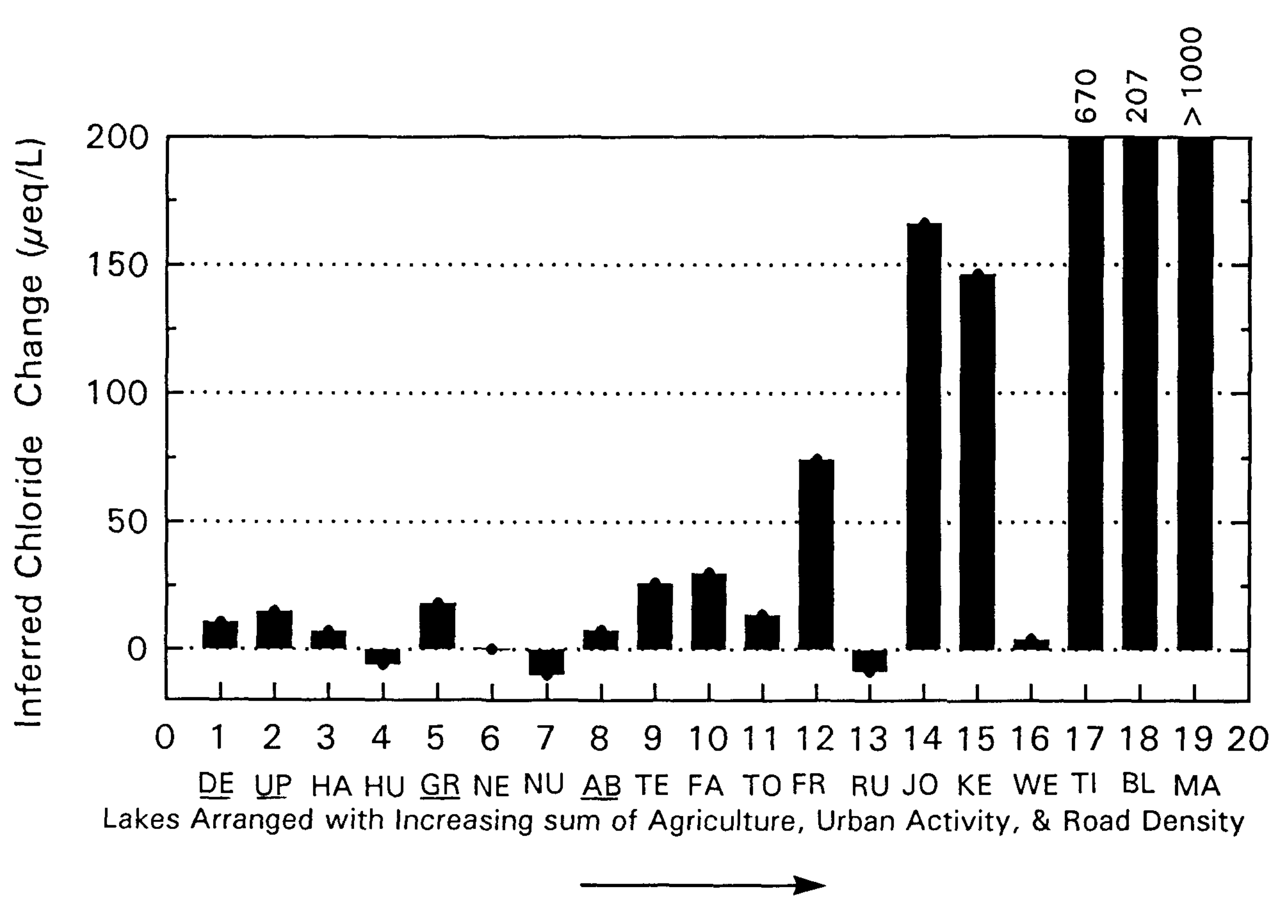


Figure 2-10. Diatom-inferred chloride change in indicator lakes. Underlined lakes are reference lakes. Agricultural and urban activities and road densities were determined from a land use database.

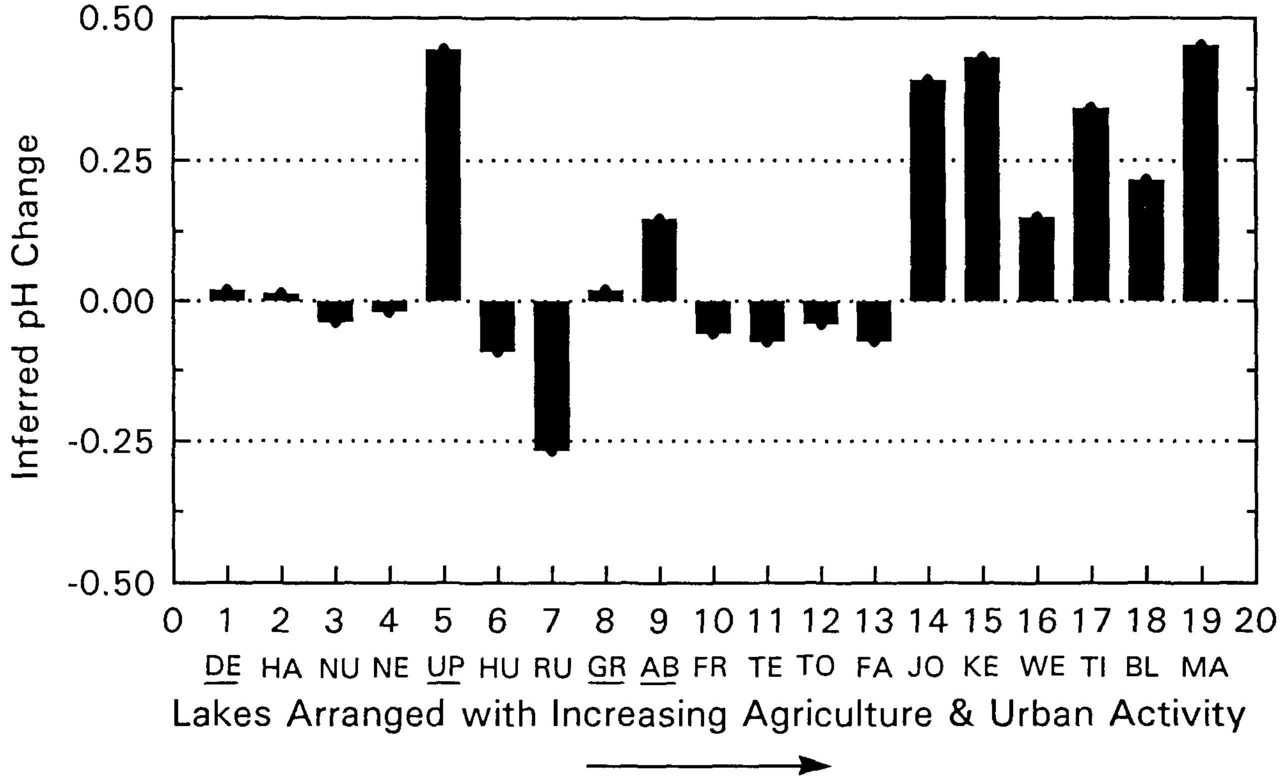


Figure 2-11. Diatom-inferred pH change in indicator lakes. Underlined lakes are reference lakes. Agricultural and urban activities were determined from a land use database.

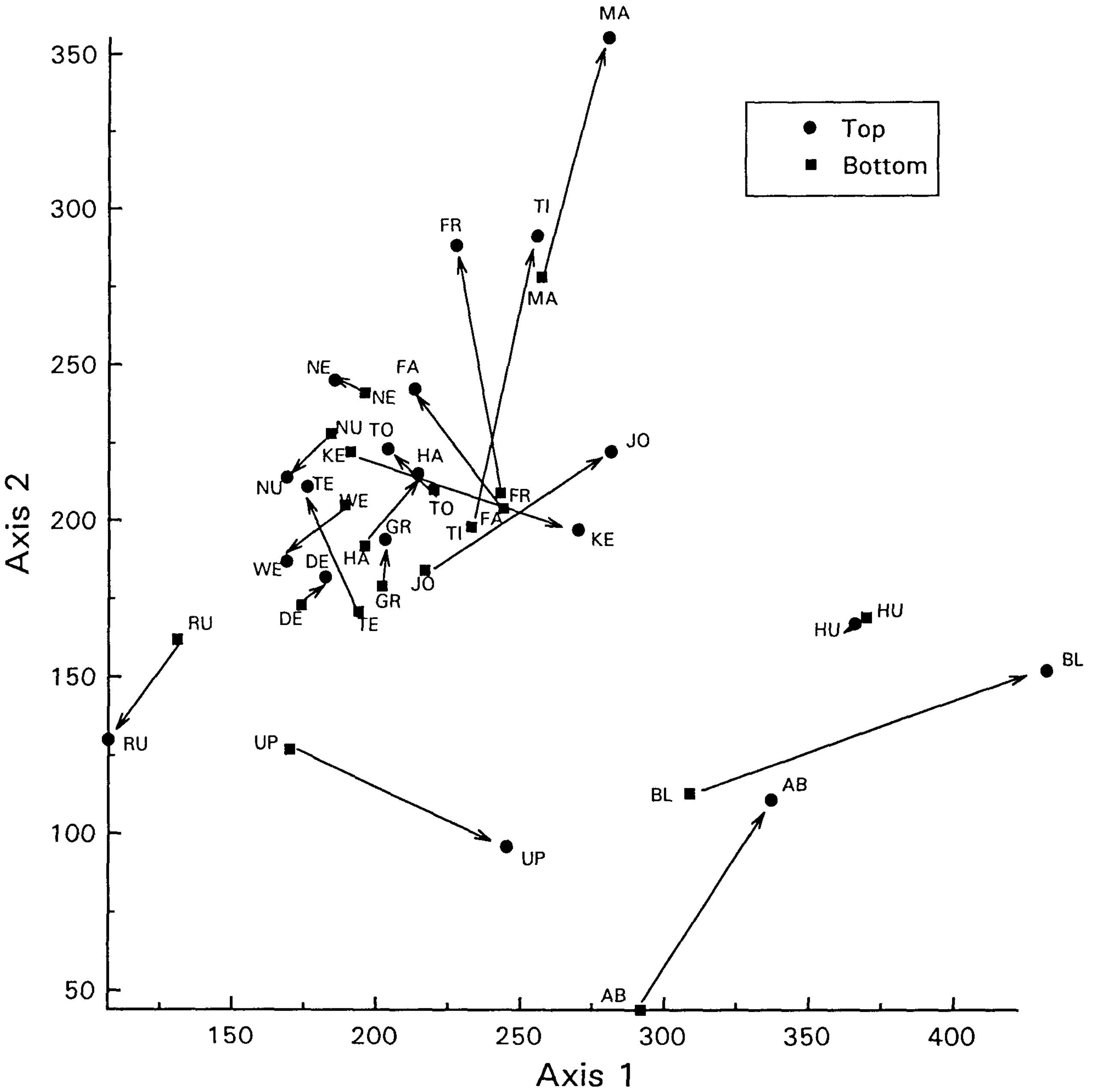


Figure 2-12. Detrended correspondence analysis of diatom assemblages. The line lengths represent the amount of change between the assemblages present at the surface and bottom sections of a 40-cm sediment core.

and 2-11, RU was acidified, and TE, FA, FR, JO, KE, TI, BL, and MA showed considerable catchment disturbance. MA and BL were enriched considerably, and AB, TI, KE, and JO were slightly enriched; although, given their current TP concentrations, it is unlikely these increases substantially altered the flora. This metric distinguished 3 of 4 reference lakes and 10 of 15 disturbed lakes. An even clearer picture may have resulted had longer cores been used. It is also possible that diatom taxa were replaced by others as a result of interspecific competition unrelated to catchment disturbance.

The number of taxa in the top and bottom sections of the sediment cores showed little relationship to lake area or lake volume, presumably because of the small habitat space required by individuals, or what is known as the paradox of the plankton (Hutchinson, 1961). Small habitat requirements are easily met by small or large lakes because both volumes are many orders of magnitude greater than those required by a microscopic organism (Smith, 1950). The top sections of sediment cores from BL, TI, and MA (small warm lakes), TO, JO, KE, and FA, (large warm lakes), and NU and NE (large cold lakes) all contained markedly fewer taxa than the maxima for their lake classes (Table 2-12). Therefore, taxa richness distinguished all 4 reference sites and 9 of 15 disturbed sites. The change in taxa richness between bottom and top sections was much lower than the maximum in each class for all lakes but AB (small warm lake), JO (large warm lake), RU and TE (small cold lakes), and DE (large cold lake). Richness change distinguished 2 of 4 minimally disturbed reference lakes and 12 of 15 disturbed lakes, and was therefore slightly more affected by the level of watershed disturbance than was present species richness.

Change in the diatom disturbance index revealed a number of disturbances similar to the number for richness change (Table 2-12). Index change was substantially greater in BL, FR, TI, and MA than in AB (small warm lakes), in GR, FA, JO, and KE than in TO (large warm lakes), in UP, TE, and HU than in RU, and in HA, NU, and NE compared to DE. The change in index values was able to distinguish 2 of 4 reference lakes and 12 of 15 lakes with disturbed catchments.

The diatom results indicate the need to (1) ensure that the cores are long enough to collect presettlement diatom assemblages and (2) date the bottom layers to substantiate presettlement times. The lengths of our cores ranged from 30 to 40 cm, which was long enough to recover presettlement sediments from oligotrophic lakes, but not from all eutrophic systems. In areas of the country like New England that were deforested in the 18th and 19th centuries (Thompson, 1853; Marsh 1864), the major enrichment of lakes may have occurred at times earlier than those represented by the bottom layers of our cores (some as recent as 125 years ago). If this is the case,

Table 2-12. Sedimentary Diatom Taxa Richness and Disturbance Index of 19 Indicator Lakes

Lake	Top		Bottom		Change	
	Taxa[a]	Index[b]	Taxa[a]	Index[b]	Taxa[a]	Index[b]
Small Warm						
AB[c]	64	2.5	52 *	2.3	+12	0.2
BL	38 *	4.5 *	56	2.0	−18 •	2.5 •
FR	53	3.0	59	2.2	−6 •	0.8 •
TI	42 *	3.7 *	69	2.3	−27 •	1.4 •
MA	29 •	4.8 *	34 •	3.7 *	−5 •	1.1 •
Large Warm						
GR[c]	96	1.7	85	1.5	+11 *	0.2 •
TO	72 *	1.8	79	1.7	−7 •	0.1
FA	48 •	2.7 *	57 *	2.0	−9 •	0.7 •
JO	72 *	3.0 *	55 *	2.0	+17	1.0 •
KE	64 *	3.2 *	60 *	1.7	+4 •	1.5 •
WE	75	1.8	69	1.7	+6 •	0.1
Small Cold						
UP[c]	82	1.8 *	89	1.7	−7 •	0.1 •
HU	64	2.7 •	63 *	2.8 •	+1 •	−0.1 •
RU	84	1.0	67 *	1.3	+17	−0.3
TE	81	1.8 *	64 *	1.8	+17	0.0 •
Large Cold						
DE[c]	91	1.7	79	1.7	+12	0.0
HA	76	1.8	73	1.7	+3 •	0.1 •
NE	36 •	2.3	60 *	2.0	−24 •	0.3 •
NU	44 •	2.2	48 *	2.0	−4 •	0.2 •

[a] * = 0.5–0.75X the maximum for a lake class; • = ≤ 0.5X the maximum.
[b] * = 1.5–2X the minimum for a lake class; • = ≥ 2X the minimum.
[c] Underlined lakes are minimally disturbed reference lakes.

the bottom layers we collected may depict disturbed, rather than reference, conditions where lakes, watersheds, and riparian zones have been recovering and are now largely forested. Enriched lakes may require cores 60 to 100 cm long.

Although well established as a diagnostic indicator and often used to evaluate historical change, this assemblage requires more research before indicators can be chosen. Future work will involve additional literature review and data analyses to develop and evaluate metrics such as metal tolerance/intolerance, benthic/pelagic, turbid/clearwater, coldwater/ warmwater, oligotrophic/ eutrophic, generally sensitive/insensitive species, and organic and toxic tolerance/intolerance. In place of DCA, we will examine assemblage similarity indices such as Pinkham and Pearson's B, which Boyle et al. (1990) found to be the most sensitive, stable, and consistent among 16 commonly used numerical indices. Work will also continue on refining the diatom disturbance index.

Zooplankton. Three metrics (species richness, the ratio of large to small zooplankton, and the number of trophic links) were evaluated for this assemblage, and 2 new rotifer species and 12 new range records for New England were revealed in this pilot. Species richness was of interest for its own sake, the large/small ratio was developed to assess size classes, and trophic links were examined to evaluate trophic specialization and complexity. Species richness tended to increase with lake area and often with temperature and disturbance (Figure 2-13). Because of this relationship, expected scores for zooplankton (and later, fish and bird) species richness were estimated by extending a line through the top scores of the small and large lakes, similar to the method recommended by Karr et al. (1986). AB, TI, MA, and BL supported fewer species than FR (small warm lakes); GR had markedly fewer taxa than FA (large warm lakes), RU had markedly fewer taxa than TE (small cold lakes), and NE had markedly fewer taxa than DE (large cold lakes). This metric was affected by the disturbances at 5 of 15 disturbed lakes and distinguished 2 of 4 reference lakes. This lack of sensitivity to watershed disturbances and fish stocking may partly explain why zooplankton species richness is not commonly used as an indicator.

Ordinations of zooplankton assemblages showed that individual body size was correlated with much of the variation among lakes. A preliminary metric was developed based on the ratio of macro- to micro-zooplankton (Table 2-13). Substantially lower values were obtained for AB, TI, FR, and MA compared with BL (small warm lakes). Among large warm lakes, KE, TO, GR, JO, and WE had substantially lower values than FA. RU and UP had substantially lower ratios than TE (small cold lake). The ratios for DE and NU were markedly lower than for NE (large cold lakes). Zooplankton size ratio detected perturbation in 9 of 15 disturbed lakes, but distinguished none of the minimally disturbed lakes from the others. Although zooplankton size structure is known to be responsive to nutrient enrichment (Tessier and Horwitz, 1990) and fish predation (Brooks and Dodson, 1965), it was not sensitive to watershed development or fish stocking. This may have occurred because watershed changes did not sufficiently increase nutrient levels or because stocking did not change planktivory rates.

The number of trophic links in the crustacean food web (Table 2-13) was found to be no more revealing of the effects of catchment disturbance on zooplankton. AB and MA both had markedly

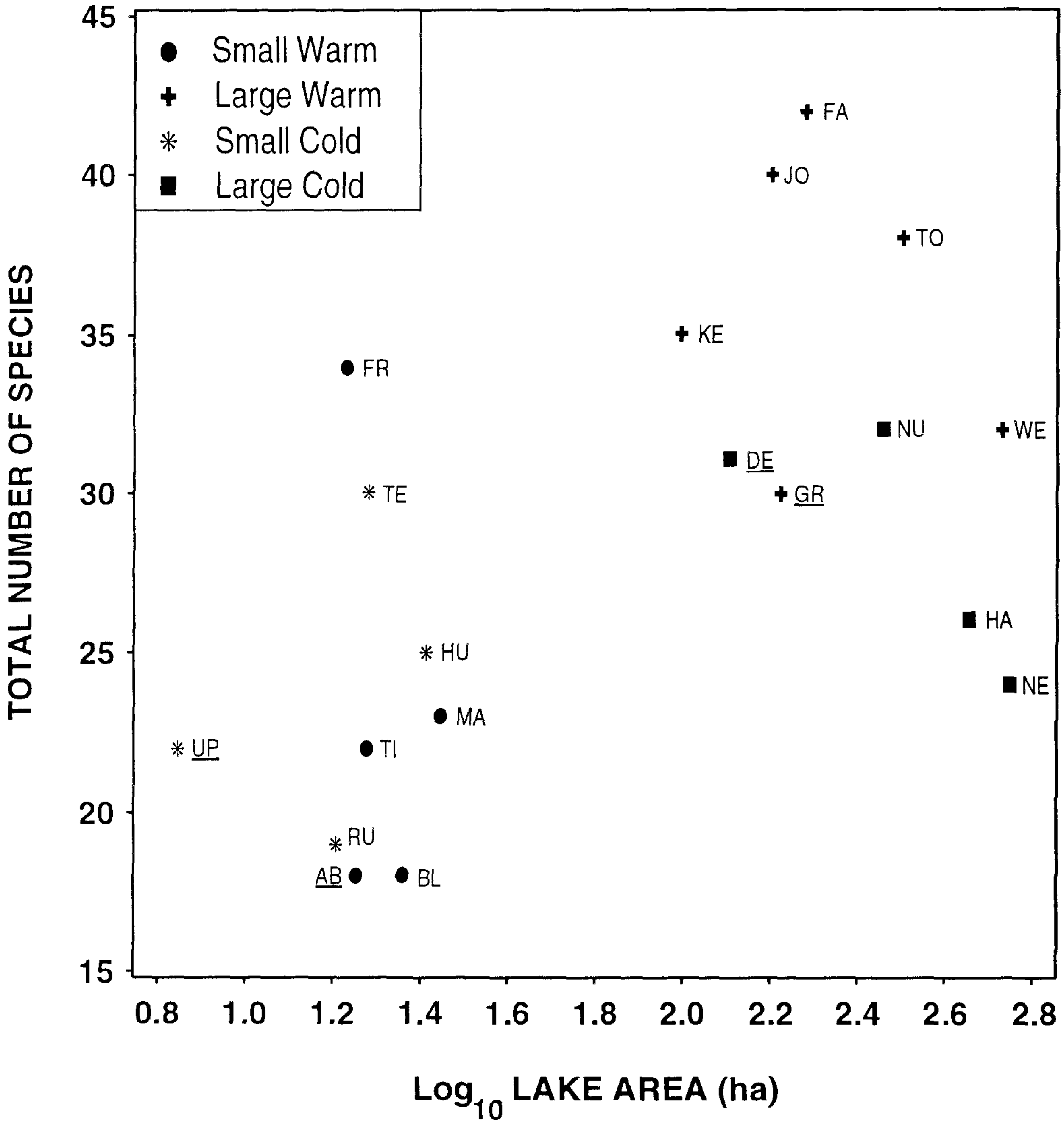


Figure 2-13. **Zooplankton species richness versus lake area. Underlined lakes are reference lakes.**

Table 2-13. Ratio of Large to Small Zooplankton and Number of Trophic Links at 19 Indicator Lakes

Lake	Ratio[a]	Trophic Links[a]
Small Warm		
AB[b]	1.94 ●	27 *
BL	7.89	39
FR	1.97 ●	45
TI	1.46 ●	37
MA	0.79 ●	24 *
Large Warm		
GR[b]	3.30 ●	45
TO	2.38 ●	53
FA	12.71	51
JO	3.98 ●	50
KE	0.23 ●	28 *
WE	0.58 ●	33 *
Small Cold		
UP[b]	3.60 *	46
HU	5.27	21 ●
RU	2.45 *	12 ●
TE	6.80	32 *
Large Cold		
DE[b]	6.08 *	63
HA	8.01	65
NE	10.21	35 *
NU	2.68 *	69

[a] * = 0.5–0.75X the maximum per lake class; ● = ≤ 0.5X the lake class maximum.

[b] Underlined lakes are minimally disturbed reference lakes.

fewer links than FR (small warm lakes), whereas WE and KE supported markedly fewer than TO (large warm lakes). All three disturbed small cold lakes (HU, RU, TE) had substantially fewer links than UP, but only NE among the large cold lakes had a marked reduction in links. The trophic link metric was affected by the disturbances at 7 of 15 disturbed lakes and distinguished 3 of 4 reference lakes, thus it was the most sensitive zooplankton metric. The zooplankton assemblage clearly requires considerably more metric development before it will be useful for assessing acceptable biological integrity. Continued research on trophic and thermal guilds have potential and the stimulatory effects of both fish predation and nutrient enrichment on microzooplankton require further study.

Benthos. A set of metrics based on those in Plafkin et al. (1989) for streams was selected, evaluated separately, and combined into a Biotic Index (BI). The metrics included taxa richness and two redundancy measures: mean number of individuals/taxon and % of individuals contributed by the dominant taxon. Higher scores for the redundancy metrics and BI indicate disturbance (Table 2-14). No relationship was evident between taxonomic richness and lake area, probably because the greater habitat complexity of larger lakes is not reflected in profundal sediments, which tend to be homogeneous (Forbes, 1925). Among small warm lakes, AB, TI MA, and FR had substantially lower richness than BL. KE, FA, GR, JO, and WE all supported fewer than half the taxa of TO (large warm lakes). HU, RU, and TE had fewer species than UP (small cold lake). DE, HA, and NU contained fewer taxa than NE (large cold lake), possibly because the large amounts of sunken logging debris provided additional habitat in NE. This metric showed the effects of disturbance in 12 of 15 disturbed lakes, but distinguished only 1 minimally disturbed lake out of 4.

Among the small warm lakes, BI scores were substantially higher for FR, TI, and MA, relative to AB; there was no marked difference in scores among the large warm lakes. HU and TE had higher BI scores than UP (small cold lakes), and NU and NE had markedly higher scores than DE. Although the Biotic Index distinguished 4 of 4 minimally disturbed reference lakes, it indicated the effects of disturbance in only 7 of 15 disturbed lakes.

Both redundancy metrics demonstrated a similar pattern. The mean number of individuals/taxon was substantially higher for BL, FR, TI, and MA than for AB (small warm lakes), for TO, FA, KE, and WE than for JO (large warm lakes), for HU compared with TE (small cold lakes), and for DE, NE, and NU compared with HA (large cold lakes). The % of individuals as the dominant taxon was greater in BL, FR, TI, and MA than in AB (small warm lakes), in TO, FA, KE, and WE than in JO (large warm lakes), in HU compared with RU (small cold lakes), and in DE, NE, and NU compared with HA (large cold lakes). Both metrics distinguished 3 of 4 reference lakes and 11 of 15 disturbed lakes.

Thus, species richness and the two redundancy measures distinguished the expected effects of lake disturbance levels more often than the more complex biological index. Further work will include evaluation of other index components, such as % of particular taxonomic groups and % intolerants. We will also examine various options for interpreting scores for the metrics and indices to increase their usefulness.

Table 2-14. **Benthic Macroinvertebrate Metric Scores for 19 Indicator Lakes**

Lake	Taxa Richness[a]	Biotic Index[b]	Mean Number of Individuals per Taxon[b]	% Individuals as Dominant Taxon[b]
Small Warm				
AB[c]	15 •	23	2.5	16
BL	41	31	23.8 •	54 •
FR	8 •	48 •	13.4 •	49 •
TI	13 •	38 *	8.5 •	28 *
MA	3 •	49 •	6.0 •	56 •
Large Warm				
GR[c]	24 •	25	4.9	16
TO	67	33	19.2 •	59 •
FA	29 •	30	8.6 •	32 •
JO	24 •	26	3.5	14
KE	20 •	25	26.3 •	59 •
WE	18 •	36	54.3 •	80 •
Small Cold				
UP[c]	43	23	9.4	30
HU	16 •	36 *	13.4 *	36 *
RU	25 *	29	7.3	21
TE	12 •	37 *	7.2	28
Large Cold				
DE[c]	18 *	19	3.0 *	26 •
HA	13 *	23	1.9	12
NE	25	28 *	8.6 •	41 •
NU	15 *	29 *	9.7 •	46 •

[a] * = 0.5–0.75X the maximum for a lake class; • = ≤ 0.5X the lake class maximum.
[b] * = 1.5–2X the minimum for a lake class; • = ≥ 2X the lake class minimum.
[c] Underlined lakes are minimally disturbed reference lakes.

Fish. Five metrics were examined for this assemblage: total and native species richness, % generally intolerant species, % omnivores, and % piscivores. All five metrics were suggested by Fausch et al. (1990) as generally appropriate indicators and they have been used successfully in large rivers (Hughes and Gammon, 1987; Oberdorff and Hughes, 1992).

Fish species richness and number of native species were directly related to lake area (Figures 2-14 and 2-15), as predicted from the theory of island biogeography (MacArthur and Wilson,

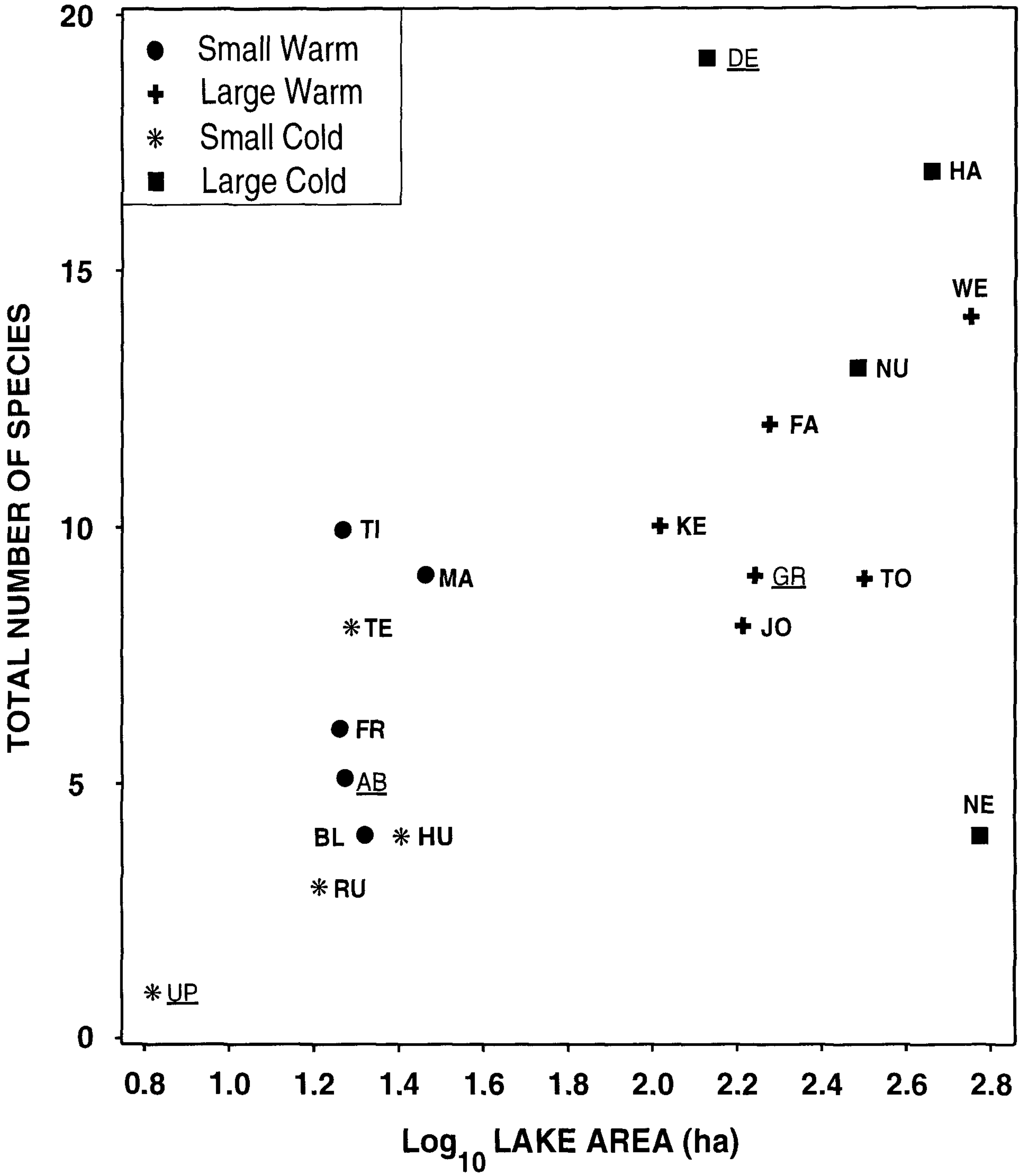


Figure 2-14. Total fish species richness versus lake area. Underlined lakes are reference lakes.

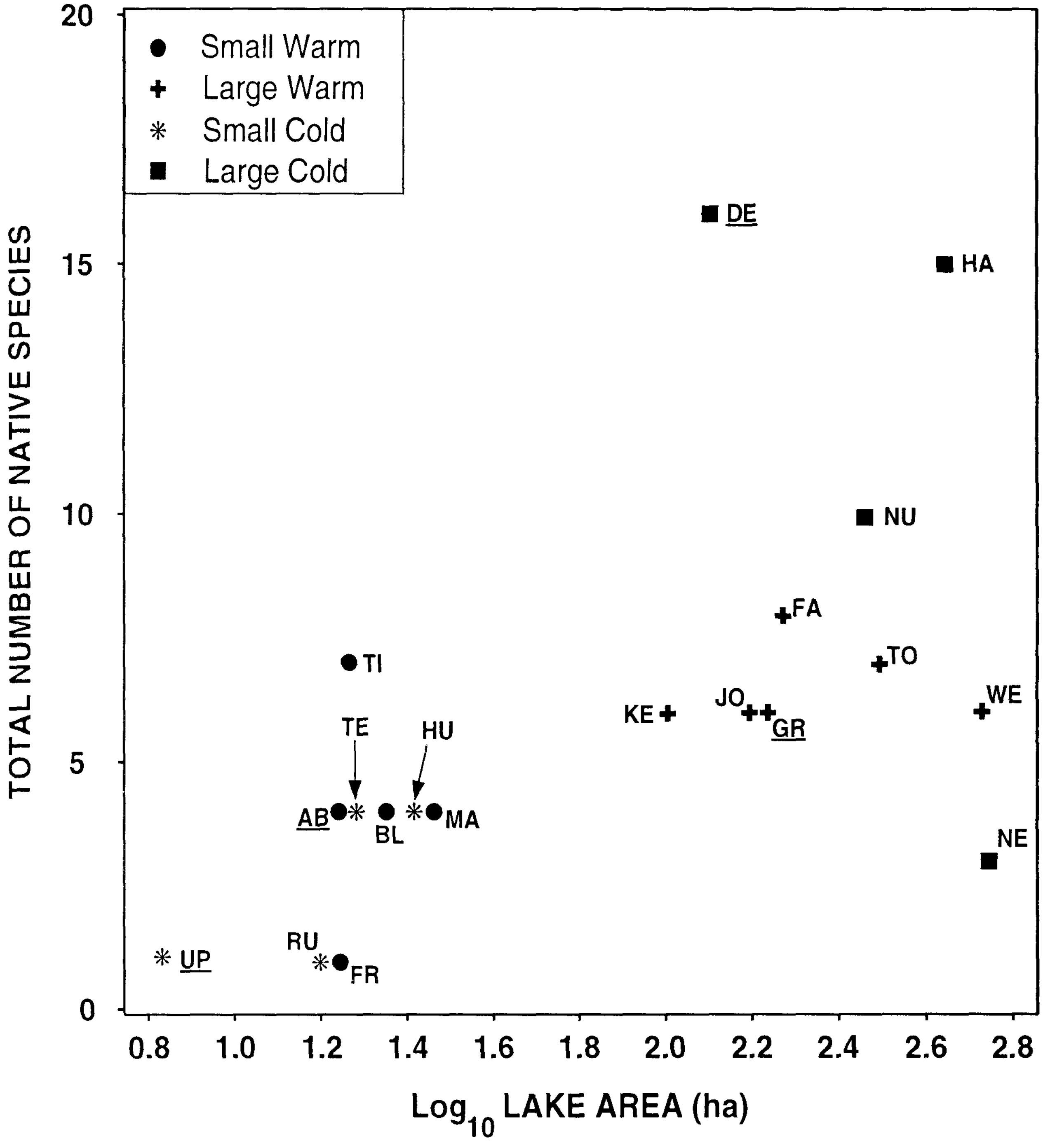


Figure 2-15. Species richness of native fishes versus lake area. Underlined lakes are reference lakes.

1967; Barbour and Brown, 1974), in which lakes function as islands in the terrestrial landscape. NU, HA, and NE had substantially fewer species than DE (large cold lakes). Among small cold lakes, UP, HU, and RU had fewer species than TE. AB, FR, and BL supported markedly fewer species than TI (small warm lakes). GR, TO, and JO held fewer fish species than expected (large warm lakes). The species richness metric detected 9 of the 15 disturbed lakes, but only 1 of the 4 minimally disturbed lakes. The number of native species metric performed no better. HA, NE, and NU had fewer natives than DE (large cold lakes); UP and RU had fewer native species than TE (small cold lakes); AB, BL, FR, and MA supported fewer native fish species than TI (small warm lakes); and GR, JO, and TO held fewer natives than FA (large warm lakes). The number of native species metric distinguished 10 of 15 disturbed lakes but only 1 of 4 reference lakes. These results differ from those found in large rivers, where species richness (Gammon, 1991; Oberdorff and Hughes, 1992) and native species richness (Hughes and Gammon, 1987) were sensitive to disturbance.

Fish species tolerances were determined largely from descriptions in Scott and Crossman (1973). The % of fish species *intolerant* of warming, turbidity, and low DO was substantially lower in AB, TI, FR, and MA than in BL (small warm lakes) (Table 2-15). In large warm lakes, the % of intolerants was markedly lower in KE, TO, GR, JO, and WE than in FA. In the small coldwater lakes, HU, RU, and TE had a substantially lower % of intolerant species than UP. NE had higher % intolerant scores than DE and NU (large cold lakes). The % of intolerant fish species distinguished 11 of 15 disturbed lakes, but only 1 of 4 reference lakes. Hughes and Gammon (1987) and Oberdorff and Hughes (1992) found the number of intolerant fish species much more sensitive than this to chemical and physical habitat disturbances in large rivers.

Trophic guilds of fishes were determined largely from species descriptions in Scott and Crossman (1973). The % of omnivorous species was substantially higher in BL, FR, TI, and MA than in AB (small warm lakes), in GR, TO, JO, and KE than in WE (large warm lakes), in HU and TE than in UP (small cold lakes), and in DE, HA, and NU compared to NE (large cold lakes). The % omnivores metric thus distinguished 11 of 15 disturbed lakes, but only 1 of 4 reference lakes.

The % of piscivorous fish species was markedly lower in BL and TI than in AB (small warm lakes), in HU, RU, and TE than in UP (small cold lakes), and in HA and NE than in NU (large cold lakes). There were no marked differences in % piscivores among large warm lakes. The % piscivores metric distinguished only 7 of 15 disturbed lakes, but all 4 reference lakes. Calculating trophic guild percentages is admittedly less sensitive when species are used rather than individuals;

Table 2-15. Fish Metric Scores for 19 Indicator Lakes

Lake	% Intolerant Species[a,b]	% Omnivorous Species[c,d]	% Piscivorous Species[a,e]
Small Warm			
AB[f]	0 •	0	50
BL	50	20 •	20 •
FR	0 •	33 •	50
TI	9 •	28 •	27 *
MA	0 •	22 •	44
Large Warm			
GR[f]	11 *	22 *	44
TO	10 *	20 *	50
FA	17	17	50
JO	13 *	25 *	50
KE	0 •	20 *	50
WE	8 •	13	47
Small Cold			
UP[f]	100	0	100
HU	50 •	17 •	16 •
RU	0 •	0	33 •
TE	13 •	25 •	38 •
Large Cold			
DE[f]	22 •	16 •	32
HA	40	25 •	25 *
NE	50	0	25 *
NU	31 •	14 •	36

[a] * = 0.5–0.75X the maximum for a lake class; • = ≤ 0.5X the lake class maximum.
[b] Intolerant fish: brook trout, fallfish, lake chub, lake trout, lake whitefish, pearl dace, slimy sculpin.
[c] * = 1.5–2X the minimum for a lake class; • = ≥ 2X the lake class minimum.
[d] Omnivores: northern redbelly dace, brown bullhead, yellow bullhead, common carp, white sucker, fathead minnow.
[e] Piscivores: brook trout, brown trout, lake trout, rainbow trout, black crappie, largemouth bass, smallmouth bass, yellow perch, chain pickerel, burbot, white perch, American eel.
[f] Underlined lakes are minimally disturbed reference lakes.

however, Hughes and Gammon (1987) and Oberdorff and Hughes (1992) found that the % of piscivorous and omnivorous individuals was also relatively insensitive to disturbance in large rivers.

Although the fish results are preliminary, it appears that all five metrics are roughly comparable in their ability to distinguish the effects of watershed disturbance and fishery management practices.

Karr et al. (1986) predicted for streams that an intolerance metric would be most effective at the highest levels of biological integrity, but that trophic guilds would be sensitive from the middle to high levels. We did not find either case true. Additional metric development will focus on habitat guilds, reproduction metrics, and various options for interpreting metric scores in order to develop a multimetric fish assemblage index. Improved indexing methods will also allow the use of metrics based on % and number of individuals, as well as % and number of species.

Birds. Four bird assemblage metrics were examined through use of the pilot data: species richness, number of individuals generally tolerant and intolerant, and principal components analysis factor 1 derived from an evaluation of 65 potential metrics. The number of species was correlated with lake area (r^2 = 0.744) and warmwater lakes supported more species than cold (Figure 2-16), except for MA (small warm lakes) and WE (large warm lakes). There were no markedly lower numbers of species among lakes in the small or large coldwater classes. Bird species richness distinguished all 4 reference lakes, but only 2 of 15 disturbed lakes. In forested areas, an increase in species number may actually indicate disturbance, because forest openings and development make new habitats available to a collection of species preferring brushy, grassland, or agricultural habitats.

Principal components analysis was used to investigate how the foraging, dietary, and migratory guild metrics interacted. Factor 1 of a PCA accounted for 67% of total variance and was correlated with % seed eating species (–0.90), % seed eating/omnivorous/ground gleaning species (–0.90), % ground gleaning species (–0.89), % foliage gleaning species (0.78), % foliage gleaning individuals (0.76), % insectivorous species (0.76), and % insectivorous individuals (0.70). Factor 1 scores were negatively correlated with riparian disturbance and showed great discriminating power for all but the large coldwater lakes (Table 2-16). BL, TI, FR, and MA had lower scores than <u>AB</u> (small warm lake); TO, FA JO, KE, and WE had lower scores than <u>GR</u> (large warm lakes); HA and NU had lower scores than NE (large cold lakes); and HU, TE, and RU had lower scores than <u>UP</u> (small cold lakes). PCA factor 1 distinguished 14 of 15 disturbed sites and all 4 reference sites. Croonquist and Brooks (1991) also found avian guilds useful indicators of anthropogenic impacts in riparian areas.

To develop another metric describing avian response to human impacts, we reviewed the literature and compiled a list of New England bird species that were either tolerant or intolerant of human disturbance on their nesting territories. To ensure a clear signal, we avoided using wetland dependent species and adaptable or ubiquitous species. Interior forest birds that were

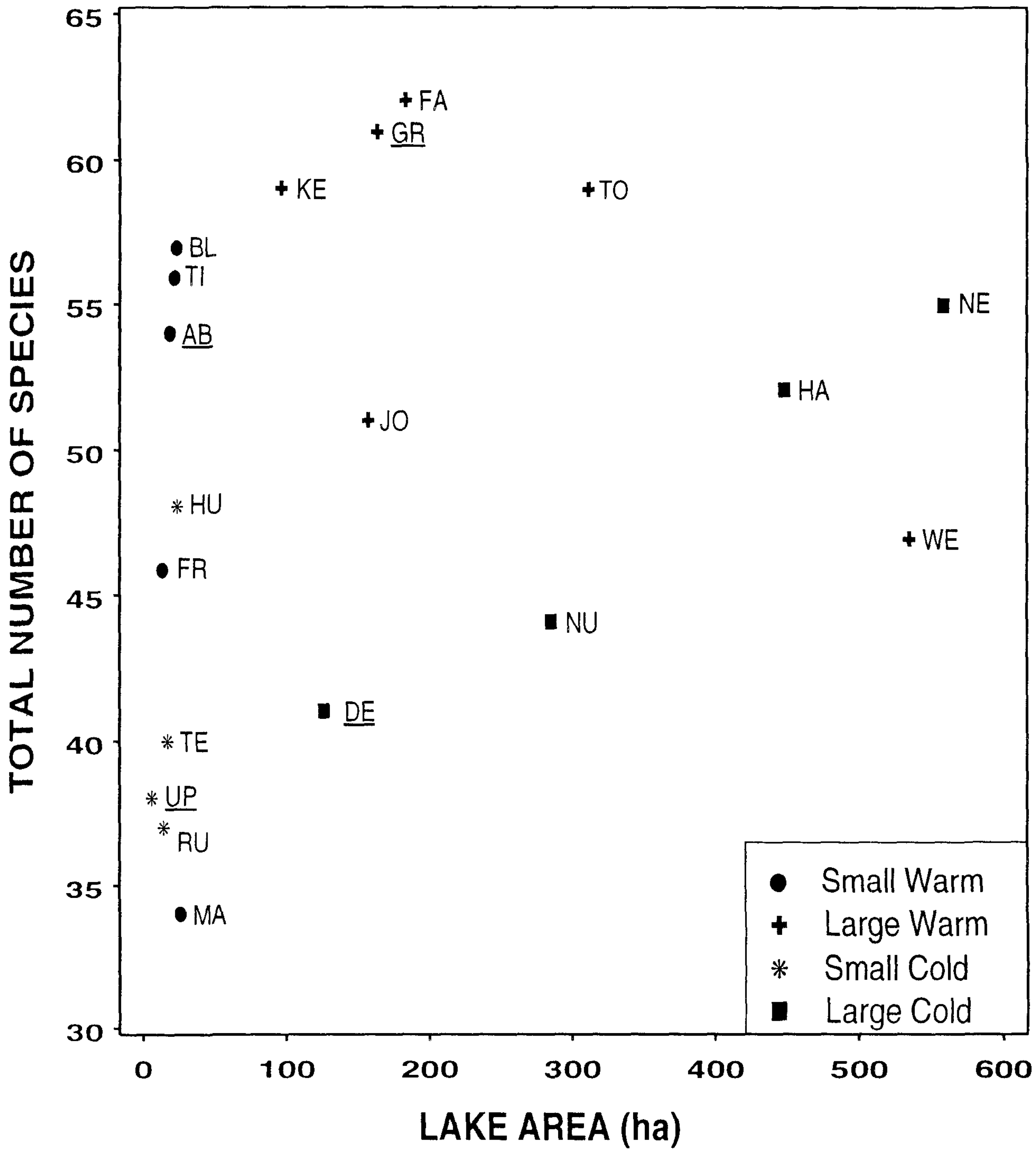


Figure 2-16. Bird species richness versus lake area. Underlined lakes are reference lakes.

Table 2-16. **Principal Components Analysis Scores and Number of Tolerant and Intolerant Birds at 19 Indicator Lakes**

Lake	Factor 1[a]	Number of Birds	
Small Warm		**Tolerant[b]**	**Intolerant[c]**
AB[d]	0.4	50	6 •
BL	0.0 •	97 *	34
FR	–0.7 •	126 •	10 •
TI	–0.8 •	105 •	2 •
MA	–2.6 •	155 •	0 •
Large Warm			
GR[d]	0.3	86 •	9 *
TO	0.1 •	32	12
FA	–0.2 •	74 •	8 *
JO	–0.9 •	105 •	6 •
KE	–1.1 •	94 •	3 •
WE	–1.5 •	189 •	0 •
Small Cold			
UP[d]	1.5	16	50
HU	0.9 *	43 •	25 •
RU	0.7 •	15	12 •
TE	0.6 •	64 •	4 •
Large Cold			
DE[d]	0.8	8	44
HA	0.6 *	13 *	10 •
NE	0.9	71 •	15 •
NU	0.2 •	27 •	25 *

[a] PCA factor 1 was positively correlated with foliage gleaning and insectivorous species and individuals; it was negatively correlated with ground gleaning, seed eating, and omnivorous species. * = 0.5–0.75X the lake class maximum; • = ≤ 0.5X the lake class maximum.

[b] Tolerant birds: American goldfinch, American robin, brownheaded cowbird, bobolink, barn swallow, chipping sparrow, common grackle, chestnut-sided warbler, European starling, field sparrow, house finch, house sparrow, house wren, magnolia warbler, Nashville warbler, northern mockingbird, rock dove, Savannah sparrow, white-throated sparrow, yellow-throated vireo. * = 1.5–2X the lake class minimum; • = ≥ 2X the lake class minimum.

[c] Intolerant birds: blackburnian warbler, boreal chickadee, black-throated green warbler, Cape May warbler, common loon, evening grosbeak, golden-crowned kinglet, gray jay, hairy woodpecker, hermit thrush, pine siskin, pileated woodpecker, ruby-crowned kinglet, solitary vireo. * = 0.5–0.75X the lake class maximum; • = ≤ 0.5X the lake class maximum.

[d] Underlined lakes are minimally disturbed reference lakes.

intolerant of habitat changes and seral birds preferring clearcut or developed habitats were identified in the database.

Birds, unlike fish, could be sampled quantitatively and with one consistent method; thus, we could use data based on numbers of individuals in place of numbers of species. The numbers of individuals/lake that were generally intolerant of physical and biological disturbance, or forest interior birds common to old growth forest, were lower than expected for AB, TI, FR and MA compared with BL (small warm lakes), for KE, FA, GR, JO, and WE compared with TO (large warm lakes), for HU, RU, and TE compared with UP (small cold lakes), and for HA, NE, and NU compared with DE (large cold lakes) (Table 2-16). The number of tolerant birds was higher than expected at BL, TI, FR, and MA than at AB (small warm lakes), at KE, GR, FA, JO, and WE than at TO (large warm lakes), at HU and TE than at RU (small cold lakes), and at HA, NU, and NE than at DE (large cold lakes). Number of intolerant individuals distinguished 2 of 4 reference lakes and 13 of 15 disturbed lakes, whereas number of tolerants distinguished 3 reference and 13 disturbed lakes.

Bird tolerance and ordination (foraging guilds) metrics better distinguished lake disturbance and reference sites than did species richness. Evidently, knowledge of species' habitat and foraging requirements is more useful than simple taxonomic identity. Patrick (1972) also described several stream assemblages with consistent species richness, but variable species composition, through time. Future work on the bird data will focus on evaluating the PCA factor 1 component metrics separately and on developing an assemblage index composed of trophic, habitat, and species richness metrics.

Physical Habitat Structure. We examined the relationships between the physical habitat structure and vertebrate assemblages. Data on canopy complexity and extent and intensity of anthropogenic shoreline disturbances were aggregated into a lake ranking of 1 to 19 and plotted against the % of intolerant bird species (Figure 2-17a). A similar plot was developed for fish, including fish cover complexity in the habitat ranking, and substituting the % native fish species for birds (Figure 2-17b). Both plots reveal a direct relationship between habitat quality and the metric ratios. We propose to further analyze habitat/exposure data in order to evaluate response indicator metrics, to develop diagnostic models for explaining unacceptable condition, and to distinguish probable cause as chemical, physical, or both, as suggested by Plafkin et al. (1989).

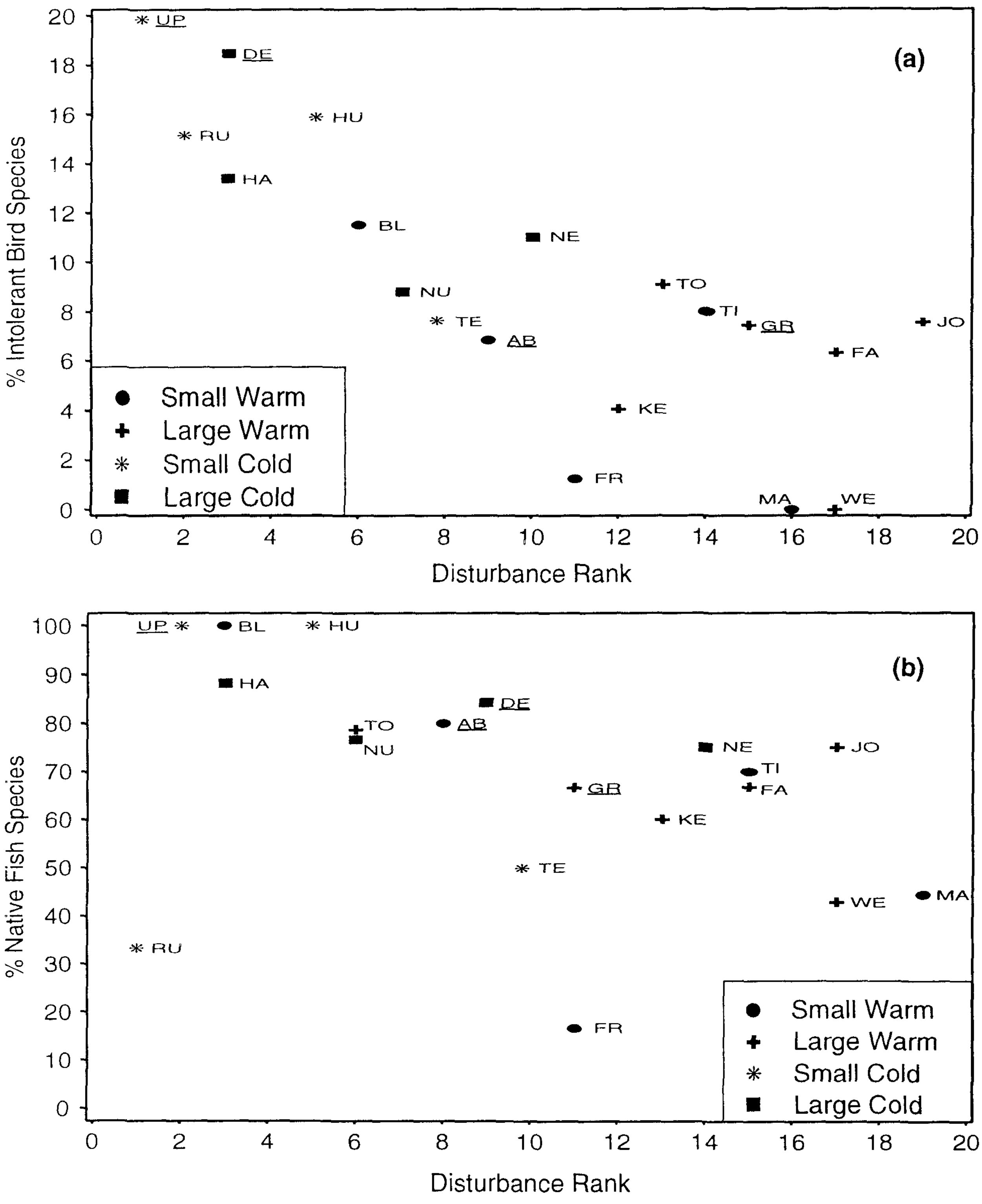


Figure 2-17. Relationship between quality of physical habitat structure and (a) % intolerant bird species, and (b) % native fish species. Physical habitat quality was ranked 1 through 19 by level of canopy complexity and extent and intensity of shoreline disturbance in (a) and by all three plus fish cover complexity in (b). Underlined lakes are reference lakes.

2.4.5.2 Metric and Assemblage Comparisons from Indicator Lake Results

The relative scores for the reported metrics of the biological response indicators are summarized in Table 2-17. Metric scores were rated as explained in Section 2.4.4 and given a black dot (•) if unacceptable and an asterisk (*) if marginal. The dots and asterisks were tallied by metric and assemblage to evaluate the responsiveness of the candidate metrics and assemblages to catchment perturbation. When metrics indicated different perturbation of the reference and disturbed lakes than predicted, it may have been due to erroneous classification of lake type or disturbance level, to inappropriate response indicator metrics or score interpretation, or to the fact that no single metric is equally sensitive to all perturbations. The last explanation has been reported elsewhere (Angermeier and Karr, 1986; Karr et al., 1987; Oberdorff and Hughes, 1992; Figure 2-9) and justifies using a multimetric index. The misclassification of disturbance level and lake type are discussed in Section 2.5.2, along with the ability of the combined set of metrics to assess predicted lake condition. Our objective in this section of the report is to evaluate individual metrics and assemblages for their responsiveness.

None of the metrics or assemblages consistently scored the minimally disturbed and more highly disturbed lakes, but some were affected more by the disturbances than others. Nine metrics (diatom DCA change, diatom disturbance index change, diatom richness change, benthos richness, benthos individuals/taxon, benthos % as dominant taxon, bird PCA factor 1 scores, and bird number of tolerant and intolerant individuals) scored as expected in 14–18 of the 19 lakes, suggesting that they could potentially detect the effects of such disturbances in other lakes at least 75% of the time. The diatom enrichment metric, all three zooplankton metrics, and the fish and bird richness metrics all scored as expected in only 10 or fewer cases out of 19, indicating a success rate of approximately 50%, too low to be useful for detecting the effects of such disturbances as studied in this pilot. This result suggests that when used alone, without further interpretation, these metrics are unlikely to respond to land use and fish stocking activities as much as to other habitat or exposure conditions. It also suggests that they should be used with care when combined with other metrics. Clearly, additional research is needed to further develop and refine metrics and their scoring interpretation.

We assessed assemblage responsiveness by examining how well the set of metrics for an assemblage distinguished disturbed from reference lakes (Table 2-17). We carried out this procedure for disturbed lakes by tallying the lakes at which half the metrics or more for an assemblage showed the effects of disturbance. For reference lakes, we tallied the lakes at which

Table 2-17. Assemblage, Metric, and Lake Scores for 19 Indicator Evaluation Lakes[a]

		SMALL WARM					LARGE WARM						SMALL COLD				LARGE COLD			
Indicator	Tally	<u>AB</u>	BL	FR	TI	MA	<u>GR</u>	TO	FA	JO	KE	WE	<u>UP</u>	HU	RU	TE	<u>DE</u>	HA	NE	NU
Diatoms	2, 12																			
DCA change	3, 10		*					*	•	•	•	*	•		•	•		•		*
Chloride change	1, 10		•	•	•	•	•	•	•	•	•		•			•	•	•		
Enriched	3, 5	•	•		•	•				•	•									
Richness change	2, 12		•	•	•	•	*	•	•		•	•	•	•				•	•	•
Disturbance index change	2, 12		•	•	•	•	•		•	•	•		•	•		•		•	•	•
Zooplankton	2, 6																			
Richness	2, 5	*	*		*	*	*								*				*	
Size Ratio	0, 9	•		•	•	•	•	•		•	•	•	*		•		*			•
Trophic links	3, 7	*				*					*	*		•	•	*			*	
Benthos	3, 12																			
Richness	1, 12	•		•	•	•	•		•	•	•	•		•	*	•	*	*		*
BI	4, 7			•	*	•								*		*			*	*
Individuals/taxon	3, 11		•	•	•	•		•	•		•	•		*			*		•	•
% as dominant taxon	3, 11		•	•	*	•		•	•		•	•		*			•		•	•
Fish	2, 12																			
Richness	1, 9	•	•	*			*	*		*			•	•	•			*	•	*
Natives	1, 10	*	*	•		*	*	*		*		*	*		•			*	•	*
% Intolerant	1, 11	•		•	•	•	*	*		*	•	•		•	•	•	•			*
% Omnivore	2, 11		•	•	•	•	*	*		*	*			•		•	•	•		•
% Piscivore	4, 7		•		*									•	•	•		•	•	
Birds	3, 14																			
Richness	4, 2					*						*								
PCA-1	4, 14		•	•	•	•		•	•	•	•	•		*	•	•		*		•
# Intolerant	2, 13	•		•	•	•	*		*	•	•	•		•	•	•		•	•	*
# Tolerant	3, 13		•	•	•	•	•		•	•	•	•		•		•		*	•	•
Total Impact		8	13	15	15	16	9	9	10	12	14	11	6	12	10	11	6	10	11	12
Acceptability		A	M	M	M	U	A	A	A	A	M	A	A	U	M	M	A	M	M	U

[a] Underlined lakes are reference lakes. Depending on the desirable direction of change (see text), * = 1.5–2X the lowest score, or 0.5–0.75X the highest score for a lake class; • = ≥ 2X the lowest score, or ≤ 0.5X the highest score in a lake class. Tallies are added by the *'s and •'s for the metric or assemblage for the 4 reference and 15 disturbed lakes. Impact is determined by adding, with *'s = 0.5, •'s = 1.0, and the total rounded up to the nearest whole number. A = acceptable, M = 1.5X lowest total, U = 2X lowest total.

fewer than half the metrics showed the effects of disturbance. Based on this method of interpretation, birds, benthos, fish, and diatom assemblages distinguished reference from disturbed lakes in 14–17 of the cases, or at least 75% of the time. The zooplankton assemblage distinguished the lakes only 42% of the time. This difference in sensitivity among assemblages may have resulted from the relative level of metric development, the method by which the assemblages were evaluated, or the types of lake disturbance. Further metric and index development is expected to improve the sensitivity. Development of a multimetric index for each assemblage may improve score interpretation. Also, evaluation of different types of disturbance in future pilots may reveal different assemblage sensitivities.

2.4.5.3 Indicator Evaluation Conclusions

These results indicate the need for five types of additional research: (1) Continued metric development is needed for all five assemblages to complete the evaluation of listed metrics and to suggest and evaluate additional metrics. (2) Additional reference lake selection and study is necessary to examine the amount of assemblage variation expected in minimally disturbed conditions. This information is needed both for indicator development and for defining acceptable lake conditions. (3) Disturbance gradients and lake classes more likely to affect zooplankton should be evaluated to test their sensitivity, because the 1991 disturbances may have been biased toward the other assemblages. (4) It is also probable that apparent disturbance levels of the indicator lakes may differ from actual disturbance and that there are multiple disturbances. We plan additional data assessments this year to examine these possibilities. (5) Candidate metrics must be evaluated for indexing variance, which is a combination of crew, measurement, index period, and index station variances. This was the major focus of the 1992 pilot, which will facilitate evaluation of among-lake versus within-lake variance for diatoms, zooplankton, fish, fish tissue contamination, and habitat structure. Indexing variance for benthos and birds will be conducted in the 1993 demonstration study.

2.5 DISTINGUISHING ACCEPTABLE (NOMINAL) FROM UNACCEPTABLE (SUBNOMINAL) CONDITIONS

As mentioned earlier, EMAP's primary objective is to describe the condition (as measured with the selected indicators) and trends in condition of ecological resources. However, the information will be useful to a wider audience if we can set criteria to distinguish between acceptable (nominal) and unacceptable (subnominal) conditions for particular lake or stream types within selected geographic locations. Hunsaker and Carpenter (1990) and Messer et al. (1991) outline the issue. Messer et al. state: "Operational criteria must be developed for each response indicator to identify the transition from acceptable or desirable (nominal) to unacceptable or undesirable (subnominal) condition...Criteria could be based on attainable conditions under 'best management practices' as observed at regional reference sites..., or on theoretical grounds or management targets." Clearly, this is not strictly a technical issue. We emphasize that our objective in approaching this issue is not to make this decision about thresholds independent of others, but rather to contribute to the process of deciding how to resolve these issues from a technical perspective. We are not advocating any particular management decision or action or societal decision. We do believe, however, that it is important to develop as sound a scientific basis as possible for such decisions.

This section describes options for assessing acceptable and unacceptable condition, demonstrates how reference lakes could be used for such purposes, and discusses some concerns regarding that approach.

2.5.1 Describing the Reference Condition

We will report the condition of populations of surface waters by plotting cumulative distribution functions (CDFs) and histograms from metric scores or from indices that combine multiple metrics and assemblages. However, EMAP is also charged with estimating the proportion of waterbodies that are in acceptable or unacceptable condition (Hunsaker and Carpenter, 1990; Messer et al., 1991). This task requires some sort of benchmark or reference condition against which the sample waterbodies can be compared. Several approaches have been suggested for estimating reference condition: reference sites, pristine sites, historical data, paleoecological data, experimental results, ecological models, empirical distributions, and expert consensus. We describe the advantages and disadvantages of each in the following paragraphs.

Regional reference sites (Hughes et al., 1986; Karr et al., 1986) incorporate historical information, assemblage-habitat knowledge, and expert consensus in selecting waters of major waterbody types that are minimally disturbed by humans. Allan et al. (1992) called for increased monitoring and preservation of minimally disturbed waters in order to distinguish various sources of variability, particularly the effects of climate change. The assumption is that such lakes and streams represent conditions that incorporate the natural limits of the region and that lower levels of watershed, riparian, and in-lake disturbance support less impaired or healthier communities.

Major disadvantages of reference site monitoring are the additional monitoring costs required, the effort needed to locate minimally disturbed sites dispersed across a region, and questions of their representativeness and acceptable level of disturbance in extensively disturbed regions. The representativeness of reference sites for populations of diverse surface waters is also of concern. Representativeness can be examined by using ordination analyses to test whether the habitat indicators for grid and reference sites are similar when both have similar response indicator values. Such comparisons are more problematic for the many sites at which response indicator values differ from those of reference sites. In such cases, habitat indicator values plotted against response indicator values for sites in the same ecoregion are useful (Plafkin et al., 1989). Despite these shortcomings, reference sites remain EMAP's primary method for estimating reference conditions (Paulsen et al., 1991).

Pristine sites are a special case of reference site with the same advantages and disadvantages, except they are presumably undisturbed. However, they do not exist in most regions and are not representative of waters in markedly different regions. Given the global extent of atmospheric pollution, the prospects of continued global climate change (Ehrlich et al., 1977), and the prevalence of fish stocking, there may be no truly pristine sites.

Historical data offer a perspective of species once present across a region. This approach requires no additional sampling and can occasionally provide data on waters at the time of settlement. Such data are available from some natural history museums or the grey literature. There are several problems with historical data. For many regions of the country and many assemblages, the sites selected are sparse and the data are incomplete or were collected with markedly different methods and during various seasons. Information about the level of disturbance at the collection sites is rare. Typically, historical data sets offer little information on species abundances and many collections focus on particular species, rather than entire assemblages. Occasionally, museum curators are reluctant to release such data and the grey literature is particularly troublesome to locate and examine. Although EMAP may not be able to use such data everywhere to determine unacceptable condition, historical data will be very useful in range finding and classification, and should be acquired and analyzed wherever possible.

Paleoecological information preserved in lake sediments is another form of historical data. If cores are long enough, they offer estimates of presettlement or preindustrial water quality conditions (Cumming et al., 1992). We propose to use sedimentary diatoms for this purpose. A key assumption in this approach is that historical sediments represent the desired or equilibrium state, but additional retrospective research, in addition to studying diatom assemblages preserved at the top and bottom of the sediment cores, is required to determine this.

Experimental results from laboratory studies are useful for assessing toxic levels and mechanisms; however, they are unreliable for determining reference conditions for entire communities. The strength of experimental tests is in reducing stressor and response variables, but this limits their applicability in systems composed of multiple stressors, multiple assemblages with individuals of multiple ages, and various types of habitats.

Models developed from historical, paleoecological, experimental, and EMAP-Surface Waters grid data could be very useful for developing standards for determining acceptability. Such models should be increasingly useful as EMAP databases increase in size. However, models developed from incomplete data or from data collected at disturbed sites can yield only incomplete or undesirably low acceptability standards.

Empirical distributions of EMAP results may show substantial discontinuities, or we could simply select an index score arbitrarily. This approach would apply just the data from probability sites and is least expensive, but its arbitrary nature is politically troublesome; also, discontinuities may not appear or they may simply indicate different lake and stream types. There is a high probability that even the highest quality probability sites may be impaired where diffuse pollution is extensive (e.g., in intensively farmed regions such as the Corn Belt, Central California, or the Mississippi Alluvial Plain).

Expert panels could be formed to seek consensus on index scores that would be acceptable or unacceptable. Such reference conditions would be the product of their professional expertise and regional experience, the quality of the available data, and personal biases. Panelists would require information on reference condition from the other approaches. A similar process has recently been used to estimate the effects of varying forest harvest levels on threatened and endangered species of Pacific Northwest forests.

Regardless of how reference data are acquired, there will be a range of scores, so we will be left with the dilemma of determining the indicator score at which a waterbody is acceptable or unacceptable. Box plots of reference versus grid data and nonparametric tests such as the Wilcoxon rank test can be used to test for significant differences from reference conditions. We could also set an arbitrary percentile of the reference data. For example, Ohio EPA (1988) uses the 25th percentile, focusing attention on the most impaired waters, while Plafkin et al. (1989) delimit impairment at the 80th percentile, where waters are slightly impaired. Hughes and Noss (1992) demonstrate how these differences may result in stream impairment rates of 64% versus 95% in highly disturbed regions. Because EMAP is also a trends monitoring program, we can also simply track trends in a direction we consider undesirable (e.g., loss of species, increase of pest species, increased anomalies).

2.5.2 Pilot Results

A reference lake or gradient approach to comparing assemblage (and the derivative metric) responsiveness is based on the assumptions that (1) lakes most lacking human disturbances represent systems with greater biological integrity than those with moderate or substantial disturbances and (2) subtle natural differences in lake morphology and chemistry among lake classes in an ecoregion are of less consequence than the level of human disturbance. A similar approach has been successfully used for streams (Hughes et al., 1986; Karr et al., 1986; Ohio EPA, 1988; Plafkin et al., 1989). We also chose to compare assemblage and metric responsiveness along a disturbance gradient, to determine how this process might work for lakes. We examined the results to see if all metrics together distinguished between lakes with disturbed and undisturbed catchments. We emphasize that this is one process by which we intend to develop and select indicators, and that we will need much more information than that resulting from this pilot before we can do so. The stressor gradient and reference lake design enabled us to (1) examine the process for determining acceptable condition from the combined set of assemblages and metrics and (2) evaluate difficulties in selecting reference lakes.

The acceptable/unacceptable status of the indicator lakes was evaluated by adding the metric scores from Table 2-17 for each lake (dot = 1, asterisk = 0.5) and ranking a lake as unacceptable or marginal if its tally was ≥ 2X or 1.5–2X that of the lake with the lowest total in the same lake class, respectively. These criteria represent the uncertainty in assessment given the single reference lake in each class, and were selected as being tallies that several workshop participants agreed were substantially different from expected values (Thornton, pers. comm.). Given these criteria, lakes with substantially impaired biological integrity are MA, NU, and HU. Marginal lakes are BL, FR, TI, KE, RU, TE, HA, and NE. The only lakes in acceptable condition are <u>AB</u>, <u>GR</u>, TO, FA, JO, WE, <u>UP</u>, and <u>DE</u>. These lake evaluations demonstrate the manner by which a set of carefully selected reference lakes can be used to determine whether or not a probability lake is in acceptable condition. However, a lake that is unacceptable or marginal may be so because of natural events or anthropogenic activities.

The lake tallies for entire communities, if not individual assemblages, generally confirmed the state biologists' preclassifications of lake disturbance and the land use databases. Reference lakes had low sums, highly disturbed lakes in a lake class had high sums, and moderately disturbed lakes usually had intermediate values. Exceptions were KE, WE, and HU. We also had difficulty selecting a large warm reference lake that was truly minimally disturbed, as shown by the high level of shoreline disturbance at GR (Table 2-7). This problem indicates the need to select several reference lakes, both to obtain a measure of reference lake variance and to be able to reject candidate reference lakes disturbed to such a degree that they are unsuitable. In ecoregions with high levels of human disturbance, the search effort for reference lakes must be increased. In addition, we must develop a method of ranking various catchment and riparian disturbances differently and obtain more accurate land use data. Clearly, before reference lakes are finalized, they must be more completely evaluated than was the case for this pilot.

The pilot indicates a process for using preliminary sets of reference lakes and assemblages to assess whether similar types of lakes are in acceptable or unacceptable condition. However, lake types other than those evaluated in this pilot may need examination. For example, EMAP will sample shallow 1–5 ha ponds, for which other metrics may be needed, and some lakes will not fall neatly into distinct cold/warm lake classes. For example, three warmwater lakes (FA, JO, and FR) supported a small coldwater zooplankton fauna, but apparently their water was not cold and oxygenated enough for trout reproduction. We might also develop lake classes from their community characteristics, as Tonn et al. (1983) and Schneider (1981) did for fish assemblages. This would have to be done for the entire community, not a single assemblage; it is, therefore, unlikely to be accomplished without considerably more data and interpretation. Finally, numerous reference lakes are needed for confident estimates of the reference condition for all the grid lakes and lake classes examined in the northeastern United States. Walters et al. (1988) showed that for several management questions, agencies need one reference site for every four to five modified sites to distinguish spatial and temporal trends. This would mean sampling 10-15 reference sites annually in the northeastern United States. During the 4-year EMAP rotation, the 40–60 carefully selected reference lakes sampled could equate to 8–20 for 3–5 ecoregions or 8–12 for 5 lake classes. Of course, some of the reference lakes will come from the probability sample, as long as their catchments are minimally disturbed. Preliminary estimates suggest that about half can be located this way.

In summary, we found that we could locate minimally disturbed lakes to serve as references for estimating acceptable and unacceptable condition. The process requires numerous conversations with local biologists, examination of remote sensing data, and site visits. Minimally disturbed warm water lakes will require more effort to locate than cold water lakes because of the greater rarity of the former in the Northeast.

2.6 SUMMARY AND CONCLUSIONS

2.6.1 Indexing

The indicator pilot fulfilled several research objectives. We were successful in developing an indexing protocol in evaluating sampling gear for all assemblages except benthos. We determined that indexing protocols based on a systematic sample of the littoral and riparian zones were appropriate for fish, birds, and physical habitat structure. No single gear type or index location was adequate for sampling fish; sampling was stratified by littoral and pelagic habitat, and different types of gear were needed in each. Benthos indexing methods require further data analyses and pilot studies to determine number and locations of samples and appropriate gear. Additional research is also needed on the effectiveness of lightweight electrofishers.

2.6.2 Indicator Development

Preliminary metrics enabled us to assess the effects of four different types of common disturbance (silviculture, agriculture, residential development, fish stocking) along an intensity gradient. Several metrics were responsive to the above disturbances, particularly changes in diatom species richness, a diatom disturbance index, diatom DCA change, benthos richness, two benthos redundancy measures, the bird PCA factor 1 score and the numbers of tolerant and intolerant birds. When multiple metrics of an entire assemblage were assessed, birds, benthos, fish, and diatoms were most affected by the disturbances studied. Bird data were the most quickly available for analysis, and bird and fish assemblages were responsive to changes in physical habitat structure.

2.6.3 Acceptable/Unacceptable Condition

The aggregated metrics appeared sensitive to the disturbance gradients of agriculture, residential development, fish stocking rates, and silviculture. In other words, they were able to distinguish acceptable, marginal, and unacceptable conditions for four lake classes and types of disturbance through use of a single reference lake in each class. The assemblages in the large warmwater reference lake appeared more affected by disturbances than those in another lake in the same class with greater catchment disturbance. This suggests a need to improve the accuracy of the landscape database or the response indicators or both.

2.6.4 Concluding Remarks

The process of selecting, developing, evaluating, and rejecting indicators is a continuum with numerous iterations and few clear breaks (Figure 2-8). We plan to determine the most appropriate assemblages, given the EMAP objectives, and to develop indicators from measurements on those assemblages. These assemblage indicators in turn must be integrated into a multi-assemblage or community assessment, which requires indicators to be complementary, without obvious contradiction, and to have minimal redundancy. However, some redundancy is useful in diagnosing cause and effect relationships when the changes are subtle. In the case of RU, for example, fish, benthos, zooplankton, and diatoms all indicated acidification, but any single assemblage would have been unconvincing. Insufficient biological evidence of impact to counter apparently acceptable chemical concentrations is a concern for low ANC lakes, because midlake summer chemistry underestimates acidification pulses that occur in the littoral area during spring snowmelt (Eshleman, 1988; Gubala et al., 1991).

The analyses described in this chapter focused on calibrating metric response to known disturbance gradients, but EMAP must also be able to assess unknown disturbances. To do this, we must know more about the basic nature of assemblage change, particularly changes in species richness, guilds, and processes that indicate fundamental shifts in assemblage integrity. We should also be tracking taxa likely to be sensitive to climate change and continued land use changes. We must develop a clearer understanding of what we believe is unacceptable change or else set an arbitrary index score. This task requires identifying the most important changes in biological integrity, trophic condition, and fishability about which enlightened citizens have concerns. In addition, our reports must clearly state that unacceptable lake integrity does not mean that such lakes are biological wastelands, they are simply unacceptable deviations from reference conditions.

Further metric development is necessary, particularly for zooplankton. A promising area for zooplankton is trophic complexity. Benthos metric development will focus on score interpretation. Bird and fish metric development will continue to focus on various guilds, and we are analyzing three large fish databases to further estimate sampling variability.

Further evaluation of six essential criteria for indicator selection (responsiveness, high signal/noise, high information/cost, high societal value, high among-site/within-site variability, low among-year variability) will be necessary in the Northeast and elsewhere before we decide which assemblages are best for national monitoring. Indicator assemblages will be evaluated against these criteria for several years in the northeastern United States before any final decisions are made for that region. Additional pilot studies in other regions are needed before the final set of assemblages is selected for nationwide monitoring.

Future demonstration projects will expand our determination of variance components (Chapter 4). Reasonable estimates of the year component require about 5 years of sampling. These variance studies will also provide data for quantifying signal/noise and cost/information. Cost estimates will be refined once the indexing protocols for all assemblages are decided, probably in 1994, after the results of the methodological research for benthos and fish are available.

Another critical aspect of indicator development involves coordination with other federal and state agencies. The U.S. Geological Survey and U.S. Fish and Wildlife Service are both developing new national monitoring initiatives. The interest of the U.S. Forest Service, Bureau of Land Management, Bureau of Reclamation, and Army Corps of Engineers in monitoring ecological conditions continues to expand. Each agency has different information needs that sound similar when expressed at a general level. State agencies are required by the U.S. EPA to develop biological criteria based on aquatic assemblages for surface waters. Although each program has a different set of objectives, and consequently a different design, it would be valuable to each agency if a common set of core indicators and sampling methods were adopted. An intergovernmental task force on water quality monitoring is currently convening to promote just such coordination.

APPENDIX 2A

ANALYTES TO BE MEASURED IN FISH TISSUE

Analyte (CAS Number)	Detection Limits (ppm)
Aldrin (309-00-2)	0.00025
Aluminum (7429-90-5)	10.0
Arsenic (7440-38-2)	2.0
Cadmium (7440-43-9)	0.2
Chlordane-cis (5103-71-9)	0.00025
Chromium (7440-47-3)	0.1
Copper (7440-50-8)	5.0
2,4' -DDD (53-19-0)	0.00025
4,4'-DDD (72-54-8)	0.00025
2,4'-DDE (3424-82-6)	0.00025
4,4'-DDE (72-55-9)	0.00025
2,4'-DDT (789-02-6)	0.00025
4,4'-DDT (50-29-3)	0.00025
Dieldrin (60-57-1)	0.00025
Endrin (72-20-8)	0.00025
Heptachlor (76-44-8)	0.00025
Heptachlor Epoxide (1024-57-3)	0.00025
Hexachlorobenzene (118-74-1)	0.00025
Hexachlorocyclohexane [Gamma-bhc/Lindane] (58-89-9)	0.00025
Lead (7439-92-1)	0.1
Mercury (7439-97-6)	0.01
Mirex (2385-85-5)	0.00025
Nickel (7440-02-0)	0.5
trans-Nonachlor (3765-80-5)	0.00025
PCB Isomers	
2,4-Dichlorobiphenyl (34883-43-7)	0.001
2,2',5-Trichlorobiphenyl (37680-65-2)	0.001
2,4,4'-Trichlorobiphenyl (7012-37-5)	0.001
2,2',5,5'-Tetrachlorobiphenyl (35693-99-3)	0.001
2,2',3,5'-Tetrachlorobiphenyl	0.001
2,3',4,4'-Tetrachlorobiphenyl	0.001
2,2',4,5,5'-Pentachlorobiphenyl (37680-73-2)	0.001
2,3',4,4',5-Pentachlorobiphenyl (31508-00-6)	0.001
2,2',4,4',5,5'-Hexachlorobiphenyl (35065-27-1)	0.001
2,3,3',4,4'-Pentachlorobiphenyl	0.001
2,2',3,4,4',5-Hexachlorobiphenyl (35065-28-2)	0.001
2,2',3,4',5,5',6-Heptachlorobiphenyl (52663-68-0)	0.001
2,2',3,3',4,4'-Hexachlorobiphenyl (38380-07-3)	0.001
2,2',3,4,4',5,5'-Heptachlorobiphenyl (35065-29-3)	0.001
2,2',3,3',4,4',5-Heptachlorobiphenyl (35065-30-6)	0.001
2,2',3,3',4,4',5,6-Octachlorobiphenyl (52663-78-2)	0.001
2,2',3,3',4,4',5,5',6-Nonachlorobiphenyl (40186-72-9)	0.001
Decachlorobiphenyl (2051-24-3)	0.001
Silver (7440-22-4)	0.01
Tin (7440-31-5)	0.05
Zinc (7440-66-6)	50.0

APPENDIX 2B

DIATOM, ZOOPLANKTON, BENTHOS, FISH, BIRD AND SEDIMENT TOXICITY METHODS

Variable	Method	Reference
Diatoms	K-B corer, 20–50-cm core ^{210}Pb dating of top, bottom, and 10 and 20 cm depths in eutrophic lakes. Identify 500 valves to species at 1250 X magnification.	Glew, 1989; Dixit et al., 1992; Smol and Glew, 1992.
Zooplankton	Paired 48-μm and 202-μm mesh nets. Vertical tow from 1/2 m off bottom of deepest site to surface at 5-6 s/m. Samples anesthetized and preserved, and both macro- and microzooplankton identified to species.	McCauley, 1984.
Benthos	Petite PONAR grab, K-B corer in profundal sites. 595 μm sieve to remove sediments from grab samples. Corer sediments will be cleared by digestion. Sweep nets at littoral sites. Specimens identified to genus or species.	Weber, 1973. Klemm et al., 1990
Fish	Experimental gill nets and eel pots in pelagic and profundal zones. Indiana trap nets, boat electrofishing, beach and short seines, minnow traps, and eel pots in littoral zone. Passive gear fished overnight and active gear fished after sunset. Specimens identified to species and size class and examined for anomalies. Voucher specimens sent to Harvard Museum of Comparative Zoology for confirmation.	Nielsen and Johnson, 1983; Hocutt and Stauffer, 1980; Kolz and Reynolds, 1991.
Birds	Circular plot count with 100-m radius 10 m from shore. 20-24 plots. Identified to species visually and aurally.	Reynolds et al., 1980; Moors and O'Connor, 1992
Sediment toxicity	Petite PONAR Grab. 7-D *Hyallela azteca* test using 20 animals 2-3 days old with 100 mL sediment and 400 mL reconstituted water with daily replacement. Measure survival and dry weight.	Klemm and Lazorchak, 1992

APPENDIX 2C

PHYSICAL HABITAT STRUCTURE FIELD METHODS

Variable	Method	Reference
Canopy layer Mid layer Ground cover Shore substrate Human influence Bank type Fish cover Aquatic Macrophytes Bottom substrate General observations (sediment color, odor, surface scum, mats)	From boats anchored 10 m offshore, crews made observations of riparian vegetation structure and human disturbances on plots 15 m wide extending 15 m back from the shore. Littoral measurements or observations of shoreline substrate, bottom substrate, near-shore fish cover, and human disturbances were made along the 15-m shoreline plots and in the water between the shore and the boat. Field forms standardized most observations as absent, present (< 10% areal cover), subdominant (10–40% cover), or dominant (> 40% cover).	Tallent-Halsell and Merritt, 1991

APPENDIX 2D

WATER QUALITY METHODS

Variable	Summary of Method	Reference
Depth	Sonar measurement or calibrated line.	Chaloud et al., 1989
Temperature, in situ	Measured at various depths using thermistor probe.	Chaloud et al., 1989
Dissolved oxygen, in situ	Measured at various depths using membrane electrode and meter.	Chaloud et al., 1989
Transparency	20-cm diameter black and white Secchi disk with calibrated line; estimated as depth of disappearance plus depth of reappearance, divided by 2.	Lind, 1979; Chaloud et al., 1989
pH, closed system	Sample collected and analyzed without exposure to atmosphere; electrometric determination (pH meter and glass combination electrode).	U.S. EPA, 1987
pH, equilibrated	Equilibration with 300 ppmV CO_2 for 1 hr prior to analysis; electrometric determination (pH meter and glass combination electrode).	U.S. EPA, 1987
Acid Neutralizing Capacity (ANC)	Acidimetric titration to pH $\leq$ 3.5, with modified Gran plot analysis.	U.S. EPA, 1987
Carbon, dissolved inorganic (DIC), closed system	Sample collected and analyzed without exposure to atmosphere; acid-promoted oxidation to CO_2 with detection by infrared spectrophotometry.	U.S., EPA, 1987
Carbon, dissolved organic (DOC)	UV-promoted persulfate oxidation, detection by infrared spectrophotometry.	U.S. EPA, 1987
Conductivity	Electrolytic (conductance cell and meter).	U.S. EPA, 1987
Aluminum, total dissolved	Atomic absorption spectroscopy (graphite furnace).	U.S. EPA, 1987
Aluminum, monomeric	Collection and analysis without exposure to atmosphere. Colorimetric analysis (automated pyrocatechol violet).	APHA, 1989; U.S. EPA, 1987
Calcium	Atomic absorption spectroscopy (flame).	U.S. EPA, 1987
Magnesium	Atomic absorption spectroscopy (flame).	U.S. EPA, 1987
Potassium	Atomic absorption spectroscopy (flame).	U.S. EPA, 1987
Sodium	Atomic absorption spectroscopy (flame).	U.S. EPA, 1987
Ammonium	Colorimetric (automated phenate).	U.S. EPA, 1987
Chloride	Ion Chromatography.	U.S. EPA, 1987
Nitrate	Ion Chromatography.	U.S. EPA, 1987
Sulfate	Ion Chromatography.	U.S. EPA, 1987
Phosphorus, total	Acid-persulfate digestion with automated colorimetric determinaton (molybdate blue).	Skougstad et al., 1979 U.S. EPA, 1987
Nitrogen, total	Alkaline persulfate digestion with determination of nitrate by cadmium reduction and determination of nitrite by automated colorimetry (EDTA/sulfanilimide).	U.S. EPA, 1987
Silica, dissolved	Automated colorimetric (molybdate blue).	U.S. EPA, 1987
Turbidity	Nephelometric	U.S. EPA, 1987
Total Suspended Solids	Gravimetric	U.S. EPA, 1983 APHA, 1989
True Color	Visual comparison to calibrated glass color disks.	APHA, 1989 U.S. EPA, 1987
Chlorophyll-*a*	Filtration (glass fiber) in field; extraction of filter into acetone; analysis by spectrophotometry (trichromatic).	APHA, 1989

CHAPTER 3
OVERVIEW OF SURVEY DESIGN AND LAKE SELECTION

This chapter presents the rationale for the overall EMAP design and its specific application to lakes. Information the reader should derive from this chapter includes:

- Rationale for the EMAP design strategy
- Lake selection process (Tiers 1 and 2)
- Differences in 1991 and 1992 lake selection procedures
- Inclusion probabilities and expansion factors—their importance and use
- Lessons learned

3.1 PURPOSE OF SURVEY DESIGN

A primary EMAP objective is to monitor the condition of the nation's ecological resources and provide unbiased estimates, with known confidence, of the status of, and trends in, indicators of condition. To carry out this task, EMAP relies on a probability-based survey of natural resources, from which inferences about the condition of natural resources can be drawn. Without a statistically sound survey design for monitoring, there is no assurance of obtaining unbiased estimates of the ecological condition of targeted resources with known confidence or to detect trends in ecological indicators for this target resource population (Overton et al., 1991; Stevens, pers. comm.). A probability-based survey establishes a firm foundation for estimating resource (i.e., lakes, streams, forests, wetlands) characteristics or attributes, just as an experimental design provides a basis for testing hypotheses in comparative scientific research studies.

As in any field, some of the terminology specific to the field of survey design may be unfamiliar to those in other fields, or may have somewhat different meanings. The terminology used here will be most familiar to those trained in survey theory and applications and may be unfamiliar to those trained solely in experimental design. Terms such as *populations* and *probability samples* are used extensively, but often have different connotations in the ecological and limnological disciplines. The *population*, in the survey sense, is the collection of units or items to be described (e.g., all lakes, streams, or wetlands in a defined area, such as the conterminous United States). Samples are drawn, with known probability (thus a probability sample), from the population. Measurements made on the samples are used to infer the properties of the population. Populations can be defined inclusively, such as all lakes in the United States, or the focus can be on subsets. The terms *target population* and *subpopulation* are sometimes used to represent a subset of particular interest about which inferences or conclusions will be drawn (e.g., all publicly owned lakes between 1 and 1,000 ha in area; blue-line streams on 1:100,000-scale maps; all wetlands > 0.4 ha).

Survey design and probability sampling theory have been rigorously defined, developed, and documented in the statistical literature. The EMAP survey design is similar in concept to the approaches used by the USDA National Agricultural Statistics Survey (NASS) (Cotter and Nealon, 1987) and the Forest Service Inventory Analysis (FIA) (Bickford et al., 1963; Hazard and Law, 1989) in monitoring the condition and productivity of agricultural and forest resources. Similar approaches are used by the Census Bureau and in Gallup polls to determine household characteristics and the opinions (i.e., attributes) of a representative cross-section of the U.S. population.

3.1.1 Probability Sampling Designs

Probability sampling is the general term applied to sampling plans in which:

- Every member of the population (i.e., the total assemblage from which individual sample units can be selected) has a known probability (> 0) of being included in the sample.
- The sample is drawn by some method of random selection, or systematic selection with random start, consistent with these probabilities.
- The probabilities of selection are used in making inferences from the sample to the target population (Snedecor and Cochran, 1967).

An advantage of probability-based surveys is their minimal reliance on assumptions about the underlying structure of the population (e.g., normal distribution). In fact, one of the goals of probability-based surveys is to describe the underlying population structure. Randomization is an aspect of probability-based surveys worth emphasizing. Probability sampling designs use randomization in the sample selection process as a way to select unbiased samples with known inclusion probabilities. If randomization is not incorporated into the design, it is impossible to know how well the sample represents the population. What are the biases? What proportion of the target population does the sample represent? A common problem in many water resource surveys is identifying what the samples represent, because randomization procedures have not been used in selecting the samples. Without probability sampling, each sample often is *assumed* to have equal representation in the target population, even though selection criteria clearly indicate this is not the case. Without the underlying statistical design and probability samples, the representativeness of an individual sample is unknown.

Drawing inferences from samples selected without randomization and without incorporating inclusion probabilities (expansion factors) can yield misleading conclusions. To provide policy-relevant information, not only is the ecological condition of the target population important but also the proportion of the resource that is in a particular state or condition such as *good* or *degraded*. Very different policy and management alternatives might be evaluated if 25% rather than < 1% of the target resource was found to be in a degraded ecological condition.

3.1.2 Need for a Common Design

EMAP is an integrated program, including all ecological resources—forests, arid lands, agroecosystems, wetlands, lakes, streams, Great Lakes, estuaries, and near-coastal systems—which means we must be able to combine and compare information within and across ecological resources and geographic regions. A common survey design among ecological resources contributes to this integration. The design, therefore, applies to more than just aquatic resources. Design goals include:

- Consistent representation of the environment and ecological resources by use of probability samples.
- The ability to include all ecological resources and environmental entities.
- Provision for the capacity to respond quickly to emerging environmental questions and issues.
- Representation of the spatial distribution of the ecological resource across the United States.
- Provision for linking the results with results from other probability-based national resource surveys.
- Ability to revisit sampling sites for trend detection.

The common design selected for EMAP requires explicitly defined target populations from which samples are drawn with known probability to estimate population attributes rigorously and without bias. This design is capable of sampling any spatially distributed and well-defined ecological resource and can accommodate sampling resources that are discrete, such as lakes, or extensive, such as forests. A common design may not be the optimal design for a focused, resource-specific question, but it is an adequate design for all ecological resources. In addition, results from it can be combined with results from other probability-based monitoring networks, if the other monitoring programs meet the basic criteria for design-based surveys. The common design also has the flexibility to accommodate multiple resources and multiple environmental problems at different scales, including adaptations to address emerging environmental issues and problems.

3.2 EMAP SAMPLING GRID

EMAP has selected and designed a triangular grid structure as the basis for selecting sampling locations from which to estimate the status of and temporal trends in indicators of ecological resource condition. This grid structure also permits estimates of the spatial extent or area of target ecological resources. Grid requirements that satisfy the EMAP objectives include the following:

- An equal-area sampling structure using a regular placement of sampling locations.
- A hierarchical structure for enhancing or decreasing the grid density.
- A realization of the grid on a single planar surface for the entire conterminous United States.

The grid serves primarily as a tool for achieving spatial coverage. It does not imply a belief that ecological resources are distributed uniformly.

3.2.1 Base Grid

The baseline grid used in EMAP, including EMAP-Surface Waters (EMAP-SW), has about 12,600 grid points distributed over the conterminous United States (Figure 3-1). Grid points are in a triangular array about 27 km equidistant. Large, contiguous hexagons can be scribed around each grid point, each with an area of 635 km^2 (e.g., the large hexagon in Figure 3-2). The large hexagons completely cover the areas of interest but are too large for cost-effective sampling. Therefore, the requirement for equal area sampling is further accommodated by ascribing a smaller hexagonal area centered around each grid point; each small hexagon has an area of 40 km^2 (e.g., the small hexagon in Figure 3-2). Preliminary evaluations indicate that the 40-km^2 hexagon around the grid point, combined with the number of grid points, probably is adequate for characterizing the extent and distribution of the different ecological resources (e.g., forests, arid lands, lakes, streams, estuaries, wetlands, agricultural systems) in the United States. These 12,600 small hexagons uniformly cover 1/16, or about 6%, of the area of the conterminous United States; sometimes this is called a 1/16 area sample of the United States.

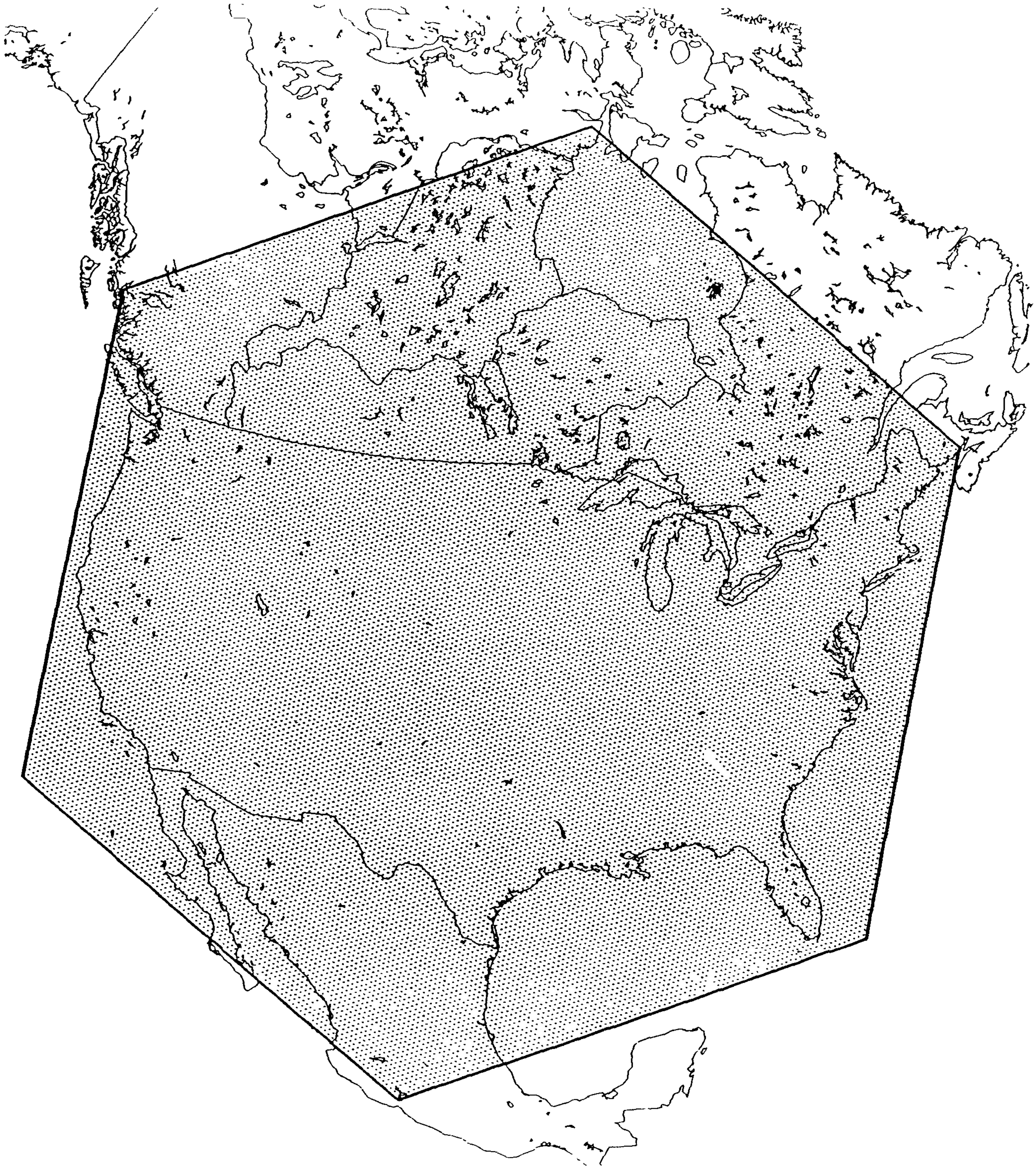


Figure 3-1. The base grid overlaid on North America. There are about 12,600 points in the conterminous United States.

Grid Density Enhancement

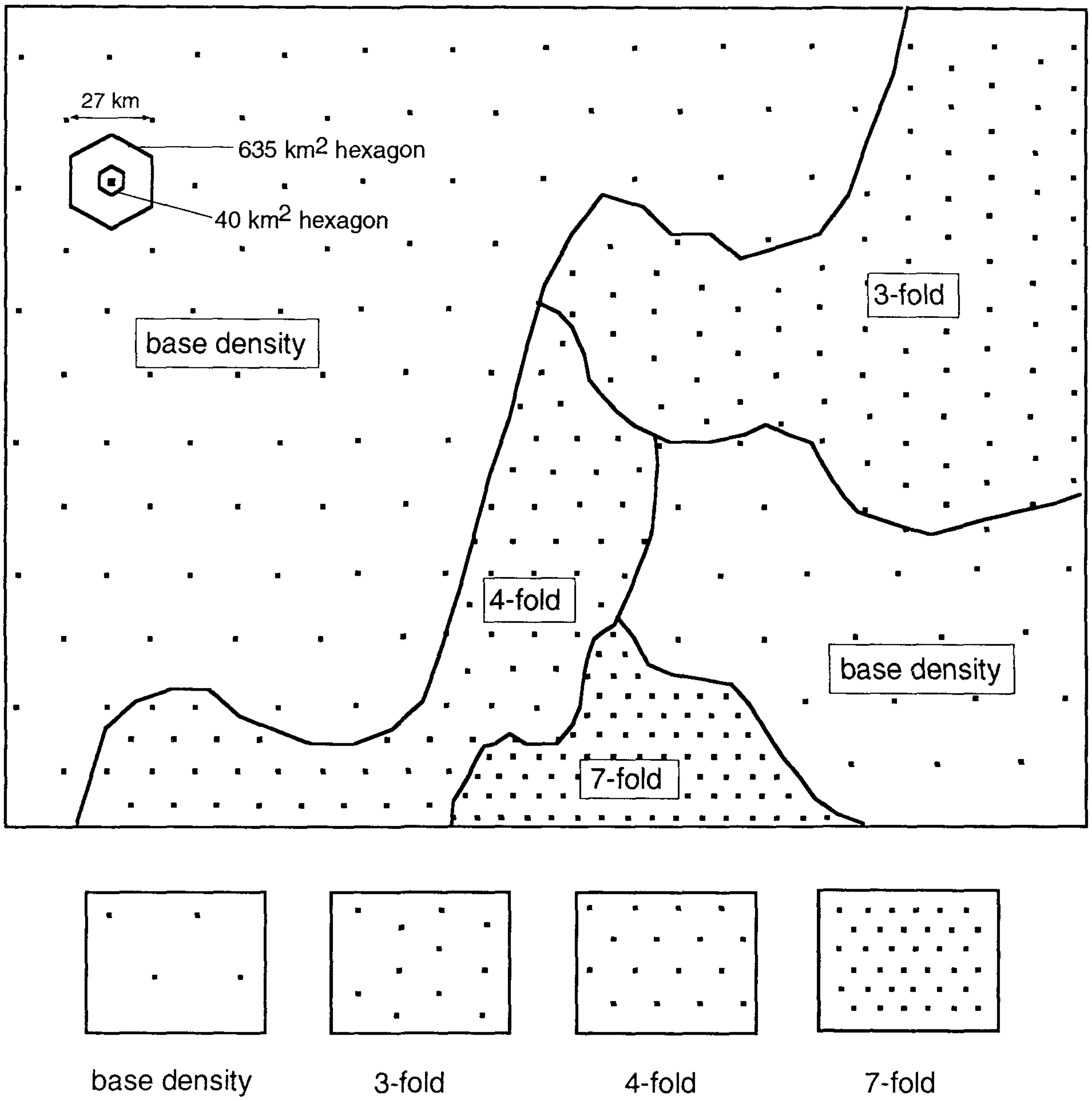


Figure 3-2. **EMAP grid structure showing the relationship between the 635 km^2 hexagons and the embedded 40-km^2 hexagon. The figure also illustrates the three-, four-, and sevenfold grid enhancements for increased sampling density.**

Initial randomization of the grid on the United States establishes the systematic sample (i.e., uniform and regular grid points and small hexagons) as a probability sample. This randomization occurs when we place the grid on a map of the United States, then shift it random distances in east-west and north-south directions. The grid structure reflects the importance of achieving geographic coverage of ecological resources. The uniformity of spatial coverage provided by a grid (Figure 3-1) ensures that each ecological resource can be sampled in proportion to its geographic presence in the United States and that all ecological resources can be included in the monitoring program (e.g., lakes, streams, prairie pothole wetlands, forests, grasslands, deserts, estuaries, managed agricultural ecosystems, and the Great Lakes). We can also place the same grid structure over maps of Hawaii, Alaska, and the Caribbean and use it as a basis for a consistent international survey design.

The grid density also can be increased to sample rare resources such as the California redwoods, additional resource classes, or smaller areas such as individual states or ecoregions. The grid arrangement also makes it easy to either increase or decrease the grid density (Figure 3-2) and retain the basic triangular structure important for consistent spatial coverage. Specific multiple factors (e.g., three-, four-, and sevenfold) are available to increase or decrease the base grid density and still maintain the sampling design requirements (Figure 3-2). The triangular nature of the grid allows greater flexibility in these enhancement factors than a square or rectangular grid.

3.2.2 Hierarchical or Tier Structure

The sampling structure is hierarchical, presently defined with four levels or tiers, two of which are used for routine monitoring. The different stages in the hierarchy reflect the desire to incorporate measurements that can be obtained for different costs. For example, we can obtain information on the distribution and extent of resources rather inexpensively with remote imagery, so many sites can be assessed. On the other hand, we can obtain detailed, time intensive data for only a few resources; the cost of such monitoring prevents sampling very many sites.

In EMAP, a Tier 1 sample of lakes consists of all lakes in the 12,600 40-km^2 hexagons. This Tier 1 sample provides good spatial coverage of the U.S. lake resource, and some measurements, such as lake size, can be made easily on the entire resource. But it is not necessary to sample *every* Tier 1 lake to achieve robust estimates of condition, and it would be prohibitively expensive to do so. Therefore, we select a subsample of the Tier 1 sample, called the Tier 2 sample. However, the greater Tier 1 density permits more precise estimates of resource extent, distribution, and characterization than can be obtained using only Tier 2 sites. Tier 1 characterization also decreases the time required for EMAP to respond to emerging environmental problems, because information about ecological resources acquired at that level can be used to refine or focus the Tier 2 sample. The Tier 2 sites are selected so they also satisfy the design criteria of good spatial and geographic distribution with known inclusion probabilities.

EMAP also establishes a four-year sampling cycle, with assignment of a particular year in the four-year cycle to each hexagon. This four-year rotational cycle is more fully described in Section 3.5.

3.3 FRAME AND TIER 1 SAMPLE SELECTION

In the following sections, we incorporate specific information about using the design for the selection of lakes. We include information on how we selected Tier 2 samples (the set of lakes to be field visited) for both the 1991 and 1992 field seasons. In general, we describe the process used during the 1991 selection and note how it was altered for the 1992 selection. Some differences arose, related both to what we learned during the 1991 selection process and to a slightly different focus during the 1992 field season.

3.3.1 Lake Sampling Frame and Tier 1 Sample

A *sampling frame* is an explicit representation of a population from which a sample can be selected. As the representation, or frame, of the population of lakes, we used the USGS 1:100,000-scale map series in digital format (DLGs) and the modification of the DLGs represented by the U.S. EPA River Reach File, Version 3 (RF3), which established edge matching and directionality in the DLG files. The DLG (or RF3) files contain a representation of the spatial distribution of all lakes and streams as recorded on the USGS 1:100,000-scale topographic map series. The lake population of interest to EMAP-SW in the current series of pilot studies consists of all lakes in the conterminous United States with a surface area ≥ 1 ha, excluding the Laurentian Great Lakes.

One difficult issue has been to provide an operational definition of a lake that field crews could use in deciding whether to sample a waterbody chosen via the lake selection process. After examining Cowardin's wetland classification system (Cowardin et al., 1979), which mainly distinguishes deep water from shallow water habitats, and after discussing the issue at length, EMAP-SW staff settled on an operational definition that combines information from the 1:100,000-scale maps and information available only from field visits.

The first level of selection included waterbodies identified as lakes > 1 ha in surface area in the DLG database (also as recorded in the DLGs). The second level of selection specified a criterion for field visitation. If the waterbody had > ≈ 1,000 m^2 of open water and an estimated maximum depth > 1 m, it would be sampled for lake indicators. If not, the waterbody would be recorded as a noninterest waterbody for the purpose of measuring lake indicators.

A Tier 1 sample consists of all lakes captured by the 40-km^2 hexagons; this 6% area sample therefore captures approximately 6% of the lakes across the region of interest. Operationally, each lake is uniquely represented by a point with specific coordinates in the DLGs or in RF3, generally at the centroid of the lake. Each point has associated attributes such as latitude/longitude, unique identification, lake surface area, and other available information. Because the DLGs and RF3 are in electronic form, it is also feasible to create an inventory of lakes in specific regions of interest.

For the 1991 field sampling season, we confined our selection process to lakes in the northeast, the region of the country selected for the initial pilot studies. We also restricted the selection to lakes < 2,000 ha for the 1991 sample, because we were not confident that we could effectively apply the index sampling process to large lakes. However, since subsequent refinements to RF3 allowed us to draw a national Tier 1 sample (lower 48 states) for the 1992 field season, we did so in view of the possibility of conducting a national survey on a small set of lake indicators. We also included larger lakes. Table 3-1 summarizes the inventory and Tier 1 sample of lakes for both years. The summaries are organized by size class; different size classes were selected for each year, as described in Section 3.3.3.

Evaluation of the 1992 Tier 1 national sample revealed that the number of lakes > 500 ha captured by the 40-km^2 hexagons was too small to form an adequate Tier 2 sample. Any lakes ≤ 500 ha that fell within the 40-km^2 hexagons were selected as Tier 1 lakes (each lake is uniquely represented by a point with specific coordinates in RF3). The complete list of lakes > 500 ha was drawn from the frame and the Tier 2 sample was drawn from this list. It was more efficient to select a probability sample directly from the lake list than to increase the density of the grid to obtain an adequate sample of these larger lakes. The spatial distribution of these larger lakes was preserved by associating each large lake with its nearest hexagon and then using the normal selection procedure described in Section 3.4 and Appendix 3A.

Table 3-1. Numbers of Lakes in the Population and Tier 1 Sample from the Base Grid[a]

1991 (Northeast)					
Size Class (ha)	Population	Tier 1 (4 yrs)	Tier 1 (1991)	Tier 1 Rejects	Tier 1 Screened
1–5	10,791	662	158	35	123
5–20	5,969	371	92	10	82
20–500	3,444	197	57	3	54
500–2,000	174	9	4	0	4
Total	20,378	1,239	311	48	263

1992 (Lower 48 states)		
Size Class (ha)	Population	Tier 1 (4 yrs)
1–5	161,616	10,101
5–10	43,744	2,734
10–50	41,648	2,603
50–500	11,712	732
500–5,000	1,661	NA[b]
> 5,000	257	NA[b]
Total	260,638	

[a] Population numbers come from the Digital Line Graph summaries; 1991 covers the northeast pilot and 1992 represents the conterminous United States.

[b] The Tier 2 lakes were selected directly from the target population of lakes, so no Tier 1 sample was drawn.

3.3.2 Frame Characterization

The frame is not a completely accurate representation of the population of lakes of interest. Various errors creep in. Also, a waterbody identified by the USGS as a lake for mapping purposes does not always correspond to a limnologically defensible definition of a lake. Therefore, the lake frame requires characterization to identify a combination of mapping errors, such as (1) parts of lakes (e.g., coves and bays), estuaries, and wide sections of rivers designated as lakes, (2) changes in the landscape since the time the maps were compiled (some lakes and reservoirs have been drained, others created) and (3) waterbodies that will be tracked but not included in the sample to be visited. Thus the characterization activity identifies two categories: *frame errors* and *noninterest (or nontarget)* waterbodies. We use the term *noninterest* solely in the context of identifying types of waterbodies that will not be monitored in the field and that will be excluded from initial population characterization in our pilot studies. In keeping with EMAP's goal of leaving no orphan ecosystems, we do not presume to exclude these waterbodies from ever being considered for EMAP monitoring. One of the advantages of survey designs such as EMAP is that the proportion or number of waterbodies not chosen for consideration can be estimated and tracked.

We evaluated frame errors and noninterest waterbodies for the 1991 Northeast Pilot by examining larger scale maps (7.5-minute topographic and larger scale county maps) and talking with local experts. The majority of the frame errors and noninterest waterbodies were associated with lakes < 20 ha. About 20% of the Tier 1 lakes from 1 to 5 ha represented on the DLGs and 10% of the lakes from 5 to 20 ha were determined to be either frame errors or noninterest waterbodies such as cranberry bog reservoirs, wide spots in rivers, or waterbodies that were arms or bays of larger lakes. No (0) noninterest lakes were found among lakes with surface areas > 50 ha. Chapter 5 (Annual Statistical Summary) contains additional details. For the 1992 pilot, resources needed to perform a Tier 1 screen on the lower 48 states were unavailable, so we drew a Tier 2 sample directly from the unscreened Tier 1 sample, while accounting for the expected low-interest lakes.

Identifying lakes that are of interest but not represented in the frame (i.e., not in the DLG file or on USGS topographic maps) is more difficult. Methods considered for future identification of lakes not represented in the frame include using remote aerial imagery/photography and relying on local experts to provide detailed area knowledge. Results of these efforts can then be compared to the lake frame, either in conjunction or separately. Our initial impression for lakes in the Northeast is that the frame over-represents the target population; that is, it is more likely to classify noninterest waterbodies as lakes than to exclude interest lakes. Evaluating the accuracy of the Tier 1 and Tier 2 lake selection process represents a major activity that will occur over the next several years.

3.3.3 Classification Strategies

To ensure an adequate representation of various subpopulations in the sample, lakes could be classified by subpopulations as part of the Tier 1 activity. The question is whether or not to develop a lake classification scheme, other than one based on size, to stratify the Tier 1 sample, then select the Tier 2 samples from each stratum. For example, the lakes could be classified as oligotrophic, mesotrophic or eutrophic; warmwater or coldwater; deep or shallow; impoundment or natural lake, etc. Because of the variety of overlapping lake classification approaches and the potential loss of precision in case of misclassification, we decided to minimize stratification and rely on post-classification as a more appropriate part of the evaluation of results. However, as part of the lake frame evaluation activities, we will analyze the Tier 1 and Tier 2 lake characteristics that can readily be obtained from maps, atlases, or similar references in order to initially determine the different subpopulations of lakes in the target population.

A second question, however, was whether we could afford not to classify lakes on the basis of surface area. A nonstratified, random sample would yield lakes in proportion to their abundance, so most of the lakes selected would be small. Over 75% of lakes in the frame have surface areas between 1 and 10 ha (Table 3-1). Equal probability allocation would produce Tier 2 samples with proportional numbers of small lakes. The importance of lakes to society is related both to numbers of lakes and to their sizes. If we strictly apportioned Tier 2 selection on the basis of numbers of lakes, we would impair our ability to describe the condition of the larger lakes. There is considerable public, policy, and regulatory interest in lakes > 10 ha. But if we strictly apportioned sampling according to area, we would impair our ability to describe the condition of small lakes. Thus, we sought to balance the selection of Tier 2 lakes. One approach to obtaining a more equitable distribution of lakes by size would be to allocate equal numbers of samples along a logarithmic or square root transformation of surface area, so that more large lakes would be selected. However, allocating samples along a continuous scale requires the use of variable inclusion probabilities, substantially complicating variance estimation (Overton, pers. comm.). The issue of whether the advantages of using variable inclusion probabilities outweigh the disadvantages of complicated variance estimation has not been completely resolved, but we decided on the approach of classifying by lake area. We used an iterative process, varying size classes and sample sizes among classes, to select samples about equally among lake size classes, approximating what would have been achieved by logarithmic or square root transformations. For the 1991 pilot, we based our sample selection only on the population of lakes in the Northeast and restricted the size strata to lakes between 1 and 2,000 ha. For the 1992 pilot, we used estimates for the lower 48 states and included lakes > 2,000 ha. As a result, slightly different size classes emerged from the process, summarized in Table 3-1, along with the numbers of lakes in each size class. Over the next several years, we will evaluate the relative allocation between size classes. These size classes are not intended to convey any ecological information, but were selected as a way to spread the Tier 2 sample more evenly among lake sizes than would have been done by equal probability sampling.

3.4 LAKE SELECTION

Evaluation of the Tier 1 sample allows us to refine the selection of the Tier 2 sample. Characterizing the Tier 1 sample by lake size (Section 3.3.3) focuses the Tier 2 sample so that it better characterizes larger lakes than could be done otherwise. It also prevents oversampling the small lakes beyond what is necessary to characterize their condition. The selection of Tier 2 lakes met two criteria: (1) we used probability methods to select the lakes from the Tier 1 sample, and (2) we preserved the spatial distribution of the lakes in the Tier 2 sample.

The details of the Tier 2 lake selection process appear in Appendix 3A. The outcome of the Tier 2 selection process is a set of lakes on which measurements of lake condition will be taken during the appropriate index period. Table 3-2 summarizes the sets of Tier 2 lakes selected for the 1991 and 1992 sampling seasons. The sample of lakes represents the larger population of lakes of interest; each lake is selected with a known inclusion probability reflecting the probability of selecting any particular lake at the Tier 1 level and at the Tier 2 level. Reciprocals of inclusion probabilities are the multipliers by which the sample attribute is expanded to its portion of the population. This multiplier is called an *expansion factor.* For example, if the attribute *surface area* for a particular lake in the Tier 2 sample were 10 ha, and that lake's expansion factor were 100 (inclusion probability = 0.01), the lake would have a weight of 1,000 ha. Summation of the weights across the sample of lakes produces an estimate of the total surface area of lakes in the population. Inclusion probabilities and expansion factors for each size class are included as part of Table 3-2, and the spatial distribution of these lakes is illustrated in Figures 3-3 and 3-4.

3.5 TEMPORAL SAMPLING SCHEDULE

EMAP is designed both to describe the current status of and to detect trends in the condition of ecological resources. Status is best estimated by allocating sampling effort among as many different sample units (lakes) as possible. However, trend is best detected by incorporating repeat visits but not necessarily annual revisits. EMAP's design balances both objectives by establishing a four-year rotational cycle in which each grid point is assigned a year. This rotational design increases the number of lakes visited over the four-year cycle to a higher number than would be the case under annual revisits. The increased effective sample size

Table 3-2. Tier 2 Yearly Sample Sizes for the Northeast, with Corresponding Inclusion Probabilities and Expansion Factors

Size Class (ha)	Tier 2 Sample	Overall Inclusion Probability	Expansion Factor
1991			
1–5	16	.001953	512.000
5–20	20	.003906	256.000
20–500	24	.007422	134.737
500–2,000	4	.01562	64.000
Total	64		
1992			
1–5	9	0.00125	800.000
5–10	31	0.00938	106.667
10–50	20	0.00469	213.333
50–500	6	0.00859	116.364
500–5,000	10	0.0625	16.000
> 5,000	1	0.0500	20.000
Total	77		

improves estimation of status (over the four-year interval) and trend detection, compared to annual revisits. This assignment preserves the triangular grid structure, so that in each year consistent spatial coverage is maintained. During the first year, one quarter of the hexagons are selected (schematically represented by "♦" in Figure 3-5). Only lakes in these hexagons are candidates for field sampling during the first year. Recall that the lakes captured by the 40-km^2 hexagons comprise the Tier 1 sample and that only a portion of the Tier 1 sample will be visited.

During the next year, the second quarter of the lakes, those captured by hexagons designated as "second-year" ("+"), are available for sampling, and so on. In this manner, all Tier 2 lakes are sampled during a four-year period. Any individual lake, therefore, is sampled once every four

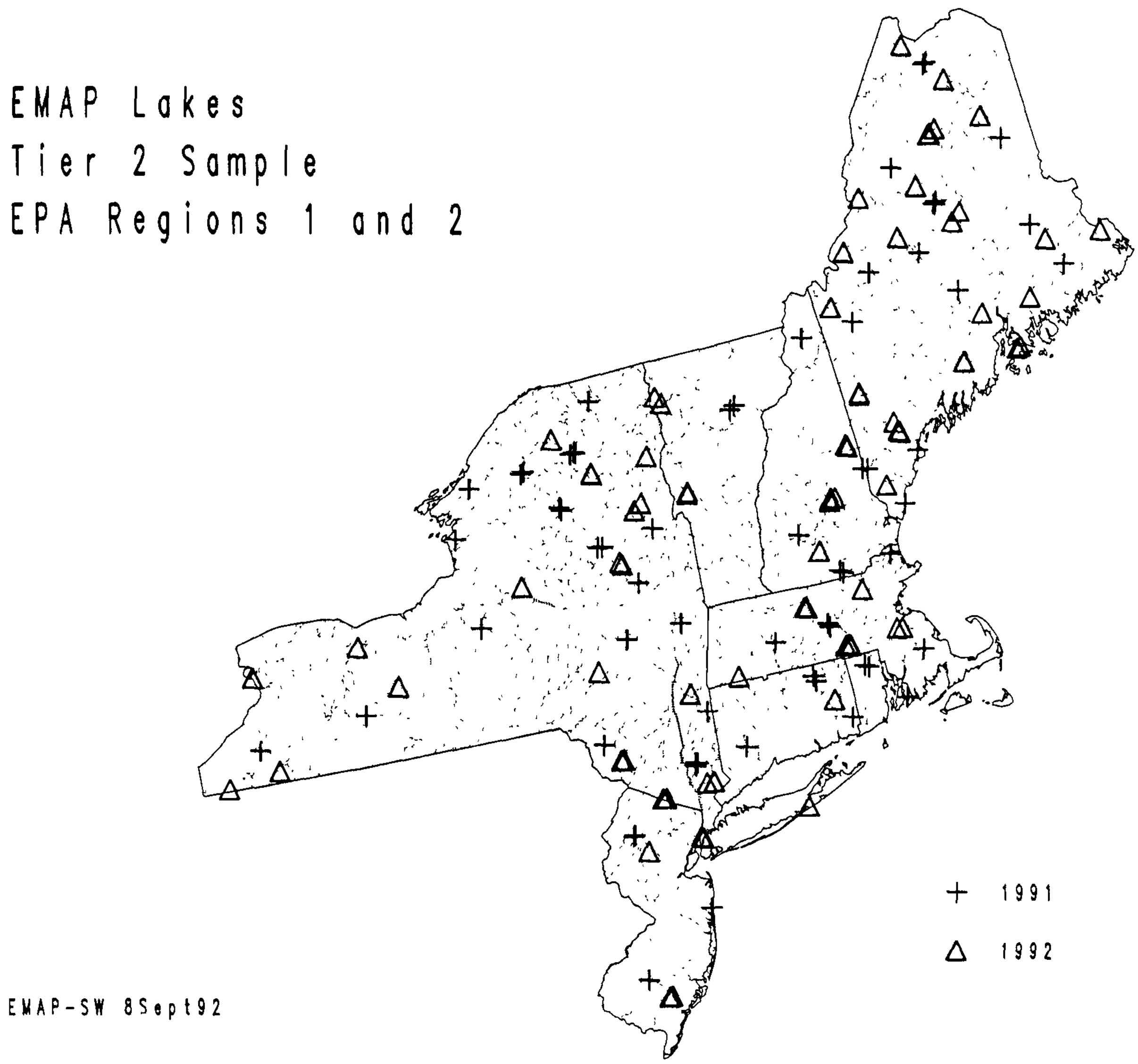


Figure 3-3. **Map illustrating the spatial distribution of the 1991 and 1992 Tier 2 samples of lakes selected from the Digital Line Graph frame.**

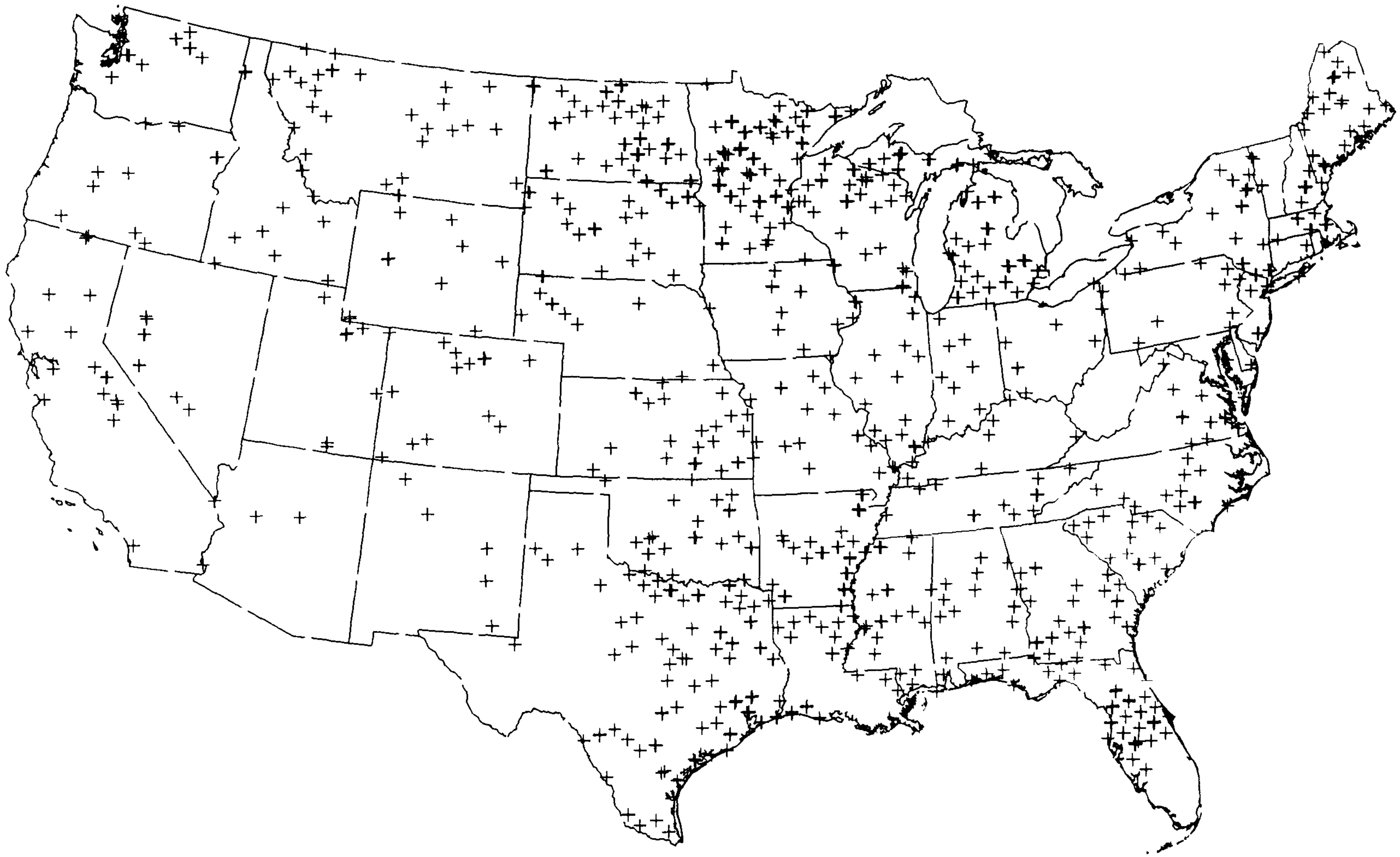


Figure 3-4. Distribution of selected 1992 Tier 2 lakes for the conterminous United States. This distribution has not been corrected for noninterest lakes.

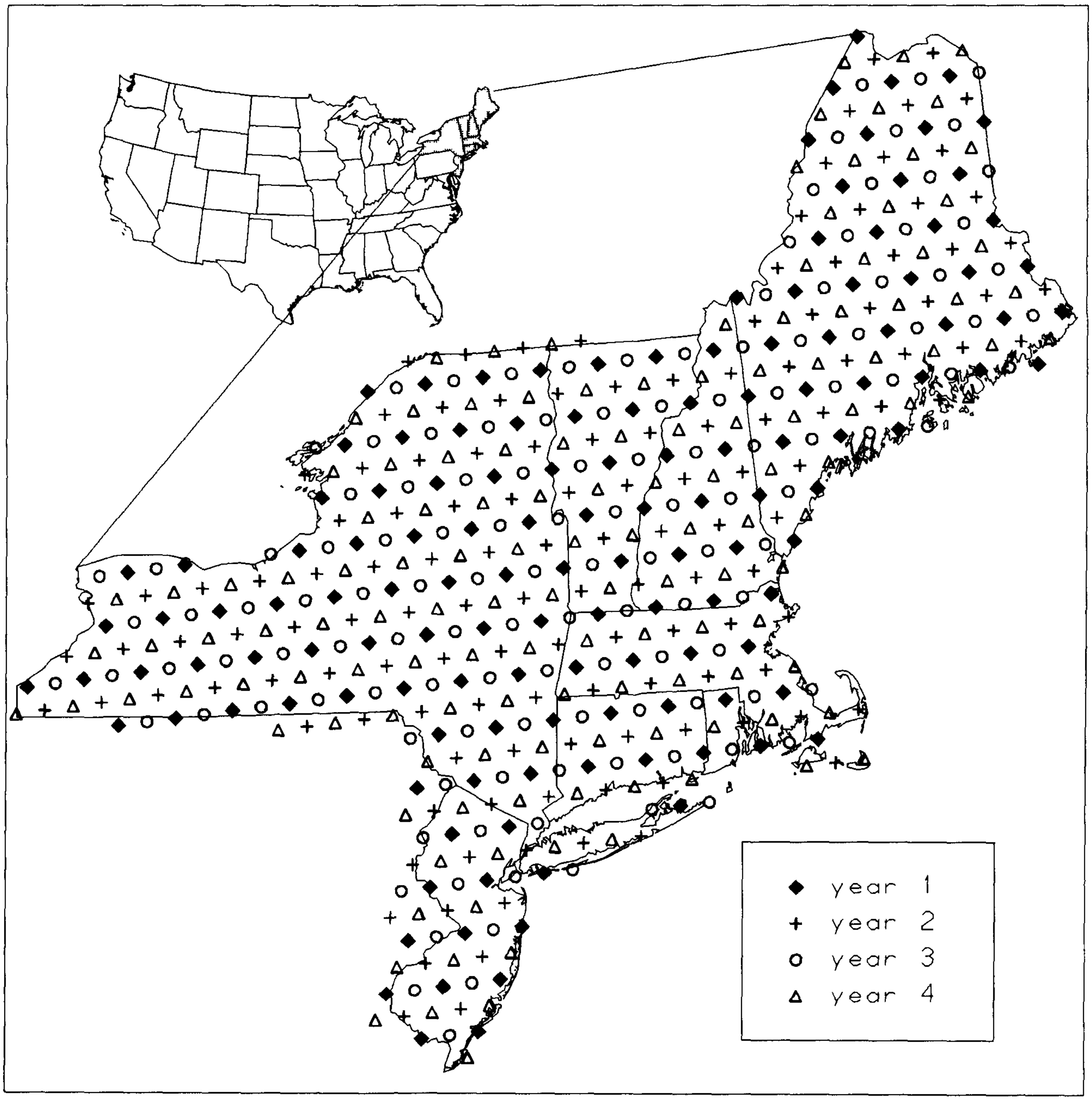


Figure 3-5. Temporal sampling schedule for all EMAP resources illustrating the uniform spatial coverage for every year.

years. A second monitoring cycle begins in the fifth year by revisiting the first year ("♦") lakes; this process could continue indefinitely. The temporal sequence of hexagons has been fixed for EMAP. Therefore, the lakes associated with each hexagon are also associated with a particular sampling year. This ensures that all the EMAP resources sampled in that hexagon will be sampled during the same year. To permit comparisons among ecological resource condition in the same year, the hexagon sequence for any year is fixed and cannot be substituted for that of another year.

The temporal sampling schedule and design have several advantages. Each year's sample provides, in itself, both national and regional estimates of ecological condition, with uniform spatial coverage from year to year. Annual estimates of population parameters are provided for every geographic region and every identifiable population, no matter how dispersed. Revisiting sites on a four-year cycle provides sufficient time for recovery from measurement stress and allows time for subtle trends to be expressed. The design is well adapted for detecting persistent, gradual change on disperse subpopulations and representing spatial patterns in ecological resources.

3.6 ANNUAL REVISIT SITES

The EMAP design foresees that lakes will be revisited every four years. Over eight or more years, this gives good sensitivity to trend detection. The sensitivity to trends during the initial years of monitoring can be increased substantially by annual revisits to a few lakes (Urquhart et al., 1991). Ten of the lakes visited in 1991 were randomly selected, subject to spatial coverage, to be revisited in 1992 and subsequent years. The process used for achieving random selection, subject to spatial balance, is described in Appendix 3A.

3.7 SUMMARY AND LESSONS LEARNED

During this past year of design work, we were actually able to complete two design cycles. The first allowed us to put our design concepts into practice in selecting the 1991 lakes for sampling. The process, although not completely automated the first time, worked reasonably well. The selection resulted in an acceptable spatial distribution of lakes in the Northeast. We gathered valuable information from the Tier 1 screening on the lake frame errors and the noninterest lakes. This permitted a more efficient Tier 2 selection that contained relatively few errors and noninterest lakes. Because of the initial grid density, the size of the selected Tier 2 sample, and the natural distribution of lakes, two or more Tier 2 lakes occurred in the same hexagon many times in our 1991 sample. This has both positive and negative aspects. On the positive side, we may be able to estimate some of the spatial autocorrelation effects sometimes observed among lakes close to one another. The negative aspect is that we do not achieve as broad a spatial coverage as might otherwise be possible, a desired characteristic in Tier 2. The only ways to decrease the frequency of this multiple occurrence of lakes are to reduce the Tier 2 sample size or to increase the Tier 1 density. We are in the process of evaluating the consequences of these alternatives.

During the second design cycle, which resulted in the selection of the 1992 samples, we were able to apply our design concepts and experience to a national (conterminous U.S.) sample. This was an important step in improving our understanding of the overall design and spatial distribution of the resulting Tier 2 sample. If the relatively simple approach applied to the Northeast resulted in an acceptable sample (i.e., adequate coverage of multiple subpopulations) nationwide, then we would continue along the same course. That is, if the same inclusion probabilities (within each size class) could be applied nationally, the analyses resulting from the data and comparison across regions would be greatly simplified. Given the magnitude of the national Tier 1 sample, we chose not to do the initial Tier 1 screening for frame errors and noninterest waterbodies the second year. Instead, we increased the Tier 2 sample size to account for the expected losses and achieve the desired Tier 2 sample size. Table 3-2 and Figures 3-3 and 3-4 show 1992 size categories, inclusion probabilities, and the spatial pattern of selected Tier 2 lakes. In general, the selection produces few surprises. There is greater density of lakes in the east than in the west. In the east, there is a greater density in the north than in the south. However, the distribution of samples raises some concerns and questions. For example, the density of lakes in the upper midwest reflects the greater density of lakes in this region; we are not sure, however, that a sample size that large is necessary in the upper midwest to adequately characterize the region. Conversely, the sparser sample selection of lakes in the west clearly reflects their lower density; but we are concerned that we may not adequately characterize some lake subpopulations of interest, such as the alpine lakes. Alternatives for dealing with these situations are currently under discussion.

Perhaps the final lesson learned from the 1992 sample selection process has been the need to perform the Tier 1 screening. An effective evaluation will in fact make the selection of Tier 2 a much more efficient process both in terms of what we are able to say about lakes simply from the Tier 1 sample and also from the perspective of a more efficient use of time by those doing the field reconnaisance on the Tier 2 samples. This Tier 1 screening will be incorporated in upcoming years.

Our experience with the lake selection process indicates that, given a lead time of 4 to 6 months, we can select a set of lakes for monitoring that meets the general design criteria for a probability sample with spatial balance. Refinements will be necessary to define important lake subpopulations and to improve descriptions of the small lakes (< 10 ha).

APPENDIX 3A

DETAILS OF THE LAKE SELECTION PROCESS

The lake selection process is illustrated schematically as Figure 3A-1; details of the selection process are as they occurred for selecting the 1992 national lake sample. Where appropriate, the text that follows identifies the differences in the process used for the 1991 sample. The first three steps are straightforward and produce the Tier 1 sample, summarized earlier in Table 3-1.

The selection of the Tier 2 sample from the Tier 1 sample recognizes the same basic survey criteria met by the Tier 1 sample selection:

- The sample is drawn with known probability.
- The sample is drawn to preserve the spatial pattern of the population of interest.

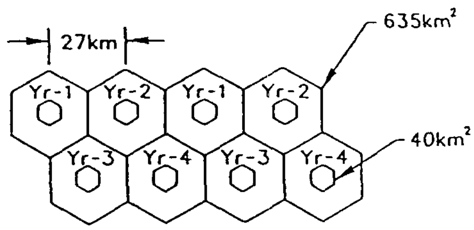
In order to meet these criteria, conceptually, we start the Tier 2 selection process by dividing the region of interest into smaller compact clusters of hexagons, then randomly selecting lakes within each cluster. For the 1991 Northeast Pilot, the clusters were developed solely for the Northeast (New York, New Jersey, and the New England states) and were subjectively drawn. For the 1992 pilot, the clusters were developed nationally using a computer algorithm. The clustering met the following criteria:

- The sums of the inclusion probabilities in each cluster for each year were as similar as possible across clusters.
- Compact, "round" clusters were preferable to long, narrow clusters.
- All lakes in a 40-km^2 hexagon were in the same cluster.

The clusters were constructed so that they contained a total inclusion probability of at least 2 in each year of the 4-year cycle. Delineating the clusters and selecting lakes within each cluster randomly also assured that the spatial distribution of the lakes would be preserved. Because the primary function of the clusters was to distribute the sampling effort in proportion to the spatial distribution of lakes, the actual dimensions and boundaries of the clusters were not critical for the next steps. It is desirable for at least one lake to be selected from each cluster, but the advantages of the clusters are defeated if more than three or four lakes are selected from any one cluster. Therefore, the target was to define clusters such that at least two lakes from each cluster would be selected. Step 4 in Figure 3A-1 shows an example of this clustering.

1. Common EMAP grid frame for all U.S. ecological resources.

2. Four-year sampling cycle established for EMAP. Hexagons designated by year. After random displacement of the base grid.



3. EPA River Reach File (RF3) used to develop sampling frame.

   a. Identify all lakes in 40-km^2 hexagons based on unique point locators for each lake.
   b. Extract lakes ≥ 1 ha using GIS.
   c. For lakes ≤ 500 ha, keep grid structure.
   d. For lakes > 500 ha, inventory or list the lakes.
   e. These are Tier 1 lakes.

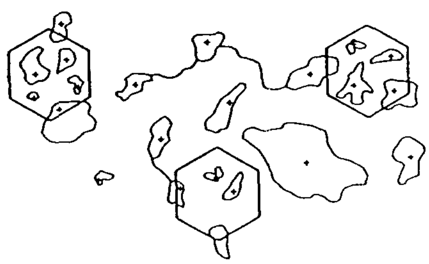
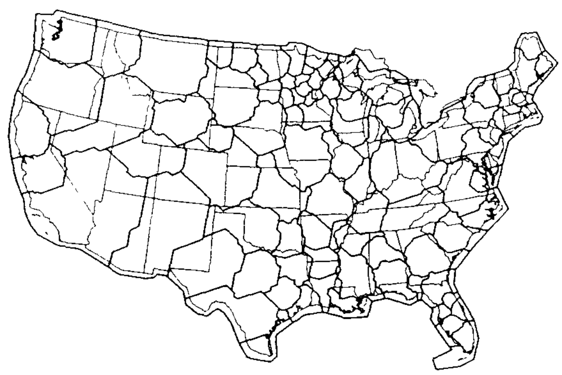


Figure 3A-1. Schematic representation of the lake selection process (page 1 of 2).

4. Classify lakes by surface area into 6 size categories: 1–5, 5–10, 10–50, 50–500, 500–5,000, and > 5,000 ha.

   a. Determine inclusion probabilities for lakes by size class.
   b. Establish lake clusters so that sum of inclusion probabilities in each cluster is about the same, regardless of size class.



5. Separate the hexagons, and associated lakes, by sampling year in each cluster. Start with Year 2 hexagons - for 1992.

   a. Randomly order clusters, hexagons within clusters, and lakes within hexagons.
   b. Randomly order list of lakes > 500 ha.
   c. Assign length proportional to inclusion probability to each lake in array.
   d. Select random number to initiate process and incrementally increase by length that results in selection of desired number of sample lakes.
   e. Follow similar process for inventory of lakes > 500 ha.

	Lake	Hex	Cluster
	Lake 18,3,2		CLUSTER 18
r ⇒	Lake 18,3,1	Hex 18,3	
	Lake 18,3,4		
r+1 ⇒	Lake 18,3,3		
	Lake 18,1,7		
r+2 ⇒	Lake 18,1,1		
	Lake 18,1,4	Hex 18,1	
r+3 ⇒	Lake 18,1,3		
	Lake 18,1,2		
r+4 ⇒	Lake 18,1,5		
r+5 ⇒	Lake 18,1,6		
	Lake 18,2,2		
r+6 ⇒	Lake 18,2,1	Hex 18,2	
	Lake 4,7,6	Hex 4,7	CLUSTER 4
	•	•	
	•	•	
	•	•	

6. Repeat process for hexagons and lakes in the other three years of the sampling cycle.

7. Perform QA on lake list to identify nontarget lakes.

Selected Lakes	QA
Lake 18,3,1	Lake
Lake 18,1,7	Lake
Lake 18,1,1	Wetland
Lake 18,1,3	Wide spot in river
Lake 18,1,6	Lake
Lake 18,2,1	Lake drained

Figure 3A-1. Schematic representation of the lake selection process (page 2 of 2).

Drawing the Tier 2 Sample

For the 1991 Tier 2 sample, the Tier 1 sample was screened for frame error and noninterest lakes before the Tier 2 sample was drawn (See Table 3-1). For the 1992 Tier 2 sample, no Tier 1 screening was performed because the time required to perform a national screen exceeded the time available to produce the Tier 2 sample. For future work, efforts to estimate Tier 1 errors will be factored into the process.

A tentative national yearly Tier 2 sample of 800 for each resource unit has been identified as a target for each EMAP resource group in each year for the lower 48 states. We translated this into 800 lakes and 800 streams per year as design targets for EMAP-SW. Based on our Tier 1 screening of the 1991 sample and subsequent field visits that identified other nontarget lakes in our 1991 sample, we overselected lakes in the 1–5 and 5–10 ha classes. Our best estimate, based on the 1991 experience, was that about half of the 1–5 ha and about one-third of the 5–10 ha Tier 1 sample (unscreened) would be errors or nontarget lakes. Therefore, target goals were established to select 200 lakes in the 1–5 ha size class and 375 lakes in the 5–10 ha size class with the expectation of achieving about 100 lakes in the 1–5 ha size class and 200 lakes in the 5–10 ha size class. If our projections were correct, the resulting sample of interest lakes should approximate 800 each year, distributed among size classes as in Table 3A-1.

The process outlined here was followed in selecting the 1992 Tier 2 sample from the Tier 1 set of lakes:

1. A conditional inclusion probability representing the fraction of the Tier 1 sample to be selected as the Tier 2 sample for each size class was assigned to each Tier 1 lake.

2. Clusters were delineated based on their total conditional inclusion probability.

3. The hexagons for the first, second, third, and fourth years of sampling were identified along with their corresponding lakes. A four-year rotation sampling cycle has been established for EMAP, so the hexagons to be sampled, if the resource exists in that hexagon, have been designated as year 1, 2, 3, 4 (Figure 3A-1).

4. Three arrays were established by randomly ordering the clusters, randomly ordering the 40-km^2 hexagons for a specific sampling year (e.g., year 1, year 2) within each cluster, and randomly ordering the lakes within each 40-km^2 hexagon. This resulted in an order for all the lakes in the sample; the lakes within each 40-km^2 hexagon were adjacent to each other, and the 40-km^2 hexagons within each cluster were also adjacent to one another (See Figure 3A-1). Each cluster consisted of a contiguous group of hexagons, for which the total conditional inclusion probability was at least 2 for each of the four years of hexagons in the cluster. In order to achieve this condition, some cluster/year combinations substantially exceeded the target value of 2.

Table 3A-1. Target Sample Sizes, Actual Sample Sizes, Inclusion Probabilities, and Expansion Factors for the 1992 National Lake Sample

1992 National Lake Sample				
Size Class	**Target Sample Size**	**Actual Sample Size**	**Inclusion Probability**	**Expansion Factor**
1–5	200 (100)[a]	188	0.00125	800.000
5–10	375 (250)[a]	413	0.00938	100.667
10–50	200	197	0.00469	213.333
50–500	100	97	0.00859	116.364
500–5,000	100	111	0.0625	16.000
> 5,000	50	10	0.0500	20.000
Totals	1,025	1,016		

[a] Because of frame errors and noninterest waterbodies in these two groups, we selected a higher target sample size than is normally required. Desired numbers of target lakes are in parentheses.

5. Within this array, each lake was assigned a length equal to its conditional (Tier 2) inclusion probability [e.g., lakes 50–500 ha in area (with conditional inclusion probability = 0.55) would have an array length about equal to the lakes 5–10 ha in area (with conditional inclusion probability = 0.6), about twice the length of the lakes 10–50 ha in area (with conditional inclusion probability = 0.3), and about seven times as long as the lakes 1–5 ha in area (with conditional inclusion probability = 0.08)]. This is illustrated in Figure 3A-1.

6. A random number, r, between 0 and 1.0 was chosen, and the lake located or overlapping the corresponding distance on the line was selected as the first Tier 2 lake. This random number then was incrementally increased by 1 and each corresponding lake encountered along the array was selected as belonging to the Tier 2 sample (Figure 3A-1).

7. A similar procedure was followed for the lakes > 500 ha. The list of lakes was ordered geographically so four lists were developed based on the sampling year of the hexagon with which the larger lake was associated. Each lake was assigned a length proportional to its inclusion probability, and the selection process (step 6) was repeated.

8. The actual number of Tier 2 samples selected annually is shown in Table 3-2 for both years.

The Tier 1 and Tier 2 lakes selected, with coordinate locations and the corresponding EPA Region, were distributed to the EPA Regional Environmental Services Division Offices for quality assurance to identify frame errors and noninterest lakes. Figure 3-4 shows the national distribution of Tier 2 lakes for all years. This basic procedure was used both years, with slight differences. For example, clusters were manually determined in 1991, whereas they were drawn with a computer algorithm in 1992. Also, 1991 sample lakes were drawn from lakes of 1–2,000 ha, with none above 2,000 ha, whereas lakes with areas of 1–500 ha were drawn separately from lakes > 500 ha in area in 1992.

Sample Expansion Factors

We have asserted several times that a fundamental feature of the design is the selection of each sample unit (lake in this case) with known probability. Why is this important? What is done with the information? In the end, the scores or attributes for each lake are weighted; the weights are expansion factors allowing us to make inferences to the entire population based on the sampled lakes.

Sample expansion factors can be interpreted as the number of lakes in the population represented by each sample lake. Therefore, the sum of the expansion factors estimates the total number of lakes in the population, or in a defined subpopulation if the summation of expansion factors is limited to sample lakes in the subpopulation. Expansion factors are calculated as the reciprocal of the lake's probability for inclusion. For example, if there are 1,000 lakes in the population and each is selected with probability of 0.05, then the sample expansion factors all equal 20 (i.e., 1/0.05). Under this scheme, the number of lakes actually selected for sampling varies around an expected value of 50 (= 0.05*1000).

If instead, 100 of the 1,000 lakes are selected with probability 0.14, while 900 are selected with probability 0.04, the expected number of lakes would still be 50 = 0.14 x 100 + 0.04 x 900. The expansion factors for the two groups of lakes would be 7.14 = 1/0.14 and 25 = 1/0.04; each sample lake in the first group would represent about 7 lakes in the population, whereas each sample lake in the second set would represent 25 lakes in the population.

Tier 1 Probability

In practice, how are inclusion probabilities determined? The probability of selecting a lake in Tier 1 is given by the proportion of the land area captured by the EMAP sample areas (usually 40-km^2

hexagons) relative to the area of the base tessellation (usually 635-km^2 hexagons). Thus, by the initial randomization of the grid:

$$\text{Tier 1 inclusion probability, (1X hexagons)} = P_{T11,1X} = 1/16$$

Because we sample on a four-year cycle, each year one-fourth of the Tier 1 sample is selected. Thus:

$$\text{Tier 1 Year 1 inclusion probability (1X hexagons)} = P_{T1Y1,1X} = (1/16)/4 = 0.015625$$

This applies for any one of the four years.

Tier 2 Probability for 1991

Tier 2 lakes are selected for field visitation as described in Section 3.4. The probability of selecting a lake in the Tier 2 sample from the Tier 1 sample reflects the *proportion* of lakes from the Tier 1 sample that will be selected for the Tier 2 sample. For example, if the Tier 1 sample consisted of 1,000 lakes and our target Tier 2 sample size were 50, the conditional Tier 2 selection probability would be 0.05. Overall Tier 1/Tier 2 inclusion probabilities are the product of the Tier 1 inclusion probability and the conditional Tier 2 inclusion probability. The sample expansion factor is then calculated as the reciprocal of this overall selection probability:

$$\text{1X expansion factor} = \frac{1}{(P_{T1Y1,1X} * P_{T2})}$$

For example, the probability of selecting a lake of 20–500 ha is (1/64)*(0.475) = 0.007421875, and its associated expansion factor is 134.737. Thus, that lake represents 134.737 lakes in the population and any indicator scores for that lake will be expanded to the population using this weighting factor.

Because we selected different proportions of lakes from the Tier 1 sample for the Tier 2 sample in each of the size classes, the conditional probabilities differ among size classes, as illustrated here for 1991:

Lake Size Class (ha)	Tier 2 Probability (P_{T2})	Expansion Factors
1–5	0.125	512.000
5–20	0.25	256.000
20–500	0.475	134.737
500–2,000	1.00	64.000

Tier 2 Probability for 1992

For the 1992 sample, the inclusion probabilities reflect a design goal of 800 Tier 2 lakes across the conterminous United States, with overselection in the two smallest size categories, as noted earlier. All lakes > 500 ha were selected for sampling at the Tier 1 level, so the Tier 1 probability for that size class is 1.00 over the four-year cycle, or 0.25 for each year. Table 3A-1 summarizes the 1992 National Lake Sample.

CHAPTER 4
A FRAMEWORK FOR EVALUATING THE SENSITIVITY OF THE EMAP DESIGN

4.1 INTRODUCTION

We continue to refine our ability to evaluate the sensitivity of the proposed EMAP design for estimating the condition, and changes in condition, of lakes and streams. Preliminary descriptions of design capabilities appear in the EMAP-Surface Waters Research Plan (Paulsen et al., 1991). Since that time, we have developed a general framework for investigators (indicator leads) to use in evaluating how well candidate indicators can be expected to perform. A variety of important components of variance affect our ability to describe status and estimate trends. The framework, based on a description and estimation of these components of variance, can be used to analyze descriptions of population and subpopulation status and to evaluate how extraneous variance might distort those descriptions. Structured as a sensitivity analysis, it also allows investigators to explore the effects of components of variance, sample sizes, years of monitoring, and levels of Type I and Type II errors on the ability to detect trends. Combined with an analysis of variance components on carefully selected available datasets and on data derived from EMAP pilot studies during the first several years, this evolving framework will provide insight into how well we can expect to describe status and to detect trends. It will also guide us in selecting indicators and estimating their sensitivity.

This chapter addresses the following:

- Components of variance that are important to consider when evaluating design sensitivity
- Estimation of components of variance using a general linear model as the framework
- Influence of extraneous variance on descriptions of population status
- Influence of extraneous variance on sensitivity to trend detection

Good, long-term data records from which we can estimate all the variance components needed are generally not available. For indicators of trophic condition, some statewide datasets can be used to illustrate the important components of variance. Additional effort will be devoted to describing these variance components for other biological indicators derived from EMAP pilots and a search for other data sets. This chapter provides examples, primarily of Secchi disk transparency (SD), total phosphorus (TP), and chlorophyll-a (chl-a). As outlined here, we can now describe the kinds of data sets that would be most beneficial for evaluating candidate indicators. We can also design appropriate studies into the pilot programs that will supply the necessary variance estimates during the first few years of the program (at least four years for estimation of some variance components derived solely from the pilots).

4.2 IMPORTANT COMPONENTS OF VARIANCE

It is critical that we evaluate the magnitude and effects of different components of variance to determine the utility of particular indicators in estimating status and trends. Furthermore, the choices for minimizing variance components, or their effects on estimates of status and trends, depend upon their relative magnitudes. In some cases, it might be cost effective to reduce the variance components associated with sample collection and processing; in other cases, it might be effective to increase the number of visits to individual lakes, or to increase the number of lakes visited at the expense of revisiting individual lakes. The following sections outline components of variance that could have important effects on status description and trend detection.

4.2.1 Population Variance (σ^2_{lake})

Population variance describes the measured differences among lakes in a regional population or subpopulation during the index period, that is, the status of the lake population. We use cumulative distribution functions (CDFs), histograms, or other population descriptors to express it. If no other forms of variation interfere with the sampling process, index snapshots derived from the randomly chosen sample of lakes would unambiguously express population variance.

4.2.2 Extraneous Variance

Extraneous variance comprises all variance components that are not a part of population variance, but that reduce the precision of estimates of status and trends. Some of these variance components are of interest themselves as possible indicators, but for the most part, they inhibit our ability to describe the population characteristics. Extraneous variance can be decomposed into several pieces as follows (components estimated here are symbolically identified):

4.2.2.1 Year Variance (σ^2_{year}; year effects)

Year variance measures the amount by which all lakes in a population or subpopulation are high or low in a particular year. The condition of regional populations of lakes fluctuates around a central value in the absence of a trend. In the presence of a trend, this variance component measures the year-to-year variation from the trend line or curve. Trend detection is fundamental to EMAP, and trend detection capability is very sensitive to the relative magnitude of this component of variance. This population level variation is sometimes called a year effect, as it is a measure of the amount by which all lakes in a particular year are above or below a long-term central value.

This common pattern of variation among lakes is caused by regional-scale factors affecting the population in a consistent way, such as regional-scale climatological conditions (e.g., wet years or dry years). We are unaware of any regional-scale databases on indicators of interest to EMAP-Surface Waters (EMAP-SW) from which we could obtain estimates of this component of variance. To our knowledge, no consistent biological monitoring programs on lakes and streams have been conducted over broad regions for many years. Some state-level sampling programs give insight into the magnitude of this component of variance at a statewide scale of resolution, that is, on a scale more local than the scales at which EMAP intends to monitor.

4.2.2.2 Lake-Year Interaction ($\sigma^2_{lake*year}$; interaction effects)

The condition of an individual lake fluctuates from year to year around its central value or around a trend for that lake. These fluctuations are responses to local effects operating at the individual lake level and are independent among lakes. This component of variance specifically describes that part of a lake's year-to-year variation not already accounted for by the year component. This interaction variance is a natural feature of the populations under study and may be of interest in itself as an indicator of stress or change in an ecosystem. It can be estimated by repeat visits to lakes across years. However, if lakes are sampled each year without revisiting during the index period, the variance estimated from the year-to-year differences confounds two components of variance ($\sigma^2_{lake*year}$ and σ^2_{index}, described next). There is no way to estimate this component of variance other than to revisit multiple lakes for several years.

4.2.2.3 Index Variance (σ^2_{index}; index effects)

Index variance is the aggregate variation seen by repeated applications of a sampling protocol for the indicator of interest during the index period. Index variance consists of several components that can be estimated. If index variation is unacceptably high, it is worth evaluating the components of index variance to determine a strategy for minimizing its effect. A useful strategy is to estimate σ^2_{index} by repeat visits during the index period and evaluate its magnitude relative to other variance components. If σ^2_{index} is unacceptably high, then determine which of the subcomponents described here can be reduced. We used this strategy in our pilot studies: First, identify the magnitude of σ^2_{index} and then decide if additional information is needed. The following is a partial list of components of index variance that might be targets for reduction in overall index variance.

- *Measurement variance* (σ^2_{meas}) arises from the measurements made on the samples collected. Most quality assurance/quality control (QA/QC) protocols target measurement variance in documenting data quality. This component includes variation introduced anywhere in the sequence of events from the point of sample collection to the final resting place of the validated/verified data point in the database.

- *Team-to-team variance* comes from differences introduced by different sampling teams that use the same protocol. In regional-scale monitoring programs, it is not feasible for a single team to visit all lakes selected for monitoring. However, the magnitude of team-to-team variance can be reduced or controlled by using as few crews as possible. Training, experience, and follow-up audits during sampling are also effective.

- *Protocol (or local spatial) variance* expresses the variance that arises from applying the same protocol instantaneously. Suppose a protocol requires random selection of three sites over the profundal zone of a lake; protocol variance expresses the variation seen in the index values derived from repeated applications of this protocol during a single sampling visit. Modification and standardization of protocols can minimize this component if reduction in its magnitude is cost effective relative to minimizing other components. If the index sampling protocol calls for collecting samples at various locations in a lake and consistent spatial patterns can be detected, their effect can be removed.

- *Temporal variance and trends* within the index period describe the variation and consistent temporal patterns in the indicator of interest during the index period. Identification and characterization of variance and trends within the index period can lead to removing the effects.

- *Remaining index variance* consists of all other unaccounted for variance, the indicator noise exhibited during the index period.

Neither $\sigma^2_{lake*year}$ nor σ^2_{year} are subject to reduction by methodological improvements as might be possible with index variance. These components of variance are natural features of the ecosystems we study, and just as we characterize the condition of lakes, we must also characterize these components of variance. We must both estimate their magnitudes and evaluate the resulting influence on the precision of estimates of status and trends. If these components are unacceptably high, the only way to improve precision is to increase the number of lakes sampled ($\sigma^2_{lake*year}$) or the number of years of monitoring (σ^2_{year}).

Figure 4-1 illustrates a design that allows separation and estimation of the major variance components identified in the foregoing subsections. Observations in each cell of the lakes/years matrix come from measurements made each year during the index period. The year effect

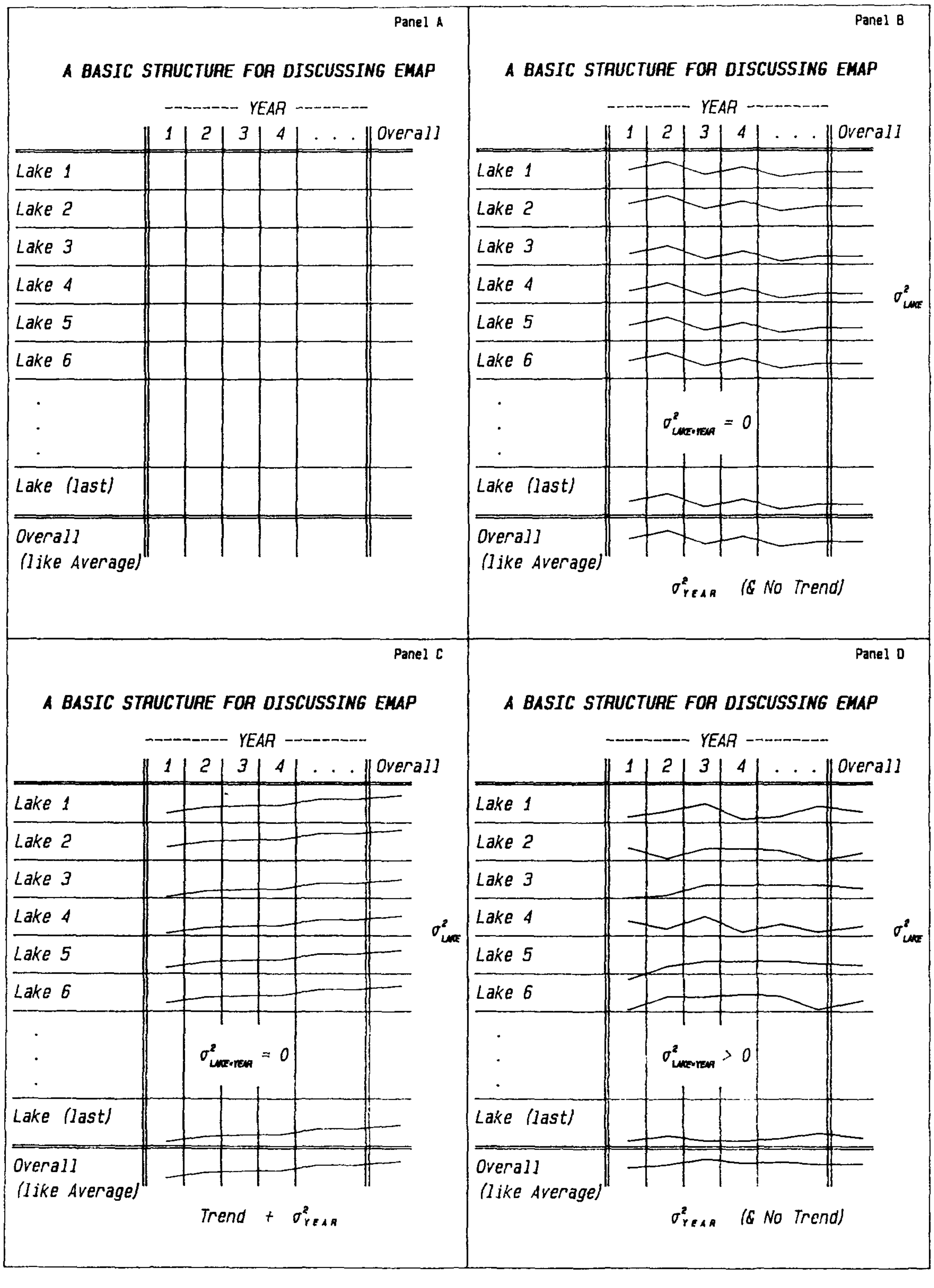


Figure 4-1. A schematic illustrating the sources of variation described in Section 4.2. Measurements made on each lake each year during the sampling index window fill cells in Panel A. Panels B, C, and D illustrate the relationships among σ^2_{lake} (population variation), σ^2_{year} (year effect), and $\sigma^2_{lake*year}$ (lake-year interaction). Index variance is zero for these illustrations.

(σ^2_{year}) is seen as the amount of variation exhibited by the lakes in concert (panels B and C); the lake-by-year interaction arises from the independent excursions of each lake (compare panel D with panels B and C). The EMAP-SW pilot studies are intended to characterize one or more of these sources of extraneous variation. The EMAP-SW staff would like to know about any databases derived from similar sampling programs for physical, chemical, and, most important, biological indicators of lake and stream condition (discussed in Chapter 2).

In the remainder of this chapter, we use indicators of trophic condition to illustrate the design, components of variance of interest, and status and trends sensitivity. We have begun to obtain information about the broader range of indicators of the condition of lakes and streams and will continue to do so.

4.3 A LINEAR STATISTICAL MODEL FOR ESTIMATING VARIANCE COMPONENTS

A convenient way to summarize the description of components of variance and to set the stage for estimating the magnitude of these components is to use a linear statistical model:

$$Y_{ij} = \mu + L_i + T_j + E_{ij} \tag{1}$$

where i indexes lakes, $i = 1, 2, \cdots \ell$; j indexes years, $j = 1, 2, \cdots t$; Y_{ij} is the condition of a lake at any particular time; μ is the grand average condition of lakes across the region of interest during the time interval of interest; L_i is the average difference of a particular lake from the grand average during the interval of interest [$L_i \sim (0, \sigma^2_{lake})$]; T_j is the difference between the regional average lake condition during any year and the grand average [$T_j \sim (0, \sigma^2_{year})$]; E_{ij} is the residual term [$E_{ij} \sim (0, \sigma^2_{res})$; $\sigma^2_{res} = \sigma^2_{lake*year} + \sigma^2_{index}$].

In terms of the model components, individual index observations on each lake can be expressed as illustrated in Table 4-1. The year means, expressing the regional snapshot each year, are:

$$\bar{Y}_{\cdot j} = \mu + \bar{L}_{\cdot} + T_j + \bar{E}_{\cdot j}$$

and the lake means, expressing the average condition of an individual lake across years, are:

$$\bar{Y}_{i\cdot} = \mu + L_i + \bar{T}_{\cdot} + \bar{E}_{i\cdot}$$

Table 4-1. Individual Index Observations (as in Figure 4-1) Expressed in Terms of a Linear Statistical Model

Symbols are defined in the text. Each cell in the matrix consists of one value; each value may be derived from repeat measurements during the index period or from several measurements taken to index the site.

Lake	Year					
	1	**2**	•••	**j**	•••	***t***
1	$\mu + L_1 + T_1 + E_{11}$	$\mu + L_1 + T_2 + E_{12}$	•••	$\mu + L_1 + T_j + E_{1j}$	•••	$\mu + L_1 + T_t + E_{1t}$
2	$\mu + L_2 + T_1 + E_{21}$	$\mu + L_2 + T_2 + E_{22}$	•••	$\mu + L_2 + T_j + E_{2j}$	•••	$\mu + L_2 + T_t + E_{2t}$
•••	•••	•••	•••	•••	•••	•••
i	$\mu + L_i + T_1 + E_{i1}$	$\mu + L_i + T_2 + E_{i2}$	•••	$\mu + L_i + T_j + E_{ij}$	•••	$\mu + L_i + T_t + E_{it}$
•••	•••	•••	•••	•••	•••	•••
ℓ	$\mu + L_\ell + T_1 + E_{\ell 1}$	$\mu + L_\ell + T_2 + E_{\ell 2}$	•••	$\mu + L_\ell + T_j + E_{\ell j}$	•••	$\mu + L_\ell + T_t + E_{\ell t}$

Variance of (lake means) is given by: $\sigma^2_{lake} + \sigma^2_{year}/t + \sigma^2_{res}/t$; variance of (year means) is given by: $\sigma^2_{lake}/\ell + \sigma^2_{year} + \sigma^2_{res}/\ell$; variance of (overall mean) is given by: $\sigma^2_{lake}/\ell + \sigma^2_{year}/t + \sigma^2_{res}/\ell t$. The overall mean and population variance can be regarded as the status of the population of lakes if there is no trend across years.

Under ideal circumstances, a two-way, balanced, replicated design (with r observations on each lake during the index interval each year), such as the one illustrated in Figure 4-1, would allow unambiguous estimation of the major components of variance outlined in Section 4.2. Data can be summarized with an analysis of variance table that highlights sources of variation, degrees of freedom, and mean square error terms, as in Table 4-2. The linear model appropriate for Table 4-2 has an interaction term not included in the foregoing equation. Variance components can be estimated by equating observed mean squares to expected mean squares, expressed in terms of estimates rather than parameters (Table 4-2). The resulting system of equations is solved to obtain estimates of the variance components.

Table 4-2. Analysis of Variance Table Summarizing Important Components of Variance for EMAP-SW Lake Monitoring

TWO-WAY, BALANCED, REPLICATED		
Source	df	E (Mean Squares)
Lakes	$\ell - 1$	$\sigma^2 + r\sigma^2_{lake*year} + rt\sigma^2_{lake}$
Years	$t - 1$	$\sigma^2 + r\sigma^2_{lake*year} + rt\sigma^2_{year}$
Lake*Year	$(\ell - 1)(t - 1)$	$\sigma^2 + r\sigma^2_{lake*year}$
Residual	$\ell t(r - 1)$	σ^2

Symbols are defined in Section 4.3.

$\hat{\sigma}^2_{res}$ = MS Residual

$\hat{\sigma}^2_{lake*year}$ = (MS Lake*Years – MS Residual)/r

$\hat{\sigma}^2_{lake}$ = (MS Lakes – MS Lake*Years)/rt

$\hat{\sigma}^2_{year}$ = (MS Years – MS Lake*Years)/$r\ell$

Evaluating the potential sensitivity of the EMAP design requires estimates of these variance components. Few databases contain the spatial and temporal coverage needed for useful estimates, but occasionally, we find databases from which most components of variance can be estimated. In the near term, we may need to estimate some components from independent data sets. The Vermont Department of Environmental Conservation adopted a lake monitoring program that has developed a database on total phosphorus (TP), chlorophyll-*a* (chl-a), and Secchi disk transparency (SD) that covers many years (Smeltzer et al., 1989). In general, TP measurements are obtained once each year during the spring; chl-a and SD are obtained weekly or biweekly during the summer months. Site replicates are also collected, except for TP. Not all lakes are visited each year, which produces an unbalanced design, but for many years a substantial number of lakes was visited continuously, so that most cells in a design such as the one shown in Figure 4-1 can be filled.

We used this database to isolate as many components of variance as we could estimate. We selected sets of lakes for which data were available over many consecutive years (at least four) and then selected data from periods corresponding to our July-August index period. We transformed both chl-a and TP to $\log_e$, because variances increased in direct proportion to levels of chl-a and TP. We estimate and report on the resulting variance components, with one exception. In Vermont, TP was measured during a spring index period and no repeat measurements were made during this index period, preventing separation of $\sigma^2_{lake*year}$ and σ^2_{index}. Many state programs monitor TP in springtime to characterize trophic condition. Because so many lakes were monitored over many years, we considered it important to estimate the variance components we could and compare them with future summer estimates. We estimated variance components using results provided by SAS's General Linear Model (GLM) procedure, summarized in Table 4-3.

For the EMAP-SW 1991 pilot survey, the only major components of variance that could be estimated were the aggregate index variance and the population variance (confounded by lake-year interaction). Other components will be estimated as the monitoring program progresses across years (Table 4-4). Appendix 4A summarizes sampling and analytical methodologies for the indicators of trophic condition.

Table 4-3. **Summary of Estimates of Components of Variance for Secchi Disk Transparency (SD), Chlorophyll-*a* (Chl-a), and Total Phosphorus (TP) Derived from the Vermont Lake Monitoring Database[a]**

VARIANCE SOURCE	VARIANCE ESTIMATES		
	SD	Chl-a	TP
σ^2_{lake} (Population Variance)	4.203	0.229	0.412
σ^2_{year} (Year Variance	0.007	0.002	0.018
$\sigma^2_{lake*year}$ (Lake-Year Interaction)	0.628	0.178	0.089[b]
σ^2_{index} (Index Variance)	0.660	0.230	
σ^2_{meas} (Measurement Variance, a component of Index Variance	0.191	0.111	

[a] Grand means: SD = 4.82 m, chl-a = 10.0 μg/L, and TP = 11.7 μg/L. Both chl-a and TP were $\log_e$ transformed before calculating variance components.

[b] Only the aggregate of $\sigma^2_{lake*year}$ and σ^2_{index} could be calculated because single samples were obtained each year.

Table 4-4. **Estimates of Variance Components Derived from the EMAP 1991 Pilot Survey[a]**

Variable	No. of Lakes	No. of Observations	$\sigma^2_{lake}+\sigma^2_{lake*year}$	σ^2_{index}
Secchi disk transparency	75	99	4.61	0.092
Chlorophyll-*a*	75	99	1.23	0.046
Total Phosphorus[b]	74	97	1.084	0.214

[a] Lakes have been monitored only one year, so only two variance components can be estimated. Chlorophyll-*a* and total phosphorus were $\log_e$ transformed before calculating variance components.

[b] One lake had an extraordinarily high TP value of 8,740 μg/L; this lake was excluded from the analysis.

4.4 EFFECTS OF VARIANCE ON ESTIMATES OF STATUS

There are two types of uncertainty in our ability to describe the status of a lake population. The first occurs as a result of drawing a sample from a population, through the sampling process. This type of uncertainty is characterized by confidence bounds (e.g., 95% upper and lower confidence bounds around the CDF) calculated with estimates of variance associated with the Horwitz-Thompson algorithm as the primary method. A description of this type of uncertainty estimate is not covered in this chapter, because the general approach has been documented elsewhere (Kaufmann et al., 1988; Kaufmann et al., 1991). Modifications pertinent to EMAP-SW will be documented in a future publication. The second type of uncertainty arises from the effects of extraneous variance on estimation of status. These effects are described here.

One way to illustrate the effects of extraneous variance on estimates of status is through a series of graphs. EMAP will use CDFs (cumulative distribution functions) to characterize populations of lakes. Under ideal conditions, the CDFs display population variation unencumbered by the effects of extraneous variation (curve E of Figure 4-2) but with uncertainty associated with the sampling process, not shown here. An alternative expression of this distribution, with which most are familiar, is shown in Figure 4-3. The initial illustrations show normal distributions. One of the goals of the sampling program is to describe the shape of the distribution functions; a normal distribution might be unusual.

For the yearly snapshot of the condition of a regional population of lakes, the effects of extraneous variation can be seen as a distortion of the variance structure of the regional population; larger and larger amounts of extraneous variation spread the CDFs out further, corresponding to increasing ($\sigma^2_{lake} + \sigma^2_{year} + \sigma^2_{res}$), illustrated as curves A through D in Figures 4-2 and 4-3. Observe that the mean values remain the same (located where the curves cross in Figure 4-2, although this occurs only for symmetric base distributions). Also, the precision of estimating the mean decreases as extraneous variance increases:

$$e.g.,\ var(mean) \cong \sigma^2/\sqrt{n}\ .$$

The effects of this distortion are most prominent at one to two standard deviations away from the mean illustrated by the maximum vertical differences between the distorted curves and the undistorted curves [see Overton (1989), from which this illustration is derived; more detail is also given in the EMAP-SW Research Strategy (Paulsen et al., 1991)].

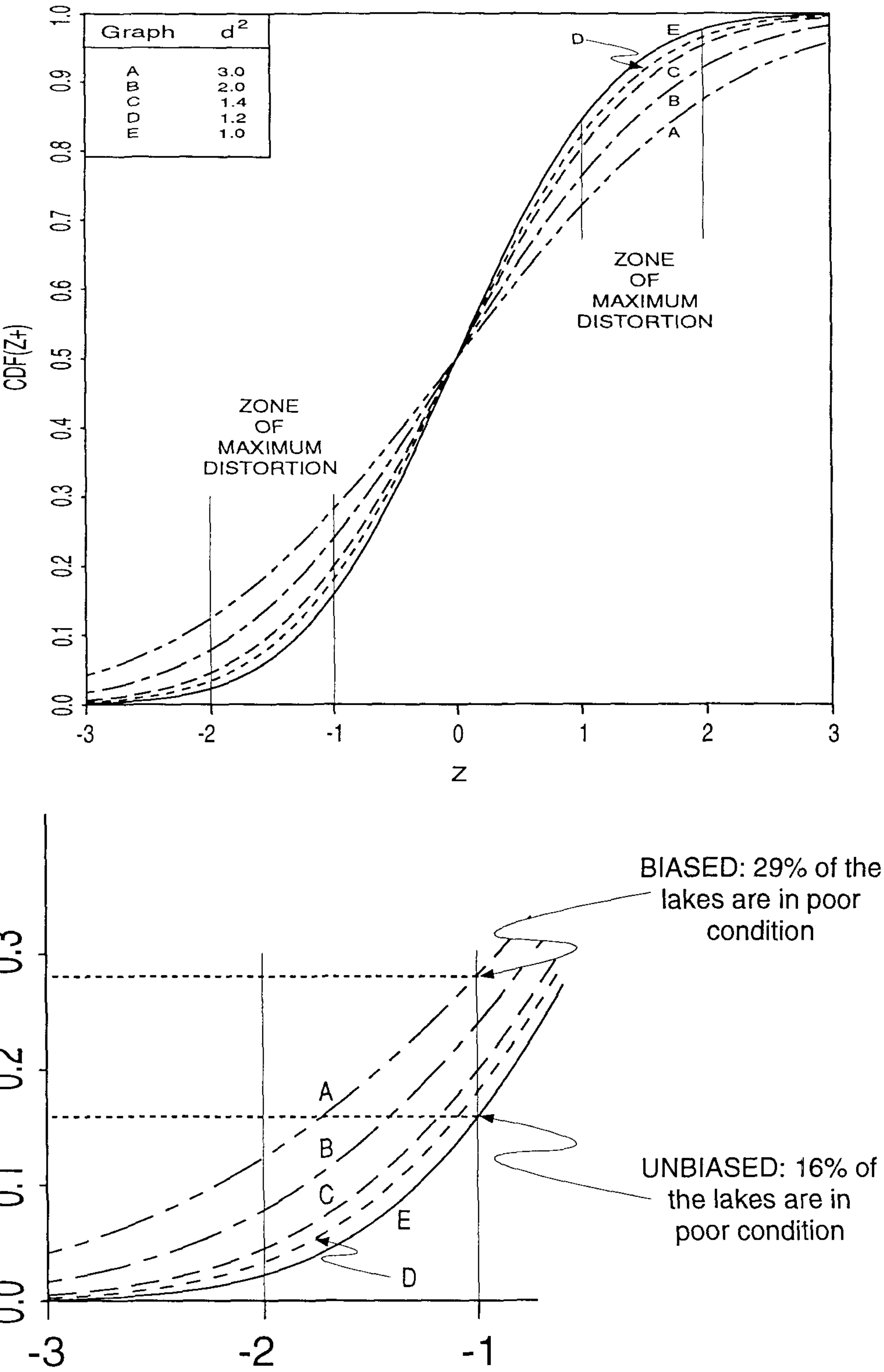


Figure 4-2. **Cumulative distribution functions illustrating the effects of increased extraneous variance on estimates of population status, where $d^2 = (\sigma^2_{lake} + \sigma^2_{res})/\sigma^2_{lake}$ (from Overton, 1989).**

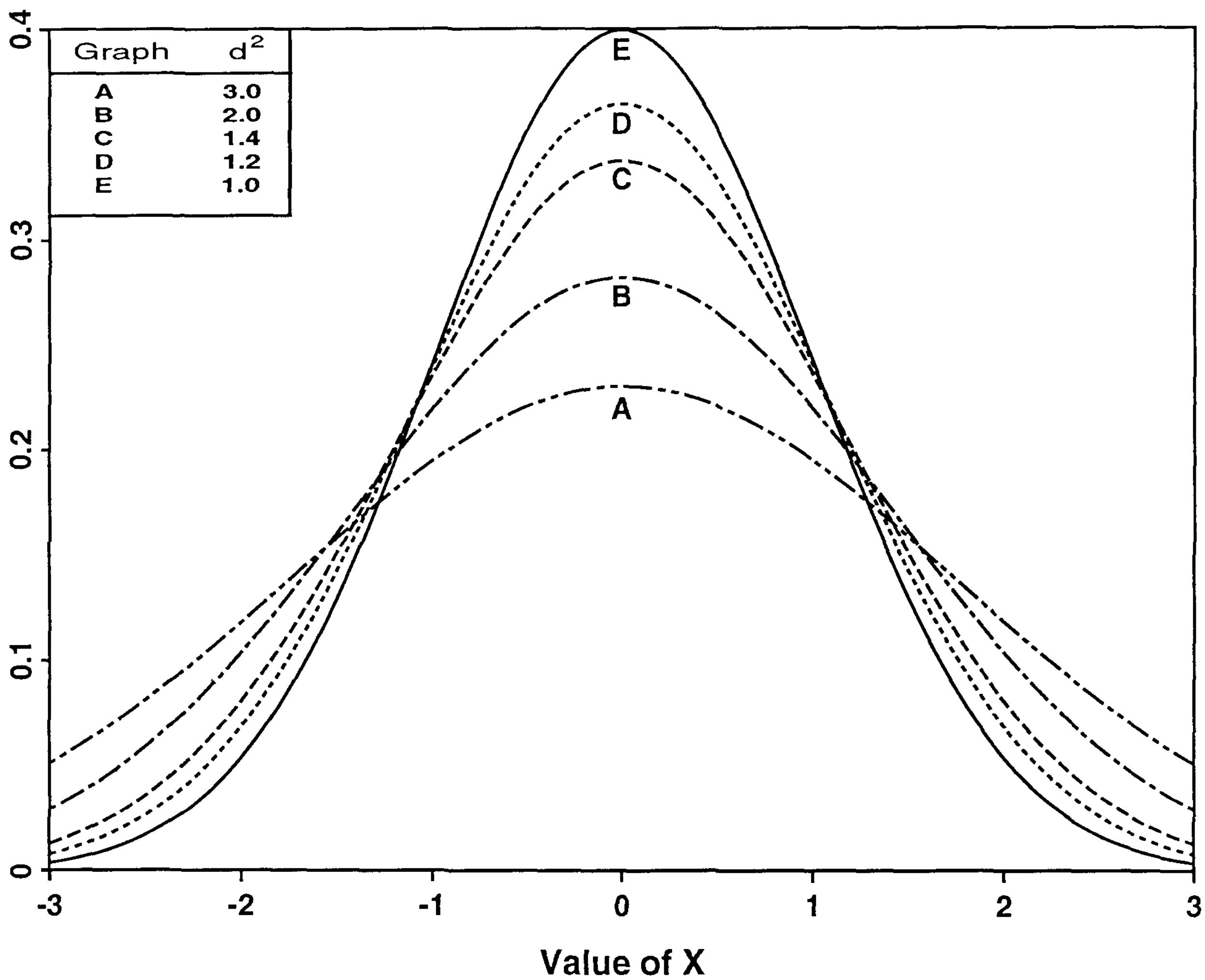


Figure 4-3. **Increasing extraneous variance spreads and flattens the normal distribution. Curves as in Figure 4-2.**

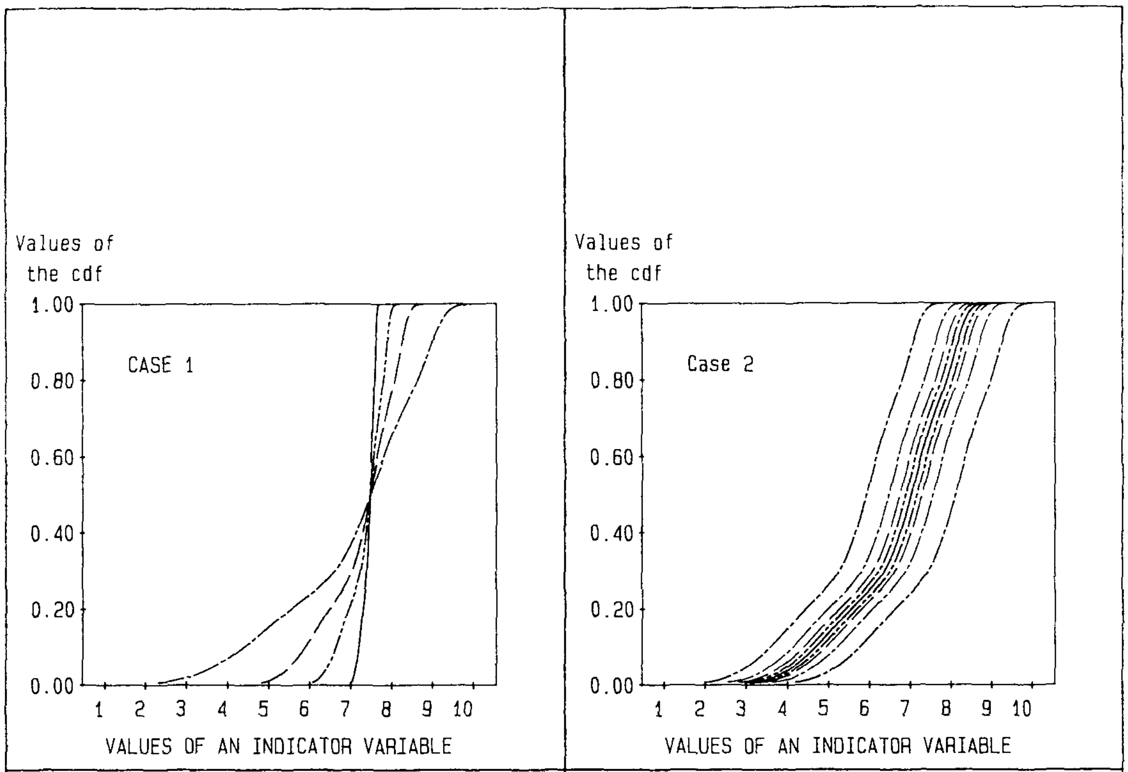
One way the effects of the distortion appear is by overestimates of the proportion of lakes with a condition less than a specified indicator score for scores below the mean, and by underestimates for scores above the mean. For example, consider the proportion of lakes falling below the condition specified by −1. In the undistorted situation, 16% of the lakes fall below this condition, but for the distorted situation (curve A, high extraneous variation), this percentage increases to 29%, a significant overestimate of lakes in hypothetically poor condition. This distortion of the CDFs means that greater proportions of the lakes appear in the extremes than in the undistorted situation.

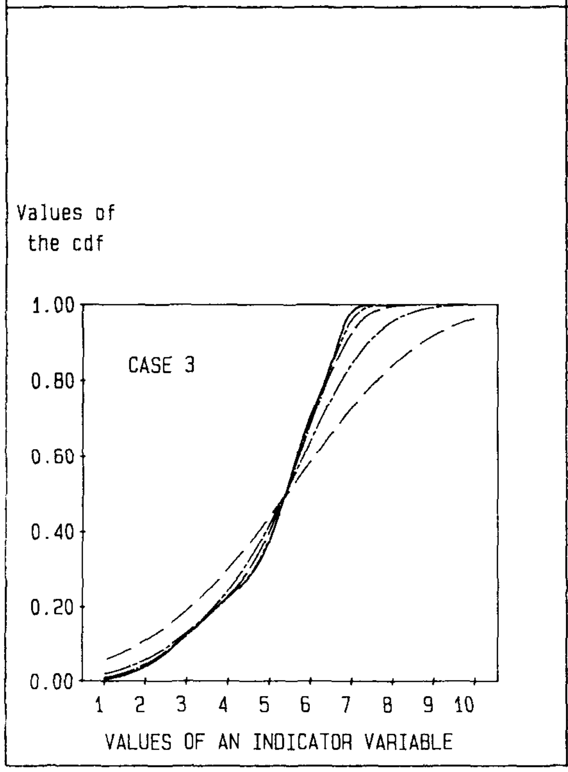
The effects of the distortion on a CDF are a function of the relative magnitudes of population and extraneous variance. A particular amount of extraneous variance might have a negligible effect on the CDF representing a subpopulation with large population variance, but the same amount of extraneous variance might impart significant distortion to the CDF for a subpopulation with low population variance. We will not know the relative magnitudes of these components of variance until we begin to obtain information on lakes selected in the probability sample and on important subpopulations of these lakes.

More complex scenarios can be modeled (see Appendix 4B), with the advantage that we do not have to make assumptions about the variance structure of the population or the extraneous variance. For illustration, we use an arbitrary, imaginary population of lakes consisting of three discrete subpopulations, one with relatively low indicator scores but with high extraneous variance, one with moderate indicator scores and moderate extraneous variance, and a third with high indicator scores but low extraneous variance. Population variance overlaps among the subpopulations. A series of panels (Figure 4-4) illustrates the effects of the components of variance on population descriptions (CDFs).

- **Case 1.** Year effects (σ^2_{year}) and residual variation ($\sigma^2_{lake*year} + \sigma^2_{index}$) are constant (or zero); we use these types of simulations to illustrate the effect of increasing population variance on the shape of the CDFs, which are the descriptors of the population characteristics to be depicted by EMAP-SW status summaries. The consequence of increasing population variance is stretching of the CDFs, as expected (Figure 4-4, Case 1). If the population exhibited no variation (all indicator scores identical), a vertical line would characterize the population.

- **Case 2.** Year effects (σ^2_{year}) operating on all lakes of a subpopulation shifts the entire CDF identically from its central location, for example, as might be observed after several years of monitoring (assuming no trends are present; Figure 4-4, Case 2). The greater the σ^2_{year} value, the broader the band within which the CDFs migrate from year to year. A particular yearly snapshot locates the CDF somewhere in this band; only after years of monitoring can we develop a reasonable picture of the status of the population.





CASE	$\sigma^2_{POPULATION}$	σ^2_{YEAR}	$\sigma^2_{RESIDUAL}$
1	0.11 1.78 7.11 28.44	0	0
2	2.13	0 0.25 1.00 4.00 16.00	0
3	2.13	0	0 0.077 0.3 1.16 3.25

Figure 4-4. Impact of variance components on the shape of cumulative distribution functions. Case 1 varies σ^2_{lake}; Case 2 varies σ^2_{year}; Case 3 varies σ^2_{res}. Parameter values used are summarized in the lower righthand panel.

- **Case 3.** As illustrated for the case of normal distributions, increases in residual variation (σ^2_{res}) flatten the CDFs in a pinwheel fashion, and also smooth the sub-population effects. However, if normal distributions are poor descriptors of the population characteristics, the pinwheel effect is not necessarily centered at the population mean (Figure 4-4, Case 3).

The design of the pilot survey allows calculation of the approximate distortion to the population CDFs for the indicators of trophic condition. A subset of lakes was resampled during the index period; the results from the revisits are used to estimate σ^2_{index}, a part of extraneous variance distorting the CDFs. However, $\sigma^2_{lake{\star}year}$ cannot be separated from σ^2_{lake} from a single year's survey, so estimates of distortion are conservative; distortion is likely to be larger than illustrated here. Variance components are summarized in Table 4-4. For SD, index differences remained constant across the range of SD measured and the distribution of SD among lakes was approximately normal, so comparisons with Figure 4-2 are appropriate; estimated $d^2 = 1.1$, indicating only a slight amount of distortion.

For chl-a and TP, index differences increased with the magnitude of the measurements and population distributions were more closely lognormal than normal; thus, in both cases data were log (natural) transformed before variances were estimated. Results cannot be compared with Figure 4-2, which is based on normal distributions. Instead, distorted and undistorted curves were simulated with population and variance structure like that estimated from the pilot survey (Table 4-4 and Figure 4-5). For both TP and chl-a, distortion appears to be relatively minor.

What options are available if extraneous variance is too large? One is to examine the components of index variance to determine whether improved sampling or measurement protocols allow significant reduction in index variance. A second option is to routinely increase the frequency of revisits to the lake during the index period to refine the estimate of lake condition during the index period by averaging revisits, but this is only useful if σ^2_{index} is large. A third option is to estimate the extraneous variance by revisiting a portion of the lakes during the index period for several years, until an adequate estimate of the components of extraneous variance is obtained. Then this estimate is used to deconvolute the distorted CDFs, estimating the true shape of the population distribution in absence of extraneous variance. Until we know the relative magnitude of extraneous variance and its influence, we would be unwise to implement a study on the individual components of extraneous variance or to select a strategy to reduce it or account for its effects. During pilot studies, repeat visits to lakes will be routinely included to estimate this variance component.

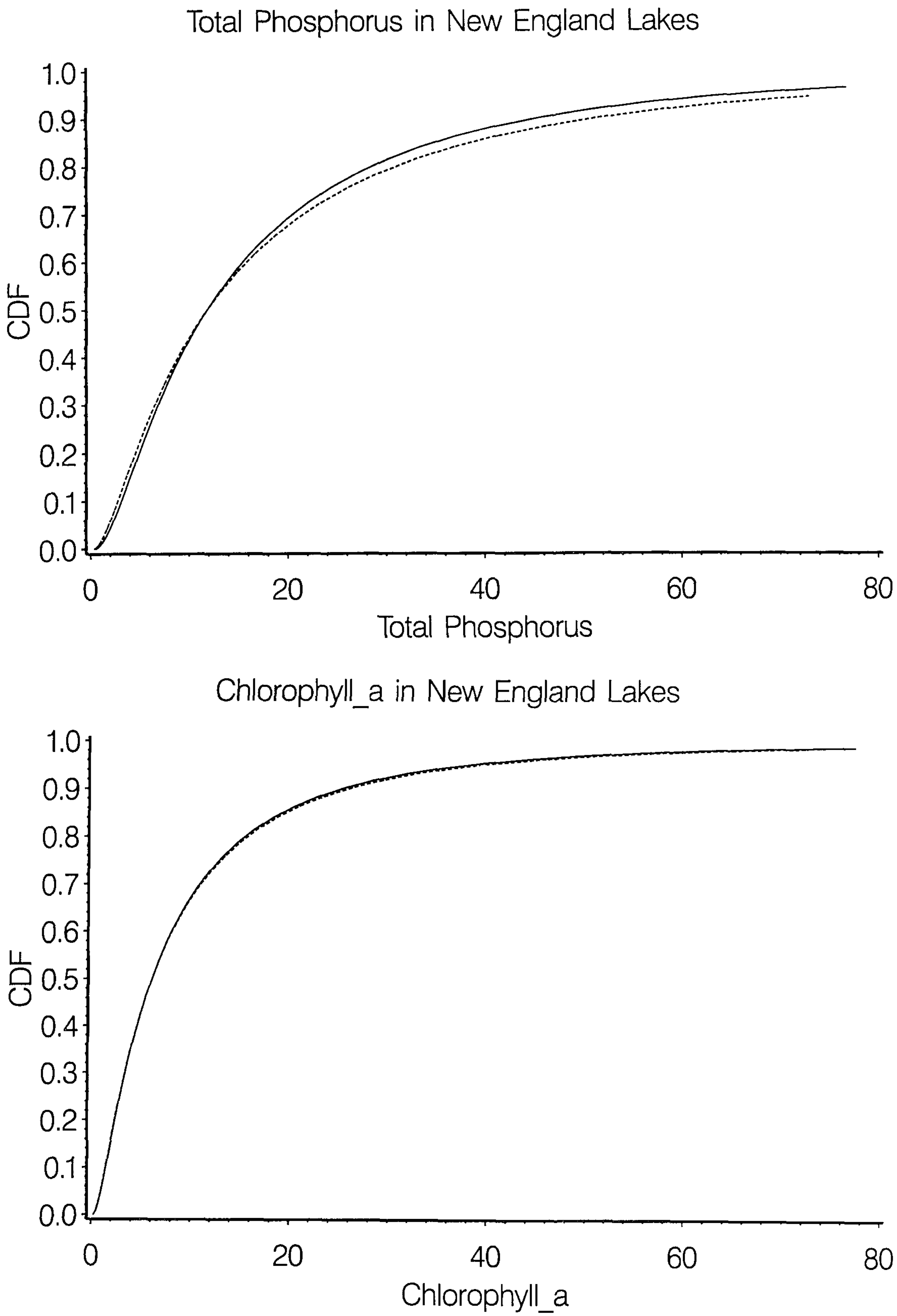


Figure 4-5. **Illustration of the approximate distortion to population CDFs resulting from observed extraneous variance estimated during the 1991 EMAP pilot survey. Solid line represents the approximate true CDF; dotted line represents the distorted CDF.**

4.5 EFFECTS OF VARIANCE ON TREND DETECTION

Extraneous variance also affects our ability to detect trends. Trends of interest to EMAP may not be evident until after at least 8 years of EMAP monitoring, although some trends might be detectable within 5 years, after the first set of lakes is revisited; trends of major magnitude that could be detected sooner do not need a design such as EMAP for detection. The following discussion focuses on a time frame of 8 to 12 years. After 8 years, two EMAP cycles will have been completed, and after 12 years, three cycles. For trend detection, population variance is of no concern; it is removed by the repeat visits to lakes in the 4-year cycles. This is analogous to the improved efficiency gained from the use of paired difference tests in experimental protocols, compared to an unpaired protocol. The following explanation first shows the basis of trend detection in EMAP and then makes several points about the effects of the variance components. At this point, it ignores some of the technical details, but these are described in Appendix 4B and incorporated into the explanation and the illustrations later. The version of the model used here includes annual revisits until year 5 to a subset of year 1 lakes.

To approximate the sensitivity of the design to trend detection, we assume the possibility of a linear trend and wish to estimate its slope, β. Refer to Section 4.3 for a description of the symbols. Consider the EMAP design after 8 years. Each set of lakes would have been visited twice. For any particular set of lakes, the yearly difference between the first and second visits (year 5) can be expressed as [symbols defined with Eq. (1)]:

$$(Y_{i5} - Y_{i1})/4 = (T_5 - T_1 + E_{i5} - E_{i1})/4$$

Averaging these differences across lakes in the first lake set (to be visited in years 1 and 5) produces an estimate of β:

$$(\bar{Y}_{\cdot 5} - \bar{Y}_{\cdot 1})/4 = (T_5 - T_1 + \bar{E}_{\cdot 5} - \bar{E}_{\cdot 1})/4$$

with variance of:

$$\begin{aligned} \mathrm{var}(\bar{Y}_{\cdot 5} - \bar{Y}_{\cdot 1})/4 &= (2\sigma^2_{year} + 2\sigma^2_{res}/\ell)/4^2 \\ &= (\sigma^2_{year} + \sigma^2_{res}/\ell)/8 \end{aligned}$$

Combining years 2 and 6, 3 and 7, and 4 and 8 similarly and averaging those results yields:

$$\hat{\beta} = [(\bar{Y}_{\cdot 5} + \bar{Y}_{\cdot 6} + \bar{Y}_{\cdot 7} + \bar{Y}_{\cdot 8}) - (\bar{Y}_{\cdot 1} + \bar{Y}_{\cdot 2} + \bar{Y}_{\cdot 3} + \bar{Y}_{\cdot 4})]/4 \times 4$$

$$\mathrm{var}(\hat{\beta}) = (\sigma^2_{year} + \sigma^2_{res}/\ell)/8 \times 4 \qquad (2)$$

In this illustration, the denominator in $\mathrm{var}(\hat{\beta})$ was $32 = 8*4$. The factor 4 comes from averaging the four estimates derived from each set of repeat visits. The factor 8 comes from $\Sigma(X_i - \bar{X})^2$, where X_i expresses years relative to some starting year, as in the variance of an estimated slope in simple linear regression. In the case of a 4-year interval, that becomes $(-2)^2 + 2^2 = 8$. For 12 years, namely 3 complete cycles of revisits, it becomes $32 = (-4)^2 + 4^2$, and after 16 years, it becomes $80 = (-6)^2 + (-2)^2 + 2^2 + 6^2$. The denominator of the variance term for trend estimation clearly increases dramatically over time; this is the source of the power, in designs such as EMAP, to detect progressively smaller trends the longer the monitoring program is in place.

Inspection of Eq. (2) reveals the relative importance of year effects (σ^2_{year}) and residual effects (σ^2_{res}) for describing the potential sensitivity of the EMAP design for trend detection. It allows rapid approximation of the effectiveness of choices for reducing or accommodating these variance components. Year effects cannot be controlled by any methodological improvements or by increased sample sizes. If trend detection sensitivity is controlled primarily by the magnitude of year effects, trend detection sensitivity increases only as the period of record increases for that indicator. In this case, expending effort reducing or accommodating σ^2_{res} would not be cost effective, because it would not substantially influence the outcome on trend detection. Another choice, of course, is to select other indicators less influenced by year effects, or develop explanations of the year effects, such as dependence on precipitation, which can be estimated and removed.

On the other hand, if residual effects are large relative to year effects, one option for increasing trend detection sensitivity is to increase the number of lakes monitored; however, a point of diminishing returns is reached as (σ^2_{res}/ℓ) declines with increasing numbers of lakes sampled. Another option is to evaluate the components of variance comprising σ^2_{index} to determine the extent to which methodological improvements will decrease its magnitude. Recall that $\sigma^2_{res} = \sigma^2_{year*lake} + \sigma^2_{index}$; σ^2_{index} is the only component subject to methodological improvements. The effects of $\sigma^2_{year*lake}$ can be overcome only by increasing sample size or years of monitoring.

If σ^2_{index} is relatively large, one possible option is to increase the number of times during an index period that individual lakes are sampled in order to estimate more precisely the lake mean condition during the index period. This option is worth considering only if the costs of repeat visits to lakes is substantially lower than the cost of sampling additional lakes. Eq. (2) may be slightly rewritten as:

$$\mathrm{var}(\hat{\beta}) = \{\sigma^2_{year} + [\sigma^2_{lake*year} + (\sigma^2_{index}/r)]/\ell\}/8 \times 4 \tag{3}$$

to illustrate the relative merits of revisits compared to sampling additional lakes. Here, r is the number of visits to individual lakes; r = 1 in Eq. (2). The sensitivity of var($\hat{\beta}$) to several reasonable scenarios of ℓ and r is illustrated in Figure 4-6. In this example, for situations in which the cost of revisits is about the same as the cost for sampling additional lakes, the choice clearly is to sample additional lakes. If the cost of revisits is 1/3 or less than the cost of additional lakes, it may be worth considering revisits instead of adding more lakes for trend detection purposes. In any case, with reasonable estimates of variance components, trend sensitivity can be estimated and options for increasing sensitivity can be evaluated. Tradeoffs such as these are also discussed in the EMAP-Surface Waters Research Strategy (Paulsen et al., 1991).

We can use the estimates of variance components derived from the Vermont dataset to illustrate the sensitivity of the EMAP design under a base set of conditions (e.g., 50 lakes per year over a period of 8 years with a variance structure like that of the Vermont lakes) and to evaluate sensitivity to some possible design choices for reduction in variance components. Eq. (2) yields an estimate of the variance of the slope. This can be translated into an estimate of the confidence interval for the slope, having a half-length of 2x the standard error of the slope for a 95% confidence interval. Cast as a null hypothesis, the test is whether the slope = 0. For the null hypothesis to be rejected, the slope must be > 2x its standard error to be detected at $\alpha = 0.05$; we should be reasonably certain of detecting trends of greater magnitude. In other words, reject $H_o{:}\beta = 0$ in favor of $H_A{:}\beta \neq 0$ if

$$\xi = \left| \frac{\hat{\beta} - o}{s.e.\,(\hat{\beta})} \right| > 2$$

with $\alpha = 0.05$; s.e. is the standard error.

Table 4-5 presents 2 s.e.($\hat{\beta}$), a measure of minimum observed trend, which implies real trend, or least significant trend. Consequently, it summarizes alternatives for increasing trend sensitivity by increasing the number of lakes visited or by revisiting lakes during the index period to better estimate individual lake condition. Assuming that the cost of revisiting is about the same as the

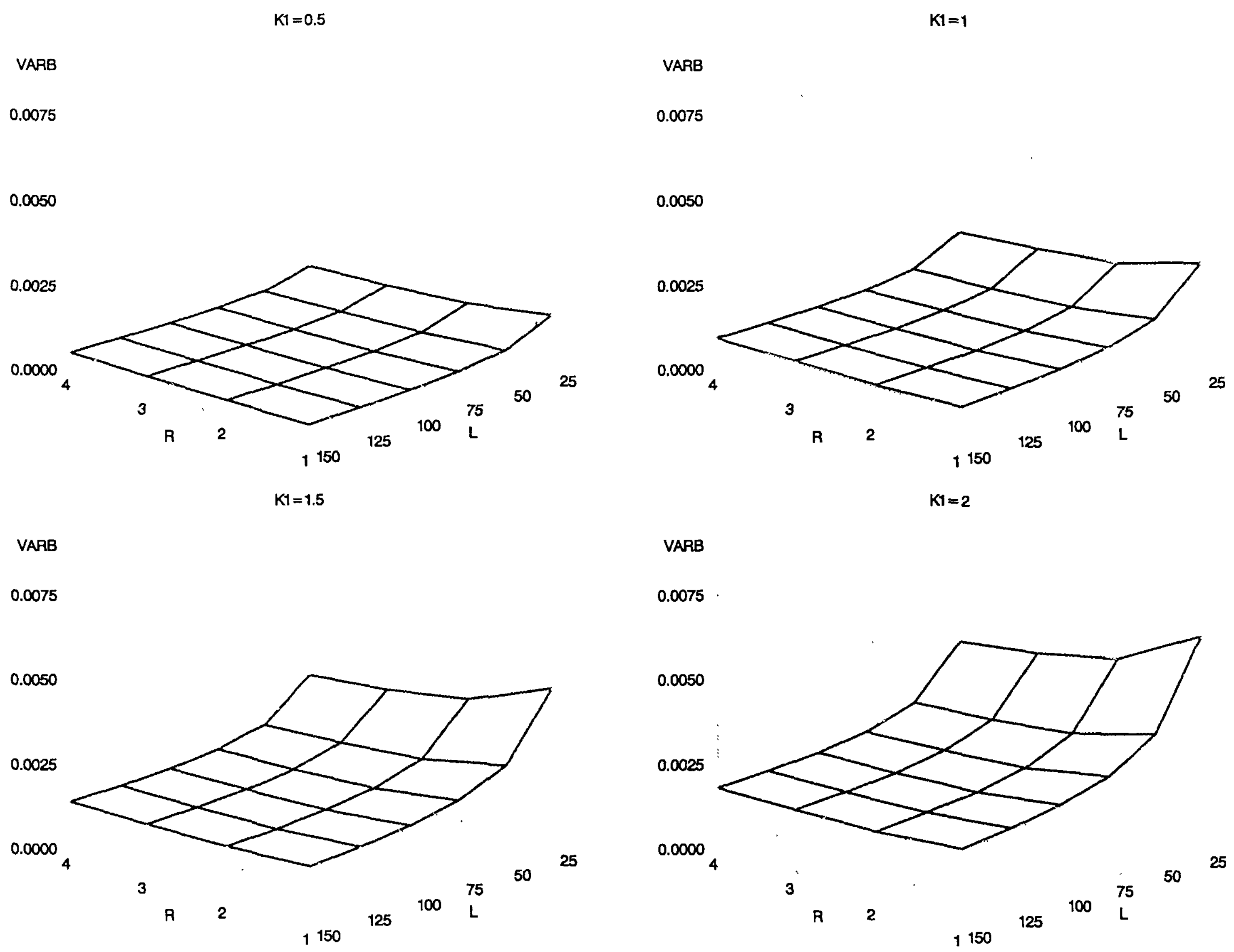


Figure 4-6. Trend detection sensitivity as a function of numbers of lakes visited (L) and numbers of revisits to individual lakes within an index period (R), from Eq. (3). Various panels reflect increasing levels of extraneous variance (σ^2_{year} + $\sigma^2_{lake*year}$ + σ^2_{index}), where σ^2_{year} = 0.02K1, $\sigma^2_{lake*year}$ = σ^2_{index} = 0.5 K1. Variances used approximate the Vermont Secchi disk transparency variances estimated (Table 4-3).

Table 4-5. **Least Significant Trend (m/yr), as Defined in the Text, as Allocating Sampling Effort to Revisits Versus Additional Lakes[a]**

r	ℓ	50	100	150
1		0.064	0.050	0.044
2		0.057	0.046	0.041
3		0.055	0.044	0.040

[a] An 8-year time frame is used; trends are Secchi Disk Transparency (SD). This example uses the variance components derived from the Vermont Secchi Disk transparency measurements. Units are: *m/yr = meters per year; ℓ = number of* lakes visited annually; r = number of revisits within a year.

cost of visiting additional lakes and that the variance components are as summarized in Table 4-3 for Vermont Secchi Disk transparency, revisiting lakes clearly would be an inefficient way of increasing trend detection sensitivity (compare sensitivity at ℓ = 50, r = 2 with sensitivity at ℓ = 100, r = 1 (Table 4-5).

Table 4-5 gives least significant trend, but cannot incorporate the ability of the design to detect trend. This quantity depends on the actual trend present in reality. Table 4-6 summarizes trend detection power for Secchi disk transparency, chlorophyll-*a*, and total phosphorus for 8 and 12 years of monitoring and for 50 and 100 lakes visited. EMAP posted a general goal to be able to detect trends of 1–2%/yr within 10 years, if such a trend were present, with $\alpha = 0.2$ and $\beta = 0.3$ (power = 0.7; McKenzie memo dated 7/13/92). Consequently, the reference values for amount of trend were chosen to be whole or half percentage values as close to 2%/yr as possible and for which power was not very small or extremely close to 1 (power = probability of observing a significant trend when there is a real trend). In contrast, the values in Table 4-5 are based on a useful approximation, associated with a consistent and fairly good estimator of trend. However, this approximation is (statistically) inefficient because it does not incorporate information about annual revisits to lakes, nor does it completely account for the effect of years. The model described in Appendix 4A provides a smaller s.e. $(\hat{\beta})$ than that given by Eq. 3.

These results imply that if variance structure is like that seen for the VT data sets, trends of the desired magnitude with the desired power should be easily detectable for SD; for chl-a, the target appears feasible, but it may take slightly longer than 10 years. For TP, although trends of the

Table 4-6. **Sensitivity of EMAP Design for Detecting Trends (%/yr) in Secchi Disk Transparency, Chlorophyll-a, and Total Phosphorus, Based on Variance Components Derived from the Vermont Lake Monitoring Database[a]**

Attribute	Number of Years	Number of Lakes Sampled 50	100
Secchi disk transparency	8	2.0 (0.95)	1.5 (0.94)
	12	1.0 (0.94)	0.5 (0.72)
Chlorophyll-*a*	8	2.5 (0.44)	2.0 (0.58)
	12	1.5 (0.62)	1.0 (0.54)
Total phosphorus	8	3.0 (0.53)	3.0 (0.55)
	12	2.0 (0.65)	2.0 (0.67)

[a] Variance components are used to estimate trend detection capabilities are from Table 4-3. Results are summarized as a %/yr change; power is included in parentheses.

desired magnitude may be detectable within 12 years, power is substantially lower than the desired level; more years of monitoring will be required to confirm such trends should they be present. A similar evaluation will be required for each indicator proposed for routine monitoring. TP demonstrates a case we must be alert for: $\sigma^2_{years} = 0.2\sigma^2_{res}$, so sensitivity increases only slightly with more lakes visited.

Another choice for improving trend detection is to execute our studies to minimize residual variance (σ^2_{res}). About half of σ^2_{res} (for the Vermont SD data) consists of $\sigma^2_{lake*year}$, a natural attribute of the lakes; of σ^2_{index}, > 70% comes from variation during the index period, also a natural attribute of the lakes under study. Relatively little variance is associated with the measurement process compared with these other components; thus, in this case, refining the measuring process will yield little return (Table 4-3). It may be possible to reduce σ^2_{index} if within-lake spatial and temporal patterns can be consistently detected. Then the effects of variance associated with these patterns can be removed.

The improvement in trend detection derived from sampling more lakes must be evaluated against the improvement in estimation of status derived from repeat visits conducted to deconvolute the CDFs. During initial phases of EMAP, it will be essential to allocate some effort to revisiting lakes to estimate this component of variance. It would be reasonable to drop revisits after enough years have passed to be confident that this variance component is stable; or, it may be important enough to continue to monitor this variance component as an indicator in itself.

The model described in Appendix 4B can be used to more thoroughly analyze the sensitivity of the design for trend detection. Sensitivity can be illustrated by describing changes in standard errors of estimates; the resulting curves typically approximate decaying exponentials. However, of greater utility in communicating the sensitivity for trend detection is a set of power curves, generalizing the values in Table 4-6, that demonstrate the probability of detecting a trend of a specified magnitude under specified levels of α, variance, and sample size. To illustrate trend sensitivity, we again use the variance components estimated for Vermont Secchi depth transparency but cast them as if they were representative of a regional population or subpopulation of lakes sampled in an EMAP context.

As was the case for illustrating the influence of components of variance on descriptions of status, we use a series of illustrations that vary different elements of the model individually (Figures 4-7, 4-8, and 4-9). The population characteristics and variance structure used here are derived from Vermont's SD database and from the variance components we have described. Population variance has no effect on trend detection because its effects are removed by the repeat visits in the design; α is set at 0.05 for the illustrations.

- **Case 1.** Power clearly increases as the magnitude of the trend to be detected increases (Figure 4-7). These kinds of curves can tell us how many years may be required to detect specified trends for variance structures characteristic of selected indicators and subpopulations of interest. We can also use these curves to demonstrate the magnitude of trend to be detected within a specified number of years at a selected power. In some cases, trends detected at low power might be enough to trigger management actions because the risk associated with delay in action if such trends are present might be high. On the other hand, if risks are low, it would be reasonable to continue monitoring to confirm the weakly detected trend.

- **Case 2.** Year effects (σ^2_{year}) have a pronounced influence on the ability to detect trends (Figure 4-8), and few design alternatives allow us to minimize these effects. If we select indicators for their monitoring importance, then high year effects mean that we either monitor for longer periods before making decisions about trend detection or accept lower power associated with any potential trends detected. It is likely that some indicators will be more sensitive to year effects than others, in which case, other factors being equal, indicators exhibiting lower year effects are desirable. It is also possible that within populations of interest, there may be subpopulations whose variance structure minimizes year effects. For trend detection purposes, it will be important to determine these subpopulations.

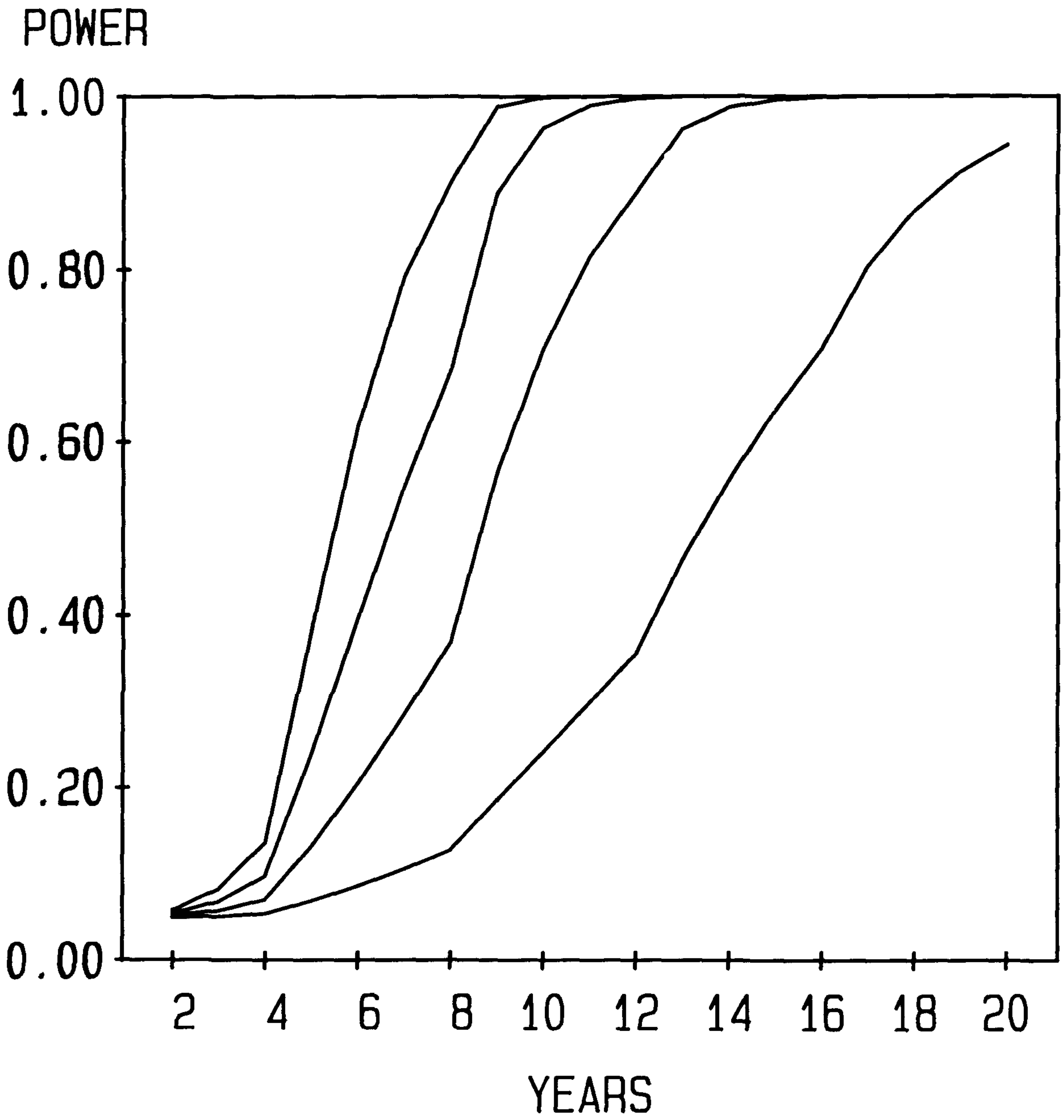


Figure 4-7. An illustration of the relationship between power and magnitude of trend detectable and years of monitoring. From left to right, curves are for trends of approximately 2%/yr, 1.5%/yr, 1%/yr, and 0.5%/yr.

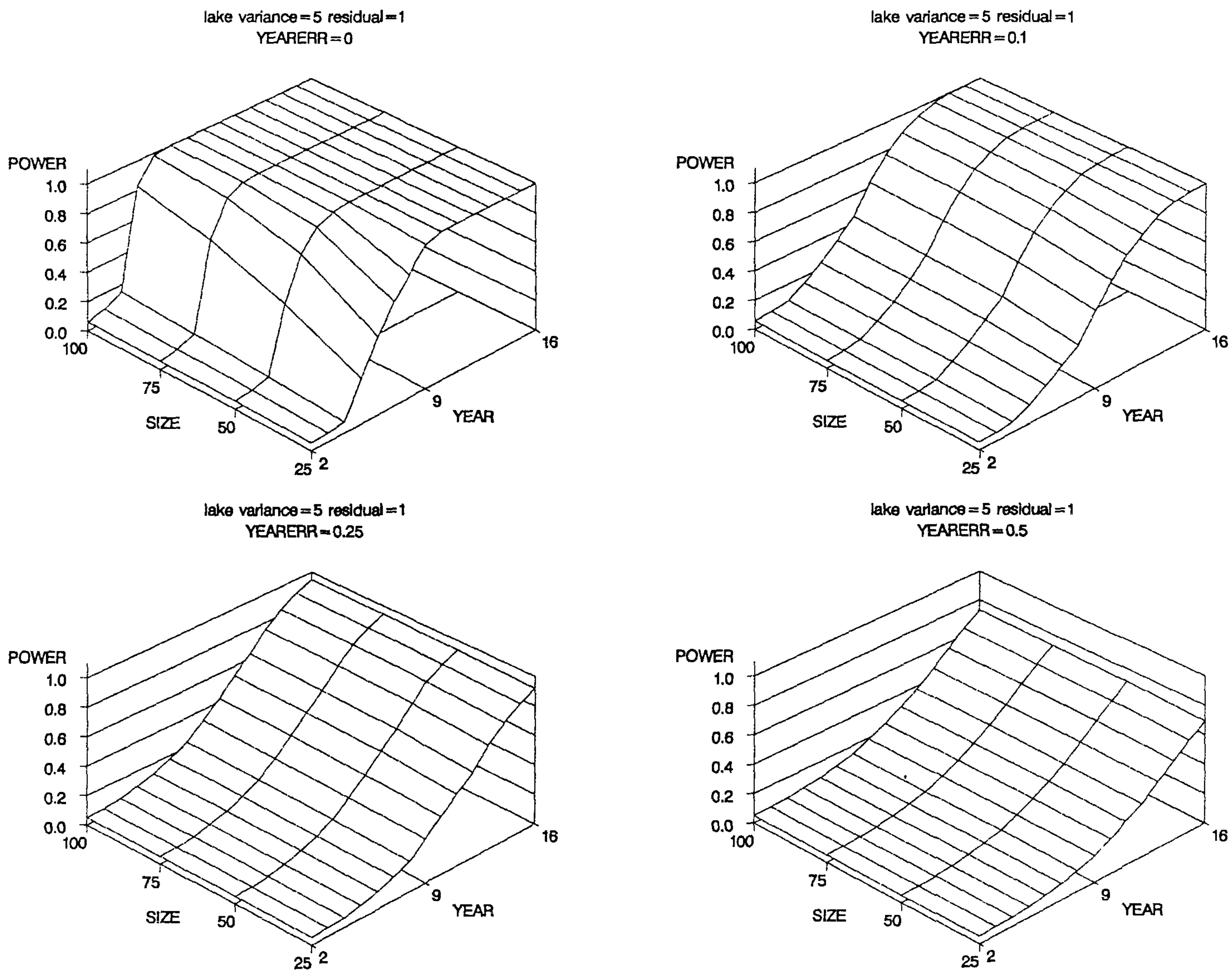


Figure 4-8. **Power to detect a 1%/year trend in Secchi Disk transparency. σ^2_{year} varies; σ^2_{res} is held constant.**

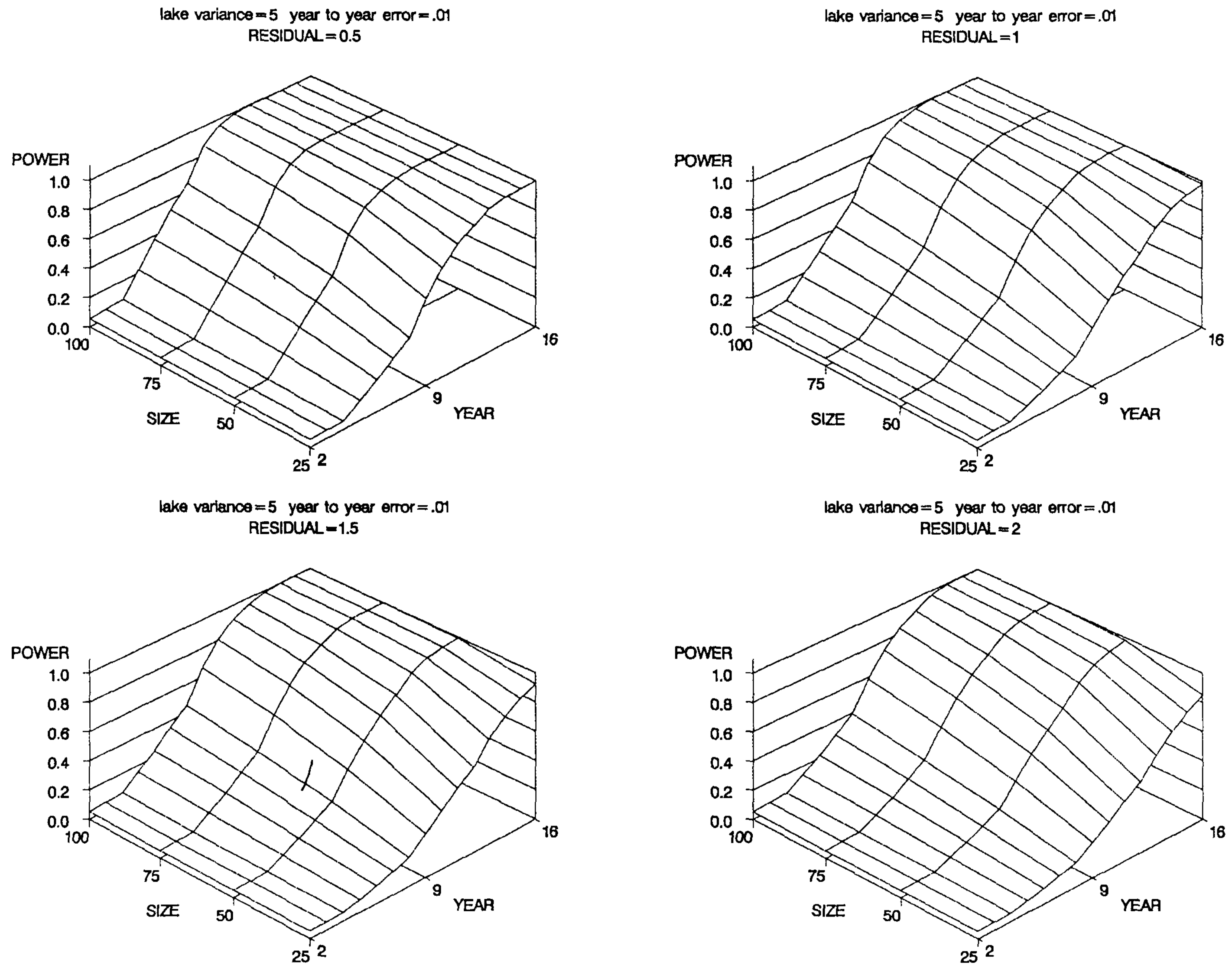


Figure 4-9. **Power to detect a 1%/year trend in Secchi Disk transparency. σ^2_{year} is held constant; σ^2_{res} varies.**

- **Case 3.** The power of trend detection is not as sensitive to the magnitude of residual variance as it is to year effects. For example, compare YEARERR = 0.5 in Figure 4-8 with RESIDUAL = 0.5 in Figure 4-9. Also compare YEARERR = 0.1 in Figure 4-8 with RESIDUAL = 2 in Figure 4-9. These cases, as well as other scenarios, indicate the extent to which power is dramatically more sensitive to σ^2_{year} than to σ^2_{res}.

The basic features of the trend detection model used in the foregoing scenarios have been assembled, as described in Appendix B. The next step is to convert the model into a form that indicator leads can use to explore the sensitivity of the design and design options for particular candidate indicators. A future step will then be to evaluate the responsiveness of the various candidate indicators together as a component in the determination of the set of indicators that would be used for routine monitoring.

4.6 A NOTE ON AVAILABLE DATABASES

Throughout the development of EMAP-SW, we have been aware of the need to identify databases that could be used to assist us in evaluating the potential for various biological indicators to serve as candidates for routine monitoring. Up to now, we have not performed a broad search for such databases because we had not described the kinds of data most useful for our EMAP studies and were unwilling to expend much effort along these lines until we could more explicitly define what kinds of databases would be best and how we were likely to use them. Now we are closer to specifying those types of databases, as described earlier in this chapter. We must also be cognizant of the costs associated with acquiring databases that require considerable modification and clean-up; for example, electronic databases are desirable; data contained in file cabinets or miscellaneous reports are undesirable; databases derived from consistent monitoring protocols over many years are desirable; databases compiled from many separate studies with differing purposes and unknown or undocumented quality are less desirable.

Typical of the kinds of databases available and that we have begun to acquire are those developed by state agencies, or other institutions with long-term monitoring interests, such as the Academy of Natural Sciences at Philadelphia, or individual research scientists who have been able to maintain long-term monitoring programs. A brief description of the kinds of databases we have acquired follows.

In the northeast, most states monitor lakes for indicators of trophic condition. For this report, we relied on a database acquired from the Vermont Department of Environmental Conservation. Other states in the northeast have also sent us copies of their databases. These are being compiled and evaluated for their utility in estimating the important variance components. Typical of these monitoring programs is the lack of data on a consistent set of lakes over many years; often state agencies target a selected set of lakes for one year or a few years, then switch lakes as interests and priorities change. Our strategy is to cull these databases for sets of lakes that have been monitored consistently over many years, then conduct the variance analyses on these subsets. The number of lakes and years over which data are available varies by indicator and by state. It would be ideal if we could extract a set of lakes from the various state databases that could be analyzed across the entire northeast.

During the early 1970s, the U.S. EPA conducted the National Eutrophication Survey (NES). The survey was limited to one year in any region, preventing its use for estimating some components of variance; however, multiple sites within a lake were sampled and most lakes were sampled three times during the year. This database is being examined to estimate site scale variance and seasonal variance. In general, the repeat visits within the year occurred over a longer time window than that anticipated for conducting EMAP lake monitoring, so variance estimates derived from an analysis of the NES data should exceed the estimates we encounter for EMAP.

A few states have ongoing monitoring programs that target biological indicators. To our knowledge, the Ohio Environmental Protection Agency (OEPA) has the most complete and extensive biological monitoring program among states. In combination with routine physical and chemical monitoring of streams throughout the state, OEPA also collects macroinvertebrate and fish assemblage data. These biological data are compiled into various indices of condition of the streams. This monitoring program has been in place for more than 10 years, covering more than 2,000 stream sites. We have worked with this database for many years for various purposes. With regard to its utility for EMAP, this database has one major shortcoming. The majority of sites are monitored only during a single year, so we are unable to estimate σ^2_{year}, an important component for trend detection. Furthermore, population variance is confounded by an unknown contribution by the stream*year interaction effects, separable only by revisits to streams across years. However, the database allows us to estimate index variance because repeat visits during the summer months are part of the routine fish monitoring. We have encouraged OEPA to begin collecting data on a subset of streams across years; data from such studies will be valuable for the states as well as for EMAP-SW.

Many other state agencies are beginning to monitor biological indicators of stream condition on a routine basis, stimulated by EPA's recent policy on the development of biological criteria. A few other states, besides OEPA, have been monitoring biological indicators of stream condition routinely, for example, Maine and North Carolina, but with the same shortcomings, for EMAP purposes, as the OEPA databases. As we begin our pilot studies on streams, we will acquire more of these databases.

The Academy of Natural Sciences at Philadelphia (ANSP) conducts biological monitoring programs on a variety of aquatic ecosystems (David Hart, pers. comm.). Some individual systems have been monitored routinely for decades. We recently completed an inventory of the monitoring programs ANSP has conducted and distributed an electronic database to appropriate indicator leads for further investigation of its utility for individual indicators.

Individual investigators often compile impressive databases for particular indicators on particular systems. These databases tend to be limited in geographic extent, thus they do not allow estimation of expected population variation or calculation of variance components across a broad range of sites. One such database that we are now compiling pertains to the Wabash River (primarily mainstem) in Indiana (Dr. James Gammon, pers. comm.). In this case, fish assemblages at many sites along the Wabash mainstem were monitored for more than 10 years. In addition, a variety of metrics have been calculated based on the fish assemblage raw data and are part of the database. This type of database will be quite useful for estimating all the variance components; the survey design is like the one described in Figure 4-1. For EMAP-SW purposes, the only limitation is the relatively limited geographic scope of the survey. We will use this database as an example of the kinds of databases most useful for our purposes.

We will continue to identify and acquire databases that appear useful for EMAP-SW purposes. Our experience so far has been that state agencies and individual investigators are quite willing to make the data sets available and encourage us to use them. We recognize the extensive amount of effort that has gone into obtaining the data and compiling the databases and we respect the proprietary interests of the principal investigators.

4.7 IMPLICATIONS FOR PILOT SURVEYS AND INDICATOR SELECTION

The preceding discussion develops a framework that identifies and describes components of variance that are important for estimating status and detecting trends in indicators of condition of ecosystems measured on sites selected with a probability-based survey design such as the one for EMAP. This framework can be used by indicator leads and others as a guide to aid in efficiently targeting research and pilot activities on elements important for surveys such as EMAP. In particular, the framework:

- Illustrates the kinds of data sets most useful for estimating the important variance components. Therefore, indicator leads can search the literature and cull their contacts for information on such data sets. It is possible that surveys of the type needed have been conducted by scientists and agencies in other countries.

- Suggests that available data sets will give only general insight into the magnitude of the components of variance of importance to EMAP monitoring. The magnitude of some variance components may be specific to region or lake type, so estimates derived from other regions or other types of lakes and streams might be misleading. In addition, the use of different methodologies for data collection and analytical procedures may result in misleading variance estimates, especially in attempts to combine data collected by various investigators into combined data sets. There is a risk of unknown magnitude if decisions are made based on data obtained from studies that are not like EMAP.

- Establishes a set of tools that indicator leads can use to evaluate how well particular indicators can be expected to perform for status estimation and trend detection. These tools allow the dissection of variance components with an objective of allocating sampling effort most efficiently, and can be used to evaluate current pilot results and make necessary modifications as needed. These tools will also assist in the design of particular studies, as for example, a focused study on important subpopulations of lakes and streams for detection of trends in acid neutralizing capacity associated with reduction in sulfate emissions mandated by the Clean Air Act Amendments.

- Indicates the need to factor into the routine pilot and demonstration studies collection of data to estimate variance magnitudes. It will be necessary to estimate the major components of variance as a fundamental part of the ongoing monitoring activities until enough data have been gathered to estimate the stability of the variance components.

Some variance components can be estimated only as part of ongoing monitoring activities based on a consistent survey design run over many years. These variance components are natural features of the lakes and streams, and, although they interfere with status estimation and trend detection, it is important to estimate them in their own right. Some will be estimated by virtue of the routine sampling design (year effects; lake-year interactions). It will be necessary to factor special studies into the pilot surveys to answer specific variance questions not answered as part of the basic survey structure. As the routine monitoring unfolds, and variance estimation stabilizes, only then will we have a sound basis for determining how well status can be characterized and trends detected. It is therefore important that pilot studies proceed to the monitoring of probability-selected lakes and streams with all due expediency. Results of evaluating available data, such as the Vermont data set, give us insight into what to expect. We have performed some preliminary analyses on other state databases that contain survey results of indicators of trophic condition. These results will be available in future reports.

4.8 SUMMARY

This chapter describes the variance components that impart important influences on EMAP's ability to describe status and detect trends. A general linear model is used as a framework within which we characterize and estimate the variance components; illustrations show how different variance components distort estimates of status and affect trend detection. A model for evaluating trend detection capability (expressed as power to detect trends of specified magnitudes), described in this chapter, can be used to examine the potential of the EMAP design to detect status and trends under different combinations of variance structure, years of monitoring, magnitude of trend, and Type I and Type II errors. A database on indicators of trophic condition available from the Vermont Department of Environmental Conservation is used throughout the chapter to illustrate the estimates of variance and their influence on status estimation and trend detection as a real world example.

The analyses presented suggest that the EMAP objective of detecting a 2%/yr change in 10 years (with $\alpha = 0.2$ and power $= 0.7$) is feasible for both Secchi disk transparency and chlorophyll-*a*, but several more years will be required before such trends can be detected in total phosphorus, if the variance structure of data collected in regions of interest to EMAP is similar to that of the Vermont database used in the illustrations.

APPENDIX 4A

Methods for Total P, Chlorophyll-*a*, and Secchi Depth for EMAP-SW 1991 Data

Analyte	Method[a]	Reference	Method Detection Limit
Total P[b]	Acid-persulfate digestion with colorimetric (molybdate blue) determination	Skougstad et al. (1979); U.S.EPA (1987)	0.51 μg/L[c]
Chlorophyll-a[d]	Field filtration of 250 to 1000 mL of water; spectrophotometric analysis	APHA Standard Methods (1989), Method 10200 H	2 μg/L[e]
Secchi Depth	20-cm black and white Secchi disk with calibrated line. Estimated as depth of disappearance plus depth of reappearance, divided by two.	Lind (1979); Chaloud et al. (1989)	---

[a] Field methods are described in Peck (1991).

[b] Precision of total P data was estimated from field performance evaluation samples and laboratory replicates as either the standard deviation (for concentrations < 20 μg/L) or as the coefficient of variation (CV) (standard deviation divided by the mean). For 1991 data, the mean CV of laboratory replicates was 2.1% and the mean standard deviation of field performance evaluation samples was 1.4 μg/L. Bias was estimated from quality control check samples as the difference between the mean value of repeated measurements and the target value of the QC sample, divided by the target value. Bias from the 1991 data for total P was -0.45%.

[c] Method detection limits were estimated from low concentration level quality control check samples and are calculated as a Students' *t* value at a significance level of 0.01 and n-1 degrees of freedom times the standard deviation of the repeated measures of the low-level QC sample (equation 5-1 in Peck, 1991).

[d] Precision of chlorophyll-*a* data was estimated with field performance evaluation samples and laboratory quality control samples as the CV; CV for the 1991 data ranged from 10% to 20%.

[e] Method detection limit was based on the minimum absorbance for a 250-mL sample.

APPENDIX 4B
A FLEXIBLE LINEAR MODEL FOR STATUS ESTIMATION AND TREND DETECTION

In this discussion, we use a flexible linear model and the tools of experimental design to illustrate the effects of various sources of variability on descriptions of the status of a population of a resource type, such as a population of lakes, and on trend detection capability. In doing so, we do not suggest that the approach exactly depicts every detail of the sampling situation. Nevertheless, the flexibility of the model allows us to explore the importance of various assumptions of the basic EMAP sampling design and characteristics of the indicators and resource populations of interest that influence status description and trend detection.

Consider an indicator of ecological condition (Y) evaluated at a lake in a particular year. Because the lakes have been randomly selected, we assume their effects are random in the sense of a random effect in a linear model. This random contribution of each individual lake may remain constant across years, but in other situations, part of a lake's effect may decay across years. For example, consider a deep water lake in which only 5% of the water volume is replaced annually. If such a lake is above the regional average one year, it is very likely to remain above the regional average the next year and for several following years. Likewise, because the years selected are of no intrinsic interest, we also assume this factor has a random effect, but years are consecutive. Part of the effect of years will be a consequence of atmospheric conditions, which may have some (auto)correlation through time. For both lakes and years, the correlation between responses will diminish, the farther apart in time the responses are. We model their covariance structure as a stationary Markov process. Machin (1975) used this covariance structure; Morrison (1970) investigated more general aspects of repeated measurements using the same covariance structure; Rao and Graham (1964) even used this structure in a study of sampling designs.

If lakes are revisited in some organized manner, as we anticipate here, sets of lakes will be visited in the same pattern of years. A reasonable plan, and the one assumed here, is that lakes to be visited will be partitioned into disjoint sets for which the visit pattern will be the same. Thus, if any lake in a lake set is visited in a year, all lakes in that set will be visited that year. Let i index these sets of lakes, $i = 1, 2, \cdots, \ell$; let j index years, $j = 1, 2, \cdots, t$; and let k *index lakes within a lake set,* $k = 1, 2, \cdots, n_i$. If the lake effects remain constant across years, then the random contributions to an indicator of ecological condition (Y_{ijk}) at a particular lake (i, k) in a particular year (j) can be modeled as:

$$Y_{ijk} = L_{ik} + T_j + E_{ijk}. \tag{1}$$

The three terms on the right in Eq. (1) represent the lake, time (year), and residual effects, respectively. The residual term includes all effects not otherwise modeled, including any lake by year interaction; more generally, the residual effect includes all random contributions specific to a lake during a particular year, including measurement error. If the effect of lakes decays with years, the lake term in Eq. (1) must be indexed also by *j*; this subscript will be enclosed in parentheses, however, to preserve the identification of lakes by their two subscripts, *i* and *k*, yet allow a correlation of lake effects across years of less than the unity implied by Eq. (1):

$$Y_{ijk} = L_{i(j)k} + T_j + E_{ijk}. \tag{2}$$

Suppose the random components in Eqs. 1 and 2 have variances σ^2_{lake}, σ^2_{year}, and σ^2_{res}, and are mutually uncorrelated except for those correlations among the $S_{i(j)k}$ and among the T_j explicitly discussed later. For now, we assume the lake components are uncorrelated between lakes; possible effects of an anticipated spatial correlation between lakes is the topic of a subsequent consideration. Define the matrix function $\Gamma_n(\rho)$ by:

$$\boldsymbol{\Gamma}_n(\rho) = \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \rho^3 & \rho^2 & \rho & \dots & \rho^{n-4} \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ . & . & . & \dots & . \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix}$$

If we denote the autocorrelation between year effects by ρ_T, then the covariance among the year contributions (T_1, T_2, $\cdots$, T_t) can be modeled as $\boldsymbol{\Sigma}_{year} = \sigma^2_{year}\boldsymbol{\Gamma}_t(\rho_T)$; likewise, the covariance among the responses at the same lake in consecutive years (namely $L_{i(1)k}$, $L_{i(2)k}$, $\cdots$ $L_{i(t)k}$) can be modeled as $\boldsymbol{\Sigma}_{lake} = \sigma^2_{lake}\boldsymbol{\Gamma}_t(\rho_{lake})$. This assumed structure is equivalent to requiring that:

$$L_{i(j)k} = \rho_S L_{i(j-1)k} + \sqrt{1-\rho_s^2}\ F_{ijk}, \tag{3}$$

where the F_{ijk} represent new independent parts of yearly random contributions to the lake effects satisfying $var(L_{i(1)k}) = var(F_{ijk}) = \sigma^2_{lake}$, but the F_{ijk} are otherwise uncorrelated. This correlation structure provides a way of investigating the impacts of possible correlations under our expectation that correlations between $L_{i(j)k}$'s (or T_j's), if present, will decline as the $L_{i(j)k}$ (or T_j) become farther apart in time. Our use of this structure does not imply that we believe it exactly models the correlation structure that exists in reality; it provides a suitable approximation that will support an investigation of possible effects of correlation. The case $\rho_T = 0$ and $\rho_S = 1$ deserves special notice because it corresponds to a randomized complete block experimental design with additive treatment (L_{ik}) and block (or year, T_j) effects. Thus, in practice, we expect ρ_T to be fairly small, say less than 0.25, and ρ_{lake} to be fairly close to 1.

If all lakes in a lake set are visited in the same year, the overall behavior of status and trend can be evaluated from the means in a table indexed by lake set and year. The averaging described here could be taken over the entire population, or over a relevant subpopulation. Thus, consider the vector of lake set-year means organized by lake set within year:

$$\bar{Y}' = (\bar{Y}_{11}, \bar{Y}_{21}, \cdots, \bar{Y}_{l1}, \bar{Y}_{12}, \bar{Y}_{22}, \cdots, \bar{Y}_{l2}, \cdots, \bar{Y}_{lt})'.$$

The general variance-covariance among these means can be represented in scalar form, using the Kronecker δ, as:

$$cov(\bar{Y}_{ij}, \bar{Y}_{i'j'}) = \delta_{ii'}\sigma^2_{lake}\rho_{lake}^{|j-j'|}/n_i + \sigma^2_{year}\rho_{year}^{|j-j'|} + \delta_{ii'}\delta_{jj'}\sigma^2_{res}/n_i, \tag{4}$$

or in matrix form, using the Kronecker product of matrices, as:

$$cov(\bar{Y}) = \phi = \Sigma_{lake} \otimes D^{-1}(n_i) + \Sigma_{year} \otimes 1_{lake} D 1_{lake}' + \sigma^2_{res} I_t \otimes D^{-1}(n_i), \tag{5}$$

where 1_n denotes a column of n ones and $D(\cdot)$ denotes a diagonal matrix having the indicated values on its diagonal. This variance-covariance matrix clearly depends on all of σ^2_{lake}, σ^2_{year}, σ^2_{res}, ρ_{lake}, ρ_{year}, and $n_1, n_2, \cdots, n_{lake}$, and the pattern of cells containing means.

Now consider the estimation of status and trend as if the trend were the same at all lakes and use generalized least squares incorporating the variance-covariance matrix given by Eq. (5). If we assume that the $E(Y_{ijk})$ are finite, then $\overline{Y}$ are *linearly sufficient* in the sense of Barnard (1963). Consequently, the generalized least squares estimator, based on all of the observations Y_{ijk}, actually can be expressed as a linear combination of the lake set-year means $\overline{Y}_{ij}$. Let $\mathbf{X}$ denote a (regression) matrix having as many rows as $\overline{\mathbf{Y}}$ has means, a column of ones, and a column of the year index associated with each mean. Estimates of the regression coefficients are given by:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\boldsymbol{\Phi}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Phi}^{-1}\overline{\mathbf{Y}}, \tag{6}$$

and their variances and covariances are given by:

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\boldsymbol{\Phi}^{-1}\mathbf{X})^{-1}. \tag{7}$$

The second element of $\hat{\boldsymbol{\beta}}$, the slope, clearly gives a measure of the trend. Status can be estimated by the value of the fitted line at the most recent time, namely, at time *t*. If the 2x2 matrix $\mathbf{C}_0$ has a first row containing 1 and t, and a second row consisting of 0 and 1, then $\mathbf{C}_0\hat{\boldsymbol{\beta}}$ estimates status and trend and has a variance-covariance matrix given by:

$$\mathbf{C}_0\,(\mathbf{X}'\boldsymbol{\Phi}^{-1}\mathbf{X})^{-1}\mathbf{C}_0'. \tag{8}$$

When a sampling plan is imposed on this context, it specifies the n_i's and the years in which specific lake sets are visited. Specifically, n_{ij}, the number of observations taken in each cell of the conceptual mean table, is either 0 or n_i, as specified by the sampling design. Consequently, a proposed sampling design can be defined in terms of the number of lake sets, the n_i, and the years in which each lake set is to be visited. If only some of the $n_{ij} > 0$, all of $\overline{\mathbf{Y}}$, $\mathbf{X}$, and $\boldsymbol{\Phi}$ need to be replaced by $\overline{\mathbf{Y}}^*$, $\mathbf{X}^*$, and $\boldsymbol{\Phi}^*$, where these represent merely that subset of rows (and columns, if appropriate) of the respective vector and matrices corresponding to cells where $n_{ij} > 0$.

Given values for σ^2_{lake}, σ^2_{year}, σ^2_{res}, ρ_{lake}, ρ_{year}, variances of the estimate of the status and trend can be obtained from Eq. (8). We used values from the Vermont set to choose values for the variance components: $\rho_{lake} = 1$ and $\rho_{year} = 0$, reasonable limiting values.

CHAPTER 5
PROTOTYPE ANNUAL REPORT

5.1 INTRODUCTION

The primary objective of the Environmental Monitoring and Assessment Program (EMAP) is to describe the condition of our nation's ecological resources on a regional scale. The Program provides data on the current status and extent of resources, and on changes and trends in condition. The production of monitoring data is not an end in itself, however; the data must be available as useful information in a timely manner. EMAP has set a goal of reporting the results of each year's field efforts, within 9-12 months. These annual statistical summaries must be more than extensive data tables. The information must be reported in a way that is readily understandable by policy makers, scientific audiences, and the public, including Congress. In order to produce reports within the specified time constraints, however, detailed interpretation of the results is not possible. A balance must be struck between summaries that can be produced in a relatively automated manner, yet be useful over time, and detailed interpretive reports that often take years to produce.

This chapter of the pilot report is intended to be a *prototype* of an annual statistical summary. It is limited to a few response indicators sampled at lakes selected by EMAP's probability design during 1991. It also contains only one year's data. A fairly complete description of condition will be based on data from the full four-year cycle. We have included the introductory material that we expect to produce every year with the annual statistical summary. Comments on the format and clarity of the presentation from various user groups will be most welcome and of tremendous value in improving these summaries[1].

5.2 OVERVIEW OF EMAP

The design and implementation of EMAP resulted from a recommendation by EPA's Science Advisory Board that EPA (1) implement a program to monitor ecological status and trends and (2) develop innovative methods for anticipating emerging problems before they reach crisis proportions. EMAP is therefore being designed to estimate the location, extent, and magnitude of

[1] Send comments to Dr. Steven G. Paulsen, U.S. EPA Environmental Research Laboratory, 200 SW 35th Street, Corvallis, OR 97333.

degradation or improvement in indicators of the condition of ecological resources at regional and national scales. This integrated monitoring program is addressing the following types of critical questions:

- What is the current status and extent of our ecological resources (e.g., estuaries, lakes, streams, forests, arid lands, wetlands), and how are they distributed geographically?
- What percentages of the resources appear to be adversely affected by pollutants or other human-induced environmental stresses?
- Which resources are degrading, where, and at what rate?
- What are the relative magnitudes of the most likely causes of adverse effects?
- Are adversely affected ecosystems improving as expected in response to control and mitigation programs?

5.2.1 EMAP Goals and Objectives

To provide the information necessary to answer these questions, EMAP has three major objectives:

1. To estimate the current status, changes, and trends in the extent (e.g., kilometers of streams, hectares of wetlands) and in indicators of the condition of the nation's ecological resources on a regional basis with known confidence.
2. To 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.
3. To provide periodic statistical summaries and interpretive reports on ecological status and trends to the Administrator and the public.

The EMAP monitoring design allows researchers to make statistically unbiased estimates of the status of ecological resources, trends in ecological condition, and associations among indicators of ecological condition. The probability design allows quantification of uncertainty over regional and national scales for periods of years to decades. The sampling approach offers a snapshot of conditions during an index period, rather than a detailed view of cycles within a year. This approach places some constraints on indicator selection and requires clear definition of assessment objectives and strategies for using information from multiple indicators.

5.2.2 EMAP Indicator Strategy

Society values several aspects of our ecological resources. These societal interests are sometimes described as environmental goods and services, or beneficial uses of ecological resources.

Biodiversity, sustainability, aesthetics, and production of food and fiber are examples of these *societal interests.* Society wishes to protect, preserve, and restore these kinds of biologically important, often conflicting interests. However, they are often difficult to quantify (e.g., sustainability, aesthetics, biodiversity). Indicators, as used in EMAP, are objective, well-defined, and quantifiable surrogates for important environmental and ecological values.

A long-term goal in EMAP is to measure biological indicators to determine whether resources of interest are in nominal or subnominal (acceptable or unacceptable) condition relative to a set of environmental or ecological values. Specific values of EMAP-Surface Waters (EMAP-SW) are described later. As an end result, EMAP should estimate the extent (i.e., number, length, surface area) of the resource of interest (e.g., lakes, streams) and the proportion of the resource in acceptable or unacceptable condition. In addition, the probable or possible cause of the poor conditions should be identified. Within EMAP, we define four categories of indicators to accomplish this goal—response, exposure, habitat, and stressor indicators.

1. *Response* indicators quantify the overall biological conditions of ecological resources by measuring either organisms, populations, communities, or ecosystem processes as they relate to values of concern.

2. *Exposure* indicators quantify the levels of pollutants to which the biota might be exposed, such as toxics, nutrients, or acidity.

3. *Habitat* indicators represent conditions on a local or landscape scale that are necessary to support a population or community (e.g., availability of snags, rocky stream bottoms, or adequate acreage or connectivity of wooded patches).

4. *Stressor* indicators reflect activities or occurrences that cause changes in exposure or habitat conditions and include pollutant export, management activity, and natural process indicators.

EMAP staff will use the response indicators to determine the condition of the systems, as described by the terms *nominal* and *subnominal*. Subnominal implies an unacceptable condition, and if a certain proportion (as yet undefined) of the systems are in subnominal condition, remedial action might be initiated, such as a specific management action focused on the types of lakes or streams affected. We will rely upon the exposure and habitat indicators to suggest possible cause for subnominal conditions on a regional scale. Stressor indicators can then be used to confirm or support the suggestion of probable cause, or they can be used alone as causal evidence for poor condition. It has been clear from the outset that defining nominal/subnominal conditions for most systems is a long-term process that will require several years to achieve.

5.2.3 Reporting in EMAP

An important objective of EMAP is to publish annual statistical summaries of indicators of the conditions of ecological resources and the exposure to stressors, as well as periodic interpretive reports. EMAP will produce data and information at varying levels of detail. Often monitoring programs or surveys report only the most detailed level of data possible. The reader is then left with massive amounts of data, but little real information, and must condense and synthesize the data into information independently. There are distinct advantages to this approach, but it is useful to a limited audience. On the other hand, as we summarize data into various levels of information, increasing amounts of interpretation are required. This condensation requires careful analyses and evaluation of the data in order to prevent development of spurious associations and erroneous conclusions. This level of care requires several years of effort and is not conducive to providing the users of the information with timely results.

EMAP will balance these conflicting needs by producing annual statistical summaries, timely, yet with enough data reduction to be useful to a range of audiences, and by producing periodic (probably every four years) interpretive reports with the full synthesis and interpretation needed and completed in a reasonable time frame.

5.3 EMAP-SURFACE WATERS

5.3.1 Legislative Mandate

Recognizing the extensive degradation of surface waters, Congress established their protection in 1972, 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. 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)].

Several other sections of the Act relate to EMAP objectives. Section 105(d)(3) requires 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 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, provide criteria sufficient to protect those uses, and establish an antidegradation policy protective of waters of outstanding quality.

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 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. Section 319 mandates identification of waters that cannot protect balanced aquatic communities without nonpoint source pollution controls.

Other legislation requires assessment of environmental risk to aquatic communities. Of particular importance are stressors with potential regional impacts that fall under the Clean Air Act Amendments, the Federal Lands Policy and Management Act, the National Environmental Policy Act, the Resource Conservation and Recovery Act, and the Federal Insecticide, Fungicide, and Rodenticide Act. 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.

There is certainly a need to evaluate the effectiveness of each of the laws and regulations, but there is also a need to determine if all of these regulations in the aggregate are resulting in the desired effect of protection of our aquatic resources. By describing the condition of our lakes and streams, EMAP-SW will determine the cumulative effectiveness of our protection efforts.

5.3.2 Issues and Problems

The variety of hazards to inland surface waters can be grouped into four broad categories for ease of summary and interpretation; several common types of hazards fall within each of these groups.

- Physical-Chemical alterations
  - Nutrient enrichment
  - Contamination (both point and nonpoint, as well as toxic and nontoxic)
  - Atmospheric pollutants (e.g., acidic deposition, other air toxics)
  - Thermal alteration
  - Global warming
- Physical habitat alterations
- Hydrologic alterations
  - Flow modification
- Biological alterations
  - Introduced species
  - Harvest imbalance (overstocking or overharvesting)

Each of these general categories of perturbations presents a potential threat to surface water condition, in the context both of traditional concepts of beneficial uses and of ecological integrity.

5.3.3 Surface Water Indicator Strategy

To be effective, monitoring data must be related to perceptions regarding aquatic condition and represent issues of concern to aquatic scientists and the public. In ecological risk assessment, these perceptions are called endpoints. One approach to determining these endpoints is to identify designated beneficial uses of aquatic resources. A *designated use* is attained if the system is being used for its intended purpose, such as habitat for aquatic life, fishing, water sports, aesthetics, navigation, or water supply. This approach has been ineffective for protecting biological integrity, because many designated uses, and the indicators for assessing use impairment, have little relationship to biological condition or integrity. Often the uses are poorly defined. For example, "warmwater fish" or "fishable" may be considered attained whether the water supports a few carp or abundant smallmouth bass, and "aquatic life" may likewise mean bluegreen algae or arctic char.

A host of perspectives and opinions exists concerning the appropriate environmental or ecological values on which a monitoring program such as EMAP should be based. Instead of selecting "beneficial uses" as the endpoints or values on which to base assessments and indicator selection, primary consideration was given to those ecological attributes or environmental values for which we manage surface waters. These were perceived to be:

- Biological integrity
- Trophic condition
- Fishability

The protection and restoration of the value "biological integrity" is mandated in P.L. 92-500, but biological integrity was not defined by the legislation. However, biological integrity has been described as "a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitat in the region" (Karr and Dudley, 1981). Fishability and trophic condition also appear as important values expressed in P.L. 92-500. These three values, though not parallel in perspective, are identified separately because managing for one is often to the detriment of the others; this is otherwise known as conflicting beneficial uses.

Ideally, a single indicator of each of the listed endpoints would form the basis for the annual statistical reports on the condition of aquatic resources. There would be one indicator for the overall condition of biotic integrity, one indicator for trophic condition, and one indicator for fishability. However, at this time, the statistical summaries report on some subindicators, such as chlorophyll-*a* (as an indicator of trophic condition). Much work has to be done over the coming years to integrate the various proposed measures that describe biological integrity and fishability.

By adopting the perspective that we are striving for a single indicator or index (aggregation of other indicators) for each of the ecological values, EMAP-SW concentrates on major values for which we manage systems and recognizes that they often compete. We may eventually be able to demonstrate with more consistent data the ways in which they conflict. We have recognized from the outset that we are not in a position to adequately describe condition with respect to these three endpoints, especially the nominal/subnominal aspects, or to compartmentalize the categories of possible cause as outlined here. It is, however, important to keep these long-term goals in mind while we evaluate the results of individual indicators presented in the annual summaries.

For fishability and biological integrity, the general categories of perturbations that have been discussed (e.g., chemical, hydrologic, biological, and habitat alterations) are the categories for probable cause of subnominal condition. Development of habitat, exposure, and stressor indicators is progressing in concert with the development of response indicators.

5.4 EXTENT OF SURFACE WATERS

One of the objectives of EMAP is to estimate the extent and spatial distribution of the ecological resources of interest such as lakes or streams. Estimates of the number, length, or surface area of these resources will be generated along with estimates of uncertainty. One advantage of a probability design such as EMAP is that maximum flexibility is retained for post-stratification to evaluate the results. In EMAP-SW, we will consistently report the results by both geopolitical regions (i.e., EPA regions) and some form of more meaningful ecological regions.

5.4.1 Lakes

The lake component of EMAP-SW will be the first to be implemented. For extent, the aspects of interest are the number, total area, and spatial distribution of lakes. As a lower limit of resolution, the program has chosen lakes of one hectare in surface area. Thus any estimates of extent or condition will apply only to lakes greater than one hectare.

Whenever a particular resource type is described, complications in definition always arise. The definition should be consistent across the United States; however, we recognize that the issue is not clear cut. There are regional differences in what local residents and legal authorities consider a lake. The definitions are somewhat dependent upon the abundance, or lack thereof, of lakes in that region. For example, a waterbody of 1–2 ha may be of great significance to local residents and wildlife in arid regions of the west, but barely recognized in areas of the northeast or upper midwest where there are many lakes. Generally, lakes are easily recognized, although small shallow lakes might also be considered wetlands. For operational purposes, the lake population targeted by EMAP-SW includes standing bodies of water > 1 ha, with 1,000 m^2 of open water and a maximum depth > 1 m.

5.4.2 National Estimates

Table 5-1 contains one estimate of the number of lakes in the lower 48 states, presented by each EPA region and as an aggregate by the entire area. This summary represents lakes drawn on the 1:100,000-scale U.S. Geological Survey (USGS) topographic series maps. These maps have been digitized and incorporated into EPA's River Reach File, version RF3. The waterbodies identified as lakes in these digital files comprise the estimates. These national estimates have not been verified other than by the USGS in its original mapping. Thus the estimates are based

Table 5-1. **Regional and National Estimates of Lake Number and Area[a]**

EPA Region	Lake Number	Lake Area (ha)
Region I and II[b]	21,725	963,330
Region III	10,739	149,761
Region IV	63,829	1,756,055
Region V	64,863	2,323,056
Region VI	46,034	1,635,451
Region VII	24,350	360,955
Region VIII	41,385	1,497,363
Region IX[c]	8,224	812,491
Region X[d]	9,060	567,385
NATIONAL TOTALS	290,209	10,065,847

[a] Data derived from U.S. EPA Office of Water, River Reach File 3 (U.S. EPA, 1991). These data represent estimates of waterbodies found on the USGS 1:100,000-scale map series and have received the quality assurance provided in the original cartography. These estimates will be evaluated during the EMAP-SW lake selection process (see Table 5-2) and will change as improvements are incorporated into River Reach File 3.

[b] Data for Regions I and II combined; does not include Puerto Rico or the Virgin Islands.

[c] Data for Region IX does not include Guam or Hawaii.

[d] Data for Region X does not include Alaska.

solely on the digital information available. They have been compiled and reported by the U.S. EPA (1991) Office of Water in an effort to provide consistency across the states in reporting on water quality. An extensive national effort to verify these national estimates, similar to that described in the next section for EPA Regions I and II, will be developed by the States and the U.S. EPA (Office of Water, Office of Research and Development, and Regions). Refined estimates can be reported when verification of the maps and routine monitoring of the extent of lakes is included as part of the basic landscape characterization of the spatial distribution of ecological resources.

5.4.3 Regional Estimates

During the 1991 pilot study, EMAP-SW considered only lakes between 1 and 2,000 ha in EPA Regions I and II. By taking a probability sample of all lakes, we can estimate the actual number of lakes present for comparison with the estimates from RF3. During the first stage of screening, we evaluated 311 lakes to identify map errors and waterbodies of noninterest for this phase of the monitoring program. We accomplished this by evaluating maps of different resolution and discussing our concerns with local managers. Forty-eight lakes were identified. The probability design allows inferring the size of the population of lakes of interest for the region of interest (Northeast). Further screening of the second-stage sample by field visits indicated that another 18 did not meet the definitional criteria. Table 5-2 lists the estimated number and area of lakes meeting the definitional criteria, based on one-fourth of the expected sample size over the four-year cycle; 11,455 waterbodies in the northeast between 1 and 2,000 ha meet our operational definition of a lake. The majority of the frame errors and noninterest lakes were 1–20 ha in size.

Table 5-2. Lake Number and Lake Area for Lakes 1–2,000 ha for the Northeast (EPA Region I and II) and Selected Subregions, Estimated from the Tier 2 Sample[a]

Class	Sample Size (# Lakes)	Estimated Number of Lakes	SE of Lake Number Estimate	Estimated Lake Area (km^2)	SE of Lake Area Estimate
Tier 2 sample lakes[b]	92	16,795	563	4,221	830
Tier 2 non-lakes	3	1,109	723	27	17
Tier 2 waterbodies that were too shallow, or wetlands	12	3,840	1,197	116	37
Tier 2, not visited	3	391	230	49	32
Tier 2 target lakes	74	**11,455**	1,251	4,030	814
Adirondacks	26	1,506	285	1,082	395
New England Uplands	29	5,669	1,206	2,099	758
Coastal/Lowlands/Plateau	19	4,280	1,048	850	254

[a] Estimates of population size (for both area and number) are accompanied by the standard error (SE) of the estimates.

[b] Includes additional lakes selected for acidification studies.

It is also important to describe the extent of the lake resource in specific subregions of the Northeast. One way is to identify ecologically important regions, such as those delineated by Omernik (1987), within which extent and condition can be described. Due to a limited sample size this first year, only three subregions were delineated: the Adirondacks, the New England Uplands, and an aggregation of the coastal, lowland, and plateau areas (Figure 5-1). Estimates for the extent of lake resources in these three regions are also displayed in Table 5-2.

5.5 CONDITION OF RESOURCES

5.5.1 Indicator Data for 1991

The following section contains results of data collected on specific response and exposure indicators during EMAP-SW field sampling in lakes of the northeast during July-August 1991. As outlined in Section 5.2, the long-term assessment goal for EMAP is to describe the proportion of populations in nominal and subnominal condition. The decision of what constitutes nominal and subnominal for specific indicators varies by region and requires extensive evaluations of "reference conditions" in a region along with a variety of subjective judgments. The statistical summaries will present data in a manner that minimizes the extent of subjective interpretation that precedes presentation of the data. Again, recall that these results, while based on 1991 data, are presented as a prototype; expect modifications as we refine the information contained in the annual statistical summaries.

One of the key tools for presenting information in the summaries will be the cumulative distribution function (Figure 5-2). The CDF is a snapshot of the complete population variation and allows us to estimate the proportion of the population above or below a particular indicator score. It provides the complete data for the population, with uncertainty estimates, with few if any value judgments imposed. Different readers can evaluate the same data with different criteria (Figure 5-2a) and come to their own conclusions.

If, as in Figure 5-2b, we select a score of 40 or less to distinguish poor condition for a particular indicator, we can estimate that 45% of the total resource (by number, area, or length) has a value of 40 or less and thus would be considered in poor condition. The uncertainty about the estimate, in this case, $\pm$15%, can also be estimated. Although this may appear complicated to many readers, it is a relatively simple way of presenting the complete data set with no judgment about particular indicator scores of concern. The reader can determine which scores are of

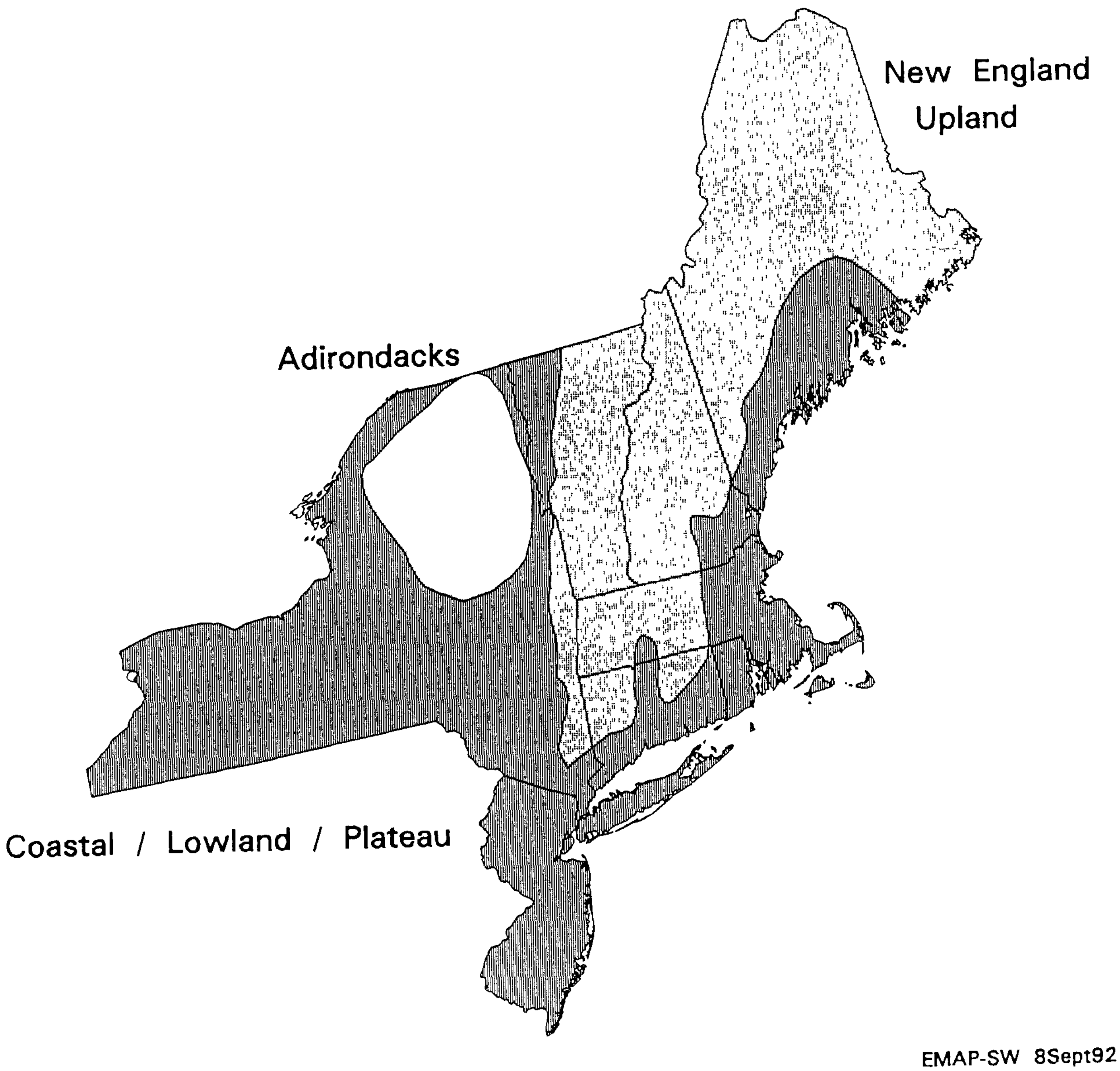


Figure 5-1. A map illustrating three ecological regions on which statistical summaries can be based for routine reporting.

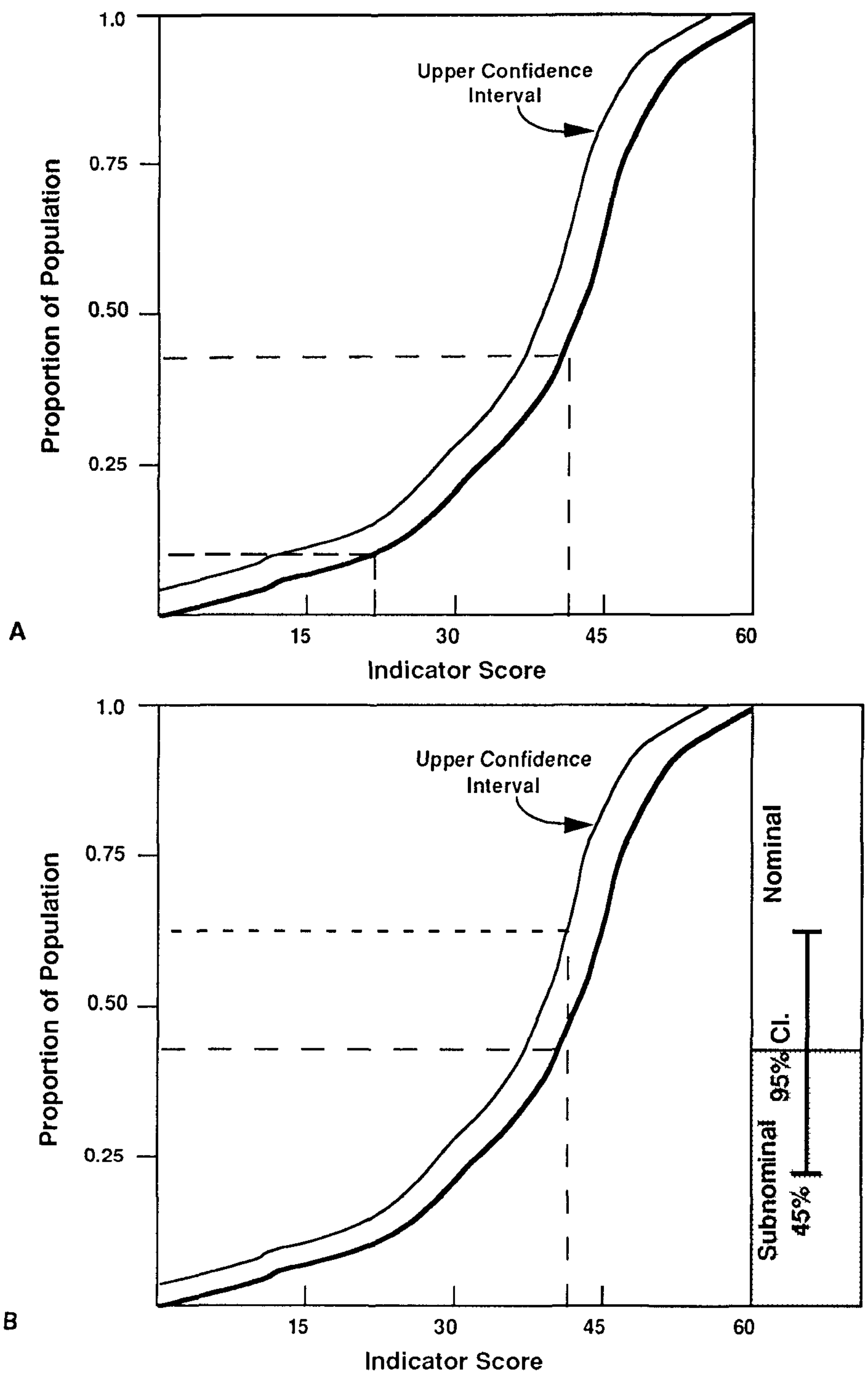


Figure 5-2. Cumulative distribution functions (CDFs) will be the primary format for reporting the condition of ecological resources. CDFs can be interpreted by reading the proportion of the resource (number of lakes, kilometers of streams) below any selected score as in A. If a particular score is chosen to delimit poor condition, the proportion (with confidence limits) of the resource in poor condition can be displayed as in B.

interest and evaluate the population from that perspective. In some instances, the CDF summaries will be accompanied by a histogram that reduces the information into some relatively "value free" classes, such as various trophic designations for lakes. After criteria are developed to classify lakes with respect to their acceptable/unacceptable condition, these criteria will be used to classify lakes into nominal and subnominal classes.

5.5.2 Trophic Condition of Lakes

Trophic condition in lakes has been one of the dominant concerns in surface water monitoring and the subject of extensive research for the last 25 years. Concerns range from the more aesthetic, which generally center around water clarity or nuisance floating algal mats, to the more biological, for example, excessive algal production that leads to oxygen depletion and subsequent fish kills during the summer or winter. Although lakes undergo a natural aging process that leads to increasingly eutrophic conditions, it is clear that in many instances this process has been accelerated by anthropogenic disturbances such as point and nonpoint discharges of organic and inorganic nutrients. A variety of indices have been developed, for example, Carlson's Trophic State Index (TSI) (Carlson, 1977), based on a combination of nutrient concentrations such as total phosphorus and water transparency taken from Secchi disk readings, and algal biomass as estimated by concentrations of algal chlorophyll-*a*. Until we have final concurrence on a trophic state index appropriate for all or most regions of the country, we will present the fundamental information that goes into indices of this type. Figure 5-3 shows the data for chlorophyll-*a* and total phosphorus for EPA Regions I and II. The data are presented as cumulative proportion of lake number [F(x)] and lake area [G(x)]. In this figure, the quartiles for the data are also presented. For example, the 25th percentile for chlorophyll-*a* is 2.41 μg/L. This means that 25% of the lakes in Regions I and II have a chlorophyll-*a* concentration of 2.4 μg/L or less. Figures 5-4, 5-5, and 5-6 present similar data, but broken down by the three regions outlined in Figure 5-1, the Adirondacks, New England Uplands, and the Coastal/ Lowland/Plateau.

Using criteria developed by NALMS (1988), we can summarize these data and estimate the proportion of lakes in various trophic categories (Figure 5-7). Using the NALMS criteria for oligotrophic (< 10 μg/L TP) and eutrophic (> 30 μg/L TP), we estimate that 21% of the lakes between 1 and 2,000 ha in Regions I and II are eutrophic. Based on one year's sample of 74 lakes, this estimate may be as high as 33% or as low as 9%. The regional differences (Figure 5-7) are quite distinct, although the confidence intervals are much larger given the smaller sample sizes. Figure 5-7 also presents the results as number of lakes rather than proportion of lakes, allowing the reader to see the actual number of lakes affected.

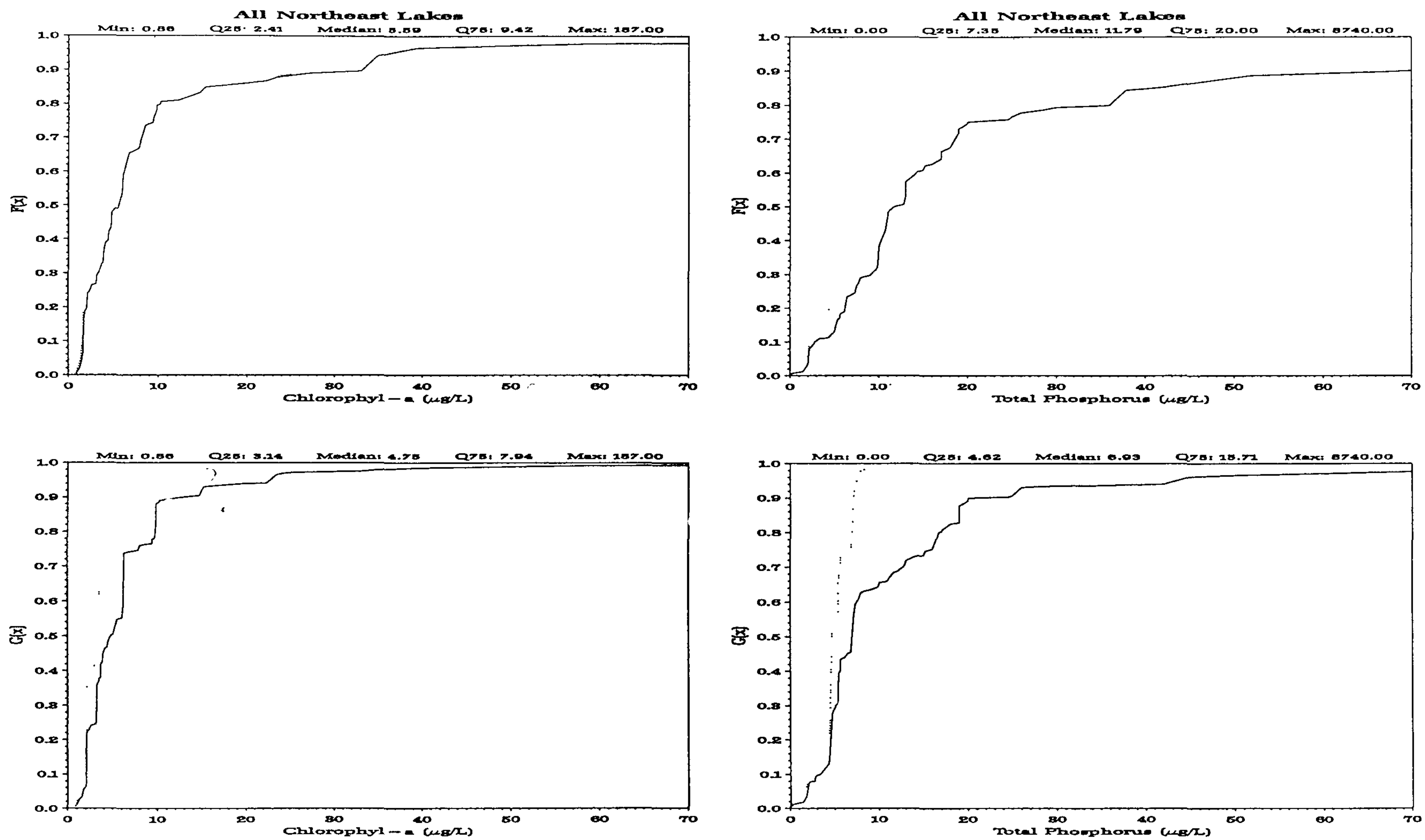


Figure 5-3. Estimated chlorophyll-*a* and total phosphorus cumulative distribution functions (CDF) for all lakes between 1 and 2,000 ha in the Northeast. Solid lines are the estimated CDF and dotted lines are the upper 95% confidence interval for the CDF. Top panels [F(x)] are CDFs for proportion of lake number having specific indicator scores. Bottom panels [G(x)] are CDFs for proportion of lake area with specific scores.

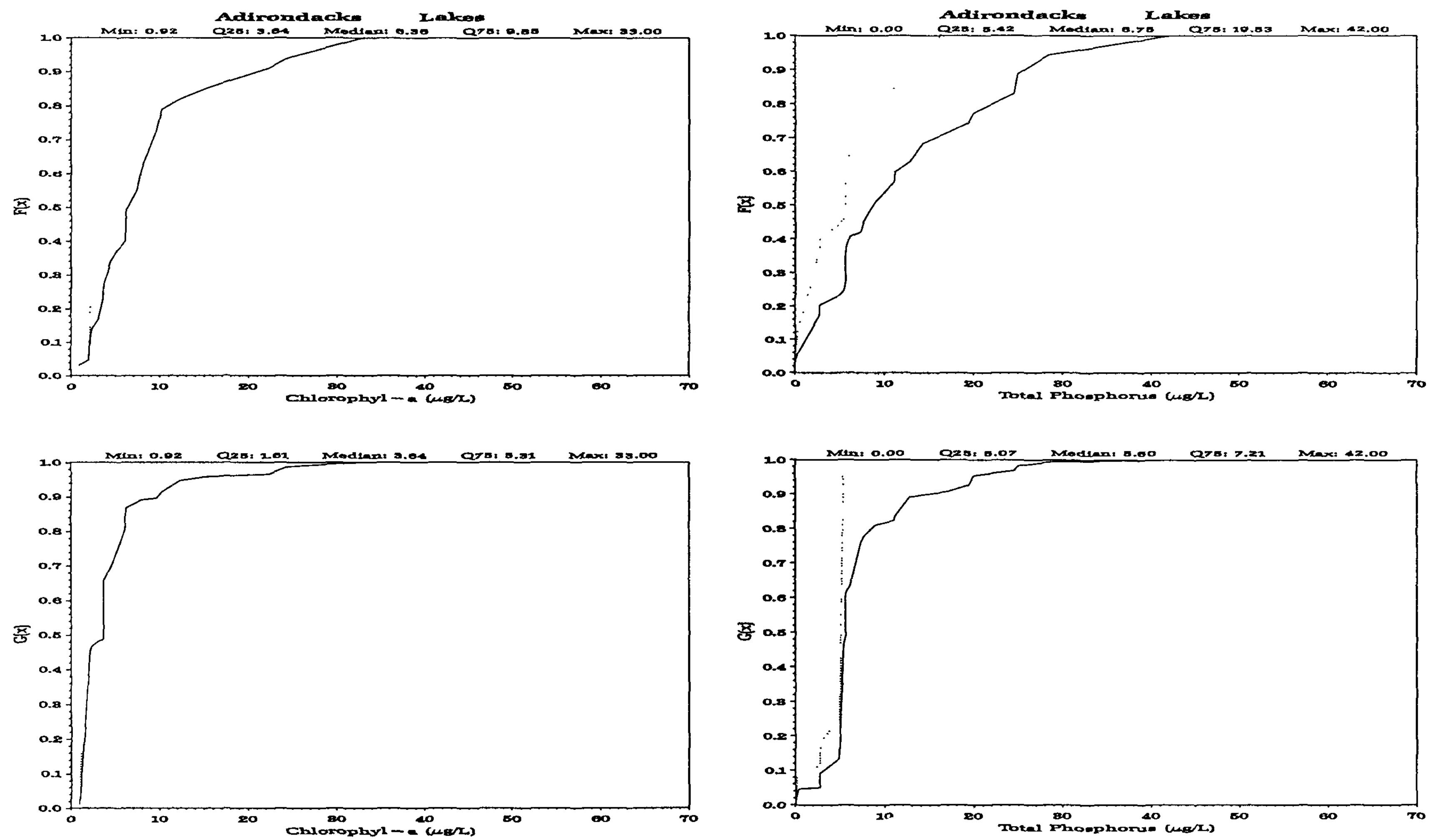


Figure 5-4. Estimated chlorophyll-a and total phosphorus cumulative distribution functions (CDF) for lakes between 1 and 2,000 ha in the Adirondacks (see Figure 5-1). Solid lines are the estimated CDF and dotted lines are the upper 95% confidence interval for the CDF. Top panels [F(x)] are CDFs for proportion of lake number having specific indicator scores. Bottom panels [G(x)] are CDFs for proportion of lake area with specific scores.

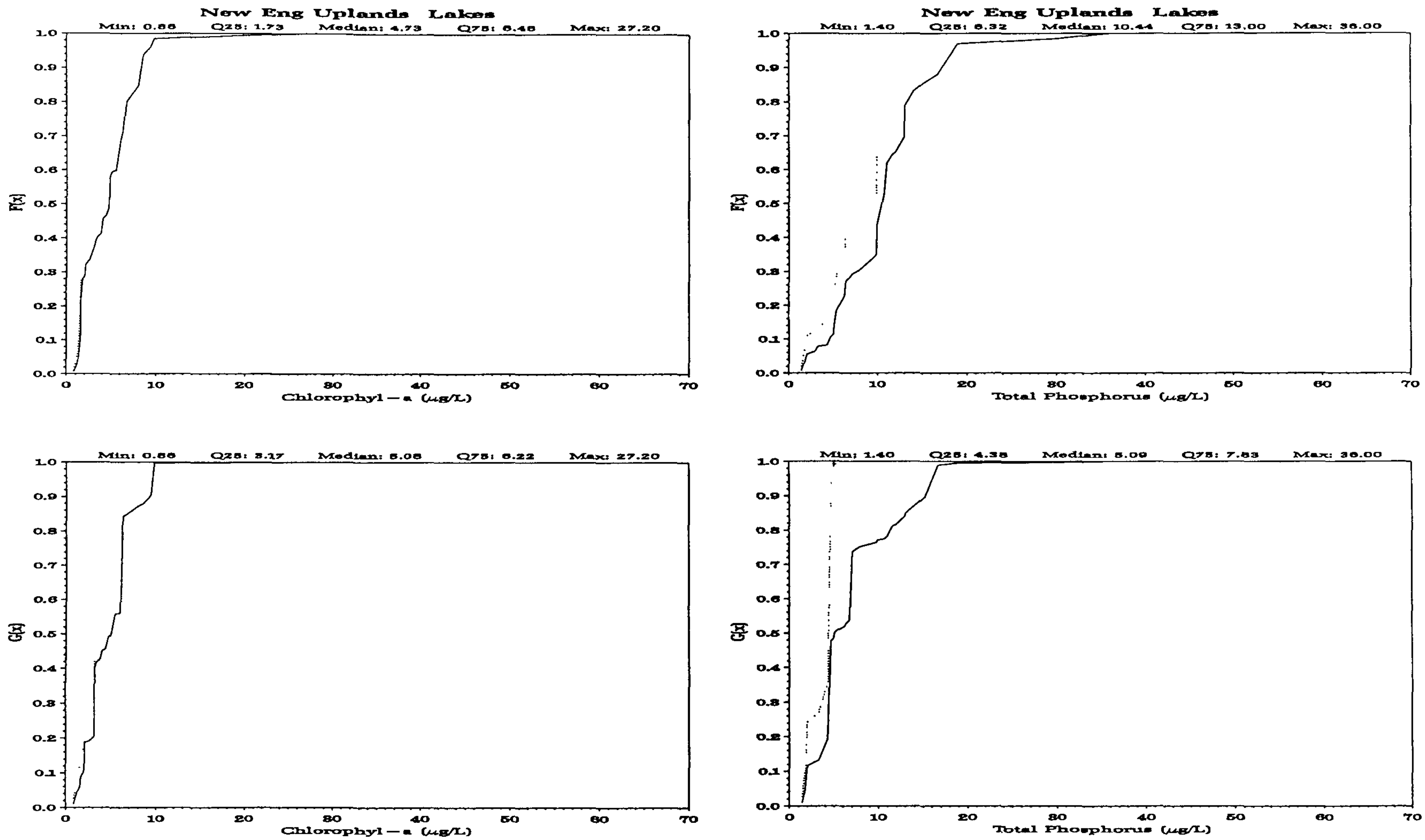


Figure 5-5. Estimated chlorophyll-*a* and total phosphorus cumulative distribution functions (CDF) for lakes between 1 and 2,000 ha in the New England Uplands (see Figure 5-1). Solid lines are the estimated CDF and dotted lines are the upper 95% confidence interval for the CDF. Top panels [F(x)] are CDFs for proportion of lake number having specific indicator scores. Bottom panels [G(x)] are CDFs for proportion of lake area with specific scores.

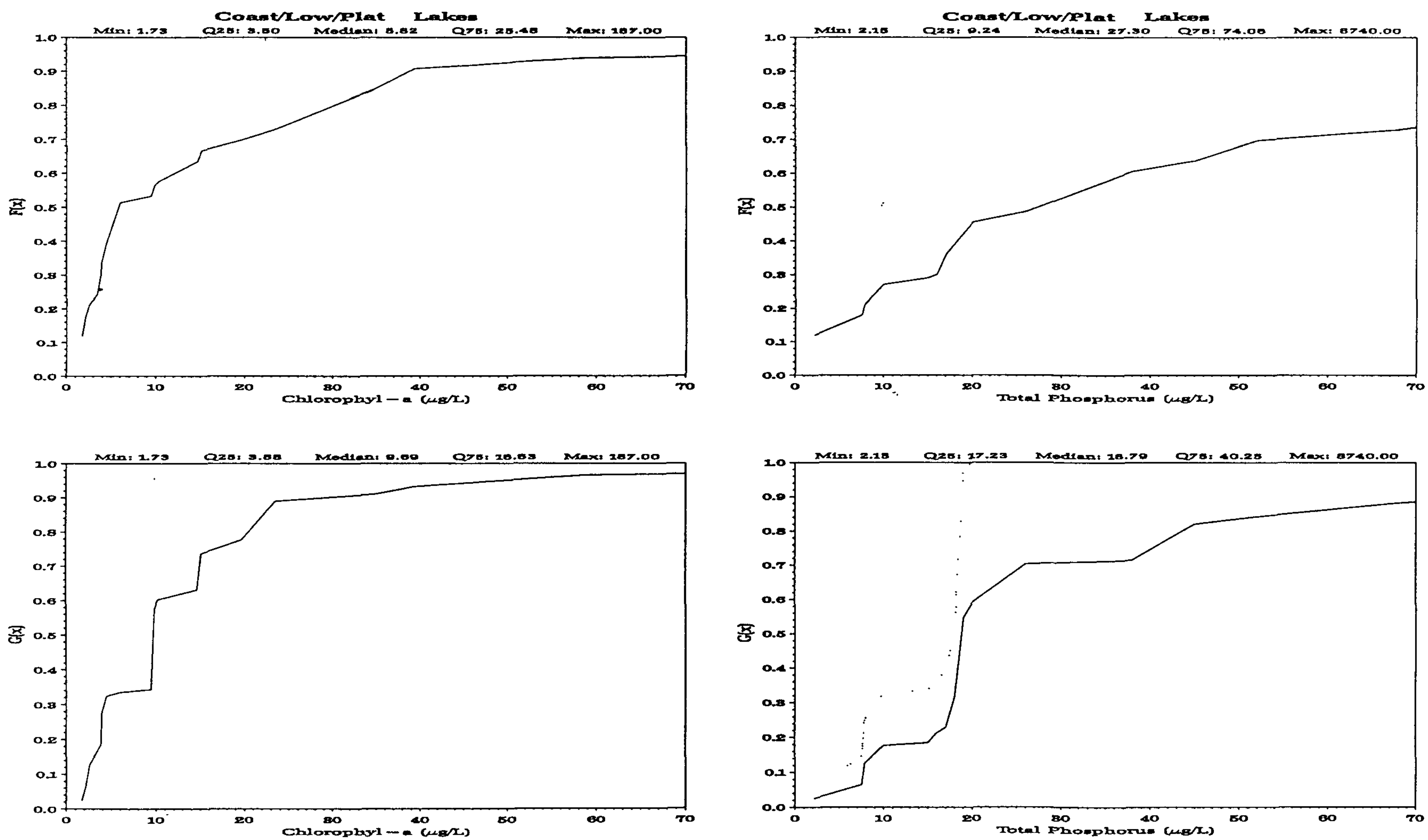


Figure 5-6. Estimated chlorophyll-*a* and total phosphorus cumulative distribution functions (CDF) for lakes between 1 and 2,000 ha in the Coastal/Lowlands/Plateaus (see Figure 5-1). Solid lines are the estimated CDF and dotted lines are the upper 95% confidence interval for the CDF. Top panels [F(x)] are CDFs for proportion of lake number having specific indicator scores. Bottom panels [G(x)] are CDFs for proportion of lake area with specific scores.

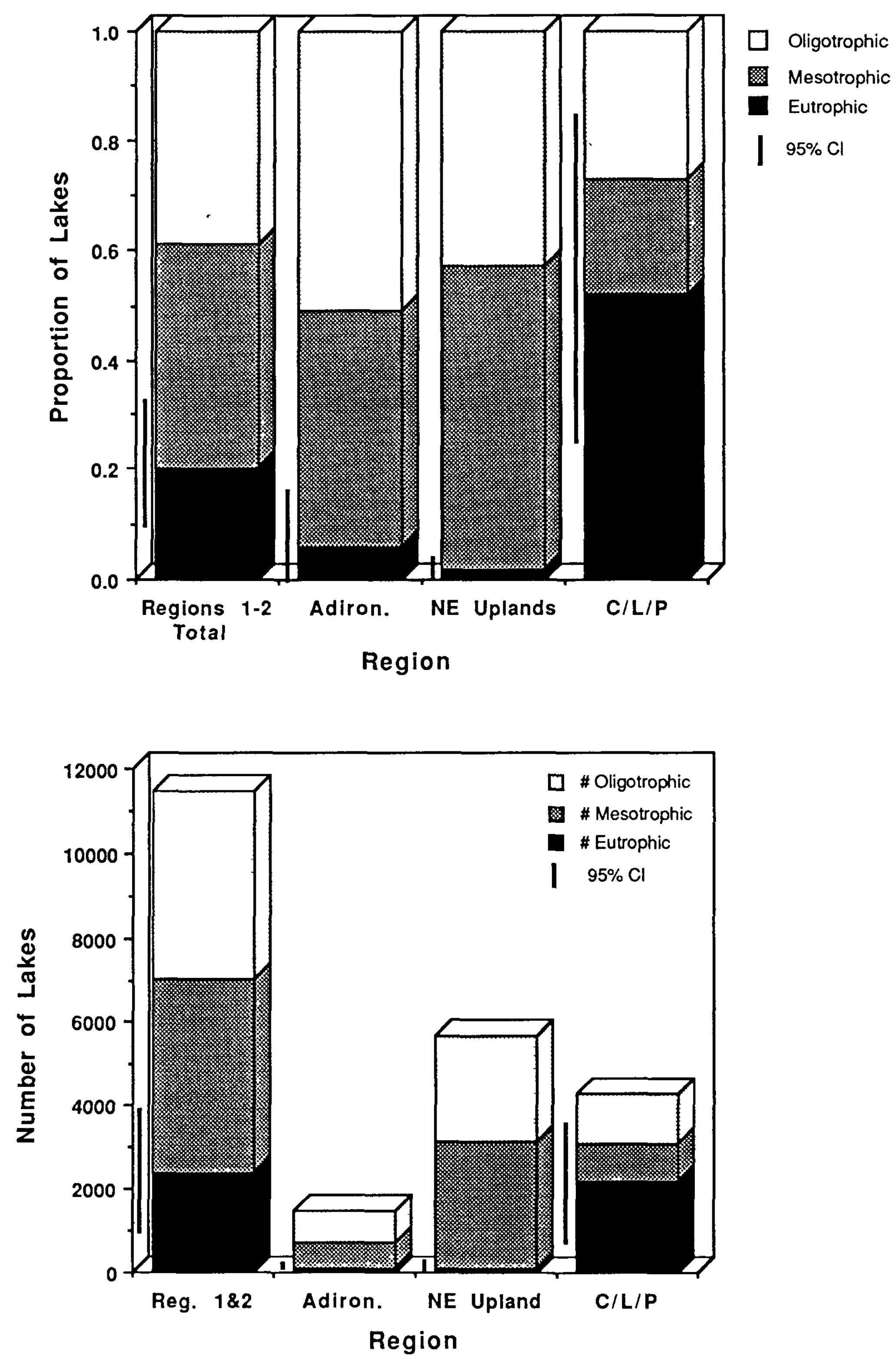


Figure 5-7. **Histogram of estimated proportion and number of eutrophic, mesotrophic, and oligotrophic lakes between 1 and 2,000 ha in the Northeast. Black bars represent the 95% confidence interval for the estimates.**

5.6 CONCLUSIONS

The primary objective of the Annual Statistical Summary is to report significant results annually. Given this objective, the data presented in the Annual Summaries will focus on the EMAP response indicators that can be currently interpreted. In 1991, trophic state measurements and zooplankton were the only response indicators collected at all probability lakes. The zooplankton indicator will be undergoing development over the next several years and thus is not ready for effective summarization in the annual statistical summaries. During the 1992 pilot, regional data on fish and riparian birds, as well as zooplankton and diatoms are being collected. We anticipate presenting the results of these additional indicators in later years.

In addition, much debate has occurred on the subject of presenting the exposure, habitat, and stressor indicator summaries in the annual statistical summaries. Given the focus of EMAP on its first objective, that is, describing status and trends in conditions, and the role of the exposure, habitat, and stressor indicators in diagnostics, we believe it is premature to commit to extensive inclusion of these additional indicators in the annual summaries. Thus, the extensive data on water chemistry and physical habitat will not be presented in the annual summaries, but will be used in the interpretive reports to explain patterns seen in response indicators. However, thoughtful comments and perspectives on this subject from users would be appreciated.

CHAPTER 6
LOGISTICAL OPERATIONS SUMMARY

6.1 INTRODUCTION

The complexity of EMAP leads to several major logistical issues, such as personnel identification, procurement of equipment, and site access. One purpose of the 1991 Northeast Lakes Pilot Survey was to assess the difficulty of assembling, training, and deploying sampling teams and required equipment to obtain lake measurements and samples within the necessary time period. We also needed to know if it is possible to coordinate the work of multiple field crews well enough for smooth, consistent sample collection and field recording. Can samples and field data collection forms be shipped and tracked to appropriate analytical facilities or data processors in a timely manner and without loss of samples or data? The logistical planning, implementation, and post-season reporting activities for the 1991 pilot focused on providing answers to these questions for specific indicators and for the total field efforts.

The 1991 pilot survey had two major components (Table 6-1). The probability lakes component focused on logistical and variance questions related to sampling lakes at a regional scale. This component had two subcomponents—one to test whether or not selected indicators can be effectively sampled in an index mode and the second, in response to the mandates of the Clean Air Act, to incorporate the Temporally Integrated Monitoring of Ecosystems (TIME) Program.

EMAP-Surface Waters (EMAP-SW) staff selected 64 regional probability lakes from the grid design to evaluate EMAP's ability to efficiently collect samples within the index period for indicators for which there is consensus about methodology and usefulness. Field plans included 1 visit to each of these lakes, a revisit to 32 of the lakes to collect data for estimating index period variability, and 1 visit to each of the additional 28 lakes selected for TIME. Each of the 3 field crews visited 4 of the original 64 lakes to provide data for crew comparability studies. Of these 12 visits, 4 were counted as initial visits. Thus, the 64 initial visits, 32 repeat visits, 28 additional TIME lakes, and 8 crew comparability visits resulted in 132 possible lake site visits. Table 6-1 lists the kinds of samples taken during each of these visits.

The indicator lakes component (Table 6-1) addressed the program's ability to obtain a cost-effective index sample for fish, benthic macroinvertebrates, and riparian birds, and a more intensive evaluation of physical habitat. This component of the pilot activity focused primarily on obtaining enough information to select adequate sampling protocols to be used in later surveys, identify

Table 6-1. Components of the Pilot Activities Planned for EMAP-Surface Waters in 1991

<table>
<tr><th colspan="5">EMAP SURFACE WATERS
1991 NORTHEAST LAKES PILOT SURVEY</th></tr>
<tr><th colspan="4">PROBABILITY LAKES
(132 visits[a])</th><th>INDICATOR LAKES</th></tr>
<tr><td colspan="3">Logistics and Variability Study and TIME Regional Lake Study
(64 lakes)</td><td>Additional TIME lakes (28 lakes)</td><td rowspan="2">(20 lakes)</td></tr>
<tr><td>Visit 1
(64 visits)</td><td>Visit 2
(32 visits)</td><td>Crew Comparability
(8 visits)</td><td>(28 visits)</td></tr>
<tr><td><u>Parameters</u>

Water chemistry
Trophic state
Zooplankton
Profundal benthos
Sediment diatoms
Sediment toxicity
Physical habitat</td><td><u>Parameters</u>

Water chemistry
Trophic state
Zooplankton
Profundal benthos</td><td><u>Parameters</u>

Water chemistry
Zooplankton
Profundal benthos</td><td><u>Parameters</u>

Water chemistry
Trophic state</td><td><u>Parameters</u>

Water chemistry
Trophic state
Zooplankton
Sediment diatoms
Sediment toxicity
Profundal and littoral benthos
Physical habitat (intensive)
Fish assemblage
Fish tissue contaminants
Fish biomarkers
Riparian bird assemblages
Benthic macroinvertebrates</td></tr>
</table>

[a] The 132 possible probability lake visits include the 64 initial visits and 32 revisits to the selected Tier 2 grid lakes, visits to the 28 additional lakes drawn from the intensified Tier 2 grid for TIME, and 12 visits to 4 lakes by each of the 3 crews for crew comparability evaluation (4 of these counted as initial visits). The TIME database includes water chemistry and trophic state data collected during the first visits to the original 64 lakes and the visits to the additional 28 lakes.

major habitat types, and estimate and evaluate the logistical and time requirements for implementing the complete suite of indicators in the field. Twenty indicator lakes were handpicked from outside the EMAP grid design to ensure that the variety of lake sizes and impacts expected to occur during annual surveys could be evaluated for sampling feasibility and variations in sampling protocols.

Three different types of field crews collected data for the indicator lakes component of the survey. One fisheries crew collected data and samples for indicators for fish, water chemistry, trophic state, zooplankton, sediment diatoms, sediment toxicity, and physical habitat. A second team conducted the riparian bird survey to address specific concerns inherent in bird survey studies. The third team collected littoral and profundal samples to assess benthic sampling protocols. Each of these crews traveled and worked independently of the others, and their logistical planning, implementation, and reporting activities will be discussed in other reports; some aspects are described in Chapter 2.

6.2 PLANNING ACTIVITIES

A logistics team was established early in the planning stages of the study to ensure that all field activities were conducted efficiently and that the data collected were of high quality. The logistics team was concerned primarily with the limnological field activities for the probability lakes. In some cases, simultaneous planning and preparation activities included the fisheries and benthological field operations. Members of the logistics team developed plans related to quality assurance (Peck, 1991), information management, and a field operations and training manual. Specific planning steps and activities are described in the following subsections.

6.2.1 Lake Verification

In October 1990, the design team provided the logistics team with a list of the 311 lakes in the Tier 1 frame for the northeast United States. In order to assess the quality of the selection process initiated by the design team, the logistics team manually plotted the point coordinates on 7.5′ or 15′ USGS topographic maps. This lake verification process revealed several potential selection errors. For example, lakes and rivers with transecting bridges may appear in the base frame as multiple lakes. The digitization process that electronically stores the map information cannot differentiate between a natural barrier and a human-made structure, such as a bridge that does not restrict cross flow of a waterbody. Other errors included identifying water bodies as lakes when actually they were ephemeral pools found in flood plains, marshes, and coastal bays, or pools in rivers generated by beaver dams or human-made dams later destroyed. Once this plotting effort was completed, the logistics team was able to evaluate the sample points selected in terms of whether or not a site met the definition for an EMAP lake. After examining the maps and telephoning the communities in which the lakes were located when necessary, the logistics team stored the resulting information in a database that was transmitted to the design team.

From the original 311 lakes drawn from the lake frame, the logistics team identified 48 sites that did not meet the criteria for EMAP lakes (a body of standing water < 1 ha in surface area, > 1 m in depth, and containing 1,000 m^2 of continuous open water) and 51 sites lacking sufficient data. Using this information, the design team removed the 48 nontarget sites from the list of 311 and then selected 64 probability lakes for the pilot study. Another 28 lakes were selected from an enhanced grid for the Temporally Integrated Monitoring of Ecosystems (TIME) study (for a total of 92 TIME lakes).

Time and weather restricted the reconnaissance conducted to investigate apparent access problems and to determine if these sites met the EMAP definition of lakes. Attempts to obtain reconnaissance information by telephone had limited success and information obtained from some local sources was not always accurate. At least one crew hiked to a site, only to discover that a road was present that provided much easier access.

In order to investigate apparent access problems, the limnological field coordinator traveled to Maine in May to participate in flyovers of about eight lakes that appeared to present access problems. The flyover effort was successful in verifying the presence and size of the lakes and clarifying some access questions. The flyover observations indicated that five of the lakes did not meet EMAP criteria and summer field crew visits confirmed these conclusions.

6.2.2 Site Access

Local government officials were contacted by telephone to identify landowners and to collect as much information as possible about a lake site. In addition, representatives of several state government agencies furnished lake information. Most of the sites selected for sampling were privately owned and thus the number of sites on public lands was limited. There were no difficulties in obtaining permission to access public lands during the 1991 survey. In some cases, especially in Maine, it was not necessary to acquire site access permission because of great pond laws which specify, for example, that ponds containing more than 10 acres lie in common for public use.

Before the logistics team contacted private landowners by telephone, EPA Regions I and II sent a letter, approved by EPA headquarters, requesting permission to access the property. It took a significant amount of time to obtain access permission and accurate details about the status of a site by telephone or mail to state and local contacts and landowners. In addition, the information obtained was often derived from subjective evaluation of a site. At times, the state or local contact would identify the wrong lake or consider the selected site to be a marsh or wetland, although the site fell within the EMAP definition of a lake. In one case, a landowner reported in a telephone conversation that his lake was greater than a hectare in surface area and greater than a meter in depth. In fact, when the site was visited for sampling, no lake was present, only a stream running through a field. Because some of the USGS topographic maps used to determine locations had not been updated recently, they sometimes misrepresented roads or did not show more recently developed roads. In some cases, access permission letters were not received before the crews reached the field, but permission letters were always in hand before the field team visited the site.

6.2.3 Protocol Development, Training, and Mobilization

The EMAP-SW team adopted sampling protocols for limnological, fishery, and benthological sampling that were originally selected by individual indicator leads. The indicator leads forwarded the methods to the logistics team for inclusion in a field operations and methods manual. Some of the methods were fully developed; others were in the formative stage, with the pilot survey serving as the means to field test the proposed methodology.

Personnel at the U.S. EPA laboratory in Las Vegas conducted protocol development and training sessions, called dry runs, at Lake Mead, Nevada, and Pahranagat National Wildlife Refuge in Alamo, Nevada. These sessions enabled individuals to work in small groups, identify problems and resolutions, and concentrate on issues pertinent to the pilot study.

Training for limnological, benthological, and fisheries sampling was held at the Pahranagat National Wildlife Refuge in Nevada from June 17 to 24. Trainers had extensive experience in training scientific personnel in collection techniques and field operations. During this week-long training period, groups of two field samplers worked in hands-on sessions with trainers who introduced and demonstrated all aspects of the field activities. Indicator leads participated in the training. The sessions also covered activities such as shipping, sample tracking, and equipment maintenance. Each sampler had to demonstrate competence in the practical aspects of field activities, such as handling boats, operating global positioning systems (GPS) and dissolved oxygen (DO) meters, locating sampling sites, completing data forms, understanding and demonstrating safety procedures, and performing general administrative tasks.

Additional practice runs were conducted on July 8 in Nashua, New Hampshire. At the conclusion of training, crews began field work by participating in a crew comparability study. Each of the three crews sampled two lakes at different times over a three-day period.

Some indicator leads visited the field teams early in the sampling period. These visits were extremely useful, providing opportunities for the leads to work with the field crews to improve their understanding of the need for particular protocol elements, solve problems with implementing the protocols, and correct minor errors in the protocols.

Limnological, benthological, and fisheries crew members moved to New Hampshire from Las Vegas with their equipment, leaving June 29 with two 14-foot rental trucks. Four-wheel drive vehicles were leased in the northeast for the field work. The crews assembled equipment and began full implementation on July 10. The fisheries and benthological teams operated independently of the limnological team. The crews used equipment that had been obtained for previous surveys and stored in Las Vegas. Although this equipment served the purpose during the 1991 survey, by the end of the field work it was obvious that most of it would need to be replaced before the next field season.

6.2.4 Field Crew Personnel

Field crew members for the limnological, fisheries, and benthological crews were drawn from U.S. EPA laboratories and regions and from contract personnel. Scientists from the U.S. EPA Environmental Monitoring Systems Laboratories in Las Vegas and Cincinnati served as field samplers for the full field season. Scientists from U.S. EPA Regions I and II also served as limnological crew members for the entire season. Participation by Region I included providing GPS units, as well as training the crew members in the use of these units. For the entire field season period, at least one U.S. EPA regional and one laboratory participant was in the field. Thirteen individuals filled the six limnological crew positions (three crews, two people each) at different times. The length of time a sampler stayed in the field varied from one to nine weeks. One person served as field coordinator for the eight-week sample collection effort.

Early efforts were directed at getting the states involved in the planning and field activities. Representatives of each of the New England states and New York and New Jersey participated in the 1991 survey at different levels. A majority of the states provided ownership and access information. New Hampshire, Vermont, and New York personnel accompanied field crews to learn more about EMAP and about the EMAP-SW effort. In general, though, the widespread depressed economic circumstances prevented the states from assuming a larger role in field activities.

6.3 FIELD OPERATIONS OVERVIEW

Three field crews of two people each and a base coordinator conducted field operations between July 8 and August 28, 1991. The crews used 4-wheel drive vehicles pulling trailers carrying 12-foot boats to travel between base sites and sampling sites. It took approximately 2 to 4 hours to sample the EMAP probability lakes and the TIME lakes. To reach 21 lakes located in remote areas, crews had to hike in (15 sites; 6 were not EMAP target sites) or use a float plane (6 sites). The crews worked out of 14 centralized base sites. The following subsections describe daily sampling and support operations.

6.3.1 Daily Sampling Operations: Probability and TIME Lakes

The field crews were scheduled to collect measurements [watershed characteristics, site location (GPS), site depth, temperature profile, dissolved oxygen profile, Secchi disk transparency] and samples (water chemistry, chlorophyll-*a*, zooplankton tow, sediment core, sediment toxicity, profundal benthos) during 132 visits to probability and TIME lakes (Table 6-1). Specific water and sediment sampling activities carried out at the lakes are shown in Figures 6-1 and 6-2.

The three crews began activities each morning by calibrating instruments and using equipment checklists to ensure that all necessary equipment and supplies were loaded into the vehicles before departing for the sample site. The field coordinator remained at the motel base site to organize logistics support (e.g., picking up equipment and supplies, shipping and tracking samples, transmitting data, contacting landowners) and maintain a communications link with program management. The field coordinator also served as a backup sampler. Using a central base site sometimes required long drives by the field crews to the lake site and then resulted in long work days.

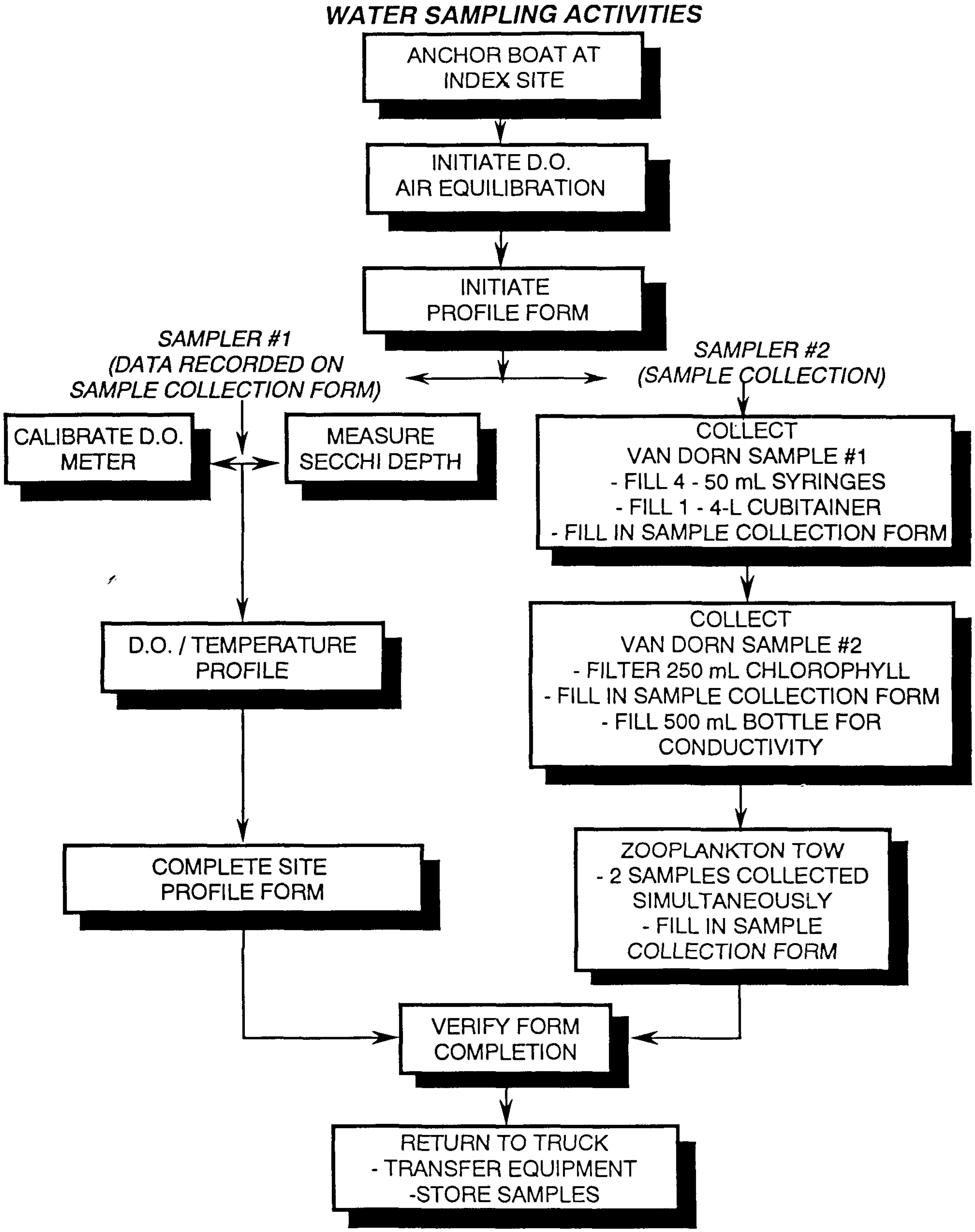


Figure 6-1. Daily water sampling activities at 1991 EMAP Northeast Lakes Pilot sites.

SEDIMENT SAMPLING ACTIVITIES

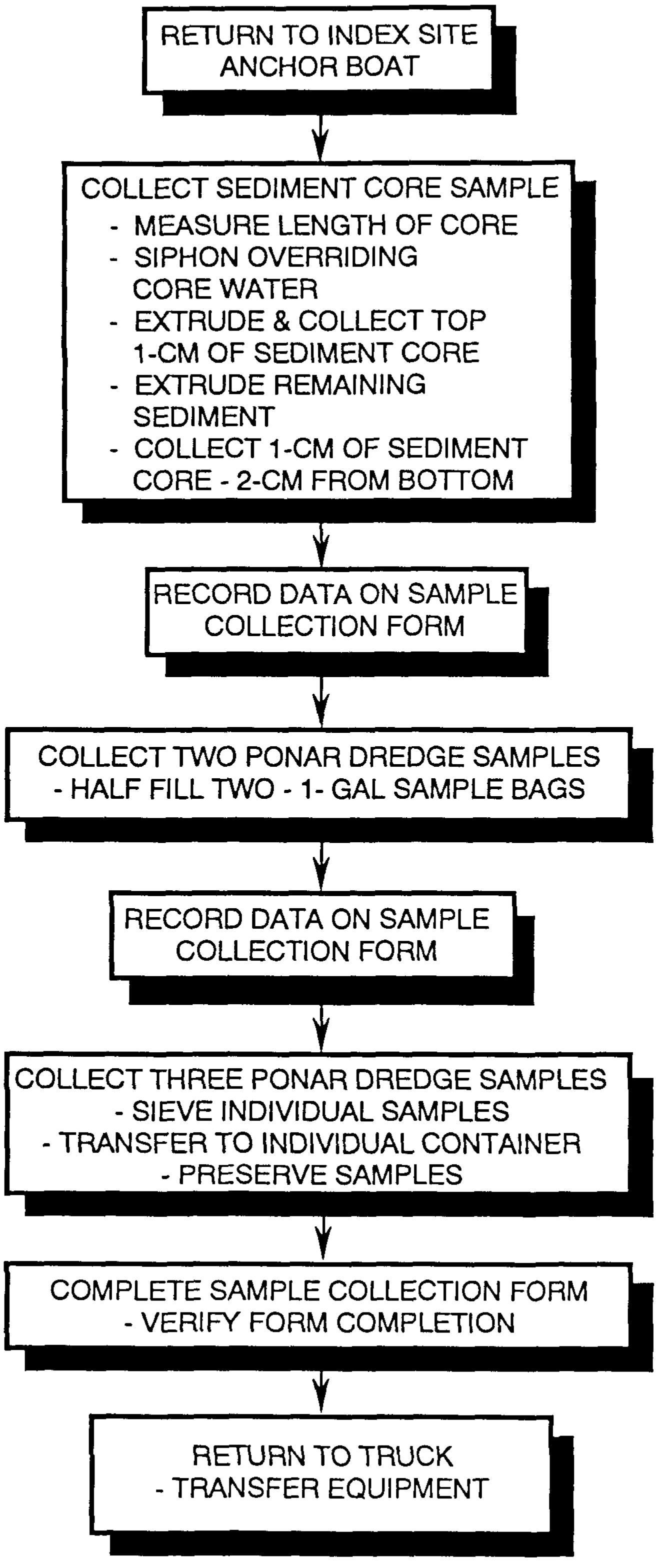


Figure 6-2. Daily sediment sampling activities at 1991 EMAP Northeast Lakes Pilot sites.

Upon arrival at the site, the field crew verified the location based on landscape features, topographic maps, and GPS information (if coverage was available at that time). After loading the equipment and launching the boat, the crew used sonar to determine the index site for sample collection at the deepest part of the lake. All samples were preserved (when appropriate), labeled, and packed for return to the base site, and the boat was returned to shore. The boat was then re-outfitted for collection of sediment samples. All samples and forms were checked for completeness against a check list and then packed for return to the base site.

The GPS units proved to be successful locational devices when connections to the satellite system were available. This system will be expanded in the future and the GPS units should become more useful.

At the base site, the field coordinator conducted a debriefing with the sample crews and checked the data forms, sampling labels, and condition of the samples. Selected data forms were transmitted by FAX to the Las Vegas communications center for data entry after they were reviewed by the field coordinator. The sampling crews and the field coordinator then cleaned and prepared equipment and supplies for the next day.

Water chemistry samples were shipped to the analytical laboratory on the morning after sampling, whereas sediment toxicity and chlorophyll-*a* samples were held for shipment for up to one week. The preserved benthic invertebrate, zooplankton, and sediment diatom samples were stored for biweekly shipment.

Measurements and samples were collected on 105 of the scheduled 132 site visits. The 27 remaining sites were not sampled for the following reasons:

- 1 access was denied when crew reached site
- 1 access was restricted (a locked gate)
- 1 site was misidentified (map typographic error; a lake has never existed at that site)
- 1 site visit was not accomplished during crew comparability lake visits
- 14 sites did not meet the definition of a lake (e.g., too shallow, wetland)
- 9 site revisits were originally scheduled to the 14 non-EMAP lakes

In total, 1,321 samples were collected and shipped to analytical support sites. Sample collection and shipment of the samples to the laboratory was 100% complete.

6.3.2 Support Operations

The communications center, located in Las Vegas, Nevada, was monitored by two people working in shifts. Answering and facsimile machines and a toll-free telephone number were available for communications throughout the field season. The field coordinator also had a facsimile machine to facilitate communications. The communications center monitored all aspects of field sampling, including coordinating and tracking support requests for supplies and maintaining a central contact point for information exchange among the limnological, benthological, and fisheries field crews; the management team; the information management team; the analytical laboratory; and the public. An information pamphlet was also available for distribution in the field to interested parties. The field crews did not have communication equipment and, when necessary, had to locate public telephones or return to the base site.

Electronic entry of field data was not implemented during the 1991 pilot survey and all field data were manually entered on paper forms. There were budget and time constraints and most data were derived from sample analyses. A portable data recorder was evaluated at the end of the field season. Several disadvantages were noted, including difficulty in entering comments and the need to download the unit on a daily basis. Further investigations into other kinds of recording units were recommended for future field seasons.

Development of the field forms was a combined effort involving the indicator leads, information management, logistics, and quality assurance team members. Thoroughly reviewed before the field season, these forms provided the means to record accurate data and site information. The forms were checked by the field coordinator before they were transmitted to the communications center; they were checked again at the communications center. Information on the data forms was entered into the database in Las Vegas.

The sample tracking system was enhanced by the use of barcoded labels. These labels, applied to all samples at the base site, enabled the sample identification to be electronically entered into the sample tracking system. The site identification, data, crew identification, and type of sample were stored in a database that could provide sample-specific information on sample shipping and tracking records. Data transfer was restricted to overnight shipment of diskettes and field forms. Use of a modem to transmit data electronically was curtailed by inadequate telephone systems in remote regions of the Northeast.

6.4 RESULTS AND RECOMMENDATIONS

In general, the 1991 Northeast Lakes Pilot Survey was successful in assembling, training, and deploying sampling teams and equipment to obtain lake measurements and samples within the required index period. The work of multiple field crews was coordinated well enough to result in efficient and consistent sample collection and field records. Field visits determined that 24 of the 132 planned site visits were to nontarget sites, leaving a possible 108 sites from which EMAP data could be collected. Of these, field crews completed operations during 105 visits for a completion rate of over 97%. In addition samples and data collection forms were successfully shipped, tracked, and received at the appropriate destination without any loss.

Problems encountered in the planning and implementation process were solved in a way that did not prevent the successful completion of the field operations. Recommendations for future surveys include:

- Site Access
  - Develop coordination and cooperation with the regional offices for site verification and access activities.
  - Identify possible sites earlier in the season to provide more time to conduct verification and access investigations.
- Reconnaissance
  - Conduct additional reconnaissance before the field operations begin to reduce the number of field crew trips to non-EMAP sites.
  - Consider conducting reconnaissance for 1993 sites during the 1992 field season.
- Field Crew Personnel
  - Cooperate with other agencies in field efforts already in place to identify crew members.
- Training
  - Conduct training in the area where field operations will take place.
  - Develop short courses, to be completed before training, for requirements such as boat handling, map orienteering, and survival skills.
- Mobilization
  - Develop warehouse facilities in the regions where field operations will take place.
  - Use existing motor pools.

- Field Activities

  - Identify base coordinator on each field crew and do not establish central base sites.
  - Schedule field crew visits to sites within small geographical area.

- Field Quality Assurance

  - Estimate precision and accuracy of measurements.
  - Improve field quality control measurements.
  - Make minor modifications in forms and labels.
  - Conduct additional field audits.
  - Improve procedures for preventive maintenance and calibration.
  - Refine field logbooks.

- Communications

  - Provide field crews with field communications capability.

- Information Management

  - Evaluate grid pad portable data recorder.
  - Use bar codes for field samples.
  - Investigate using modem to track samples.

In summary, the logistical element of the 1991 Northeast Lakes Pilot Survey was able to plan, implement, and successfully complete field operations and to identify recommendations for future surveys.

CHAPTER 7

LITERATURE CITED

Ackerman, J. 1992. Tracking a river of birds. *Nature Conservancy* 42:23-27.

Allan, D. and 21 coauthors. 1992. Report of the NABS FWI (Freshwater Imperative) working group. *Bull. N. Am. Benthol. Soc.* 9:266-271.

American Public Health Association (APHA). 1989. *Standard Methods for the Examination of Water and Wastewater.* Seventeenth Edition. American Public Health Association, Washington, D.C.

Angermeier, P.L., and J.R. Karr. 1986. Applying an index of biotic integrity based on stream fish communities: considerations in sampling and interpretation. *N. Am. J. Fish. Manage.* 6:418-429.

Bailey, R.G. 1976. *Ecoregions of the United States.* U.S. Forest Service, Ogden, UT.

Barbour, C.D., and J.H. Brown. 1974. Fish species diversity in lakes. *Am. Nat.* 108:473-489.

Bickford, C.A., C.E. Mayer, and K.D. Ware. 1963. An efficient sampling design for forest inventory: the Northeast Forest Resurvey. *J. For.* 61:826-833.

Bovee, K.D. 1982. *A Guide to Stream Analysis Using the Instream Flow Incremental Methodology.* Instream Flow Information Paper 12. FWS/OBS 82/26. U.S. Fish & Wildlife Service, Fort Collins, CO.

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.

Brinkhurst, R.O. 1974. *The Benthos of Lakes.* MacMillan, London.

Brooks, J.L., and S.I. Dodson. 1965. Predation, body size, and composition of plankton. *Science* 150:28-35.

Brooks, R.P., M.J. Croonquist, E.T. Silva, J.E. Gallagher, and D.E. Arnold. 1991. Selection of biological indicators for integrating assessments of wetland, stream and riparian habitats. Pages 81-89 in *Proceedings Symposium on Biological Criteria.* U.S. Environmental Protection Agency, Washington, D.C.

Brussock, P.P., A.V. Brown, and J.C. Dixon. 1985. Channel form and stream ecosystem models. *Water Resour. Bull.* 21:859-866.

Callicott, J.B. 1991. Conservation ethics and fishery management. *Fisheries* 16:22-28.

Carlson, R.E. 1977. A trophic state index for lakes. *Limnol. Oceanogr.* 22:361-368.

Carson, R. 1962. *Silent Spring.* Fawcett, Greenwich, CN.

Chaloud, D.C., J.M. Nicholson, B.P. Baldigo, C.A. Hagley, and D.W. Sutton. 1989. *Handbook of Methods for Acid Deposition Studies: Field Methods for Surface Water Chemistry.* EPA/600/4-89/020. U.S. Environmental Protection Agency, Washington, D.C.

Chapman, P.M., and R.O. Brinkhurst. 1984. Lethal and sublethal tolerances of aquatic oligochaetes with reference to their use as a biotic index of pollution. *Hydrobiologia* 115:139-144.

Charles, D.F., and 12 coauthors. 1990. Paleoecological investigation of recent lake acidification in the Adirondack Mountains, NY. *J. Paleolimnol.* 3:195-241.

Charles, D.F., S.S. Dixit, B.F. Cumming, and J.P. Smol. 1991. Variability in diatom and chrysophyte assemblages and inferred pH: paleolimnological studies of Big Moose Lake, New York, USA. *J. Paleolimnol.* 5:267-284.

Cole, G.A. 1975. *Textbook of Limnology.* C.V. Mosby, Saint Louis, MO.

Cotter, J., and J. Nealon. 1987. *Area Frame Design for Agricultural Surveys.* Area Frame Section, Research and Applications Division, National Agricultural Statistics Service, U.S. Department of Agriculture.

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. U.S. Fish and Wildlife Service, Washington, D.C. 131 pp.

Croonquist, M.J., and R.P. Brooks. 1991. Use of avian and mammalian guilds as indicators of cumulative impacts in riparian-wetland areas. *Environ. Manage.* 15:701-714.

Cumming, B.F., J.P. Smol, J.C. Kingston, D.F. Charles, H.J.B. Birks, K.E. Camburn, S.S. Dixit, A.J. Uutala, and A.R. Selle. 1992. How much acidification has occurred in Adirondack lakes (New York, USA) since preindustrial times? *Can. J. Fish. Aquat. Sci.* 49:128-141.

Dixit, S.S., J.P. Smol, J.C. Kingston, and D.F. Charles. 1992. Diatoms: powerful indicators of environmental change. *Environ. Sci. Technol.* 26:23-33.

Doudoroff, P., and C.E. Warren. 1957. Biological indices of water pollution with special reference to fish populations. Pages 144-163 in *Biological Problems in Water Pollution.* U.S. Public Health Service, Cincinnati, OH.

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 rivers symposium. *Can. Spec. Publ. Fish. Aquat. Sci.* 106.

Ehrlich, P.R., A.H. Ehrlich, and J.P. Holdren. 1977. *Ecoscience: Population, Resources, Environment.* Freeman, San Francisco, CA.

Ellis, M.M. 1937. Detection and measurement of stream pollution. *Bull. Bureau Fish.* 48:365-437.

Eshleman, K.N. 1988. Predicting regional episodic acidification of surface waters using empirical models. *Water Resour. Res.* 24:1118-1126.

Fausch, K.D., J. Lyons, J.R. Karr, and P.L. Angermeier. 1990. Fish communities as indicators of environmental degradation. *American Fisheries Society Symposium* 8:123-144.

Forbes, S.A. 1925. The lake as a microcosm. *Illinois Natural History Survey Bulletin* 15:537-550.

Ford, J. 1989. The effects of chemical stress on aquatic species composition and community structure. Pages 99-144 in S.A. Levin, M.A. Harwell, J.R. Kelly, and K.D. Kimball, eds. *Ecotoxicology: Problems and Approaches.* Springer-Verlag, New York.

Frey, D.G. 1977. Biological integrity of water—a historical approach. Pages 127-140 in R.K. Ballentine and L.J. Guarria, eds. *The Integrity of Water.* U.S. Environmental Protection Agency, Washington, D.C.

Frissell, C.A., W.J. Liss, C.E. Warren, and M.D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. *Environ. Manage.* 10:199-214.

Gammon, J.R. 1991. *The Environment and Fish Communities of the Middle Wabash River.* Annual Report. Dept. of Biology, DePauw University, Greencastle, IN.

Gammon, James R., DePauw University, Greencastle, Indiana (personal communication).

Gilbertson, M. 1992. PCB and dioxin research and implications for fisheries research and resource management. *Can. J. Fish. Aquat. Sci.* 49:1078-1079.

Glew, J.R. 1989. A new trigger mechanism for sediment samplers. *J. Paleolimnol.* 2:241-243.

Goulding, M., M.L. Carvalho, and E.G. Ferreira. 1988. *Rio Negro: Rich Life in Poor Water.* SPB Academic Publ., The Hague, The Netherlands.

Gregory, S.V., F.J. Swanson, W.A. McKee, and K.W. Cummins. 1991. An ecosystem perspective of riparian zones. *BioScience* 41:540-551.

Gubala, C.P., C.T. Driscoll, R.M. Newton, and C.L. Schofield. 1991. Chemistry of a near-shore lake region during spring snowmelt. *Environ. Sci. Technol.* 25:2024-2030.

Hall, R.I., and J.P. Smol. 1992. A weighted-averaging regression and calibration model for inferring total phosphorus concentration from diatoms in British Columbia (Canada) lakes. *Freshw. Biol.* 27:417-434.

Halsell, N. 1990. Lake indicator workshop. U.S. Environmental Protection Agency, Las Vegas, Nevada.

Hart, David D., Academy of Natural Sciences, Philadelphia, Pennsylvania (personal communication).

Hazard, J.W., and B.E. Law. 1989. *Forest Survey Methods Used in the USDA Forest Service.* EPA/600/3-89/065. NTIS PB89-220-594/AS. U.S. EPA Environmental Research Laboratory, Corvallis, OR.

Herlihy, A.T., D.H. Landers, R.F. Cusimano, W.S. Overton, P.J. Wigington, A.K. Pollack, and T.E. Mitchell-Hall. 1990. *Temporal Variability in Lakewater Chemistry in the Northeastern United*

States: Results of Phase II of the Eastern Lake Survey.* EPA/600/3-91/012. U.S. Environmental Protection Agency, Corvallis, OR.

Hilborn, R. 1992. Can fisheries agencies learn from experience? *Fisheries* 17:6-14.

Hilsenhoff, W.L. 1987. An improved biotic index of organic stream pollution. *Great Lakes Entomol.* 20:31-39.

Hocutt, C.H., and J.R. Stauffer, Jr. (eds) 1980. *Biological Monitoring of Fish.* D.C. Heath, Lexington, MA.

Hughes, R.M., and J.R. Gammon. 1987. Longitudinal changes in fish assemblages and water quality in the Willamette River, Oregon. *Trans. Am. Fish. Soc.* 116:196-209.

Hughes, R.M., and R.F. Noss. 1992. Biological diversity and biological integrity: current concerns for lakes and streams. *Fisheries* 17:11-19.

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.

Hunsaker, C.T., and D.E. Carpenter. 1990.*Environmental Monitoring and Assessment Program Ecological Indicators.* EPA/600/3-90/060. U.S. Environmental Protection Agency, Research Triangle Park, NC.

Hunsaker, C., D. Carpenter, and J. Messer. 1990. Ecological indicators for regional monitoring. *Bull. Ecol. Soc. Am.* 71(2):165-172.

Hutchinson, G.E. 1961. The paradox of the plankton. *Am. Nat.* 95:137-146.

Hynes, H.B.N. 1975. The stream and its valley. *Verh. Internat. Verein. Limnol.* 19:1-15.

Karr, J.R. 1987. Biological monitoring and environmental assessment: a conceptual framework. *Environ. Manage.* 11:249-256.

Karr, J.R., and D.R. Dudley. 1981. Ecological perspective on water quality goals. *Environ. Manage.* 5:55-68.

Karr, J.R., K.D. Fausch, P.L. Angermeier, P.R. Yant, and I.J. Schlosser. 1986. *Assessing Biological Integrity in Running Waters: A Method and Its Rationale.* Illinois Natural History Survey Spec. Publ. No.5., Champaign, IL.

Karr, J.R., P.R. Yant, K.D. Fausch, and I.J. Schlosser. 1987. Spatial and temporal variability of the index of biotic integrity in three midwestern streams. *Trans. Am. Fish. Soc.* 116:1-11.

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.

Kaufmann, P.R., A.T. Herlihy, M.E. Mitch, and W.S. Overton. 1991. Chemical characteristics of streams in the Eastern United States: I. Synoptic survey design, acid-base status, and regional chemical patterns. *Water Resour. Res.* 27:611-621.

Keller, W., and N.D. Yan. 1991. Recovery of crustacean zooplankton species richness in Sudbury area lakes following water quality improvements. *Can. J. Fish. Aquat. Sci.* 48:1635-1644.

Klemm, D.J., and J.M. Lazorchak. 1992. *Environmental Monitoring and Assessment Program (EMAP) Field and Laboratory Methods Manual Used on Probability Freshwater Lakes.* U.S. Environmental Protection Agency, Cincinnati, OH.

Klemm, D.J., P.A. Lewis, F. Folk, and J.M. Lazorchak. 1990. *Macroinvertebrate Field and Laboratory Methods for Evaluating the Biological Integrity of Surface Waters.* EPA/600/4-90/030. U.S. Environmental Protection Agency, Cincinnati, OH.

Knight, J.G., M.B. Bain, and K.J. Scheidegger. 1991. *Ecological Characteristics of Fish Assemblages in Two Seasonally Inundated Palustrine Wetlands.* Project Report. U.S. Fish and Wildlife Service, Auburn University, Auburn, AL.

Kolz, A.L., and J.B. Reynolds. 1991. *Principles and Techniques of Electrofishing.* U.S. Fish and Wildlife Service, Kearneysville, WV.

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: Population Descriptions and Physico-Chemical Relationships.* EPA/600/3-86/054a. U.S. Environmental Protection Agency, Washington, D.C.

Landres, P.B., J. Verner, J.W. Thomas. 1988. Ecological uses of vertebrate indicator species: a critique. *Conserv. Biol.* 2:316-328.

Larsen, D.P., D.L. Stevens, A.R. Selle, and S.G. Paulsen. 1991. Environmental Monitoring and Assessment Program, EMAP-Surface Waters: A Northeast Lakes Pilot. *Lake Reserv. Manage.* 7(1):1-11.

Lewis, W.M. 1983. A revised classification of lakes based on mixing. *Can. J. Fish. Aquat. Sci.* 40:1779-1787.

Lien, L., G.G. Raddum, and A. Fjellheim. 1992. *Critical Loads of Acidity to Freshwater Fish and Invertebrates.* Report No. 0-89185. Norwegian Institute for Water Research, Oslo, Norway.

Likens, G.E., and F.H. Bormann. 1974. Linkages between terrestrial and aquatic ecosystems. *Bioscience* 24:447-456.

Lind, O.T. 1979. *Handbook of Common Methods in Limnology.* Second edition. C.V. Mosby Co., St. Louis. 199 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.

Lowe, R.L. 1974. *Environmental Requirements and Pollution Tolerance of Freshwater Diatoms.* EPA/670/4-74/005. U.S. Environmental Protection Agency, Cincinnati, OH.

MacArthur, R.H., and E.O. Wilson. 1967. *The Theory of Island Biogeography.* Princeton University Press, Princeton, NJ.

Machin, D. 1975. Note 385: On a design problem in growth studies. *Biometrics* 31:749-753.

Magnuson, J.J., B.J. Benson, and T.K. Kratz. 1990. Temporal coherence in the limnology of a suite of lakes in Wisconsin, U.S.A. *Freshwater Biol.* 23:145-159.

Margalef, R. 1963. On certain unifying principles in ecology. *Am. Nat.* 97:357-374.

Marsh, G.P. 1864. *Man and Nature.* Harvard University Press, Cambridge, Massachusetts.

McCauley, E. 1984. The estimation of the abundance and biomass of zooplankton in samples. Pages 228-265 in J.A. Downing and F.H. Rigler, eds. *A Manual on Methods for Secondary Productivity in Fresh Waters.* Blackwell, Oxford, U.K.

Messer, J.J. 1990. EMAP indicator concepts. Pages 2.1–2.26 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.

Messer, J.J., R.A. Linthurst, and W.S. Overton. 1991. An EPA program for monitoring ecological status and trends. *Environ. Monit. Assess.* 17:67-78.

Miller, R.R., J.D. Williams, and J.E. Williams. 1989. Extinctions of North American fishes during the past century. *Fisheries* 13:22-38.

Miller R.R., J.D. Williams, and J.E. Williams. 1991. Extinctions of North American fishes during the past century. *Fisheries* 14:22-38.

Moors, A., and R.J. O'Connor. 1992. *Development of an Avian Index of Biotic Integrity from Selected New England Lakes.* Final Report on CR 818179-01-0. Department of Wildlife, University of Maine, Orono.

Morrison, D.F. 1970. The optimal spacing of repeated measurements. *Biometrics* 26:281-290.

Moyle, J.B. 1956. Relationships between the chemistry of Minnesota surface waters and wildlife management. *J. Wildl. Manage.* 20:303-320.

Naiman, R.J., D.G. Lonzarich, T.J. Beechie, and S.C. Ralph. 1992. General principles of classification and the assessment of conservation potential in rivers. Pages 93-123 in P.J. Boon, P. Calow, and G.E. Petts, eds. *River Conservation and Management.* John Wiley & Sons, London.

Nehlsen, W., J.E. Williams, and J.A. Lichatowich. 1991. Pacific salmon at the crossroads: stocks at risk from California, Oregon, Idaho, and Washington. *Fisheries* 16:4-21.

Newell, A.D., and R. Hjort. 1991. Part II: Long-Term Monitoring Project Data Dictionary. In *Data User's Guide for the U.S. EPA Long-Term Monitoring Project: Quality Assurance Plan and Data Dictionary.* EPA/600/3-91/072. U.S. EPA Enviornmental Research Laboratory, Corvallis, Oregon.

Nielsen, L.A., and D.L. Johnson, eds. 1983. *Fisheries Techniques.* American Fisheries Society, Bethesda, MD.

Noss, R.F. 1990. Indicators for monitoring biodiversity: a hierarchical approach. *Conserv. Biol.* 4:355-364.

Oberdorff, T., and R.M. Hughes. 1992. Modification of an index of biotic integrity based on fish assemblages to characterize rivers of the Seine Basin, France. *Hydrobiologia* 228:117-130.

O'Connor, R.J., and M. Shrubb. 1986. *Farming and Birds.* Cambridge University Press, Cambridge, New York. 290 pp.

Ohio EPA. 1988. *Biological Criteria for the Protection of Aquatic Life.* Ohio Environmental Protection Agency, Columbus, OH.

Olsen, A.R. (ed). 1992. *The Indicator Development Strategy for the Environmental Monitoring and Assessment Program.* EPA/600/3-91/023. U.S. Environmental Protection Agency, Corvallis, OR.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. *Ann. Assoc. Am. Geogr.* 77:118-125.

Overton, W.S. 1989. *Effects of Measurement and Other Extraneous Errors on Estimated Distribution Functions in the National Surface Water Surveys.* Technical Report 129. Department of Statistics, Oregon State University, Corvallis, OR.

Overton, W. Scott, Department of Statistics, Oregon State University, Corvallis, Oregon (personal communications).

Overton, W.S., D. White, and D.L. Stevens, Jr. 1991. *Design Report for EMAP, the Environmental Monitoring and Assessment Program.* EPA/600/3-91/053. U.S. Environmental Protection Agency, Washington, D.C.

Paine, R.T. 1966. Food web complexity and species diversity. *Am. Nat.* 100:65-75.

Patalas, K. 1990. Diversity of the zooplankton communities in Canadian lakes as a function of climate. *Verh. Internat. Verein. Limnol.* 24:360-368.

Patrick, R. 1972. Aquatic communities as indices of pollution. Pages 93-100 in W.A. Thomas, ed. *Indicators of Environmental Quality.* Plenum, New York.

Paulsen, S.B., D.P. Larsen, P.R. Kaufmann, T.R. Whittier, J.R. Baker, D.V. Peck, J. McGue, R.M. Hughes, D. McMullen, D.L. Stevens, J.L. Stoddard, J. Lazorchak, W. Kinney, A.R. Selle, and R. Hjort. 1991. *The Environmental Monitoring and Assessment Program (EMAP) - Surface Waters Monitoring Research Plan, Fiscal year 1991.* EPA/600/3-91/022. U.S. Environmental Protection Agency, Corvallis, OR.

Payne, B.R., and R.M. Degraaf. 1975. Economic values and recreational trends associated with human enjoyment of mongame birds. Pages 6-10 in *Proceedings Symposium on Management of Forest and Range Habitats for Nongame Birds.* Gen. Tech. Rept. WO-1. U.S. Forest Service.

Peck, D.V. (ed). 1992. *Environmental Monitoring and Assessment Program: Integrated Quality Assurance Project Plan for the Surface Waters Resource Group, Fiscal Year 1991.* EPA/600/X-91/080. U.S. Environmental Protection Agency, Las Vegas, Nevada.

Pister, E.P. 1987. A pilgrim's progress from group a to group b. Pages 221-232 in J.B. Callicott, ed. *Companion to a Sand County Almanac.* University of Wisconsin Press, Madison.

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.* EPA/444/4-89-001. U.S. Environmental Protection Agency, Washington, D.C.

Platts, W.S., W.F. Megahan, and G.W. Minshall. 1983. *Methods for Evaluating Stream, Riparian, and Biotic Conditions.* Gen. Tech. Rep. INT-138. U.S. Forest Service, Ogden, UT.

Pollard, J.E., and K.M. Peres, eds. 1991. *Environmental Monitoring and Assessment Program: Surface Waters Implementation Plan – Northeast Pilot Lake Survey, Summer 1991.* U.S. Environmental Protection Agency, Las Vegas, NV.

Power, M.E., W.J. Matthews, and A.J. Stewart. 1985. Grazing minnows, piscivorous bass, and stream algae: dynamics of a strong interaction. *Ecology* 66:1448-1456.

Rao, J.N.K., and J.E. Graham. 1964. Rotation designs for sampling on repeated occasions. *J. Am. Stat. Assoc.* 59:492-509.

Rawson, D.S. 1939. Some physical and chemical factors in the metabolism of lakes. Pages 9-26 in E.R. Moulton, ed. *Problems of Lake Biology.* Publ. 10. Amer. Assoc. Advance. Sci. Washington, D.C.

Redford, K.H. 1992. The empty forest. *Bioscience* 42:412-422.

Reynolds, J.B., and A.L. Kolz. 1988. Electrofishing injury to large rainbow trout. *N. Amer. J. Fish. Manage.* 8:516-518.

Reynolds, R.T., J.M. Scott, and R.A. Nussbaum. 1980. A variable circular-plot method for estimating bird numbers. *Condor* 82:309-313.

Robbins, C.S., D. Bystrak, and P.H. Geissler. 1986. *The Breeding Bird Survey: Its First Fifteen Years, 1965-1979.* U.S. Fish and Wildlife Service Resource Publ. 157. Washington, D.C.

Ruttner, F. 1963. *Fundamentals of Limnology.* Translated by D.G. Frey and F.E.J. Fry. University of Toronto Press, Toronto, Ontario.

Schindler, D.W. 1987. Detecting ecosystem responses to anthropogenic stress. *Can. J. Fish. Aquat. Sci.* 44:6-25.

Schmidt, R.E. 1985. Zoogeography of the northern Appalachians. Pages 137-160 in C.H. Hocutt and E.O. Wiley, eds. *The Zoogeography of North American Freshwater Fishes.* John Wiley & Sons, New York.

Schneider, J.C. 1981. *Fish Communities in Warmwater Lakes.* Michigan Department of Natural Resources, Fisheries Research Report 1890. Lansing, MI.

Scott, W.B., and E.J. Crossman. 1973. Freshwater fishes of Canada. *Fish. Res. Bd. Can. Bull.* 184. Ottawa, Ontario.

Senner, S.E. 1986. Federal research on migratory nongame birds: is the U.S. Fish and Wildlife Service doing its job? *Am. Birds* 40:413-417.

Sharber, N.G., and S.W. Carothers. 1988. Influence of electrofishing pulse shape on spinal injuries in adult rainbow trout. *N. Amer. J. Fish. Manage.* 8:117-122.

Sharrock, J.T.R. 1974. The changing status of breeding birds in Britain and Ireland. Pages 204-220 in D.L. Hawksworth, ed. *The Changing Flora and Fauna of Britain.* Academic Press, New York.

Siegfried, C.A., J.A. Bloomfield, and J.W. Sutherland. 1989. Planktonic rotifer community structure in Adirondack, New York, U.S.A. lakes in relation to acidity, trophic status and related water quality characteristics. *Hydrobiologia* 175:33-48.

Skougstad, M.W., M.J. Fishman, L.C. Friedman, D.E. Erdman, and S.S. Duncan (eds.). 1979. Method I-4600-78, Automated phosphomolybdate colorimetric method for total phosphorus. In: *Methods for Determination of Inorganic Substances in Water and Fluvial Sediments: Techniques of Water Resources Investigations of the United States Geological Survey.* Book 5, Chapter A1. U.S. Government Printing Office, Washington, D.C.

Smeltzer, E., V. Garrison, and W.W. Walker, Jr. 1989. Eleven years of lake eutrophication monitoring in Vermont: a critical evaluation. *Enhancing States' Lake Management Programs*: 53-62.

Smith, G.M. 1950. *The Freshwater Algae of the United States.* McGraw-Hill, New York.

Smith, R.L. 1977. *Elements of Ecology and Field Biology.* Harper & Row, New York.

Smith, S.H. 1968. Species succession and fishery exploitation in the Great Lakes. *J. Fish. Res. Board Can.* 25:667-693.

Smol, J.P., and J.R. Glew. 1992. Paleolimnology. Pages 551-564 in W.A. Nierenberg, ed. *Encyclopedia of Earth System Science.* Academic Press, San Diego, CA.

Snedecor, G.W., and W.G. Cochran. 1967. *Statistical Methods.* 6th ed. Iowa State University Press, Ames.

Spellerberg, I.F. 1991. *Monitoring Ecological Change.* Cambridge University Press, New York.

Sprules, W.G. 1975. Midsummer crustacean zooplankton communities in acid-stressed lakes. *J. Fish. Res. Bd. Can.* 32:389-395.

Stevens, Donald L., Jr., ManTech Environmental Technology, Inc., U.S. EPA Environmental Research Laboratory, 200 SW 35th Street, Corvallis, OR 97333 (personal communication).

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.

Tallent-Halsell, N.G., and G.D. Merritt. 1991. *Environmental Monitoring and Assessment Program: Surface Waters Component Northeast Lakes Pilot Study Field Training and Operations Manual.* U.S. Environmental Protection Agency, Las Vegas, NV.

Terborgh, J. 1989. *Where Have All the Birds Gone?* Princeton University Press.

Tessier, A.J., and R.J. Horwitz. 1990. Influence of water chemistry on size structure of zooplankton assemblages. *Can. J. Fish. Aquat. Sci.* 47:1937-1943.

Tessier, A.J., and R.J. Horwitz. 1991. *Analysis and Interpretation of Zooplankton Samples Collected during Phase II of the National Lakes Survey.* U.S. EPA/600/R-92/012. Environmental Protection Agency, Corvallis, OR.

Thompson, Z. 1853. *Natural History of Vermont.* C.E. Tuttle, Rutland, VT.

Thornton, Kent W. FTN Associates, 3 Innwood Circle, Wuite 220, Little Rock, AR, 72211 (personal communication).

Tonn, W.M., J.J. Magnuson, and A.M. Forbes. 1983. Community analysis in fishery management: an application with northern Wisconsin lakes. *Trans. Am. Fish. Soc.* 112:368-377.

Underhill, J.C. 1985. The fish fauna of the Laurentian Great Lakes, the St. Lawrence lowlands, Newfoundland and Labrador. Pages 105-136 in C.H. Hocutt and E.O. Wiley, eds. *The Zoogeography of North American Freshwater Fishes.* John Wiley & Sons, New York.

University of Maine. 1991. *Laboratory Operations and Quality Assurance Report: EMAP/TIME Summer 1991.* Draft report. Environmental Chemistry Laboratory, Sawyer Environmental Research Center, University of Maine, Orono.

Urquhart, N.S., W.S. Overton, and D.S. Birkes. 1991. *Comparing Sampling Designs for Monitoring Ecological Status and Trends: Impact of Temporal Patterns.* Technical Report 149. Department of Statistics, Oregon State University, Corvallis.

U.S. EPA. 1983. *Handbook for Chemical Analysis of Water and Wastes.* EPA 600/4-79-020. U.S. Environmental Protection Agency, Cincinnati, Ohio.

U.S. EPA. 1987. *Handbook of Methods for Acid Deposition Studies: Laboratory Analyses for Surface Water Chemistry.* EPA/600/4-87/026. U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C.

U.S. EPA. 1989. *Handbook of Methods for Acid Deposition Studies: Field Methods for Surface Water Chemistry.* EPA/600/4-89/020. U.S. Environmnental Protection Agency, Washington, D.C.

U.S. EPA. 1990. *Biological Criteria: National Program Guidance for Surface Waters.* EPA-440/5-90-004. U.S. Environmental Protection Agency, Washington, D.C.

U.S. EPA. 1991. *Total State Waters: Estimating River Miles and Lake Acreages for the 1992 Water Quality Assessments 305(b) Reports.* Draft.

U.S. FWS. 1981. *Standards for the Development of Habitat Suitability Index Models.* 103 ESM. U.S. Fish & Wildlife Service, Washington, D.C.

Walters, C.J., J.S. Collie, and T. Webb. 1988. Experimental designs for estimating transient responses to management disturbances. *Can. J. Fish. Aquat. Sci.* 45:530-538.

Ward, R.C. 1989. A systems approach to design. Pages 37-46 in R.C. Ward, J.C. Loftis, and G.B. McBride, eds. *Design of Water Quality Information Systems.* Information Series No. 61. Colorado Water Resources Research Institute, Colorado State University, Fort Collins.

Weber, C.I. (ed.). 1973. *Biological Field and Laboratory Methods.* EPA-670/4-73-001. U.S. Environmental Protection Agency, Cincinnati, OH.

Whittier, Thomas R., ManTech Environmental Technology, Inc., U.S. EPA Environmental Research Laboratory, 200 SW 35th Street, Corvallis, OR 97333 (personal communication).

Williams, J.E., J.E. Johnson, D.A. Hendrickson, S. Contreras-Balderas, J.D. Williams, M. Navarro-Mendoza, D.E. McAllister, and J.E. Deacon. 1989. Fishes of North America endangered, threatened, or of special concern: 1989. *Fisheries* 14:2-20.

★ U.S. GOVERNMENT PRINTING OFFICE: 1993—750-002 / 60147

United States
Environmental Protection Agency
Center for Environmental Research Information
Cincinnati, OH 45268

Official Business
Penalty for Private Use
$300

EPA/620/R-93/003

Please make all necessary changes on the below 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.