RESEARCH TRIANGLE INSTITUTE

Contract No. 68-C9-0013
March 1990

# Background Paper: Use Support Assessment Methods


Prepared by

Randy Dodd
Mike McCarthy
Keith Little
Pat Cunningham
Julie Duffin
Research Triangle Institute
P.O. Box 12194
Research Triangle Park, NC 27709

Wayne Praskins
U.S. Environmental Protection Agency
Assessment and Watershed Protection Division
401 M Street
Washington, DC 20460

Ann Armstrong
School of Forestry and Environmental Studies
Duke University
Durham, NC 27706

Contract No 68-C9-0013

March 1990

# Background Paper: Use Support Assessment Methods

Prepared by

Randy Dodd
Mike McCarthy
Keith Little
Pat Cunningham
Julie Duffin
Research Triangle Institute
PO Box 12194
Research Triangle Park, NC 27709

Wayne Praskins
U S Environmental Protection Agency
Assessment and Watershed Protection Division
401 M Street
Washington, DC 20460

Ann Armstrong
School of Forestry and Environmental Studies
Duke University
Durham, NC 27706

# CONTENTS



## APPENDIXES

## FIGURES



## TABLE

# 1. INTRODUCTION

## The Current Process

Section 305(b) of the Clean Water Act requires each State, Territory, and Interstate Commission (hereafter collectively referred to as States) to submit a report to the EPA every 2 years that describes the quality of its surface and ground waters. In 1988, States reported on 29 percent of total stream miles in the United States, 41 percent of total lake acres, and approximately 76 percent of total estuarine miles. These §305(b) reports are summarized by EPA and transmitted to Congress. This reporting process is designed to provide each State with a comprehensive, systematic water quality assessment and EPA with information needed to assess water quality nationwide

In preparing their biennial §305(b) assessments, States are encouraged to compile a wide variety of chemical, toxicological, ecological, and other data describing the condition of their waters (and drainage areas) States are increasingly using data collected in biological surveys to complement more traditional physical/chemical measurements. States interpret the data to report the degree to which the designated uses specified in the State water quality standards are met. Four categories of use support are reported: fully supporting uses, partially supporting uses, not supporting uses, and fully supporting uses but threatened. Threatened waters are identified by downward trends in water quality or other evaluations indicating possible future degradation (U S. EPA, 1989)

## The Problem

EPA has provided only limited, and fairly simplistic, guidance on how to make use support determinations:

- States are asked to judge the quality of the data they use by classifying assessed waters as "monitored waters" or "evaluated waters "
- Section 305(b) guidelines recommend that chemical data be used to classify waters as fully, partially, or not supporting their uses based on the percentage of measurements exceeding criteria (These recommendations, included in this document as Figure 1-1, are referred to as "Figure 1" in the 1990 §305(b) guidance They were developed in the "States' Evaluation of Progress" project in 1981 )

| Assessment Basis | Support of Designated Use | | | |
|---|---|---|---|---|
| | **Assessment Description** | **Fully Supporting** | **Partially Supporting** | **Not Supporting** |
| Evaluated | No site-specific ambient data. Assessment is based on land use, location of sources, citizen complaints, etc. Predictive models use estimated inputs; are not calibrated/verified. | No sources (point or non-point) are present that could interfere with the use, or sources present but information indicates uses fully attained. Criteria attainment predicted. | Sources are present and information indicates uses are partially supported or uncertainty about use support. Complaints on record. | Sources are present and information clearly indicates use not supported. Criteria exceedences predicted. |
| Monitored (Chemistry) | Fixed station sampling or survey sampling. Chemical analysis of water, sediment, or biota. | For all pollutants, criteria exceeded in ≤10% of measurements and mean of measurements is less than criteria. Pollutants not found at levels of concern, where criteria not available. | For any one pollutant, criteria exceeded 11-25% and mean of measurements is less than criteria; or criteria exceeded ≤10% and mean is greater than criteria. Pollutants not found at levels of concern, where criteria not available. | For any one pollutant, criteria exceeded >25% or criteria exceeded 11-25% and mean of measurements is greater than criteria. Pollutants found at levels of concern, where criteria not available. |
| Monitored (Biology) | Site visit by qualified biological personnel. Rapid bioassessment protocols may be used. | Use fully supported; no evidence of modification of community (within natural range of control/ecoregion). | Some uncertainty about use support; some modification of community noted. | Use clearly not supported; definite modification of community. |

Classification Guidelines for Multiple Use Waterbodies

Fully supporting = All uses are fully supported

Partially supporting = One or more uses partially supported and remaining uses are fully supported

Not supporting = One or more uses not supported

**Figure 1-1. Criteria for designated use support classification.**

- If any one pollutant implies nonattainment of water quality standards, the waterbody is reported as not supporting its use
- The lack of national guidance on how to make use support determinations has contributed to inconsistencies among States in how use support determinations are made. Therefore, interstate comparisons of assessment results are suspect, and results in national summary statistics are of unknown validity.
- Some States rely exclusively on monitoring data to make use support determinations (e g., Massachusetts, Minnesota, Ohio, South Carolina, Texas, Virginia), whereas other States rely largely on "evaluative" information (e.g., New York, Mississippi, Alabama, Montana).
- Some States report waters where evaluative information indicates a problem as threatened; other States ignore evaluative information.
- Some States use a wide variety of chemical and biological monitoring data; others rely solely on fixed-station water column chemistry sampling.
- Many States use the "Figure 1" guidance to analyze their chemical data; others rely on professional judgment.
- Several States assign partial support status to waters experiencing severe fish kills; other States assign nonsupport status.
- Five States report 90 percent or more of their waters fully supporting uses; whereas one State reports only 1 percent fully supporting uses, reflecting differences in water quality **and** use support methodology. (Table 1-1 presents summary statistics from 1988 State §305(b) reports.)

The lack of guidance is a problem not only for EPA, but for those States that seek technical assistance to improve the quality of their §305(b) reports. States have requested assistance on how to deal with apparent conflicts between different types of data, how to account for frequency and duration specifications in water quality criteria, how to interpret unfamiliar types of data, and other topics.

## Project Objectives

This background paper was prepared for the EPA/State "305(b) Consistency Workgroup " The workgroup was created to assist EPA in achieving two goals:

- Improve national reporting by increasing consistency among States in data collection, data analysis methods, total waters reported, and/or the fraction of waters assessed
- Improve the accuracy and coverage of State §305(b) assessments by providing additional guidance on data collection and/or analysis methods.

## Table 1-1. Designated Use Support in Rivers and Streams

| State | Total River Miles | Miles Assessed: Total | Miles Assessed: Percent Evaluated | Miles Assessed: Percent Monitored | Percent Miles Fully Supporting | Percent Miles Partially Supporting | Percent Miles Not Supporting |
|---|---|---|---|---|---|---|---|
| Alabama | 40,600 | 11,174 | 85 | 15 | 91 | 6 | 4 |
| Arizona | 6,671 | 2,279 | — | — | 69 | 9 | 21 |
| Arkansas | 11,508 | 4,107 | 46 | 54 | 42 | 1 | 58 |
| California | 26,970 | 9,885 | — | — | 67 | 22 | 11 |
| Colorado | 14,655 | 10,000 | 54 | 46 | 86 | 7 | 7 |
| Connecticut | 8,400 | 880 | 33 | 68 | 66 | 27 | 7 |
| Delaware | 500 | 467 | 0 | 100 | 60 | 33 | 7 |
| Delaware River Basin | 206 | 206 | — | — | 94 | 0 | 6 |
| District of Columbia | 36 | 26 | 0 | 100 | 0 | 0 | 100 |
| Florida | 12,659 | 7,943 | 27 | 73 | 67 | 25 | 8 |
| Georgia | 20,000 | 20,000 | 66 | 34 | 97 | 2 | 1 |
| Hawaii | 349 | 349 | 28 | 72 | 76 | 23 | 1 |
| Illinois | 14,080 | 12,970 | 23 | 77 | 45 | 54 | 1 |
| Indiana | 90,000 | 5,181 | 28 | 72 | 68 | 19 | 13 |
| Iowa | 18,300 | 8,235 | 75 | 25 | 1 | 79 | 20 |
| Kansas | 19,791 | 6,888 | 57 | 43 | 58 | 11 | 31 |
| Kentucky | 18,465 | 8,653 | 63 | 37 | 71 | 10 | 18 |
| Louisiana | 14,180 | 8,483 | — | — | 68 | 25 | 7 |
| Maine | 31,672 | 31,672 | — | — | 99 | 0 | 1 |
| Maryland | 9,300 | 9,300 | 84 | 16 | 93 | 5 | 2 |
| Massachusetts | 10,704 | 1,646 | 0 | 100 | 43 | 36 | 20 |
| Michigan | 36,350 | 36,350 | — | — | 98 | 0 | 2 |
| Minnesota | 91,944 | 4,443 | 0 | 100 | 35 | 13 | 52 |
| Mississippi | 15,623 | 15,623 | 87 | 13 | 89 | 9 | 3 |
| Missouri | 19,630 | 19,630 | 77 | 23 | 52 | 48 | 0 |
| Montana | 20,532 | 19,505 | 85 | 15 | 63 | 34 | 3 |
| Nebraska | 10,212 | 5,690 | — | — | 57 | 21 | 22 |
| New Hampshire | 14,544 | 1,331 | 77 | 23 | 71 | 16 | 13 |
| New Mexico | 3,500 | 1,152 | — | — | 50 | 48 | 2 |
| New York | 70,000 | 69,988 | 95 | 5 | 76 | 12 | 12 |
| North Carolina | 37,378 | 33,275 | 45 | 55 | 67 | 28 | 5 |
| North Dakota | 11,284 | 9,850 | 44 | 56 | 69 | 31 | 0 |
| Ohio | 43,917 | 7,045 | 0 | 100 | 32 | 21 | 47 |
| Ohio River Valley | 981 | 981 | 17 | 83 | 0 | 100 | 0 |
| Oklahoma | 19,791 | 9,248 | 36 | 64 | 36 | 38 | 26 |
| Oregon | 90,000 | 27,738 | — | — | 45 | 31 | 24 |
| Pennsylvania | 50,000 | 13,242 | 39 | 61 | 73 | 13 | 14 |
| Puerto Rico | 5,373 | 5,373 | 67 | 33 | 46 | 21 | 33 |
| Rhode Island | 724 | 581 | 43 | 57 | 84 | 2 | 13 |
| South Carolina | 9,900 | 3,795 | 0 | 100 | 74 | 10 | 15 |
| South Dakota | 9,937 | 3,750 | 18 | 82 | 37 | 34 | 29 |
| Tennessee | 19,124 | 9,428 | — | — | 63 | 26 | 10 |
| Texas | 80,000 | 13,998 | 0 | 100 | 87 | 0 | 13 |
| Vermont | 5,162 | 5,162 | 83 | 17 | 88 | 7 | 5 |
| Virginia | 27,240 | 3,532 | 0 | 100 | 34 | 40 | 26 |
| Washington | 40,492 | 4,621 | 22 | 78 | 50 | 35 | 16 |
| West Virginia | 28,361 | 14,301 | 46 | 54 | 20 | 71 | 9 |
| Wyoming | 19,437 | 19,437 | 67 | 33 | 83 | 17 | 0 |
| Totals | 1,150,482 | 519,413 | | | 70 | 20 | 10 |

— Not reported

Source: 1988 State Section 305(b) reports

The workgroup will be asked to recommend practical, technically sound improvements to the §305(b) process on the type or quality of data that are used and on methods for interpreting data

## What Is in This Document?

This document focuses on:

- A review of 1988 use support approaches
- An overview of both monitored and evaluated assessment methods
- Detailed discussions and examples of several specific topics in appendixes: use of reference distributions for providing a more site-specific assessment of water quality criteria; water quality indices; use of flowcharts; trend analysis; and the utility of pollutant-specific, biological, toxicological, and remote sensing methods for determining use support.

# 2. ASSESSMENT METHODS USED BY STATES IN 1988 §305(b) REPORTS

As a result of limited guidance, and diversity in States' waters and programs, §305(b) assessment results vary widely. Table 1-1 presents data on the percent of assessed rivers and streams reported as fully supporting or impaired (partially or not supporting) by each State in 1988. Although interstate variability resulting from differences in types of waterbodies, criteria, available resources, pollution problems, percent of total waters assessed, and other factors is to be expected, it is desirable for national assessments (i e., §305[b]) to contain minimal variability due to different methods.

A review of 1988 use support approaches used in rivers, lakes and estuaries is summarized below A more complete analysis is provided for each Region and State in Appendix A.

## 2.1 Rivers and Streams

Of the 53 States, 18 reported using the EPA Figure 1 approach, 12 States reported using a modified EPA Figure 1 approach, 10 States reported using some type of water quality index (WQI) based on chemical data (including 3 States that used a WQI in conjunction with a modified Figure 1 approach), and 15 States provided insufficient information to determine the method used in their use support determinations. The U.S. Virgin Islands has no permanent streams for which to determine use support

Of the 18 States that reported using the EPA Figure 1 approach, many mentioned Figure 1 only once and gave no further indications that the approach was consistently used or that the results of this analysis would override best professional judgment

Twelve States used a modified EPA Figure 1 approach That is, they adopted the concept of making use support determinations based on evaluated vs monitored information, percent of samples exceeding water quality criteria, or other features of the EPA Figure 1 approach, but they took liberties with the exact selection criteria used Again, how rigorously some of these States applied their Figure 1-style decision criteria is unknown Professional judgment rather than systematic analyses may feature more prominently in use support determinations than the §305(b) reports indicate

## 2.2 Lakes

Review of the 1988 §305(b) reports does not provide a clear picture of how States make use support determination for lakes. Of the 53 States, 14 used the EPA Figure 1 approach or a modified EPA Figure 1 approach, 10 States used Carlson's Trophic State Index (TSI), 15 States used other trophic indices or other methods, and 12 States did not provide sufficient information to determine the method used in their use support determinations. It is not clearly stated, however, if States using trophic indices actually correlated the TSI values to use support determinations. Hawaii and the U.S. Virgin Islands have no lakes for which to determine use support.

A few States mentioned using toxics data in their assessments. States using the EPA Figure 1 approach probably screen for toxics exceedances during their STORET runs States that used only trophic status information may not consider toxics in their routine lake assessments, but may incorporate toxics concerns through the use of best professional judgment.

Of the 14 States that reported using the EPA Figure 1 approach, many mentioned Figure 1 only once and provided no further information as to how this approach was actually applied or whether the results of this analysis overrode best professional judgment.

## 2.3 Estuaries/Coastal Waters

Little specific information on methods used in estuarine use support determination is available in the §305(b) reports Of the 26 States having coastal waters, 13 States gave insufficient information to classify their assessment methods Twelve States reported using Figure 1 or a modification, and one (Florida) used a TSI.

# 3. OVERVIEW OF ASSESSMENT METHODS

Section 305(b) guidelines distinguish between monitored and evaluated approaches. Monitored assessments are those based on recent ambient, physical/chemical/bacteriological data, ambient biological data, or toxicological data. Biological data include quantitative information such as the relative abundance of different invertebrate or fish species Toxicological data consist of effluent bioassay results or, less frequently, ambient toxicity testing results.

Evaluated assessments are those based on information other than recent monitoring data Such information includes fish kill records; remote sensing; water quality and land use data; questionnaire and survey results; public input; and professional judgment The distinction between monitored and evaluated assessments can be expected to conflict in certain circumstances. This might occur where monitoring results fail to identify a perceived loss of uses (i e , where monitoring results and professional judgment are in conflict), or where monitoring data exist only for pollutants and media without water quality criteria, such as for metals in sediments

## 3.1 "MONITORED" METHODS

### 3.1.1 Interpreting Pollutant-Specific Data

Virtually all States compare pollutant-specific ambient data to water quality standards to determine use support. Standards are generally not intended to be fixed limits but to specify levels not to be exceeded for a specified duration at a specified frequency. Newer EPA criteria for toxic pollutants specify two values, each with its own duration and frequency specifications:

- The acute value should protect aquatic life if the **1-hour average concentration** does not exceed the criterion value more than **once every 3 years on the average**
- The chronic value should protect aquatic life if the **4-day average concentration** does not exceed the criterion value more than **once every 3 years on the average**

The duration and frequency aspects reflect the random or probabilistic nature of water quality variables

One fairly simplistic approach to accounting for the random behavior of water quality variables is the "EPA Figure 1" approach (see Figure 1-1), which recommends that the determination of use support be based on the number of measurements exceeding a criterion A waterbody is "fully supporting" its uses if fewer than 10 percent of the measurements exceed the criterion; "partially supporting" if between 10 and 25 percent of the measurements exceed the criterion; and "not supporting" if greater than 25 percent of the measurements exceed the criterion. The figure does not make use of the new two-number criteria. One State that has modified the Figure 1 approach to consider both acute and chronic criteria is Rhode Island. They report partial use support, for example, if between 10 and 33 percent of samples exceed the chronic criterion but no samples exceed the acute criterion.

Other techniques for interpreting pollutant-specific data are discussed in Appendixes B and C. The approach described in Appendix B improves on "EPA Figure 1" by replacing the arbitrary categories (0 to 10 percent, 10 to 25 percent, over 25 percent) with site-specific or region-specific categories The categories are calculated from reference probability distributions determined from data describing clean water sites. Appendix C discusses the use of trend analyses to determine threatened waters.

Another issue raised by Figure 1 is how to determine use support when measurements of several different parameters are available at a single site. Figure 1 implies a "worst case approach" in which an exceedance of a criterion for any one pollutant implies less than full use support An alternative approach is to use a water quality index that weights each parameters (see Section 3.1.2)

The Center for Exposure Assessment has recently developed support for a model (DYNTOX) that predicts frequency and duration of criteria exceedance for a given stream and waste source. The model addresses the problem of extrapolating monitoring data and enables the user to choose from three approaches: a continuous simulation approach (requiring continuous flow and concentration data), a Monte Carlo approach, and a log normal approach, which requires daily mean and variance flow and concentration data. Although this model should be useful for interpreting instream data when supported by intensive studies in high-priority waterbodies, it probably does not provide a practical broad-scale solution to determining the frequency and duration of criteria exceedances given current monitoring resources

The advantages of using pollutant-specific data to assess use support include:

- Numerical criteria are often available
- Standardized sampling and laboratory methods are available
- Ambient conditions can be related to sources

- Historical data are available with reasonable spatial and temporal coverage
- Standardized data base management systems are available (e g., STORET, WATSTORE)

Physical/chemical measurements can also have less-than-desirable attributes for use support determinations:

- Pollutants of interest must be known
- Criteria for the pollutants of interest often do not exist or are below analytical detection levels
- Ambient data are highly variable, resulting in an inability to test hypotheses at high confidence levels
- Single-pollutant grab sampling may be difficult to relate to aquatic criteria (e g., 3-day-average chronic criteria) or public health effects because of issues of exposure, synergism/antagonism, and biomagnification
- Monitoring results can be difficult to communicate to laypersons
- A substantial number of chemical and physical parameters may be needed to adequately represent water quality

Another major difficulty for national use is the difference in water quality standards among States. This issue manifests itself as bias (where two States have different standards for the same parameters) and as unequal coverage (where one State has a standard and another does not)

In addition to water column data, physical/chemical parameters for monitored assessments include toxics in sediment and fish tissue These measurents offer the advantage of integrating the effects of pollutants over time, thus reducing the impact of high water column variability. Their major disadvantages to assessments is the lack of a clear relationship between chemical levels and use support (i e., the lack of criteria)

### 3.1.2 Physical/Chemical Indices

Physical/chemical water quality indices (WQIs) provide a means of reducing large amounts of information into a single value that can ease communication between technical water quality experts and laypersons Indices can be constructed in a "worst case" approach (i e , they can signal nonsupport when any one parameter falls below an acceptable limit), but more typically they individually weight and aggregate several parameters Appendix D discusses frequently used indices and their development in some detail

A simple arithmetic WQI has the formula:

$$\sum_{i=1}^{n} q_i w_i \quad ,$$

where

- n = number of parameters (pollutants)
- i = the given parameter
- $q_i$ = the rating value for the $i^{th}$ parameter (from a special curve for the $i^{th}$ parameter relating instream concentrations to quality ratings for a particular use)
- $w_i$ = weighting factor for the $i^{th}$ parameter (i.e., how important the $i^{th}$ parameter is to the designated use).

Indices vary considerably in the number and type of parameters incorporated and in weighting and aggregation schemes.

The advantages of physical/chemical WQIs include the following:

- They can reflect the opinions of a large group of experts as to necessary weighting factors or rating curves, as well as deciding what constitutes support of designated uses.
- Once an index is developed, results are reproducible and minimal professional judgment is required.
- Programs are available to carry out WQI calculations automatically from STORET data.

Potential disadvantages of WQIs include:

- A loss of information in the aggregation process
- A lengthy development process (although several good WQIs already exist)
- A tendency to obscure individual high pollutant measurements
- Problems due to missing data or discontinued stations in some WQIs
- Difficulty in making a connection to an ecological or human health endpoint.

### 3.1.3 Interpreting Biological Data

Many States now supplement their physical/chemical monitoring with biological surveys of fish, macroinvertebrates, periphyton, plankton, or other biological communities Biological survey data are usually expressed as the number or condition of organisms of each species present. The data may be compared to lists of "key" or "indicator" species or may be reduced using a biological index A general discussion of biological methods for aquatic communities is given in Appendix E

Recent guidance on rapid bioassessment protocols (Plafkin et al., 1989) argues that biological (biosurvey) measures are the best methods for determining aquatic life impairment, whereas physical/chemical and toxicity testing (bioassay) data are more useful for detecting and assessing sources The principal advantages of biological methods are their ability to integrate a variety of processes and effects into a single direct measure of aquatic life use support and their cost-effectiveness.

### 3.1.4 Interpreting Toxicological Data

Indications from State §305(b) reports are that toxicological methods are not as widely used in making use support determinations as physical/chemical or biological data. As pointed out in the "draft Ecopolicy" (see Section 3 1.5), these methods can be an important tool in use support assessment by predicting the toxicity of receiving waters. For example, a chronic toxicity test might measure the minimum concentration of effluent causing reproductive effects in laboratory test organisms; this endpoint concentration then can be used to calculate whether indigenous organisms would experience chronic effects at low stream flows or other flows of interest.

The limitations of toxicological methods include the fact that toxicity test results, like chemical-specific ambient data, are indirect estimators of biological integrity That is, they assess the suitability of the aquatic environment to support a healthy community, but they do not assess the community itself Like biosurveys, toxicity tests address only the aquatic-life uses of a waterbody, while other methods may be more appropriate for other uses A general discussion of toxicological methods is given in Appendix F

### 3.1.5 Combining Different Types of Monitoring Data

EPA has approved a new policy ("Ecopolicy") on the use of ecological assessment information, which states that·

> chemical-specific analyses, toxicity testing, and biosurveys can each provide a valid and independent assessment of designated aquatic life use impairment When any one of the three types of assessments demonstrates that the standard is not attained, it is EPA's policy that appropriate action should be taken to achieve attainment

"Appropriate action" can be interpreted to include classifying a waterbody as not supporting or partially supporting uses. In contrast, North Carolina decided to use biological ratings preferentially over chemical ratings, which in turn were preferred over qualitative evaluations based on a series of public workshops and professional staff judgment for the 1988 §305(b) report. (Biological ratings disagreed with physical/chemical ratings [using an index] at about 30 percent of all sites examined ) Ohio used a similar decision approach. For their 1990 §305(b) reports, both Ohio and North Carolina are further integrating biological and physical/chemical approaches by attempting to validate physical/chemical indices with biological ratings (see Appendix D) This approach is essentially an alternative to the EPA Figure 1 approach

Another approach for integrating physical/chemical and biological data is to develop an integrated index. An example of this concept is the Ohio Lake Condition Index (LCI), which is also discussed in Appendix D. The index is a weighted sum of 13 quantitative and qualitative measures of water quality, including the Index of Biotic Integrity and general physical/chemical parameters.

## 3.2 "EVALUATED" METHODS

Evaluated assessments are made in the absence of current ambient chemical or biological data. Sources of "evaluated" information can include fish kill reports, fishery status reports, water supply closure reports, old assessments, citizen complaint logs, remote sensing data, and information on potential nonpoint source inputs. Professional judgment is used to assess the reliability of the information, determine the severity of the problems indicated, and rate use support. States often rely upon evaluated information to assess use support because economic factors limit the spatial, temporal, and/or parametric coverage of monitoring networks Information on nonpoint source impacts is often lacking For example, Oregon monitors approximately 4 percent of its stream miles, which receive 90 percent of the State's point source discharges (Oregon Department of Environmental Quality, 1988).

States can assess a greater percentage of their waters by using evaluated information For example, Virginia expects to increase its percentage of assessed waters from 20 percent in the 1988 §305(b) report to almost 100 percent for the 1990 report by including evaluated assessments (Virginia Water Control Board, 1989) The following discussion will describe advantages and drawbacks of determining use support with various types of evaluated information

### 3.2.1 Use Restrictions

The main limitation of information on use restrictions (e g., drinking water restrictions, fishing advisories or bans, shellfish harvesting closures, and swimming prohibitions), is a lack of standardized criteria for imposing restrictions For example, criteria for closing shellfish waters are based on measurements of fecal coliforms (Oregon, North Carolina), total coliforms (Massachusetts), enterococcus bacteria (Delaware), and *E. coli* bacteria (Maine) The variety of criteria can result in inconsistent application of use restrictions and subsequent use support determinations.

States also take different approaches to assigning a use support category to specific restrictions Maryland assesses chronic shellfish closures as not supporting designated uses, while conditionally approved waters are partially supporting their uses. In contrast, North Carolina assigns a partial support assessment to all shellfish closures because other uses are still supported (e.g , aquatic life and contact recreational uses).

Within-State variation is perhaps a more serious drawback associated with swimming advisories and beach closures, which are often inconsistently monitored and imposed by local health agencies Some States are reluctant to base assessments on swimming restrictions because regions that seldom sample for bacterial violations (and therefore infrequently impose swimming advisories) will appear to have better water quality than regions that strictly enforce closure criteria Despite inconsistent application of swimming restrictions, this approach does provide the most direct measure of attainment of Clean Water Act (CWA) swimmable goals

State or local health officials are responsible for implementing drinking water closures resulting from violations of Federal maximum contaminant levels (MCLs) and State standards States are required to report public supply closures in the §305(b) report, and the EPA Office of Water tracks these closures in the Federal Reporting Data System. Therefore, the information is accessible

### 3.2.2 Fishery Status

Fisheries information can indicate nonattainment of fishable goals or impairment of aquatic habitat and fishery use designations The extent and severity of fish kills, diseases, and abnormalities can be acquired from fish kill reports submitted to the EPA and by surveying State fishery biologists, commercial fishermen, and recreational fishermen

Fish kill characteristics, such as size and frequency of kill events, may indicate partial and nonsupport of aquatic life survival and propagation uses In Maryland, isolated fish kills caused by spills are interpreted as partial support of designated uses; areas affected by recurring fish kills are assessed as not supporting designated uses Establishing strict numerical criteria to rate the severity or frequency of fish kills may result in misleading assessments The species affected and cause of the fish kill (some are natural) must also be considered

Accurate numerical fish kill data may not be available because of the variability in the States' abilities to quickly investigate reported kills. Rapid response to fish kill discovery is important to establish size, severity, and causes, but may not be possible Discovery will also be biased toward reporting kills near developed areas where people are available to identify and report incidents. Natural conditions, abandoned mines, and forestry activities might also cause fish kills that could go undetected in isolated areas.

Occurrences of fish diseases and abnormalities (i.e., lesions, sores, tumors, or eroded fins) are also difficult to associate with use support status. Cause and effect relationships between contaminants and abnormalities are often unknown, susceptibility to abnormalities and diseases varies among species, and background frequencies of abnormalities are unknown. In some instances, the majority of a catch exhibits stress, and it is apparent that aquatic life use is impaired. However, local conditions may not be responsible for the abnormalities. Fish exhibiting stress symptoms may be exposed to contaminants elsewhere during their migratory life cycle.

States use different approaches to associate qualitative fishery status measures with use support assessments. For example, Colorado considers severe and frequent fish kills an indicator of partial support, while Delaware considers such waters to be not supporting designated use Inconsistency is aggravated by the lack of criteria for defining "frequent and severe" fish kills. Improved reporting requirements might provide an initial step toward improving the quality and reproducibility of evaluated assessments.

### 3.2.3 Presence of Sources

The presence of land-disturbing activities, or large wastewater inputs relative to stream flow, can indicate waters with a high probability of not fully supporting designated uses. However, States hesitate to assess designated uses based on this information Nonetheless, a number of States used land use and land cover information to prepare their 319 nonpoint source assessments. Many States relied upon State and Federal soil conservation agents to identify areas with highly erodable conditions and agricultural activities likely to degrade nearby waterbodies. Soil conservation agents acquire specific knowledge of localized problems through frequent field visits Local planning officials and foresters can also map land uses that might degrade waters

### 3.2.4 Public Input

Public perceptions of waterbody value and degradation may not correspond to water quality data but can provide valuable input to decisionmakers User surveys, citizen complaints, and public input at meetings indicate directly to natural resource managers user perceptions of use support, which can be associated with water quality measurements Public input can also notify State personnel of localized water quality problems

The public can identify impaired waters with complaints of foul odors, floating scum, water discoloration, and fish kills and diseases. A drawback to relying upon citizen complaints for assessing use support is that the information may be inaccurate and often the State staff are informed after the suspected problem has dissipated. A followup investigation may not be warranted if the complaint was improperly reported and the State's ability to perform investigations is limited by staff restrictions Kentucky has improved the reliability of public complaints by developing a massive public education program called Water Watch Water Watch projects train citizen volunteers to investigate suspected spills and illegal discharges and to monitor local streams. The program has sucessfully identified impaired waters and unpermitted discharges (Cooke, personal communication). Citizen monitoring has also been used effectively by other States such as Rhode Island, Illinois, and New York Guidance highlighting these programs is currently under preparation by the EPA Assessment and Watershed Protection Division

Public workshops played a prominent role in the development of a number of State NPS Assessments North Carolina held 14 NPS Assessment workshops across the State and invited Federal, State, municipal, county, and general public representatives to attend. The attendees identified many impaired waterbody segments that were not monitored Oregon's Department of Environmental Quality also held public meetings and distributed questionnaires to citizens while developing a Clean Water Strategy

## 3.3 COMBINING EVALUATED AND MONITORED INFORMATION

From discussions with State officials and examinations of §305(b) reports, it appears that professional judgment based on evaluated data is sometimes used to overrule monitored assessments For example, this might occur where water column monitoring has not detected criteria exceedances, but observations by fisheries biologists indicate severe biological impacts EPA guidance does not address this issue

It would be difficult or impossible to develop a standardized ranking for sources of evaluated information because of the variation in their reliability Some States have special units to quickly respond to citizen complaints and fish kill reports, elevating the reliability of such information Other States may have access to refined land use information on a Geographic Information System Each State may have to prioritize its own information sources, or even integrate all available information for each waterbody individually. Some States rely on the judgment of a single §305(b) Report Coordinator to make the final use support judgments Other States (e g , Virginia and Tennessee) request regional staff to develop consensus use support judgments for waterbodies in their jurisdiction

Special committees are another option developed by several States while preparing NPS Assessments, but it is not clear whether committees can reach a consensus about use support determinations. Idaho formed the Technical Advisory Committee with Federal, State, and local agency personnel, as well as representatives of interested citizen groups. Committee members were solicited for information about NPS impacts, the sources of the information, and the reliability of the information. All monitored and evaluated data were entered into a data base that generated lists of sometimes contradictory information The Technical Advisory Committee avoided ranking the information by presenting a range of NPS impacts for each waterbody rather than resolving conflicts in the information and consolidating the varied data into a single NPS assessment.

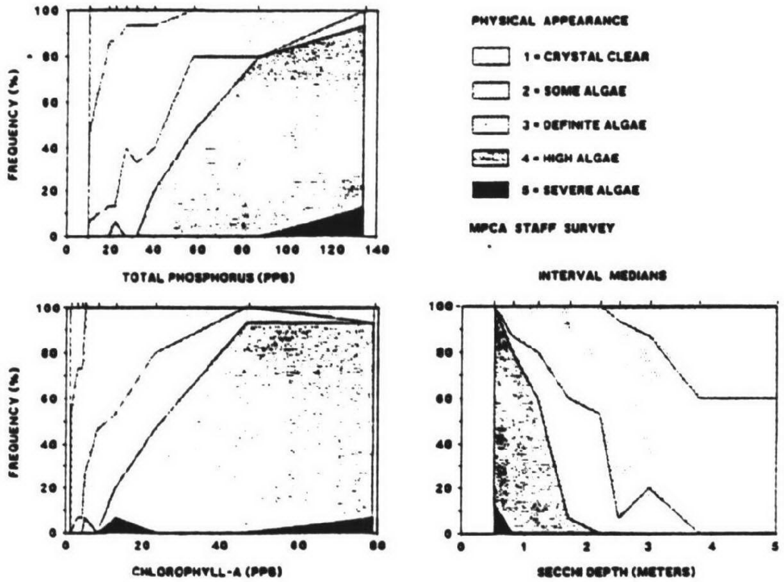
Another approach is to use surveys to test public perceptions of resource value against monitored information For example: Minnesota circulated a questionnaire asking the respondent to rate physical condition and recreational suitability of a lake, based primarily upon visual observations of algal abundance (Figure 3-1). Perceptions of recreational value were plotted against monitored trophic state indicators (total phosphorus concentration, Secchi disk depth, and chlorophyll *a* levels) The results presented the percentage of respondents who perceived each level of use support at lakes with varying trophic indicator measurements. The results could suggest ranges of trophic indicator values that correspond to use support categories, or conversely associate future survey results with trophic state and use support status. For example, fully supporting assessments could be assigned to the range of Secchi disk depths that 80 percent of the survey respondents perceive as fully supporting swimming use.

A. Please circle the one number that best describes the physical condition of the lake water today:

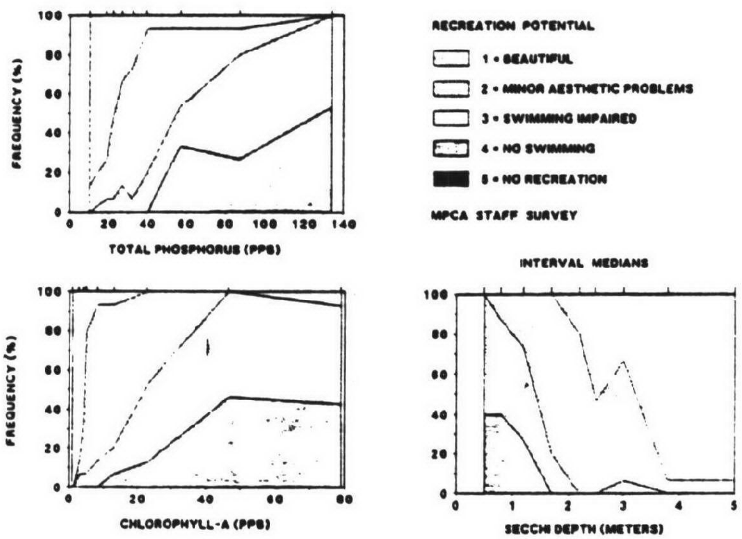
1. Crystal clear water.
2. Not quite crystal clear, a little algae present/visible.
3. Definite algal green, yellow, or brown color apparent.
4. High algal levels with limited clarity and/or mild odor apparent.
5. Severely high algae levels with one or more of the following: massive floating scums on lake or washed up on shore, strong foul odor or fish kill.

B. Please circle the one number that best describes your opinion on how suitable the lake water is for recreation and aesthetic enjoyment today:

1. Beautiful, could not be any nicer.
2. Very minor aesthetic problems; excellent for swimming, boating, and enjoyment.
3. Swimming and aesthetic enjoyment slightly impaired because of algae levels.
4. Desire to swim and level of enjoyment of the lake substantially reduced because of algae levels (would not swim, but boating is okay).
5. Swimming and aesthetic enjoyment of the lake nearly impossible because of algae levels.




Physical appearance ratings vs. lake water quality measurements.




Recreation potential ratings vs. lake water quality measurements.

Source: Heiskary, S. A., and W. W. Walker (1988)

**Figure 3-1. Minnesota's lake observer survey.**

# 4. REFERENCES


Cooke, Ken, Kentucky Division of Water Personal communication.

Heiskary, S.A and W W. Walker, Jr., 1988. Developing Phosphorus Criteria for Minnesota Lakes. Lake and Reservoir Management, 4(1):1-10

Oregon Department of Environmental Quality, 1988. *Water Quality Assessment and Program Plan* (Draft) Salem, Oregon.

Plafkin, J.L , M.T. Barbour, K.D. Porter, S K Gross, and R M Hughes 1989 Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish Assessment and Watershed Protection Division, Office of Water, U S. Environmental Protection Agency, Washington, DC 20460. Publication No. EPA 444/4-89-001, May

U S EPA, 1989 Guidelines for the Preparation of the 1990 State Water Quality Assessment (305(b) Report). Office of Water, Washington, DC, February.

Virginia Water Control Board. 1989. *305(b) Procedures Manual for the 1990 Edition.* Pages 4(1) - 4(2). Richmond, VA.

Walker, W W., 1985 Statistical Bases for Mean Chlorophyll *a* Criteria. In *Lake and Reservoir Management*, North American Lake Management Society

**Appendix A**

# APPENDIX A

# 1988 USE SUPPORT APPROACHES

# APPENDIX A. 1988 USE SUPPORT APPROACHES

This appendix describes assessment methods used by selected States in their 1988 305(b) reports States that reported using a straight EPA Figure 1 approach are not mentioned here

## A.1 Rivers and Streams

### Variations on the Figure 1 Approach

#### Region 1

Rhode Island is one of the few States that distinguished between chronic and acute criteria for priority pollutants For example, for partial support, priority pollutants could exceed **chronic** water quality criteria (WQC) in 10 to 33 percent of samples, but could not be present at acute concentrations

#### Region 6

New Mexico used a modified Figure 1 approach that treated toxics differently from other parameters For full support, New Mexico allowed a maximum exceedance for any parameter of < 150 percent of WQC, no parameter could exceed WQC in > 10 percent of samples; and no 307(a) toxics could be present at levels of concern We assume this means no exceedances were allowed for any 307(a) toxic

#### Region 7

Iowa referred to a methods document by the Iowa Department of Natural Resources for information on their assessment approach To be used, a station had to have > 75 percent of the number of samples expected from a 2-year quarterly sampling period for conventional pollutants and ammonia, for metals, all data points from 1982 to 1987 had to be present

Missouri's modified Figure 1 approach added a "physical criteria" description that had to be met regarding the type of substrate

Nebraska used both a modified EPA Figure 1 approach and a water quality index (WQI). The State did not describe how the WQI was incorporated into use support determinations The Figure 1 type matrix used separate frequency-of-violation criteria for toxics vs. other pollutants. For example, for toxics, full support allowed a maximum of a single parameter value in exceedance of WQC and no fish tissue values in excess of U.S Food and Drug Administration (FDA) or National Academy of Sciences (NAS)/NAE action levels.

### Region 8

Utah used a modified EPA Figure 1 approach that calculated a "severity index" for each parameter and an "overall index" for each waterbody. This approach may be further explained in the Region 8 WQI documentation.

## Water Quality Index Approach

Ten States reported using some type of WQI based on chemical data. Four States (Idaho, Illinois, New Jersey, Wisconsin) used the Region 10 WQI, and one (Tennessee) used the Region 8 WQI At least four States—Ohio, Maine, North Carolina, and Kansas—relied heavily on biological standards and/or biological indexes Several other States mentioned using biological data to designate use support, but gave no specific details on their implementation Several WQIs and their application in State use support assessments are mentioned below WQIs are discussed in more detail in Appendix D

### Region 4

Florida used a WQI that incorporates six water quality categories (water clarity, dissolved oxygen, oxygen demanding substances, bacteria, nutrients, and biological diversity) The 6 categories comprise 13 parameters that are arithmetically averaged to create the WQI. Biological diversity is assessed for macroinvertebrates (collected on both natural and artificial substrates) using the Shannon-Wiener Index and Beck's Biotic Index

Tennessee reported using the Region 8 WQI, which allows an "unlimited number of parameters and multiple types of criteria."

### Region 5

Illinois used the Region 10 WQI and Biological Stream Characterizations (BSCs) in a Figure 1 type matrix The BSC is used to determine use attainment in streams The stream classification system is predicated largely on attributes of lotic fish communities. In the absence of fisheries data, macroinvertebrate data or physical habitat descriptors may be used The BSC method is driven primarily by an assessment of fish community structure as represented by the Index of Biotic Integrity (IBI) for fish. The IBI method incorporates 12 measures (metrics) of fish community structure, including species composition and species richness, trophic state composition, and abundance and condition A Macroinvertebrate Index used in Illinois is a modification of the method of Hilsenhoff (1982)

Ohio gave excellent documentation of the methods used to determine use support A habitat index was used to determine whether habitat was causing partial or nonsupport. For biological data, results for three indices (IBI and Index of Well Being [Iwb] for fish, Invertebrate Community Index [ICI] for macroinvertebrates) were compared to State biological water quality standards For example, partial attainment was reported when one or two indices did not meet the ecoregion criteria, but did not suggest severe toxic impact. For chemical data, a modified Figure 1 approach was used that distinguished between acute and chronic criteria. Biological data overrode chemical data when both were available

### Region 6

Louisiana referred to a use impairment index but did not clearly explain its application The State used a very modified Figure 1 approach that employed only a primary determinant parameter and a secondary parameter for each use classification. A percent exceedance was then calculated for these parameters

## A.2 Lakes

### Region 4

Florida set up its own lake index by averaging the Carlson values for chlorophyll *a*, Secchi depth, and the limiting nutrient (nitrogen or phosphorus) No specifics were given on the correlation between trophic state index (TSI) and use support.

### Region 5

Minnesota used STORET to analyze 11 years' worth of data and calculate Carlson TSIs for (1) epilimnetic concentrations of total phosphorus and chlorophyll *a* and (2) summer Secchi disk transparency. These TSIs were then compared to ranges of TSI values corresponding to levels of use support For example, TSIs ≤50 indicated full use support for swimming and aesthetics; TSIs of 51 to 59 indicated threatened use support TSIs of 60 to 65 indicated partial support and TSIs of >65 indicated nonsupport The TSI ranges were selected based on a "user perception survey" (see Chapter 3).

**Region 7**

Kansas used the Carlson TSI, which was calculated based on chlorophyll *a* data The degree of aquatic life use support was estimated from the TSI. TSIs < 50 were considered fully supporting TSIs from 50 to 59 were considered partially supporting, and TSIs > 59 were considered nonsupporting.

**Other Trophic Indexes or Other Methods**

Fifteen States used other trophic indexes or other methods in use support determination for lakes Examples of those using other indexes or other methods are discussed below

**Region 1**

Massachusetts used a severity index based on measurements of six critical parameters. The six parameters included: hypolimnetic dissolved oxygen, Secchi disk reading, phytoplankton count (or chlorophyll *a*), total ammonia, nitrate-nitrogen, total phosphorus, and aquatic macrophyton

**Region 2**

New York used a PC program to analyze all lakes data for selected parameters and determine trophic status The method for determining use support was not presented

Puerto Rico suggested using special trophic status criteria from the Pan American Health Organization for determining trophic status, and presumably, use support

**Region 3**

Delaware used three Carlson Indexes based on chlorophyll *a*, transparency, and total phosphorus to characterize the trophic state of lakes surveyed Because the State's lakes contain extensive macrophytic and filamentous algae, total nitrogen and oxygen deficit were added to the trophic state evaluation How the TSI was used to determine use support was not reported.

**Region 4**

Kentucky clearly stated its use support criteria for lakes The criteria were based not on a TSI, but rather on the qualitative degree of fish kills, hypolimnetic and epilimnetic oxygen depletion, nuisance algal blooms or macrophytes, taste and odor problems, treatability problems, and suspended sediment problems

South Carolina appears to have used three approaches: a modified EPA Figure 1 approach, Carlson TSI, and the National Eutrophication Survey Index How these three approaches were used together was not reported, nor was their relation to use support determinations

**Region 5**

Illinois used a Lake Impairment Index that evaluates TSI values and the severity of impairment from sediment, algae, and macrophytes. Illinois also used biological data, field observations and best professional judgment in lake assessments.

Ohio proposed the use of a Lake Condition Index (Ohio LCI) composed of 13 parameters selected to provide a holistic evaluation of lake conditions and to meet 305(b) guidance These parameters included both monitored and evaluated biological, chemical, physical, and aesthetic information The biological parameters included the IBI for fish, nuisance growths of macrophytes, fecal coliform bacteria contamination, primary productivity based on chlorophyll *a*, and fish tissue contamination For the biological parameters, monitoring data were available primarily for the nuisance growths of macrophytes, fecal coliform, and primary productivity based on chlorophyll *a* values The chemical parameters included nonpriority pollutants, priority organics, priority metals, nutrients based on spring total phosphorus, sediment contamination, and acid mine drainage For the chemical parameters, monitoring data were available primarily for nonpriority pollutants, priority metals, total phosphorus, and acid mine drainage The physical parameter, volume loss due to sedimentation, had been monitored for some lakes, and the public perception of lake condition (aesthetics) when monitored was a measure of eutrophication based on chlorophyll *a* The 13 parameters used to evaluate use support incorporated best professional judgment and results of a lake questionnaire Some of the parameters/matrices were subjective in nature, but this represented an innovative attempt to apply many types of information to use support assessments, while reducing the level of purely subjective judgement

Wisconsin has attempted to use LANDSAT data to measure trophic status A description of how this related to use support was not provided

**Region 6**

Oklahoma mentioned using both statistical and TSI techniques to classify lakes as to trophic state The link between trophic state and use support was not explained

**Region 10**

Idaho developed a TSI to classify a subpopulation of its lakes through a one-time sampling during peak productivity using a linear-weighed sum of 11 water quality variables

**Reference**

Hilsenhoff, W.L., 1982 Using a Biotic Index to Evaluate Water Quality in Streams. Technical Bulletin No. 132. Department of Natural Resources, Madison, Wisconsin.

# Appendix B

# APPENDIX B

# USE OF REFERENCE DISTRIBUTIONS TO ASSESS WATER QUALITY CRITERIA VIOLATIONS

# APPENDIX B. USE OF REFERENCE DISTRIBUTIONS TO ASSESS WATER QUALITY CRITERIA VIOLATIONS

## B.1 Introduction

For purposes of determining whether waterbodies support their designated beneficial uses, the U.S. Environmental Protection Agency (EPA) has recognized this shortcoming of traditional WQC by providing Section 305(b) guidance based on percentile statistics. In "Guidelines for the Preparation of the 1990 State Water Quality Assessment (305(b) Report" (U.S EPA, 1989), a designated use is considered "fully supported" if at least 90 percent of the chemical measurements comply with the relevant WQC Designated uses are "partially supported" if at least 75 percent of the data comply and "not supported" if less than 75 percent of the data comply (See Figure 1 of the 305(b) guidance ) A shortcoming of these recommended acceptable risks is that they are neither site- nor chemical-specific and therefore may not represent true natural risks of criteria exceedances.

The purpose of this discussion is to present a statistically based approach to assessing WQC "violations" and to demonstrate highlights of the method by means of a case study.

## B.2 Proposed Approach

The proposed approach is quite simple, yet powerful, and is based on concepts previously described in the literature (Loftis and Ward, 1981; Loftis, Ward, and Smillie, 1983) The approach uses site-specific historical water quality data to construct a "reference distribution" for the constituent under consideration as shown for an idealized case (Figure B-1; all figures appear at the end of this appendix) In the figure, water quality data for an arbitrary constituent X have been used to estimate the probability distribution function (PDF) for X The horizontal axis represents concentration while the vertical axis gives the probability or relative frequency of concentration. Data used to construct the reference distribution would be selected so as to exclude man-induced pollution events to the extent possible For example, the data may come from clean waters upstream of the station in question, from a clean stream in the same ecoregion, or from the station itself with outliers thought to be related to unnatural causes removed

The WQC has been imposed on the PDF in Figure B-1 to identify the natural risk of violation, represented by the shaded area under the PDF to the right of the WQC. For example, if X is total dissolved solids (TDS) with a WQC of 250 mg/L, the natural risk might be, say, 1 percent. "Violations" then detected in new monitoring data would be compared to this natural risk to determine if, indeed, they exceed this natural risk level. If so, then the WQC has in fact been violated; if not, the new data can be considered as not significantly different from the reference distribution data. A single TDS sample of, for example, 270 mg/L could be considered as having no more than a 1 percent chance of occurring due only to natural variation; thus, one could conclude a pollution "event" with at least 99 percent confidence.

For the more common situation where a set of samples, say N samples, are to be compared to the reference distribution, it could be expected that no more than 1 percent of them would exceed the WQC due to natural variation alone If, for example, 3 percent of them are in violation, it could be concluded that 3 minus 1 percent, or 2 percent, represent real pollution events and the WQC has been violated. (Note that it would not be possible, without additional information such as knowledge of spills, upsets, etc , to differentiate between those 2 percent that were the actual pollution events and the 1 percent that were due to natural violations.)

## B.3 Ohio Ecoregion Case Study

### B.3.1 Introduction

The State of Ohio has implemented the ecoregion concept in defining and assessing water quality conditions. Ohio comprises portions of five ecoregions The State has established a set of reference monitoring stations in each one, selected on the basis of being the least impacted locations in the ecoregions. These reference stations are used primarily for biomonitoring from which biological criteria are developed. However, chemical data are also monitored. Ohio's reference stations should closely approximate natural, background conditions for the ecoregions. They also apply ideally for the development of reference distributions with which to assess WQC attainment for nonreference stations

### B.3.2 Results

#### B.3.2.1 Nonconditional Reference Distributions

Figure B-2 ((a) and (b)) contains frequency histograms for dissolved oxygen (DO) and lead concentrations in the Huron-Erie Lake Plain (HELP) ecoregion in northwestern Ohio. These histograms were developed using only reference station data and thus can be considered as reference distributions from which background risks of WQC violations can be estimated Within the HELP ecoregion there are three reference stations: 500080, 500290, and

500820. Data from these reference stations were downloaded from STORET and comprised monthly samples from 1973 through most of 1989, except for station 500820, which also includes some 1968 data. A few months are missing in each station's record. In all, 683 monthly DO samples were used to generate the DO reference distribution in Figure B-2 while 607 monthly lead samples were used to chart the lead reference distribution.

Of the 683 samples used in the DO reference distribution of Figure B-2, 14 (2.1 percent) were less than the Ohio WQC for DO of 4 5 mg/L. Thus, the natural risk level for DO criterion violations is estimated at 2.1 percent. Of the 607 samples used in the lead reference distribution of Figure B-2, only 4 (0.66 percent) exceeded the U S EPA maximum contaminant level (MCL) for lead of 50 µg/L. The natural risk of lead WQC exceedance is then estimated at 0 66 percent.

Figure B-2 ((c) and (d)) also contains frequency histograms for DO and lead at station 500510, a nonreference station arbitrarily selected out of three nonreference stations in ecoregion HELP Samples are again monthly samples from the period 1973-1989, with some missing months A total of 173 DO samples were available of which 5 (2.8 percent) were below the 4 5 mg/L WQC. Based on the natural risk of 2.1 percent from Figure B-1, one would expect only 3 6 violations (0 021 x 173) out of 173 observations due to background variation alone Although confidence limits have not been constructed here (this should be done in practice), it is very unlikely that a single additional violation beyond the 3.6, say 4, expected violations is statistically significant and one could reasonably conclude from this analysis that DO WQC exceedances at station 500510 are within the natural range of variation and that no DO problems are in evidence

The lead data show similar results. Graph (d) in Figure B-2 is the frequency histogram for lead at the nonreference station 500510. Of 150 samples, it would be expected using the natural risk of 0.66 percent that approximately 1 sample (0.0066 x 150) was in excess of 50 µg/L, under the null hypothesis of no difference in distributions Indeed a single sample is all that was observed to exceed 50 µg/L and the conclusion from this analysis is that lead at station 500510 does not "violate" the WQC

### B.3.2.2 Conditional Reference Distributions

While the foregoing analyses are correct, they can be criticized for ignoring seasonal differences in water quality because the data from which the distributions were generated were selected without regard to sampling time-of-year Seasonal differences in many water quality variables, especially DO, can be substantial due to seasonally related factors such as flow and temperature To improve the seasonal resolution of the analyses, it was decided to split the data sets into two seasons--summer, from May through October, and winter, from November through April These selections were arbitrary and it may be that a different scheme, perhaps with more seasons, is more appropriate. These split data sets were then used to generate "conditional" frequency histograms (conditional on the season) for DO and lead both at the reference stations and at station 500510. These conditional distributions are shown in Figure B-3.

As shown in diagram (a) in Figure B-3, the natural risk for summer DO violations is 1 9 percent (7 of 362) Interestingly, this is a marginally lesser risk than the 2 1 percent resulting from the nonconditional analysis (Figure B-1) but the difference is not likely to be significant. At station 500510, diagram (b) in Figure B-3 shows that only 1.1 percent (1 of 90) of the summer samples were below 4.5 mg/L, suggesting that summer DO conditions at this station are even better than those at the reference station. (Based on these data, it would be tempting to reduce the natural risk to around 1 percent.)

For winter DO violations, the conditional analysis yields a natural risk of 2.2 percent (7 of 321), as shown in graph (c) in Figure B-3, approximately equal to the 2.1 percent risk obtained from the nonconditional analysis. From this risk level, it would be expected that, of the 83 winter samples at station 500510, less than 2 would be below 4 5 mg/L if only natural variation were present. Instead, as shown in graph (d) of Figure B-3, four samples were below the WQC, suggesting that some problem may exist at this station in the winter. Notice that the enhanced resolution provided by the winter-conditioned distribution has detected a possible DO problem that was not apparent from the nonconditional analysis. One explanation might be that relatively greater ice cover may occur, thereby inhibiting atmospheric reaeration. (Again, it should be noted that this analysis has not considered placing confidence limits on the natural risk levels. Such limits might have revealed here that these four violations are not significantly different from the natural risk.)

The conditional analyses for lead (Figure B-4) do not alter the conclusions from the prior, nonconditional analyses that lead is not a problem at station 500510 Natural summer risks of exceeding the MCL are 1.21 percent (4 of 332) with only 1 of 79 (1 27 percent) summer lead samples at station 500510 in violation. This number, slightly higher than expected, is not likely to be statistically significant Natural winter risks are estimated from these data at 0 percent; i e., none of the 275 samples was in exceedance. Likewise, of the 71 winter samples at station 500510, none was in exceedance, indicating that winter lead levels are also not a problem at this station.

## B.4 Discussion

### B.4.1 Bootstrapped Reference Distributions

Ohio's use of ecoregions with relatively unimpacted reference monitoring stations provides an ideal application for the natural risk approach to assessing WQC attainment. However, Ohio is one of only a few States known to use such an approach for defining attainable water quality (others include Arkansas, Oregon, and Minnesota (Hughes and Larsen, 1988)) The remaining States must estimate natural risks from data that may be less well suited for this purpose. Indeed, in many cases, data from the particular monitoring station itself for which WQC attainment is being assessed will be the only data available from which natural risks can be identified. The reference distribution must then be "bootstrapped" from these data.

Figure B-5 illustrates such a bootstrapped reference distribution for winter DO at station 500510 Data used to construct the distribution were winter data from station 500510 prior to 1986. Data collected during and after 1986 could be used to test against the reference distribution to assess WQC violations, in a manner analogous to what might have happened in preparation of the 1988 305(b) use attainment analysis The natural risk estimate from the bootstrapped reference distribution is 6.5 percent (4 of 62). The 1986 and later data include no WQC violations and would be judged as fully supporting the DO WQC regardless of the natural risk estimate.

This particular set of data was selected because winter DO data at station 500510 were found to exhibit DO WQC violations in excess of the number expected as defined by the true natural risk (from the reference stations) The point of the example is to illustrate how bootstrapped distributions may over-estimate natural risk, thereby underestimating actual problems. The bootstrapped estimate of natural risk, 6.5 percent, exceeds the "true" natural risk of 2.2 percent as given by the actual winter reference distribution (Figure B-3, diagram (c)) The problems, of course, are that there is no easy way to screen the data to eliminate "unnatural" violations from "natural" violations and that including these unnatural violations tends to distort the reference distribution In some cases, State staff may have personal knowledge of plant upsets, spills, or other anthropogenic causes of poor water quality and can screen the data accordingly, but it is expected that this would be the exception and not the rule.

One method that might mitigate the overestimate of natural risk from bootstrapped reference distributions would be to perform a trend analysis on the reference distribution's candidate data If the candidate data were to show a trend of declining water quality, then it would be appropriate to use only that portion of the record (if any) that did not include the trend. At a minimum, the use of trend-free data would ensure that the estimated risk, albeit perhaps larger than natural risk, would serve as a status quo, or baseline, that could not be exceeded in future years. That is, the baseline risk would preclude further degradation even if it did not define the waterbody's potential water quality

### B.4.2 Confidence Limits on the Reference Distribution

Each reference distribution presented here was estimated from a particular data set. Because that data set is merely a sample of the entire population of constituent concentrations that completely define the reference distribution, the derived distribution is only an estimate of the true, but unknown, reference distribution. It is therefore desirable to explicitly acknowledge this source of error by placing confidence limits on the estimated distributions. These limits would be presented as upper and lower confidence limits and would be used to define when the number of observed WQC exceedances is significantly different from the expected number of exceedances so as to constitute a true violation Both parametric and nonparametric methods for determining confidence limits are discussed by Loftis and Ward (1981).

### B.4.3 How to Avoid Zero Risk

Occasionally, an empirically derived reference distribution, such as those presented here, will indicate that none of the data exceeded the WQC and, consequently, that the empirical distribution shows no natural risk Such was the case for the winter lead distribution depicted in diagram (c) of Figure B-4 It is very unlikely that there is zero risk of winter lead exceeding 50 µg/L due to natural variation; it is more likely that the data set simply was not large enough to contain such exceedances.

One method to avoid this problem is to fit a suitable, theoretical probability model (e g., normal, lognormal) to the data. The theoretical model would then be used to estimate the natural risk by extending the distribution's tail out beyond the WQC. It should be noted that this method has the disadvantage of uncertainty with regard to selection of the appropriate theoretical model, especially for small data sets In addition, State expertise and resources may discount the practicality of this approach.

Bayesian statistics offers another approach to the zero-risk problem by incorporating into the risk estimator the subjective notion that just because something did not happen in the past does not mean that it will not happen in the future Box and Tiao (1973) discuss Bayesian estimators.

### B.4.4 Conditioning for Other Factors

The use of reference distributions conditioned on time of year has been presented. As discussed by Loftis, Ward, and Smillie (1983), this same technique may be used to account for other factors that might affect water quality such as flow For example, it might be that low summer DO levels are more strongly correlated to low flows than to time of year per se. Accordingly, it would be more appropriate to condition the DO reference distribution on flow than on time of year

## B.5 References

Box, G.E.P and G.C Tiao, 1973 Bayesian Inference in Statistical Analysis, Addison-Wesley Publishing Co , Reading, MA

Hughes, R M and D.P Larsen, 1988. Ecoregions: An Approach to Surface Water Protection. Journal WPCF, 60(4), April

Loftis, J C and R C Ward, 1981. Evaluating Stream Standard Violations Using a Water Quality Data Base. Water Resources Bulletin, 17(6), December

Loftis, J.C., R.C. Ward, and G.M. Smillie, 1983 Statistical Models for Water Quality Regulation Journal WPCF, 55(8), August

U S. EPA, 1989 Guidelines for the Preparation of the 1990 State Water Quality Assessment (305(b) Report). Office of Water, Washington, DC, February

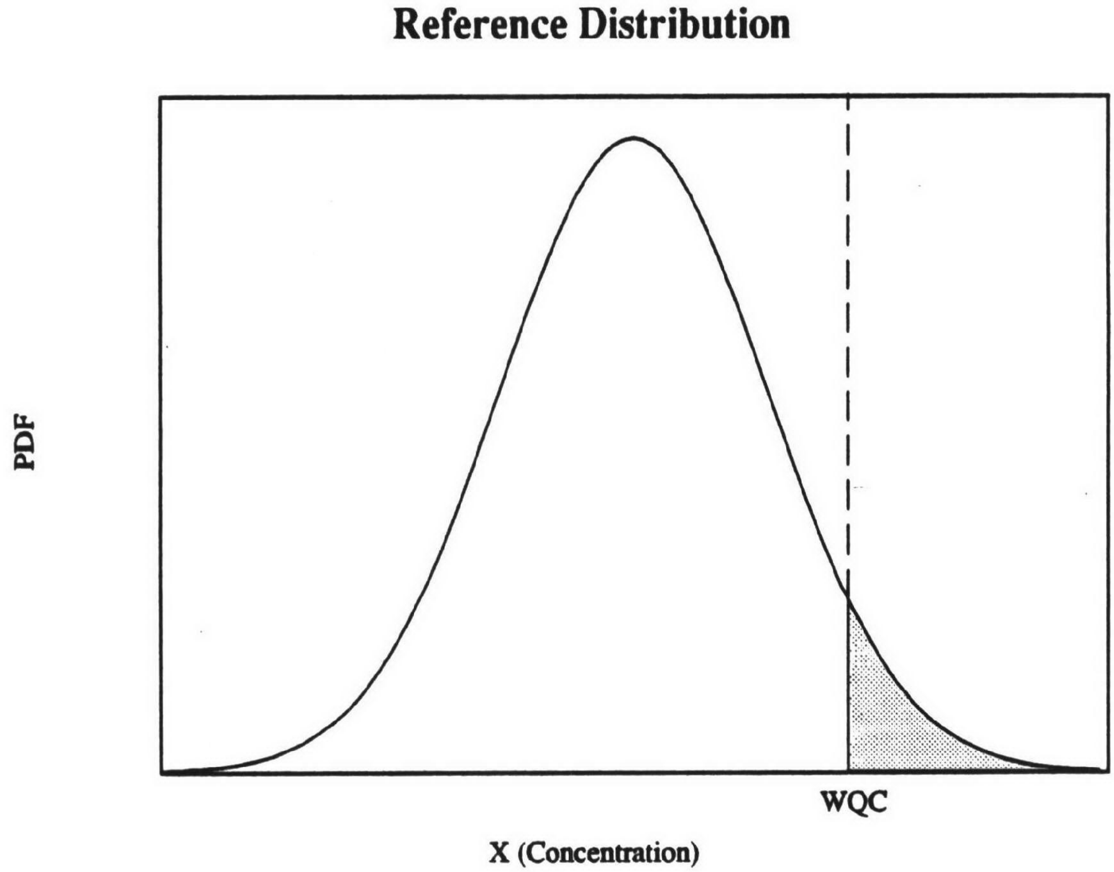



**Figure B-1. Reference probability distribution function (PDF) for pollutant X, with the indicated water quality criteria (WQC).**

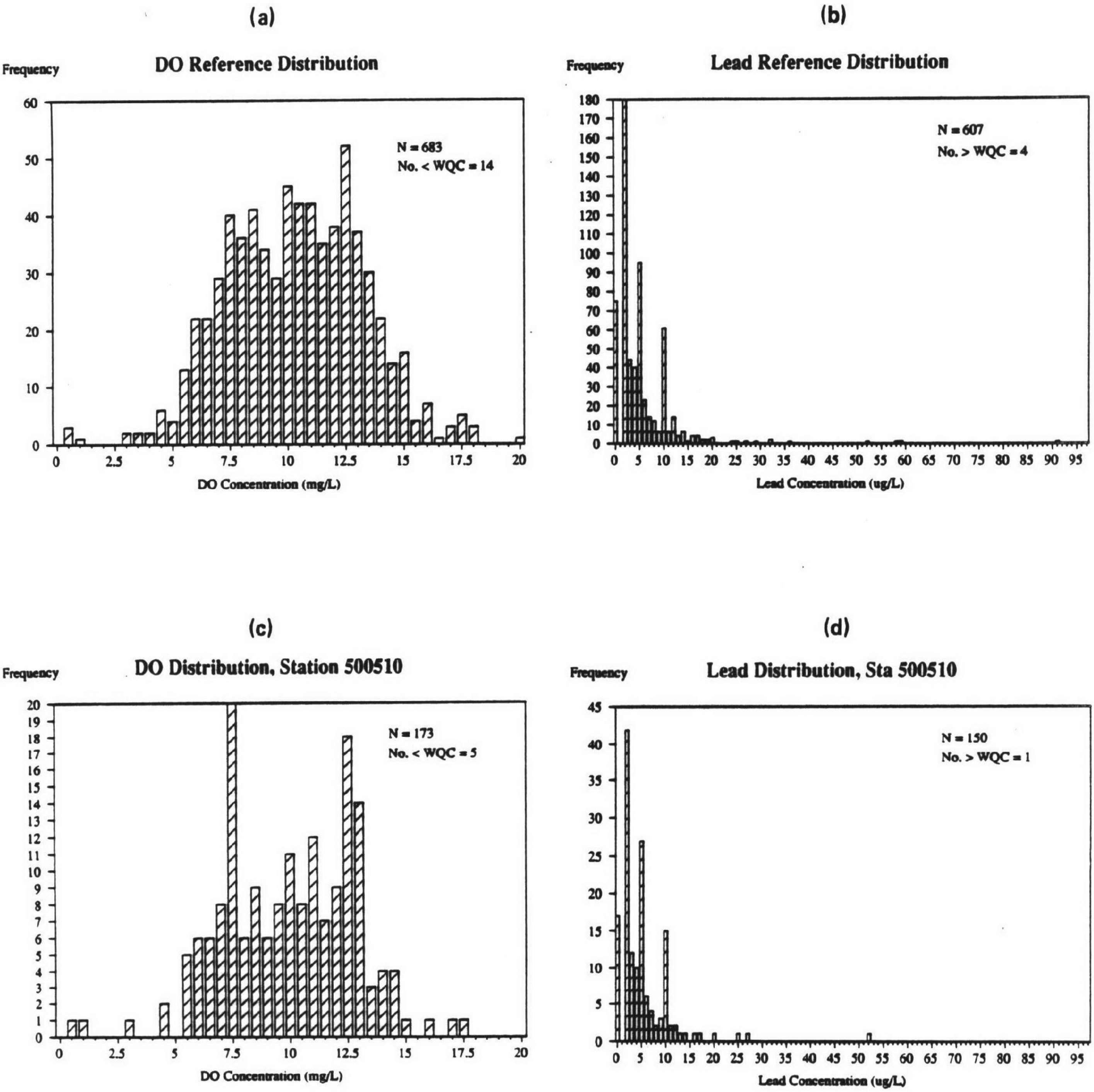



**Figure B-2. Nonconditional distributions for Huron-Erie Lake Plain ecoregion in Ohio.**

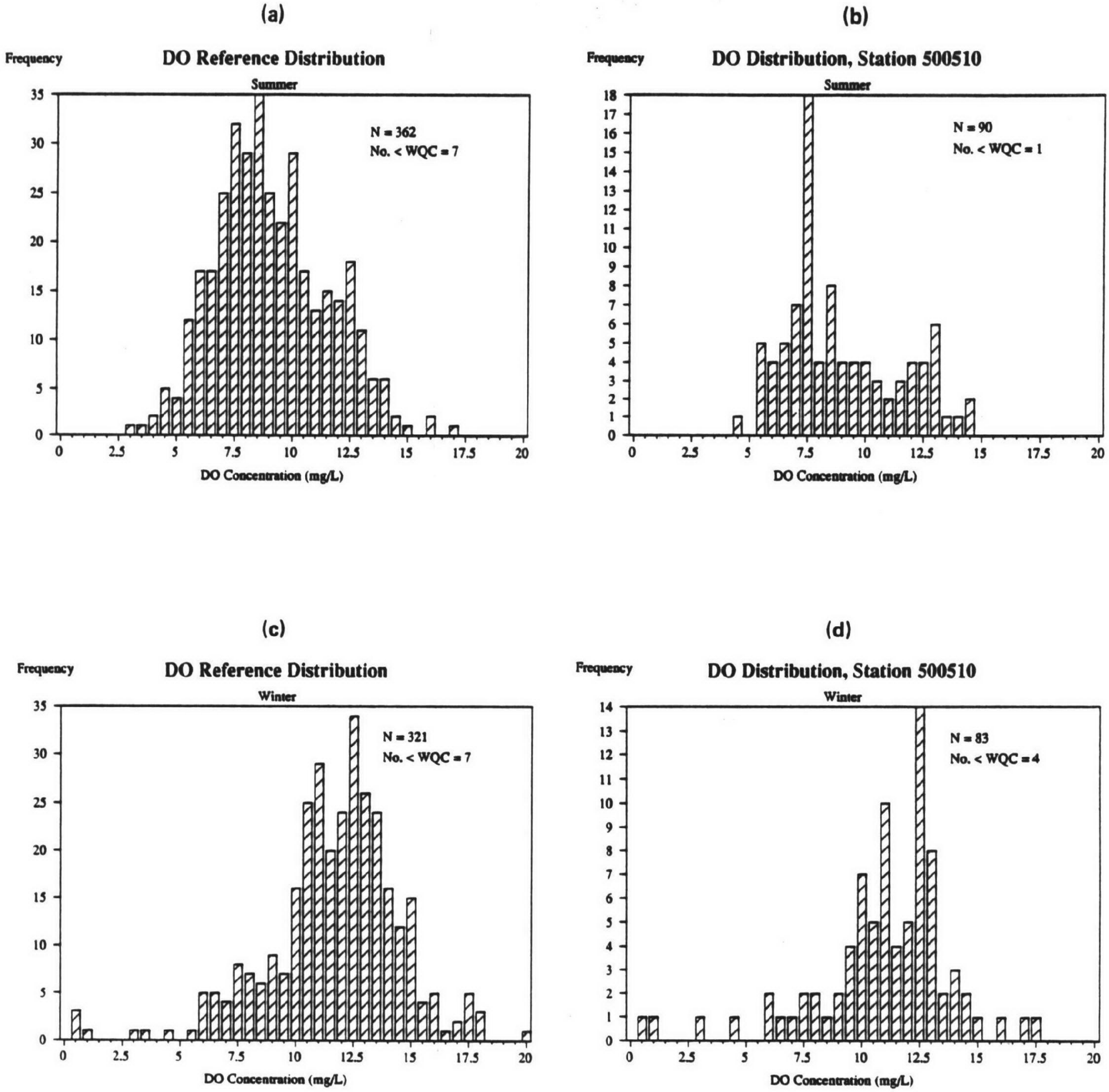



**Figure B-3. Seasonal dissolved oxygen (DO) distributions for Huron-Erie Lake Plain ecoregion in Ohio.**

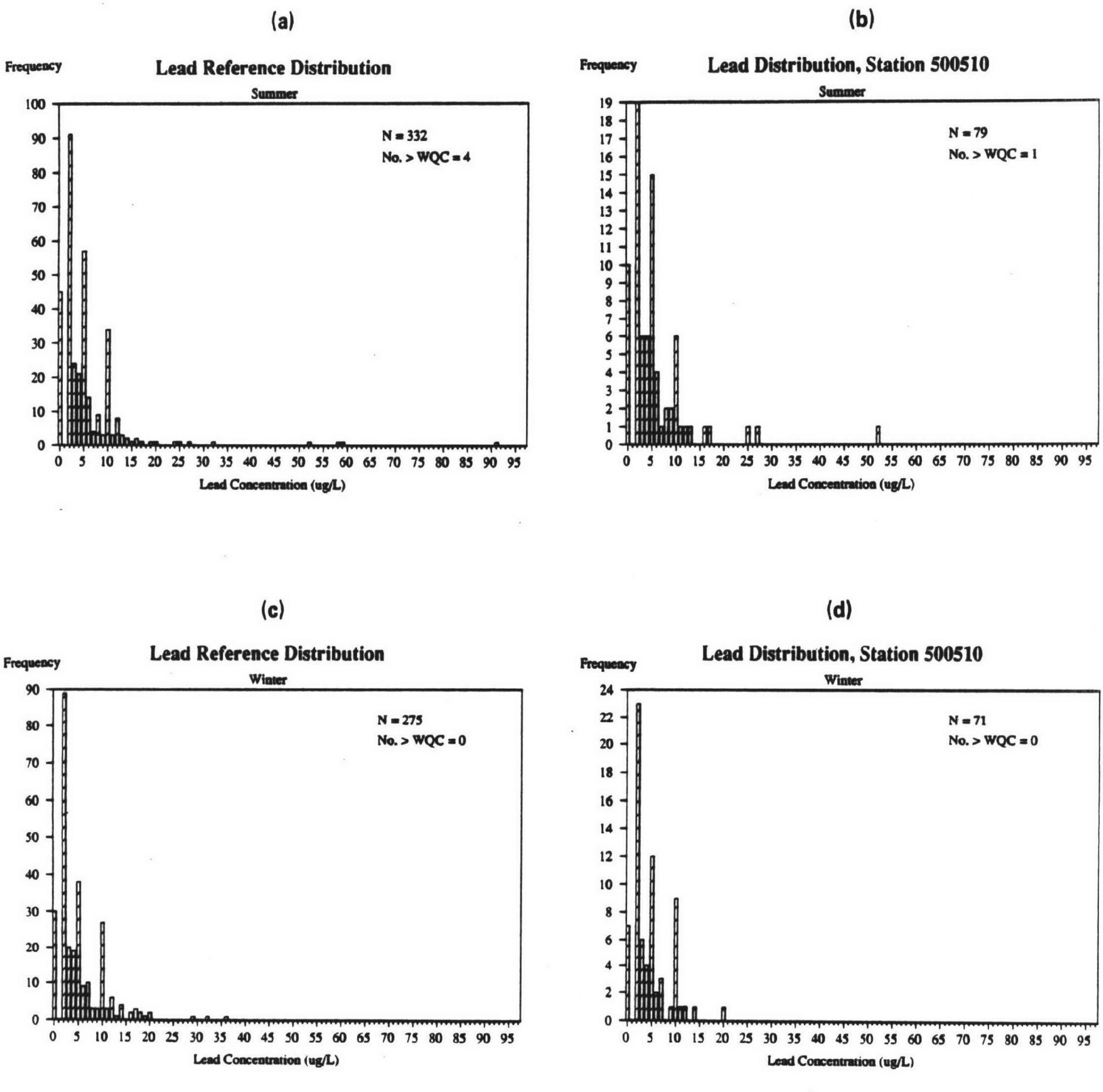



**Figure B-4. Seasonal lead distributions for the Huron-Erie Lake Plain ecoregion in Ohio.**

Appendix C

# APPENDIX C

# TREND ASSESSMENTS AND THREATENED USES

# APPENDIX C. TREND ASSESSMENTS AND THREATENED USES

The EPA's 1990 305(b) guidance document defines a "threatened" waterbody as one that is currently fully supportive of designated uses but poses the threat of not fully supporting those uses in the future due to some adverse temporal trend.

Detection of time trends in water quality data can be accomplished either by statistical methods such as the Seasonal Kendall Tau Test, or by essentially subjective methods such as visual inspection of a time series plot When adverse trends are clear and unambiguous, such trends can be detected satisfactorily by either method and the waterbody can be judged threatened with little possibility of error However, when adverse trends are not so clear, the possibility of decision errors becomes proportionately greater For these situations, the subjective methods suffer the considerable disadvantage of being unable to quantify these errors

With either method, two types of decision errors are possible A "Type I" error occurs when a waterbody is judged threatened that in fact is not (no trend is present) Conversely, a "Type II" error occurs when a waterbody that truly does have an adverse trend is judged nonthreatened The implication of either type of error is a misallocation of monitoring or control resources In the former case, a falsely identified threatened waterbody may unknowingly shift resources away from higher priority waterbodies In the latter case, a resource shift should be made that will not occur because of the failure to identify the threatened use

It is an unfortunate fact that the possibility of these two errors can never be completely eliminated, regardless of how many data are available or how carefully they were selected. Moreover, these two errors are not independent, indeed, there is a tradeoff between them in the sense that, for a given amount of data, a monitoring program design cannot be chosen that will reduce the probability of both errors simultaneously A reduction in Type I error occurs at the expense of an increase in Type II error and vice versa Only by collecting more data can both be reduced at the same time

Despite the fact that water quality managers must learn to live with the consequences of these two decision errors, statistical theory does at least offer a means by which they can be quantified Type I error is quantified by the "p" value, or "significance" level, used in conventional hypothesis tests

Type I error is a single number (probability) corresponding to the fact that there is a single null hypothesis (no trend) that is either accepted or not. Type II error is not a single number because, unlike the null hypothesis, an infinite number of alternative hypotheses exist and a Type II error is associated with each of them Consequently, Type II error is quantified by constructing curves that give Type II error probabilities as a function of specific alternative hypotheses These curves are commonly called "power" or "operating characteristics" curves

For data from an established monitoring program design, the probability of each error type can be determined, but it cannot be changed Significantly, for new monitoring programs, the relative magnitude of each error can be predetermined by its design For example, a policy decision might be that the detection of truly threatened waters with high probability outweighs the cost of judging some waters to be threatened when they are not. In this case, a design would be selected that would reduce Type II at the expense of Type I error.

In summary, statistical methods of trend analysis to identify threatened uses offer significant advantages over subjective methods due to their ability to quantify decision errors, not because they can somehow avoid them This advantage is particularly important when trends are ambiguous due to noisy and/or sparse data An important issue for analysis of existing data, and for the design of new monitoring programs, is the relative importance of the two types of decision errors

# APPENDIX D

# WATER QUALITY INDICES

1) Indices developed for rivers and lakes, but not for estuaries, wetlands
2) Parameters do not include metals, WET,
3) Limitations in collection of data

# APPENDIX D. WATER QUALITY INDICES

## Selection and Development

Indices can be categorized as general or specialized. General indices use the same equation for different uses but have differing ranges for use support classification (example in Figure D-1) Specific indices are tailor-made to best represent the water use that is being indexed Both the parameters and the weights for the indices are specialized for the use. This is beneficial for comparing waterbodies with the same use but does not allow comparison between two different uses

## Parameter Choice

Parameter choice is one of the most important steps in developing an index because the index will only reflect the water quality and use support status of the waterbody if the parameters are actually important to the intended use. For example, North Carolina has a use designation for Trout Waters. If parameters were chosen which placed a heavy emphasis on nutrients and pesticides and little or no emphasis on temperature and dissolved oxygen content, then the results of the index could, potentially, incorrectly indicate that the stream was supporting its use as a trout fishery For the general water quality indices (WQIs), conventional parameters such as turbidity, temperature, pH, dissolved oxygen, chemical oxygen demand (COD), and nutrients are included to try to give a general idea as to the quality of water in all of the uses. Specific use indices often contain these parameters and, in addition, also contain specialized parameters which best indicate the status of the water for the given use

Another important consideration for parameter choice is data availability Often, because of budget considerations or sampling techniques, data availability is highly variable. A third guideline for parameter choice is the existence of water quality standards or criteria for the parameter Indices are designed to be objective representations of water quality, and as a result, the frame of reference provided by standards and criteria is quite important

One technique which has been employed in the past to increase objectivity is the use of a DELPHI questionnaire process The true DELPHI method:

<table>
<tr><th rowspan="2">Level of Pollution (100 = Best)</th><th colspan="6">WATER USES</th></tr>
<tr><th>Public Water Supply</th><th>Recreation</th><th>Fish</th><th>Shellfish</th><th>Agricultural</th><th>Industrial</th></tr>
<tr><td>100</td><td>Purification Not Necessary</td><td rowspan="3">Acceptable for All Water Sports</td><td rowspan="3">Acceptable for All Fish</td><td rowspan="3">Acceptable for All Shellfish</td><td>Purification Not Necessary</td><td>Purification Not Necessary</td></tr>
<tr><td>90</td><td>Minor Purification Required</td><td rowspan="2">Minor Purification for Crops Requiring High Quality Water</td><td rowspan="2">Minor Purification for Industries Requiring High Quality Water</td></tr>
<tr><td>80</td><td rowspan="3">Necessary Treatment Becoming More Extensive</td></tr>
<tr><td>70</td><td rowspan="2">Becoming Polluted Still Acceptable Bacteria Count</td><td>Marginal for Sensitive Fish</td><td>Marginal for Sensitive Shellfish</td><td rowspan="2">No Treatment Necessary for Most Crops</td><td rowspan="2">No Treatment Necessary for Normal Industry</td></tr>
<tr><td>60</td><td>Doubtful for Sensitive Fish</td><td>Doubtful for Sensitive Shellfish</td></tr>
<tr><td>50</td><td>Doubtful</td><td>Doubtful for Water Contact</td><td>Hardy Fish Only</td><td>Hardy Shellfish Only</td><td rowspan="2">Extensive Treatment for Most Crops</td><td rowspan="2">Extensive Treatment for Most Industry</td></tr>
<tr><td>40</td><td rowspan="4">NOT ACCEPTABLE</td><td>Only Boating, No Water Contact</td><td>Coarse Fish Only</td><td>Coarse Shellfish Only</td></tr>
<tr><td>30</td><td>Obvious Pollution Appearing</td><td rowspan="3">NOT ACCEPTABLE</td><td rowspan="3">NOT ACCEPTABLE</td><td>Use Only for Very Hardy Crops</td><td>Rough Industrial Use Only</td></tr>
<tr><td>20</td><td rowspan="2">Obvious Pollution<br>Not Acceptable</td><td rowspan="2">Not Acceptable</td><td rowspan="2">Not Acceptable</td></tr>
<tr><td>10<br>0</td></tr>
</table>

Source: Dinius, 1987.

**Figure D-1. General rating scale for water quality.**

> is an effective method of integrating the opinions of experts without the disadvantageous effects of the committee process .A series of sequential questionnaires, separated by additional information and response feedback from earlier questionnaires, replaces committee debate. Important elements of the technique are anonymity of response, statistical analysis of responses, and increasingly refined feedback The technique is designed to control negative committee characteristics such as pressure from dominant individuals and irrelevant and often lengthy discussions.. Each round [of questionnaires] should bring the panel member's views on the questions closer together so that by the final round they are in approximate accord (Dinius, 1987, p 834)

Many of the studies involved in creating indices use a "modified DELPHI"; these are shortened versions of the principle In most modified DELPHI questionnaires there are only one or two rounds and the reaching of a consensus is not as crucial. This method is often shortened in the interest of time For instance, the Tennessee Valley Authority (TVA) DELPHI procedure took 7 months to run three rounds Experience has shown that if there are too many rounds the experts will begin to drop out, leaving the index developer with no statistically significant results

If the parameters are to be developed into a large-scale, widely usable index, they must be widely monitored on a regular basis Problems occur, on a State level, when budget cuts require the discontinuation of monitoring at some stations and parameters from the regularly monitored schedule. Indices can be extremely sensitive to missing data and as a result, in order to get defensible results, typically no more than one or two parameters may be missing at a given time Index developers have approached this problem from a variety of standpoints. The National Sanitation Foundation (NSF) index requires a perfect data set (no missing data); the index developed by House allows for up to two of the minor parameters but none of the major parameters to be missing; and Region 10 allows for wholesale substitution of parameters to ensure a complete data set

Complications also evolve from seasonally monitored data Parameters such as fecal coliform and chlorophyll *a* are important indicators of microbiological quality and trophic state Both of these are commonly only monitored during the summer From a monitoring standpoint this schedule is fine, but from an indexing standpoint, it impedes their use in annual or longer-term measurements One solution to this is the development of indices which run only for the summer and, therefore, include these parameters

There are several methods of dealing with missing data for indices (Landwehr, 1989, personal communication). The first is to skip over that sampling date for that station, as NSF has The second is to try to calculate a value based on historical values This is impractical for large-scale use from two standpoints: (1) for parameters that are monitored seasonally, there are no good historical data for the rest of the year; and (2) the effort involved in reentering each missing parameter for each date for each station for an entire State may be prohibitive A third approach, use of regression techniques, has the same drawbacks as the historical extrapolation. Once again, the time required to complete the data set may be greater than that of perusing it and there is the danger of actually creating, through regression, a large percentage of the entire data set. This could occur, for example, if many of the parameters are monitored only on a quarterly basis.

## Parameter Transformation

Parameters have different values and ranges based on the units that they are measured in and the natural levels found in water. These differences require indices to transform the raw data before aggregating it into a single index value While each index uses a slightly different method of standardizing its results, most use some type of rating curve transformation The methods of developing rating curves differ based on the objective of the index. For example, "standard based" indices incorporating toxicants use the standard or final chronic value (FCV) and assign a rating between 50 and 100 or 5 and 10, based on the rating scale (Armstrong, 1989; House, 1989, personal communication). The final acute value (FAV) is also assigned a rating value ranging from 25 (or 2.5) to 0 to demonstrate the low end of the rating curve. A line is formed using nonparametric functions which connects the points and results in a simple transformation from a data value to a standardized rating. In most cases, rating curves do not differ between general and specific use indices developed by the same person or agency

There may be a great deal of professional judgment in the placement of coordinates (e g., the FCV and FAV) on the rating curve The effort is made to define a range of water quality from "clearly impacted" (FAV) to "clearly unimpacted," e.g., the detection limit An effort is also made not to lose the detail of the water quality which became evident through sampling. In most cases the transformation is the most complex mathematical step to the index, but once established, it is also the most stable aspect of the index An example of a rating curve for a new North Carolina index is shown later in this appendix (Figure D-3).

## Weightings

All indices have some form of weighting factors on the parameters Those indices which claim to be unweighted actually assign an equal weight of one to all of the parameters. This "unweighted" system is not used widely in the indices currently in use because it is unreasonable to say that, for example, color is as important as temperature to the quality of water for cold water fishery uses. A second and widely used method is unequal weights of all the parameters based on some expert's or group of experts' opinion of the importance of the parameter for the given water use. In most instances weights are determined using either a DELPHI survey or one-time survey In either case, a panel of water quality experts is chosen and questioned about the importance of each parameter for indicating water quality. The results are then compiled using a method as simple as an arithmetic mean The weighting process, too, is highly subjective but the survey approach is designed to reduce the subjectivity as much as possible.

## Aggregation Methods

General and specialized indices have been developed using both arithmetic and multiplicative aggregation. The arithmetic mean formula is:

$$\sum_{i=1}^{n} q_i w_i \quad ,$$

where:

- n = number of parameters,
- i = the given parameter,
- $q_i$ = the rating value for the ith parameter, and
- $w_i$ = weighting factor for the ith parameter.

This formula is widely used both because it is easy to understand and because all weighted values are equally important.

The multiplicative mean takes the form of:

$$\prod_{i=1}^{n} q_i w_i \quad .$$

This formula has the characteristic of being more sensitive to extreme values for a given parameter.

## Examples of Indices

The majority of indices have been developed, or at least verified, for streams and rivers The scope of indices and the number available make the task of reviewing them quite a challenge, so we have reviewed only indices for which we have a clear, technical explanation Ground water indices are not widely published at this point The general applicability of indices to estuaries and wetlands is currently unstudied. One inventory, six river quality indices, and two lake indices are evaluated below. A review of trophic-state indices is not included.

### The Maryland Inventory

The State of Maryland uses a modified Figure 1 inventory system to classify its streams and lakes for the Section 305(b) reports Their method has a high degree of accuracy and fits the Clean Water Act's requirements of status toward the fishable-swimmable goals. The State uses a combination of monitored water quality data and evaluations where chemical data are not available. The inventory relies quite heavily on the State biologists for incident reports of fish kills and blooms as well as for judgment regarding general quality of all of the streams and rivers. Data are compiled onto worksheets (Figure D-2) and the severity of water quality impact is assessed based on the percentage of exceedance (for chemical data) and professional judgment (*Maryland Water Quality Inventory* 1985-1987, p. B5). Use support decisions are made based on a set of "rules of assessment."

This method is thorough and clear in its treatment of conventional pollutants, sources, and (most notably) evaluated information It can also be used, with the same accuracy, for different waterbody types Potential disadvantages include labor requirements and reproducibility (a second expert examining the same data may come up with a different view of use support) This approach may be a useful alternative to an index for those who are very opposed to indices.

## River Water Quality Indices

### Dinius' Water Quality Index

Dinius has refined the Social Accounting System originally published in 1972 (Ott, 1978a, p 219), resulting in a scientifically based water quality index The objective of this index is to provide a quantified method of determining the extent of increasing water pollution in order to estimate the cost of pollution prevention and control The index uses a geometric mean (multiplicative mean) to aggregate data and uses a full four-round DELPHI questionnaire to determine both the standards and the weights for the index It was developed for 6 water uses: public water supply, recreation, fish, shellfish, agriculture, and industry; and for 12 parameters. dissolved oxygen (DO), BOD-5, total coliform, fecal coliform, alkalinity, hardness, chloride, specific conductance, pH, nitrate, temperature, and color (Dinius, 1987, p 840) Dinius formulated rating curves and made data transformations based on the DELPHI panel's responses

This index provides a very good template for developing an index The process is well documented, with the exception of the rating curves A

## 1985-1987 Maryland Water Quality Inventory
## Water Quality Assessment Summary

Basin: ____________ Name: ____________

Class: ____________ Priority: ____________

| | | | |
|---|---|---|---|
| River (mi) | Monitored: ______ | Evaluated: ______ | Total: ______ |
| Estuary (mi) | Monitored: ______ | Evaluated: ______ | Total: ______ |
| Lake (ac) | Monitored: ______ | Evaluated: ______ | Total: ______ |
| Coast (mi) | Monitored: ______ | Evaluated: ______ | Total: ______ |
| Wetland (ac) | Monitored: ______ | Evaluated: ______ | Total: ______ |
| Groundwater | Monitored: ______ | Evaluated: ______ | Total: ______ |

**LAND USE**

Urban: ______ Agricultural: ______ Forest: ______ Wetland: ______ Mines: ______

**MONITORED**

| | | State WQ Standards | | | | | Subjective | | |
|---|---|---|---|---|---|---|---|---|---|
| Station | WB Type | Temp | DO | pH | Turb | Bact | N | P | Chla |

| STP | WB Type | Compliance | MGD | Pretreat |
|---|---|---|---|---|

| Industry | WB Type | Compliance | MGD | Pretreat |
|---|---|---|---|---|

| CSO/Stormwater | WB Type | Compliance | MGD | Pretreat |
|---|---|---|---|---|

## 1985-1987 Maryland Water Quality Inventory
## Water Quality Assessment Summary

Basin: ____________ Name: ____________

**EVALUATED**

**Swimming Ban**

| Area | WB Type | Source | Cause |
|---|---|---|---|

**Fish Kill**

| Area | WB Type | Source | Cause |
|---|---|---|---|

**SHELLFISH CLOSURES/CONDITIONAL ACRES**

**Permanent**

| Area | WB Type | Source | Cause |
|---|---|---|---|

**Conditional**

| Area | WB Type | Source | Cause |
|---|---|---|---|

**Fish Advisory**

| Area | WB Type | Source | Cause |
|---|---|---|---|

**Algae Bloom**

| Area | WB Type | Source | Cause |
|---|---|---|---|

**Agricultural Water Pollution Complaints**

| Area | WB Type | Source | Cause |
|---|---|---|---|

**Mine Impact**

| Area | WB Type | Source | Cause |
|---|---|---|---|

**Figure D-2. Example of watershed worksheet for water quality assessment.**

disadvantage in using the index for use support determinations is that the rating curves are not based on the standards of any State or on EPA criteria. Another potential problem is the choice of parameters (e.g., toxicants are not included for drinking water). Unfortunately, parameters may not be readily substituted into and out of this index due to the lengthy process needed to weight the parameters

**House's Water Quality Index**

House has developed several indices to be used by the United Kingdom Public Water company as both an enforcement and an assessment tool for the country's river waters. House has proposed both multiplicative and arithmetic indices but has found, through verification, that the arithmetic index best represents the quality of the streams based on the national standards. This index is based not only on Great Britain's standards but also on the European Economic Community's (EEC's) and EPA's standards and criteria. The House index will be implemented in Great Britain starting in early 1990, and as a result, it has been streamlined a great deal. The Water company has developed software that automatically computes the water quality index value for a given sampling date (House, 1989, personal communication).

The "index" is actually a family of indices which House reports may be used either separately or combined. The first index is the General Water Quality Index, which has nine parameters. These are: DO, BOD-5, $NH_3$-N, total coliform, suspended solids, pH, nitrates, chlorides, and temperature. In the general index each parameter is assigned the same weighting for all classifications. The general index is an arithmetic mean index using rating curves developed from the available standards and criteria and weights based on a modified DELPHI questionnaire House deals with missing data by allowing the program to run if all of the first five parameters (listed above) are present and if no more than one of the others is missing (House, 1989, personal communication). In addition to the General Water Quality Index, she has also developed three other indices which are based on the specific parameters important to human drinking water, aquatic toxicity, and drinking water standards Because in each of these indices the exceedance of a standard is unacceptable, a geometric mean aggregation was used to emphasize any exceedance problems. Also, the rating curves were based on a 0 to 10 range instead of the 10 to 100 range of the general index. Another major difference is that the specific indices were not given any weighting factors because House felt that if there was an exceedance in one parameter it was equally toxic to humans and aquatic organisms

The indices developed by House suit many of the needs for a national index because they are standard-based and designed to be applied to a relatively large range of geographic areas and classifications. Another advantage to the House index is that it is very well documented. House did not use complex math or a full DELPHI questionnaire and as a result, modifications to this index are the easiest and most legitimate of the indices reviewed Other advantages to this index include. the rating curves are both easy to construct and adjust, and the index has been fully statistically tested and verified in Great Britain Disadvantages include a bias towards drinking water parameters because Great Britain does not have a designated class for cold water fisheries Along the same line, many States do not have standards for parameters such as chlorides in any classification other than drinking water Also, the rating curves are based primarily on Great Britain's standards and as a result would need modification before being adopted in the United States House does not use the "fully supported," "partially supported," and "not supported" categories to rate waters within a classification; however, it would not be difficult to add these delineations into the index results and in this way address the use support question.

**Region 8 Water Quality Index**

The Region 8 Water Quality Index is a chemical index designed to express water quality in terms of use impairment The first step in using the index is to determine the frequency and amount of standard exceedance for individual parameters A severity index (SI) is calculated from the exceedance data for each applicable use Station SI values within a reach are averaged to determine a reach SI A use impairment value (UIV) is then calculated for each reach as the sum of the SI values for different uses. EPA calculates UIV values The States then augment these values with subjective, evaluated information for 305(b) reports

**Region 10 Water Quality Index**

Region 10 has developed a comprehensive index based on a complex arithmetic aggregation This index was designed to be highly adaptable for each of the States in Region 10 and as a result it has addressed many of the problems faced by index developers The index is based on the following categories of parameters: temperature, oxygen, pH, bacteria, trophic state, aesthetic state, solids, organic toxicity, and metal toxicity. This index is applied to all streams, regardless of classification, and ratings of blue, yellow, and red are produced The ratings may be interpreted for uses as fully, partially, and nonsupporting. The report on the index states that

> the WQI compares measured water quality with recommended "fishable/swimmable" Federal criteria by default These criteria are a synthesis of national criteria, State standards, information in the technical literature, and professional judgment....An overall WQI number is calculated for every selected water quality sampling station with sufficient data The overall WQI number (calculated monthly) is an aggregation of subindices for 10 pollution categories which are weighted by the relative severity of criteria exceedances for each group (Peterson and Bogue, 1989)

Each of the parameters is weighted by a value assigned to each use category of water The subindex value, $z_{jk}$, is calculated by (Ott, 1978b, p. 293):

$$z_{jk} = Q_{j,k} * F_{j,k} * W_k \quad , \tag{1}$$

where:

$Q_{j,k}$ = quality rating for station j and category k

$F_{j,k}$ = frequency of violation of recommended limit

$W_k$ = weight for the $k^{th}$ category.

$$I_k = 1/2 \sum_{j=1}^{n-1} \left(Z_{j+1,k} + Z_{j,k}\right) * \left(d_{j+1} - d_j\right) \tag{2}$$

where:

$I_k$ = water quality subindex for category k along the reach

$z_{j,k}$ = water quality value of the $k^{th}$ category at the $j^{th}$ station in the reach

$d_j$ = river mile distance of the $j^{th}$ station ($d_j < d_{j+1}$)

n = number of stations.

These values are then aggregated using a summation for each of the categories for the index value (I) for the reach

$$I = \sum_{i=1}^{k} I_k$$

This index addresses the problem of missing values by calculating replacement values for a group of parameters, not just a single parameter. For example, if nitrite were the parameter required for the index, but it was not sampled, then the STORET retrieval program would look for any of five other nitrogen parameters; if one of these were present, it would be scaled and substituted Rating ("severity") curves are employed to adjust all of the parameters into the same scale; documentation of these curves has not been received to date

This index has many attributes which are not seen in the other indices. The first is that it is designed to use STORET data and as a result is also designed to deal with factors such as changes in monitoring schedules. Second, it has been used in the United States for at least 10 years and as a result many of the problems of the newer, untested indices have been worked out A third major advantage is the flexibility of the program to be adapted to different States' needs (Bogue, 1989, personal communication). Several questions and potential disadvantages remain. The documentation that has been reviewed includes very little technical information as to the aggregation technique, the rating curves, and the origin of the weights. The use of linear interpolation to fill gaps in the data that cannot be filled by direct substitution may be misleading in some instances (e.g., seasonal data sets). Another concern is that metals data cannot fall into the highest category, regardless of lab results. In many cases the cutoff for the middle (yellow) category is also the detection limit, which means that index values may indicate that a support is threatened because of the inability to measure lower levels. A final concern is the large-scale substitution which takes place in order to maintain a complete data set. In the case of nitrogen, for example, ammonia is more toxic (at the same concentration) than nitrate, yet the index has the ability to calculate an index value based on either compound. Where parametric coverage is not uniform, assurances need to be made that this degree of substitution is not leading to the comparison of "apples and oranges" in the index results.

**National Sanitation Foundation Water Quality Index**

The NSF index is one of the older indices (1973) and as a result has had many modifications made to it The index is based on nine parameters: DO, fecal coliform, pH, BOD-5, NO3, PO4, temperature (degrees centigrade from the equilibrium), turbidity, and total solids. There are three use categories which this index considers: public water supply, fish and wildlife propagation, and contact recreation (Brown and McClelland, 1974, p 3) The parameter selection and the weightings of the parameters were both accomplished using a full DELPHI questionnaire and five rounds of questioning As part of the DELPHI questionnaire the experts were asked to create rating curves with levels of water quality from 0 to 100 These responses were then averaged together to form the curves. The early papers published on the index used an arithmetic aggregation; however, a later report on the index, "WQI Enhancing Appreciation of Quality of Improvement" (McClelland et al., 1973) reports that the aggregation process is a multiplicative one. The change occurred as a result of a Ph D dissertation presented by Landwehr (1974) which tested both the arithmetic and multiplicative aggregation processes on the NSF index and found better correlation between the multiplicative and experts' opinions of the water quality at given sites One additional characteristic which this index has is the ability to work within data availability constraints. This excerpt from the 1976 report (McClelland, Brown, and Deininger) has been included because it is true for all of the American indices which use STORET:

> Naturally it is disappointing to those who developed and maintain STORET that so little of the water quality data it contains lends itself to retrospective determination of index values by means of WQI or any other proposed formulation. This is not critical of STORET, but of the inconsistency and inadequacy of data (for the index development purposes) entered by agencies supplying the data. Data deficiencies can frequently be handled by applying estimation techniques. They can be entirely overcome in the future by adopting a few simple but important policy decisions relating to parameter selection and use of STORET (McClelland, Brown, and Deininger, 1976, p. 24).

The index could be adjusted to a use support index simply by altering their current scale of "very bad" to "excellent" into, for example, use supporting = excellent and good; partially supporting = medium; and nonsupporting = bad and very bad. This is advantageous from two standpoints: first, these designations were determined as part of the questionnaires and as a result do not represent arbitrary scorings; and second, the scaling process is helpful when coupled with EPA's support scorings to lend meaning to "use supporting" for the layperson. This index was developed using the opinions of experts from all over the United States. As a result of the experts' distribution, the parameters chosen also are useful indicators throughout the United States. This is the only index which has published thorough accounts of its validation process and as a result it has demonstrated itself as a sound index with few reasoning errors and a strong user-support network. The index is in a finished, verified state and as a result, would require little or no alteration for future 305(b) use.

**Proposed North Carolina Water Quality Index**

The proposed North Carolina index is based on House's general index, with modifications to better suit the State's needs The index being developed will include 12 parameters for both conventional water quality indicators and toxics This index is a monitor of the water column only and as a result does not include parameters that rapidly adsorb to particulates (e g., pesticides) Similarly, sediment and fish tissue data will not be incorporated For fresh waters this index will use 12 parameters: DO, total solids, temperature, turbidity, arsenic, cadmium, total chromium, copper, lead, mercury, nickel, and zinc. These parameters will be applied to all waters except Nutrient Sensitive Waters since nutrients and chlorophyll have not been included. Salt waters will have the same parameters, with the exception of the toxic metals because of analytical difficulties in measuring water column metals in saline waters

The index, like many of the others, used a modified DELPHI questionnaire to determine both the parameters which are included and the weights of each parameter Rating curves are being developed using the lowest concentration measured in North Carolina (or the detection limit), the State standard for the parameter, the EPA FAV concentration and the concentration which falls at 95 percent on a distribution curve for concentrations measured in North Carolina This method of determining the high score is designed to choose the true highest recorded score and not anomalies resulting from lab error an example showing development of a rating curve for zinc is given in Figure D-3

The greatest deviation from House's index is in the weightings In the North Carolina index the weights are based on the parameter **and** on the the classification This modification was made to allow the importance of individual parameters to vary based on the projected use of the water. On the other hand, to allow the comparison of unlike use classifications, a constant set of parameters was used in each classification

Once data have been retrieved from STORET and transformed through the rating curves, aggregation takes place. Like House's index, this index uses the Solway aggregation method to total the parameters' values for each date This formula is:

$$WQI = 1/2 * \left(\sum_{i=1}^{n} q_i w_i\right)^2 .$$

For a given station there would be anywhere from 3 to 13 of these values for a 2-year period, and in order to aggregate them to a single index value for that station for that time period, a distribution curve is created. The final index value for that station is the 5-percent value on the distribution curve This aggregation method is chosen based on House's work Many authors have noted that the most obvious approach of taking the mean of the aggregates overestimates the quality of water when using an arithmetic approach. Taking the 5-percent value eliminates this overestimation.

North Carolina's index is still in the early stages of development and has not undergone the degree of verification of the other indices discussed The State does plan to use the index, in its unfinished state, for the 1990 Section 305(b) report This index has several advantages over older indices: the rating curves and standards are based on the most recent toxicological data, the index is using the experience and criticisms of the older indices to avoid many of the problems of index development, and it is being written in SAS within the STORET environment so that there is no need to actually handle the raw data at all This is a tremendous time saver and allows less technically oriented staff members to work with the index The index has the ability to have 7 more parameters added onto it, with the same statistical significance as the first 12, when consistently monitored data become available This index is designed to address the use support classifications directly as a result of being developed for the 305(b) report and therefore is ideal for the goals of the EPA To a large extent this is mathematically the simplest

**Figure D-3. Development of a rating curve for Zinc.**

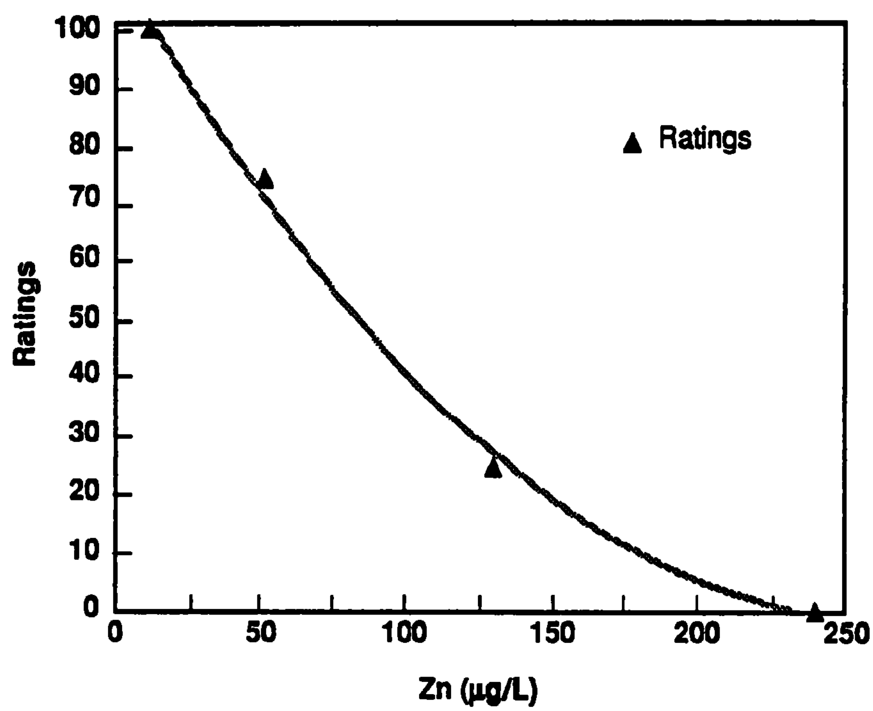
Step 1: STORET retrieval to find distribution of concentrations.
Step 2: Isolate the 5% and 95% concentrations.
Step 3: Set up table:

| Rating | USE Water Supply |
|---|---|
| 100 | 25 (Detection Limit) |
| 75 | 50 (Aquatic Life Standard) |
| 25 | 130 (FAV) |
| 0 | 240 (95% Recorded in North Carolina) |

Step 4: Graph and find equation for curve.
Step 5: Test on actual data to ensure that concentrations fall within the 100-0 range.
Step 6: Apply the rating to the index to find a subindex value.

Source: Proposed North Carolina WQI

**Rating Curve for Zinc**



index discussed and as a result is the easiest to manipulate to fit other States' needs. A final important distinction is that this index, unlike the others, is being calibrated and verified using biological data, not a panel of experts.

The chemical index values will fall in the range of 100 to 0 with use-category cutoffs (calibration) being determined by the biological data and criteria Verification will take place in a similar manner using different stations. The most notable exception to this is the calibration for water supply waters. Macrobenthos data may indicate pristine water conditions in water supplies which are not supporting their use due to coliform contamination. As a result, a panel of experts will be used to designate the cutoff ratings for water supply

## Lake Water Quality Indices

### The Ohio Lake Condition Index

This index was introduced in November 1988 and is titled as a "new approach to lake classification." One of the main objectives of this index is to address overall water quality and not just trophic state. The index is aggregated using an arithmetic mean and is reportedly easily modified to accommodate revisions of water quality standards and changes in management objectives. The objectives of this index are to meet the Water Quality Act of 1987 and as a result the use support designation has been included in this index from the beginning. The index is based on 13 parameters or matrices which include: index of biotic integrity (IBI), nuisance growths of macrophytes, fecal coliforms, primary productivity based on chorophyll *a*, fish tissue contamination, nonpriority pollutants, toxic organics, toxic metals, sediment contamination, nutrients based on spring total phosphorous, acid mine drainage, volume loss due to sediment, and aesthetics Those parameters which had standards were based on the standards while those that did not, like aesthetics, were based on best professional judgment Each matrix was judged as either supporting, partially supporting, or nonsupporting for each sampling date for each station The support categories were each assigned a number of points with the best quality being one and the worst being ten To avoid the problem of missing data, this index totals up all of these subindex scores and divides them by the number of matrices assessed. The resulting index value is a three-part number which identifies the overall condition of the lake, any "support threatened" or "nonsupporting" matrices, and finally the number of matrices which are missing This index does not make use of rating curves per se and does allow for more data to be missing than any of the other indices

Potentially this index shows more information than just a number in its results, but the final value is complex enough that laypersons may need more guidance to understand its meaning The index is based on standards and does address the use category approach of water quality determination

Potential drawbacks are that it is very labor intensive to apply and allows for up to 50 percent of the data to be missing.

### The Tennessee Valley Authority (TVA) Water Quality Index

This index is essentially the same as Dinius' index with some modifications made to the rating curves to include water quality standards and additional parameters. This index was developed to allow TVA to simplify its data so that it could be interpreted by laypersons without losing its technical validity. It has been developed specifically for the Southeast. This index is designed for shallow southeastern lakes and reservoirs with weak summer stratification and short retention times. Butkus (1989) used a three-round DELPHI questionnaire to determine which parameters to include in the index as well as weightings and ratings. The parameters included were: suspended solids, fecal coliform, true color, DO, chlorophyll *a*, temperature, and pH. The index was designed to examine the quality of water for four water uses: fish and aquatic life, recreation, domestic water supply, and industrial water supply

This index is still in its development stages. In fact, this review is based on the draft sent out to the experts taking part in the questionnaire for the index. It has been written up in a program to be loaded onto a PC Once again, this may be a factor for STORET users because data must be manually downloaded in order to use the program. In terms of application for use support determinations, this index has the same drawbacks as that of Dinius: it is not based on the standards of a State and the water uses are not necessarily those of any State. The problem of missing data was not addressed in this draft of the index and as a result further development and documentation are needed to fully explain its applicability

## References

Armstrong, Ann, 1989. The Development of an Arithmetic Mean Index for Determining North Carolina's Water Quality. Master's Thesis Draft

Bogue, Bill, 1989 Personal communication, August.

Brown, R M. and N.I McClelland, 1974 Up from Chaos: The Water Quality Index as an Effective Instrument in Water Quality Management National Sanitation Foundation, Ann Arbor, MI

Brown, R M , N I McClelland, R A Deininger, and J.M Landwehr, 1973 Validating the WQI Presented at ASCE National Meeting on Water Resources Engineering, Washington, DC

Butkus, Steve, 1989 Personal communication, May

Butkus, Steve and Donald W Anderson, 1989. Development of a Reservoir Water Quality Index DRAFT from Tennessee Valley Authority, Chattanooga, TN

Davic, Robert D and Jeffery E Deshon, 1988 The Ohio Lake Condition Index: A New Multiparameter Approach to Lake Classification State of Ohio Environmental Protection Agency, Division of Water Quality Monitoring Assessment

Dinius, S H , 1987. Design of an Index of Water Quality Water Resources Bulletin, 23(5):833-43

Entzminger, Tom. Region 8 WQI EPA Region 8, Denver

EPA Region 8, 1981. EPA Water Quality Trends, Region 8

Florida DER, 1988. A Guide to the Interpretation of Metal Concentrations in Estuarine Sediments Tallahassee, FL, April.

House, Margaret, 1989 Personal communication, July

House, M and J B Ellis, 1980 Water Quality Indices: An Additional Management Tool? Prog Wat Tech , 13, pp 413-23

House, M A and D H. Newsome, 1989. Water Quality Indices for the Management of Surface Water Quality Wat Sci Tech , 13, pp. 1137-1148

House, M A , 1990. A Water Quality Index for Use in the Operational Management of River Water Quality in Europe, to be published in IAWPRC Series: Advances in Water Pollution Series, Watershed '89: The Future of Water Quality in Europe, V2.

Landwehr, J , 1974 Water Quality Indices--Construction and Analysis Ph.D Dissertation.

Maryland Water Quality Inventory, 1985-1987 Volume II of the Section 305(b) Report.

McClelland, Nina I , Robert M Brown, Rolf A Deininger, and Jurate M Landwehr, 1973 Water Quality Index Application in the Kansas River Basin Presented at 46th Annual Conference, Water Pollution Control Federation, Cleveland, OH

McClelland, Nina I , Robert M Brown, and Rolf A Deininger, 1976 WQI Enhancement Appreciation of Quality Improvement Presented at American Chemical Society, New York, NY

Ott, Wayne R , 1978a Environmental Indices--Theory and Practice Ann Arbor Science Publishers, Inc

Ott, Wayne R , 1978b Water Quality Indices: A Survey of Indices Used in the United States U.S Environmental Protection Agency, Washington, D C Publication No EPA-600/4-78-005

Peterson, Ray and Bill Bogue, 1989 Water Quality Index (Used in Environmental Assessments) EPA Region 10, Seattle, Washington

Research Triangle Institute, 1989 Summary of State Use Support Methodologies. Prepared for EPA AWPD, April

Saylor, Charles and Edwin M Scott, Jr., 1987. Application of the Index of Biotic Integrity to Existing TVA Data Tennessee Valley Authority, Norris, TN.

Schaeffer, D.J. and K G. Janardan, 1977. Communicating Environmental Information to the Public: A New Water Quality Index Journal of Environmental Education, 8(4), Summer.

Tyson, J.M. and M A. House, 1989. The Application of a Water Quality Index to River Management. Wat. Sci. Tech., 21, pp. 1149-1159.

U S EPA, 1985. Technical Support Document for Water Quality-Based Toxics Control Office of Water, Washington, DC, September.

U S EPA, 1989a. Water Quality Program Highlights: Ohio EPA's Use of Biological Survey Information. Office of Water Regulations and Standards, May. Draft

U S. EPA, 1989b. Resource Document for Workshop on Environmental Indicators for Surface Water Program, Alexandria, VA, March.

U S EPA, 1989c. Guidelines for the Preparation of the 1990 State Water Quality Assessment (305(b) Report) Office of Water, Washington, DC, February.

# Appendix E

# APPENDIX E

# BIOLOGICAL METHODS

# APPENDIX E. BIOLOGICAL METHODS

Historically, most aquatic life criteria have been chemical-specific limits based on laboratory toxicity tests These methods have been heavily relied on to protect designated aquatic life uses, in part because proven measures of natural communities had not been developed In a sense, the chemical-specific results have been used as surrogates for measures of the actual biological health of aquatic communities However, many direct measures of community health have now been developed, tested and presented in national guidance. These biological methods surpass chemical-specific methods in certain areas, including assessing impacts of habitat modification, of complex toxic inputs, and of nontoxic pollutants (e g , nutrients and sediment from nonpoint sources).

*Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish* (Plafkin et al , 1989) presents tested procedures for conducting biosurveys and analyzing results These macrobenthos and fish methods measure community structure and detect impairment at various levels of certainty and technical effort. Other methods are available for periphyton, plankton, and macrophyton communities The advantages of using biological assessment methods in concert with physical/chemical methods are described in Section 2.3 and in EPA's draft *Policy on the Use of Biological Assessments and Criteria in the Water Quality Program* (the "draft Ecopolicy")

## E.1 PHYTOPLANKTON

Because of their position at the base of food chains, phytoplankton monitoring may be a very useful assessment tool, particularly in standing or slow-flowing waters Phytoplankton populations, under favorable growth conditions, may reach "bloom" conditions where they may cause taste and odor problems and extreme diurnal dissolved oxygen fluctuations leading to fish kills

## E.2 PERIPHYTON AND MACROPHYTON

The periphyton, those organisms attached to underwater substrates, are an assemblage of a wide variety of organisms The principal components are algae, fungi, bacteria and protozoa In the shallow waters of lakes, ponds, and streams, the periphyton are often the dominant primary producers Because they are fixed in place, they provide a time-integrated record of the quality of the water passing by More subtle changes, such as species composition, can indicate subtle changes in water quality

The diatoms, which are often the dominant algae of the periphyton, are particularly useful because there are many species and each can be identified The environmental requirements and pollution tolerance of several hundred species are documented Because of this, diatoms have been widely used for high sensitivity water quality monitoring. The silicon cell walls of diatoms preserve very well and make them useful for monitoring environmental changes through longer time periods (decades or centuries). Recent research on the effects of acid rain has used diatom data extensively. The primary disadvantage is the high level of expertise and time required to make the population counts and species identifications.

Macrophyton have been used much less than plankton and periphyton for water quality monitoring. The macrophyton community has many of the attributes of the periphyton community, but the diversity is much lower, consisting of only a few dozen species of vascular plants. Macrophytes have provided an estimate of habitat quality, particularly in the southern United States

## E.3 BENTHIC MACROINVERTEBRATES

Macroinvertebrate community measures include indicators of abundance, biomass, species composition, trophic (feeding) levels, pollution tolerances, and life history requirements of benthos. They are used extensively to characterize and evaluate the biological integrity of surface waters. The benthos are relatively sedentary, inhabit the bottom sediment and other benthic substrates for their life functions, and are sensitive to both long and short-term changes in the sediment and water quality of various aquatic ecosystems

Use of macroinvertebrate community measures reflects overall ecological integrity (chemical, physical, and biological), integrates the present stress and, over time, provides an ecological measure of fluctuating environmental conditions, and is relatively inexpensive. A limitation of macroinvertebrate methods is the uneven distribution of the organisms in the sediment and on substrates resulting in sampling data that may reflect more variability than is generally accepted with chemical data

Rapid bioassessment protocols (RBPs) have been developed for conducting biological assessments of lotic ecosystems and were designed as inexpensive screening and/or intensive tools for determining if a stream is or is not supporting a designated aquatic life use (Plafkin et al 1989) Rapid bioassessment techniques are gaining in popularity because they serve as cost-effective screening procedures for determining if a stream is or is not supporting the designated use These macroinvertebrate community measurement methods are described in methods manuals developed by both government and private standards-setting organizations. State biomonitoring programs specify the use of many of these techniques for regular surveillance activities Rapid bioassessment techniques are most useful in wadeable streams (streams sampled without a boat) Larger streams and rivers require other types of sampling equipment and more intensive resources.

## E.4 FISH

Fish are good indicators of long-term (several years) effects and broad habitat conditions because they are relatively long-lived and mobile. Fish community measures are appropriate direct measures of aquatic life resources

Several measures are employed to evaluate the biological integrity of rivers and streams. Top carnivore population size and structure can be used to relate directly to the "fishable" goal of the Clean Water Act. Relative abundances of juveniles, adults, and large adults are used to measure a site for its adequacy for rearing and reproduction and the potential impact of bioaccumulation Where piscivorous (fish-eating) birds, reptiles, and mammals are expected, the relative abundance of fish offers insights for further food chain effects and provides an important assessment of species commonly valued by the public. Fish are sensitive to over-harvest, acid deposition, contaminants, and physical habitat/flow modification.

Another fish community measure commonly used is fish gross pathology Stressors generally produce functional change in particular cells, tissues, and organs If the duration or intensity of stressors is sufficient, fish structural changes occur, followed by changes in fish populations and assemblages Typically, as the level of contaminants and pathogens rises, pathology increases

Eutrophication has been found to affect gills and fin structure, acid deposition has been found to affect sex organs, and contaminants affect the liver and gills One particular fish community measure used in one form or another by more than 30 States in the United States is the Index of Biotic Integrity (IBI) The IBI integrates several fish community measures such as top carnivore size and structure and fish pathology with others such as age class, number of tolerant species, number of native fish species, darter/benthic-intolerant species, omnivorous individuals, insectivorous (insect-eating) individuals, and a number of exotic/hybrid individuals

Fish community measures such as the IBI are good indicators of whether designated uses are being attained The IBI can provide a measure of health and complexity of a fish assemblage relative to those of a series of minimally impacted sites of similar size from the same ecological region Several advantages of using fish community measures have already been mentioned Others include: fish communities generally include a range of species that represent a variety of trophic levels and therefore tend to integrate effects of lower trophic levels and integrated environmental health; fish are at the top of the food chain and are consumed by humans, making them important subjects in assessing contamination; fish are easy to collect and identify; environmental requirements of common fish are comparatively well known, life history information is extensive for most species, and information on fish distributions is commonly available; aquatic life uses (water quality standards) are typically characterized in terms of fisheries (coldwater, warmwater, sport, forage); and fish account for nearly half of the endangered vertebrate species and subspecies in the United States.

The limitations of fish community measures are that they may be more difficult to obtain and use in large rivers, estuaries, and lakes/reservoirs. This is due mainly to the logistics of capturing fish and the fact that fish communities in large systems are the least understood. Coldwater fisheries such as trout fisheries are more difficult to evaluate because of their low species richness and lower relative abundance Variance is high in perturbed habitats, and temporal and spatial variability affects larger systems more than it affects smaller streams

Protocols for two fish community assessment techniques have been published by the USEPA in the aforementioned RBPs The use of the IBI is described in this document. The protocols also include examples and discussions of how various States are using fish community assessments in their monitoring programs

## E.5 HABITAT

Evaluations of habitat quality is critical to any assessment of biological integrity. The information needed is usually collected during biological surveys. Examples of physical habitat measures are: general land use and physical stream or lake/reservoir characteristics (width, depth, flow, and substrate for streams; average depth, residence time, and average surface acreage for lakes and reservoirs) Physical characterization starts with the riparian zone (stream bank or lake shore features, and drainage area) and proceeds in-lake or in-stream to sediment/substrate descriptions Such information can provide insight as to what organisms should be present or are expected to be present.

Predominant surrounding land use information is useful because of its potential effect on water quality Local watershed erosion, existing or potential, is characterized because detachment of soil from within the local watershed and its movement into a stream or lake will alter the physical habitat and reduce the expected species richness and abundance. Other nonpoint source pollution that affects habitat alteration, such as agricultural and urban runoff that contributes to factors other than siltation (feedlots, wetlands, septic systems, dams and impoundments, and/or mine seepage and tailings), is also identified and characterized.

Physical integrity or habitat is an important aspect of the abiotic component of any ecosystem and can be evaluated qualitatively or quantitatively It is well recognized by past and recent nonpoint source national assessments and State 305(b) reports that habitat alteration is the major cause for reduced or damaged ecological condition in streams, rivers, and lakes Guidance for evaluating use attainment in EPA's National Water Quality Standards program includes habitat assessment More recent policy development and guidance for establishing biocriteria for lakes and streams include a habitat assessment protocol (Platkin et al., 1989).

Habitat assessments of small and medium-sized streams and rivers are expected to be straightforward with few problems Large rivers and lakes and reservoirs will need research for the development of standardized assessment and quantification methods. Physical habitat alteration due to sedimentation, flow alteration, and channelization must be determined so that impacts to biological communities can be separated into those caused by discharges of contaminants and those related to physical habitat damage

Some modifications of existing habitat evaluation assessment methods may be necessary depending upon the predominant fish species present (coldwater vs warmwater) For example, assessments of trout fishery habitats would emphasize different substrate requirements than would assessments of warmwater fishery habitats for species such as bass or sunfish Along with these fishery differences, benthic community assemblages will also vary and, therefore, change the habitat characteristics emphasized in a physical habitat assessment Many States have already accounted for such considerations, and their techniques should be readily usable by States having similar fisheries and benthic communities

Several States are known to use habitat assessments to supplement macroinvertebrate and fish surveys The States of Arkansas, Maine, Missouri, New York, North Carolina, Ohio, and Virginia use some aspect of habitat evaluation to assist in their monitoring efforts Federal Agencies such as the Bureau of Recreation, U S Forest Service, Bureau of Land Management, and U S. Fish and Wildlife Service utilize habitat assessments to help sort out differences between water column effects and physical habitat alterations

## References


Ohio EPA (Environmental Protection Agency). 1987. Biological Criteria for the Protection of Aquatic Life: Volumes I, II, and III. The Role of Biological Data in Water Quality Assessment (Final Draft). User's Manual for Biological Assessment of Ohio Surface Waters. Standardized Biological Field Sampling and Laboratory Methods for Assessing Fish and Macroinvertebrate Communities Ohio EPA, Columbus, Ohio 43212.

Plafkin, J. L., M. T. Barbour, K. D. Porter, S. K. Gross, and R. M Hughes 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish. Assessment and Watershed Protection Division, Office of Water, U.S. Environmental Protection Agency, Washington, DC 20460. Publication No. EPA-444/4-89-001 May These protocols include a habitat assessment method.

# Appendix F

# APPENDIX F

# TOXICOLOGICAL METHODS

# APPENDIX F. TOXICOLOGICAL METHODS

## F.1 GENERAL

Toxicity tests are used to estimate one or more of the following: (1) the toxicity of an effluent collected at the end of the discharge pipe and tested with a standard dilution water; (2) the toxicity of effluents collected at the end of the discharge pipe and tested with dilution water consisting of nontoxic receiving water collected upstream or beyond the influence of the outfall, or with other uncontaminated surface water or standard dilution water having approximately the same hardness or salinity as the receiving water, depending on the nature of the receiving water (fresh or saline) and test organisms; (3) the toxicity of diluted effluent in the receiving water downstream or at increasing distance from the outfall; and (4) the effects of multiple discharges on the quality of the receiving water

## F.1 ACUTE TOXICITY TESTS

Acute toxicity tests are used to predict potential acute and chronic toxicity in the receiving water, based on the LC50 (lethal concentration to 50 percent of exposed organisms) or EC50 (effective concentration for death, immobilization, or other adverse effect for 50 percent of exposed organisms) and appropriate dilution, application, and persistence factors

Two types of acute toxicity tests, static and flowthrough, are available

Static tests include (1) nonrenewal tests in which the test organisms are exposed to the same test solution for the duration of the test, and (2) renewal tests in which the test organisms are exposed to a fresh solution of the same concentration of test solution every 24 hours or another prescribed interval, either by transferring the test organisms from one test chamber to another or by replacing all or a portion of the test solution in the test chambers The renewal system is preferred because interfering factors such as toxicant adsorption on the walls of the test chambers, volatilization, uptake by test organisms, and metabolism may affect toxicity

Two approaches to flow-through testing are available: (1) the test solution is pumped continuously from the sampling point directly to the diluter system; and (2) grab or composite samples are collected periodically, placed in a tank adjacent to the test laboratory, and pumped continuously from the tank to the diluter system

Standard acute toxicity tests involve the exposure of 20 test organisms to each of five toxicant concentrations and a control water. The test duration depends on the test species and ranges from a few hours for screening tests to 96 hours for definitive tests with fish and some invertebrates. The results of the test are reported as acute toxicity (LC50 or EC50) usually expressed as a percent, which is the concentration of test solution causing death (or immobilization, or other adverse effect) in 50 percent of the test organisms.

These methods are applicable to all freshwater and near coastal marine waters for assessing impacts from both point and nonpoint sources. In general, the advantages of acute test methods are that they are short, inexpensive, and easily reproducible In addition, they measure a direct endpoint, such as lethality or immobilization, which is easily determined by the investigator. A limitation of these tests is that they cannot be used directly to measure long-term impact of toxic chemicals or effluents. The single laboratory and multilaboratory precision of acute toxicity tests with several common test species and reference toxicants ranges from 38 to 50 percent.

## F.3 CHRONIC TOXICITY TESTS

The endpoints generally used in chronic tests are growth and reproduction. The effects include the synergistic, antagonistic, and additive effects of all the chemical, physical, and biological components that adversely affect the physiological and biochemical functions of the test organisms. The results of the short-term chronic toxicity tests are expressed as the NOEC, which is the highest percent concentration at which no adverse effect on survival, growth, or reproduction is observed.

Freshwater methods to measure the chronic toxicity of effluents and receiving waters to three freshwater species include an invertebrate, a plant, and a fish: the cladoceran, *Ceriodaphnia dubia*; the alga, *Selenastrum capricornutum*; and the fathead minnow, *Pimephales promelas* The methods are 4- to 7-day tests using the various test endpoints observed

Methods to measure the chronic toxicity of effluents and receiving waters to five marine species include two fish species, two invertebrates and an alga: the sheepshead minnow, *Cyprinodon variegatus*, and the inland silverside, *Menidia beryllina*; the mysid, *Mysidopsis bahia*; the sea urchin, *Arbacia punctulata*; and the red macroalga, *Champia parvula*. The methods are 1-hour to 9-day tests using the various test endpoints as described by Weber and Peltier (1988)

These methods are applicable to all freshwater and near coastal marine waters for assessing impacts from both point and nonpoint sources. In general, the advantages of these chronic methods are that they are shorter than most life-cycle tests while demonstrating good correlations of sensitivity with the longer-cycle tests proposed earlier In addition, the tests are cost-effective and are easily reproducible A limitation of these tests is that they are short-term estimates for predicting chronic toxicity that need to be checked periodically as to their sensitivity, compared to longer term, full life-cycle tests. These methods can be used to determine receiving water impacts from toxic effluents and evaluate single-chemical toxicity for developing water quality criteria. The sensitivity of the tests depends in part on the number of replicates, the probability level selected, and the type of statistical analysis chosen. Precision depends on the dilution factor chosen

## References

U S EPA, 1989 Guidelines Establishing Test Procedures for the Analysis of Pollutants Under the Clean Water Act Proposed rule. Deliverable No 0749[A]/EMSL-Cincinnati No. 1141, "Report Promulgate Rapid Chronic Toxicity Tests in Section 304(h) 40 CFR 136" Environmental Monitoring System Laboratory, Cincinnati, OH

Weber, C I., W B Horning, II, D J. Klemm, T W Nieheisel, P A Lewis, E L Robinson, J R. Menkedick, and F A Kessler (eds.). 1988. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms Environmental Monitoring and Support Laboratory, U.S Environmental Protection Agency, Cincinnati, Ohio Publication No EPA/600/4-87/028

Weber, C.I , W H Peltier, T J. Norberg-King, W B Horning, II, F A. Kessler, J.R. Menkedick, T.W Neiheisel, P A Lewis, D J Klemm, Q.H. Pickering, E.L. Robinson, J M Lazorchak, L.J Wymer, and R W Freyberg. 1989 Short-term Methods for Estimating the Chronic Toxicity of Effluents and Surface Waters to Freshwater Organisms, Second Edition Environmental Monitoring Systems Laboratory, U S Environmental Protection Agency, Cincinnati, Ohio Publication No. EPA/600/4-89/001

# Appendix G

# APPENDIX G

# REMOTE SENSING

# APPENDIX G. REMOTE SENSING

Remote sensing technology consists of aerial photography as well as airborne and satellite digital (multispectral scanners, or MSS) sensors. Photography is typically used to map spills, features associated with landfills and hazardous waste sites, and wetland delineations. MSS have been used to assess quality; to detect changes over time; to map algal, macrophyte, intertidal, and wetland species and extent; to detect thermal plumes; and to record point and nonpoint runoff in lakes and rivers throughout the United States Products for management applications include aerial photography prints, digital image map products, and geographical information systems (GIS) data bases

Aerial cameras or MSS are appropriate for (1) mapping large areas, (2) obtaining emergency or simultaneous views of an area, (3) developing complete aerial coverage rather than point data, (4) developing a digital data sets, and (5) mapping areas that are inaccessible or remote.

Several measures of water quality and character may be acquired using remote sensing devices Because light penetration in water is directly related to the amount of suspended material in the water column, measures of turbidity and clarity may be made using remote sensors If a relationship between sediments and other parameters of concern can be established, then sediment concentration may be used as a surrogate measure for such parameters as conductivity, nutrients and chemicals In waters with low turbidity, passive remote sensing can be used to characterize subsurface features Under optimal conditions, dense stands of submerged macrophytes and shallow water (< 10 m) bottom types can be delineated. Additionally, remote sensing devices may give a synoptic view of the watershed Such a view is particularly useful because water quality problems seldom originate at the shoreline or within a waterbody itself

The advantages of remote sensing techniques are: they provide synoptic views of a resource; they permit data for large areas to be acquired simultaneously or over a very short period of time; and they cost little in comparison to intensive field investigations. Perhaps the major question with using remote sensing for use support determinations is whether MSS data can be correlated to a particular measurement of interest Several studies have demonstrated that MSS data correlate directly with water clarity However, for other factors such as dissolved organic compounds, pH, and conductivity, direct measurements cannot be made by MSS and indirect correlations alone may or may not adequately represent the measurement of interest Regardless of the features of being assessed by cameras or MSS, some field data must be collected to define the accuracy of the final product.

# APPENDIX H

# USE OF FLOW DIAGRAMS

## APPENDIX H. USE OF FLOW DIAGRAMS

One of the major deficiencies in national reporting for the 1988 §305(b) reports was that it was very difficult to follow the use support decision process through the data analysis and interpretation steps. One simple method for improving reporting and assisting in structuring the decision process is to incorporate flow charts or decision trees in the reporting process.

A schematic of an idealized use support decision process for protecting aquatic life is presented in Figure H-1. For each waterbody, the process could be framed as follows:

1 "Direct" measures of use support are evaluated Has a fish kill occurred? Are sublethal impacts known to have occurred (e g , fish/ shellfish disease, noxious algal blooms, etc.)? If such events are sufficiently documented, a preliminary determination (before review of ambient monitoring information) that a waterbody is not fully supporting could be reached Ambient biological monitoring data are reviewed, and tentative use support determinations are made

2 All other recent "indirect" measures of use support are reviewed. If data are available, this step could include:

- Reviewing water column data
- Calculating a water quality index value
- Determining if fish or sediments are contaminated
- Assessing habitat integrity
- Toxicity data analysis.

"Weights" are assigned to the different measures of water quality within these categories. This assigning requires implicit or explicit judgments of the importance of, for example, toxic vs. conventional pollutants, and bioassay results relative to pollutant-specific results Indices might be used for aggregating data into an overall measure of water quality. Decisions on how to integrate direct and indirect use support measures are made

3 Lacking recent or sufficient monitoring data, "evaluated" information is used to make the use support decision

This flowchart (Figure H-1), shows that a single use support decision could require a complex series of decisions based on analyses of diverse data sets

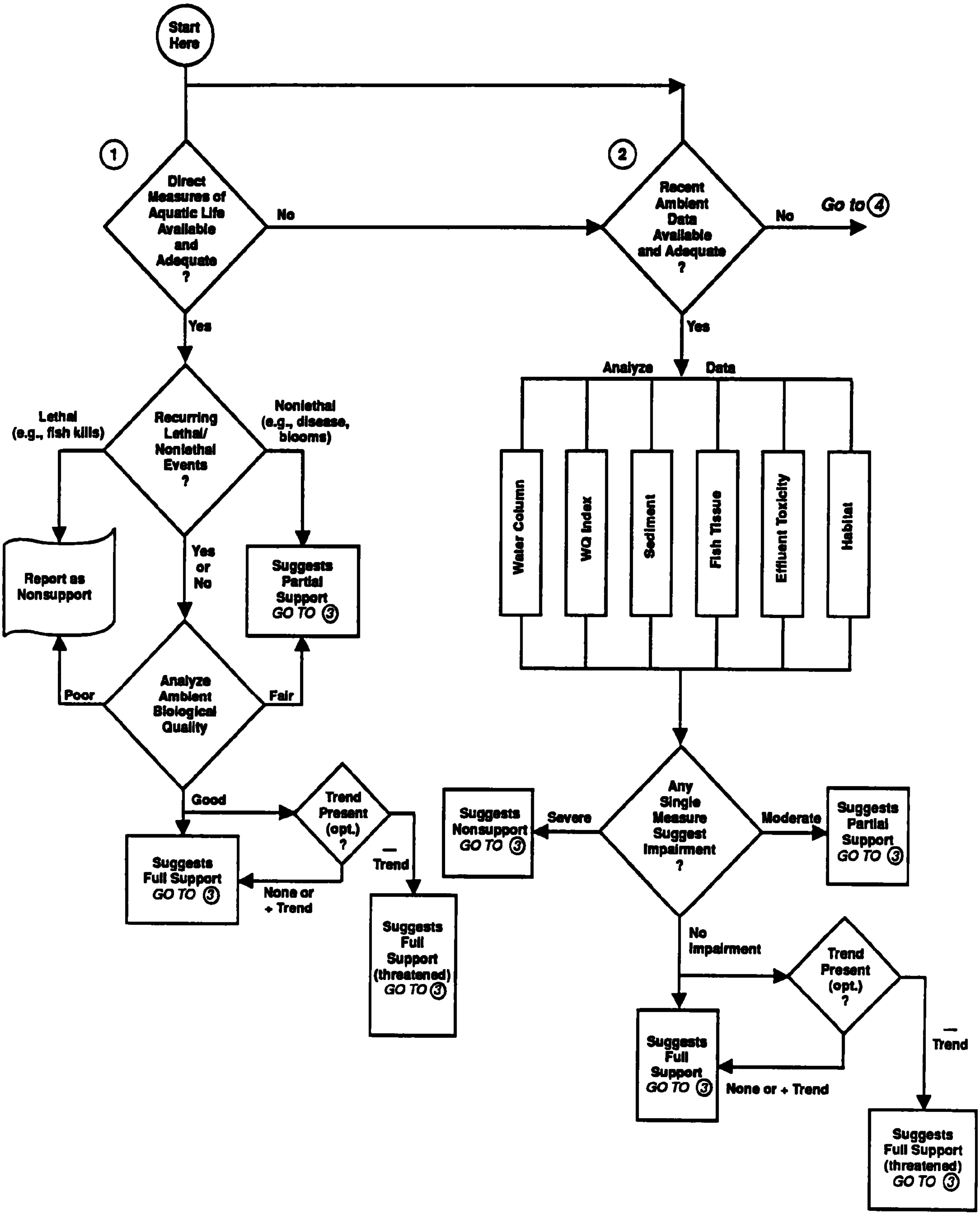



Figure H-1. Hypothetical use support decision process for aquatic life uses.

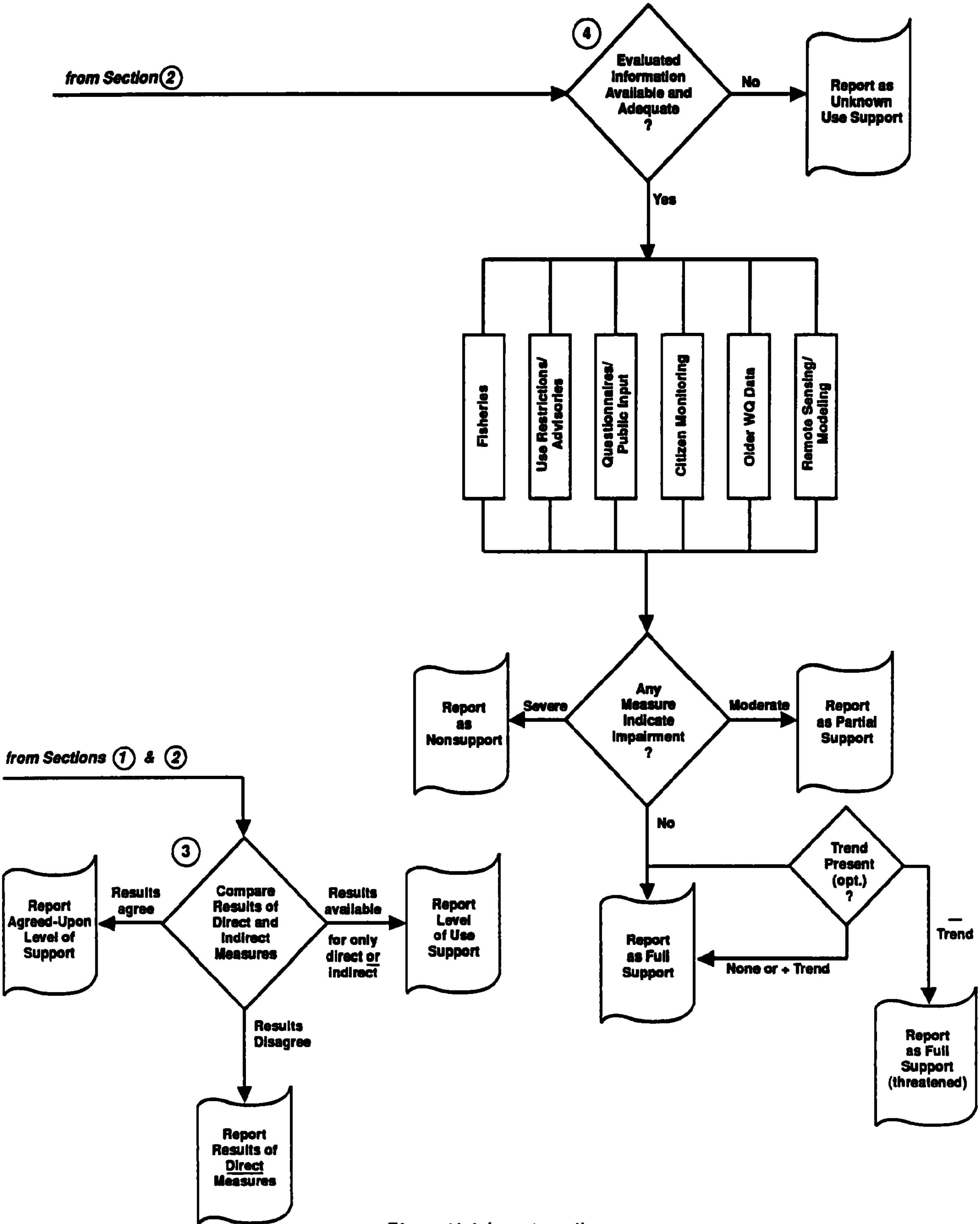



Figure H-1 (continued).

Also at issue is whether or not different methods should be developed for different uses. Using the flow chart as a starting point, a hypothetical State could develop a list of indicators to use for use support determinations (see Figure H-2).

**I. Aquatic Life Survival and Propagation**

A. Direct Measures
- Fish kills
- Disease and stress indicators
- Algal blooms

B Comprehensive measures (biological, chemical, physical)

C. Community Measures
- Benthic invertebrates
- Fish
- Submerged aquatic vegetation

D. Habitat Measures
- Sedimentation
- Flow alterations

E. Toxics Measures
- Toxicity
- Tissue contamination
- Sediment contamination
- 304(l) lists

F Chemical and Physical Measures
- Indices
- Parameter-specific methods
  - Toxicant assessments
  - Nontoxicant assessments (DO, nutrients, etc.)

G Evaluated Measures
- Public and professional opinion (questionnaires, hearings, correspondence)
- 319 lists

**II. Public Water Supply**

A. Drinking Water (Raw)
- Direct evaluations
  - Water system closures
  - Waterborne disease
  - Algal blooms
- Pathogen analysis
- Toxicant criteria analysis
- Aesthetic measures
  - Color
  - Turbidity/suspended solids
  - Taste and odor

**III. Contact Recreation**

A. Recreation Closures

B. Pathogen Measurements

C. Aesthetic Measures
- Turbidity/suspended solids
- Algal blooms
- Oil and grease
- Scums and floating debris

**IV. Fish and Shellfish Consumption**

A. Closures

B Tissue contamination

C Pathogens

**V. Outstanding Resource Waters**

A Water Quality Measures Based on Other Uses

B Resource Value Measures for Identified Resource
- Fishery production
- Water-based recreational use

**Figure H-2. Examples of methods for determining designated use support.**