United States Environmental Protection Agency
Office of Water 4304
EPA 822-B-00-013 December 2000

Ambient Water Quality Criteria Recommendations

Information Supporting the Development of State and Tribal Nutrient Criteria

Lakes and Reservoirs in Nutrient Ecoregion XII

EPA 822-B-00-013

AMBIENT WATER QUALITY CRITERIA RECOMMENDATIONS

INFORMATION SUPPORTING THE DEVELOPMENT OF STATE AND TRIBAL NUTRIENT CRITERIA

FOR

LAKES AND RESERVOIRS IN NUTRIENT ECOREGION XII

Southern Coastal Plains

including all or parts of the States of:

Alabama, Florida, Mississippi, Georgia

and the authorized Tribes within the Ecoregion

U.S. ENVIRONMENTAL PROTECTION AGENCY

OFFICE OF WATER
OFFICE OF SCIENCE AND TECHNOLOGY
HEALTH AND ECOLOGICAL CRITERIA DIVISION
WASHINGTON, D.C.

DECEMBER 2000

FOREWORD

This document presents EPA's nutrient criteria for **Lakes and Reservoirs in Nutrient Ecoregion XII.** These criteria provide EPA's recommendations to States and authorized Tribes for use in establishing their water quality standards consistent with section 303(c) of CWA. Under section 303(c) of the CWA, States and authorized Tribes have the primary responsibility for adopting water quality standards as State or Tribal law or regulation. The standards must contain scientifically defensible water quality criteria that are protective of designated uses. EPA's recommended section 304(a) criteria are not laws or regulations – they are guidance that States and Tribes may use as a starting point for the criteria for their water quality standards.

The term "water quality criteria" is used in two sections of the Clean Water Act, Section 304(a)(1) and Section 303(c)(2). The term has a different impact in each section. In Section 304, the term represents a scientific assessment of ecological and human health effects that EPA recommends to States and authorized Tribes for establishing water quality standards that ultimately provide a basis for controlling discharges or releases of pollutants or related parameters. Ambient water quality criteria associated with specific waterbody uses when adopted as State or Tribal water quality standards under Section 303 define the level of a pollutant (or, in the case of nutrients, a condition) necessary to protect designated uses in ambient waters. Quantified water quality criteria contained within State or Tribal water quality standards are essential to a water quality-based approach to pollution control. Whether expressed as numeric criteria or quantified translations of narrative criteria within State or Tribal water quality standards, quantified criteria serve as a critical basis for assessing attainment of designated uses and measuring progress toward meeting the water quality goals of the Clean Water Act.

EPA is developing section 304(a) water quality criteria for nutrients because States and Tribes consistently identify excessive levels of nutrients as a major reason why as much as half of the surface waters surveyed in this country do not meet water quality objectives, such as full support of aquatic life. EPA expects to develop nutrient criteria that cover four major types of waterbodies – lakes and reservoirs, rivers and streams, estuarine and coastal areas, and wetlands – across fourteen major ecoregions of the United States. EPA's section 304(a) criteria are intended to provide for the protection and propagation of aquatic life and recreation. To support the development of nutrient criteria, EPA is publishing Technical Guidance Manuals that describe a process for assessing nutrient conditions in the four waterbody types.

EPA's section 304(a) water quality criteria for nutrients provide numeric water quality criteria, as well as procedures by which to translate narrative criteria within State or Tribal water quality standards. In the case of nutrients, EPA section 304(a) criteria establish values for causal variables (e.g., total nitrogen and total phosphorus) and response variables (e.g., turbidity and

chlorophyll *a*). EPA believes that State and Tribal water quality standards need to include quantified endpoints for causal and response variables to provide sufficient protection of uses and to maintain downstream uses. These quantified endpoints will most often be expressed as numeric water quality criteria or as procedures to translate a State or Tribal narrative criterion into a quantified endpoint.

EPA will work with States and authorized Tribes as they adopt water quality criteria for nutrients into their water quality standards. EPA recognizes that States and authorized Tribes require flexibility in adopting numeric nutrient criteria into State and Tribal water quality standards. States and authorized Tribes have several options available to them. EPA recommends the following approaches, in order of preference:

(1) Wherever possible, develop nutrient criteria that fully reflect localized conditions and protect specific designated uses using the process described in EPA's Technical Guidance Manuals for nutrient criteria development. Such criteria may be expressed either as numeric criteria or as procedures to translate a State or Tribal narrative criterion into a quantified endpoint in State or Tribal water quality standards.

(2) Adopt EPA's section 304(a) water quality criteria for nutrients, either as numeric criteria or as procedures to translate a State or Tribal narrative nutrient criterion into a quantified endpoint.

(3) Develop nutrient criteria protective of designated uses using other scientifically defensible methods and appropriate water quality data.

Geoffrey H. Grubbs, Director
Office of Science and Technology

DISCLAIMER

This document provides technical guidance and recommendations to States, authorized Tribes, and other authorized jurisdictions to develop water quality criteria and water quality standards under the Clean Water Act (CWA) to protect against the adverse effects of nutrient overenrichment. Under the CWA, States and authorized Tribes are to establish water quality criteria to protect designated uses. State and Tribal decision-makers retain the discretion to adopt approaches on a case-by-case basis that differ from this guidance when appropriate and scientifically defensible. While this document contains EPA's scientific recommendations regarding ambient concentrations of nutrients that protect aquatic resource quality, it does not substitute for the CWA or EPA regulations; nor is it a regulation itself. Thus it cannot impose legally binding requirements on EPA, States, authorized Tribes, or the regulated community, and it might not apply to a particular situation or circumstance. EPA may change this guidance in the future.

EXECUTIVE SUMMARY

Nutrient Program Goals

EPA developed the National Strategy for the Development of Regional Nutrient Criteria (National Strategy) in June 1998. The strategy presents EPA=s intentions to develop technical guidance manuals for four types of waters (lakes and reservoirs, rivers and streams, estuaries and coastal waters, and wetlands) and produce section 304(a) criteria for specific nutrient ecoregions by the end of 2000. In addition, the Agency formed Regional Technical Assistance Groups (RTAGs) which include State and Tribal representatives working to develop more refined and more localized nutrient criteria based on approaches described in the waterbody guidance manuals. This document presents EPA=s current recommended criteria for total phosphorus, total nitrogen, chlorophyll *a*, and turbidity for lakes and reservoirs in Nutrient Ecoregion XII (Southern Coastal Plain) which were derived using the procedures described in the Lakes and Reservoirs Nutrient Criteria Technical Guidance Manual (U.S. EPA, 2000a).

EPA's ecoregional nutrient criteria are intended to address cultural eutrophication-- the adverse effects of excess nutrient inputs. The criteria are empirically derived to represent conditions of surface waters that are minimally impacted by human activities and protective of aquatic life and recreational uses. The information contained in this document represent starting points for States and Tribes to develop (with assistance from EPA) more refined nutrient criteria.

In developing these criteria recommendations, EPA followed a process which included, to the extent they were readily available, the following elements critical to criterion derivation:

- **Historical and recent nutrient data in Nutrient Ecoregion XII**
  Data from Legacy STORET were used to assess nutrient conditions from 1990 to 1998. A number of scientific publications were useful in classifying Florida lakes within Ecoregion 12 including work by Brezonik and Shannon (1971. OWRR Project No. B-004-FLA, Publ. No. 13, Water Research center, University of Florida) and Canfield (1983. Water Resources Bull. 19:255-262).

- **Reference sites/reference conditions in Nutrient Ecoregion XII**
  Reference sites/reference conditions in Nutrient Ecoregion XII were based on the lake population distribution approach using a representative sample of all lakes within the

Ecoregion (see Nutrient Criteria Technical Guidance Manual-Lakes and Reservoirs, April 2000, EPA-822-B00-001**.** The Aggregate Ecoregion and Level 3 ecoregion values are similar because this Ecoregion contains only one sub-ecoregion. The State and Tribes are urged to determine their own reference sites for lakes and reservoirs within the ecoregion at different geographic scales and to compare them to EPA's reference conditions.

! **Models employed for prediction or validation**
EPA did not identify any specific models used in the ecoregion to develop nutrient criteria. States and Tribes are encouraged to identify and apply appropriate models to support nutrient criteria development.

! **RTAG expert review and consensus**
EPA recommends that when States and Tribes prepare their nutrient criteria, they obtain the expert review and consent of the RTAG.

! **Downstream effects of criteria**
EPA encourages the RTAG to assess the potential effects of the proposed criteria on downstream water quality and uses.

In addition, EPA followed specific **QA/QC procedures** during data collection and analysis: All data were reviewed for duplications. All data are from ambient waters that were not located directly outside a permitted discharger. Florida and Georgia indicated that their data were sampled and analyzed using either Standard methods or EPA approved methods.

The following tables contain a summary of Aggregate and level III ecoregion values for TN, TP, water column chl *a*, and turbidity:

BASED ON 25th PERCENTILES ONLY

Nutrient Parameters	Aggregate Nutrient Ecoregion XII Reference Conditions (same as subecoregion 75)
Total phosphorus (µg/L)	10.0
Total nitrogen (mg/L)	0.52
Chlorophyll *a* (µg/L) (Spectrophotometric method)	2.6
Secchi depth (meters)	2.1

NOTICE OF DOCUMENT AVAILABILITY

This document is available electronically to the public through the INTERNET at: (http://www.epa.gov/OST/standards/nutrient.html). Requests for hard copies of the document should be made to EPA's National Service Center for Environmental Publications (NSCEP), 11029 Kenwood Road, Cincinnati, OH 45242 or (513) 489-8190 or toll free (800) 490-9198. Please refer to EPA document number **EPA-822-B-00-013.**

ACKNOWLEDGMENTS

The authors thankfully acknowledge the contributions of the following State and Federal reviewers: EPA 4; the States of Florida, Georgia, Alabama and Mississippi; the Tribes within the Ecoregion; EPA Headquarters personnel from the Office of Wetlands, Oceans and Watersheds, Office of Wastewater Management, Office of General Counsel, Office of Research and Development, and the Office of Science and Technology. EPA also acknowledges the external peer review efforts of Eugene Welch (University of Washington), Robert Carlson (Kent State University), Steve Heiskary (Minnesota Pollution Control Agency), Greg Denton and Sherry Wang (Tennessee Department of Environment and Conservation), and Gerhard Kuhn (U.S. Geological Survey).

LISTS OF TABLES AND FIGURES

Figures

Tables

TABLE OF CONTENTS

Page

1.0 INTRODUCTION

Background

Nutrients are essential to the health and diversity of our surface waters. However, in excessive amounts, nutrients cause hypereutrophication, which results in overgrowth of plant life and decline of the biological community. Excessive nutrients can also result in potential human health risks, such as the growth of harmful algal blooms - most recently manifested in the *Pfiesteria* outbreaks of the Gulf and East Coasts. Chronic nutrient overenrichment of a waterbody can lead to the following consequences: low dissolved oxygen, fish kills, algal blooms, overabundance of macrophytes, likely increased sediment accumulation rates, and species shifts of both flora and fauna.

Historically, National Water Quality Inventories have repeatedly shown that nutrients are a major cause of ambient water quality use impairments. EPA's 1996 National Water Quality Inventory report identifies excessive nutrients as the leading cause of impairment in lakes and the second leading cause of impairment in rivers (behind siltation). In addition, nutrients were the second leading cause of impairments reported by the States in their 1998 lists of impaired waters. Where use impairment is documented, nutrients contribute roughly 25-50% of the impairment nationally. The Clean Water Act establishes a national goal to achieve, wherever attainable, water quality which provides for the protection and propagation of fish, shellfish, and wildlife and recreation in and on the water. In adopting water quality standards, States and Tribes designate uses for their waters in consideration of the Clean Water Act goals, and establish water quality criteria that contain sufficient parameters to protect those uses. To date, EPA has not published information and recommendations under section 304(a) for nutrients to assist States and Tribes in establishing numeric nutrient criteria to protect uses when adopting water quality standards.

In 1995, EPA gathered a set of national experts and asked the experts how to best deal with the national nutrient problem. The experts recommended that the Agency not develop single criteria values for phosphorus or nitrogen applicable to all water bodies and regions of the country. Rather, the experts recommended that EPA put a premium on regionalization, develop guidance (assessment tools and control measures) for specific waterbodies and ecological regions across the country, and use reference conditions (conditions that reflect pristine or minimally impacted waters) as a basis for developing nutrient criteria.

With these suggestions as starting points, EPA developed the National Strategy for the Development of Regional Nutrient Criteria (National Strategy), published in June 1998. This strategy presented EPA's intentions to develop technical guidance manuals for four types of waters (lakes and reservoirs, rivers and streams, estuaries and coastal waters, and wetlands) and, thereafter, to publish section 304(a) criteria recommendations for specific nutrient ecoregions. Technical guidance manuals for lakes/reservoirs and rivers/streams were published in April 2000 and July 2000, respectively. The technical guidance manual for estuaries/coastal waters will be published in spring 2000 and the draft wetlands technical guidance manual will be published by

December 2001. Each manual presents EPA's recommended approach for developing nutrient criteria values for a specific waterbody type. In addition, EPA is committed to working with States and Tribes to develop more refined and more localized nutrient criteria based on approaches described in the waterbody guidance manuals and this document.

Overview of the Nutrient Criteria Development Process

For each Nutrient Ecoregion, EPA developed a set of recommendations for two causal variables (total nitrogen and total phosphorus) and two early indicator response variables (chlorophyll *a* and some measure of turbidity). Other indicators such as dissolved oxygen and macrophyte growth or speciation, and other fauna and flora changes are also deemed useful. However, the first four are considered to be the best suited for protecting designated uses.

The technical guidance manuals describe a process for developing nutrient criteria that involves consideration of five factors. The first of these is the Regional Technical Assistance Group (RTAG), which is a body of qualified regional specialists able to objectively evaluate all of the available evidence and select the value(s) appropriate to nutrient control in the water bodies of concern. These specialists may come from such disciplines as limnology, biology, natural resources management-- especially water resource management, chemistry, and ecology. The RTAG evaluates and recommends appropriate classification techniques for criteria determination, usually physical within an ecoregional construct.

The second factor is the historical information available to establish a perspective of the resource base. This is usually data and anecdotal information available within the past ten-twenty five years. This information gives evidence about the background and enrichment trend of the resource.

The third factor is the present reference condition. A selection of reference sites chosen to represent the least culturally impacted waters of the class existing at the present time. The data from these sites is combined and a value from the distribution of these observations is selected to represent the reference condition, or best attainable, most natural condition of the resource base at this time.

A fourth factor often employed is theoretical or empirical models of the historical and reference condition data to better understand the condition of the resource.

The RTAG comprehensively evaluates the other three elements to propose a candidate criterion (initially one each for TP, TN, chl *a*, and some measure of turbidity).

The last and final element of the criteria development process is the assessment by the RTAG of the likely downstream effects of the criterion. Will there be a negative, positive, or neutral effect on the downstream waterbody? If the RTAG judges that a negative effect is likely,

then the proposed State/Tribal water quality criteria should be revised to ameliorate the potential for any adverse downstream effects.

While States and authorized Tribes would not necessarily need to incorporate all five elements into their water quality criteria setting process (e.g., modeling may be significant in only some instances), the best assurance of a representative and effective criterion for nutrient management decision making is the balanced incorporation of all five elements, or at least all elements except modeling.

Because some parts of the country have naturally higher soil and parent material enrichment, and different precipitation regimes, the application of the criterion development process has to be adjusted by region. Therefore, an ecoregional approach was chosen to develop nutrient criteria appropriate to each of the different geographical and climatological areas of the country. Initially, the continental U.S. was divided into 14 separate ecoregions of similar geographical characteristics. Ecoregions are defined as regions of relative homogeneity in ecological systems; they depict areas within which the mosaic of ecosystem components (biotic and abiotic as well as terrestrial and aquatic) is different than adjacent areas in a holistic sense. Geographic phenomena such as soils, vegetation, climate, geology, land cover, and physiology that are associated with spatial differences in the quantity and quality of ecosystem components are relatively similar within each ecoregion.

The Nutrient ecoregions are aggregates of U.S. EPA=s hierarchal level III ecoregions. As such, they are more generalized and less defined than level III ecoregions. EPA determined that setting ecoregional criteria for the large scale aggregates is not without its drawbacks - variability is high due to the lumping of many waterbody classes, seasons, and years worth of multipurpose data over a large geographic area. For these reasons, the Agency recommends that States and Tribes develop nutrient criteria at the level III ecoregional scale and at the waterbody class scale where those data are readily available. Data analyses and recommendations on both the large aggregate ecoregion scale as well as more refined scales (level III ecoregions and waterbody classes), where data were available to make such assessments, are presented for comparison purposes and completeness of analysis.

Relationship of Nutrient Criteria to Biological Criteria

Biological criteria are quantitative expressions of the desired condition of the aquatic community. Such criteria can be based on an aggregation of data from sites that represent the least-impacted and attainable condition for a particular waterbody type in an ecoregion, subecoregion, or watershed. EPA's nutrient criteria recommendations and biological criteria recommendations have many similarities in the basic approach to their development and data requirements. Both are empirically derived from statistical analysis of field collected data and expert evaluation of current reference conditions and historical information. Both utilize direct measurements from the environment to integrate the effects of complex processes that vary

according to type and location of waterbody. The resulting criteria recommendations, in both cases, are efficient and holistic indicators of water quality necessary to protect uses.

States and authorized Tribes can develop and apply nutrient criteria and biological criteria in tandem, with each providing important and useful information to interpret both the nutrient enrichment levels and the biological condition of sampled waterbodies. For example, using the same reference sites for both types of criteria can lead to efficiencies in both sample design and data analysis. In one effort, environmental managers can obtain information to support assessment of biological and nutrient condition, either through evaluating existing data sets or through designing and conducting a common sampling program. The traditional biological criteria variables of benthic invertebrate and fish sampling can be readily incorporated to supplement a nutrient assessment. To demonstrate the effectiveness of this tandem approach, EPA has initiated pilot projects in both freshwater and marine environments to investigate the relationship between nutrient overenrichment and apparent declines in diversity indices of benthic invertebrates and fish.

2.0 BEST USE OF THIS INFORMATION

EPA recommendations published under section 304(a) of the CWA serve several purposes, including providing guidance to States and Tribes in adopting water quality standards for nutrients that ultimately provide a basis for controlling discharges or releases of pollutants. The recommendations also provide guidance to EPA when promulgating Federal water quality standards under section 303(c) when such action is necessary. Other uses include identification of overenrichment problems, management planning, project evaluation, and determination of status and trends of water resources.

State water quality inventories and listings of impaired waters consistently rank nutrient overenrichment as a top contributor to use impairments. EPA's water quality standards regulations at 40 CFR §131.11(a) require States and Tribes to adopt criteria that contain sufficient parameters and constituents to protect the designated uses of their waters. In addition, States and Tribes need quantifiable targets for nutrients in their standards to assess attainment of uses, develop water quality-based permit limits and source control plans, and establish targets for total maximum daily loads (TMDLs).

EPA expects States and Tribes to address nutrient overenrichment in their water quality standards, and to build on existing State and Tribal initiated efforts where possible. States and Tribes can address nutrient overenrichment through establishment of numerical criteria or through use of new or existing narrative criteria statements (e.g., free from excess nutrients that cause or contribute to undesirable or nuisance aquatic life or produce adverse physiological response in humans, animals, or plants). In the case of narrative criteria, EPA expects that States and Tribes establish procedures to quantitatively translate these statements for both assessment and source control purposes.

The intent of developing ecoregional nutrient criteria is to represent conditions of surface waters that are minimally impacted by human activities and thus protect against the adverse effects of nutrient overenrichment from cultural eutrophication. EPA's recommended process for developing such criteria includes physical classification of waterbodies, determination of current reference conditions, evaluation of historical data and other information (such as published literature), use of models to simulate physical and ecological processes or determine empirical relationships among causal and response variables (if necessary), expert judgement, and evaluation of downstream effects. To the extent allowed by the information available, EPA has used elements of this process to produce the information contained in this document. The values for both causal (total nitrogen, total phosphorus) and biological and physical response (chlorophyll *a*, turbidity) variables represent a set of starting points for States and Tribes to use in establishing their own criteria in standards to protect uses.

In its water quality standards regulations, EPA recommends that States and Tribes establish numerical criteria based on section 304(a) guidance, section 304(a) guidance modified to reflect site-specific conditions, or other scientifically defensible methods. For many pollutants, such as toxic chemicals, EPA expects that section 304(a) guidance will provide an appropriate level of protection without further modification in most cases. EPA has also published methods for modifying 304(a) criteria on a site-specific basis, such as the water effect ratio, where site-specific conditions warrant modification to achieve the intended level of protection. For nutrients, however, EPA expects that, in most cases, it will be necessary for States and authorized Tribes to identify with greater precision the nutrient levels that protect aquatic life and recreational uses. This can be achieved through development of criteria modified to reflect conditions at a smaller geographic scale than an ecoregion such as a subecoregion, the State or Tribe level, or specific class of waterbodies. Criteria refinement can occur by grouping data or performing data analyses at these smaller geographic scales. Refinement can also occur through further consideration of other elements of criteria development, such as published literature or models.

The values presented in this document generally represent nutrient levels that protect against the adverse effects of nutrient overenrichment and are based on information available to the Agency at the time of this publication. However, States and Tribes should critically evaluate this information in light of the specific designated uses that need to be protected. For example, more sensitive uses may require more stringent values as criteria to ensure adequate protection. On the other hand, overly stringent levels of protection against the adverse effects of cultural eutrophication may actually fall below levels that represent the natural load of nutrients for certain waterbodies. In cases such as these, the level of nutrients specified may not be sufficient to support a productive fishery. In the criteria derivation process, it is important to distinguish between the natural load associated with a specific waterbody and current reference conditions, using historical data and expert judgement. These elements of the nutrient criteria derivation process are best addressed by States and Tribes with access to information and local expertise. Therefore, EPA strongly encourages States and Tribes to use the information contained in this document and to develop more refined criteria according to the methods described in EPA's technical guidance manuals for specific waterbody types.

To assist in the process of further refinement of nutrient criteria, EPA has established ten Regional Technical Advisory Groups (experts from EPA Regional Offices and States/Tribes). In the process of refining criteria, States and authorized Tribes need to provide documentation of data and analyses, along with a defensible rationale, for any new or revised nutrient criteria they submit to EPA for review and approval. As part of EPA's review of State and Tribal standards, EPA intends to seek assurance from the RTAG that proposed criteria are sufficient to protect uses.

In the process of using the information and recommendations contained in this document, as well as additional information, to develop numerical criteria or procedures to translate narrative criteria, EPA encourages States and Tribes to:

- Address both chemical causal variables and early indicator response variables. Causal variables are necessary to provide sufficient protection of uses before impairment occurs and to maintain downstream uses. Early response variables are necessary to provide warning signs of possible impairment and to integrate the effects of variable and potentially unmeasured nutrient loads.
- Include variables that can be measured to determine if standards are met, and variables that can be related to the ultimate sources of excess nutrients.
- Identify appropriate periods of duration (i.e., how long) and frequency (i.e., how often) of occurrence in addition to magnitude (i.e., how much). EPA does not recommend identifying nutrient concentrations that must be met at all times, rather a seasonal or annual averaging period (e.g., based on weekly measurements) is considered appropriate. However, these seasonal or annual central tendency measures should apply each season or each year, except under the most extraordinary of conditions (e.g., a 100 year flood).

3.0 AREA COVERED BY THIS DOCUMENT

The following sections provide a general description of the aggregate ecoregion and its geographical boundaries. Descriptions of the level III ecoregions contained within the aggregate ecoregion are also provided.

3.1 Description of Aggregate Ecoregion XII - Southern Coastal Plain

The hot, low-lying Southern Coastal Plain contains concentrations of swamps, marshes, and lakes. The region is nearly level; it has more lakes than the neighboring Southeastern Temperate Forested Plains and Hills (IX) or the Southern Florida Coastal Plain (XIII) and it is flatter than Region IX. It is underlain by limestone and has a sandy mantle of varying thicknesses. Sand hills reach over 200 feet elevation and are nutrient-poor. Karst topography occurs and is particularly extensive in the Big Bend area from eastern Wakulla County south to Pasco County. Thousands of lakes dot the region and have varying trophic states; they are far more numerous than in adjacent nutrient regions. Woodlands, forests, citrus orchards, vegetable

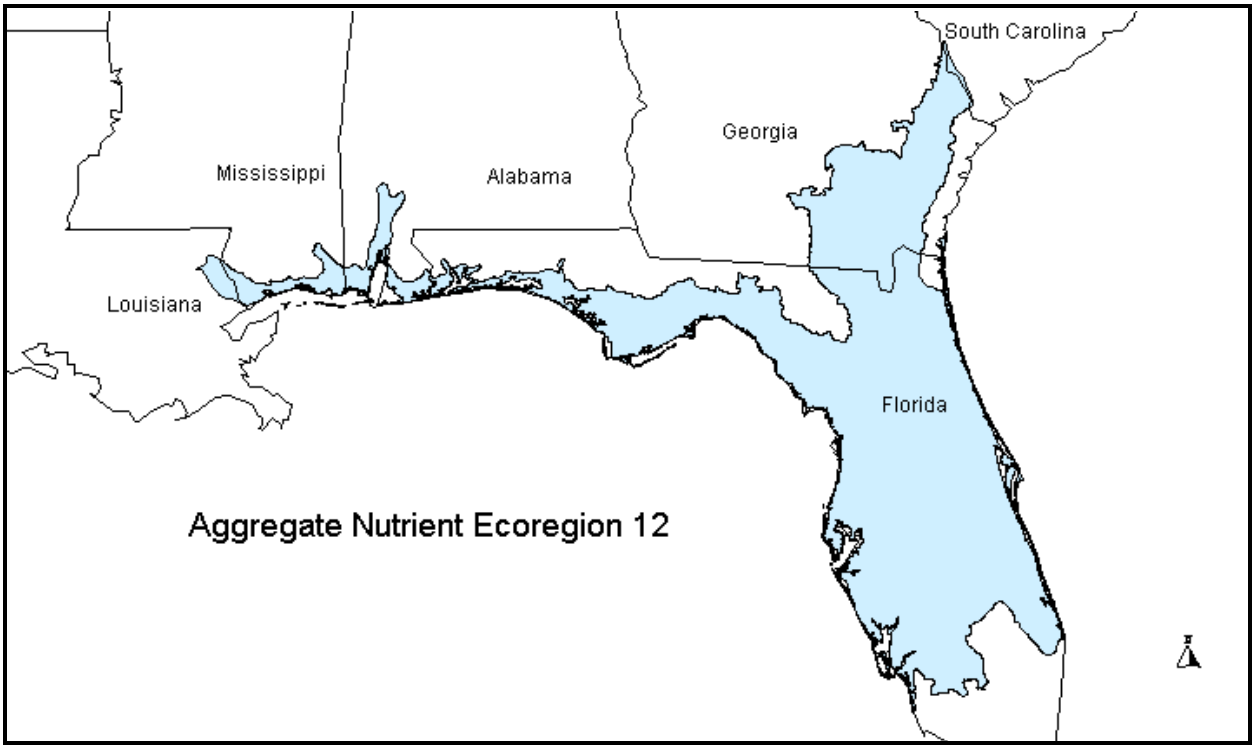

Figure 1. Aggregate Ecoregion XII.

farming, and pastures dominate most of the region. Locally, urban and suburban areas are common and have grown rapidly in the last fifty years. Surficial water quality has been significantly affected by human activity including urban development, industry, agriculture, silviculture, water management activity, and mining. Stream nutrient levels have been increased by runoff from sewage treatment plants, weathered rock, phosphate mines, fertilizer plants, citrus orchards, and other farms. Dissolved oxygen concentrations have been lowered by the effects of nutrient enrichment. Suspended sediment has been added to streams by agriculture and logging. Lake quality varies throughout the region. Highest median total phosphorus levels and total nitrogen concentrations are in the lakes of the southwest.

3.2 Geographic Boundaries of Aggregate Ecoregion XII

Ecoregion XII encompasses the southeast corner of Geogia (excluding the immediate coastline) and a large segment of central and Gulf of Mexico coastal Florida.

3.3 Level III Ecoregions Within Aggregate Ecoregion XII

75. *Southern Coastal Plain*

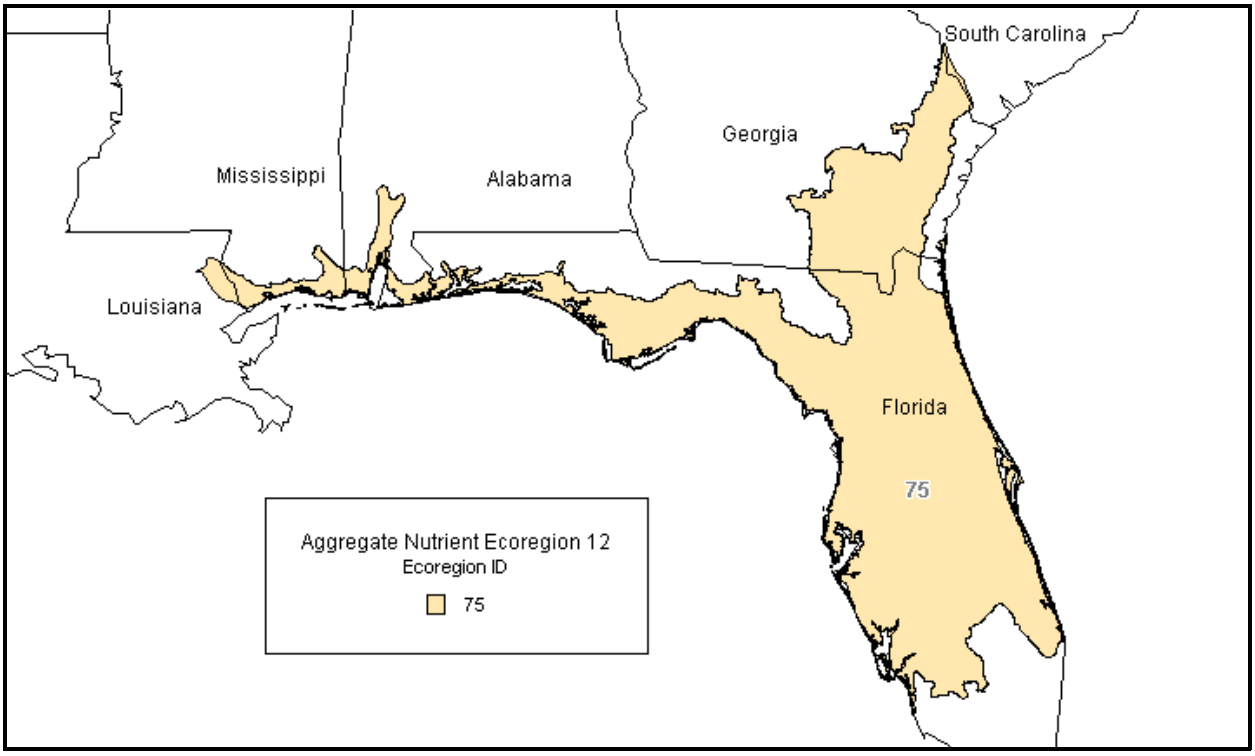

Figure 2. Aggregate Ecoregion XII with level III Sub-ecoregion shown. Note this Ecoregion contains only one sub-ecoregion (Sub-ecoregion 75).

This Aggregate Ecoregion XII consists only of a single Sub-Ecoregion (Sub-Ecoregion 75). The Southern Coastal Plain consists of mostly flat plains with numerous swamps, marshes and lakes. This ecoregion is warmer, more heterogeneous, and has a longer growing season and coarser textured soils than the Middle Atlantic Coastal Plain. Once covered by a forest of beech, sweetgum, southern magnolia, slash pine, loblolly pine, white oak, and laurel oak, land cover in the region is now mostly longleaf-slash pine forest, oak-gum-cypress forest in some low lying areas, pasture for beef cattle, and urban development.

Suggested Ecoregional subdivisions or adjustments.

The following sections provide a general description of the aggregate ecoregion and its geographical boundaries. Descriptions of the level III ecoregions contained within the aggregate ecoregion are also provided.

4.0 DATA REVIEW FOR LAKES AND RESERVOIRS IN AGGREGATE ECOREGION XII

The following section describes the nutrient data EPA has collected and analyzed for this

Ecoregion, including an assessment of data quantity and quality. The data tables present the data for each causal parameter-- total phosphorus and total nitrogen (both reported and calculated from TKN and nitrite/nitrate), and the primary response variables-- some measure of turbidity and chlorophyll *a*. These are the parameters which EPA considers essential to nutrient assessment because the first two are the main causative agents of enrichment and the two response variables are the early indicators of system enrichment for most of the surface waters
(See Chapter 5 of the Lakes and Reservoirs Nutrient Criteria Guidance Manual (U.S. EPA, 2000a) for a complete discussion on choosing causal and response variables.)

4.1 Data Sources

Data sets from Legacy STORET were used to assess nutrient conditions from 1990 to 1999. EPA recommends that the RTAGs identify additional data sources that can be used to supplement the data sets listed above. In addition, the RTAGs may utilize published literature values to support quantitative and qualitative analyses.

4.2 Historical Data from Aggregate Ecoregion XII (TP, TN, Chlor a and Secchi Depth)

The long-term trend over the past 50 years has been for major urbanization to occur in this ecoregion with attendant high potential for increases in nutrient enrichment. Limestone (Karst) topography over much of the area has resulted in a relative high frequency for groundwater contamination by nitrates. Nonpoint sources of nutrients continue to dominate nutrient loads in a number of waterbodies. Historical nutrient data are consistent with a trend for increased water quality-based impairments. To gain additional perspective on more recent trends, it is recommended that States and Tribes assess nutrient trends over the last 10 years (e.g., what do seasonal trends indicate?)

4.3 QA/QC of data sources

An initial quality screen of data was conducted using the rules presented in Appendix C. Data remaining after screening for duplications and other QA measures (e.g., poor or unreported analytical records, sampling errors or omissions, stations associated with outfalls, storm water sewers, hazardous waste sites) were the data used in the statistical analyses.

States within Ecoregion XII were contacted regarding the quality of their data. The following States provided information on the methods used to sample and analyze their waters: Florida and Georgia. Other States in the Ecoregion did not provide information at the time this document was published.

4.4 Data for all Lakes/Reservoirs within Aggregate Ecoregion XII

The map in Figure 3 shows the location of the sampling stations within Ecoregion XII

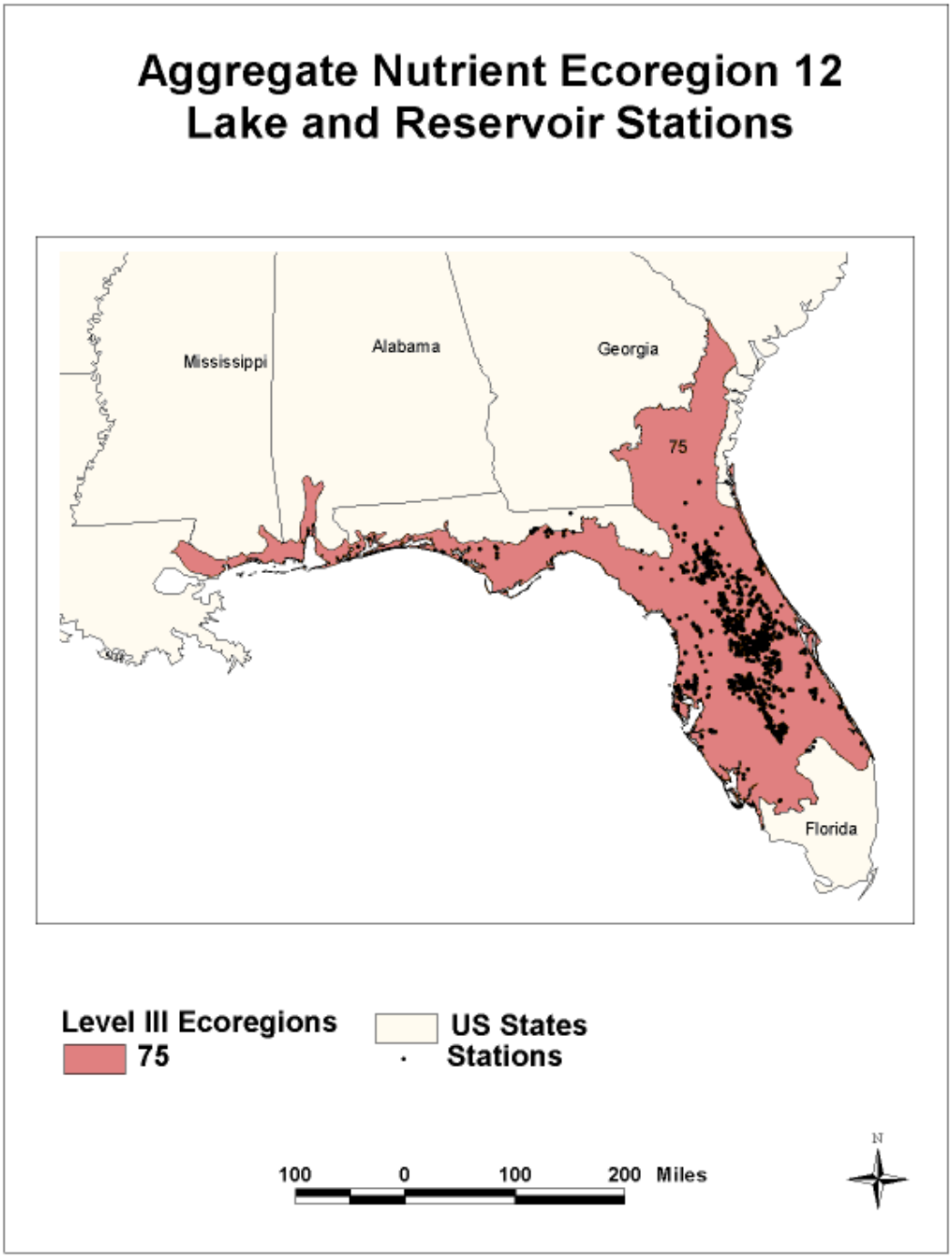

Figure 3 Sampling locations within each level III ecoregion

(Sub-Ecoregion 75). Most of the sampling took place in the centeral Florida peninsula where the bulk of lakes occur. Lake and reservoir sampling within central Florida appears to be reasonably representative of lake/reservoir distribution with some clustering around metropolitan areas. The

relatively small portion of Ecoregion XII located in the State of Georgia appears to be somewhat under-sampled. Table 1 presents all data records for Sub-Ecoregion 75.

.

Table 1. Lake records for Aggregate Ecoregion XII - Southeastern Coastal Plains

	Aggregate Ecoregion XII	Sub ecoR 75
# of Lake Systems	943	943
# of Lake Stations	2869	2869
Key Nutrient Parameters (listed below)		
- # of records for Secchi depth	53,524	53,524
- # of records for Chlorophyll *a* (all methods)	62,008	62,008
- # of records for Total Kjeldahl Nitrogen (TKN)	15,043	15,043
- # of records for Nitrate + Nitrite (NO_2 + NO_3)	9,038	9,038
- # of records for Total Nitrogen (TN)	49,492	49,492
- # of records for Total Phosphorus (TP)	58,974	58,974
Total # of records for key nutrient parameters	248,079	248,079

Definitions used to complete Table 1:

1. # of records refers to the total count of observations for that parameter over the entire decade (1990-1999) for that particular aggregate or subecoregion. These are counts for all seasons over that decade.

2. # of lake stations refers to the total number of lake and reservoir stations within the aggregate or subecoregion from which nutrient data were collected. Since lakes and reservoirs can cross ecoregional boundaries, it is important to note that only those portions of a lake or reservoir (and data associated with those stations) that exist within the ecoregion are included within this table.

4.5 Statistical Analysis of Data

EPA's Technical Guidance Manual for Developing Nutrient Criteria for Lakes and Reservoirs describes two ways of establishing a reference condition. One method is to choose the upper 25th percentile (75th percentile) of a reference population of lakes. This is the preferred method to establish a reference condition. The 75th percentile was chosen by EPA since it is likely associated with minimally impacted conditions, will be protective of designated uses, and provides management flexibility. When reference lakes are not identified, the second method is to determine the lower 25th percentile of the population of all lakes within a region. The 25th percentile of the entire population was chosen by EPA to represent a surrogate for an actual reference population. Data analyses to date indicate that the lower 25th percentile from an entire population roughly approximates the 75th percentile for a reference population (see case studies for Minnesota lakes in the Lakes and Reservoirs Nutrient Criteria Technical Guidance Document [U.S. EPA, 2000a], the case study for Tennessee streams in the Rivers and Streams Nutrient Criteria Technical Guidance Document [U.S. EPA, 2000b], and the letter from Tennessee Department of Environment and Conservation to Geoffrey Grubbs [TNDEC, 2000]). New York State has also presented evidence that the 25th percentile and the 75th percentile compare well based on user perceptions of water resources (NYSDEC, 2000).

The following tables 2 and 3a, present the potential reference conditions for both the aggregate ecoregion and the subecoregions using both methods. However, the reference lake column is left blank because EPA does not have observed reference data and anticipates that States will provide information on reference lakes. Appendix A provides a complete presentation of all descriptive statistics for both the aggregate ecoregion and the level III subecoregion.

Table 2. Reference conditions for aggregate ecoregion XII lakes.

Parameter	No. of Lakes N ++	Reported values		25th Percentiles based on all seasons data for the Decade	Reference Lakes **
		Min	Max	P25* all seasons +	P75 all seasons
TKN (mg/L)	401	0.04	5.09	**0.6**	
NO2 + NO3 (mg/L)	261	0.00	1.28	**0.002**	
TN (mg/L) - calculated	NA	0.04	6.37	**0.60**	
TN (mg/L) - reported	545	0.07	4.73	**0.52**	
TP (*u*g/L)	692	0.00	980.0	**10.0**	
Secchi (meters)	688	0.02	5.6	**2.1**	
Chlorophyll *a* (*u*g/L) - F	--	--	--	**--**	
Chlorophyll *a* (*u*g/L) - S	205	0.0	135.0	**2.6**	
Chlorophyll *a* (*u*g/L) - T	649	0.3	146.0	**3.5**	

P25: 25th percentile of all data
P75: 75th percentile of all data
** as determined by the Regional Technical Assistance Groups (RTAGs)
+ Median for all seasons' 25th percentiles. E.g. this value was calculated from four seasons' 25th percentiles. If the seasonal 25th percentile (P25) TP values are - spring 10*u*g/L, summer 15*u*g/L, fall 12*u*g/L, and winter 5*u*g/L, the median value of all seasons P25 will be 11*u*g/L.
++ N = largest value reported for a decade / Season.
TN calculated is based on the sum of TKN + NO_2+NO_3.
TN reported is actual TN value reported in the database for one sample.
F Chlorophyll *a* measured by Fluorometric method with acid correction.
S Chlorophyll *a* measured by Spectrophotometric method with acid correction.
T Chlorophyll *a b c* measured by Trichromatic method.
NA Not Applicable

Table 3a present the potential reference conditions for lakes and reservoirs in the Level III subecoregions within the Aggregate Ecoregion. The footnotes for Table 2 apply to table 3a.

Table 3a. Reference conditions for level III ecoregion 75 lakes.

Parameter	No. of Lakes N ++	Reported values		25th Percentiles based on all seasons data for the Decade	Reference Lakes **
		Min	Max	P25* all seasons +	P75 all seasons
TKN (mg/L)	401	0.04	5.09	0.6	
NO2 + NO3 (mg/L)	261	0.00	1.28	0.002	
TN (mg/L) - calculated	NA	0.04	6.37	0.60	
TN (mg/L) - reported	545	0.07	4.73	0.52	
TP (*u*g/L)	692	0.00	980.0	10.0	
Secchi (meters)	688	0.02	5.6	2.1	
Chlorophyll *a* (*u*g/L) - F	--	--	--	--	
Chlorophyll *a* (*u*g/L) - S	205	0.0	135.0	2.6	
Chlorophyll *a* (*u*g/L) - T	649	0.3	146.0	3.5	

Definitions used in filling Tables 2 and 3 - Reference Condition tables

1. Number of Lakes in Table 2 refers to the largest number of lakes and reservoirs for which data existed for a given season within an aggregate nutrient ecoregion.

2. Number of Lakes in Table 3 refers to the number of lakes and reservoirs for which data existed for the summer months since summer is generally when the greatest amount of nutrient sampling is conducted. If another season greatly predominates, notification is made (s=spring, f=fall, w=winter).

3. Medians. All values (min, max, and 25th percentiles) included in the table are based on waterbody medians. All data for a particular parameter within a lake for the decade were reduced to one median for that lake. This prevents over-representation of individual waterbodies with a great deal of data versus those with fewer data points within the statistical analysis.

4. 25th percentile for all seasons is calculated by taking the median of the 4 seasonal 25th percentiles. If a season is missing, the median was calculated with 3 seasons of data. If less than 3 seasons were used to derive the median, the entry is flagged (z).

5. A 25th percentile for a season is best derived with data from a minimum of 4 lakes/season. However, this table provides 25th percentiles that were derived with less than 4 lakes/season in order to retain all information for all seasons. In calculating the 25th percentile for a season with less than 4 lake medians, the statistical program automatically used the minimum value within the less-than-4 population. If less than 4 lakes were used in developing a seasonal quartile and or all-seasons median, the entry is flagged (zz).

Preferred Data Choices and Recommendations When Data Are Missing

1. Where data are missing or are very low in total records for a given parameter, use 25th percentiles for parameters within an adjacent, similar subecoregion within the same aggregate nutrient ecoregion or when a similar subecoregion can not be determined, use the the 25th percentile for the Aggregate ecoregion or consider the lowest 25th percentile from a subecoregion (level III) within the aggregate nutrient ecoregion. The rationale being that without data, one may assume that the subecoregion in question may be as sensitive as the most sensitive subecoregion within the aggregate.

2. TN calculated: When reported Total Nitrogen (TN) median values are lacking or very low in comparison to TKN and Nitrate/Nitrite-N values, the medians for TKN and nitrite/nitrate-N were added, resulting in a calculated TN value. The number of samples (N) for calculated TN is not filled in since it is represented by two subsamples of data: TKN and nitrite/nitrate-N. Therefore,

N/A is placed in this box.

3. TN reported: This is the median based on reported values for TN from the database.

4. Chlorophyll *a*: Medians based on all methods are reported, however, the acid corrected medians are preferred to the uncorrected medians. In developing a reference condition from a particular method, it is recommended that the method with the most observations be used. Fluorometric and Spectrophotometric are preferred over all other methods. However, when no data exist for Fluorometric and Spectrophotometric methods, Trichromatic values may be used. Data from the variance techniques are not interchangeable.

5. Periphyton: Where periphyton data exist, record them separately For periphyton-dominated streams, a measure of periphyton chlorophyll is a more appropriate response variable than planktonic chlorophyll *a*. See Table 4, p. 101 of the Rivers and Streams Nutrient Technical Guidance Manual (U. S. EPA, 2000b) for values of periphyton and planktonic chlorophyll *a* related to eutrophy in streams.

6. Secchi depth: The 75th percentile is reported for Secchi depth since this is the only variable for which the value of the parameter increases with greater clarity. (For lakes and reservoirs only.)

7. Turbidity units: All turbidity units from all methods are reported. FTUs and NTUs are preferred over JCUs. If FTUs and NTUs do not exist, use JCUs. These units are not interchangeable. Turbidity is chosen as a response variable in streams since it can be an indicator of increasing algal biomass due to nutrient enrichment. See pages 32 -33 of the Rivers and Streams Nutrient Technical Guidance Manual for a discussion of turbidity and correlations with algal growth.

8. Lack of data: A dash (-) represents missing, inadequate, or inconclusive data. A zero (0) is reported if the reported median for a parameter is 0 or if the component value is below detection.

4.6. Classification of Lake/Reservoir Type

It is anticipated that assessing the data by lake type will further reduce the variability in the data analysis. There were no readily available classification data in the National datasets used to develop these criteria. States and Tribes are strongly encouraged to classify their lakes before developing a final criterion.

4.7. Summary of Data Reduction Methods

All descriptive statistics were calculated using the medians for each lake within ecoregion XII, for which data existed. For example, if one lake had 300 observations for phosphorus over the decade or one year's time, one median resulted. Each median from each lake was then used in calculating the percentiles for phosphorus for the aggregate nutrient ecoregion/subecoregion (level III ecoregion) by season and year (Figure 4a & b).

Observations for All Lakes/Reservoirs

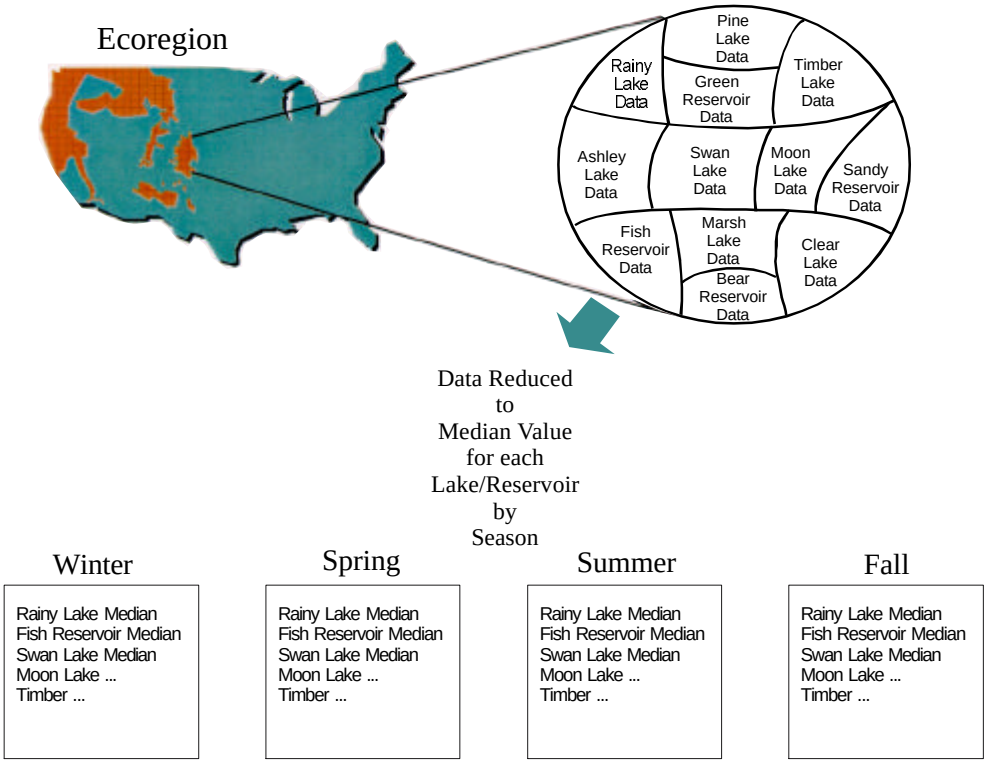

Figure 4a. Illustration of data reduction process for lake data.

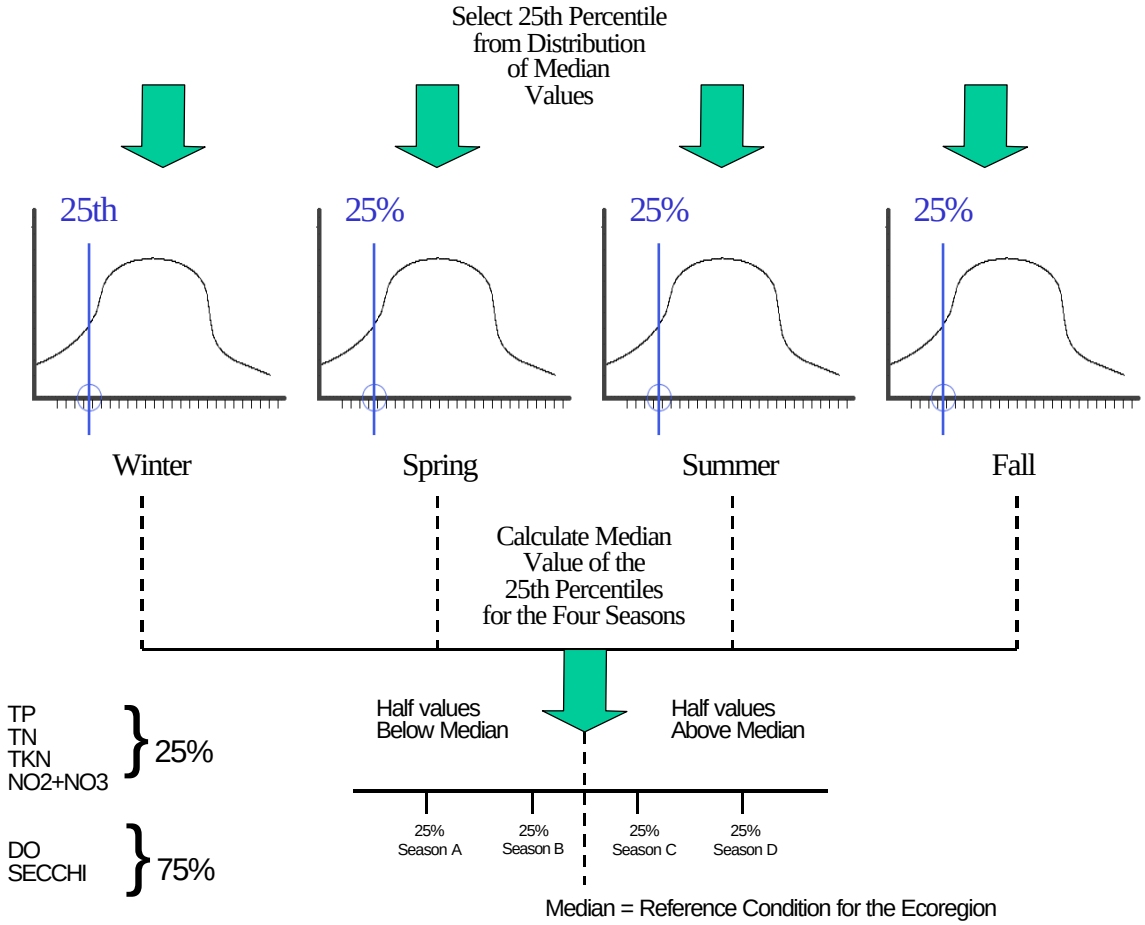

Figure 4b. Illustration of reference condition calculation.

5.0 REFERENCE SITES AND CONDITIONS IN AGGREGATE ECOREGION XII

Reference conditions represent the natural, least impacted conditions or what is considered to be the most attainable conditions. This section compares the different reference conditions determined from the two methods and establishes which reference condition is most appropriate.

A priori determination of reference sites. The preferred method for establishing reference condition is to choose the upper percentile of an a priori population of reference lakes. States and Tribes are encouraged to identify reference conditions based on this method.

Statistical determination of reference conditions (25th percentile of entire database.) See Tables 2 and 3a in section 4.0.

RTAG discussion and rationale for selection of reference sites and conditions in Ecoregion II. The RTAG should compare the results derived from the two methods described above and present a rationale for the final selection of reference sites.

6.0 MODELS USED TO PREDICT OR VERIFY RESPONSE PARAMETERS

The RTAG is encouraged to identify and apply relevant models to support nutrient criteria development. The following are three scenarios under which models may be used to derive criteria or support criteria development.

- Models for predicting correlations between causal and response variables
- Models used to verify reference conditions based on percentiles
- Regression models used to predict reference conditions in impacted areas

7.0 FRAMEWORK FOR REFINING RECOMMENDED NUTRIENT CRITERIA FOR LAKES AND RESERVOIRS IN AGGREGATE ECOREGION XII

Information on each of the following six weight of evidence factors is important to refine the criteria presented in this document. All elements should be addressed in developing criteria, as is expressed in our nutrient criteria technical guidance manuals. It is our expectation that EPA Regions, States, and Tribes (as RTAGs) will consider these elements as States/Tribes develop their criteria. This section should be viewed as a work sheet (sections are left blank for this purpose) to assist in the refinement of nutrient criteria. If many of these elements are ultimately unaddressed, EPA may rely on the proposed reference conditions presented in Table 3a and other literature and

information readily available to the HQ nutrient team to develop nutrient water quality recommendations for this ecoregion.

7.1 Example Worksheet for Developing Aggregate Ecoregion and Subecoregion Nutrient Criteria

- *Literature sources*

- *Historical data and trends*

- *Reference condition*

- *Models*

- *RTAG expert review and consensus*

- *Downstream effects*

7.2 Tables of Refined Nutrient Water Quality Criteria for Aggregate Ecoregion II and Level III Subecoregions for TP, TN, Chl a, Turbidity (where sufficient data exist)

Aggregate Ecoregion XII-Southern Coastal Plain	**Proposed Criterion**
Total Phosphorus (µg/L)	
Total Nitrogen (mg/L)	
Chlorophyll a (µg/L or mg/m^2)	
Secchi depth (meters)	
Other (Index; other parameter such as DO)	

7.3 Setting Seasonal Criteria

The criteria presented in this document are based in part on medians of all the 25^{th} percentile seasonal data (decadal), and as such are reflective of all seasons and not one particular season or year. It is recommended that States and Tribes monitor in all seasons to best assess compliance with the resulting criterion. States/Tribes may choose to develop criteria which reflect **each** particular season or a **given year** when there is significant variability between seasons/years or designated uses that are specifically tied to one or more seasons of the year (e.g., recreation, fishing). Using the tables in Appendix A and B, one can set reference conditions based on a particular season or year and then develop a criterion based on each individual season. Obviously, this option is season-specific and would also require increased monitoring within each season to assess compliance.

7.4 When Data/Reference Conditions are Lacking

When data are unavailable to develop a reference condition for a particular parameter(s) within a subecoregion, EPA recommends one of three options: 1. Use data from a similar neighboring subecoregion. E.g., If data are few or nonexistent for the northern cascades, consider using the data and reference condition developed for the cascades; or 2. Use the 25^{th} perecentiles for the Aggregate ecoregion or 3. Consider using the lowest of the yearly medians for that parameter calculated for all the subecoregions within the Aggregate Ecoregion.

7.5 Site-specific Criteria Development

Criteria may be refined in a number of ways. The best way to refine criteria is to follow the

critical elements of criteria development as well as to refer to the Lakes and Reservoirs technical guidance manual (U. S. EPA, 2000a).

The Lakes and Reservoirs Nutrient Criteria Technical Guidance Manual presents sections on each of the following factors to consider in setting criteria

- refinements to ecoregions (Chapter 3)
- classification of waterbodies (Chapter 3)
- setting seasonal criteria to reflect major seasonal climate differences (Chapter 7)
- accounting for significant or cyclical rainfall events - high flow/low flow conditions (Chapter 7)
- setting criteria for reservoirs only (The technical guidance manual recommends that data be separated for lakes and reservoirs and treated independently if possible because of differing physical conditions that occur in lakes and reservoirs. In this document all data from both reservoirs and lakes were considered together since STORET does not allow for the differentiation of data except by waterbody name.)

8.0 LITERATURE CITED

NYSDEC (New York State Department of Environment and Conservation). 2000. Memorandum from Scott Kishbaugh to Jay Bloomfield, September 26, 2000, regarding reference lakes for nutrient criteria.

TNDEC (Tennessee Department of Environment and Conservation). 2000. Letter to Geoff Grubbs, October 5, 2000, containing comments on draft nutrient criteria recommendations.

U.S. EPA. 2000a. Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs, U.S. Environmental Protection Agency, Washington, DC. EPA-822-B00-001.

U.S. EPA. 2000b. Nutrient Criteria Technical Guidance Manual: Rivers and Streams, U.S. Environmental Protection Agency, Washington, DC. EPA-822-B00-002.

9.0 APPENDICES

A. Descriptive Statistics Data Tables for Aggregate Ecoregion

B. Descriptive Statistics Data Tables for Level III Subecoregions within Aggregate Ecoregion

C. Quality Control/Quality Assurance Rules

APPENDIX A

Descriptive Statistics Data Tables for Aggregate Ecoregion

Aggregate Nutrient Ecoregion: XII

1

Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter Chla_Phyto_Spec_A_ug_L_Median

SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
FALL	180	23.4	.000	155.60	31.0	2.31	133	0.97	3.39	8.80	32.4	104
SPRING	205	18.8	.000	119.30	25.7	1.80	137	0.48	1.87	6.46	26.1	72.2
SUMMER	179	20.8	.000	150.60	29.2	2.18	140	0.94	3.48	7.67	25.0	97.3
WINTER	125	16.6	.000	112.20	24.2	2.16	146	0.30	1.63	5.45	20.9	62.6

Aggregate Nutrient Ecoregion: XII

2

Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter Chla_Tric_U_ug_L_Median

SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
FALL	646	19.1	.685	127.43	22.2	0.87	116	2.00	4.50	10.0	26.3	68.0
SPRING	652	14.9	.000	140.55	21.2	0.83	142	1.00	3.00	6.71	17.9	58.0
SUMMER	649	16.8	.250	151.50	20.8	0.81	123	1.75	4.00	8.25	21.4	57.2
WINTER	604	15.2	.360	151.55	19.9	0.81	131	1.00	3.00	7.00	20.0	52.0

Aggregate Nutrient Ecoregion: XII

3

Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter DO_mg_L_Median

SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
FALL	30	6.94	2.40	10.50	1.50	0.27	22	4.30	6.20	7.05	7.80	9.30
SPRING	34	8.68	6.00	11.80	1.28	0.22	15	6.65	7.80	8.58	9.40	11.4
SUMMER	35	6.93	4.00	12.70	1.59	0.27	23	4.65	6.00	7.00	7.40	10.6
WINTER	24	8.39	6.10	10.10	1.03	0.21	12	6.50	7.75	8.45	9.20	9.40

Aggregate Nutrient Ecoregion: XII

4

Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter NO2_NO3_mg_L_Median

SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
FALL	262	0.02	.000	0.41	0.04	0.00	230	0.00	0.00	0.01	0.02	0.08
SPRING	259	0.03	.000	0.76	0.07	0.00	252	0.00	0.00	0.01	0.02	0.10

SUMMER	261	0.02	.000	1.81	0.11	0.01	520	0.00	0.00	0.01	0.01	0.06
WINTER	197	0.15	.000	5.21	0.62	0.04	411	0.00	0.01	0.02	0.06	0.44

Aggregate Nutrient Ecoregion: XII

5

Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter SECCHI_m_Median

SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
FALL	681	1.53	.025	5.49	1.01	0.04	66	0.38	0.80	1.22	2.02	3.66
SPRING	684	1.60	.025	5.03	1.05	0.04	65	0.31	0.83	1.27	2.13	3.66
SUMMER	668	1.63	.025	6.13	1.08	0.04	66	0.38	0.85	1.30	2.13	3.81
WINTER	627	1.68	.076	5.70	1.10	0.04	65	0.46	0.91	1.31	2.24	3.96

Aggregate Nutrient Ecoregion: XII

6

Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter TKN_mg_L_Median

SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
FALL	389	1.03	.000	5.00	0.62	0.03	60	0.41	0.65	0.88	1.24	2.12
SPRING	398	1.04	.070	5.13	0.74	0.04	71	0.30	0.60	0.85	1.25	2.58
SUMMER	401	0.98	.000	5.15	0.64	0.03	65	0.29	0.60	0.85	1.19	2.14
WINTER	319	0.98	.100	5.06	0.66	0.04	68	0.38	0.60	0.80	1.14	2.29

Aggregate Nutrient Ecoregion: XII

7

Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter TN_mg_L_Median

SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
FALL	531	0.88	.070	4.86	0.56	0.02	64	0.24	0.54	0.76	1.09	1.82
SPRING	511	0.88	.080	4.94	0.63	0.03	72	0.24	0.51	0.70	1.04	2.11
SUMMER	545	0.87	.065	4.29	0.58	0.03	67	0.24	0.53	0.74	1.05	2.02
WINTER	524	0.88	.070	4.60	0.62	0.03	71	0.22	0.53	0.74	1.03	2.03

Aggregate Nutrient Ecoregion: XII

Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter TP_ug_L_Median

SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
FALL	686	43.9	1.50	1010.00	76.6	2.93	175	6.00	10.0	20.0	50.0	140
SPRING	695	45.2	.000	1100.00	82.0	3.11	181	5.00	10.0	20.0	47.5	170
SUMMER	692	39.2	.000	787.50	64.7	2.46	165	5.50	10.0	20.0	40.0	135
WINTER	638	41.9	.000	950.00	70.2	2.78	168	6.00	10.0	20.0	45.0	155

APPENDIX B

Descriptive Statistics Data Tables for Level III Subecoregions within Aggregate Ecoregion

Aggregate Nutrient Ecoregion: XII
Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter Chla_Phyto_Spec_A_ug_L_Median

Eco_Level_III	SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
75	FALL	180	23.4	.000	155.60	31.0	2.31	133	0.97	3.39	8.80	32.4	104
75	SPRING	205	18.8	.000	119.30	25.7	1.80	137	0.48	1.87	6.46	26.1	72.2
75	SUMMER	179	20.8	.000	150.60	29.2	2.18	140	0.94	3.48	7.67	25.0	97.3
75	WINTER	125	16.6	.000	112.20	24.2	2.16	146	0.30	1.63	5.45	20.9	62.6

Aggregate Nutrient Ecoregion: XII
Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter Chla_Tric_U_ug_L_Median

Eco_Level_III	SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
75	FALL	646	19.1	.685	127.43	22.2	0.87	116	2.00	4.50	10.0	26.3	68.0
75	SPRING	652	14.9	.000	140.55	21.2	0.83	142	1.00	3.00	6.71	17.9	58.0
75	SUMMER	649	16.8	.250	151.50	20.8	0.81	123	1.75	4.00	8.25	21.4	57.2
75	WINTER	604	15.2	.360	151.55	19.9	0.81	131	1.00	3.00	7.00	20.0	52.0

Aggregate Nutrient Ecoregion: XII
Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter DO_mg_L_Median

Eco_Level_III	SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
75	FALL	30	6.94	2.40	10.50	1.50	0.27	22	4.30	6.20	7.05	7.80	9.30
75	SPRING	34	8.68	6.00	11.80	1.28	0.22	15	6.65	7.80	8.58	9.40	11.4
75	SUMMER	35	6.93	4.00	12.70	1.59	0.27	23	4.65	6.00	7.00	7.40	10.6
75	WINTER	24	8.39	6.10	10.10	1.03	0.21	12	6.50	7.75	8.45	9.20	9.40

Aggregate Nutrient Ecoregion: XII
Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter NO2_NO3_mg_L_Median

Eco_Level_III	SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
75	FALL	262	0.02	.000	0.41	0.04	0.00	230	0.00	0.00	0.01	0.02	0.08
75	SPRING	259	0.03	.000	0.76	0.07	0.00	252	0.00	0.00	0.01	0.02	0.10

75	SUMMER	261	0.02	.000	1.81	0.11	0.01	520	0.00	0.00	0.01	0.01	0.06
75	WINTER	197	0.15	.000	5.21	0.62	0.04	411	0.00	0.01	0.02	0.06	0.44

Aggregate Nutrient Ecoregion: XII 5
Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter SECCHI_m_Median

Eco_ Level_ III	SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
75	FALL	681	1.53	.025	5.49	1.01	0.04	66	0.38	0.80	1.22	2.02	3.66
75	SPRING	684	1.60	.025	5.03	1.05	0.04	65	0.31	0.83	1.27	2.13	3.66
75	SUMMER	668	1.63	.025	6.13	1.08	0.04	66	0.38	0.85	1.30	2.13	3.81
75	WINTER	627	1.68	.076	5.70	1.10	0.04	65	0.46	0.91	1.31	2.24	3.96

Aggregate Nutrient Ecoregion: XII 6
Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter TKN_mg_L_Median

Eco_ Level_ III	SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
75	FALL	389	1.03	.000	5.00	0.62	0.03	60	0.41	0.65	0.88	1.24	2.12
75	SPRING	398	1.04	.070	5.13	0.74	0.04	71	0.30	0.60	0.85	1.25	2.58
75	SUMMER	401	0.98	.000	5.15	0.64	0.03	65	0.29	0.60	0.85	1.19	2.14
75	WINTER	319	0.98	.100	5.06	0.66	0.04	68	0.38	0.60	0.80	1.14	2.29

Aggregate Nutrient Ecoregion: XII 7
Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter TN_mg_L_Median

Eco_ Level_ III	SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
75	FALL	531	0.88	.070	4.86	0.56	0.02	64	0.24	0.54	0.76	1.09	1.82
75	SPRING	511	0.88	.080	4.94	0.63	0.03	72	0.24	0.51	0.70	1.04	2.11
75	SUMMER	545	0.87	.065	4.29	0.58	0.03	67	0.24	0.53	0.74	1.05	2.02
75	WINTER	524	0.88	.070	4.60	0.62	0.03	71	0.22	0.53	0.74	1.03	2.03

Aggregate Nutrient Ecoregion: XII
Lakes and Reservoirs
Descriptive Statistics by Decade and Season
Parameter TP_ug_L_Median

Eco_ Level_ III	SEASON	N	MEAN	MIN	MAX	STDDEV	STDERR	CV	P5	P25	MEDIAN	P75	P95
75	FALL	686	43.9	1.50	1010.00	76.6	2.93	175	6.00	10.0	20.0	50.0	140
75	SPRING	695	45.2	.000	1100.00	82.0	3.11	181	5.00	10.0	20.0	47.5	170
75	SUMMER	692	39.2	.000	787.50	64.7	2.46	165	5.50	10.0	20.0	40.0	135
75	WINTER	638	41.9	.000	950.00	70.2	2.78	168	6.00	10.0	20.0	45.0	155

APPENDIX C

Quality Control/Quality Assurance Rules

Support for the Compilation and Analysis of National Nutrient Data

15 Nutrient Ecoregion/Waterbody Type Summary Chapters

Prepared for:

Robert Cantilli
Environmental Protection Agency
OW/OST/HECD

Prepared by:

INDUS Corporation
1953 Gallows Road
Vienna, Virginia 22182

Contract Number: **68-C-99-226**
Task Number: **04**
Subtask Number: **4**

August 8, 2000

CONTENTS

1.0 BACKGROUND

The Nutrient Criteria Program has initiated development of a national Nutrient Criteria Database application that will be used to store and analyze nutrient data. The ultimate use of these data will be to derive ecoregion- and waterbody-specific nutrient criteria ranges. EPA converted STOrage and RETrieval (STORET) legacy data, National Stream Quality Accounting Network (NASQAN) data, National Water-Quality Assessment (NAWQA) data, and other relevant nutrient data from universities and States/Tribes into the database. The data imported into the Nutrient Criteria Database will be used to develop national nutrient criteria ranges.

1.1 Purpose

The purpose of this deliverable is to provide EPA with information regarding the data used to create the statistical reports which will be used to derive ecoregion- and waterbody-specific nutrient criteria ranges for Level III ecoregions. There are fourteen aggregate nutrient ecoregions. Each aggregate nutrient ecoregion is divided into smaller ecoregions referred to as Level III ecoregions. EPA will determine criteria ranges for the waterbody types and Level III ecoregions within the following aggregate nutrient ecoregions:

- Lakes and Reservoirs
  - Aggregate Nutrient ecoregions: 2, 6, 7, 8, 9, 11, 12, 13

- Rivers and Streams
  - Aggregate Nutrient ecoregions: 2, 3, 6, 7, 9, 11, 12, 14

1.2 References

This section lists documents that contain baselines, standards, guidelines, policies, and references that apply to the data analysis. Listed editions were valid at the time of publication. All documents are subject to revision, but these specific editions govern the concepts described in this document.

Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs (Draft). EPA, Office of Water, EPA 822-D-99-001, April 1999.

Nutrient Criteria Technical Guidance Manual: Rivers and Streams (Draft). EPA, Office of Water, EPA 822-D-99-003, September 1999.

Guidance for Data Quality Assessment: Practical Methods for Data Analysis. EPA, Office of Research and Development, EPA QA/G-9, January 1998.

2.0 QA/QC PROCEDURES

In order to develop nutrient criteria, EPA needed to obtain nutrient data from the states. EPA requested nutrient data from the states and forwarded the data sets to INDUS via e-mail and/or US mail. In addition, EPA tasked INDUS to convert data from three national data sets. EPA provided INDUS with a Legacy STORET extraction to convert into the database. The United States Geologic Survey (USGS) sent INDUS a CD-ROM with NASQAN data to convert. INDUS downloaded NAWQA files from the USGS Web site to convert the data. In total, INDUS converted and imported the following national and state data sets into the Nutrient Criteria Database:

- Legacy STORET
- NAWQA
- NASQAN
- Region 1
- Region 2 - Lake Champlain Monitoring Project
- Region 2 - NYSDEC Finger Lakes Monitoring Program
- Region 2 - NY Citizens Lake Assessment Program
- Region 2 - Lake Classification and Inventory Survey
- Region 2 - NYCDEP (1990-1998)
- Region 2 - NYCDEP (Storm Event data)
- Region 2 - New Jersey Nutrient Data (Tidal Waters)
- Region 5
- Region 3
- Region 3 - Nitrite Data
- Region 3 - Choptank River files
- Region 4 - Tennessee Valley Authority
- Region 7 - Central Plains Center for BioAssessment (CPCB)
- Region 7 - REMAP
- Region 2 - Delaware River Basin Commission (1990-1998)
- Region 3 - PA Lake Data
- Region 3 - University of Delaware
- Region 10
- University of Auburn

As part of the conversion process, INDUS performed a number of Quality Assurance/Quality Control (QA/QC) steps to ensure that the data was properly converted into the Nutrient Criteria Database. Section 2 explains the steps performed by INDUS to convert the data.

2.1 National Data Sets

INDUS converted three national data sets into the Nutrient Criteria Database: Legacy STORET data, NASQAN data, and NAWQA data. A previous EPA contractor performed the extraction of Legacy STORET data and documented the QA/QC procedures used on the data. This documentation is included in Appendix A. INDUS performed minimal QA/QC on the Legacy STORET data set because the previous contractor completed the steps outlined in Appendix A. INDUS and EPA also agreed to convert the NAWQA and NASQAN data sets with minimal QA/QC on the assumption that the source agency, the USGS, QA/QC'd the data.

For each of the three national data sets, INDUS ran queries to determine if 1) samples existed without results and 2) if stations existed without samples. Per Task Order Project Officer (TOPO) direction, these records were deleted from the system. For analysis purposes, EPA determined that there was no need to keep station records with no samples and sample records with no results. INDUS also confirmed that each data set contained no duplicate records.

In addition, INDUS deleted all composite results from the Legacy STORET data. Per TOPO direction, it was decided that composite sample results would not be used in the statistical analysis.

2.2 State Data

Each state data set was delivered in a unique format. Many of the data sets were delivered to INDUS without corresponding documentation. INDUS analyzed each state data set in order to determine which parameters should be converted for analysis. INDUS obtained a master parameter table from EPA and converted the parameters in the state data sets according to those that were present in the EPA parameter table. INDUS converted all of the data elements in the state data sets that mapped directly to the Nutrient Criteria Database; data elements that did not map to the Nutrient Criteria Database were not converted. In some cases, state data elements that did not directly map into the Oracle database were inserted into a comment field within the database. Also, INDUS maintained an internal record of which state data elements were inserted into the comment field.

As part of the data clean-up efforts, INDUS determined whether or not there were any duplicate records in the state data sets and deleted the duplicate records. INDUS checked the waterbody, station, and sample entities for duplicate records. In addition, INDUS deleted station records with no samples and sample records with no results. INDUS also deleted waterbody records that were not associated with a station. In each case, INDUS maintained an internal record of how many records were deleted.

If INDUS encountered referential integrity errors, such as samples that referred to stations that did not exist, or if INDUS was unsure of whether a record was a duplicate, INDUS contacted the agency directly via e-mail or phone to resolve any issues that arose. INDUS saved an electronic

copy of each e-mail correspondence with the states to ensure that a record of the decision was maintained. INDUS also contacted each agency to determine which laboratory methods were used for each parameter.

Finally, INDUS examined the remark codes of each result record in the state data sets. INDUS mapped the remark codes to the STORET remark codes listed in Table 2 of Appendix A. If any of the state result records were associated with remark codes marked as "Delete" in Table 2 of Appendix A, the result records were not converted into the database.

2.3 Laboratory Methods

Many of the state data sets did not contain laboratory method information. In addition, laboratory method information was not available for the three national data sets. In order to determine missing laboratory method information, EPA tasked another contractor to contact the data owners to obtain the laboratory method. In some cases, the data owners responded and the laboratory methods were added to the database.

2.4 Waterbody Name and Class Information

A large percentage of the data did not have waterbody-specific information. The only waterbody information contained in the three national data sets was the waterbody name, which was embedded in the station 'location description' field. Most of the state data sets contained waterbody name information; however, much of the data was duplicated throughout the data sets. Therefore, the waterbody information was cleaned manually. For the three national data sets, the 'location description' field was extracted from the station table and moved to a temporary table. The 'location description' field was sorted alphabetically. Unique waterbodies were grouped together based on name similarity and whether or not the waterbodies fell within the same county, state, and waterbody type. Finally, the 'location description' field was edited to include only waterbody name information, not descriptive information. For example, 110 MILE CREEK AT POMONA DAM OUTFLOW, KS PO-2 was edited to 110 MILE CREEK. Also, if 100 MILE CREEK was listed ten times in New York, but in four different counties, four 100 MILE CREEK waterbody records were created.

Similar steps were taken to eliminate duplicate waterbody records in the state data sets. If a number of records had similar waterbody names and fell within the same state, county, and waterbody type, the records were grouped to create a unique waterbody record.

Most of the waterbody data did not contain depth, surface area, and volume measurements. EPA needed this information to classify waterbody types. EPA attempted to obtain waterbody class

information from the states. EPA sent waterbody files to the regional coordinators and requested that certain class information be completed by each state. The state response was poor; therefore, EPA was not able to perform statistical analysis for the waterbody types by class.

2.5 Ecoregion Data

Aggregate nutrient ecoregions and Level III ecoregions were added to the database using the station latitude and longitude coordinates. If a station was lacking latitude and longitude coordinates or county information, the data were not included in the statistical analysis. Appendix B lists the steps taken to add the two ecoregion types (aggregate and Level III) to the Nutrient Criteria Database. The ecoregion names were pulled from aggregate nutrient ecoregion and Level III ecoregion Geographical Information System (GIS) coverages. In summary, the station latitude and longitude coordinates were used to determine the ecoregion under the following circumstances:

- The latitude and longitude coordinates fell within the county/state listed in the station table.
- The county data was missing.

The county centroid was used to determine the ecoregions under the following circumstances:

- The latitude and longitude coordinates were missing, but the state/county information was available.
- The latitude and longitude coordinates fell outside the county/state listed in the station table. The county information was assumed to be correct; therefore, the county centroid was used.

If the latitude and longitude coordinates fell outside the continental US county coverage file (i.e., the point fell in the ocean or Mexico/Canada), the nearest ecoregion was assigned to the station.

3.0 STATISTICAL ANALYSIS REPORTS

Aggregate nutrient ecoregion tables were created by extracting all observations for a specific aggregate nutrient ecoregion from the nutrient criteria database. Then, the data were reduced to create tables containing only the yearly median values. To create these tables, the median value for each waterbody was calculated using all observations for each waterbody by Level III ecoregion, year, and season. Tables of decade median values were created from the yearly median tables by calculating the median for each waterbody by Level III ecoregion by decade and season.

The Data Source and the Remark Code reports were created using all observations (all reported values). All the other reports were created from either the yearly median tables or the decade median tables. In other words, the descriptive statistics and regressions were run using the median values for each waterbody and not the individual reported values.

Statistical analyses were performed under the assumption that this data set is a random sample. If this assumption cannot be verified, the observations may or may not be valid. Values below the 1st and 99th percentile were removed from the Legacy STORET database prior to the creation of the national database. Also, data were treated according the Legacy STORET remark codes in Appendix A.

The following contains a list of each report and the purpose for creating each report:

- Data Source Created to provide a count of the amount of data and to identify the source(s).
- Remark Codes Created to provide a description of the data.
- Median of Each Waterbody by Year This was an intermediate step performed to obtain a median value for each lake to be used in the yearly descriptive statistics reports and the regression models.
- Median of Each Waterbody by Decade This was an intermediate step performed to obtain a median value for each lake to be used in the decade descriptive statistics.
- Descriptive Statistics Created to provide EPA with the desired statistics for setting criteria levels.
- Regression Models Created to examine the relationships between biological and nutrient variables.

Note: Separate reports were created for each season.

3.1 Data Source Reports

Data source reports were presented in the following formats:

- The number and percentage of data from each data source were summarized in tables for each aggregate nutrient ecoregion by season and waterbody type.

- The number and percentage of data from each data source were summarized in tables for each Level III ecoregion by season and waterbody type.

The 'Frequency' represents the number of data values from a specific data source for each parameter by data source. The 'Row Pct' represents the percentage of data from a specific data source for each parameter.

3.2 Remark Code Reports

Remark code reports were presented in the following formats:

- The number and percentage of data associated with a particular remark code for each parameter were summarized in tables by Level III ecoregion by decade and season.

- The number and percentage of data associated with a particular remark code for each parameter were summarized in tables by Level III ecoregion by year and season.

The 'Frequency' represents the number of data values corresponding to the remark code in the column. The 'Row Pct' represents the percentage of data that was associated with the remark code in that row.

In the database, remark codes that were entered by the states were mapped to Legacy STORET remark codes. Prior to the analysis, the data were treated according to these remark codes. For example, if the remark code was 'K,' then the reported value was divided by two. Appendix A contains a complete list of Legacy STORET remark codes.

Note: For the reports, a remark code of 'Z' indicates that no remark codes were recorded. It does not correspond to Legacy STORET code 'Z.'

3.3 Median of Each Waterbody

To reduce the data and to ensure heavily sampled waterbodies or years were not over represented in the analysis, median value tables (described above) were created. The yearly median tables and decade median tables were delivered to the EPA in electronic format as csv (comma separated value or comma delimited) files.

3.4 Descriptive Statistic Reports

The number of waterbodies, median, mean, minimum, maximum, 5th, 25th , 75th , 95th percentiles, standard deviation, standard error, and coefficient of variation were calculated. The tables (described above) containing the decade median values for each waterbody for each parameter were used to create descriptive statistics reports for:

- Level III ecoregions by decade and season
- Aggregate nutrient ecoregions by decade and season

In addition, the tables containing the yearly median values for each waterbody for each parameter were used to create descriptive statistics reports for:

- Level III ecoregions by year and season

3.5 Regression Models

Simple linear regressions using the least squares method were performed to examine the relationships between biological and nutrient variables in lakes and reservoirs, and rivers and streams. Regressions were performed using the yearly median tables. Chlorophyll(s) in micrograms per liter (ug/L), secchi in meters (m), dissolved oxygen in milligrams per liter (mg/L), turbidity, and pH were the biological variables in these models. When there was little or no data for chlorophyll, then pH or dissolved oxygen was substituted for chlorophyll. Secchi data were used in the lake and reservoir models, and turbidity data were used in the river and stream models. The nutrient variables in these models include: total phosphorus in ug/L, total nitrogen in mg/L, total kjeldahl nitrogen in mg/L, and nitrate and nitrite in mg/L. Regressions were also run for total nitrogen and total phosphorus for ecoregions where both these variables were measured.

Note: At the time of creation of this document only regressions for aggregate nutrient ecoregion 7 for lakes and reservoirs were delivered to the EPA. Regressions for the remaining aggregate nutrient ecoregions will be delivered in August 2000.

4.0 TIME PERIOD

Data collected from January 1990 to December 1999 were used in the statistical analysis reports. To capture seasonal differences, the data were classified as follows:

- Aggregate nutrient ecoregions: 6, 7, and 8

  - Spring: April to May
  - Summer: June to August
  - Fall: September to October
  - Winter: November to March

- Aggregate nutrient ecoregions: 1, 2, 9, 10, 11, 12, and 13

  - Spring: March to May
  - Summer: June to August
  - Fall: September to November
  - Winter: December to February

5.0 DATA SOURCES AND PARAMETERS FOR THE AGGREGATE NUTRIENT ECOREGIONS

This section provides information for the nutrient aggregate ecoregions that were analyzed by waterbody type. Each section lists the data sources for the aggregate nutrient ecoregion including: 1) the data sources, 2) the parameters included in the analysis, and 3) the Level III ecoregions within the aggregate nutrient ecoregions.

Note: For analysis purposes, the following parameters were combined to form Phosphorous, Dissolved Inorganic (DIP):

Phosphorus, Dissolved Inorganic (DIP)
Phosphorus, Dissolved (DP)
Phosphorus, Dissolved Reactive (DRP)
Orthophosphate, dissolved, mg/L as P
Orthophosphate (OPO4_PO4)

5.1 Lakes and Reservoirs

5.1.1 Aggregate Nutrient Ecoregion 2

Data Sources:

Legacy STORET
EPA Region 10

Parameter:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
Phosphorus, Total Reactive	(ug/L)
SECCHI	(m)
pH	

Level III ecoregions:

1, 2, 4, 5, 9, 11, 15, 16, 17, 19, 21, 23, 41, 77, 78

5.1.2 Aggregate Nutrient Ecoregion 6

Data Sources:

Legacy STORET

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
SECCHI	(m)

Level III ecoregions:

46, 47, 48, 54, 55, 57

5.1.3 Aggregate Nutrient Ecoregion 7

Data Sources:

LCMPD
Legacy STORET
NYCDEP
EPA Region 1

Parameters:

Chlorophyll A, Fluorometric Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric, Uncorrected	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)

Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Orthophosphate, Total as P	(ug/L)
Phosphorus, Total (TP)	(ug/L)
SECCHI	(m)

Level III ecoregions:

51, 52, 53, 56, 60, 61, 83

5.1.4 Aggregate Nutrient Ecoregion 8

Data Sources:

LCMPD
Legacy STORET
NYCDEP
NYCDEC
EPA Region 1
EPA Region 3

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric, Uncorrected	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Chlorophyll B	(ug/L)
Chlorophyll C	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
SECCHI	(m)

Level III ecoregions:

49, 50, 58, 62, 82

5.1.5 Aggregate Nutrient Ecoregion 9

Data Sources:

Auburn University
Legacy STORET
EPA Region 4

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Pheophytin	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric, Uncorrected	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
SECCHI	(m)

Level III ecoregions:

29, 33, 35, 37, 40, 45, 64, 65, 71, 72, 74

5.1.6 Aggregate Nutrient Ecoregion 11

Data Sources:

Auburn University
Legacy STORET
NYSDEC
EPA Region 3
EPA Region 4

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Pheophytin	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric, Uncorrected	(ug/L)

Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
SECCHI	(m)

Level III ecoregions:

36, 38, 39, 66, 67, 68, 69, 70

5.1.7 Aggregate Nutrient Ecoregion 12

Data Sources:

Legacy STORET

Parameters:

Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
SECCHI	(m)

Level III ecoregions:

75

5.1.8 Aggregate Nutrient Ecoregion 13

Data Sources:

Legacy STORET

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
SECCHI	(m)

Level III ecoregions:

76

5.2 Rivers and Streams

5.2.1 Aggregate Nutrient Ecoregion 2

Data Sources:

Legacy STORET
NASQAN
NAWQA
EPA Region 10

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, chromotographic- fluorometric	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Chlorophyll B, Phytoplankton, chromotographic- fluorometric	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Phosphorus, Orthophosphate, Total as P	(ug/L)
Phosphorus, Total (TP) Reactive	(ug/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
Turbidity	(FTU)

Turbidity	(JCU)
Turbidity	(NTU)

Level III ecoregions:

1, 2, 4, 5, 8, 9, 11, 15, 16, 17, 19, 21, 23, 41, 77, 78

5.2.2 Aggregate Nutrient Ecoregion 3

Data Sources:

Legacy STORET
NASQAN
NAWQA
EPA Region 10

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, chromotographic- fluorometric	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Chlorophyll B, Phytoplankton, chromotographic- fluorometric	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
Turbidity	(FTU)
Turbidity	(JCU)
Turbidity	(NTU)

Level III ecoregions:

6, 10, 12, 13, 14, 18, 20, 22, 24, 79, 80, 81

5.2.3 Aggregate Nutrient Ecoregion 6

Data Sources:

Legacy STORET
NASQAN
NAWQA
EPA Region 5
EPA Region 7

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, chromotographic- fluorometric	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Chlorophyll B, Phytoplankton, chromotographic- fluorometric	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Organic, Phosphorus	(ug/L)
Phosphorus, Total (TP)	(ug/L)
Phosphorus, Orthophosphate, Total as P	(ug/L)
Turbidity	(FTU)
Turbidity	(JCU)
Turbidity	(NTU)

Level III ecoregions:

46, 47, 48, 54, 55, 57

5.2.4 Aggregate Nutrient Ecoregion 7

Data Sources:

LCMPD
Legacy STORET
NASQAN
NAWQA
NYCDEP

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric, Uncorrected	(ug/L)
Chlorophyll A, Phytoplankton, chromotographic- fluorometric	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Chlorophyll B, Phytoplankton, chromotographic- fluorometric	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Organic, Phosphorus	(ug/L)
Phosphorus, Orthophosphate, Total as P	(ug/L)
Phosphorus, Total (TP)	(ug/L)
Turbidity	(FTU)
Turbidity	(JCU)
Turbidity	(NTU)

Level III ecoregions:

51, 52, 53, 56, 60, 61, 83

5.2.5 Aggregate Nutrient Ecoregion 9

Data Sources:

Auburn University
Legacy STORET
NASQAN
NAWQA
EPA Region 3
EPA Region 5
EPA Region 7

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, chromotographic- fluorometric	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric, Uncorrected	(ug/L)

Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Chlorophyll B, Phytoplankton, chromotographic- fluorometric	(ug/L)
Chlorophyll B, Phytoplankton, Spectrophotometric	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Organic, Phosphorus	(ug/L)
Phosphorus, Orthophosphate, Total as P	(ug/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
Turbidity	(FTU)
Turbidity	(JCU)
Turbidity	(NTU)

Level III ecoregions:

29, 33, 35, 37, 40, 45, 64, 65, 71, 72, 74

5.2.6 Aggregate Nutrient Ecoregion 11

Data Sources:

Auburn University
Legacy STORET
NASQAN
NAWQA
EPA Region 3
EPA Region 5
EPA Region 7

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, chromotographic- fluorometric	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric, Uncorrected	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Chlorophyll B, Phytoplankton, chromotographic- fluorometric	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Organic, Phosphorus	(ug/L)

Phosphorus, Orthophosphate, Total as P	(ug/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
Turbidity	(FTU)
Turbidity	(JCU)
Turbidity	(NTU)

Level III ecoregions:

36, 38, 39, 66, 67, 68, 69, 70

5.2.7 Aggregate Nutrient Ecoregion 12

Data Sources:

Legacy STORET
NASQAN
NAWQA

Parameters:

Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric, Uncorrected	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Chlorophyll B, Phytoplankton, Spectrophotometric	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Phosphorus, Orthophosphate, Total as P	(ug/L)
Phosphorus, Total (TP)	(ug/L)
Turbidity	(FTU)
Turbidity	(NTU)

Level III ecoregions:

75

5.2.8 Aggregate Nutrient Ecoregion 14

Data Sources:

Legacy STORET
NASQAN
NAWQA
NYCDEP
EPA Region 1
EPA Region 3

Parameters:

Chlorophyll A, Fluorometric, Corrected	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric Acid	(ug/L)
Chlorophyll A, Phytoplankton, Spectrophotometric, Uncorrected	(ug/L)
Chlorophyll A, Trichromatic, Uncorrected	(ug/L)
Phosphorous, Dissolved Inorganic (DIP)	(ug/L)
Dissolved Oxygen (DO)	(mg/L)
Nitrite and Nitrate, (NO2+NO3)	(mg/L)
Phosphorus, Orthophosphate, Total as P	(ug/L)
Nitrogen, Total Kjeldahl (TKN)	(mg/L)
Nitrogen, Total (TN)	(mg/L)
Phosphorus, Total (TP)	(ug/L)
Turbidity	(FTU)
Turbidity	(JCU)
Turbidity	(NTU)

Level III ecoregions:

59, 63, 84

APPENDIX A

Process Used to QA/QA the Legacy STORET Nutrient Data Set

1. STORET water quality parameters and Station and Sample data items were retrieved from USEPA's mainframe computer. Table 1 lists all retrieved parameters and data items.

TABLE 1: PARAMETERS AND DATA ITEMS RETRIEVED FROM STORET

Parameters Retrieved (STORET Parameter Code)	Station Data Items Included (STORET Item Name)	Sample Data Items Included (STORET Item Name)
TN - mg/l (600) TKN - mg/l (625) Total Ammonia (NH3+NH4) - mg/l (610) Total NO2+NO3 - mg/l (630) Total Nitrite - mg/l (615) Total Nitrate - mg/l (620) Organic N - mg/L (605) TP - mg/l (665) Chlor *a* - ug/L (spectrophotometric method, 32211) Chlor *a* - ug/L (fluorometric method corrected, 32209) Chlor *a* - ug/L (trichromatic method corrected, 32210) Secchi Transp. - inches (77) Secchi Transp. - meters (78) +Turbidity JCUs (70) +Turbidity FTUs (76) +Turbidity NTUs field (82078) +Turbidity NTUs lab (82079) +DO - mg/L (300) +Water Temperature (degrees C, 10/degrees F, 11)	Station Type (TYPE) Agency Code (AGENCY) Station No. (STATION) Latitude - std. decimal degrees (LATSTD) Longitude - std. decimal degrees (LONGSTD) Station Location (LOCNAME) County Name (CONAME) State Name (STNAME) Ecoregion Name - Level III (ECONAME) Ecoregion Code -Level III (ECOREG) Station Elevation (ELEV) Hydrologic Unit Code (CATUNIT) RF1 Segment and Mile (RCHMIL) RF1ON/OFF tag (ONOFF)	Sample Date (DATE) Sample Time (TIME) Sample Depth (DEPTH) Composite Sample Code (SAMPMETH)

+ If data record available at a station included data only for this or other such marked parameters, data record was deleted from data set.

The following set of retrieval rules were applied to the retrieval process:

- Data were retrieved for waterbodies specified only as 'lake', 'stream', 'reservoir', or 'estuary' under "Station Type" parameter. Any stations specified as 'well,' 'spring,' or 'outfall' were eliminated from the retrieved data set.

- Data were retrieved for station types described as 'ambient' (e.g., no pipe or facility discharge data) under the "Station Type" parameter.

- Data were retrieved that were designated as 'water' samples only. This includes 'bottom' and 'vertically integrated' water samples.

- Data were retrieved that were designated as either 'grab' samples and 'composite' samples (mean result only).

- No limits were specified for sample depths.

- Data were retrieved for all fifty states, Puerto Rico, and the District of Columbia.

- The time period specified for data retrieval was January 1990 to September 1998.

- No data marked as "Retired Data" (i.e., data from a generally unknown source) were retrieved.

- Data marked as "National Urban Runoff data" (i.e., data associated with sampling conducted after storm events to assess nonpoint source pollutants) were included in the retrieval. Such data are part of STORET's 'Archived' data.

- Intensive survey data (i.e., data collected as part of specific studies) were retrieved.

2. Any values falling below the 1st percentile and any values falling above the 99th percentile were transformed into 'missing' values (i.e., values were effectively removed from the data set, but were not permanently eliminated).

3. Based on the STORET 'Remark Code' associated with each retrieved data point, the following rules were applied (Table 2):

TABLE 2: STORET REMARK CODE RULES

STORET Remark Code	Keep or Delete Data Point
blank - Data not remarked.	Keep
A - Value reported is the mean of two or more determinations.	Keep
B - Results based upon colony counts outside the acceptable ranges.	Delete
C - Calculated. Value stored was not measured directly, but was calculated from other data available.	Keep
D - Field measurement.	Keep

E - Extra sample taken in compositing process.	Delete
F - In the case of species, F indicates female sex.	Delete
G - Value reported is the maximum of two or more determinations.	Delete
H - Value based on field kit determination; results may not be accurate.	Delete
I - The value reported is less than the practical quantification limit and greater than or equal to the method detection limit.	Keep, but used one-half the reported value as the new value.
J - Estimated. Value shown is not a result of analytical measurement.	Delete
K - Off-scale low. Actual value not known, but known to be less than value shown.	Keep, but used one-half the reported value as the new value.
L - Off-scale high. Actual value not known, but known to be greater than value shown.	Keep
M - Presence of material verified, but not quantified. Indicates a positive detection, at a level too low to permit accurate quantification.	Keep, but used one half the reported value as the new value.
N - Presumptive evidence of presence of material.	Delete
O - Sample for, but analysis lost. Accompanying value is not meaningful for analysis.	Delete
P - Too numerous to count.	Delete
Q - Sample held beyond normal holding time.	Delete
R - Significant rain in the past 48 hours.	Delete
S - Laboratory test.	Keep
T - Value reported is less than the criteria of detection.	Keep, but replaced reported value with 0.

U - Material was analyzed for, but not detected. Value stored is the limit of detection for the process in use.	Keep, but replaced reported value with 0.
V - Indicates the analyte was detected in both the sample and associated method blank.	Delete
W - Value observed is less than the lowest value reportable under remark "T."	Keep, but replaced reported value with 0.
X - Value is quasi vertically-integrated sample.	No data point with this remark code in data set.
Y - Laboratory analysis from unpreserved sample. Data may not be accurate.	Delete
Z - Too many colonies were present to count.	Delete
If a parameter (excluding water temperature) value was less than or equal to zero and no remark code was present, the value was transformed into a missing value. Rationale - Parameter concentrations should never be zero without a proper explanation. A method detection limit should at least be listed.	

4. Station records were eliminated from the data set if any of the following descriptors were present within the "Station Type" parameter:

- **MONITR -** Source monitoring site, which monitors a known problem or to detect a specific problem.
- **HAZARD -** Site of hazardous or toxic wastes or substances.
- **ANPOOL -** Anchialine pool, underground pools with subsurface connections to watertable and ocean.
- **DOWN -** Downstream (i.e., within a potentially polluted area) from a facility which has a potential to pollute.
- **IMPDMT -** Impoundment. Includes waste pits, treatment lagoons, and settling and evaporation ponds.
- **STMSWR -** Storm water sewer.
- **LNDFL -** Landfill.
- **CMBMI -** Combined municipal and industrial facilities.
- **CMBSRC -** Combined source (intake and outfall).

Rationale - these descriptors potentially indicate a station location that at which an ambient water sample would not be obtained (i.e., such sampling locations are potentially

biased) or the sample location is not located within one of the designated water body types (i.e, ANPOOL).

5. Station records were eliminated from data set if the station location did not fall within any established cataloging unit boundaries based on their latitude and longitude.

6. Using nutrient ecoregion GIS coverage provided by USEPA, all station locations with latitude and longitude coordinates were tagged with a nutrient ecoregion identifier (nutrient region identifiers are values 1 - 14) and the associated nutrient ecoregion name. Because no nutrient ecoregions exist for Alaska, Hawaii, and Puerto Rico, stations located in these states were tagged with "dummy" nutrient ecoregion numbers (20 = Alaska, 21 = Hawaii, 22 = Puerto Rico).

7. Using information provided by TVA, 59 station locations that were marked as 'stream' locations under the "Station Type" parameter were changed to 'reservoir' locations.

8. The nutrient data retrieved from STORET were assessed for the presence of duplicate data records. The duplicate data identification process consisted of three steps: 1) identification of records that matched exactly in terms of each variable retrieved; 2) identification of records that matched exactly in terms of each variable retrieved except for their station identification numbers; and 3) identification of records that matched exactly in terms of each variable retrieved except for their collecting agency codes. The data duplication assessment procedures were conducted using SAS programs.
Prior to initiating the data duplication assessment process, the STORET nutrient data set contained:

41,210 station records
924,420 sample records

- <u>Identification of exactly matching records</u>
All data records were sorted to identify those records that matched exactly. For two records to match exactly, all variables retrieved had to be the same. For example, they had to have the same water quality parameters, parameter results and associated remark codes, and have the same station data item and sample data item information. Exactly matching records were considered to be exact duplicates, and one duplicate record of each identified matching set were eliminated from the nutrient data set. A total of 924 sample records identified as duplicates by this process were eliminated from the data set.

- <u>Identification of matching records with the exception of station identification number</u>
All data records were sorted to identify those records that matched exactly except for their station identification number (i.e., they had the same water quality

parameters, parameter results and associated remark codes, and the same station and sample data item information with the exception of station identification number). Although the station identification numbers were different, the latitude and longitude for the stations were the same indicating a duplication of station data due to the existence of two station identification numbers for the same station. For each set of matching records, one of the station identification numbers was randomly selected and its associated data were eliminated from the data set. A total of 686 sample records were eliminated from the data set through this process.

- Identification of matching records with the exception of collecting agency codes
  All data records were sorted to identify those records that matched exactly except for their collecting agency codes (i.e., they had the same water quality parameters, parameter results and associated remark codes, and the same station and sample data item information with the exception of agency code). The presence of two matching data records each with a different agency code attached to it suggested that one agency had utilized data collected by the other agency and had entered the data into STORET without realizing that it already had been placed in STORET by the other agency. No matching records with greater than two different agency codes were identified. For determining which record to delete from the data set, the following rules were developed:

  - If one of the matching records had a USGS agency code, the USGS record was retained and the other record was deleted.
  - Higher level agency monitoring program data were retained. For example, federal program data (indicated by a "1" at the beginning of the STORET agency code) were retained against state (indicated by a "2") and local (indicated by values higher than 2) program data.
  - If two matching records had the same level agency code, the record from the agency with the greater number of overall observations (potentially indicating the data set as the source data set) was retained.

  A total of 2,915 sample records were eliminated through this process.

As a result of the duplicate data identification process, a total of 4,525 sample records and 36 individual station records were removed from the STORET nutrient data set. The resulting nutrient data set contains the following:

41,174 station records
919,895 sample records

APPENDIX B

Process for Adding Aggregate Nutrient Ecoregions and Level III Ecoregions

Steps for assigning Level III ecoregions and aggregate nutrient ecoregion codes and names to the Nutrient Criteria Database (performed using ESRI's ARCView v 3.2 and its GeoProcessing Wizard). This process is performed twice; once for the Level III ecoregions and once for the aggregate nutrient ecoregions:

- Add the station .dbf data table, with latitude and longitude data, to project by 'Add Event Theme'
- Convert to the shapefile format
- Create 'stcojoin' field, populate the 'stcojoin' field with the following formula: 'County.LCase+State.LCase'
- Add field 'stco_flag' to the station shapefile
- Spatially join the station data with the county shapefile (cntys_jned.shp)
- Select 'stcojoin' (station shapefile) field = 'stco_join2' (county shapefile) field
- Calculate stco_flag = 0 for selected features
- Step through all blank stco_flag records, assign the appropriate stco_flags, see list on the following page
- Select all stco_flags = 4 or 7, switch selection
- Calculate ctyfips (station) to cntyfips (county)
- Stop editing and save edits, remove all joins
- Add in 2 new fields 'x-coord1' and 'y-coord1' into station table
- Select all stco_flags = 1, 2, and 6
- Link county coverage with station coverage
- Populate 'x-coord1' and 'y-coord1' with 'x-coord' and 'y-coord' from county coverage
- Select all stco_flags = 1, 2, and 6, export to new .dbf file
- Add new .dbf file as event theme
- Convert to shapefile format
- Add the following fields to both tables (original station and station126 shapefiles): 'eco_omer', 'name_omer', 'dis_aggr', 'code_aggr', 'name_aggr'
- Spatially join station126 and eco-omer coverage
- Populate the 'eco_omer' field with the 'eco' value
- Repeat the previous step using the nearest method (line coverage) to determine ecoregion assignment for the line coverage, if some records are blank
- Spatially join the ecoregion line coverage to station coverage, link the LPoly# (from the spatially joined table) to Poly# (of the ecoregion polygon coverage)
- Populate the Eco fields with the appropriate information.
- Follow the same steps to the Rpoly#
- Remove all table joins
- Link the useco-om table with station126 table and populate 'name-omer' field
- Spatially join station aggr coverage and populate the rest of the fields. Follow the same procedures as outlined above
- Remove all joins
- Make sure the new Eco field added into the station126 shapefile are different than

the ones in the original station shapefile
- Join station126 and station coverage by station-id
- Populate all the Eco fields in the original station coverage
- Remove all joins
- Save table
- Make sure that all ctyfips records are populated; the county shapefile may have to be joined to populate the records, if the stco_flag = 4
- Create 2 new fields, 'NewCounty' and 'NewState'
- Populate these new fields with a spatial join to the county coverage
- Select by feature (ecoregion shapefile) all of the records in the station shapefile
- Switch selection (to get records outside of the ecoregion shapefile)
- If any of the selected records have stco_flag = 0 (they are outside the ecoregion shapefile boundary), calculate them to stco_flag = 3

stco_flags (state/county flags in order of importance)

0 The state and county values from the data set matched the state and county values from the spatial join.
(Ecoregions were assigned based on the latitude/longitude coordinates.)

1 The state and county values from the data set did not match the state and county values from the spatial join, but the point was inside the county coverage boundary.
(Ecoregions were assigned based on the county centroid.)

2 The state and county values from the data set did not match the state and county values from the spatial join because the point was outside the county coverage boundary; therefore, there was nothing to compare to the point (i.e., the point falls in the ocean/Canada/Mexico). This occurred for some coastal samples.
(Ecoregions were assigned based on the county centroid.)

3 The state and county values from the data set matched the state and county from the spatial join, but the point was outside the ecoregion boundary.
(Ecoregions were assigned to the closest ecoregion to the point.)
(No ecoregions were assigned to AK, HI, PR, BC, and GU.)

4 Latitude/longitude coordinates were provided, but there was no county information.
(Ecoregions were assigned based on the latitude/longitude coordinates.)

5 The state and county values from the data set did not match the state and county values from the spatial join due to spelling or naming convention errors. The matches were performed manually.
(Ecoregions were assigned based on the latitude/longitude coordinates.)

6 No latitude/longitude coordinates were provided, only state and county information was available.
(Ecoregions were assigned based on the county centroid.)

7 No latitude/longitude coordinates were provided, only state information was available; therefore, no matches were possible.
(Ecoregions were not assigned. Data is not included in the analysis.)

APPENDIX C

Glossary

Coefficient of Variation- Equal to the standard deviation divided by the mean multiplied by 100.

Maximum- The highest value.

Mean- The arithmetic average.

Median- The 50th percentile or middle value. Half of the values are above the median, and half of the values are below the median.

Minimum- The lowest value.

Standard Deviation- Equal to the square root of the variance with the variance defined as the sum of the squared deviations divided by the sample size minus one.

Standard Error- Standard error of the mean is equal to the standard deviation divided by the square root of the sample size.