Regional water quality models for the prediction of eutrophication endpoints

Anindita Das
Louisiana State University and Agricultural and Mechanical College, adas1@lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses
Part of the Environmental Sciences Commons

Recommended Citation

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master’s Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
REGIONAL WATER QUALITY MODELS FOR THE PREDICTION OF EUTROPHICATION ENDPOINTS

A Thesis

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College In partial fulfillment of the Requirements for the degree of Master of Science

in

The Department of Environmental Studies

by

Anindita Das
B.S., University of Calcutta, 1991
December, 2003
ACKNOWLEDGEMENTS

I am extremely grateful to my major professor, Dr. E. Conrad Lamon III, for his patience and guidance. He was always there to explain and give valuable suggestions for the improvement of this thesis. I would also like to express my sincere appreciation to my committee members, Dr. Walter K. Keithly, Dr. Michael Wascom and Dr. Margaret Reams, for their support (course-related as well as moral) throughout the two years in this department.

I would like to thank my husband, Neel Das, who has always stood by me and was always there to reassure me when I was feeling disheartened. I am forever grateful to my parents for being my pillar of strength and for encouraging me in all that I do. Thanks also to my friends and family for their positive attitude, understanding and support.
# TABLE OF CONTENTS

**ACKNOWLEDGEMENTS** ................................................................. ii  
**LIST OF TABLES** ........................................................................ vi  
**LIST OF FIGURES** .................................................................... vii  
**ABSTRACT** ................................................................................ viii  
**CHAPTER 1. INTRODUCTION** .................................................... 1  
**CHAPTER 2. EPA’S WATER QUALITY INVENTORY** ..................... 6  
  Sources of Nutrient Pollution .................................................... 7  
  Limiting Factors in Eutrophication .......................................... 9  
**CHAPTER 3. ESTABLISHING NUTRIENT CRITERIA** ................. 11  
**CHAPTER 4. DEVELOPMENT OF ECOREGIONS** ................. 15  
  Formation of Nutrient Criteria Database ............................... 16  
  Dealing with Quality of Historical Data ............................... 18  
**CHAPTER 5. MATERIALS AND METHODS** ............................ 21  
  Regression Criticism ............................................................... 25  
  Residuals Used ................................................................. 26  
  Residual Analysis ............................................................... 27  
  Exploratory Data Analysis Techniques ............................... 28  
**CHAPTER 6. RESULTS** ............................................................. 30  
**CHAPTER 7. DISCUSSION** ....................................................... 44  
  Limitations of the Models ..................................................... 44  
  Other Approaches ............................................................... 47  
  Summary/Conclusion ............................................................ 48  
**REFERENCES** ............................................................................ 52  
**APPENDIX A. CONSENT FORM** ............................................. 56  
**APPENDIX B. SAS ANOVA OUTPUT FOR MODELS** ............ 59  
**APPENDIX C. TABLES WITH PARAMETER VALUES FOR EACH MODEL** .......................................................... 68  
**APPENDIX D. BOXPLOTS OF DISTRIBUTION OF CHLOROPHYLL A FOR EACH MODEL BY ECOREGION** .......................................................... 73
LIST OF TABLES

Table 1: Regression equations for TN.................................................................22
Table 2: Total number of data and spatial spread of models.............................24
Table 3: Comparison of the six models .............................................................42
Table 4: Weighted ranks of each model.............................................................43
LIST OF FIGURES

Figure 1: Percentage of lakes assessed .................................................................7

Figure 2: The bar charts, (a) and (b), present the leading sources and the number of lake, reservoir, and pond acres impacted ..........................................................8

Figure 3: Fourteen nutrient Ecoregions as delineated by Omernik (2000) .............17

Figure 4: Summary plots of the residuals for Model 1 ...........................................31

Figure 5: Predicted chlorophyll $a$ vs. residual plot for Model 1 ............................31

Figure 6: Cook’s Distance plot for Model 1 ........................................................32

Figure 7: Summary plots of the residuals for Model 2 ..........................................33

Figure 8: Predicted chlorophyll $a$ vs. residual plot for Model 2 ............................33

Figure 9: Cook’s Distance plot for Model 2 ........................................................34

Figure 10: Summary plots of the residuals for Model 3 ........................................35

Figure 11: Predicted chlorophyll $a$ vs. residual plot for Model 3 ...........................35

Figure 12: Cook’s Distance plot for Model 3 ........................................................36

Figure 13: Summary plots of the residuals for Model 4 ........................................37

Figure 14: Predicted chlorophyll $a$ vs. residual plot for Model 4 ............................38

Figure 15: Cook’s Distance plot for Model 4 ........................................................39

Figure 16: Summary plots of the residuals for Model 5 ........................................39

Figure 17: Predicted chlorophyll $a$ vs. residual plot for Model 5 ............................40

Figure 18: Cook’s Distance plot for Model 5 ........................................................40

Figure 19: Summary plots of the residuals for Model 6 ........................................41

Figure 20: Predicted chlorophyll $a$ vs. residual plot for Model 6 ............................41

Figure 21: Cook’s Distance plot for Model 6 ........................................................42
ABSTRACT

Eutrophication is a process by which a waterbody progresses from its origin to its extinction. During this period, there is a gradual accumulation of nutrients and organic biomass, accompanied by a decrease in average depth of the water due to sediment accumulation, and an increase in primary productivity, usually in the form of dense algal blooms. Cultural eutrophication occurs when humans, through their various activities, greatly accelerate this process. Eutrophication can cause loss in species diversity, fish kills, and decrease the aesthetic value of a waterbody. The EPA is trying to prevent cultural eutrophication by setting standards for water quality criteria for each of the fourteen ecoregions in the United States. Nutrients are the most common pollutants affecting waterbodies. The EPA considers total phosphorous and total nitrogen as the two causal variables and chlorophyll $a$ and Secchi depth as the two early indicator response variables. There are models that predict the relationship of chlorophyll $a$ to phosphorous and chlorophyll $a$ to nitrogen, but there are very few that combine phosphorous and nitrogen to predict chlorophyll $a$ at a cross-sectional level. This study is concerned with fitting a linear model for the prediction of chlorophyll $a$, using phosphorous and nitrogen, for the fourteen ecoregions. Six combinations of the three variables have been tested (because of the different methods used to obtain each variable) to find out which model is the best with respect to model fit, number of observations, and geographical coverage. The best model can then be used in further studies to determine eutrophication end points at smaller and more homogeneous divisions of the ecoregion for better management of water quality in lakes.
CHAPTER 1. INTRODUCTION

Eutrophication is a process by which a waterbody progresses from its origin to its extinction (Novotny and Olem, 1994). Natural eutrophication occurs over thousands of years during which lakes gradually age and become more productive. During this period, there is a gradual accumulation of nutrients and organic biomass, accompanied by a decrease in average depth of the water due to sediment accumulation, and an increase in primary productivity, usually in the form of dense algal blooms. These algal blooms become the dominant species in the water body and overshadow the flora and fauna in the deeper water column, leading to a loss of diversity.

The EPA characterizes as eutrophic waterbodies that have decreasing hypolimnetic dissolved oxygen concentrations, increasing nutrient concentrations, increasing suspended solids, especially organic material, progress from a diatom population to a population dominated by blue-green or green algae, decreasing light penetration, and increasing phosphorous concentrations in the sediments.

Cultural eutrophication occurs when humans, through their various activities, greatly accelerate this process. This might be beneficial in some aquatic systems. For example, in aquaculture, ponds are deliberately fertilized to increase the production of fish or shellfish. In general, though, cultural eutrophication causes problems when the increased production levels, and the effects associated with this, are not compatible with the intended uses of the waterbody. This can drastically reduce the life span of a lake through accelerated aging.

Cultural eutrophication affects the value of aquatic systems in three major ways. First, the species associated with the eutrophic system are undesirable. For example, as
blue-green algae become more dominant due to nutrient enrichment, the fish species present may be commercially less valuable, so, the monetary value of the waterbody declines.

The second reason is that in highly eutrophic systems, oxygen concentrations vary over a wide range due to increased productivity and plant biomass decomposition. Many organisms cannot stand such fluctuations (which may result in fish kills). Large-scale fish kills of desirable species are a serious problem associated with eutrophication.

Finally, increased phytoplankton populations (due to nutrient enrichment) decrease the aesthetic value of water by making it appear turbid. Large plant biomasses decay, creating a rotten smell. These may make a waterbody unsuitable for water supply, contact recreation and navigation.

The Environmental Protection Agency’s (EPA) mandate is to protect the country’s waters by setting standards for the management of water quality criteria in the United States. Section 304 of the Clean Water Act deals with a scientific assessment of ecological and human health effects recommended by EPA to the States and Tribes for establishing water quality standards. These would serve as a basis for control of discharges or release of pollutants (USEPA 2000). EPA intends to use this scientific assessment to develop the default nutrient criteria in Section 304(a) of the Clean Water Act for the all the ecoregions in the country. This process requires good policy, which requires good science (Reckhow, 1994). But there are many uncertainties in science (e.g., natural variability, modeler bias, sampling bias, knowledge base, etc.). According to Reckhow (1994), one way to approach this problem is to first build a decision planning framework that identifies management objectives, attributes to calculate achievement of
those objectives, and plans that have to be executed to attain those objectives. This would help in making management decisions, taking into account significant scientific uncertainty.

For example, control of eutrophication might be a management objective. One way to approach this is the control of algal blooms. Chlorophyll $a$ is a surrogate measure of the amount of algal biomass present in a waterbody. Phosphorous and nitrogen are attributes that can be used to predict the amount of chlorophyll $a$ present in a waterbody. Chlorophyll $a$ would be considered an attribute which can be used as a measure for achieving an objective (control of eutrophication). The next step in the framework is to define the type of scientific decision support to attain the objective. Decision support includes predictive methods (e.g., expert judgments, simulation models) and information needs (e.g., research, monitoring, experiments) (Reckhow, 1994).

In this regard, many models have been proposed, like WASP4 (Ambrose et al, 1988), QUAL2E (Brown et al, 1987), EUTROMOD (Reckhow, 1992). Chlorophyll $a$ is considered a measure of eutrophication because the amount of chlorophyll $a$ can be used to estimate the amount of algal biomass present in a waterbody. Presence of phosphorous and nitrogen are considered the primary causes of eutrophication (USEPA, 2000). There are several models that deal with phosphorous and nitrogen loading (Vollenweider, 1969; Portielje and Van der Molen, 1999). There are also models that predict the relationship of chlorophyll $a$ to phosphorous (Jones et al, 1998; Walker and Havens, 1995) and chlorophyll $a$ to nitrogen (De Vries et al, 1998; Mineeva, 1993), but there are very few that combine phosphorous and nitrogen to predict chlorophyll $a$ (Lamon, 1995; Lamon
and Clyde, 2000). This combination is important because taken together, these two variables might be able to explain more variation in chlorophyll \( a \) in the waterbodies.

The Environmental Protection Agency (EPA) is endeavoring to set standards for water quality criteria for each of the fourteen ecoregions (USEPA, 2002) in the United States. The first step in addressing this issue is designing models required for decision support. Data availability for these models is the next step. In this regard, EPA has developed a Nutrient Criteria Database containing data for many variables as a beginning point for development of models. However, these models also require consistency in terms of the variables being used. For example, prediction of chlorophyll \( a \) from phosphorous and nitrogen requires that all three variables be measured simultaneously in all the observations and that each parameter be measured using a single method over time (because it is recommended by EPA that data for the same variable cannot be interchanged if they have been obtained by different methods).

This study is concerned with fitting a linear model for the prediction of chlorophyll \( a \) using phosphorous and nitrogen as predictors, for the fourteen ecoregions (recommended by EPA). The Nutrient Criteria Database will be assessed to find whether or not it is consistent when it comes to using the same method for measuring a variable or in measuring all the variables at the same time from a single sample of water. Six combinations of the three variables mentioned above will be tested (because of the different methods used to obtain each variable) to find out which model gives the best estimate of chlorophyll \( a \) with respect to model fit, number of observations, and geographical coverage. This can then be treated as a preliminary step in the scientific decision support of the planning framework for management of eutrophication. The
sections that follow will discuss the EPA’s water quality inventory, development of nutrient criteria, formation of ecoregions, materials and methods used, results, and discussion.
CHAPTER 2. EPA’S WATER QUALITY INVENTORY

EPA’s ecoregional nutrient criteria (USEPA, 2000) address mainly cultural eutrophication. The Clean Water Act states that all waters should be able to provide for recreation and the protection and propagation of aquatic life. EPA sets water quality standards to protect the nation’s waterbodies. EPA’s water quality standards have three elements: designated uses, water quality criteria and antidegradation policy.

Designated uses include, but are not limited to, drinking water supply, fish consumption, ground water recharge, wildlife habitat, shellfish harvesting and agriculture. Each designated use has its own set of water quality criteria that must be met for the use to be realized.

Water quality criteria may be either numeric or narrative. Numeric criteria are used to establish thresholds for physical conditions, chemical concentrations, and biological attributes required to support a beneficial or designated use. Narrative criteria describe, instead of enumerate, conditions that must be maintained to support a designated use. For example, a narrative criterion might be “Waters must be free of substances that are toxic to humans, aquatic life, and wildlife” (National Water Quality Inventory, EPA, 2000).

Antidegradation policies are narrative statements used to protect existing uses and to prevent waterbodies from deteriorating, even if their water quality is better than the “fishable” and “swimmable” goals of the Act (National Water Quality Inventory, EPA, 2000).

In 2000, EPA assessed 43% of the lakes in the United States. Of these, 45% were declared impaired and 55% were declared unimpaired (Figure 1). According to the EPA
report, “The Quality of Our Nation’s Waters”, nutrients are the most common pollutants affecting assessed lakes. Nutrient impairment is found in 22% of the assessed lakes and contributes to 50% of reported water quality problems in impaired lakes (Figure 2(a)).

**Figure 1:** Percentage of lakes assessed. Source: National Water Quality Inventory, 2000 Report, EPA.¹

**Sources of Nutrient Pollution**

There are many sources of nutrient pollution including agriculture, urbanization, hydromodification, and urban runoff and storm sewers (Figure 2(b)). According to EPA, agriculture is the leading source of pollution in assessed lakes. Agricultural pollution problems affect 18% of the assessed lakes and contribute to 41% of reported water quality problems in impaired lakes. The main factor that causes increased transport of pollutants in agriculture is disturbing the soil by tillage. This greatly increases sediment loss compared to undisturbed soils. As much as 90% of nutrient loss (phosphorous and nitrogen) are associated with this sediment loss (Alberts, *et al* 1978). Nutrient losses represent only a small percentage of applied fertilizers, but their addition as run-off into water bodies greatly increases the effects of eutrophication. The resultant accumulation in

---

¹ Copy of permission in Appendix A
Figure 2: The bar charts, (a) and (b), present the leading sources and the number of lake, reservoir, and pond acres impacted. The percent scales on the upper and lower x-axes of the bar chart provide different perspectives on the magnitude of the impact of these sources. The lower axis compares the acres impacted by the source to the total assessed acres. The upper axis compares the acres impacted by the source to the total impaired acres.

Source: National Water Quality Inventory, 2000 Report, EPA.²

---

² Copy of permission in Appendix A
surface waters can cause fishkills (like the recent fish kills in the university lakes (Advocate, 2003).

Urbanization affects 8% of the assessed lake acres and 18% of the impaired lake acres (EPA, 2000). It is considered to have caused the most adverse change in water quality (Novotny and Olem, 1994). It modifies atmospheric composition, the hydrology of the watershed, receiving streams and other waterbodies, and soil. Urbanization has increased emission of wastes from a variety of sources like industries, households, transportation, sewage conveyance and disposal (landfills and incinerators). Increased imperviousness of soils decreases the capacity of the soil to store runoff water and this tends to make runoff peak levels higher. Urbanization of watersheds increases imperviousness of the soils making the surface flows peak at higher levels and also increases the volume (Novotny and Olem, 1994).

Examples of hydromodification are flow regulation and modification, channelization, dredging and construction of dams (which mainly affects rivers). This modification changes the natural habitat in such a way that it can no longer support fish and other desired flora and fauna.

In unsewered urban development, sewage is usually disposed into soils (e.g., septic tanks). When the adsorption capacity of this soil disposal system is exhausted, nutrients enter ground water. But during storm runoff, the hydraulic load exceeds the infiltration capacity and surface waters are contaminated by the sewage.

**Limiting Factors in Eutrophication**

It has been found, through various experiments, that the limiting factors in primary production are nitrogen (Ryther and Dunstan, 1971) and phosphorous (Schindler,
Generally, phosphorous is the limiting factor in fresh water systems and nitrogen in marine systems (Laws, 1981). Phosphorous becomes the limiting factor because many species of phytoplankton (e.g., blue-green algae) are capable of fixing atmospheric nitrogen and so may compensate for nitrogen deficiency in the water. This is not possible with phosphorous. All essential phosphorous has to come from outside inputs or sediment recycling in the water body. Also, the negative phosphate ion (PO$_4^{3-}$), forms insoluble compounds with many positive ions like Al$^{3+}$, Ca$^{2+}$ and Fe$^{3+}$, the main compound being ferric phosphate. These sink to the bottom as precipitates, trapping phosphate in the sediments, and hence make this phosphorous unavailable to the phytoplankton. In marine systems, iron concentration is much lower than in fresh water and therefore precipitation of ferric phosphate is not important in marine phosphorous cycles. Also, blue-green algae are a small proportion of the phytoplankton population of marine systems, so nitrogen fixation becomes less important.
CHAPTER 3. ESTABLISHING NUTRIENT CRITERIA

Trophic state variables comprise measures of nutrient concentration (like total phosphorous, soluble reactive phosphorus, total nitrogen, total Kjeldahl nitrogen), plant (macrophyte or algal) biomass (e.g., organic carbon, chlorophyll \(a\), Secchi depth), and watershed attributes like land use (USEPA, 2000). All of these could be used to establish criteria to deal with eutrophication concerns, but only a few are feasible as candidates for early warning variables. The factor that limits plant biomass may change seasonally or over longer periods of time, vary depending on the land use, or vary regionally. So, it does not make sense to construct a single nutrient criterion when that nutrient may not necessarily limit a target lake or lakes. This is why EPA emphasizes the development of nutrient criteria based on both the nutrient inputs (cause) and the biological response (effect).

The EPA considers total phosphorous and total nitrogen as the two causal variables and chlorophyll \(a\) and Secchi depth as the two early indicator response variables among other variables like dissolved oxygen, macrophyte, benthic algal growth or speciation, and other fauna and flora changes. The causal variables (phosphorus and nitrogen) are necessary criteria because they will be the limits required to establish management objectives and are usually directly related to discharge runoff abatement efforts by the states. Dissolved oxygen is also an important parameter to be considered. Dissolved oxygen is necessary for protecting aquatic life. This is especially important in the case of fishes because different species of fishes have different oxygen tolerance levels. However, nutrients have a marked but indirect effect on dissolved oxygen. Increased levels of nutrients affect the dissolved oxygen balance by increased growth of
flora and decomposing biomass. Also, dissolved oxygen levels vary diurnally, and this important variability is not likely to show up in monthly observations. In this study, dissolved oxygen has not been considered.

Different forms of phosphorus can be measured to determine trophic state. Of these, total phosphorus (TP) is a measure of all forms of dissolved or particulate phosphorus in a sample. TP concentrations in runoff or areal exports can also be readily related to watershed land use (Reckhow and Simpson, 1980; Walker, 1985a). This makes it a superior variable for addressing point and nonpoint source loads from the watershed. This is why TP has been used throughout North America as a basis for setting trophic state criteria and in developing related models (NALMS, 1992), and the reason why it was chosen as a causal variable in this thesis.

Control of nitrogen sources is more difficult than phosphorous because nitrogen can be assimilated directly from the atmosphere by several types of organisms, including some species of Cyanophyta (blue-green algae). Nitrogen is not as often limiting to plant growth, thus the focus on phosphorous as the major factor considered in eutrophication.

The most common forms of nitrogen that are of concern in eutrophication evaluation are nitrite, nitrate, ammonia, and organic nitrogen (as measured by total Kjeldahl nitrogen (TKN)). Total nitrogen (TN) is considered to be the sum of ammonia, nitrate, nitrite, and TKN. Usually, nitrate, nitrite, and ammonia are present at very low levels in lakes or reservoirs unless there are some relatively recent loadings in runoff from the watershed, or if nitrogen is not the limiting factor of algal growth in that particular water body. These forms are rapidly used by algae and aquatic plants or
converted to other forms of nitrogen. The most useful measurement from a modeling standpoint is either TN or TKN (USEPA, 2000).

Chlorophyll $a$ is the major photosynthetic pigment in plants. It is an important variable when one wants to estimate the photosynthetic capacity of an ecosystem. It is a surrogate measure of algal density, which is costlier to measure. Therefore, it is the chosen variable when an estimate of the primary productivity of an ecosystem is required. Chlorophyll $a$ is also preferred as an indicator because there are lakes where TP is not the sole or primary limiter of algal production or biomass, for example, lakes with high inorganic turbidity or high flushing rates (USEPA, 2000). The relationship between chlorophyll and phosphorus and its linkage to algal biomass, makes chlorophyll $a$ major component of trophic state indices (Carlson, 1977) and water quality criteria.

In addition to the use of chlorophyll $a$ in classification, the chlorophyll interval frequency, or bloom frequency, have been predicted based on regression equations developed by Walker (1985b) relating blooms to phosphorus. These chlorophyll $a$ intervals can be related to varying user perceptions of lake condition. The projected frequency of these extreme events, as a result of increased phosphorus loading, can be readily understood by citizens and decision-makers (Heiskary and Walker, 1988).

EPA encourages the development of mechanistic or empirical models for identification of overenrichment problems, management planning, and determination of status and trends of water resources. The causal and biological and physical response variables represent only a set of starting points for States and Tribes to use in establishing their own criteria. This is because control of causal variables would help to protect uses before impairment occurs and to maintain downstream uses. Early response variables
would warn of possible impairment and help to integrate the effects of variable and potentially unmeasured nutrient loads.
CHAPTER 4. DEVELOPMENT OF ECOREGIONS

The establishment of a single, national nutrient criteria for lakes is not a sensible goal when one considers the significant variability of water bodies that exists across the country in a variety of climates, geographic locations, and ecosystems. Individual lakes and reservoirs are affected by varying degrees of development, and user perceptions of water quality throughout the country can differ even over small distances (USEPA, 2000). Consequently, EPA bases its nutrient criteria development process on an approach that takes into account the geographic differences in lakes across the country and uses a classification system to explain those differences. The initial classification scheme used by EPA is the ecoregion approach (Omernik, 1987, 1988, 1995).

EPA defines lakes as natural and artificial impoundments with a surface area greater than 10 acres and a mean water residence time of 14 or more days. Man-made lakes with the same characteristics are viewed as part of the same system. Reservoirs are man-made lakes for which the primary purpose of the impoundment is other than recreation (e.g., boating, swimming) or fishing, and the water retention time and water body depth and volume vary widely. This definition of lakes has been used by EPA for collecting data to set reference conditions for water quality in lakes.

EPA identified geographic divisions as part of a hierarchical classification procedure with the purpose of grouping similar lakes together. Classification of lakes was used in order to reduce the variability of lake-related measures (e.g., physical, biological, or water quality variables) within classes and maximize the variability among classes. This helps to group lakes together that under ideal conditions would have similar characteristics (e.g., biological, ecological, physical). Classification was restricted to
those characteristics of lakes that are intrinsic, or natural, and are not the result of human activities. Measures like size, maximum or mean depth, detention time, in lake phosphorous and nitrogen, and shape are incorporated.

Ecoregions are a mapped classification system of ecological regions, that is, regions with assumed relative homogeneity of ecological characteristics (Omernik, 1987). EPA has developed maps of ecoregions of the United States at various levels of resolution and aggregation (Omernik, 1987). The most commonly used is the Level III ecoregions, consisting of 79 ecoregions in the conterminous United States. Ecoregions were based on analysis of the spatial coincidence in all geographic phenomena that are the source of or indicate differences in ecosystem patterns. These phenomena consist of geology, physiography, vegetation, climate, soils, land use, wildlife, and hydrology. The relative importance of each characteristic varies among ecoregions regardless of the hierarchical level (USEPA, 2000).

Level III ecoregions were aggregated to describe broad areas, which are generally comparable in quality and types of ecosystems as well as in natural and anthropogenic characteristics that have an effect on nutrients. A map of these ecoregion aggregations was made for the National Nutrient Criteria Program (USEPA, 2000). This aggregation resulted in fourteen ecoregions (Figure 3). The regions are meant to furnish a geographic framework for guidance and reporting for the National Nutrient Criteria Program and can form the basis for initial development of nutrient criteria.

**Formation of Nutrient Criteria Database**

The Nutrient Criteria Database contains data from STORET, the National Eutrophication Survey (NES), the National Surface Water Survey (NSWS), the
Figure 3: Fourteen nutrient Ecoregions as delineated by Omernik (2000). Ecoregions were based on geology, land use, ecosystem type, and nutrient conditions. Source: EPA.\(^3\)

\(^3\) Copy of permission in Appendix A
Environmental Monitoring and Assessment Program (EMAP), the Clean Lakes Program, Volunteer Monitoring Programs, State Monitoring Programs, the U.S. Army Corps of Engineers and other sources.

Dealing with Quality of Historical Data

The quality of older historical data sets is usually a problem because the data quality is often unknown. This is because objectives, methods, and investigators may have changed many times over the years. The most reliable data are those collected by a single agency using the same protocol for a limited number of years. When “mining” from large heterogeneous data repositories such as STORET, EPA investigators screened data for acceptance considering a number of factors like location, variables and analytical methods, laboratory quality control, collecting agencies, time period, index period and representativeness.

Location

STORET data are georeferenced. These data can be used to select specific locations or specific USGS hydrologic units. For selection of lakes within a geographic region, it is important to know the underlying principle and methods of site selection by the original investigators. This information may be included in STORET metadata.

Variables and Analytical Methods

Thousands of variables are recorded in STORET records. Each separate analytical method yields a unique variable. Methods differ in accuracy, precision, and detection limits, so, it is generally not sensible to mix methods in the same analysis. According to EPA, selection of a particular “best” method may result in very few observations, so it suggests that it may be prudent to select the most frequently used method in the database.
Laboratory Quality Control

Laboratory quality control data (blanks, spikes, replicates, known standards, etc.) are normally not accounted in the larger data repositories. EPA suggests that it is more cost-effective to accept or reject all data of the collecting agency or laboratory based on overall confidence of their quality control. Sometimes, eliminating lower quality data can be counterproductive, because the increase in variance caused by analytical laboratory error may be negligible compared with natural variability or sampling error.

Collecting Agencies

STORET data identifies the agency that collected the data. Selecting data only from particular agencies with known, reliable collection and analytical methods and accepted quality reduces inconsistency due to unidentified quality problems.

Time Period

Long-term records are vitally important for detecting and establishing trends. While defining reference conditions for nutrient criteria, it is important to determine if trends exist in the reference site database. For example, over time, many lakes may have improved markedly while other lakes, exposed to increased nonpoint-source runoff, may have declined in overall quality.

Index Period

An index period for approximating average concentrations should be designated if nutrient and water quality variables were measured more than once a year. The index period could represent the entire year, spring or fall mixing or the summer growing season. The most suitable index period can be determined by investigators who should consider the characteristics of the lakes of the region, the quality and quantity of data

**Representativeness**

Historical data may have been collected for specific purposes, such as developing nutrient budgets for eutrophic lakes. These data are not likely to be characteristic of the type of region or lake of interest. The investigator has to decide whether the lakes in the database are representative of the population of lakes to be characterized. If a sufficient sample of representative lakes (i.e., one large enough to characterize reference conditions) cannot be found, a new survey will be necessary (Nutrient Criteria Technical Guidance Manual, EPA, 2000).
CHAPTER 5. MATERIALS AND METHODS

Data for the study were obtained from the Nutrient Criteria Database of the EPA. The data were collected by state. Datasets from each state were then formatted in Access. These were then merged to form a single dataset and then sorted by ecoregion. SAS and S-PLUS were used in the analysis of this single dataset.

The Nutrient Criteria Database has observations for a large number of parameters (Appendix G). Of these, the parameters of choice were Chlorophyll \( a \) Fluorometric corrected (CHLA, ug/l, STORET code 32209) and Chlorophyll \( a \) Trichromatic uncorrected, (CHLAttri, ug/l, STORET code 32210), total phosphorous (TP, ug/l, STORET code 00665), total nitrogen (TN, mg/l, STORET code 00600) and Kjeldahl nitrogen (TKN, mg/l, STORET code 00625). These parameters were chosen because they were also used by EPA to formulate reference conditions for lakes.

This dataset has 593,650 observations. TP is present in only 324,325 samples, TN in 163,838 samples, TKN in 91400 samples, CHLA in 15,816 samples, and CHLAttri in 79,572 samples. Since, in a linear model observations of all variables have to be present simultaneously, only 93,894 samples could be used in the study. The ecoregions did not have equal number of samples (Appendix G).

It was found that only three ecoregions (2, 7, and 8) had observations for total nitrogen and total phosphorous with corresponding chlorophyll \( a \) (STORET code 32209) observations (Appendix C). Also, only three ecoregions (9, 12, and 13) had observations for total nitrogen and total phosphorous with corresponding chlorophyll \( a \) (STORET code 32210) observations (Appendix C). Total nitrogen consists of Kjeldahl nitrogen, nitrate
and nitrite. So, total nitrogen was regressed with total Kjeldahl nitrogen to find the extent to which total nitrogen can be predicted using Kjeldahl nitrogen.

It was found that Kjeldahl nitrogen can account for 94.58% of total nitrogen variability. Separate regressions were fit for each of the ecoregions mentioned above. The equations used are given in Table 1. This was done in the above five ecoregions because only in these was TN measured concurrently with TKN. Such a strong correlation prompted the creation of a new parameter named “newTN”. This included all the actual total nitrogen measurements along with the total nitrogen predicted from TKN in the cases where TN was missing but TKN was present. This allowed an increase in the number of observations and an increase in the spatial coverage. The relationship between chlorophyll $a$ with total phosphorous and total nitrogen is fit using log-log regression models to stabilize the variance (Lamon, 1995). This is a common procedure in many research fields and is usually the first analytical step (Hamilton, 1992; Reckhow, 1988).

**Table 1**: Regression equations for TN

<table>
<thead>
<tr>
<th>Ecoregion</th>
<th>Equation</th>
<th>(Standard Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$\log \text{TN} = -0.445172475 + 0.766388167*(\log \text{TKN})$</td>
<td>(0.077) (0.095)</td>
</tr>
<tr>
<td>7</td>
<td>$\log \text{TN} = 0.000667036 + 0.991866474*(\log \text{TKN})$</td>
<td>(0.004) (0.019)</td>
</tr>
<tr>
<td>11</td>
<td>$\log \text{TN} = 0.047977207 + 1.0255552479*(\log \text{TKN})$</td>
<td>(0.010) (0.029)</td>
</tr>
<tr>
<td>12</td>
<td>$\log \text{TN} = 0.015506943 + 0.958164113*(\log \text{TKN})$</td>
<td>(0.001) (0.003)</td>
</tr>
<tr>
<td>13</td>
<td>$\log \text{TN} = 0.026786063 + 0.997622084*(\log \text{TKN})$</td>
<td>(0.001) (0.004)</td>
</tr>
</tbody>
</table>

As with any linear model, the observations to be analyzed had to have all three variables as non-missing. Also, it was found that all the variables were not available for
all the ecoregions. Choosing just CHLA, TP and TN covered very few ecoregions (as mentioned above). So, various combinations of the five parameters stated above were made to find which combination of the three parameters (response and two predictors) had the most observations, the most spatial coverage, and the best fit.

The combinations of the five parameters produced six datasets. These datasets were analyzed, using regression, to find the best model defined by the three criteria (model fit, number of observations and spatial coverage). The six models for these datasets are given in Table 2.

The null hypothesis for these models is:

\[ H_0 : \text{There is no significant variation in chlorophyll } a \text{ due to the factors ecoregion, total phosphorous and total nitrogen. In other words, } \]

\[ H_0 : \beta_0 = \beta_1 = \beta_2 = 0. \]

The alternative hypothesis or \( H_a \) is that at least one of the \( \beta \)'s is non zero, or, there is some variation in chlorophyll \( a \) due to ecoregion, total phosphorous and total nitrogen.

The general linear model used in this study was conducted by using ecoregion as a class variable. Ecoregion was used in an interaction term in the model with the predictors TP and TN (Lamon, 1995). The model is of the form:

\[ \log_{10} \text{CHL}a = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon, \]

where, CHL\( a \) = chlorophyll \( a \)

\( \beta_0 = \) ecoregion (treated here as a dummy variable)

\( \beta_1 = \) vector of slope estimates for phosphorous

\( \beta_2 = \) vector of slope estimates for nitrogen

\( x_1 = \) interaction of the ecoregion dummy variable with log\( _{10} \) total phosphorous
$x_2 =$ interaction of the ecoregion dummy variable with $\log_{10}$ total nitrogen

$\varepsilon =$ residual

A dummy variable is a numerical variable (usually 0 and 1) used in a regression analysis to represent subgroups of a sample. In this study, ecoregion has been used as a dummy variable because it helps in using a single regression equation to represent different ecoregions in each model.

In regression, a simple fit statistic, the coefficient of determination ($R^2$) gives the proportion of the variance of $Y$ explained by regression on $X$. $R^2$ is the explained variance divided by total variance or 1- (residual variance/total variance). It ranges from zero to one. Zero means no variance is explained, and one means hundred percent of the variance is explained.

**Table 2**: Total number of data and spatial spread of models.

<table>
<thead>
<tr>
<th>Models</th>
<th>Total (N)</th>
<th>#</th>
<th>Ecoregions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1 $\log CHLA = log TP + log TN$</td>
<td>4246</td>
<td>3</td>
<td>2, 7, 8</td>
</tr>
<tr>
<td>Model 2 $\log CHLA = log TP + log TKN$</td>
<td>5069</td>
<td>7</td>
<td>1, 2, 7, 8, 9, 11, 14</td>
</tr>
<tr>
<td>Model 3 $\log CHLA = log TP + log new TN$</td>
<td>4817</td>
<td>4</td>
<td>2, 7, 8, 11</td>
</tr>
<tr>
<td>Model 4 $\log CHLA_{tri} = log TP + log TN$</td>
<td>49301</td>
<td>3</td>
<td>9, 12, 13</td>
</tr>
<tr>
<td>Model 5 $\log CHLA_{tri} = log TP + log TKN$</td>
<td>20681</td>
<td>10</td>
<td>2, 3, 6, 7, 8, 9, 10, 11, 12, 13</td>
</tr>
<tr>
<td>Model 6 $\log CHLA_{tri} = log TP + log new TN$</td>
<td>62475</td>
<td>6</td>
<td>2, 7, 9, 11, 12, 13</td>
</tr>
</tbody>
</table>

Ordinary Least Squares (OLS) was used for regression analysis because it is one of the most frequently used techniques for such analysis. This produces coefficient estimates that minimize the sum of squared residuals. The solution obtained through OLS assures “best fit” to the data. OLS works under some assumptions:

1. Fixed $X$: Theoretically, we can obtain many random samples, each with same $X$ values but different $Y_i$ due to different residual values.
2. Errors have zero mean. These two assumptions together ensure independence of errors and X variables. This is sufficient for unbiased estimation of all parameters.

3. Errors have constant variance (homoscedasticity). Heteroscedasticity leads to inefficiency and biased standard error estimates.

4. Errors are uncorrelated with each other. (No autocorrelation).

5. Errors are normally distributed. Nonnormal errors increase inefficiency and undermine the rationale for t-tests and F-tests, especially with small samples. Also, since OLS tends to track outliers, heavy-tailed error distributions can cause great sample-to-sample variation.

**Regression Criticism**

Residuals versus predicted plots provide a starting point for criticism in multiple regression. These may uncover problems such as curvilinearity, heteroscedasticity, nonnormality, or outliers.

The linear model assumptions (i.i.d.) imply certain characteristics of the model error term (the model residuals):

- Errors have identical distributions, with zero mean and same variance, for every value of x.
- Errors are independent. They are unrelated to other errors, variables or cases.
- Errors are normally distributed.

Some possible violation of assumptions of regression are:

- Nonlinear relationships: Ordinary least squares (OLS) finds the best fitting straight line, but this would be misleading if Y is actually a nonlinear function of X.
- Nonconstant error variance: i.e., heteroscedasticity, when present, can lead to misrepresentation of findings (over/under estimation) and thereby increase the possibility of a Type I error.

- Correlation among errors: The standard errors, tests, and confidence intervals assume independence among errors. But this might not always be the case, especially when observations are adjacent in time and space. Correlation among the errors could mean that there are effectively fewer observations than the original dataset and this can lead to overestimation of the standard error and the confidence intervals.

- Nonnormal errors: $F$ and $t$ tests assume normal distribution of errors. Nonnormal errors may nullify these tests and increase sample-to-sample variation of the estimates.

- Influential cases: OLS regression is not resistant; a single influential observation can pull the regression line up or down and substantially influence all results.

**Residuals Used**

In the Jackknife procedure, n-1 cases of the dataset are used for calibration and one case is used for confirmation of normality. This procedure is run n times, so each case is used as a confirming case. If the confirmation is successful according to a statistical goodness of fit criterion, then all of the data are used for the final calibration. In this procedure, the model and parameter errors are better characterized because these errors may be estimated by the net result of the “one-case-at-a-time” confirmation (Rechow and Chapra, 1983).
A Studentized residual is a residual which has been divided by its estimated standard error. This standard error is based upon fitting a statistical model using all points except the point whose residual is to be computed. It is assumed that the residuals are normally distributed, so these values approximately follow a $t$ distribution, where for large samples about 65% are between -1 and +1, about 95% are between -2 and +2, and about 99% are between -2.6 and +2.6.

**Residual Analysis**

Testing regression assumptions focuses on errors (residuals) to measure their plausibility. Residuals versus predicted values are a general-purpose diagnostic tool. These plots can show an influential case, a curvilinear relation, a nonnormal residual distribution and heteroscedasticity. Mean-median comparison provides a simple check for symmetry, a necessary condition for normality. The mean OLS residual always equals zero or is very close to zero. A non-zero median would then imply skewness.

The residuals of each model were also checked to ensure that they met the criteria for other linear model assumptions (normal i.i.d).

Compliance with these assumptions is determined by graphical methods (e.g., plots of residuals versus predicted values) and statistical tests (e.g., Kolmogorov-Smirnov test) applied to the model residuals. The sum of squared residuals (RSS) reflects the overall accuracy of predictions. The lower the RSS, the closer the fit. Another indication of goodness of fit is the residual standard deviation, $s_e$. This measures scatter or spread around a regression line.

The Kolmogorov-Smirnov test is a nonparametric test for goodness of fit. It does not assume a normally distributed population. It compares the empirical distribution of a
variable with a normal distribution that has the same mean and variance as the empirical distribution. The null hypothesis here is that the parameter has a standard normal distribution. This test is very sensitive to violations of normality in the tails.

Cooks distance ($D_i$) is a measure of influence. It measures influence on the whole model. The value of $D_i$ is an indication of the $i$th case’s influence on all the estimated regression coefficients. Two criteria for $D_i$ are:

- An observation is influential if $D_i > 1$.
- An observation is extremely influential if $D_i > 4/n$.

**Exploratory Data Analysis Techniques**

Boxplots are useful for a comparison between two or more distributions. They provide information about the central tendency, given by the median, dispersion, which is given by the interquartile range, symmetry and outliers (Appendix D).

A symmetry plot graphs the distance of the $i$th value of a dataset above the median to the $i$th value below the median. Each pair of observation defines one point on the plot. A symmetry plot of $n$ data contains $n/2$ points. If the distribution were symmetrical, the points would lie on one line. Points that do not lie on the line indicate asymmetry. Positive skew is indicated if points lie above the line and negative skew is indicated if points lie below the line.

Quantile-quantile plots are used to graph quantiles of one variable against another variable. This can be used to compare two empirical distributions, or an empirical distribution with a theoretical distribution (Hamilton, 1992). The latter is called a quantile–normal plot or normal probability plot when variables are plotted against a Gaussian
distribution that has the same mean and standard deviation as the sample. Quantile-normal plots of sample residuals can also help detect nonnormality.
CHAPTER 6. RESULTS

In all the six models, the $p$-value for the $F$-statistic is less than 0.0001 (Appendix B), so we reject the null hypothesis and can say that at least one of the $\hat{\alpha}$s is non zero, or, there is some variation in chlorophyll $a$ explained by regression on total phosphorous and total nitrogen.

Model 1 includes observations from only three ecoregions (2, 7, and 8). It explains about 28% of the variation in chlorophyll $a$ (Appendix B). The interaction effect is significant ($p < 0.0001$), i.e., at least two of the ecoregions have different slopes or intercepts. Tests for normality (Kolmogorov-Smirnov test) on the residuals show that as a whole, the model residuals are not normal ($p < 0.01$), as also seen in the difference between the mean (0.0001) and the median (0.043). Visually, the summary plots (Figure 4) show a high degree of normality, except in the tails. The scatter plots (Figure 5) of the predicted variable, chlorophyll $a$, versus the residuals, also show a random distribution. Therefore, considering the number of observations (4246) used in the model, and the fact that the Kolmogorov-Smirnov test is very sensitive in the tails, we can say that the distribution is approximately normal. The Cook’s distance graph (Figure 6) also shows that the values are well below 1, so there are no unduly influential observations.

Model 2 includes observations from seven ecoregions (1, 2, 7, 8, 9, 11 and 14). It explains about 23% of the variability in chlorophyll $a$ (Appendix B). The interaction effect is significant ($p < 0.0001$), i.e., at least two of the ecoregions have different slopes or intercepts. Tests for normality (Kolmogorov-Smirnov test) on the residuals show that as a whole, the model residuals are not normal ($p < 0.01$), as also seen in the difference between the mean (-0.00008) and the median (0.13). Visually, the summary
**Figure 4**: Summary plots of the residuals for Model 1.

**Figure 5**: Predicted chlorophyll \( a \) vs. residual plot for Model 1.
plots (Figure 7) show a high degree of normality, except in the tails. There is a slight negative skew. The scatter plots (Figure 8) of the predicted variable, chlorophyll \( a \), versus the residuals, also show a random distribution. Therefore, considering the number of observations (5069) used in the model, and the fact that the Kolmogorov-Smirnov test is very sensitive in the tails, we can say that the distribution is approximately normal. The Cook’s distance graph (Figure 9) also shows that the values are well below 1, so there are no unduly influential observations.

Model 3 includes observations from four ecoregions (2, 7, 8 and 11). It explains about 27% of the variability in chlorophyll \( a \) (Appendix B). The interaction effect is significant (\( p < 0.0001 \)), i.e., at least two of the ecoregions have different slopes or intercepts. Tests for normality (Kolmogorov-Smirnov test) on the residuals show that as a whole, the model residuals are not normal (\( p < 0.01 \)), as also seen in the difference
Figure 7: Summary plots of the residuals for Model 2.

Figure 8: Predicted chlorophyll $a$ vs. residual plot for Model 2.
Figure 9: Cook’s Distance plot for Model 2.

between the mean (0.0001) and the median (0.030). Visually, the summary plots (Figure 10) show a high degree of normality, except in the tails. The scatter plots (Figure 11) of the predicted variable, chlorophyll \(a\), versus the residuals, also show a random distribution. Therefore, considering the number of observations (4817) used in the model, and the fact that the Kolmogorov-Smirnov test is very sensitive in the tails, we can say that the distribution is approximately normal. The Cook’s distance graph (Figure 12) also shows that the values are well below 1, so there are no unduly influential observations.

Model 4 includes observations from 3 ecoregions (9, 12 and 13). It explains about 60% of the variability in chlorophyll \(a\) (Appendix B). The interaction effect is significant (p< 0.0001), i.e., at least two of the ecoregions have different slopes or intercepts. Tests for normality (Kolmogorov-Smirnov test) on the residuals show that as a whole, the model residuals are not normal (p<0.01), as also seen in the difference between the mean (-0.00001) and the median (0.08). Visually, the summary plots (Figure 13) show a high
Figure 10: Summary plots of the residuals for Model 3.

Figure 11: Predicted chlorophyll $a$ vs. residual plot for Model 3.
degree of normality, except in the tails. The scatter plots (Figure 14) of the predicted variable, chlorophyll $a$, versus the residuals, also show a random distribution. Therefore, considering the number of observations (49301) used in the model, and the fact that the Kolmogorov-Smirnov test is very sensitive in the tails, we can say that the distribution is approximately normal. The Cook’s distance graph (Figure 15) also shows that the values are well below 1, so there are no unduly influential observations.

Model 5 includes observations from 10 ecoregions (6, 7, 8, 9, 10, 11, 12 and 13). It explains about 48% of the variability in chlorophyll $a$ (Appendix B). The interaction effect is significant ($p<0.0001$), i.e., at least two of the ecoregions have different slopes or intercepts. Tests for normality (Kolmogorov-Smirnov test) on the residuals show that as a whole, the model residuals are not normal ($p<0.01$), as also seen in the difference between the mean (-0.00002) and the median (0.11). Visually, the summary plots (Figure 16) show a high degree of normality, except in the tails. The scatter plots (Figure 17) of
the predicted variable, chlorophyll \(a\), versus the residuals, also show a random distribution. Therefore, considering the number of observations (20,681) used in the model, and the fact that the Kolmogorov-Smirnov test is very sensitive in the tails, we can say that the distribution is approximately normal. The Cook’s distance graph (Figure 18) also shows that the values are well below 1, so there are no unduly influential observations.

![Boxplot](image)

**Figure 13**: Summary plots of the residuals for Model 4.

Model 6 includes observations from 6 ecoregions (2, 7, 9, 11, 12 and 13). It explains about 58% of the variability in chlorophyll \(a\) (Appendix B). The interaction effect is significant \((p<0.0001)\), i.e., at least two of the ecoregions have different slopes or intercepts. Tests for normality (Kolmogorov-Smirnov test) on the residuals show that as a whole, the model residuals are not normal \((p<0.01)\), as also seen in the difference between the mean (-0.00001) and the median (0.08). Visually, the summary plots (Figure
19) show a high degree of normality, except in the tails. The scatter plots (Figure 20) of the predicted variable, chlorophyll $a$, versus the residuals, also show a random distribution. Therefore, considering the number of observations (62,475) used in the model, and the fact that the Kolmogorov-Smirnov test is very sensitive in the tails, we can say that the distribution is approximately normal. The Cook’s distance graph (Figure 21) also shows that the values are well below 1, so there are no influential observations.

Models 1, 2, and 3 represent many common ecoregions (Table 3). Model 1 has a low $R^2$ and a small N compared to models 4, 5 and 6. Model 2 has an even lower $R^2$ and larger root mean square error, and the number of observations representing each ecoregion in this model is much less than to Model 1 and 3. Model 3 has a slightly lower $R^2$ than Model 1. Addition of the predicted TN (model 3) adds another ecoregion that is represented by this model and also increases the proportion of TN that can explain variation in chlorophyll $a$. So, based on $R^2$ and root mean square error, Model 3 is a better model than Model 1 or 2.

![Figure 14: Predicted chlorophyll $a$ vs. residual plot for Model 4.](image)
Figure 15: Cook’s Distance plot for Model 4.

Figure 16: Summary plots of the residuals for Model 5.
Figure 17: Predicted chlorophyll $a$ vs. residual plot for Model 5.

Figure 18: Cook’s Distance plot for Model 5.
Figure 19: Summary plots of the residuals for Model 6.

Figure 20: Predicted chlorophyll $a$ vs. residual plot for Model 6.
**Figure 21**: Cook’s Distance plot for Model 6.

**Table 3**: Comparison of the six models.

<table>
<thead>
<tr>
<th>Model</th>
<th>$R^2$</th>
<th>Root MSE</th>
<th>logCHLA/tri (mean)</th>
<th>No. of Observations (N)</th>
<th>No. of ecoregions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.280531</td>
<td>0.323859</td>
<td>0.585787</td>
<td>4246</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>0.230490</td>
<td>0.492696</td>
<td>0.848403</td>
<td>5069</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>0.274044</td>
<td>0.344937</td>
<td>0.574107</td>
<td>4817</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>0.607496</td>
<td>0.336146</td>
<td>0.967883</td>
<td>49301</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>0.480209</td>
<td>0.365508</td>
<td>1.146313</td>
<td>20681</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>0.583141</td>
<td>0.347155</td>
<td>0.994776</td>
<td>62475</td>
<td>6</td>
</tr>
</tbody>
</table>

Model 4 has the largest $R^2$ but it represents observations from only 3 ecoregions (9, 12 and 13). As seen in the map (Appendix E), data are restricted to Florida. Model 5 has a slightly lower $R^2$ and a slightly higher root mean square error compared to Model 4, but it has the largest spatial coverage (10 ecoregions). Of these 10 ecoregions, 3 ecoregions (2, 3 and 10) are under-represented in terms of number of observations (Appendix C). Model 6 has an $R^2$, a root mean square error and a spatial coverage of
ecoregions in between models 4 and 5. For models 4, 5 and 6, observations are dominated by ecoregions 12 and 13.

The $R^2$, number of observations, and number of ecoregions covered by each model were ranked from highest to lowest (Table 4). Weights were assigned to them. $R^2$ and number of observations were given equal weights of 0.25. The number of ecoregions was assigned a weight of 0.5. This was given twice the weight of the others because it will be useful for EPA to use a model that has the most coverage by ecoregion to set nutrient criteria standards for the ecoregions so that these can be used as a guide by States and Tribes to set their own standards.

**Table 4: Weighted ranks of each model.**

<table>
<thead>
<tr>
<th>Model</th>
<th>Rank of $R^2$ (assigned weight = 0.25)</th>
<th>Rank of Number of Observations (assigned weight = 0.25)</th>
<th>Rank of Number of Ecoregions (assigned weight = 0.5)</th>
<th>Weighted Average Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>3.5</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4.5</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3.25</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2.0</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2.25</td>
</tr>
</tbody>
</table>

The weighted average rank for Model 5 ranks the lowest (2.0) and seems to be the best model among the six models because it has a moderately large $R^2$ (third of six), it can be used to predict chlorophyll $a$ in ten ecoregions (first of six), and has a large number of observations (third of six). If equal weights were assigned to each of the criteria, then model 6 ranks the highest.
CHAPTER 7. DISCUSSION

Statistical analysis and a deterministic approach are among the many approaches that can be used to predict a dependent variable with respect to changes in independent variables. In a deterministic approach, all pertinent knowledge about the independent variables is considered in the model to predict the change in the dependent variable. Results can be obtained as probabilities when uncertainties (like natural variability and sampling bias) are taken into account. But this cannot account for uncertainties that are unknown or less understood (e.g., model specifications). In contrast, statistical analysis can be used to find confidence limits thus making an allowance for all these uncertainties. But such analysis requires sufficient data of adequate quality (Portielje and Van der Molen, 1999). Also, statistical analyses can at least quantify uncertainties, whereas the deterministic approach needs other approaches, like Monte Carlo simulation, to account for uncertainty. Regression analysis is a type of statistical analysis and has been used for determining chlorophyll \(a\) levels from phosphorous and nitrogen for a cross-sectional dataset (Reckhow, 1988). Using regression helps to explain some of the variability seen in chlorophyll \(a\) levels in the lakes and can thus provide valuable information for managing eutrophication.

Limitations of the Models

Scale should not be disregarded while making predictions from cross-sectional models (Jones et al, 1998). Data used to fit the models have come from individual samples taken from different monitoring stations in different lakes throughout an ecoregion. The analysis of the data has been done at the spatial scale of an ecoregion. The models can be treated as a basis for predicting chlorophyll \(a\) in individual lakes within an
ecoregion. But there may be spatial variability among the lakes within an ecoregion due to size, morphology, land use patterns, latitude and longitude etc., and scaling up from individual lakes to ecoregion might make the predicted chlorophyll $a$ unusable for a single lake within the ecoregion (i.e., the prediction uncertainty might be unacceptable for a single lake within an ecoregion). This uncertainty or risk would have to be quantified using a utility function (at the management level) to decide whether a model can be used for decision making purposes.

There are many factors that may affect the amount of chlorophyll $a$ in lakes, other than nitrogen and phosphorous. Land use patterns, seasonality, depth, hydraulic retention time, mixing, latitude (Dodds et al, 2002), zooplankton grazing, species organization of the algal community, higher trophic levels and distribution of submerged macrophytes (Portielje and Van der Molen, 1999) may all affect chlorophyll $a$. If these had been incorporated into the model, more of the variability in the chlorophyll $a$ may have been explained. These predictors were not included in the present study because these measurements would have to be concurrent with chlorophyll $a$ in the dataset (Appendix G) to be used in a regression. This was not the case in the Nutrient Criteria Database. Also, including all these predictors might decrease the number of observations. This was also the reason why Dodds et al (2002) could not construct more complex predictive regression models using multiple variables in their study of benthic algal biomass relations to nutrients in streams.

There is natural variability among lakes. Lakes in the same ecoregion might not have the same attributes because delineation of ecoregions is not solely based on lake attributes (e.g., size, morphology, geographic location, hydraulic residence time), but on
potential natural vegetation, physiography, soils and land use and land cover (Jennerette et al, 2002). Therefore, the deviation from normality in the tails may be due to heterogeneity in the data which could be a function of choosing ecoregion as a geographic division.

EPA has used the notion of ecoregions for determining nutrient criteria. Ecoregions have had little noticeable influence in this regard in Europe (Siep et al, 2000), but they have been shown to have some effect in the Southern Coastal Region of the United States (Dodds et al, 2002). The accuracy of prediction using models fitted to observations from a single lake and applied to that lake may be much higher than doing so for each ecoregion, but this would become very expensive. An alternative would be to delineate homogeneous groups of lakes (in terms of chlorophyll $a$, phosphorous and nitrogen), for example, further divisions of the ecoregions, for better management of water quality.

Most of the studies regarding the relation of chlorophyll $a$ to TP and TN have been done in individual lakes, steams and reservoirs (An and Park, 2002; Scasso et al, 2001; Perkins and Underwood, 2000; Attayde and Bozelli, 1998-1999; Holopainen and Letanskaya, 1999; Adams, 1998; Burkholder et al, 1998; Lamon et al, 1996; Walker and Havens, 1995). There are relatively few cross-sectional studies of lakes for the prediction of eutrophication endpoints. The cross-sectional studies that are present are limited to smaller regions (Reckhow, 1988). This cross-sectional study done by ecoregions for the whole of the continental U.S. can be treated as a step toward finding a suitable regional delineation within ecoregions or among them so that the model for that particular region can better predict chlorophyll $a$ levels as an aid to control eutrophication.
Of the 593,650 observations, the study used 93,894 observations. The ecoregions did not have equal number of samples. The variability in the number of observations in each ecoregion for each model ranged from 15 to more than 50,000 (Appendix C). Therefore, the method of sample collection could be improved by making sure that all observations are taken simultaneously, so that a larger dataset might be used for better prediction.

Other Approaches

There are instances when models show a quadratic relation of chlorophyll $a$ with TP, with chlorophyll $a$ reaching an asymptote at high TP. An and Park (2002) found that using a quadratic model explained about 45% more of the variation in their data than a linear relationship. Model 4, 5, and 6 in this study (Figures 14, 17 and 20) show some non-linearity. So, this might be another approach for finding the best model to use for prediction of chlorophyll $a$.

Another approach to dealing with non-linearity is a semiparametric model (Lamon and Clyde, 2000). This model includes the explanatory variables (e.g., TP and TN) as linear predictors or regression spline predictors, and this can account for nonlinear relationships. Bayesian model averaging was used by Lamon and Clyde (2000) for predictions that included uncertainty about inclusion of variables (model specifications).

The data in each ecoregion show great variation. Using breakpoint regression could help to find two linear relationships that might describe the highest proportion of variance in the ecoregion (Dodds et al, 2002). This could then be used to form smaller but more homogeneous regions with respect to lake type.
In a linear regression, the parameters are assumed to be constant for all observations. But in cross-sectional or time series data, the parameters might vary among the cross-sectional units or over time. This problem might be solved by explicitly modeling the unmodeled factors to make them predictable. But, there might be cases where this modeling will not be feasible (data scarce, unavailable, etc.). In this case a random coefficient regression model can be used where the parameter reflects the variance over cross-sectional data (Reckhow, 1992).

**Summary/Conclusion**

Section 304 of the Clean Water Act deals with a scientific assessment of ecological and human health effects recommended by EPA to the States and Tribes for establishing water quality standards. These serve as a basis for control of discharges or release of pollutants. The EPA has divided the continental United States into fourteen ecoregions to facilitate the management of water quality. EPA intends to use this scientific assessment to develop default Section 304(a) nutrient criteria for the all the ecoregions in the country. They have identified nutrient measures (like total phosphorous, soluble reactive phosphorus, total nitrogen, total Kjeldahl nitrogen, plant biomass and land use) that can be used to formulate standards for water quality. According to EPA, the main causal variables in establishing these standards are total phosphorous and total nitrogen, and the indicator response variables are chlorophyll \(a\) and Secchi depth.

In the nutrient criteria database, sampling for phosphorous, nitrogen and chlorophyll \(a\) have not been done following the same methods for the same variable. So, all the samples cannot be used for all ecoregions to fit a linear model. A regression analysis of several combinations of TP, nitrogen and chlorophyll \(a\) method types was
done to find the best combination that can be used in future studies to predict eutrophication endpoints. It was found that model 5 is the best in terms of number of observations, geographic coverage and model fit (R$^2$) of chlorophyll $a$ that can be explained by TP and TKN.

Since lakes vary within ecoregions due to natural and anthropogenic factors, it would be ideal to have a standard for each lake. But, it would be very expensive to collect data and develop models to predict chlorophyll $a$ for each individual lake. This is why lakes should be grouped in a way so that a particular model for prediction of chlorophyll $a$ can be used for the all the lakes in the group. The ecoregion approach is a broad approach. The differences in lakes within an ecoregion may not be well explained by the differences used to create the ecoregions. An additional subdivision of lakes can be accomplished (by conditioning on TP and TN) with these empirical models. Model 5 can be used in further studies to find the best subdivision of ecoregions to set standards for water quality.

There is evidence of non-linearity in models 4, 5, and 6 in this study between the response and predictor variables (Figures 14, 17, and 20). Residuals from Model 5 were plotted against logTKN (Figure 22). This shows a random scatter of points which indicates an approximately random pattern. Residuals were plotted against logTP (Figure 23). This shows nonlinearity as indicated by the conical shape at the end. Nonlinear quadratic relationships between chlorophyll $a$ and TP and TN, semiparametric models, breakpoint regression and a random coefficient regression model are some of the alternatives to linear models.
From the models, it is seen that nutrients can explain some of the variance in chlorophyll $a$. But many factors that affect chlorophyll $a$ (light, flushing, zooplankton...
grazing, etc.) cannot be controlled. This leaves nutrient management as the most probable tool in controlling eutrophication in water bodies (Dodds et al, 2002). Data collection methods should be improved. This could be done by analyzing all the samples for all the variables of concern. Sampling methods should be uniform for individual variables.

This work can be treated as a first step to formulating an appropriate model for the prediction of eutrophication endpoints. The limitations and uncertainties encountered while doing this analysis are important for understanding the drawbacks/gaps that need to be addressed for finding the best model that takes into account these limitations and offers a more reliable and practical approach for prediction of these endpoints.
REFERENCES


Ambrose Jr. RB, Wool TA, Connolly JP. 1988. WASP4, A hydrodynamic and water quality model-Model theory, user’s manual, and programmer’s guide. Environmental Research Laboratory. USEPA. Athens, GA.


Marlene NaaNes. Drought, algae levels take life from lake, give rotten odor. 06/03/03. Advocate (2theadvocate.com).


About Privacy and Security

1. This World Wide Web (WWW) site is provided as a public service by the Environmental Protection Agency.

2. Information presented on this WWW site is considered public information and may be distributed or copied. The U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce these documents, or allow others to do so, for U.S. Government purposes. These documents may be freely distributed and used for non-commercial, scientific and educational purposes. Commercial use of the documents available from this server may be protected under the U.S. and Foreign Copyright Laws. Individual documents on this server may have different copyright conditions, and that will be noted in those documents.

3. When you come to this web site to browse, you do so anonymously. EPA does not collect identifying information about you. We collect only summary information [see below] about the numbers of individuals who visit our web site and what those individuals look at. This government computer system uses industry-standard software to create summary statistics, which are used for such things as assessing what information is of most and least interest, determining technical design specifications, and identifying system performance or problem areas.

4. Where identifying information is asked of you (to respond to an information request, etc.) that information is used only for responding to your comment or question and is not made available for other purposes. See our comments notice of use.

5. For site security purposes and to ensure that this service remains available to all users, this government computer system employs industry-standard methods to monitor network traffic to identify unauthorized attempts to upload or change information, or otherwise cause damage.

6. No other attempts are made to identify individual users or their usage habits. Raw data logs are used for no other purposes and are scheduled for regular destruction in accordance with National Archives and Records Administration guidelines.

7. Unauthorized attempts to upload information or change information on this service are strictly prohibited and may be punishable under the Computer Fraud and Abuse Act of 1986.

If you have any questions or comments about the information presented here, please forward them to us through the address on our Comments Page.
SAS Anova Output for Model 1

The SAS System
The GLM Procedure

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>3</td>
<td>2 7 8</td>
</tr>
</tbody>
</table>

Number of observations 4246

Dependent Variable: logCHLA

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>9</td>
<td>1630.274856</td>
<td>181.141651</td>
<td>1727.06</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>4237</td>
<td>444.395785</td>
<td>0.104885</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>4246</td>
<td>2074.670641</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square    Coeff Var     Root MSE     logCHLA Mean
0.280531    55.28613      0.323859     0.585787

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>3</td>
<td>1543.387733</td>
<td>514.462578</td>
<td>4905.04</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>3</td>
<td>83.146523</td>
<td>27.715508</td>
<td>264.25</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTN*ECOREGION_ID</td>
<td>3</td>
<td>3.740600</td>
<td>1.246867</td>
<td>11.89</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>3</td>
<td>242.4560533</td>
<td>80.8186844</td>
<td>770.55</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>3</td>
<td>76.1004409</td>
<td>25.3668136</td>
<td>241.85</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTN*ECOREGION_ID</td>
<td>3</td>
<td>3.7406003</td>
<td>1.2468668</td>
<td>11.89</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

| Parameter                   | Estimate | Error     | t Value | Pr > |t| |
|-----------------------------|----------|-----------|---------|------|--------|
| ECOREGION_ID                | 2        | 1.231016797 | 0.04008926 | 30.71 | <.0001 |
| ECOREGION_ID                | 7        | 1.567993796 | 0.04959061 | 31.62 | <.0001 |
| ECOREGION_ID                | 8        | 2.160235259 | 0.11245924 | 19.21 | <.0001 |
| logTP*ECOREGION_ID          | 2        | 0.262851846 | 0.02570754 | 10.22 | <.0001 |
| logTP*ECOREGION_ID          | 7        | 0.479468690 | 0.02644438 | 18.13 | <.0001 |
| logTP*ECOREGION_ID          | 8        | 0.914703768 | 0.05350348 | 17.10 | <.0001 |
| logTN*ECOREGION_ID          | 2        | 0.048957431 | 0.03561909 | 1.37  | 0.1694 |
| logTN*ECOREGION_ID          | 7        | 0.236321104 | 0.06124053 | 3.86  | 0.0001 |
| logTN*ECOREGION_ID          | 8        | -0.271911108 | 0.06257249 | -4.35 | <.0001 |
Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>---Statistic---</th>
<th>-----p Value------</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D    0.05365</td>
<td>Pr &gt; D &lt;0.0100</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq 2.820438</td>
<td>Pr &gt; W-Sq &lt;0.0050</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq 18.43623</td>
<td>Pr &gt; A-Sq &lt;0.0050</td>
</tr>
</tbody>
</table>

SAS Anova Output for Model 2

The SAS System
The GLM Procedure

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>7</td>
<td>1 11 14 2 7 8 9</td>
</tr>
</tbody>
</table>

Number of observations 5069

Dependent Variable: logCHLA

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>21</td>
<td>4015.644699</td>
<td>191.221176</td>
<td>787.73</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>5048</td>
<td>1225.397727</td>
<td>0.242749</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>5069</td>
<td>5241.042426</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square Coeff Var Root MSE logCHLA Mean
0.230490 58.07333 0.492696 0.848403

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>7</td>
<td>3767.090731</td>
<td>538.155819</td>
<td>2216.92</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>7</td>
<td>151.471102</td>
<td>21.638729</td>
<td>89.14</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTKN*ECOREGION_ID</td>
<td>7</td>
<td>97.082865</td>
<td>13.868981</td>
<td>57.13</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>7</td>
<td>778.6225599</td>
<td>111.2317943</td>
<td>458.22</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>7</td>
<td>73.0806366</td>
<td>10.4400909</td>
<td>43.01</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTKN*ECOREGION_ID</td>
<td>7</td>
<td>97.0828654</td>
<td>13.8689808</td>
<td>57.13</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

| Parameter                  | Estimate | Std Error | t Value | Pr > |t| |
|---------------------------|----------|-----------|--------|------|---|
| ECOREGION_ID 1            | 0.481217024 | 0.65419533 | 0.74 | 0.4620 |
| ECOREGION_ID 11           | 0.738267367 | 0.11222459 | 6.58 | <.0001 |
| ECOREGION_ID 14           | 1.534680297 | 0.09550770 | 16.07 | <.0001 |
| ECOREGION_ID 2            | 0.351149711 | 0.31728827 | 1.11 | 0.2685 |
Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D 0.053683</td>
<td>&lt;0.0100</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq 5.598006</td>
<td>&lt;0.0050</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq 35.26148</td>
<td>&lt;0.0050</td>
</tr>
</tbody>
</table>

SAS Anova Output for Model 3

The SAS System
The GLM Procedure

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>4</td>
<td>11 2 7 8</td>
</tr>
</tbody>
</table>

Number of observations 4817

Dependent Variable: logCHLA

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>12</td>
<td>1803.492198</td>
<td>150.291017</td>
<td>1263.15</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>4805</td>
<td>571.705879</td>
<td>0.118981</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>4817</td>
<td>2375.198078</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square: 0.274044
Coeff Var: 60.08236
Root MSE: 0.344937
logCHLA Mean: 0.574107
### SAS Anova Output for Model 4

The SAS System

The GLM Procedure

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>3</td>
<td>12 13 9</td>
</tr>
</tbody>
</table>

Number of observations 49301
Dependent Variable: logCHLAttri

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>9</td>
<td>54805.55345</td>
<td>6089.50594</td>
<td>53892.4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>49292</td>
<td>5569.69428</td>
<td>0.11299</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>49301</td>
<td>60375.24773</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R-Square: 0.607496, Coeff Var: 34.72997, Root MSE: 0.336146, logCHLAttri Mean: 0.967883

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>3</td>
<td>46290.92148</td>
<td>15430.30716</td>
<td>136559</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>3</td>
<td>7674.39147</td>
<td>2558.13049</td>
<td>22639.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTN*ECOREGION_ID</td>
<td>3</td>
<td>840.24050</td>
<td>280.08017</td>
<td>2478.72</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>3</td>
<td>9200.368610</td>
<td>3066.789537</td>
<td>27141.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>3</td>
<td>2011.978456</td>
<td>670.659485</td>
<td>5935.36</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTN*ECOREGION_ID</td>
<td>3</td>
<td>840.240504</td>
<td>280.080168</td>
<td>2478.72</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Parameter | Estimate | Standard Error | t Value | Pr > |t|
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID 12</td>
<td>2.118677800</td>
<td>0.00760272</td>
<td>278.67</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ECOREGION_ID 13</td>
<td>1.887067346</td>
<td>0.04230991</td>
<td>44.60</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ECOREGION_ID 9</td>
<td>2.420275175</td>
<td>0.05744195</td>
<td>42.13</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID 12</td>
<td>0.621876411</td>
<td>0.00475086</td>
<td>130.90</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID 13</td>
<td>0.519484773</td>
<td>0.02692492</td>
<td>19.29</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID 9</td>
<td>0.725213006</td>
<td>0.04189370</td>
<td>17.31</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTN*ECOREGION_ID 12</td>
<td>0.646263468</td>
<td>0.00757070</td>
<td>85.36</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTN*ECOREGION_ID 13</td>
<td>0.428154791</td>
<td>0.08054478</td>
<td>5.32</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTN*ECOREGION_ID 9</td>
<td>0.722803987</td>
<td>0.06572837</td>
<td>11.00</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D 0.034347</td>
<td>Pr &gt; D  &lt;0.0100</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq 25.95156</td>
<td>Pr &gt; W-Sq &lt;0.0050</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq 182.3509</td>
<td>Pr &gt; A-Sq &lt;0.0050</td>
</tr>
</tbody>
</table>
### SAS Anova output for Model 5

The SAS System  
The GLM Procedure  

#### Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>10</td>
<td>10 11 12 13 2 3 6 7 8 9</td>
</tr>
</tbody>
</table>

#### Number of observations 20681

#### Dependent Variable: logCHLatri

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>30</td>
<td>29724.35587</td>
<td>990.81186</td>
<td>7416.46</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>20651</td>
<td>2758.89717</td>
<td>0.13360</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>20681</td>
<td>32483.25304</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### R-Square, Coefficient of Variation, Root MSE, and logCHLatri Mean

- R-Square: 0.480209  
- Coefficient of Variation: 31.88554  
- Root MSE: 0.365508  
- logCHLatri Mean: 1.146313

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>10</td>
<td>27611.61387</td>
<td>2761.16139</td>
<td>20667.9</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>10</td>
<td>1714.46846</td>
<td>171.44685</td>
<td>1283.32</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTKN*ECOREGION_ID</td>
<td>10</td>
<td>398.27353</td>
<td>39.82735</td>
<td>298.12</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>10</td>
<td>3813.891338</td>
<td>381.389134</td>
<td>2854.79</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>10</td>
<td>436.405816</td>
<td>43.640582</td>
<td>326.66</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTKN*ECOREGION_ID</td>
<td>10</td>
<td>398.273534</td>
<td>39.827353</td>
<td>298.12</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

| Parameter               | Estimate | Standard Error | t Value | Pr > |t| |
|-------------------------|----------|----------------|---------|-------|---|
| ECOREGION_ID            | 10       | 2.359215541    | 1.02278022 | 2.31  | 0.0211 |
| ECOREGION_ID            | 11       | 1.475358834    | 0.06872305 | 21.47 | <.0001 |
| ECOREGION_ID            | 12       | 1.702749554    | 0.01477653 | 115.23 | <.0001 |
| ECOREGION_ID            | 13       | 1.690133009    | 0.02537298 | 66.61  | <.0001 |
| ECOREGION_ID            | 2        | 1.519924153    | 0.39809402 | 3.82   | 0.0001 |
| ECOREGION_ID            | 3        | 1.320811280    | 0.21079561 | 6.27   | <.0001 |
| ECOREGION_ID            | 6        | 1.566600500    | 0.03556858 | 44.04  | <.0001 |
| ECOREGION_ID            | 7        | 1.811234703    | 0.02141959 | 84.56  | <.0001 |
| ECOREGION_ID            | 8        | 2.125527877    | 0.09069597 | 23.44  | <.0001 |
| ECOREGION_ID            | 9        | 1.276928827    | 0.04924022 | 25.93  | <.0001 |
| logTP*ECOREGION_ID      | 10       | 0.875672742    | 1.88953312 | 0.46   | 0.6431 |
| logTP*ECOREGION_ID      | 11       | 0.447838758    | 0.03684421 | 12.15  | <.0001 |
| logTP*ECOREGION_ID      | 12       | 0.375003092    | 0.01017208 | 36.87  | <.0001 |
### Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D</td>
<td>&lt;0.0100</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq</td>
<td>&lt;0.0050</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq</td>
<td>&lt;0.0050</td>
</tr>
</tbody>
</table>

### SAS Anova Output for Model 6

The SAS System
The GLM Procedure

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Levels</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>6</td>
<td>11 12 13 2 7 9</td>
</tr>
</tbody>
</table>

Number of observations 62475

Dependent Variable: logCHLAtri

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>18</td>
<td>7235.65913</td>
<td>4019.64773</td>
<td>3335.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Error</td>
<td>62457</td>
<td>7527.11954</td>
<td>0.12052</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncorrected Total</td>
<td>62475</td>
<td>79880.77867</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### R-Square Coeff Var Root MSE logCHLAtri Mean

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.583141</td>
<td>34.89782</td>
<td>0.347155</td>
<td>0.994776</td>
</tr>
</tbody>
</table>

### Source DF Type I SS Mean Square F Value Pr > F

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type I SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>6</td>
<td>62404.70893</td>
<td>10400.78482</td>
<td>86301.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>6</td>
<td>8900.27050</td>
<td>1483.37842</td>
<td>12308.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>lognewTN*ECOREGION_I</td>
<td>6</td>
<td>1048.67970</td>
<td>174.77995</td>
<td>1450.25</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

### Source DF Type III SS Mean Square F Value Pr > F

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Type III SS</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECOREGION_ID</td>
<td>6</td>
<td>11140.27599</td>
<td>1856.71266</td>
<td>15406.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>logTP*ECOREGION_ID</td>
<td>6</td>
<td>2261.97477</td>
<td>376.99579</td>
<td>3128.16</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>lognewTN*ECOREGION_I</td>
<td>6</td>
<td>1048.67970</td>
<td>174.77995</td>
<td>1450.25</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

### Standard Parameter Estimate Error t Value Pr > |t|

| Parameter                      | Estimate | Error     | t Value | Pr > |t|      |
|-------------------------------|----------|-----------|---------|------|-------|
| ECOREGION_ID                  | 11       | 1.466040179| 0.06509604| 22.52| <.0001|
| ECOREGION_ID                  | 12       | 2.074289570| 0.00743439| 279.01| <.0001|
| ECOREGION_ID                  | 13       | 1.680480508| 0.02517821| 66.74 | <.0001|
| ECOREGION_ID                  | 2        | 2.505844894| 0.35647245| 7.03  | <.0001|
| ECOREGION_ID                  | 7        | 1.810782528| 0.02035535| 88.96 | <.0001|
| ECOREGION_ID                  | 9        | 2.420275175| 0.05932334| 40.80 | <.0001|
| logTP*ECOREGION_ID            | 11       | 0.447838758| 0.03499419| 12.80 | <.0001|
| logTP*ECOREGION_ID            | 12       | 0.599469075| 0.00464522| 129.05| <.0001|
| logTP*ECOREGION_ID            | 13       | 0.382782137| 0.01713821| 22.34 | <.0001|
| logTP*ECOREGION_ID            | 2        | -0.118101064| 0.29346137|-0.40  | 0.6874|
| logTP*ECOREGION_ID            | 7        | 0.490158208| 0.01432299| 34.22 | <.0001|
| logTP*ECOREGION_ID            | 9        | 0.725213006| 0.04326583| 16.76 | <.0001|
| lognewTN*ECOREGION_I          | 11       | 0.194230877| 0.02250508| 8.63  | <.0001|
| lognewTN*ECOREGION_I          | 12       | 0.645501091| 0.00734834| 87.84 | <.0001|
| lognewTN*ECOREGION_I          | 13       | 0.582397945| 0.05034986| 11.57 | <.0001|
| lognewTN*ECOREGION_I          | 2        | 2.214693847| 0.46837705| 4.73  | <.0001|
| lognewTN*ECOREGION_I          | 7        | 0.677888173| 0.02677332| 25.32 | <.0001|
| lognewTN*ECOREGION_I          | 9        | 0.722803987| 0.06788116| 10.65 | <.0001|

### Tests for Normality

<table>
<thead>
<tr>
<th>Test</th>
<th>---Statistic---</th>
<th>-----p Value-----</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D 0.036567</td>
<td>Pr &gt; D &lt;0.0100</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq 36.65631</td>
<td>Pr &gt; W-Sq &lt;0.0050</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq 251.3934</td>
<td>Pr &gt; A-Sq &lt;0.0050</td>
</tr>
</tbody>
</table>
APPENDIX C
TABLES WITH PARAMETER VALUES FOR EACH MODEL
Model 1: $\log{\text{CHLA}} = \log{\text{TP}} + \log{\text{TN}}$

<table>
<thead>
<tr>
<th>Ecoregion</th>
<th>$\log{\text{CHLA}}$</th>
<th>$\beta_0$</th>
<th>$\beta_1$ logTP</th>
<th>$\beta_2$ logTN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Std. error)</td>
<td>(Std. error)</td>
<td>(Std. error)</td>
<td>(Std. error)</td>
</tr>
<tr>
<td>2</td>
<td>1.2310 (0.0400)</td>
<td>0.2628 (0.0257)</td>
<td>0.0489 (0.0356)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.5679 (0.0495)</td>
<td>0.4794 (0.0264)</td>
<td>0.2363 (0.0612)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.1602 (0.1124)</td>
<td>0.9147 (0.0535)</td>
<td>-0.2719 (0.0625)</td>
<td></td>
</tr>
</tbody>
</table>

Model 2: $\log{\text{CHLA}} = \log{\text{TP}} + \log{\text{TKN}}$

<table>
<thead>
<tr>
<th>Ecoregion</th>
<th>$\log{\text{CHLA}}$</th>
<th>$\beta_0$</th>
<th>$\beta_1$ logTP</th>
<th>$\beta_2$ logTKN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Std. error)</td>
<td>(Std. error)</td>
<td>(Std. error)</td>
<td>(Std. error)</td>
</tr>
<tr>
<td>1</td>
<td>0.4812 (0.6541)</td>
<td>-1.3050 (1.2618)</td>
<td>1.8911 (1.1046)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.3511 (0.3172)</td>
<td>-0.6846 (0.2744)</td>
<td>1.5555 (0.2485)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.0106 (0.2378)</td>
<td>0.2357 (0.1275)</td>
<td>0.1688 (0.1737)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.7771 (0.5045)</td>
<td>0.2636 (0.3410)</td>
<td>-0.0952 (0.3909)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.5398 (0.0294)</td>
<td>0.2571 (0.0204)</td>
<td>0.5598 (0.0304)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.7382 (0.1122)</td>
<td>0.1503 (0.0594)</td>
<td>0.1018 (0.0678)</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1.5346 (0.0955)</td>
<td>0.3363 (0.0571)</td>
<td>0.3004 (0.0758)</td>
<td></td>
</tr>
</tbody>
</table>
Model 3: \( \text{logCHLA} = \text{logTP} + \text{lognewTN} \)

<table>
<thead>
<tr>
<th>Ecoregion</th>
<th>( \beta_0 ) (Std. error)</th>
<th>( \beta_1 ) ( \text{logTP} ) (Std. error)</th>
<th>( \beta_2 ) ( \text{lognewTKN} ) (Std. error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.2514 (0.0408)</td>
<td>0.2557 (0.0263)</td>
<td>0.1368 (0.0354)</td>
</tr>
<tr>
<td>7</td>
<td>1.6931 (0.0502)</td>
<td>0.5529 (0.0267)</td>
<td>0.1968 (0.0574)</td>
</tr>
<tr>
<td>8</td>
<td>2.1602 (0.1197)</td>
<td>0.9147 (0.0569)</td>
<td>-0.2719 (0.0666)</td>
</tr>
<tr>
<td>11</td>
<td>0.7335 (0.0783)</td>
<td>0.1503 (0.0416)</td>
<td>0.0992 (0.0463)</td>
</tr>
</tbody>
</table>

Model 4: \( \text{logCHLA}_{tri} = \text{logTP} + \text{logTN} \)

<table>
<thead>
<tr>
<th>Ecoregion</th>
<th>( \beta_0 ) (Std. error)</th>
<th>( \beta_1 ) ( \text{logTP} ) (Std. error)</th>
<th>( \beta_2 ) ( \text{logTN} ) (Std. error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>2.4202 (0.0574)</td>
<td>0.7252 (0.0418)</td>
<td>0.7228 (0.0657)</td>
</tr>
<tr>
<td>12</td>
<td>2.1186 (0.0076)</td>
<td>0.6218 (0.0047)</td>
<td>0.6462 (0.0075)</td>
</tr>
<tr>
<td>13</td>
<td>1.8870 (0.0423)</td>
<td>0.5194 (0.0269)</td>
<td>0.4281 (0.0805)</td>
</tr>
</tbody>
</table>
**Model 5: logCHLAtri = logTP + logTKN**

<table>
<thead>
<tr>
<th>Ecoregion</th>
<th>logCHLAtri (Std. error)</th>
<th>β₀ (Std. error)</th>
<th>β₁ logTP (Std. error)</th>
<th>β₂ logTKN (Std. error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.5199 (0.3980)</td>
<td>-0.1181 (0.3089)</td>
<td>1.6973 (0.3779)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.3208 (0.2107)</td>
<td>-0.3358 (0.1937)</td>
<td>1.7856 (0.2555)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.5666 (0.0355)</td>
<td>0.3095 (0.0310)</td>
<td>0.0051 (0.0425)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.8112 (0.0214)</td>
<td>0.4901 (0.0150)</td>
<td>0.6723 (0.0279)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.1255 (0.0906)</td>
<td>0.6420 (0.0581)</td>
<td>0.1333 (0.0775)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.2769 (0.0492)</td>
<td>-0.0190 (0.0487)</td>
<td>0.5124 (0.0835)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2.3592 (1.0227)</td>
<td>0.8756 (1.8895)</td>
<td>0.9076 (1.8994)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1.4753 (0.0687)</td>
<td>0.4478 (0.0368)</td>
<td>0.1991 (0.0243)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.7027 (0.0147)</td>
<td>0.3750 (0.0101)</td>
<td>0.8381 (0.0184)</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1.6901 (0.0253)</td>
<td>0.3899 (0.0178)</td>
<td>0.6791 (0.0543)</td>
<td></td>
</tr>
</tbody>
</table>

**Model 6: logCHLAtri = logTP + lognewTN**

<table>
<thead>
<tr>
<th>Ecoregion</th>
<th>logCHLAtri (Std. error)</th>
<th>β₀ (Std. error)</th>
<th>β₁ logTP (Std. error)</th>
<th>β₂ lognewTN (Std. error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.5058 (0.3564)</td>
<td>-0.1181 (0.2934)</td>
<td>2.2146 (0.4683)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.8107 (0.0203)</td>
<td>0.4901 (0.0143)</td>
<td>0.6778 (0.0267)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2.4202 (0.0593)</td>
<td>0.7252 (0.0432)</td>
<td>0.7228 (0.0678)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1.4660 (0.0650)</td>
<td>0.4478 (0.0349)</td>
<td>0.1942 (0.0225)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2.0742 (0.0074)</td>
<td>0.5994 (0.0046)</td>
<td>0.6455 (0.0073)</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1.6804 (0.0251)</td>
<td>0.3827 (0.0171)</td>
<td>0.5823 (0.0503)</td>
<td></td>
</tr>
</tbody>
</table>
Table: Number of data in each model by ecoregion.

<table>
<thead>
<tr>
<th>Model</th>
<th>Eco1</th>
<th>Eco2</th>
<th>Eco3</th>
<th>Eco4</th>
<th>Eco5</th>
<th>Eco6</th>
<th>Eco7</th>
<th>Eco8</th>
<th>Eco9</th>
<th>Eco10</th>
<th>Eco11</th>
<th>Eco12</th>
<th>Eco13</th>
<th>Eco14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>716</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3045</td>
<td>485</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>117</td>
<td>58</td>
<td>3887</td>
<td>0</td>
<td>394</td>
<td>0</td>
<td>0</td>
<td>538</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>776</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3162</td>
<td>485</td>
<td>0</td>
<td>0</td>
<td>394</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1423</td>
<td>0</td>
<td>0</td>
<td>46349</td>
<td>1529</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>16</td>
<td>37</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>642</td>
<td>5525</td>
<td>597</td>
<td>231</td>
<td>7</td>
<td>899</td>
<td>8934</td>
<td>3793</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5525</td>
<td>0</td>
<td>1423</td>
<td>0</td>
<td>899</td>
<td>50797</td>
<td>3815</td>
<td>0</td>
</tr>
</tbody>
</table>
APPENDIX D
BOXPLOTS OF DISTRIBUTION OF CHLOROPHYLL A FOR EACH MODEL
BY ECOREGION
MODEL 3  Log CHLA
By Ecoregion

MODEL 4  Log CHLatri
By Ecoregion
APPENDIX E
MAPS OF THE SPATIAL DISTRIBUTION OF OBSERVATIONS FOR EACH MODEL
Figure: Spatial distributions of observations for Model 1
Figure: Spatial distributions of observations for Model 2
Figure: Spatial distributions of observations for Model 3.
Figure: Spatial distributions of observations for Model 4.
Figure: Spatial distributions of observations for Model 5.
Figure: Spatial distributions of observations for Model 6.
Arrangement in record format (example for Alabama):

Data TP;
set epalakes.lakeAL;
if STORET_CODE=665;
TP=REPORTED_VALUE;
keep WATERBODY_ID_ WATERBODY_TYPE_ID STATION_ID STATE COUNTY ECOREGION_ID EPA_ECOREGION_ID SAMPLE_ID AGENCY_NAME SAMPLING_DATE SAMPLE_DEPTH LATITUDE LONGITUDE HYDROLOGIC_UNIT_CODE TP;
run;
proc sort;
by STATION_ID SAMPLING_DATE SAMPLE_ID;
run;

data NO2;
set epalakes.lakeAL;
if STORET_CODE=615;
NO2=REPORTED_VALUE;
keep WATERBODY_ID_ WATERBODY_TYPE_ID STATION_ID STATE COUNTY ECOREGION_ID EPA_ECOREGION_ID SAMPLE_ID AGENCY_NAME SAMPLING_DATE SAMPLE_DEPTH LATITUDE LONGITUDE HYDROLOGIC_UNIT_CODE NO2;
run;
proc sort;
by STATION_ID SAMPLING_DATE SAMPLE_ID;
run;

data DO;
set epalakes.lakeAL;
if STORET_CODE=300;
DO=REPORTED_VALUE;
keep WATERBODY_ID_ WATERBODY_TYPE_ID STATION_ID STATE COUNTY ECOREGION_ID EPA_ECOREGION_ID SAMPLE_ID AGENCY_NAME SAMPLING_DATE SAMPLE_DEPTH LATITUDE LONGITUDE HYDROLOGIC_UNIT_CODE DO;
run;
proc sort;
by STATION_ID SAMPLING_DATE SAMPLE_ID;
run;

data TKN;
set epalakes.lakeAL;
if STORET_CODE=625;
TKN=REPORTED_VALUE;
keep WATERBODY_ID_ WATERBODY_TYPE_ID STATION_ID STATE COUNTY ECOREGION_ID EPA_ECOREGION_ID SAMPLE_ID AGENCY_NAME SAMPLING_DATE SAMPLE_DEPTH LATITUDE LONGITUDE HYDROLOGIC_UNIT_CODE TKN;
run;
proc sort;
by STATION_ID SAMPLING_DATE SAMPLE_ID;
run;

data NH3;
   set epalakes.lakeAL;
   if STORET_CODE=610;
      NH3=REPORTED_VALUE;
   keep WATERBODY_ID_ WATERBODY_TYPE_ID STATION_ID STATE COUNTY ECOREGION_ID EPA_ECOREGION_ID SAMPLE_ID AGENCY_NAME SAMPLING_DATE SAMPLE_DEPTH LATITUDE LONGITUDE HYDROLOGIC_UNIT_CODE NH3;
run;
proc sort;
by STATION_ID SAMPLING_DATE SAMPLE_ID;
run;

data NO2NO3;
   set epalakes.lakeAL;
   if STORET_CODE=630;
      NO2NO3=REPORTED_VALUE;
   if STORET_CODE=10 then TEMP=REPORTED_VALUE;
   keep WATERBODY_ID_ WATERBODY_TYPE_ID STATION_ID STATE COUNTY ECOREGION_ID EPA_ECOREGION_ID SAMPLE_ID AGENCY_NAME SAMPLING_DATE SAMPLE_DEPTH LATITUDE LONGITUDE HYDROLOGIC_UNIT_CODE NO2NO3;
run;
proc sort;
by STATION_ID SAMPLING_DATE SAMPLE_ID;
run;

data TEMP;
   set epalakes.lakeAL;
   if STORET_CODE=10;
      TEMP=REPORTED_VALUE;
   keep WATERBODY_ID_ WATERBODY_TYPE_ID STATION_ID STATE COUNTY ECOREGION_ID EPA_ECOREGION_ID SAMPLE_ID AGENCY_NAME SAMPLING_DATE SAMPLE_DEPTH LATITUDE LONGITUDE HYDROLOGIC_UNIT_CODE TEMP;
run;
proc sort;
  by STATION_ID SAMPLING_DATE SAMPLE_ID;
run;

data epalakes.lakeAL;
merge TP NO2 DO TKN3 NO2NO3 TEMP;
by STATION_ID SAMPLING_DATE SAMPLE_ID;
run;
proc means;
run;

Merging datasets of all the states to form one dataset:

data epalakes.lakeall;
by STATION_ID SAMPLING_DATE SAMPLE_ID;
run;
proc means;
run;

Regression of TN with TKN:

data epalakes.TN_TKN ;                  /*making data set with TN TKN */
set epalakes.lakeall ;
if TN ne . and TKN ne .  ;
run

data TN_TKN ;
set epalakes.TN_TKN ;
logCHLAttri = log10(CHLAttri);
logTP = log10(TP);
logTN = log10(TN);
logTKN = log10(TKN);
run;
proc sort data= TN_TKN ;
by ecoregion_id ;
run
proc means data = TN_TKN ;
by ecoregion_id ;
var TP TN TKN CHL Atri logTN logCHL Atri logTKN;
run
proc glm data= TN_TKN ;    /* general linear model*/
class ecoregion_id ;
model logTN = ecoregion_id ecoregion_id*logTKN
/ noint solution  ;   /* add these one by one to see if total number of samples vary
because SAS will take only those data that have all non-missing values. Take either TN
or TKN but not together */
id sampling_date ecoregion_id;
output out= TN_TKN predicted=pred
    residual=res
    rstudent=jackres    /* jack knife residual */
    press=press
    h=h     /* hat matrix*/
    cookd=cookd ;   /* cook's distance*/
run;

Example of code for formation of variable “newTN”:

libname epalakes "C:\Program Files\Insightful\splus6netclient\users";
data eco2;
set epalakes.lakeall;
if ecoregion_id = 2;
logTN = log10(TN);
logTKN = log10(TKN);
pTN2=0;
logTN2 = -0.445172475 + 0.766388167*(logTKN)  ;
if logTKN ne . then pTN2 = 1;
TN2=10**(logTN2);             /*(** = raise to the power of)*/
if TN ne . then newTN = TN;
if TN = . then newTN = TN2;     /* creating variable 'newTN' to include TN and
predicted TN to get more observations*/
rn;
proc sort;
by pTN2;
run;
proc means;
by pTN2;
run;
proc means;
by pTN2;
var TN logTN TKN logTKN TN2 logTN2 pTN2 newTN ;
run;
proc means;
by pTN2;
var TN logTN TKN logTKN TN2 logTN2 pTN2 newTN ;
run;
proc gplot;
plot TN*TN2;
run;

**Regression codes for each model:**

libname epalakes "C:\Program Files\Insightful\splus6netclient\users";
data CHLATPTN;  /* Checking model 1 logCHLA = logTP + logTN*/
set epalakes.lakeall;
logCHLA = log10(CHLA);
logTP = log10(TP);
logTN = log10(TN);
if logCHLA ne . ;
if logTP ne . ;
if logTN ne . ;
run;
proc sort data= CHLATPTN ;
by ecoregion_id ;
run;
proc means data = CHLATPTN ;
by ecoregion_id ;
var TP logTP TN logTN CHLA logCHLA ;
run;

proc glm data= CHLATPTN ;  /* general linear model*/
class ecoregion_id ;
model logCHLA = ecoregion_id ecoregion_id*logTP ecoregion_id*logTN
/ noint solution ;  /* add these one by one to see if total number of samples vary
because SAS will take only those data that have all non-missing values. Take either
TN or TKN but not together */
id sampling_date ecoregion_id;
output out= CHLATPTN predicted=pred
    residual=res
    rstudent=jackres  /* jack knife residual */
    press=press
    h=h  /* hat matrix*/
    cookd=cookd /* cook's distance*/
    student = student ;
run;

proc univariate data = CHLATPTN normal plot ;
var jackres student;
run;

libname epalakes "C:\Program Files\Insightful\splus6netclient\users";
data CHLATPTKN;  /* checking model 2 logCHLA = logTP + logTKN */
set epalakes.lakeall;
logCHLA = log10(CHLA);
logTP = log10(TP);
logTKN = log10(TKN);
if logCHLA ne . ;
if logTP ne . ;
if logTKN ne . ;
run;

proc sort  data= CHLATPTKN ;
by ecoregion_id ;
run;

proc means data = CHLATPTKN ;
by ecoregion_id ;
var TP logTP TKN logTKN CHLA   logCHLA ;
run;

proc glm data= CHLATPTKN ; /* general linear model*/
class ecoregion_id ;
model logCHLA = ecoregion_id  ecoregion_id*logTP  ecoregion_id*logTKN
/ noint solution ; /* add these one by one to see if if total number of samples vary because SAS will take only those data that have all non-missing values. take either TN or TKN but not together */
id sampling_date ecoregion_id;
output out= CHLATPTKN predicted=pred
   residual=res
   rstudent=jackres /* jack knife residual */
   press=press
   h=h   /* hat matrix*/
   cookd=cookd /* cook's distance*/
   student = student ;
run;

proc univariate  data = CHLATPTKN normal plot ;
var jackres student;
run;

libname epalakes "C:\Program Files\Insightful\splus6netclient\users";
data CHLATPnewTN;  /*checking model 3 logCHLA = logTP + lognewTN. (do this for ecoregions 2,7,11,12,13 with respective formulae)*/
set epalakes.lakeall;
logCHLA = log10(CHLA);
logTP = log10(TP);
logTKN = log10(TKN);
if ecoregion_id = 2 then logTN2 = -0.445172475 + 0.766388167*(logTKN) ;
if ecoregion_id = 7 then logTN2 = 0.000667036 + 0.991866474*(logTKN) ;
if ecoregion_id = 11 then logTN2 = 0.047977207 + 1.025552479*(logTKN) ;
if ecoregion_id = 12 then logTN2 = 0.015506943 + 0.958164113*(logTKN) ;
if ecoregion_id = 13 then logTN2 = 0.026786063 + 0.997622084*(logTKN) ;

if ecoregion_id = 1 then TN2 = TN;
if ecoregion_id = 2 then  TN2=10**(logTN2); /*(** = raise to the power of)*/
if ecoregion_id = 3 then TN2 = TN;
if ecoregion_id = 4 then TN2 = TN;
if ecoregion_id = 5 then TN2 = TN;
if ecoregion_id = 6 then TN2 = TN;
if ecoregion_id = 7 then  TN2=10**(logTN2);
if ecoregion_id = 8 then TN2 = TN;
if ecoregion_id = 9 then TN2 = TN;
if ecoregion_id = 10 then TN2 = TN;
if ecoregion_id = 11 then  TN2=10**(logTN2);
if ecoregion_id = 12 then  TN2=10**(logTN2);
if ecoregion_id = 13 then  TN2=10**(logTN2);
if ecoregion_id = 14 then TN2 = TN;

If TN ne . then newTN = TN;
if TN = . then newTN = TN2;
lognewTN =  log10(newTN);
if logCHLA ne . ;
if logTP ne . ;
if lognewTN ne . ;
run;

proc sort  data= CHLATPnewTN ;
by ecoregion_id ;
run;

proc means data = CHLATPnewTN ;
by ecoregion_id ;
var TP logTP newTN lognewTN CHLA logCHLA ;
run;

proc glm data= CHLATPnewTN ; /* general linear model*/
class ecoregion_id ;
model logCHLA = ecoregion_id ecoregion_id*logTP ecoregion_id*lognewTN

/ noint solution ; /* add these one by one to see if if total number of samples vary
because SAS will take only those data that have all non-missing values . take either TN
or TKN but not together */
id sampling_date ecoregion_id;
output out= CHLATPnewTN predicted=pred
    residual=res
    rstudent=jackres /* jack knife residual */
    press=press
    h=h /* hat matrix*/
    cookd=cookd /* cook's distance*/
    student = student;
run;

proc univariate data = CHLATPnewTN normal plot ;
var jackres student;
run;

libname epalakes "C:\Program Files\Insightful\splus6netclient\users";
data CHLAtriTPTN; /* checking model 4  logCHLAtri = logTP + logTN*/
  set epalakes.lakeall;
  logCHLAtri = log10(CHLAtri);
  logTP = log10(TP);
  logTN = log10(TN);
  if logCHLAtri ne .;
  if logTP ne .;
  if logTN ne .;
run;

proc sort data= CHLAtriTPTN ;
by ecoregion_id ;
run;

proc means data = CHLAtriTPTN ;
by ecoregion_id ;
var TP logTP TN logTN CHLAtri logCHLAtri ;
run;

proc glm data= CHLAtriTPTN ; /* general linear model*/
class ecoregion_id ;
model logCHLAtri  = ecoregion_id  ecoregion_id*logTP  ecoregion_id*logTN
  / noint solution ; /* add these one by one to see if if total number of samples vary
  because SAS will take only
  those data that have all non-missing values . take either TN or
  TKN but not together */
  id sampling_date ecoregion_id;
output out= CHLAtriTPTN predicted=pred
  residual=res
  rstudent=jackres /* jack knife residual */
  press=press
  h=h /* hat matrix*/

cookd=cookd /* cook's distance*/
student = student ;
run;

proc univariate data = CHLAtriTPTN normal plot; /*gives univariate with plots and
tests for normality for variables in the var statement*/
var jackres student;
run;

libname epalakes "C:\Program Files\Insightful\splus6netclient\users";
data CHLAtriTPTKN ;                  /*checking Model 5 :  logCHLAtri = logTP +
logTKN*/
set epalakes.lakeall ;
logCHLAtri =  log10(CHLAtri);
logTP = log10(TP);
logTKN = log10(TKN);
if logCHLAtri ne . ;
if logTP ne . ;
if logTKN ne . ;
run;

proc sort data= CHLAtriTPTKN ;
by ecoregion_id ;
run;

proc means data = CHLAtriTPTKN ;
by ecoregion_id ;
var TP logTP TKN logTKN CHLAtri   logCHLAtri ;
run;

proc glm data= CHLAtriTPTKN ; /* general linear model*/
class ecoregion_id ;
model logCHLAtri = ecoregion_id ecoregion_id*logTP ecoregion_id*logTKN
/ noint solution ; /* add these one by one to see if if total number of samples vary
because SAS will take only
those data that have all non-missing values . take either TN or
TKN but not together */
id sampling_date ecoregion_id;
output out= CHLAtriTPTKN  predicted=pred
    residual=res
    rstudent=jackres /* jack knife residual */
    press=press
    h=h     /* hat matrix*/
    cookd=cookd /* cook's distance*/
    student = student ;
run;
proc univariate data = CHLAtriTPTKN normal plot;
var jackres student;
run;

libname epalakes "C:\Program Files\Insightful\splus6netclient\users";

data CHLAtriTPnewTN; */checking model 6 \( \log(CHLAttri) = \log(TP) + \log(newTN) \). 
(do this for ecoregions 2,7,11,12,13 with respective formulae)*/
set epalakes.lakeall;
logCHLAttri = log10(CHLAttri);
logTP = log10(TP);
logTKN = log10(TKN);
if ecoregion_id = 2 then logTN2 = -0.445172475 + 0.766388167*(logTKN) ;
if ecoregion_id = 7 then logTN2 = 0.000667036 + 0.991866474*(logTKN) ;
if ecoregion_id = 11 then logTN2 = 0.047977207 + 1.025552479*(logTKN) ;
if ecoregion_id = 12 then logTN2 = 0.015506943 + 0.958164113*(logTKN) ;
if ecoregion_id = 13 then logTN2 = 0.026786063 + 0.997622084*(logTKN) ;

if ecoregion_id = 1 then TN2 = TN;
if ecoregion_id = 2 then TN2=10**(logTN2); /*(** = raise to the power of)*/
if ecoregion_id = 3 then TN2 = TN;
if ecoregion_id = 4 then TN2 = TN;
if ecoregion_id = 5 then TN2 = TN;
if ecoregion_id = 6 then TN2 = TN;
if ecoregion_id = 7 then TN2=10**(logTN2);
if ecoregion_id = 8 then TN2 = TN;
if ecoregion_id = 9 then TN2 = TN;
if ecoregion_id = 10 then TN2 = TN;
if ecoregion_id = 11 then TN2=10**(logTN2);
if ecoregion_id = 12 then TN2=10**(logTN2);
if ecoregion_id = 13 then TN2=10**(logTN2);
if ecoregion_id = 14 then TN2 = TN;

If TN ne . then newTN = TN;
if TN = . then newTN = TN2;
lognewTN = log10(newTN);
if logCHLAttri ne . ;
if logTP ne . ;
if lognewTN ne . ;
run;

proc sort data= CHLAtriTPnewTN ;
  by ecoregion_id ;
run;

proc means data = CHLAtriTPnewTN ;
by ecoregion_id;
var TP logTP newTN lognewTN CHLatri logCHLAtri;
run;

proc glm data=CHLatriTPnewTN /* general linear model*/
class ecoregion_id;
model logCHLAtri = ecoregion_id ecoregion_id*logTP ecoregion_id*lognewTN
/ noint solution; /* add these one by one to see if if total number of samples vary
because SAS will take only those data that have all non-missing values. Take either
TN or TKN but not together */
id sampling_date ecoregion_id;
output out=CHLatriTPnewTN predicted=pred
residual=res
rstudent=jackres /* jack knife residual */
press=press
h=h /* hat matrix*/
cookd=cookd /* cook's distance*/
student = student;
run;

proc univariate data=CHLatriTPnewTN normal plot;
var jackres student;
run;
APPENDIX G
SUMMARY STATISTICS FOR ECOREGIONS
<table>
<thead>
<tr>
<th>Codes</th>
<th>Parameter</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Temperature (C)</td>
<td>TEMP</td>
</tr>
<tr>
<td>11</td>
<td>Temperature (F)</td>
<td>TF</td>
</tr>
<tr>
<td>60</td>
<td>Mean Daily Stream Flow, cfs</td>
<td>MDS</td>
</tr>
<tr>
<td>61</td>
<td>Instantaneous Streamflow, cfs</td>
<td>ISF</td>
</tr>
<tr>
<td>70</td>
<td>Turbidity</td>
<td>JCU</td>
</tr>
<tr>
<td>76</td>
<td>Turbidity</td>
<td>FTU</td>
</tr>
<tr>
<td>77</td>
<td>Secchi (inches)</td>
<td>SECHIin</td>
</tr>
<tr>
<td>78</td>
<td>Secchi (metres)</td>
<td>SECHI</td>
</tr>
<tr>
<td>94</td>
<td>Conductivity Field, ys/cm</td>
<td>CF</td>
</tr>
<tr>
<td>95</td>
<td>Conductivity</td>
<td>COND</td>
</tr>
<tr>
<td>300</td>
<td>Dissolved Oxygen, mg/l</td>
<td>DO</td>
</tr>
<tr>
<td>301</td>
<td>Dissolved Oxygen (DO % saturated)</td>
<td>DOsat</td>
</tr>
<tr>
<td>400</td>
<td>pH whole water field</td>
<td>pH</td>
</tr>
<tr>
<td>403</td>
<td>pH whole water lab S.U.</td>
<td>pHSU</td>
</tr>
<tr>
<td>409</td>
<td>Alkalinity, ueq/l</td>
<td>ALKueq</td>
</tr>
<tr>
<td>410</td>
<td>Alkalinity Total (as CACO3) Field</td>
<td>ALK</td>
</tr>
<tr>
<td>415</td>
<td>Alkalinity, Phenolphthalein, mg/l</td>
<td>ALKP</td>
</tr>
<tr>
<td>500</td>
<td>Total Solids, mg/l</td>
<td>TS</td>
</tr>
<tr>
<td>530</td>
<td>Residual Total Nonfiltrable, mg/l</td>
<td>RTN</td>
</tr>
<tr>
<td>600</td>
<td>Nitrogen Total, mg/l</td>
<td>TN</td>
</tr>
<tr>
<td>605</td>
<td>N Org total</td>
<td>TON</td>
</tr>
<tr>
<td>608</td>
<td>Ammonia dissolved at 180 deg. mg/l as N</td>
<td>NH3D</td>
</tr>
<tr>
<td>610</td>
<td>Ammonia Nitrogen Total</td>
<td>NH3</td>
</tr>
<tr>
<td>615</td>
<td>Nitrate</td>
<td>NO2</td>
</tr>
<tr>
<td>618</td>
<td>Nitrate dissolved as N, mg/l</td>
<td>ND</td>
</tr>
<tr>
<td>620</td>
<td>Nitrate mg/l</td>
<td>NO3</td>
</tr>
<tr>
<td>623</td>
<td>Nitrogen Kjeldhal dissolved as N, mg/l</td>
<td>NKD</td>
</tr>
<tr>
<td>625</td>
<td>Nitrogen Total Kjeldhal, mg/l</td>
<td>TKN</td>
</tr>
<tr>
<td>630</td>
<td>Nitrite and Nitrate, mg/l</td>
<td>NO2NO3</td>
</tr>
<tr>
<td>631</td>
<td>Nitrite and Nitrate dissolved mg/l as N</td>
<td>NND</td>
</tr>
<tr>
<td>660</td>
<td>Orthophosphate (OPO4_PO4), mg/l</td>
<td>OPO4</td>
</tr>
<tr>
<td>665</td>
<td>Phosphorous Total, ug/l</td>
<td>TP</td>
</tr>
<tr>
<td>666</td>
<td>Phosphorous Dissolved</td>
<td>DP</td>
</tr>
<tr>
<td>669</td>
<td>Phosphorous Total Reactive, mg/l</td>
<td>PTR</td>
</tr>
<tr>
<td>671</td>
<td>Orthophosphate dissolved mg/l as P</td>
<td>OPD</td>
</tr>
<tr>
<td>915</td>
<td>Alkalinity, mg/l</td>
<td>ALKmg</td>
</tr>
<tr>
<td>955</td>
<td>Silica dissolved, mg/l</td>
<td>SIL</td>
</tr>
<tr>
<td>32209</td>
<td>Chlorophyll a Fluorometric corrected, ug/l</td>
<td>CHLA</td>
</tr>
<tr>
<td>32210</td>
<td>Chlorophyll a Trichromatic uncorrected, ug/l</td>
<td>CHLAttri</td>
</tr>
<tr>
<td>32211</td>
<td>Chlorophyll a Phytoplankton Spectrophotometric Acid, ug/l</td>
<td>CHLAsphyt</td>
</tr>
<tr>
<td>32230</td>
<td>Chlorophyll a Phytoplankton spectrophotometric uncorrected, ug/l</td>
<td>CHLAespec</td>
</tr>
<tr>
<td>49701</td>
<td>Secchi Disk Water Unfiltered, feet</td>
<td>SDunfil</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>70300</td>
<td>Dissolved Solids (residue on evaporation *180C)</td>
<td>DSE</td>
</tr>
<tr>
<td>70301</td>
<td>Dissolved Solids (sum of constituents), mg/l</td>
<td>DS</td>
</tr>
<tr>
<td>70331</td>
<td>Sediment suspended percent finer than 0.062 mm sieve diameter</td>
<td>FSS</td>
</tr>
<tr>
<td>70507</td>
<td>Phosphorous Orthophosphate Total</td>
<td>POT</td>
</tr>
<tr>
<td>80154</td>
<td>Total suspended solid sediment suspended, mg/l</td>
<td>TSS</td>
</tr>
<tr>
<td>82078</td>
<td>Turbidity (field)</td>
<td>TURBf</td>
</tr>
<tr>
<td>82079</td>
<td>Turbidity (lab)</td>
<td>TURBlab</td>
</tr>
</tbody>
</table>
## The MEANS Procedure (For complete dataset 'lakeall')

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>593648</td>
<td>994143939</td>
<td>2.6520443</td>
<td>-234057600</td>
<td>1256083200</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>324325</td>
<td>0.0787600</td>
<td>2.6520443</td>
<td>0</td>
<td>1486.00</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>20826</td>
<td>0.0088409</td>
<td>0.0185637</td>
<td>0</td>
<td>0.3300000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>123774</td>
<td>7.7206583</td>
<td>28.5233095</td>
<td>0</td>
<td>7400.00</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>91400</td>
<td>0.9422286</td>
<td>0.7932359</td>
<td>0</td>
<td>22.5000000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>67622</td>
<td>0.1222654</td>
<td>0.2336343</td>
<td>0</td>
<td>4.4000000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>78850</td>
<td>0.4270715</td>
<td>2.6124017</td>
<td>-0.0600000</td>
<td>650.0000000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>214641</td>
<td>20.0803541</td>
<td>47.0852804</td>
<td>0</td>
<td>17000.00</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>82970</td>
<td>3.3681708</td>
<td>534.4894018</td>
<td>-0.0100000</td>
<td>88888.00</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>55597</td>
<td>7.6195082</td>
<td>14.4223530</td>
<td>0</td>
<td>1340.00</td>
</tr>
<tr>
<td>TURBF</td>
<td></td>
<td>3706</td>
<td>13.9358770</td>
<td>22.6166462</td>
<td>0</td>
<td>191.3000000</td>
</tr>
<tr>
<td>CHLAttri</td>
<td></td>
<td>79572</td>
<td>20.9721048</td>
<td>27.2996871</td>
<td>0</td>
<td>160.5500000</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>225988</td>
<td>3.3124659</td>
<td>2.2534222</td>
<td>0</td>
<td>23.0000000</td>
</tr>
<tr>
<td>SECCHIn</td>
<td></td>
<td>90275</td>
<td>55.8356245</td>
<td>45.9237217</td>
<td>0</td>
<td>288.0000000</td>
</tr>
<tr>
<td>TURBLAB</td>
<td></td>
<td>11500</td>
<td>14.2437383</td>
<td>19.9257704</td>
<td>0</td>
<td>80.00000000</td>
</tr>
<tr>
<td>TF</td>
<td></td>
<td>400</td>
<td>56.6540000</td>
<td>14.1786537</td>
<td>32.00000000</td>
<td>80.0000000</td>
</tr>
<tr>
<td>CHLAphty</td>
<td></td>
<td>33416</td>
<td>21.7680154</td>
<td>28.1039865</td>
<td>0</td>
<td>254.5200000</td>
</tr>
<tr>
<td>TON</td>
<td></td>
<td>72277</td>
<td>2.7928538</td>
<td>331.0344305</td>
<td>-0.0600000</td>
<td>88888.00</td>
</tr>
<tr>
<td>TN</td>
<td></td>
<td>163838</td>
<td>2.2106112</td>
<td>311.7374630</td>
<td>-9999.00</td>
<td>88889.00</td>
</tr>
<tr>
<td>CHLA</td>
<td></td>
<td>15816</td>
<td>10.8619028</td>
<td>17.5863537</td>
<td>0</td>
<td>607.0000000</td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>106994</td>
<td>8.8843761</td>
<td>385.091860</td>
<td>0.21000000</td>
<td>88888.00</td>
</tr>
<tr>
<td>ALKueq</td>
<td></td>
<td>129</td>
<td>230.0375194</td>
<td>324.3410432</td>
<td>-51.8000000</td>
<td>1897.53</td>
</tr>
<tr>
<td>CHLAspec</td>
<td></td>
<td>24265</td>
<td>6.8579485</td>
<td>12.0245549</td>
<td>0</td>
<td>247.0000000</td>
</tr>
<tr>
<td>TSS</td>
<td></td>
<td>34147</td>
<td>18.6041170</td>
<td>503.6363311</td>
<td>-12.0000000</td>
<td>88888.00</td>
</tr>
<tr>
<td>ALKmg</td>
<td></td>
<td>34868</td>
<td>59.7922105</td>
<td>958.2821230</td>
<td>0</td>
<td>88888.00</td>
</tr>
<tr>
<td>TURBjcu</td>
<td></td>
<td>779</td>
<td>28.1765083</td>
<td>28.0546497</td>
<td>0</td>
<td>280.0000000</td>
</tr>
<tr>
<td>OPD</td>
<td></td>
<td>4554</td>
<td>2.7090051</td>
<td>10.4509987</td>
<td>0</td>
<td>335.0000000</td>
</tr>
<tr>
<td>DSE</td>
<td></td>
<td>290</td>
<td>5.8912414</td>
<td>10.9541763</td>
<td>0</td>
<td>167.0000000</td>
</tr>
<tr>
<td>DOsat</td>
<td></td>
<td>55992</td>
<td>713.947894</td>
<td>112813.08</td>
<td>0</td>
<td>24500000.00</td>
</tr>
<tr>
<td>ISF</td>
<td></td>
<td>568</td>
<td>279.7253996</td>
<td>737.2001510</td>
<td>0</td>
<td>5900.00</td>
</tr>
<tr>
<td>ALK</td>
<td></td>
<td>38134</td>
<td>16.2136201</td>
<td>20.9381481</td>
<td>0</td>
<td>273.0000000</td>
</tr>
<tr>
<td>DP</td>
<td></td>
<td>4342</td>
<td>0.0082741</td>
<td>0.0111490</td>
<td>0</td>
<td>0.4000000</td>
</tr>
<tr>
<td>POT</td>
<td></td>
<td>426</td>
<td>3.5020634</td>
<td>19.0629522</td>
<td>0</td>
<td>112.0000000</td>
</tr>
<tr>
<td>NH3D</td>
<td></td>
<td>484</td>
<td>0.0256198</td>
<td>0.0424682</td>
<td>0</td>
<td>0.5000000</td>
</tr>
<tr>
<td>NND</td>
<td></td>
<td>502</td>
<td>0.2143426</td>
<td>0.2164711</td>
<td>0</td>
<td>2.8000000</td>
</tr>
<tr>
<td>SIL</td>
<td></td>
<td>3444</td>
<td>1.5634782</td>
<td>1.1741162</td>
<td>0.05000000</td>
<td>6.7200000</td>
</tr>
</tbody>
</table>

(table contd.)
### The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>105</td>
<td>1027061486</td>
<td>2631355.10</td>
<td>948944800</td>
<td>1067558400</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>28</td>
<td>0.2535714</td>
<td>0.1177321</td>
<td>0.0800000</td>
<td>0.5400000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>25</td>
<td>9.7160000</td>
<td>1.5531366</td>
<td>7.3000000</td>
<td>12.8000000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>29</td>
<td>0.8241379</td>
<td>0.3979009</td>
<td>0.3000000</td>
<td>1.7000000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>29</td>
<td>0.0393103</td>
<td>0.0377899</td>
<td>0.0100000</td>
<td>0.1700000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>26</td>
<td>0.1392308</td>
<td>0.1699394</td>
<td>0.0100000</td>
<td>0.7000000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>96</td>
<td>19.2552083</td>
<td>6.2458385</td>
<td>3.0000000</td>
<td>30.5000000</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>28</td>
<td>44.8928571</td>
<td>37.252326</td>
<td>8.0000000</td>
<td>167.0000000</td>
</tr>
<tr>
<td>TURBf</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLAttri</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>91</td>
<td>0.8852747</td>
<td>1.0448268</td>
<td>0.1000000</td>
<td>3.7000000</td>
</tr>
<tr>
<td>SECCHIin</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TURBLAB</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TF</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLAphty</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TON</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TN</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLA</td>
<td></td>
<td>16</td>
<td>16.6625000</td>
<td>14.7921545</td>
<td>2.2000000</td>
<td>46.0000000</td>
</tr>
</tbody>
</table>

(table contd.)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>9551</td>
<td>990956274</td>
<td>180621853</td>
<td>111110400</td>
<td>1245628800</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>7189</td>
<td>0.0402291</td>
<td>0.1378137</td>
<td>0</td>
<td>9.3000000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>343</td>
<td>0.0069388</td>
<td>0.0085730</td>
<td>0</td>
<td>0.0500000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>5112</td>
<td>6.8597164</td>
<td>3.3895434</td>
<td>0</td>
<td>80.6000000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>2395</td>
<td>0.3714363</td>
<td>0.4595310</td>
<td>0</td>
<td>4.7000000</td>
</tr>
</tbody>
</table>

(table contd.)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value 1</th>
<th>Value 2</th>
<th>Value 3</th>
<th>Value 4</th>
<th>Value 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH3</td>
<td>3778</td>
<td>0.0616665</td>
<td>0.1322849</td>
<td>0</td>
<td>1.8000000</td>
</tr>
<tr>
<td>NO₂NO₃</td>
<td>1687</td>
<td>0.0680237</td>
<td>0.1841811</td>
<td>0</td>
<td>1.6000000</td>
</tr>
<tr>
<td>TEMP</td>
<td>6956</td>
<td>13.6972772</td>
<td>6.0184815</td>
<td>0</td>
<td>30.5000000</td>
</tr>
<tr>
<td>NO₃</td>
<td>759</td>
<td>0.0519368</td>
<td>0.2259156</td>
<td>0</td>
<td>5.4100000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td>953</td>
<td>2.8647429</td>
<td>5.6189367</td>
<td>0</td>
<td>82.0000000</td>
</tr>
<tr>
<td>TURBf</td>
<td>11.0000000</td>
<td>11.0000000</td>
<td>11.0000000</td>
<td>0</td>
<td>11.0000000</td>
</tr>
<tr>
<td>CHLₐtri</td>
<td>25</td>
<td>8.0580000</td>
<td>18.4325102</td>
<td>0</td>
<td>90.7000000</td>
</tr>
<tr>
<td>SECCHI</td>
<td>4139</td>
<td>3.0670742</td>
<td>2.0127659</td>
<td>0</td>
<td>16.5000000</td>
</tr>
<tr>
<td>SECCHIₐn</td>
<td>153</td>
<td>97.1960784</td>
<td>35.9930656</td>
<td>32.0000000</td>
<td>215.0000000</td>
</tr>
<tr>
<td>TURBLAB</td>
<td>105</td>
<td>2.3789524</td>
<td>3.3524534</td>
<td>0</td>
<td>18.0000000</td>
</tr>
<tr>
<td>CHLₐphyt</td>
<td>1318</td>
<td>7.6222762</td>
<td>13.1638510</td>
<td>0</td>
<td>140.3000000</td>
</tr>
<tr>
<td>TN</td>
<td>1225</td>
<td>0.3914245</td>
<td>0.4095719</td>
<td>0.0090000</td>
<td>5.5900000</td>
</tr>
<tr>
<td>CHLA</td>
<td>1717</td>
<td>9.9983286</td>
<td>14.7885104</td>
<td>0</td>
<td>123.0000000</td>
</tr>
<tr>
<td>COND</td>
<td>1643</td>
<td>120.6418138</td>
<td>292.5072725</td>
<td>-10.0000000</td>
<td>720.00</td>
</tr>
<tr>
<td>pH</td>
<td>1732</td>
<td>3.463510</td>
<td>1.8199349</td>
<td>1.0400000</td>
<td>76.0000000</td>
</tr>
<tr>
<td>ALKueq</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLₐspec</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TSS</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALKmg</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TURBjcu</td>
<td>1</td>
<td>1.0000000</td>
<td>.</td>
<td>1.0000000</td>
<td>1.0000000</td>
</tr>
<tr>
<td>OPD</td>
<td>604</td>
<td>0.0034553</td>
<td>0.0098476</td>
<td>0</td>
<td>0.1140000</td>
</tr>
<tr>
<td>DSE</td>
<td>150</td>
<td>5.2697333</td>
<td>4.7710655</td>
<td>0</td>
<td>25.0000000</td>
</tr>
<tr>
<td>DOsat</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ISF</td>
<td>16</td>
<td>29.0418750</td>
<td>23.4422272</td>
<td>0.1000000</td>
<td>77.3900000</td>
</tr>
<tr>
<td>ALK</td>
<td>494</td>
<td>17.1174069</td>
<td>17.4236228</td>
<td>0</td>
<td>116.0000000</td>
</tr>
<tr>
<td>DP</td>
<td>277</td>
<td>0.0036679</td>
<td>0.0132467</td>
<td>0</td>
<td>0.1700000</td>
</tr>
<tr>
<td>POT</td>
<td>10</td>
<td>9.4010000</td>
<td>29.7250588</td>
<td>0</td>
<td>94.0000000</td>
</tr>
<tr>
<td>NH₃D</td>
<td>188</td>
<td>0.0037234</td>
<td>0.0058473</td>
<td>0</td>
<td>0.0300000</td>
</tr>
<tr>
<td>NND</td>
<td>206</td>
<td>0.0674757</td>
<td>0.2386148</td>
<td>0</td>
<td>2.8000000</td>
</tr>
<tr>
<td>SIL</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>PTR</td>
<td>40</td>
<td>0.0116250</td>
<td>0.0403762</td>
<td>0.0030000</td>
<td>0.2500000</td>
</tr>
<tr>
<td>TS</td>
<td>29</td>
<td>89.2758621</td>
<td>8.6018958</td>
<td>79.0000000</td>
<td>100.0000000</td>
</tr>
<tr>
<td>pHSU</td>
<td>154</td>
<td>6.6701299</td>
<td>0.8887187</td>
<td>3.4000000</td>
<td>7.7000000</td>
</tr>
<tr>
<td>CF</td>
<td>122</td>
<td>174.3278689</td>
<td>115.3113961</td>
<td>24.0000000</td>
<td>540.0000000</td>
</tr>
<tr>
<td>RTN</td>
<td>91</td>
<td>7.6791209</td>
<td>7.2155621</td>
<td>0</td>
<td>31.0000000</td>
</tr>
<tr>
<td>NKD</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALKP</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ND</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
| FSS           | 10      | 161.4000000 | 13.1419599 | 150.0000000 | 191.0000000 | (table contd.)
### Ecoregion 3

The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>3001</td>
<td>1071082950</td>
<td>69049397.35</td>
<td>948153600</td>
<td>1243209600</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>2803</td>
<td>0.0810963</td>
<td>0.1439799</td>
<td>0</td>
<td>2.0300000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>55</td>
<td>0.0071091</td>
<td>0.0080385</td>
<td>0.0010000</td>
<td>0.0500000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>2330</td>
<td>7.0891502</td>
<td>2.6507863</td>
<td>0.1000000</td>
<td>14.5000000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>1529</td>
<td>0.6097646</td>
<td>0.5258962</td>
<td>0.0100000</td>
<td>5.1200000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>2578</td>
<td>0.0695132</td>
<td>0.1794231</td>
<td>0</td>
<td>2.1700000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>290</td>
<td>0.5860069</td>
<td>0.7738849</td>
<td>0.0010000</td>
<td>2.3100000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>2508</td>
<td>17.0769537</td>
<td>5.5817108</td>
<td>0.4000000</td>
<td>31.0000000</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>47</td>
<td>0.1787234</td>
<td>0.2550884</td>
<td>0.0050000</td>
<td>0.9600000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>597</td>
<td>8.0212730</td>
<td>16.2099905</td>
<td>0</td>
<td>150.000000</td>
</tr>
<tr>
<td>TURBf</td>
<td></td>
<td>15</td>
<td>18.673333</td>
<td>25.2097336</td>
<td>3.3000000</td>
<td>94.1000000</td>
</tr>
<tr>
<td>CHLAttri</td>
<td></td>
<td>37</td>
<td>23.463783</td>
<td>32.850983</td>
<td>0.6000000</td>
<td>131.000000</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>1128</td>
<td>5.9671233</td>
<td>11.2669139</td>
<td>0.5000000</td>
<td>52.0000000</td>
</tr>
<tr>
<td>TURBLAB</td>
<td></td>
<td>73</td>
<td>5.9671233</td>
<td>11.2669139</td>
<td>0.5000000</td>
<td>52.0000000</td>
</tr>
<tr>
<td>TF</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLApHyt</td>
<td></td>
<td>1045</td>
<td>11.538086</td>
<td>19.4087674</td>
<td>0</td>
<td>154.200000</td>
</tr>
<tr>
<td>TON</td>
<td></td>
<td>1</td>
<td>2.5000000</td>
<td>2.5000000</td>
<td>2.5000000</td>
<td>2.5000000</td>
</tr>
<tr>
<td>TN</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLA</td>
<td></td>
<td>36</td>
<td>24.866666</td>
<td>20.8247665</td>
<td>0.7000000</td>
<td>80.0000000</td>
</tr>
<tr>
<td>COND</td>
<td></td>
<td>64</td>
<td>423.304687</td>
<td>87.0817267</td>
<td>3.0000000</td>
<td>615.000000</td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>155</td>
<td>8.2074839</td>
<td>0.3926956</td>
<td>6.3000000</td>
<td>9.3000000</td>
</tr>
<tr>
<td>ALKueq</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLAspec</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TSS</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALKmg</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TURBjcu</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>OPD</td>
<td></td>
<td>140</td>
<td>0.0628643</td>
<td>0.0699994</td>
<td>0</td>
<td>0.6350000</td>
</tr>
<tr>
<td>DSE</td>
<td></td>
<td>140</td>
<td>6.5571429</td>
<td>14.9739685</td>
<td>2.0000000</td>
<td>167.000000</td>
</tr>
<tr>
<td>DOsat</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

(table contd.)
### Ecoregion 4

#### The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>2373</td>
<td>1055808000</td>
<td>84637860.29</td>
<td>465523200</td>
<td>1215561600</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>1525</td>
<td>0.0849016</td>
<td>0.1357922</td>
<td>0</td>
<td>1.6600000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>676</td>
<td>0.0121612</td>
<td>0.0223142</td>
<td>0.0010000</td>
<td>0.3000000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>813</td>
<td>7.6377614</td>
<td>2.3967986</td>
<td>0.1000000</td>
<td>14.4000000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>1235</td>
<td>0.7572955</td>
<td>0.7315680</td>
<td>0</td>
<td>5.2000000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>1366</td>
<td>0.1049458</td>
<td>0.2015044</td>
<td>0</td>
<td>2.0700000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>1136</td>
<td>0.2120352</td>
<td>0.3283756</td>
<td>0</td>
<td>1.9700000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>1300</td>
<td>18.9302154</td>
<td>6.9499424</td>
<td>0.1000000</td>
<td>31.0000000</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>1136</td>
<td>0.2120352</td>
<td>0.3283756</td>
<td>0</td>
<td>1.9700000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>659</td>
<td>11.1980273</td>
<td>17.6021935</td>
<td>0.3000000</td>
<td>170.0000000</td>
</tr>
<tr>
<td>TURBF</td>
<td></td>
<td>78</td>
<td>33.9910256</td>
<td>25.7105526</td>
<td>4.9000000</td>
<td>161.2000000</td>
</tr>
<tr>
<td>CHLATRI</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>817</td>
<td>1.8675765</td>
<td>1.4475689</td>
<td>0.1000000</td>
<td>7.0000000</td>
</tr>
<tr>
<td>SECCHIIIN</td>
<td></td>
<td>299</td>
<td>75.8026756</td>
<td>49.1738239</td>
<td>2.0000000</td>
<td>214.0000000</td>
</tr>
</tbody>
</table>

(table contd.)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Value</th>
<th>Value</th>
<th>Value</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TURBLAB</td>
<td>27</td>
<td>50.23</td>
<td>49.88</td>
<td>2.20</td>
<td>190.00</td>
</tr>
<tr>
<td>TF</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHLApphyt</td>
<td>220</td>
<td>16.09</td>
<td>18.71</td>
<td>0.20</td>
<td>104.70</td>
</tr>
<tr>
<td>TON</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TN</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHLA</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COND</td>
<td>17</td>
<td>7613.12</td>
<td>10677.37</td>
<td>368.00</td>
<td>45100.00</td>
</tr>
<tr>
<td>pH</td>
<td>19</td>
<td>8.0531579</td>
<td>1.0515547</td>
<td>4.42</td>
<td>9.4300000</td>
</tr>
<tr>
<td>ALKueq</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHLAspec</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALKmg</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TURBjcu</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPD</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSE</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOSat</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISF</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>ALK</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POT</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH3D</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NND</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIL</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTR</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pHSU</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RTN</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NKD</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALKP</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ND</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDunfil</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDS</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Ecoregion 5

#### The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>2732</td>
<td>1053944201</td>
<td>455932800</td>
<td>1215475200</td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>2147</td>
<td>0.1362948</td>
<td>0.4014982</td>
<td>0</td>
<td>6.8300000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>796</td>
<td>0.0202563</td>
<td>0.0293580</td>
<td>0.0010000</td>
<td>0.3000000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>1021</td>
<td>7.2124976</td>
<td>2.9037959</td>
<td>0.0500000</td>
<td>20.0000000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>1927</td>
<td>1.0244369</td>
<td>1.0396480</td>
<td>0</td>
<td>22.5000000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>1543</td>
<td>0.1231905</td>
<td>0.2208612</td>
<td>0</td>
<td>2.0000000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>1560</td>
<td>0.1998846</td>
<td>0.3824704</td>
<td>0</td>
<td>6.5000000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>1733</td>
<td>19.9669359</td>
<td>7.1293248</td>
<td>0.2000000</td>
<td>31.0000000</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>128</td>
<td>0.6865234</td>
<td>1.9325773</td>
<td>0</td>
<td>18.0600000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>621</td>
<td>18.1658615</td>
<td>25.7178429</td>
<td>0.4000000</td>
<td>180.0000000</td>
</tr>
<tr>
<td>TURBf</td>
<td></td>
<td>63</td>
<td>25.9389407</td>
<td>32.9551421</td>
<td>2.3000000</td>
<td>150.0000000</td>
</tr>
<tr>
<td>CHLAttri</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>705</td>
<td>1.3630071</td>
<td>1.2700294</td>
<td>0.1000000</td>
<td>7.3000000</td>
</tr>
<tr>
<td>SECCHIin</td>
<td></td>
<td>318</td>
<td>57.333333</td>
<td>47.2571550</td>
<td>1.0000000</td>
<td>216.0000000</td>
</tr>
<tr>
<td>TURBLAB</td>
<td></td>
<td>42</td>
<td>18.9428571</td>
<td>13.0223287</td>
<td>2.0000000</td>
<td>64.0000000</td>
</tr>
<tr>
<td>TF</td>
<td></td>
<td>9</td>
<td>66.744444</td>
<td>11.4269540</td>
<td>40.1000000</td>
<td>79.4000000</td>
</tr>
<tr>
<td>CHLApHyt</td>
<td></td>
<td>228</td>
<td>27.5338156</td>
<td>28.9253675</td>
<td>0.5000000</td>
<td>129.2000000</td>
</tr>
<tr>
<td>TON</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TN</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLA</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>COND</td>
<td></td>
<td>16</td>
<td>3276.69</td>
<td>2365.50</td>
<td>917.0000000</td>
<td>8650.00</td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>15</td>
<td>8.4893333</td>
<td>0.6620473</td>
<td>7.4100000</td>
<td>9.6000000</td>
</tr>
<tr>
<td>ALKueq</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLAspec</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TSS</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALKmg</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TURBJcu</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>OPD</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSE</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DOSat</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ISF</td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ALK</td>
<td></td>
<td>4</td>
<td>29.0000000</td>
<td>0</td>
<td>29.0000000</td>
<td>29.0000000</td>
</tr>
<tr>
<td>DP</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>POT</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>NH3D</td>
<td></td>
<td>0</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Variable</td>
<td>Label</td>
<td>N</td>
<td>Mean</td>
<td>Std Dev</td>
<td>Minimum</td>
<td>Maximum</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------</td>
<td>-----</td>
<td>------------</td>
<td>-------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>42026</td>
<td>1061710364</td>
<td>69237114.28</td>
<td>946944000</td>
<td>1216080000</td>
</tr>
<tr>
<td>TP</td>
<td>10969</td>
<td>0.1440117</td>
<td>0.2121563</td>
<td></td>
<td>0</td>
<td>2.4000000</td>
</tr>
<tr>
<td>NO2</td>
<td>629</td>
<td>0.0279396</td>
<td>0.0391561</td>
<td>0.0010000</td>
<td>0.3300000</td>
<td></td>
</tr>
<tr>
<td>DO</td>
<td>3582</td>
<td>8.1094640</td>
<td>3.2073938</td>
<td>0.1000000</td>
<td>14.7000000</td>
<td></td>
</tr>
<tr>
<td>TKN</td>
<td>8275</td>
<td>1.2752701</td>
<td>0.8740797</td>
<td></td>
<td>5.5000000</td>
<td></td>
</tr>
<tr>
<td>NH3</td>
<td>9734</td>
<td>0.1793115</td>
<td>0.2767060</td>
<td></td>
<td>2.1000000</td>
<td></td>
</tr>
<tr>
<td>NO2NO3</td>
<td>10348</td>
<td>1.4950291</td>
<td>2.3900716</td>
<td></td>
<td>10.7000000</td>
<td></td>
</tr>
<tr>
<td>TEMP</td>
<td>10980</td>
<td>18.6634699</td>
<td>7.4019453</td>
<td>0.1000000</td>
<td>31.0000000</td>
<td></td>
</tr>
<tr>
<td>NO3</td>
<td>204</td>
<td>0.7368431</td>
<td>1.2950575</td>
<td>8.6000000</td>
<td>94.0000000</td>
<td></td>
</tr>
<tr>
<td>TURBIDITY</td>
<td>4332</td>
<td>9.7290859</td>
<td>15.6668770</td>
<td>0.3000000</td>
<td>182.0000000</td>
<td></td>
</tr>
<tr>
<td>TURBf</td>
<td>426</td>
<td>35.5856808</td>
<td>35.5688410</td>
<td>0.6000000</td>
<td>180.0000000</td>
<td></td>
</tr>
<tr>
<td>CHLAttri</td>
<td>3851</td>
<td>33.9622643</td>
<td>32.8781486</td>
<td>0.0100000</td>
<td>160.0000000</td>
<td></td>
</tr>
<tr>
<td>SECCHI</td>
<td>7385</td>
<td>1.1351401</td>
<td>1.0486217</td>
<td>0.1000000</td>
<td>7.9200000</td>
<td></td>
</tr>
<tr>
<td>SECCHIin</td>
<td>25553</td>
<td>41.3122921</td>
<td>34.8701298</td>
<td>1.0000000</td>
<td>216.0000000</td>
<td></td>
</tr>
<tr>
<td>TURBLAB</td>
<td>86</td>
<td>32.9023256</td>
<td>16.1232530</td>
<td>8.6000000</td>
<td>94.0000000</td>
<td></td>
</tr>
<tr>
<td>TF</td>
<td>2</td>
<td>47.5000000</td>
<td>18.1019336</td>
<td>34.7000000</td>
<td>60.3000000</td>
<td></td>
</tr>
<tr>
<td>CHLApht</td>
<td>4645</td>
<td>32.3337524</td>
<td>31.4692459</td>
<td>0.1800000</td>
<td>159.3100000</td>
<td></td>
</tr>
<tr>
<td>TON</td>
<td>9</td>
<td>2.1566667</td>
<td>0.7172866</td>
<td>1.2000000</td>
<td>3.2000000</td>
<td></td>
</tr>
<tr>
<td>TN</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>CHLA</td>
<td>188</td>
<td>24.6914894</td>
<td>23.0184890</td>
<td>1.0000000</td>
<td>124.0000000</td>
<td></td>
</tr>
<tr>
<td>COND</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
</tbody>
</table>

Ecoregion 6

The MEANS Procedure

(continuation)
### Ecoregion 7

The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>sampling_date</td>
<td>95959</td>
<td>1066886436</td>
<td>107390156</td>
<td>286329600</td>
<td>1256083200</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>43163</td>
<td>0.0966367</td>
<td>0.1846405</td>
<td>0</td>
<td>3.8000000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>802</td>
<td>0.0119925</td>
<td>0.0202864</td>
<td>0</td>
<td>0.3200000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>23363</td>
<td>7.3870535</td>
<td>3.8782855</td>
<td>0.0500000</td>
<td>14.700000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>18115</td>
<td>1.1829616</td>
<td>0.7638099</td>
<td>0</td>
<td>5.5000000</td>
</tr>
</tbody>
</table>

(table contd.)
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NH3</td>
<td>8840</td>
<td>0.1833612</td>
<td>0.3468874</td>
<td>0</td>
<td>2.1700000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td>11717</td>
<td>0.3528088</td>
<td>1.0588926</td>
<td>-0.0600000</td>
<td>11.9000000</td>
</tr>
<tr>
<td>TEMP</td>
<td>31472</td>
<td>16.1549346</td>
<td>7.1289453</td>
<td>0</td>
<td>31.0000000</td>
</tr>
<tr>
<td>NO3</td>
<td>540</td>
<td>0.0710370</td>
<td>0.3532226</td>
<td>0.0030000</td>
<td>7.8000000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td>10124</td>
<td>6.1662339</td>
<td>10.146756</td>
<td>0.0600000</td>
<td>11.9000000</td>
</tr>
<tr>
<td>TURBF</td>
<td>645</td>
<td>15.3073298</td>
<td>17.5532979</td>
<td>1.8000000</td>
<td>196.0000000</td>
</tr>
<tr>
<td>CHLAttri</td>
<td>8853</td>
<td>22.4629789</td>
<td>25.6432150</td>
<td>0.0040000</td>
<td>159.1300000</td>
</tr>
<tr>
<td>SECCHI</td>
<td>67745</td>
<td>2.6243479</td>
<td>1.7663243</td>
<td>0</td>
<td>14.3000000</td>
</tr>
<tr>
<td>TURBf</td>
<td>645</td>
<td>15.3073298</td>
<td>17.5532979</td>
<td>1.8000000</td>
<td>196.0000000</td>
</tr>
<tr>
<td>CHLApphyt</td>
<td>9065</td>
<td>22.4629789</td>
<td>25.6432150</td>
<td>0.0040000</td>
<td>159.1300000</td>
</tr>
<tr>
<td>TON</td>
<td>42</td>
<td>1.4666667</td>
<td>0.6237834</td>
<td>0.5000000</td>
<td>2.8000000</td>
</tr>
<tr>
<td>TN</td>
<td>3908</td>
<td>0.5954831</td>
<td>3.062136</td>
<td>0.0200000</td>
<td>205.8000000</td>
</tr>
<tr>
<td>CHLA</td>
<td>6277</td>
<td>8.0731963</td>
<td>15.2203786</td>
<td>0.1100000</td>
<td>607.0000000</td>
</tr>
<tr>
<td>COND</td>
<td>199</td>
<td>184.0502513</td>
<td>69.5321616</td>
<td>47.0000000</td>
<td>570.0000000</td>
</tr>
<tr>
<td>pH</td>
<td>2550</td>
<td>7.7703804</td>
<td>0.4295751</td>
<td>4.9000000</td>
<td>9.6600000</td>
</tr>
<tr>
<td>ALKueq</td>
<td>1</td>
<td>1764.56</td>
<td>.</td>
<td>1764.56</td>
<td>1764.56</td>
</tr>
<tr>
<td>CHLA spec</td>
<td>2279</td>
<td>10.8478412</td>
<td>26.4008527</td>
<td>.</td>
<td>247.0000000</td>
</tr>
<tr>
<td>TSS</td>
<td>3459</td>
<td>1.7235230</td>
<td>2.498953</td>
<td>0</td>
<td>36.9000000</td>
</tr>
<tr>
<td>AlKmg</td>
<td>1263</td>
<td>50.8989311</td>
<td>11.3675744</td>
<td>21.3000000</td>
<td>104.5000000</td>
</tr>
<tr>
<td>TURBjcu</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>OPD</td>
<td>75</td>
<td>0.0349333</td>
<td>0.0364886</td>
<td>0.0100000</td>
<td>0.1700000</td>
</tr>
<tr>
<td>DSE</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>D0sat</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ISF</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALK</td>
<td>1137</td>
<td>52.3948989</td>
<td>20.9924049</td>
<td>1.0000000</td>
<td>273.0000000</td>
</tr>
<tr>
<td>DP</td>
<td>3675</td>
<td>0.0089189</td>
<td>0.0078251</td>
<td>0</td>
<td>0.1100000</td>
</tr>
<tr>
<td>POT</td>
<td>371</td>
<td>0.6088113</td>
<td>8.2118085</td>
<td>0</td>
<td>112.0000000</td>
</tr>
<tr>
<td>NH3D</td>
<td>37</td>
<td>0.1270270</td>
<td>0.0732145</td>
<td>0.1000000</td>
<td>0.5000000</td>
</tr>
<tr>
<td>NND</td>
<td>37</td>
<td>0.1270270</td>
<td>0.0732145</td>
<td>0.1000000</td>
<td>0.5000000</td>
</tr>
<tr>
<td>SIL</td>
<td>3443</td>
<td>1.5629448</td>
<td>1.1738694</td>
<td>0.0500000</td>
<td>6.7200000</td>
</tr>
<tr>
<td>PTR</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TS</td>
<td>3</td>
<td>139.6666667</td>
<td>26.2741952</td>
<td>124.0000000</td>
<td>170.0000000</td>
</tr>
<tr>
<td>pHSU</td>
<td>1137</td>
<td>7.1012313</td>
<td>0.3313150</td>
<td>5.8000000</td>
<td>9.1000000</td>
</tr>
<tr>
<td>CF</td>
<td>252</td>
<td>206.9880952</td>
<td>77.3295057</td>
<td>52.0000000</td>
<td>650.0000000</td>
</tr>
<tr>
<td>RTN</td>
<td>161</td>
<td>13.8807453</td>
<td>26.2592782</td>
<td>0.8000000</td>
<td>239.0000000</td>
</tr>
<tr>
<td>NDK</td>
<td>34</td>
<td>0.9647059</td>
<td>0.5515171</td>
<td>0.2000000</td>
<td>1.9000000</td>
</tr>
<tr>
<td>ALKP</td>
<td>4</td>
<td>12.0000000</td>
<td>0</td>
<td>12.0000000</td>
<td>12.0000000</td>
</tr>
<tr>
<td>ND</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
| FSS | 0 | . | . | . | . | (table contd.)
## Ecoregion 8

The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>146623</td>
<td>955810070</td>
<td>189186600</td>
<td>-234057600</td>
<td>1251158400</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>46788</td>
<td>0.0304025</td>
<td>0.1236048</td>
<td>0</td>
<td>15.0000000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>103</td>
<td>0.0087184</td>
<td>0.0094458</td>
<td>0.0010000</td>
<td>0.0700000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>6150</td>
<td>7.5126667</td>
<td>3.6835174</td>
<td>0</td>
<td>14.6000000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>9084</td>
<td>0.3994570</td>
<td>0.4115512</td>
<td>0</td>
<td>6.1000000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>6356</td>
<td>0.1050746</td>
<td>0.1383819</td>
<td>0</td>
<td>4.4000000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>9208</td>
<td>0.2328394</td>
<td>1.1898907</td>
<td>-0.0500000</td>
<td>67.0000000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>33736</td>
<td>16.4048405</td>
<td>6.7453273</td>
<td>0</td>
<td>62.0000000</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>1122</td>
<td>0.2730214</td>
<td>0.8793994</td>
<td>0</td>
<td>25.5000000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>9075</td>
<td>3.0845950</td>
<td>3.2021525</td>
<td>0</td>
<td>110.0000000</td>
</tr>
<tr>
<td>TURBF</td>
<td></td>
<td>960</td>
<td>0.8032083</td>
<td>1.1356710</td>
<td>0.0800000</td>
<td>26.7000000</td>
</tr>
<tr>
<td>CHLAttri</td>
<td></td>
<td>1995</td>
<td>10.6313183</td>
<td>13.4006272</td>
<td>0.0100000</td>
<td>155.0000000</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>102067</td>
<td>4.3748740</td>
<td>2.0893759</td>
<td>0</td>
<td>23.0000000</td>
</tr>
<tr>
<td>SECCHIin</td>
<td></td>
<td>1082</td>
<td>75.3049908</td>
<td>54.8713014</td>
<td>1.0000000</td>
<td>216.0000000</td>
</tr>
<tr>
<td>TURBLAB</td>
<td></td>
<td>85</td>
<td>1.1917647</td>
<td>0.7467727</td>
<td>0.3000000</td>
<td>4.6000000</td>
</tr>
<tr>
<td>TF</td>
<td></td>
<td>67</td>
<td>52.6716418</td>
<td>14.5439648</td>
<td>32.0000000</td>
<td>80.0000000</td>
</tr>
<tr>
<td>CHLApht</td>
<td></td>
<td>2059</td>
<td>6.4681350</td>
<td>9.1875785</td>
<td>0.0600000</td>
<td>131.0000000</td>
</tr>
<tr>
<td>TON</td>
<td></td>
<td>65</td>
<td>0.4201538</td>
<td>0.1897489</td>
<td>0.0500000</td>
<td>1.0400000</td>
</tr>
<tr>
<td>TN</td>
<td></td>
<td>560</td>
<td>-17.2944107</td>
<td>422.5588666</td>
<td>-9999.00</td>
<td>4.6500000</td>
</tr>
<tr>
<td>CHLA</td>
<td></td>
<td>2088</td>
<td>8.3959444</td>
<td>17.6082905</td>
<td>0.0600000</td>
<td>305.0000000</td>
</tr>
<tr>
<td>COND</td>
<td></td>
<td>6485</td>
<td>64.3592136</td>
<td>64.5107199</td>
<td>2.0000000</td>
<td>956.0000000</td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>15495</td>
<td>6.3067189</td>
<td>0.7599801</td>
<td>2.9600000</td>
<td>9.7300000</td>
</tr>
<tr>
<td>ALKueq</td>
<td></td>
<td>100</td>
<td>201.7628000</td>
<td>293.0419409</td>
<td>-12.4200000</td>
<td>1897.53</td>
</tr>
<tr>
<td>CHLAspec</td>
<td></td>
<td>18676</td>
<td>6.6546262</td>
<td>9.4119963</td>
<td>0</td>
<td>234.8000000</td>
</tr>
<tr>
<td>TSS</td>
<td></td>
<td>132</td>
<td>2.0107455</td>
<td>2.5201238</td>
<td>0</td>
<td>16.0000000</td>
</tr>
<tr>
<td>ALKmg</td>
<td></td>
<td>2617</td>
<td>6.1758884</td>
<td>7.4283608</td>
<td>0.1000000</td>
<td>120.0000000</td>
</tr>
<tr>
<td>TURBJcu</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>OPD</td>
<td></td>
<td>844</td>
<td>14.155339</td>
<td>20.5673612</td>
<td>0.0100000</td>
<td>335.0000000</td>
</tr>
<tr>
<td>DSE</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DOsat</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

(table contd.)
## Ecoregion 9

The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>37210</td>
<td>1049009892</td>
<td>106340822</td>
<td>326505600</td>
<td>121608000</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>20247</td>
<td>0.1102134</td>
<td>0.9063010</td>
<td>0</td>
<td>59.000000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>3513</td>
<td>0.0185989</td>
<td>0.0294614</td>
<td>0</td>
<td>0.330000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>10361</td>
<td>7.9556269</td>
<td>2.5000834</td>
<td>0</td>
<td>17.200000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>14676</td>
<td>0.6713219</td>
<td>0.5798187</td>
<td>0</td>
<td>5.420000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>15157</td>
<td>0.1212204</td>
<td>0.2372204</td>
<td>0</td>
<td>2.140000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>13728</td>
<td>0.2334268</td>
<td>0.5247336</td>
<td>0</td>
<td>10.200000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>16111</td>
<td>20.7822469</td>
<td>7.2177746</td>
<td>0.1000000</td>
<td>36.400000</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>4019</td>
<td>0.5959473</td>
<td>1.4709105</td>
<td>0</td>
<td>52.870000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>8151</td>
<td>11.7404122</td>
<td>17.5039062</td>
<td>0.3000000</td>
<td>190.000000</td>
</tr>
<tr>
<td>TURBF</td>
<td></td>
<td>434</td>
<td>14.9923963</td>
<td>20.7341550</td>
<td>1.2000000</td>
<td>191.300000</td>
</tr>
<tr>
<td>CHLatri</td>
<td></td>
<td>3474</td>
<td>26.4100835</td>
<td>29.9299757</td>
<td>0</td>
<td>160.550000</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>6510</td>
<td>1.1311982</td>
<td>0.8978058</td>
<td>0.1000000</td>
<td>7.800000</td>
</tr>
<tr>
<td>SECCHIn</td>
<td></td>
<td>14687</td>
<td>37.3150405</td>
<td>30.1701937</td>
<td>1.0000000</td>
<td>216.000000</td>
</tr>
</tbody>
</table>

(table contd.)
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TURBLAB</td>
<td>1282</td>
</tr>
<tr>
<td>TF</td>
<td>45</td>
</tr>
<tr>
<td>CHLApht</td>
<td>2944</td>
</tr>
<tr>
<td>TON</td>
<td>1565</td>
</tr>
<tr>
<td>TN</td>
<td>3768</td>
</tr>
<tr>
<td>CHLA</td>
<td>4299</td>
</tr>
<tr>
<td>COND</td>
<td>1084</td>
</tr>
<tr>
<td>pH</td>
<td>1191</td>
</tr>
<tr>
<td>ALKueq</td>
<td>0</td>
</tr>
<tr>
<td>CHLAspec</td>
<td>0</td>
</tr>
<tr>
<td>TSS</td>
<td>715</td>
</tr>
<tr>
<td>ALKmg</td>
<td>918</td>
</tr>
<tr>
<td>TURBjcu</td>
<td>23</td>
</tr>
<tr>
<td>OPD</td>
<td>0</td>
</tr>
<tr>
<td>DSE</td>
<td>0</td>
</tr>
<tr>
<td>DOsat</td>
<td>1404</td>
</tr>
<tr>
<td>ISF</td>
<td>20</td>
</tr>
<tr>
<td>ALK</td>
<td>0</td>
</tr>
<tr>
<td>DP</td>
<td>0</td>
</tr>
<tr>
<td>POT</td>
<td>0</td>
</tr>
<tr>
<td>NH3D</td>
<td>0</td>
</tr>
<tr>
<td>NND</td>
<td>0</td>
</tr>
<tr>
<td>SIL</td>
<td>0</td>
</tr>
<tr>
<td>PTR</td>
<td>0</td>
</tr>
<tr>
<td>TS</td>
<td>0</td>
</tr>
<tr>
<td>pHSU</td>
<td>0</td>
</tr>
<tr>
<td>CF</td>
<td>0</td>
</tr>
<tr>
<td>RTN</td>
<td>0</td>
</tr>
<tr>
<td>NKD</td>
<td>0</td>
</tr>
<tr>
<td>ALKP</td>
<td>0</td>
</tr>
<tr>
<td>ND</td>
<td>0</td>
</tr>
<tr>
<td>FSS</td>
<td>0</td>
</tr>
<tr>
<td>SDunfil</td>
<td>0</td>
</tr>
<tr>
<td>MDS</td>
<td>0</td>
</tr>
</tbody>
</table>
## Ecoregion 10

**The MEANS Procedure**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>1452</td>
<td>1066307861</td>
<td>74097532.35</td>
<td>947376000</td>
<td>1205107200</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>1354</td>
<td>0.1218936</td>
<td>0.1244424</td>
<td>0.0050000</td>
<td>0.1000000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>130</td>
<td>0.0235769</td>
<td>0.0255955</td>
<td>0.0050000</td>
<td>0.1000000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>1227</td>
<td>8.0116544</td>
<td>2.3434723</td>
<td>0.1000000</td>
<td>14.5000000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>1082</td>
<td>0.8502865</td>
<td>0.5355254</td>
<td>0.0500000</td>
<td>4.5900000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>463</td>
<td>0.0765983</td>
<td>0.1714552</td>
<td>0.0050000</td>
<td>1.7000000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>1274</td>
<td>0.1665581</td>
<td>0.4299151</td>
<td>0.0050000</td>
<td>9.0200000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>1251</td>
<td>21.3302158</td>
<td>7.2079510</td>
<td>3.0000000</td>
<td>31.0000000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>64</td>
<td>39.6406250</td>
<td>33.9539170</td>
<td>4.4000000</td>
<td>140.0000000</td>
</tr>
<tr>
<td>TURBf</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLatri</td>
<td></td>
<td>19</td>
<td>94.8521053</td>
<td>31.5054931</td>
<td>49.6100000</td>
<td>155.2000000</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>65</td>
<td>0.2544615</td>
<td>0.1175536</td>
<td>0.1000000</td>
<td>0.6500000</td>
</tr>
<tr>
<td>SECCHIin</td>
<td></td>
<td>709</td>
<td>26.9985896</td>
<td>21.8326001</td>
<td>1.0000000</td>
<td>172.0000000</td>
</tr>
<tr>
<td>TURBLAB</td>
<td></td>
<td>813</td>
<td>18.3217712</td>
<td>29.5015468</td>
<td>0.5000000</td>
<td>190.0000000</td>
</tr>
<tr>
<td>TF</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLAlphyt</td>
<td></td>
<td>19</td>
<td>90.7589474</td>
<td>32.8603528</td>
<td>45.3900000</td>
<td>157.5700000</td>
</tr>
<tr>
<td>TON</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TN</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLA</td>
<td></td>
<td>7</td>
<td>77.4428571</td>
<td>47.8074909</td>
<td>7.1000000</td>
<td>126.2000000</td>
</tr>
<tr>
<td>COND</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALKueq</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLAspec</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TSS</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALKmg</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TURBjcu</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>OPD</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSE</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DOSat</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ISF</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALK</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DP</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>POT</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>NH3D</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

*(table contd.)*
### Ecoregion 11

The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>19525</td>
<td>950356094</td>
<td>231976334</td>
<td>286329600</td>
<td>1222560000</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>8783</td>
<td>0.0388849</td>
<td>0.0712246</td>
<td>0</td>
<td>0.9100000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>1534</td>
<td>0.0094465</td>
<td>0.0119808</td>
<td>0</td>
<td>0.1700000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>3286</td>
<td>8.0710590</td>
<td>2.9990175</td>
<td>0</td>
<td>104.000000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>7673</td>
<td>0.4035996</td>
<td>0.4609487</td>
<td>0</td>
<td>6.9000000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>4973</td>
<td>0.1551172</td>
<td>0.2406584</td>
<td>0</td>
<td>2.1200000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>8156</td>
<td>0.4103352</td>
<td>7.2067355</td>
<td>0</td>
<td>650.000000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>16490</td>
<td>16.9344269</td>
<td>7.7076708</td>
<td>0</td>
<td>83.2000000</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>1121</td>
<td>0.4071142</td>
<td>0.7460818</td>
<td>0</td>
<td>8.6600000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>7303</td>
<td>6.4471724</td>
<td>20.0797205</td>
<td>0.1000000</td>
<td>1340.00000</td>
</tr>
<tr>
<td>TURBF</td>
<td></td>
<td>945</td>
<td>15.1442328</td>
<td>22.4935525</td>
<td>0.5000000</td>
<td>190.0000000</td>
</tr>
<tr>
<td>CHLAttri</td>
<td></td>
<td>2012</td>
<td>8.5041501</td>
<td>12.3315369</td>
<td>0.2000000</td>
<td>254.5200000</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>3039</td>
<td>0.1551172</td>
<td>0.2406584</td>
<td>0</td>
<td>2.1200000</td>
</tr>
<tr>
<td>SECCHIin</td>
<td></td>
<td>5848</td>
<td>100.4151847</td>
<td>61.5665955</td>
<td>1.0000000</td>
<td>288.0000000</td>
</tr>
<tr>
<td>TURBLAB</td>
<td></td>
<td>363</td>
<td>15.9407713</td>
<td>27.1384295</td>
<td>0.5000000</td>
<td>197.0000000</td>
</tr>
<tr>
<td>TF</td>
<td></td>
<td>100</td>
<td>0.4821000</td>
<td>0.3346830</td>
<td>0.0400000</td>
<td>1.6500000</td>
</tr>
<tr>
<td>CHLApphyt</td>
<td></td>
<td>61</td>
<td>0.7877049</td>
<td>0.6312062</td>
<td>0.1200000</td>
<td>2.6400000</td>
</tr>
</tbody>
</table>

(table contd.)
### Ecoregion 12

The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>176588</td>
<td>973999659</td>
<td>186903935</td>
<td>315964800</td>
<td>1211155200</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>150597</td>
<td>0.0700447</td>
<td>0.5718296</td>
<td>0</td>
<td>203.0000000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>7295</td>
<td>0.0050647</td>
<td>0.0065821</td>
<td>0</td>
<td>0.1900000 (table contd.)</td>
</tr>
<tr>
<td>DO</td>
<td>57259</td>
<td>7.798670</td>
<td>40.7716127</td>
<td>0</td>
<td>7400.00</td>
<td></td>
</tr>
<tr>
<td>TKN</td>
<td>13953</td>
<td>1.2242392</td>
<td>0.9169208</td>
<td>0</td>
<td>5.500000</td>
<td></td>
</tr>
<tr>
<td>NH3</td>
<td>10261</td>
<td>0.0679010</td>
<td>0.1265280</td>
<td>0</td>
<td>2.160000</td>
<td></td>
</tr>
<tr>
<td>NO2NO3</td>
<td>8531</td>
<td>0.0397799</td>
<td>0.1707230</td>
<td>0</td>
<td>7.560000</td>
<td></td>
</tr>
<tr>
<td>TEMP</td>
<td>75129</td>
<td>24.3214277</td>
<td>78.7808292</td>
<td>0</td>
<td>17000.00</td>
<td></td>
</tr>
<tr>
<td>NO3</td>
<td>58828</td>
<td>4.6217556</td>
<td>634.7541938</td>
<td>-0.0100000</td>
<td>88888.00</td>
<td></td>
</tr>
<tr>
<td>TURBIDITY</td>
<td>7286</td>
<td>4.5128054</td>
<td>6.4401157</td>
<td>0</td>
<td>158.000000</td>
<td></td>
</tr>
<tr>
<td>TURBf</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>CHLatri</td>
<td>55316</td>
<td>19.9274252</td>
<td>27.3416600</td>
<td>0</td>
<td>160.000000</td>
<td></td>
</tr>
<tr>
<td>SECCHI</td>
<td>15269</td>
<td>1.2353055</td>
<td>1.0617322</td>
<td>0.1000000</td>
<td>7.9000000</td>
<td></td>
</tr>
<tr>
<td>SECCHIin</td>
<td>37773</td>
<td>65.0422524</td>
<td>46.9782510</td>
<td>1.0000000</td>
<td>216.0000000</td>
<td></td>
</tr>
<tr>
<td>TURBLAB</td>
<td>8218</td>
<td>11.6459966</td>
<td>14.4245010</td>
<td>0</td>
<td>180.0000000</td>
<td></td>
</tr>
<tr>
<td>TF</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>CHLAphyt</td>
<td>6114</td>
<td>31.2346467</td>
<td>35.0889742</td>
<td>0</td>
<td>158.1500000</td>
<td></td>
</tr>
<tr>
<td>TON</td>
<td>53892</td>
<td>3.0692290</td>
<td>382.8935185</td>
<td>-0.0600000</td>
<td>88888.00</td>
<td></td>
</tr>
<tr>
<td>TN</td>
<td>138519</td>
<td>2.3785926</td>
<td>337.7571956</td>
<td>-0.0600000</td>
<td>88889.00</td>
<td></td>
</tr>
<tr>
<td>CHLA</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>COND</td>
<td>55803</td>
<td>533.4906486</td>
<td>2358.78</td>
<td>0</td>
<td>318000.00</td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>55399</td>
<td>10.7909163</td>
<td>535.0288588</td>
<td>0.2100000</td>
<td>88888.00</td>
<td></td>
</tr>
<tr>
<td>ALKueq</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>CHLAspec</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>TSS</td>
<td>24971</td>
<td>21.7279091</td>
<td>588.8163691</td>
<td>-12.0000000</td>
<td>88888.00</td>
<td></td>
</tr>
<tr>
<td>ALKmg</td>
<td>27488</td>
<td>63.2694836</td>
<td>1078.95</td>
<td>0</td>
<td>88888.00</td>
<td></td>
</tr>
<tr>
<td>TURBjcu</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>OPD</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>DSE</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>D0sat</td>
<td>50187</td>
<td>785.5254202</td>
<td>119158.86</td>
<td>0</td>
<td>24500000.00</td>
<td></td>
</tr>
<tr>
<td>ISF</td>
<td>518</td>
<td>305.6026969</td>
<td>767.0531637</td>
<td>0</td>
<td>5900.00</td>
<td></td>
</tr>
<tr>
<td>ALK</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>DP</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>POT</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>NH3D</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>NND</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>SIL</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>PTR</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>pHSU</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>CF</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>RTN</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>NKD</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>ALKP</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
</tbody>
</table>

(table contd.)
Ecoregion 13

The MEANS Procedure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>21437</td>
<td>977116930</td>
<td>193760378</td>
<td>316569600</td>
<td>1210464000</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>20730</td>
<td>0.1902814</td>
<td>10.3285810</td>
<td>0</td>
<td>1486.00</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>4772</td>
<td>0.0015136</td>
<td>0.0035430</td>
<td>0</td>
<td>0.1200000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>7372</td>
<td>8.2699349</td>
<td>25.6106987</td>
<td>0</td>
<td>1168.00</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>8880</td>
<td>1.4955079</td>
<td>0.5845659</td>
<td>0</td>
<td>5.4800000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>130</td>
<td>0.0531538</td>
<td>0.0969593</td>
<td>0</td>
<td>0.5500000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>8760</td>
<td>0.1121220</td>
<td>0.1470003</td>
<td>0</td>
<td>1.2700000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>8375</td>
<td>24.9861337</td>
<td>4.6287669</td>
<td>2.2000000</td>
<td>37.0000000</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>13614</td>
<td>0.1193996</td>
<td>0.2784066</td>
<td>0</td>
<td>18.7900000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>4916</td>
<td>15.9059398</td>
<td>18.4005552</td>
<td>0</td>
<td>155.0000000</td>
</tr>
<tr>
<td>TURBf</td>
<td></td>
<td>24</td>
<td>8.1875000</td>
<td>9.8232339</td>
<td>1.0000000</td>
<td>38.0000000</td>
</tr>
<tr>
<td>CHLAntri</td>
<td></td>
<td>3985</td>
<td>24.8535257</td>
<td>17.7445825</td>
<td>0</td>
<td>152.5000000</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>4296</td>
<td>0.6151304</td>
<td>0.3210948</td>
<td>0.1000000</td>
<td>4.2500000</td>
</tr>
<tr>
<td>SECCHIin</td>
<td></td>
<td>6</td>
<td>34.0000000</td>
<td>4.8989795</td>
<td>24.0000000</td>
<td>36.0000000</td>
</tr>
<tr>
<td>TURBLAB</td>
<td></td>
<td>24</td>
<td>8.1875000</td>
<td>9.8232339</td>
<td>1.0000000</td>
<td>38.0000000</td>
</tr>
<tr>
<td>TF</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLAnphyt</td>
<td></td>
<td>13</td>
<td>7.2153846</td>
<td>10.1185347</td>
<td>1.6000000</td>
<td>40.1000000</td>
</tr>
<tr>
<td>TON</td>
<td></td>
<td>16094</td>
<td>2.0627387</td>
<td>34.4829477</td>
<td>0</td>
<td>3054.00</td>
</tr>
<tr>
<td>TN</td>
<td></td>
<td>15797</td>
<td>2.2447495</td>
<td>34.8640512</td>
<td>0</td>
<td>3058.00</td>
</tr>
<tr>
<td>CHLA</td>
<td></td>
<td>6</td>
<td>7.2333333</td>
<td>6.0905391</td>
<td>2.2000000</td>
<td>18.5000000</td>
</tr>
<tr>
<td>COND</td>
<td></td>
<td>7419</td>
<td>1702.01</td>
<td>7061.09</td>
<td>0</td>
<td>54400.00</td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>7765</td>
<td>8.2044018</td>
<td>12.9973952</td>
<td>1.0000000</td>
<td>844.0000000</td>
</tr>
<tr>
<td>ALKueq</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLAspec</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TSS</td>
<td></td>
<td>4831</td>
<td>15.6642621</td>
<td>20.8519478</td>
<td>0.1000000</td>
<td>484.0000000</td>
</tr>
<tr>
<td>ALKmg</td>
<td></td>
<td>2099</td>
<td>114.2226489</td>
<td>35.2040134</td>
<td>1.2000000</td>
<td>442.0000000</td>
</tr>
<tr>
<td>TURBJcu</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>OPD</td>
<td></td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

(table contd.)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMPLING_DATE</td>
<td>SAMPLING_DATE</td>
<td>35066</td>
<td>934771180</td>
<td>17093449</td>
<td>329270400</td>
<td>1229299200</td>
</tr>
<tr>
<td>TP</td>
<td></td>
<td>8002</td>
<td>0.0242129</td>
<td>0.0416411</td>
<td>0</td>
<td>0.7800000</td>
</tr>
<tr>
<td>NO2</td>
<td></td>
<td>178</td>
<td>0.0103989</td>
<td>0.0112206</td>
<td>0</td>
<td>0.0700000</td>
</tr>
<tr>
<td>DO</td>
<td></td>
<td>1873</td>
<td>8.5911372</td>
<td>2.1183560</td>
<td>0</td>
<td>14.300000</td>
</tr>
<tr>
<td>TKN</td>
<td></td>
<td>2547</td>
<td>0.5977935</td>
<td>0.5163471</td>
<td>0.0500000</td>
<td>5.5000000</td>
</tr>
<tr>
<td>NH3</td>
<td></td>
<td>2414</td>
<td>0.0575998</td>
<td>0.1024775</td>
<td>0</td>
<td>1.9300000</td>
</tr>
<tr>
<td>NO2NO3</td>
<td></td>
<td>2429</td>
<td>1.2351770</td>
<td>1.7391695</td>
<td>0.0010000</td>
<td>10.450000</td>
</tr>
<tr>
<td>TEMP</td>
<td></td>
<td>8504</td>
<td>19.6206315</td>
<td>5.9807921</td>
<td>0.3000000</td>
<td>31.100000</td>
</tr>
<tr>
<td>NO3</td>
<td></td>
<td>2423</td>
<td>0.9927856</td>
<td>1.5705867</td>
<td>0</td>
<td>24.940000</td>
</tr>
<tr>
<td>TURBIDITY</td>
<td></td>
<td>1488</td>
<td>4.8444220</td>
<td>5.7669392</td>
<td>0.4000000</td>
<td>66.000000</td>
</tr>
<tr>
<td>TURBF</td>
<td></td>
<td>139</td>
<td>1.0847482</td>
<td>1.7020168</td>
<td>0.2000000</td>
<td>17.500000</td>
</tr>
<tr>
<td>CHLatri</td>
<td></td>
<td>5</td>
<td>26.1620000</td>
<td>31.938256</td>
<td>5.6000000</td>
<td>81.800000</td>
</tr>
<tr>
<td>SECCHI</td>
<td></td>
<td>12732</td>
<td>4.9280518</td>
<td>2.4607699</td>
<td>0.1000000</td>
<td>16.000000</td>
</tr>
</tbody>
</table>

(table contd.)
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SECCHIin</td>
<td>1126</td>
<td>38.1456</td>
<td>23.4261</td>
<td>1.0000</td>
<td>108.00</td>
</tr>
<tr>
<td>TURBLAB</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TF</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLApHyt</td>
<td>690</td>
<td>18.1453</td>
<td>27.4796</td>
<td>0.4500</td>
<td>147.00</td>
</tr>
<tr>
<td>TON</td>
<td>509</td>
<td>0.7013</td>
<td>0.5421</td>
<td>0.0200</td>
<td>3.6900</td>
</tr>
<tr>
<td>TN</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CHLA</td>
<td>745</td>
<td>15.0484</td>
<td>26.6927</td>
<td>0.4000</td>
<td>472.00</td>
</tr>
<tr>
<td>COND</td>
<td>750</td>
<td>76.6858</td>
<td>63.8364</td>
<td>2.0000</td>
<td>688.00</td>
</tr>
<tr>
<td>pH</td>
<td>16077</td>
<td>6.4529</td>
<td>0.8394</td>
<td>3.6000</td>
<td>10.7000</td>
</tr>
<tr>
<td>ALKueq</td>
<td>28</td>
<td>276.2143</td>
<td>302.4184</td>
<td>-51.8000</td>
<td>1630.89</td>
</tr>
<tr>
<td>CHLAspec</td>
<td>3310</td>
<td>5.2580</td>
<td>8.1666</td>
<td>0</td>
<td>103.40</td>
</tr>
<tr>
<td>TSS</td>
<td>39</td>
<td>3.5692</td>
<td>4.7755</td>
<td>0</td>
<td>27.6000</td>
</tr>
<tr>
<td>ALKmg</td>
<td>483</td>
<td>7.9892</td>
<td>8.4221</td>
<td>0.6000</td>
<td>83.0000</td>
</tr>
<tr>
<td>TURBjcu</td>
<td>3</td>
<td>10.8667</td>
<td>10.8247</td>
<td>2.2000</td>
<td>23.0000</td>
</tr>
<tr>
<td>OPD</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DSE</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>DOSat</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ISF</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALK</td>
<td>13660</td>
<td>12.0530</td>
<td>12.9968</td>
<td>0.0200</td>
<td>139.50</td>
</tr>
<tr>
<td>DP</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>POT</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>NH3D</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>NND</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>SIL</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>PTR</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>TS</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>pHSU</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>CF</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>RTN</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>NKD</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ALKP</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>ND</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>FSS</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>SDunfil</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>MDS</td>
<td>0</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
VITA

Anindita Das is a citizen of India. She stays in the city of Calcutta, West Bengal, India. She obtained her basic education from various schools in different parts of India. She graduated high school in 1988, from Loreto House, Calcutta. She then acquired her undergraduate degree in zoology from the University of Calcutta in 1991. In 1993, she obtained her graduate degree in zoology from the University of Calcutta. In 1997, she completed a diploma in environmental management from the Indian Institute of Social Welfare and Business Management, Calcutta. She started her masters program in the planning and management concentration in the Department of Environmental Studies in 2001. She is currently a candidate for a Master of Science degree to be awarded in December, 2003.