

2008

# A Diatom Phosphorus Inference Model for 30 Freshwater Lakes in NE Ohio and NW Pennsylvania

Kyle C. Scotese  
*Cleveland State University*

Follow this and additional works at: <https://engagedscholarship.csuohio.edu/etdarchive>

 Part of the [Environmental Sciences Commons](#)

**How does access to this work benefit you? Let us know!**

---

## Recommended Citation

Scotese, Kyle C., "A Diatom Phosphorus Inference Model for 30 Freshwater Lakes in NE Ohio and NW Pennsylvania" (2008). *ETD Archive*. 779.

<https://engagedscholarship.csuohio.edu/etdarchive/779>

This Thesis is brought to you for free and open access by EngagedScholarship@CSU. It has been accepted for inclusion in ETD Archive by an authorized administrator of EngagedScholarship@CSU. For more information, please contact [library.es@csuohio.edu](mailto:library.es@csuohio.edu).

**A DIATOM PHOSPHORUS INFERENCE MODEL FOR 30  
FRESHWATER LAKES IN NE OHIO AND NW PENNSYLVANIA**

**KYLE SCOTESE**

Bachelor of Science in Biology

John Carroll University

May, 2005

Submitted in partial fulfillment of requirement for the degree

MASTER OF SCIENCE IN ENVIRONMENTAL SCIENCE

at the

CLEVELAND STATE UNIVERSITY

December, 2008

This thesis has been approved for  
the Department of Biological, Geological,  
and Environmental Sciences and for the  
College of Graduate Studies of  
Cleveland State University

by

\_\_\_\_\_ Date: \_\_\_\_\_

*Julie Wolin, PhD. BGES/CSU*

Major Advisor

\_\_\_\_\_ Date: \_\_\_\_\_

*Jeff Johansen, PhD. BIOLOGY/JCU*

Advisory Committee Member

\_\_\_\_\_ Date: \_\_\_\_\_

*Robert Wei, PhD. CHEMISTRY/CSU*

Advisory Committee Member

## **DEDICATION**

To my loving father, who shares with me the deepest thirst for scientific knowledge. His life and work inspired me to pursue a career in biology, and his persistent support made my unique dream become a reality. To my caring mother, who graced me with the love and motivation that I needed to overcome the many obstacles. To my brothers and sister, for sharing their passions, which strengthened my drive and ambition. To my grandmothers June and Eleanor, for their nurturing and encouragement. To Erika, for the compassionate understanding and loving acceptance that carried me towards my goal.

## **ACKNOWLEDGEMENTS**

I thank Julie Wolin, Jeff Johansen, Gerald Sgro, Mike Nichols, Robert Wei, Jeff Dean, John Beaver and BSA Environmental Services, Paul Doerder, Ralph Gibson, Erika Anaya, Bruna and Noel Scotese, Erik Alexander, Ed Ostrander, Max Evert, Vincent LaSalvia, Ted Rosati, Pat Finegan, Nate and Mike Jochum, Trevor Moment, Mark Fritz, Kevin Tloczynski, Christina Znidarsic, Allison Yasik, Ynes Arocho, Denny Sampson, Brian Pilarcik, Milt Ostrofsky, Robert Carlson, Robert Davic, and the Ohio EPA.

**A DIATOM PHOSPHORUS INFERENCE MODEL FOR 30 FRESHWATER  
LAKES IN NE OHIO AND NW PENNSYLVANIA**

**KYLE SCOTESE**

**ABSTRACT**

Nutrient enrichment in the form of anthropogenic phosphorous and nitrogen inputs has occurred in lakes worldwide. In the absence of historical water chemistry data, the extent to which this disturbance has impacted lakes in the Erie/Ontario drift and lake plain and Western Allegheny Plateau ecoregions remains to be determined. The objective of this study was to develop a diatom calibration set through analysis of surface sediments and water chemistry from 30 lakes spanning a phosphorous and nitrogen gradient in the glaciated regions of northeast Ohio and northwest Pennsylvania with an additional lake in New York. No current training set exists for this unique geographic region. The relationship between diatom species and environmental variables was established using ordination techniques involving multiple regression and weighted-averaging methods. Canonical correspondence analysis (CCA) was used to determine environmental variables that have a strong influence on diatoms from the ecoregions studied. Total phosphorus, ammonia, and magnesium were the three most statistically significant variables determined through multivariate analyses, although maximum depth and nickel concentrations were also found to be important.

A total phosphorus inference model was developed from recent diatom fossil remains and contemporary water chemistry measurements. The ecological indicator values (optima and tolerances) of 40 abundant diatom species were defined using C<sup>2</sup> computer software. The root mean squared error associated with prediction of the TP

inference model was 17 ug/L, and the  $R^2$  linear coefficient of correlation between observed and diatom-inferred TP values was 0.77. The optima developed in this research match closely those constructed from calibration studies covering similar or longer TP gradients (12 ug/L-153 ug/L TP). A comparison with optima developed from other studies yields values much lower than those in this project and illustrates the need for regional calibration studies. This calibration set will be used to develop a diatom transfer function for use in regional lake management decisions.

## TABLE OF CONTENTS

ABSTRACT.....	v
LIST OF TABLES.....	ix
LIST OF FIGURES.....	x
CHAPTER	
I. INTRODUCTION.....	1
1.1 Diatom ecology.....	2
1.2 Applications – calibration studies.....	4
1.3 Tracking anthropogenic modifications.....	6
1.3.1 Lake acidification.....	7
1.3.2 Nutrient enrichment.....	8
II. CHARACTERISTICS OF SAMPLE LAKES, DATA COLLECTION, AND STATISTICAL ANALYSES.....	16
2.1 Geomorphology of sample lakes.....	19
2.2 Sample collection, preparation, and analyses.....	21
2.3 Data exploration and statistical analyses.....	22
III. RESULTS, DISCUSSION, AND CONCLUSIONS.....	26
3.1 Lake Chemistry.....	26
3.2 Diatom assemblages – community parameters.....	30
3.2.1 Planktonic taxa.....	34
3.2.2 Periphytic taxa.....	37



3.3	Results of data exploration .....	38
3.3.1	Detrended correspondence analysis.....	38
3.3.2	Canonical correspondence analysis .....	39
3.4	Ordination .....	50
3.5	Inference model development.....	54
3.5.1	Total phosphorus inference model.....	54
3.5.2	Ammonia inference model.....	58
3.5.3	Exploratory models (Mg, Ni, $Z_{\max}$ , Na).....	59
3.6	Comparison of TP optima values .....	61
3.7	Factors influencing TP inference models .....	66
3.8	Conclusion .....	68
LITERATURE CITED .....		70
APPENDICES .....		89
1.	Water column data for 30 sample lakes.....	90
2.	Description of automated analyses for metal detection and detailed metal analyses methods with literature cited .....	98
3.	Species list .....	108
4.	Species response curves.....	116

## LIST OF TABLES

Table		page
I.	Elevation, coordinates, county and state for sample lakes .....	18
II.	Environmental variables and water chemistry for sample lakes.....	27
III.	Trophic state index for sample lakes .....	28
IV.	Water column metal concentrations for sample lakes .....	29
V.	Dominant and subdominant species of sample lakes.....	31
VI.	Shannon diversity, evenness, and richness for sample lakes.....	32
VII.	Significance values for Monte Carlo permutations, eigenvalues for constrained CCA.....	41
VIII.	Pearson correlation matrix of 10 significant environmental variables.....	42
IX.	Numbers and lake codes used in CCA analysis and biplot diagrams .....	43
X.	Regression/canonical coefficients for 3 significant forward selected variables .....	49
XI.	T-values of regression coefficients .....	49
XII.	Results of CCA with 40 species and TP, NH <sub>3</sub> , and Mg.....	51
XIII.	TP optima and tolerances, ug/L for 40 abundant diatom species .....	56
XIV.	List of TP calibration studies, range of TP values, and regions examined .....	62

## LIST OF FIGURES

Figure		page
1.	Location of 30 sample lakes .....	17
2.	Ordination biplot of lakes sites and 3 significant variables .....	44
3.	Ordination biplot of species and 3 significant variables .....	45
4.	Ordination biplot of lakes sites 4 exploratory variables .....	46
5.	Ordination of species and 4 exploratory variables.....	47
6a.	Observed vs. WA residuals values for 28 lakes.....	57
6b.	Observed vs. WA diatom-inferred TP values for 28 lakes .....	57

## **CHAPTER I**

### **INTRODUCTION**

Throughout much of the world there has been an increase in environmental awareness. The shift has been realized not only in professional attitude towards the condition of natural resources, but in public perception of environmental quality. The federal government has responded to environmental degradation of air, land, and water with legislative acts. The Clean Water Act of 1972 and its amendments aim to reduce anthropogenic deterioration of surface water bodies. With legislation in place to protect further degradation, it is of interest to understand natural or pre-settlement conditions of aquatic systems. Estimates of pre-industrial nutrient conditions within a lake can provide a target for restoration efforts (Bennion et al., 1996). Without baseline information, it is impossible to determine the extent of impairment or improvement of our water resources. Unfortunately, there exists little if any historical water chemistry records for most aquatic systems in northeastern Ohio.

One approach to estimating past conditions in a lake is to study fossil remains of organisms preserved in lake sediments (Battarbee et al., 2001). The relationship between an abundant organism and the chemical conditions supporting it can provide useful qualitative information. Once contemporary species are calibrated to recent water chemistry measurements, a number of water quality parameters can be quantitatively reconstructed using remains of indicator organisms.

Organisms commonly used as indicators include copepods, chironomids, chrysophytes, and diatoms (Davis et al., 2006; Findlay et al., 1998; Garcia-Rodriguez et al., 2007; Zhang et al., 2007). Scientists since the 1900s have utilized diatoms as indicators of environmental conditions in aquatic systems when water chemistry data was absent (Kolkwitz and Marsson, 1908). Diatom research in the past has provided a wealth of ecological information (Cleve, 1894-95; Cholnoky, 1968) allowing qualitative generalizations about aquatic conditions (Hustedt, 1957; Smol, 1988), while more recent studies offer sophisticated quantitative methods of lake reconstruction (Pienitz et al., 1995; Laird and Cumming, 2008).

## **1.1 Diatom Ecology**

Diatoms are microscopic single-celled photosynthetic aquatic organisms contained within a distinctive silica frustule that can remain preserved in sediments for thousands of years. Diatoms are found nearly anywhere on earth that offers light, water, and nutrients. These photosynthetic organisms occupy a wide variety of habitats including wetlands, streams, waterfalls, seepwalls, pools, lakes, and oceans (Fritsch, 1931). Diatoms are the supporting foundation of most aquatic food webs. These primary

producers convert sunlight energy into useable calories for higher trophic organisms. Diatoms are preyed upon by zooplankton, aquatic insects, snails, and planktivorous fish (Brönmark and Hansson, 2005). Parameters affecting diatom assemblages in an ecosystem include predation and competition for space, nutrients and light with other algae and aquatic plants.

Diatoms are sensitive to changes in the chemistry of the water in which they exist. Early researchers realized this principle with respects to pH (Fritsch, 1931; Hustedt, 1937-39). Although some species were found to have wide ranges of pH tolerance, scientists noted that optimal development only occurs within a narrow range of hydrogen-ion concentrations. In more recent years phosphorus and nitrogen have been identified as components that typically limit the growth of diatoms under natural conditions. Experimental lake enrichment studies show that phosphorous loading amplifies diatom biomass, and nitrogen to phosphorous ratios strongly influences diatom assemblages (Schindler and Fee, 1974).

Diatoms can be used as indicators of environmental variables (Moser et al., 1996). Remnants of their unique cell wall provide a fossil record of diatom assemblages. Other types of algae such as green (Chlorophyceae) and blue-green (Cyanophyceae) lack the protective silica cell wall and do not preserve well in sediments. Changes in the diatom community can be related to corresponding fluctuations in water chemistry. Variables known to affect diatom distribution include salinity and trace metal levels, pH, total nitrogen and total phosphorus (Dixit et al., 1992; Reavie and Smol, 2001).

Diatoms have been used in numerous projects across the world to successfully reconstruct water chemistry (Björck et al., 1991; Christie and Smol, 1993; Dixit et al.,

1998; Reavie et al., 2002; Ramstack et al., 2003; Cubizolle et al., 2005; Haberyan and Horn, 2005; Wasylikowa et al., 2006; Vasari et al., 2007). Phosphorus reconstructions utilizing diatoms as environmental indicators have provided valuable information to lake managers and environmental agencies. Changes in other environmental factors including salinity, conductivity, total nitrogen, pH, climate, and temperature have been reconstructed using diatom fossils.

## **1.2 Applications of diatom research – calibration studies**

Early diatom ecology work identified ecological characteristics of many species from habitats worldwide (Cleve, 1894-95; Hustedt, 1956; Cholnoky, 1968). This information allowed scientists to make generalizations about habitat characteristics, but did not offer numerical approaches to water quality parameters. Qualitative assessment of aquatic ecosystems remains an important application of diatoms in paleoecological work (Battarbee et al., 2001), although a desire to report estimates for environmental variables inspired developments in quantitative diatom reconstructions. Quantitative reconstructions involve averaging indicator values for diatom species determined from a calibration data set to report a numerical estimate of historical chemical or physical parameters from a lake for which no data exists. Diatoms with narrow ecological tolerances can be used to accurately infer water chemistry, and extrapolation of historical water quality from analysis of deep sediment diatom fossils is possible.

A calibration study is the primary step in quantitative lake reconstruction. Calibration data sets are constructed from analyses of recent diatoms in shallow surface sediments and contemporary water chemistry from a number of lakes in a given region

(Bennion et al., 1996). The diatoms are then used to infer unknown water chemistry based on the ecological principle that each species has a set of defined optimum nutrient concentrations and does not exist in abundance outside of this threshold. Where a species is present, it can be assumed that the water chemistry is close to the indicator value or optimum for that variable constructed from the calibration data base. The objective is to develop optima and tolerance values for environmental variables of interest for each abundant diatom species. Although many diatoms may occur at a site, diatoms that are present in a number of sites should prove the most robust indicators of water chemistry. If for example, a unique species occurs in only two sites, the indicator values based on two water chemistry measurements are inherently unreliable.

Abundant taxa assumed to represent current flora are identified by enumerating surface sediment remains on microscope slides. Near-shore habitats are exposed to different levels of nutrients and diatom species in shallow waters respond to environmental variables in a manner distinct from planktonic taxa (Bradshaw and Anderson, 2001). Thus, surface-sediments are sampled from the deepest site in a lake. Obtaining samples from the deepest location within the basin provides a combination of near-shore periphytic, open-water planktonic, and deep-benthic diatom species and correspond to annual diatom community fluctuations more accurately than near-shore sampling. Relating these dominant planktonic forms to contemporary nutrient levels provides an understanding of the influence environmental variables have on diatom community composition within the lake.

Lake sites are chosen based on one or more nutrient gradient of interest. A more complete understanding of diatom response to environmental variables is achieved by



sampling sites covering a wide range of water chemistry. A long environmental gradient increases the probability of accurately determining diatom nutrient optima, producing indicator values more realistic than those developed from lakes with a narrow range of water chemistries. The environmental variables of primary interest in this study include nutrients that naturally limit growth but when present in large amounts stimulate blooms of diatoms.

Multivariate statistics are used to develop diatom proxies for environmental variables within the lake data set. The relationship between surface sediment assemblages and contemporary water chemistry from lakes in a specific geographic region is determined by canonical correspondence analysis (ter Braak, 1986). Parameters that explain a statistically significant percentage of variation in diatom taxonomy may be considered candidates for reconstruction and inference model development. Variables that most strongly affect diatom distribution across sample sites can be reconstructed by averaging the nutrient optima weighted by the relative abundance of each species present in a sample.

### **1.3 Tracking anthropogenic modifications**

Aquatic resource management decisions based on diatom inferred reference conditions are made possible through the use of a calibration data set. Diatoms as bioindicators provide a more detailed and comprehensive extrapolation of water chemistry than one-shot water samples, which can fluctuate dramatically spatially, daily, seasonally, and yearly (Reavie et al., 2006).

Research utilizing diatom fossils from lake sediments to quantitatively assess contemporary and historical water chemistry has benefited from nearly 30 years of dedicated investigations. Many significant findings concerning diatom ecology have emerged from research in Europe and North America. Early successful reconstruction work involved tracking lake acidification by analyzing sediments for diatom remains (Merilainen, 1967; Charles et al., 1986). In the pages that follow I review projects involving pioneering diatom research and address current applications of diatoms as environmental indicators in lake systems. Applications to lake management and restoration efforts are discussed.

### **1.3.1 Lake Acidification**

Changes in the pH of a water body can have dramatic effects on aquatic organisms. Research qualifying the response of diatoms to pH began in Finnish lakes (Nygard, 1956). Quantitative studies reconstructing pH followed in response to questions regarding the pace and onset of lake acidification throughout Finland (Merilainen and Huttunen, 1983; Merilainen and Huttunen, 1990). The Paleolimnological Investigation of Recent Lake Acidification (PIRLA project) was a quantitative project in North America involving the use of diatoms as indicators of acidification trends (Charles and Whitehead, 1986; Charles et al., 1990). Extensive research began in 1980 to examine lake acidification over the past 100 years. Lakes within four regions susceptible to atmospheric acid deposition in the eastern half of North America were sampled. Sediment analysis revealed acidification effects on the composition of diatom communities within lakes in the Adirondack region of New York and lakes in Northern

New England (Charles et al., 1986). The United States Environmental Protection Agency (EPA) realized the success of the PIRLA project and funded PIRLA II, which probed deeper into the natural state of lakes, acidification in Adirondack lakes and response to air quality improvements as made evident in the sediment record (Charles and Smol, 1990). This study expanded the number of lakes sampled and aimed to infer pre-settlement acid neutralizing capacity (ANC), pH, dissolved organic carbon and aluminum levels in lakes from diatom records. Information from the calibration set was then tested on 675 Adirondack lakes to infer pH change over time, and established baseline conditions of lakes, allowing assessment of chemical changes corresponding to industrial pollution. Results indicated that between 25 and 35% of these lakes have acidified since pre-industrial times (Cumming et al., 1992). This evidence of lake acidification empowered resource managers to make intelligent restoration efforts. These studies represent some of the early successful research involving pH reconstructions from changes in diatom assemblages over time.

### **1.3.2 Nutrient Enrichment**

Increased inputs of nitrogen and/or phosphorus can lead to lake eutrophication, resulting in poor water quality, decreased fish support, algal blooms and overall degradation of a lake. It was understood since the beginning of the 20<sup>th</sup> century that oligotrophic lakes low in nitrogen and phosphorous support different planktonic communities than nutrient-rich eutrophic lakes (Naumann, 1919). Early in the 1900s Swiss scientists correctly attributed observable increases in the productivity of a lake to the influx of untreated sewage (Fehlmann and Minder, 1920).

In paleolimnological reconstructions of nutrient status, phosphorus is of primary concern. Transfer functions created from calibration sets have been used to estimate total phosphorus (TP) and have proved reliable (Bennion et al., 1996). A number of studies have been conducted in Europe on systems comparable to those studied in this research. Europe has a long legacy of industrialization. Many lakes in Europe have experienced prolonged agricultural and municipal nutrient inputs. Lake shorelines in Europe were settled long before areas in North America. European lakes vary in water quality and range from oligotrophic to hypereutrophic, with many eutrophic lakes. Concern over the rate and extent of aquatic pollution has spurred a number of studies utilizing diatom fossil remains to reconstruct environmental conditions. It is relevant to discuss the results of a few European calibration studies.

King et al. (2000) investigated the response of epilithic diatoms (attached to rocks) to 17 environmental variables from 17 lakes in the English Lake District. Statistical analyses determined that diatom species composition was most directly influenced by total phosphorus and calcium concentrations. Total phosphorus (TP) ranged from 0.8 ug/L to 49.2 ug/L (King et al., 2000). Lakes in King's study have relatively low phosphorus concentrations and as a result, the optimum for some species may be underestimated.

A study of Danish lakes (Bradshaw et al., 2002) investigated diatom distribution over a larger phosphorus gradient. Their TP gradient ranged from 24 to 1145 ug/L in 28 lakes, representing primarily mesotrophic to hypereutrophic systems. Mean TP and mean depth were the environmental variables influencing most of the diatom distribution. Lake chemistry was sampled frequently (19 times/year) in comparison to other calibration sets,

yielding chemical measurements that are strongly representative of lake conditions. This increases the ecological significance and applicability of this data set.

Diatoms have been widely used in both contemporary and paleoecological investigations of water quality in Europe (Bennion, 1994; Bigler and Hall, 2002; Bradshaw et al., 2001; Heinsalu et al., 2007). Some studies are comparable to sampling sites in North America based on lake size, trophic classification and extent of anthropogenic impact. A review of diatom calibration studies of nutrient status from North American lakes follows.

An early diatom-inferred TP reconstruction was conducted by Fritz et al., in 1993 on 42 lakes in Michigan. Lakes in this training set span a TP gradient from 1 ug/L to 51 ug/L, representing oligotrophic and mesotrophic systems. Fritz et al. (1993) found changes in diatom communities corresponding to years with construction and logging activities in the associated watersheds. Oligotrophic taxa dominated early sediment records with an increase in mesotrophic species in the 1960s and a return to oligotrophic species in the most recent sediments. Multivariate analysis of diatom and environmental variables showed that TP, maximum depth, and surface area were the factors that most directly affected community composition. Fossil diatoms were then used to infer changes in TP concentrations in the lakes. Three of the four lakes studied exhibited TP increase concomitant with logging and settlement activities.

Ramstack et al. (2003) conducted a study in Minnesota to evaluate water quality changes since enactment of the Clean Water Act in 1972. This training set is noteworthy for the range of sites included; varying from near-pristine woodlands to heavily impacted metropolitan and agricultural lands. Fifty-five lakes spanning three ecoregions were

sampled to construct diatom inference training sets. Significant environmental variables determining diatom community composition were TP, pH, acid neutralizing capacity, chloride and color. Researchers concluded that chloride concentrations increased in many urban lakes corresponding to road-deicer application. Total phosphorus reconstructions indicated a steady rise in nutrient enrichment of lakes in agricultural and urban catchments (Ramstack et al., 2003). Other training sets utilize pre-settlement baselines but this study shows diatoms can also be used to assess other temporal aspects of change in contemporary aquatic systems.

Reconstructions based on the Minnesota transfer function indicate many lakes in the corn-belt and metropolitan areas have changed significantly since the 1970s, whereas lakes in the forested regions have not. It was expected that regulatory actions would result in TP decreases after the 1970s due to reductions in the use of phosphate detergents. However, while many of the lakes decreased in TP, some are so impacted that the decline was not significant enough to change the trophic state of the lake. This implies that land-use practices and urban/agricultural water resource management have not adequately addressed nutrient enrichment and ionic disturbance of lakes in the metropolitan and rural areas of Minnesota.

Another significant calibration study involves the inference of long-term nutrient changes in lakes of southeastern Ontario, Canada (Reavie et al., 2002). This research involved sampling 50 lakes to construct a calibration set based on diatom remains from top (recent) and bottom (historic) sediments. The calibration set was then used to reconstruct TP changes from pre-industrial to present time. Another method, the lake-shore capacity model (LCM) involved mass-balance reconstruction of phosphorus

loading associated with land-use and watershed changes and the two methods were compared to assess their performance in nutrient reconstruction.

The diatom analyses showed a shift in community assemblage from mesotrophic to eutrophic taxa indicative of nutrient inputs. Diatom inferred transfer functions showed 78% of the lakes increased in TP, while the LCM model showed only 56% of lakes increased. Statistical examination suggests that the LCM tends to underestimate elevated phosphorous levels based on outliers in the data set with extremely high TP levels. The average increase was in agreement for both methods however, and lake management efforts should involve multiple reconstructions when available to ensure accuracy. Based on both types of reconstructions, the conclusion is that lakes in southeastern Ontario have become more eutrophic corresponding with lake shore development, logging, agriculture and municipal activities. With this knowledge lake managers can approach future restoration efforts more logically.

The Great Lakes Environmental Indicators (GLEI) project was a large scale multi-taxonomic investigation of water quality in coastal areas of the Great Lakes (Reavie et al., 2006). This study represented the first large multi-disciplinary investigation of coastal zones of the Great Lakes and associated wetlands using diatoms, birds, fish, reptiles, amphibians and near-shore vegetation as biological indicators of water quality (Brazner et al., 2007). Diatom - based transfer functions for TP, TN, chloride and chlorophyll were constructed from the data (Sgro et al., 2007). Diatom - inferred environmental variables corresponded more closely to watershed land-use and development values than the snapshot water chemistry measurements taken at the same time (Reavie et al., 2008). These promising results suggest coastal areas of large and

important freshwater resources can be quickly and inexpensively assessed for nutrient status using biological organisms (Kireta et al., 2008).

Diatom-based calibration sets have been established for lakes in many locations around North America (Schmidt et al., 2004). Regional calibration sets have been developed for Minnesota, Michigan, the Adirondack region of New York and Southeastern Ontario, Canada (Ramstack et al., 2003; Fritz et al., 1993; Charles et al., 1990; Dixit et al., 1998; Christie and Smol, 1993; Reavie and Smol, 2001, respectively).

Anthropogenic inputs of phosphorus and nitrogen in the form of agricultural, municipal, and industrial pollution have impacted lakes in the Erie/Ontario-drift lake-plain ecoregion. Excess agricultural and residential fertilizer applications, municipal sewage disposal and water treatment have led to degradation of surface water bodies. Legislation has been enacted which aims to reverse the nutrient enrichment of freshwater lakes. In the absence of historical water chemistry, the extent to which pollution and restoration events have affected lakes remains to be determined.

A calibration set that is unique to the Erie-Ontario-drift lake-plain ecoregion could be used to assess water quality conditions in northeastern Ohio lakes. The calibration set needs to be specific to the region in which it will be utilized for a number of reasons. The diatoms in a calibration set from one region may be found in conditions that are not comparable to those in this region. Other calibration sets have been assembled from lakes with much lower total phosphorus values, and may underestimate historical phosphorus levels in Ohio lakes.

No diatom-based calibration sets have been developed for the Erie/Ontario drift and lake plain and Western Allegheny Plateau ecoregions. These ecoregions are



effectively distinct from the aforementioned sample sites with respect to the underlying geology, rate of polluting impacts, nutrient content, and management strategies. The underlying bedrock is composed primarily of shales and conglomerates with overlying clay and sandy glacial till (Woods et al., 2006). In addition, these lakes have experienced varying stresses from industry, agriculture, and residential development. The distinctiveness of lakes in NE Ohio and NW Pennsylvania demands the development of a relevant training set.

Data from other regions cannot be used to accurately reconstruct nutrient conditions of lakes in this particular area. Diatoms may have slightly different tolerances depending on the geographic distribution of the population and the range of water chemistries sampled. The training set developed in Minnesota is not applicable to this area for a number of reasons. That calibration set spans three different ecoregions that are geologically distinct from Ohio. The lakes in the Minnesota set exist within Canadian Shield bedrock and calcareous glacial drift (Ramstack et al., 2003). As a result, some of the lakes are naturally highly alkaline. A training set developed in NW Michigan is constructed from lakes that are alkaline, much larger in surface area, deeper, and generally less nutrient rich (Fritz et al., 1993.) The training set developed for the Adirondack region of New York (Charles et al., 1986; Charles and Smol, 1990) is based on lakes that are physically distinct from those in NE Ohio. Lakes in the Adirondacks overlay granitic gneiss and metasedimentary rocks, are low in pH, and have experienced increased sulphate and nitrate deposition (Cumming et al., 1992.). Although a SE Ontario training set may seem applicable to this region because of relative proximity, the

Canadian lakes exist on limestone, granite, or a combination of both bedrock types and are of a slightly different age (Reavie and Smol, 2001).

The objective of this thesis research is to develop a diatom calibration set and transfer function for lakes in the Erie/Ontario drift and lake plain and Western Allegheny Plateau ecoregions. This calibration set will be useful in lake management, restoration and limnological research to infer historic or baseline water conditions.

## **CHAPTER II**

### **CHARACTERISTICS OF SAMPLE LAKES, DATA COLLECTION, AND STATISTICAL ANALYSES**

#### **Introduction**

It is often desirable in ecological research to obtain reference conditions of populations in their most natural or unaffected state in order to draw conclusions regarding how a system operates. The first step in selecting reference lakes for this calibration study was to identify natural lakes. Sampling naturally formed lakes eliminates the potential of confounding variables related to reservoir systems that may affect diatom species response.

Naturally formed public access lakes were identified from an Ohio Environmental Protection Agency (OEPA) 1996 Ohio Water Resource Inventory (Davic et al., 1996). The report identifies 24 naturally formed glacial lakes out of 446 public lakes greater than 5 acres in surface area. Twenty lakes were selected that occurred in glaciated counties of NE Ohio. An additional nine lakes from NW Pennsylvania and Findley Lake in the SW corner of New York were included to broaden the nutrient gradient based on published and unpublished data (Ostrofsky and Bradley, 2006; Lathrop, personal communication, 2006; Wolin, personal communication, 2006) (Figure 1, Table I).

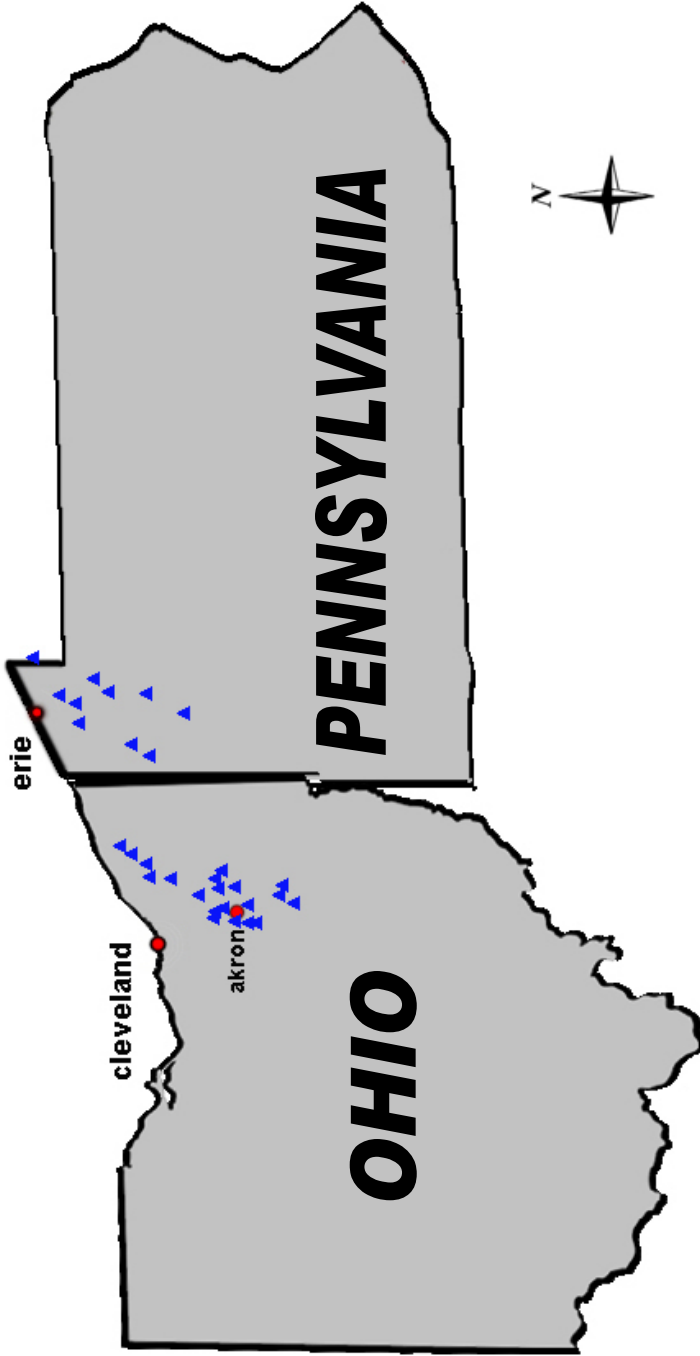


Figure 1. Location of 30 sample lakes (blue triangles), NE OH and NW PA

Lake code	Lake name	Number	Latitude	Longitude	Elevation (ft)	Surface area	County	State
NYFIN	FINDLEY	9	42.106	-79.725	1420	290	Chautauqua	NY
OHAQU	AQUILLA	1	41.548	-81.167	1134	27	Geauga	OH
OHBAS	BASS	2	41.548	-81.222	1141	169	Geauga	OH
OHBRA	BRADY	3	41.165	-81.316	1055	70	Portage	OH
OHCRY	CRYSTAL	-	41.163	-81.465	996	25	Summit	OH
OHEAS	EAST	4	41.195	-81.334	1044	67	Portage	OH
OHMEY	MEYERS	5	40.817	-81.417	1056	134	Stark	OH
OHMUZ	MUZZY	6	41.118	-81.252	993	82	Portage	OH
OHNES	NESMITH	7	41.026	-81.552	965	80	Summit	OH
OHNOR	NORTH RESERVOIR	8	41.001	-81.545	894	160	Summit	OH
OHPIP	PIPPEN	19	41.178	-81.311	1054	143	Portage	OH
OHPUN	PUNDERSON	-	41.454	-81.206	1149	100	Geauga	OH
OHSAN	SANDY	20	41.120	-81.297	1082	90	Portage	OH
OHSIL	SILVER	21	41.156	-81.460	994	91	Summit	OH
OHSIP	SIPPO	22	40.804	-81.455	1027	88	Stark	OH
OHSFP	SPRINGFIELD	23	41.028	-81.439	1078	200	Summit	OH
OHSPV	SPRING VALLEY	24	40.817	-81.444	990	58	Stark	OH
OHSUM	SUMMIT	25	41.052	-81.545	965	100	Summit	OH
OHSUN	SUNNY	26	41.297	-81.318	1103	63	Portage	OH
OHWES	WEST	27	41.198	-81.341	1045	91	Portage	OH
OHWYO	WYOGA	28	41.187	-81.487	978	60	Summit	OH
PACAN	CANADOHTA	10	41.812	-79.839	1389	170	Crawford	PA
PACON	CONNEAUT	11	41.622	-80.305	1073	929	Crawford	PA
PACRY	CRYSTAL (MUD)	12	41.554	-80.368	900	60	Crawford	PA
PAEDI	EDINBORO	13	41.885	-80.135	1197	252	Erie	PA
PALEB	LEBOEUF	14	41.933	-79.987	1166	70	Erie	PA
PAPLE	PLEASANT	15	42.005	-79.898	1302	60	Erie	PA
PASAN	SANDY	16	41.345	-80.110	1157	150	Mercer	PA
PASPR	SPRING	17	41.764	-79.908	1167	56	Crawford	PA
PASUG	SUGAR	18	41.566	-79.945	1288	90	Crawford	PA

**Table I. Sample lakes, assigned number, coordinates, elevation in feet, surface area in acres, county and state**

## 2.1 Geomorphology of sample lakes

Sample lakes exist within the Erie/Ontario drift and lake plains ecoregion and the glaciated Allegheny plateau, spanning NE OH and NW PA between 40° to 43° latitude and 79° to 82° longitude (Table I). Lakes in this region were formed during retreat of the Wisconsin glacier, between 10,000 and 16,000 years ago. This glaciation left behind large blocks of ice to form the many kettle lakes used in this research (Berg and Speck, 2006; Schiner, 1979). Parent material in these ecoregions includes primarily shales, sandstones, and conglomerates with overlying limestone, gravel, silt, and sandy lacustrine drift deposition. The majority of lakes are depressions or basins within flat, poorly-drained clay soils or end moraine. A few lakes in Summit County, OH exist on sandy glacial outwash or unsorted glacial till. The topography is generally low, with elevations between 800 and 1400 feet above sea level. Surface area of lakes sampled ranges from 25 to 300 acres with one lake at 929 acres. A number of lakes are part of drainage systems connected by rivers, as is the case with Pennsylvanian lakes in the French Creek watershed network.

Lake sites were chosen to span a phosphorus gradient. Historical TSI's were used where available to approximate relative phosphorus concentrations and ensure inclusion of sites representing low, intermediate, and high phosphorus levels. The Trophic State Index (TSI) (Carlson, 1977) is used to classify lakes based on light penetration, chlorophyll a, and total phosphorus measurements. Trophic state indices were also calculated for all lakes from data collected in this study using equations provided in Carlson, (1977) and Davic et al., (1996).

Nearly all 20 lakes in Ohio exist in mixed urban-industrial, residential and light-agricultural watersheds while 9 lakes in Pennsylvania are within mixed urban-residential and light-agricultural watersheds. Findley Lake in New York is in an agricultural and low-density residential watershed. Most lakes have some wetlands associated with the immediate lakeshore area, although a number of systems have had these drained or filled to promote dock and cottage development. Reducing the wetland buffer around a lake can have devastating effects on nutrient loads that would otherwise be sequestered or filtered through the wetland border. Many lakes in Ohio have dense shoreline residential communities that enjoy active recreation on the lakes. Systems in Pennsylvania generally have less intense housing development along the immediate lake shore, with few exceptions. Findley Lake has vacation cottages and permanent residences surrounding the basin. A few communities surrounding sample lakes at various points in time have used septic systems to process domestic sewage wastes, which can affect nutrient levels in nearby waters bodies. An effort has been made to upgrade residences on some of these lakes to sewer systems, with varying degrees of success.

Hunting and fishing is encouraged on and around the majority of lakes sampled, with most lakes in Pennsylvania offering generous boat launch ramps sponsored by the active PA Fish and Boat commission. Many systems have been stocked with commercially and recreationally attractive sport fish. These various watershed uses and management attitudes contribute to productivity differences in the sample sites.

## 2.2 Sample collection, preparation, and analysis

Surface sediment and water chemistry samples were collected between October 9<sup>th</sup> and November 19<sup>th</sup> 2006 (Appendix 1). Bathymetric maps if available were used to estimate the deepest point in each lake. Sampling the deepest point addresses the issue of spatial heterogeneity by obtaining a centralized, gravity-induced composite of diatoms present in the system. Water column profiles for pH, temperature, and dissolved oxygen were recorded using YSI model 63 pH and temperature system and YSI 550A dissolved oxygen instrument (YSI Environmental Incorporated, 1998; YSI Environmental Incorporated 2002) at 1 meter intervals, beginning at the surface and ending 1 meter above the sediment (Appendix 1). Triplicate surface-sediment samples were obtained using an (23cm<sup>3</sup>) Eckman dredge (Wetzel and Likens, 1991) and combined to yield a representative lake bottom sample. The top 1 cm of the sediment was subsampled and extruded for recent diatom remains using clear plastic pvc tubing fitted with a rubber stopper. Sediment samples were collected into Ziploc® quart size freezer bags and stored in a cooler on ice. Chlorophyll a was measured in the field within 12 hours of collection using a portable Turner design Aquafluorometer (Axler and Henneck, 2002).

Water samples were acquired using a horizontal non-metallic Van Dorn sampler (Van Dorn, 1956). Water grabs were obtained from 1 meter below the surface and 1 meter above the sediment and combined into 500ml plastic Nalgene® bottles and placed on ice to limit fluctuation of nutrient values. Water samples were shipped on ice to Upstate Freshwater Institute (UFI) in Syracuse, NY and analyzed for total phosphorus (SM18 4500PE), nitrite + nitrate and ammonia (EPA 350.1, EPA 353.2, EPA 353.2). Additional 250 mL samples were collected for silica and shipped at room temperature to



avoid distortion of silica levels (SM 18-19 4500-SI D). Total dissolved nitrogen was measured using Pyro-Chemiluminescence on an Anteck 9000 Series Nitrogen Analyzer at UFI. Analyses for lead and cadmium (SM 3113), aluminum, copper, iron, nickel, and zinc (SM 3120B), sodium, calcium, potassium, and magnesium (SM 3111 A-B) were performed at John Carroll University, University Heights, OH using atomic absorption, inductively-coupled plasma and flame atomic emission spectroscopy (see Appendix 2 for description of automated analyses and detailed metal chemistry methods). Light penetration was measured using a 20cm Secchi disk following OEPA methods (2006).

Diatom sediment samples were dried in a 200° F oven, and 0.1 g dry material was weighed into 250 ml glass beakers, and digested in 30% H<sub>2</sub>O<sub>2</sub> and K<sub>2</sub>Cr<sub>2</sub>O<sub>7</sub>. Samples were then rinsed a total of 6 times and allowed to settle for a day between each rinse. Three dilutions were prepared on square 22 mm glass cover slips and dried on a slide warmer. Cover slips were mounted to 3” x 1” Fisher brand glass microscope slides using Naphrax<sup>®</sup> mounting media. Taxonomic counts were performed at 1000x total magnification on an Olympus BX 50 microscope under brightfield conditions. A minimum of 500 valves were identified from each sample and raw counts were converted into percent abundance. Krammer & Lange-Bertalot (1991; 1997; 2004) was used for taxonomic identification. Taxonomy follows current genera descriptions as detailed in Round et al., 1990.

### 2.3 Data exploration and statistical analyses

Shannon’s index ( $H'$ ) calculated as  $(H' = -\sum_{i=1}^S p_i \ln p_i)$  was used as a measure of diversity. Evenness ( $J$ ) and richness ( $S$ ) were determined for each site, where ( $J$ ) is

$(H'/H_{max})$ ,  $(H_{max})$  is the  $\ln(S)$  and  $(S)$  is the number of species at a site. The relationship between diatom species and environmental variables was explored using multiple regression and ordination techniques in CANOCO version 4.54 (Jongman, et al., 1995; ter Braak, 2005). Ordination was used to reduce the multivariate nature of ecological relationships to a few statistically important variables.

The procedure for determining variables that impact diatom species distribution involves screening data for potential outlier sites, analyses of ordination diagrams and interpretation of statistical values. Outlier sites with extreme or suspicious water chemistry measurements should be excluded from inference model development. Covariables (two or more factors that fluctuate as a result of some underlying relationship) should be removed or reduced to a single variable to eliminate redundancies in ordination. The effect of each variable on diatom distribution is then assessed independently of other variables by performing constrained correspondence analyses. Forward selection is a process of reducing the number of variables to those that statistically explain a unique portion of variance in the distribution of diatoms across sample sites. Indicator values are then developed for variables that most strongly affect diatom distribution.

Canonical correspondence analysis (CCA) was used to determine limnological variables of importance. Fifteen variables were initially assessed for impact on species distribution. Constrained analyses were performed by using each variable as the only explanatory value. This procedure reduced the initial set of 15 explanatory variables to 10 statistically significant predictors. This set was further reduced by Monte Carlo

forward selection to 3 environmental variables that account for the majority of variance in diatom species distribution.

A simple linear regression was conducted to identify potential outlier sites that did not exhibit a clear relationship between total phosphorus and algal production inferred from light penetration. Two outlier sites with extremely high TP values (Crystal Lake<sub>(OH)</sub> and Punderson Lake<sub>(OH)</sub>) were identified as influential using leverage, Cooks distance, and the Studentized residuals in SAS statistical software version 9.1 (SAS Institute Inc., 2004) and removed prior to statistical analysis and construction of the total phosphorus inference model. Pearson product-moment correlation between variables was examined in R to remove redundant explanatory parameters (R statistical software version 2.3.1, R Foundation for Statistical Computing (2006). Detrended correspondence analysis (DCA) was run using CANOCO for Windows version 4.54 (ter Braak, 2005) to determine whether linear or unimodal ordination methods should be used.

The C<sup>2</sup> program was used to develop ecological indicator values for environmental variables found to be important in determining the dispersion of abundant diatom species (Juggins, 2007), for quantitative exploration of the relationship between species and their environment, and to create and examine various transfer functions. A transfer function is an inference model for reconstructing a particular variable from diatom microfossils that have been calibrated to contemporary water chemistry measurements. Response variable data (diatom species) and predictor data (environmental variables) are imported from a data matrix, and a transfer function is generated with desired statistical parameters specified. The options selected in this

particular analysis include weighted averaging regression with leave-one-out cross validation.

Once a transfer function is generated, the  $C^2$  output provides ecological indicator values for each species including optima and tolerances for the variable in question. The optimum and tolerance for a single species is calculated by weighted averaging (WA) (ter Braak in Jongman et al., 1995). An example of developing total phosphorus optima for *Aulacoseira granulata* is given here for clarification. Total phosphorus optimum of *A. granulata* = average of TP measured from all lakes in which *A. granulata* is present weighted (multiplied) by the abundance of *A. granulata* at each site divided by the sum of all abundance measures for *A. granulata*. Tolerance measures ecological amplitude along a gradient or range of values where a species is present, and its calculation involves weighted averaging and the standard deviation of a variable among sites in which the species is found.

**CHAPTER III**  
**RESULTS, DISCUSSION AND CONCLUSIONS**  
**DIATOM ASSEMBLAGES, ORDINATION, AND INFERENCE MODELS**

**3.1 Lake Chemistry**

Water nutrient analyses indicate that a broad nutrient gradient was captured for TP, TN and NH<sub>3</sub> in the 30 study lakes sampled (Table II). Total phosphorus ranged from 12.2 ug/L in Sandy Lake (PA) to 338.9 ug/L Crystal Lake (OH); TN ranged from 597 ug/L in Lake Aquilla (OH) to 3221 ug/L in Canadohta Lake (PA); NH<sub>3</sub> ranged from 19.7 ug/L in Silver Lake (OH) to 1871.2 ug/L in Crystal Lake (OH); and NO<sub>x</sub> values ranged from 8.1 ug/L in Silver Lake (OH) to 2335.9 ug/L in Canadohta Lake (PA). Calculated Trophic State indices showed general agreement with historic reported TSI for most of the sample lakes (Table III). Cation concentrations for Ca and Mg were consistent with lower alkalinity due to regional geological conditions (Table IV). Aluminum and Ni were the only detectable metals with Cd, Cu, Fe, Pb and Zn concentrations below detection (Table IV). Bass Lake (OH) and Sunny Lake (OH) had high levels of Al (1349 ug/L and 2425 ug/L, respectively) and Ni (420 ug/L and 422 ug/L, respectively), while most lakes had low concentrations or were below detection.

Code	TP (ug/L)	TN (ug/L)	Si (mg/L)	NH <sub>3</sub> (ug/L)	NO <sub>x</sub> (ug/L)	pH	Secchi (cm)	Zmax (m)
NYFIN	36.9	1094	3	103.7	242.5	8	237.5	8
OHAQU	34.7	597	3.6	69.9	57.1	7	142.5	4.5
OHBAS	75.1	814	4.9	176.3	160.4	7.2	32.5	1.9
OHBRA	37.2	763	1	203.8	48	6.8	195	12
OHCRY	338.9	1379	1.1	1871.2	13.9	7.4	240	14.5
OHEAS	38.5	1444	5.2	186.3	44.5	7.3	170	12
OHMEY	22.2	697	2.3	42.8	32.2	7.8	172.5	6.8
OHMUZ	122.2	983	1.1	265.8	35.1	7.3	105	8
OHNES	91.5	2282	5.6	124.6	424.1	8.3	48.5	5.8
OHNOR	88.7	703	6.9	33.9	35.8	7.9	44.5	2.5
OHPIP	24.4	1150	1.3	318.3	211.9	7.6	210	14
OHPUN	198.8	1322	3.6	728.3	68.7	7.7	130	15.5
OHSAN	141.8	1234	5.7	779.5	55.6	7.3	92.5	7.9
OHSIL	85.4	798	0.8	19.7	8.1	8.4	53.5	11
OHSIP	89.3	1515	5.1	113.7	65.9	7.3	45	3.6
OHSPR	38.9	837	1.4	51.3	85.2	8.3	105	7.1
OHSPV	25.5	641	0.9	58.7	27.4	8.1	330	3.3
OHSUM	61.8	780	5.3	170.6	119.9	8.2	86	7.8
OHSUN	153.2	1254	3.7	190.7	206.4	7.4	32.5	1.5
OHWES	59.7	971	1.8	274.1	53.6	7.3	122.5	12
OHWYO	114.1	1152	2.2	244.4	234	7.3	51	7.3
PACAN	32.1	3221	2.9	130.6	2335.9	7.1	176.5	11
PACON	26	915	4.9	148.8	134.3	7.3	195	13
PACRY	28.6	863	2.6	74.8	65.7	7.1	249	8
PAEDI	34.4	1299	2.8	50.1	836.9	7.7	125	8.5
PALEB	27.6	1028	4.1	52.5	429.1	7.4	190	8.5
PAPLE	36.3	716	3.5	44.6	107	7.6	157.5	12
PASAN	12.2	885	5.8	139	141.9	7.5	282.5	12
PASPR	56.8	803	3.8	88.5	154.8	7.4	197.5	4
PASUG	32.4	823	5.2	117.6	259.2	6.8	192.5	5

**Table II. Environmental variables and water chemistry for 30 lakes. Total phosphorus (TP ug/L), total dissolved nitrogen (TN ug/L), ammonia (NH<sub>3</sub> ug/L), and nitrate + nitrite (NO<sub>x</sub> ug/L), silica (Si mg/L), pH (pH) measured as an average of water column readings, Secchi disk depth (Secchi centimeters), maximum depth (Z<sub>max</sub> meters).**

Code	TP TSI	Secchi TSI	Chla TSI	Average TSI	Trophic State	Historic TSI	Year
NYFIN	56	48	37	52	Eutrophic	none	
OHAQU	55	55	40	55	Eutrophic	53 <sup>a</sup>	1989
OHBAS	66	76	39	71	Hypereutrophic	69 <sup>b</sup>	2002
OHBRA	56	50	32	53	Eutrophic	62 <sup>a</sup>	1992
OHCRY	88	47	62	68	Hypereutrophic	none	
OHEAS	57	52	47	55	Eutrophic	52 <sup>e</sup>	1987
OHMEY	49	52	42	50	Eutrophic	none	
OHMUZ	73	59	61	66	Eutrophic	63 <sup>a</sup>	1994
OHNES	69	70	56	70	Hypereutrophic	68 <sup>a</sup>	1993
OHNOR	69	72	55	70	Hypereutrophic	59 <sup>a</sup>	1993
OHPIP	50	49	48	50	Eutrophic	44 <sup>a</sup>	1989
OHPUN	80	56	58	68	Hypereutrophic	56 <sup>a</sup>	1989
OHSAN	76	61	56	68	Hypereutrophic	none	
OHSIL	68	69	37	69	Hypereutrophic	none	
OHSIP	69	72	51	70	Hypereutrophic	60 <sup>a</sup>	1990
OHSPR	57	59	53	58	Eutrophic	56 <sup>a</sup>	1994
OHSPV	51	43	47	47	Mesotrophic	50 <sup>a</sup>	1989
OHSUM	64	62	43	63	Eutrophic	64 <sup>a</sup>	1986
OHSUN	77	76	36	76	Hypereutrophic	66 <sup>a</sup>	1995
OHWES	63	57	49	60	Eutrophic	47 <sup>d</sup>	1979
OHWYO	72	70	43	71	Hypereutrophic	none	
PACAN	54	52	36	53	Eutrophic	52 <sup>f</sup>	1996
PACON	51	50	43	51	Eutrophic	52 <sup>f</sup>	2004
PACRY	53	47	33	50	Eutrophic	49 <sup>c</sup>	2006
PAEDI	55	57	32	56	Eutrophic	57 <sup>f</sup>	2005
PALEB	52	51	38	51	Eutrophic	60 <sup>c</sup>	2006
PAPLE	56	53	39	55	Eutrophic	50 <sup>f</sup>	1991
PASAN	40	45	37	43	Mesotrophic	42 <sup>c</sup>	2006
PASPR	62	50	38	56	Eutrophic	55 <sup>f</sup>	2000
PASUG	54	51	37	52	Eutrophic	58 <sup>c</sup>	2006

**Table III. Trophic State Index calculated from each lake following formulae in Carlson, 1977, & Davic, et al., 1996. TSI TP = 14.42 ln(TP ug/L) + 4.15. TSI Secchi = 60 – 14.41 ln(SD meters). TSI Chla = 9.81 ln(Chla ug/L) + 30.6. Historical TSI obtained from various state agencies and thesis research (<sup>a</sup>Davic et al., 1996; <sup>b</sup>Anderson and Davic, 2002; <sup>c</sup>Ostrofsky and Bradley, 2006; <sup>d</sup>Moffett, 1979; <sup>e</sup>Corkan, 1993; <sup>f</sup>Lathrop, personal communication, 2006).**

Code	Al ug/l	Ca mg/l	K mg/l	Mg mg/l	Na mg/l	Ni ug/l	Cd ug/l	Cu ug/l	Fe ug/l	Pb ug/l	Zn ug/l
NYFIN	81*	13.96	1.32	7.86	7.52	98	<0.15	<12.5	<12.5	<2	<25
OHAQU	169	11.66	2.73	5.92	7.48	140	<0.15	<12.5	<12.5	<2	<25
OHBAS	1349	19.20	3.15	8.82	11.54	420	<0.15	<12.5	<12.5	2	<25
OHBRA	112	10.00	3.23	7.88	20.43	35	<0.15	<12.5	<12.5	<2	<25
OHCRY	47*	36.38	3.14	12.58	16.05	38	<0.15	<12.5	<12.5	<2	<25
OHEAS	44*	22.96	2.02	11.84	18.05	93	<0.15	<12.5	<12.5	<2	<25
OHMEY	75*	41.84	2.13	14.10	16.26	48	<0.15	<12.5	<12.5	<2	<25
OHMUZ	69*	18.72	2.73	8.36	15.11	28	<0.15	<12.5	<12.5	<2	<25
OHNES	66*	44.78	3.54	16.10	18.73	129	<0.15	<12.5	<12.5	<2	<25
OHNOR	112	30.48	2.98	17.60	17.04	179	<0.15	<12.5	<12.5	<2	<25
OHPIP	40*	20.08	1.48	7.66	3.39	26	<0.15	<12.5	<12.5	<2	<25
OHGUN	29*	34.92	2.95	11.30	14.30	91	<0.15	<12.5	<12.5	<2	<25
OHSAN	128	41.56	3.03	12.50	11.66	140	<0.15	<12.5	<12.5	<2	<25
OHSIL	90*	25.48	2.72	9.24	21.41	10	<0.15	<12.5	<12.5	<2	<25
OHSIP	112	37.47	4.09	17.40	18.25	186	<0.15	<12.5	<12.5	<2	<25
OHSPR	90*	28.86	3.31	11.98	19.10	48	<0.15	<12.5	<12.5	<2	<25
OHSPV	75*	31.22	3.70	22.68	17.44	19	<0.15	<12.5	<12.5	<2	<25
OHSUM	71*	44.40	3.08	17.40	20.16	128	<0.15	<12.5	<12.5	<2	<25
OHSUN	2425	25.68	4.81	10.82	9.83	422	<0.15	<12.5	<12.5	<2	<25
OHWES	90*	27.18	2.95	11.38	18.37	70	<0.15	<12.5	<12.5	<2	<25
OHWYO	361	34.82	4.46	12.68	15.58	223	<0.15	<12.5	<12.5	<2	<25
PACAN	58*	5.70	2.44	5.08	4.51	137	<0.15	<12.5	<12.5	<2	<25
PACON	212	10.30	2.39	7.52	4.90	90	<0.15	<12.5	<12.5	<2	<25
PACRY	69*	13.28	2.06	6.24	8.64	191	<0.15	<12.5	<12.5	<2	<25
PAEDI	88*	11.24	2.04	6.50	7.68	96	<0.15	<12.5	<12.5	<2	<25
PALEB	103	10.32	2.27	6.40	5.27	238	<0.15	<12.5	<12.5	<2	106
PAPLE	110	25.94	1.64	9.86	6.72	119	<0.15	<12.5	<12.5	<2	<25
PASAN	79*	25.24	2.62	12.50	7.93	139	<0.15	<12.5	<12.5	<2	<25
PASPR	64*	19.04	1.62	11.98	3.48	112	<0.15	<12.5	<12.5	<2	<25
PASUG	173	2.94	1.78	4.26	4.81	133	<0.15	<12.5	<12.5	<2	<25

**Table IV. Water column metal concentrations of sample lakes. Values with an asterisk are below detection but are close to the limit (Appendix 2). Cd, Cu, Fe, Pb, & Zn values lower than detection limit unless indicated otherwise with numerical value.**



## 3.2 Diatom Assemblages

### Community parameters

A total of 15,246 valves were recorded from microscope slide counts representing 321 different taxa in 57 Genera (see Appendix 3 for species list). Planktonic diatoms were most abundant in nutrient enriched lakes with high productivity (Kuhn et al., 1981). In these lakes the water column is rich with nutrients, and light is absorbed by plankton and particulates suspended in the open water, limiting the amount of light reaching benthic organisms. Periphytic forms dominated a few systems, indicating an important distinction between lakes that support primarily plankton and those that favor attached forms (Table V). Shannon index ( $H'$ ) of diversity ranged from 1.25 to 3.56 with a mean diversity of 2.61. Evenness ( $J$ ) ranged from 0.43 to 0.82. The species richness ( $S$ ) from each lake ranged from 18 in hypereutrophic Crystal Lake (<sub>OH</sub>) to 77 in mesotrophic Lake Leboeuf (<sub>PA</sub>) with a mean of 41 species (Table VI).

Productive, impacted lakes typically show lower species diversity, evenness and richness than less enriched lakes (Williams, 1964). This may be a result of the population dynamics described by resource-competition theory (Tilman, 1977; Tilman, et al., 1982), which holds that the number of species present at a site is directly related to the quantity of resources at growth-limiting concentrations. Increased productivity of a lake can also be negatively related to phytoplankton diversity (Interland and Kilham, 2001).

If resources such as phosphorus, nitrogen and silica are supplied at rates much higher than cell demands, as is the case in hypereutrophic systems, than resource-

<b>Code</b>	<b>Dominant</b>	<b>%</b>	<b>Subdominant</b>	<b>%</b>
OHWYO	<i>Aulacoseira granulata</i>	72	<i>Cyclostephanos dubius</i>	8
OHSUN	<i>Aulacoseira granulata</i>	40	<i>Cyclostephanos dubius</i>	31
OHSIP	<i>Aulacoseira granulata</i>	36	<i>Cyclostephanos dubius</i>	29
OHAQU	<i>Aulacoseira granulata</i>	28	<i>Cocconeis placentula v. lineata</i>	11
OHBAS	<i>Aulacoseira granulata</i>	25	<i>Cocconeis placentula</i>	15
PACRY	<i>Aulacoseira ambigua</i>	41	<i>Asterionella formosa</i>	22
PASAN	<i>Aulacoseira ambigua</i>	24	<i>Cyclotella comensis</i>	11
PASUG	<i>Aulacoseira ambigua</i>	15	<i>Staurosira construens</i>	14
PALEB	<i>Aulacoseira ambigua</i>	13	<i>Fragilaria crotonensis</i>	10
OHMUZ	<i>Aulacoseira laevisissima</i>	61	<i>Cyclostephanos dubius</i>	9
OHCRY	<i>Asterionella formosa</i>	71	<i>Aulacoseira ambigua</i>	7
OHSUM	<i>Asterionella formosa</i>	18	<i>Synedra tenera</i>	11
OHEAS	<i>Asterionella formosa</i>	15	<i>Fragilaria crotonensis</i>	15
OHPIP	<i>Asterionella formosa</i>	15	<i>Fragilaria crotonensis</i>	15
OHMEY	<i>Staurosira construens</i>	50	<i>Cyclotella ocellata</i>	18
OHNOR	<i>Staurosira construens</i>	31	<i>Pseudostaurosira brevistriata</i>	13
OHSPV	<i>Staurosira construens</i>	18	<i>Cyclotella ocellata</i>	16
OHSPR	<i>Pseudostaurosira brevistriata</i>	26	<i>Cyclotella ocellata</i>	21
OHWES	<i>Pseudostaurosira brevistriata</i>	22	<i>Asterionella formosa</i>	13
OHNES	<i>Pseudostaurosira brevistriata</i>	17	<i>Fragilaria capucina v. mesolepta</i>	9
PAEDI	<i>Fragilaria capucina v. mesolepta</i>	32	<i>Cocconeis placentula v. lineata</i>	11
PAPLE	<i>Fragilaria crotonensis</i>	17	<i>Asterionella formosa</i>	14
PACAN	<i>Tabellaria flocculosa</i>	29	<i>Aulacoseira ambigua</i>	10
PACON	<i>Tabellaria flocculosa</i>	15	<i>Puncticulata bodanica</i>	11
PASPR	<i>Cyclotella ocellata</i>	39	<i>Fragilaria capucina v. mesolepta</i>	21
OHBRA	<i>Cyclotella ocellata</i>	16	<i>Achnantheidium catenatum</i>	16
OHSAN	<i>Stephanodiscus hantzschii</i>	26	<i>Stephanodiscus minutulus</i>	12
OHPUN	<i>Stephanodiscus parvus</i>	22	<i>Fragilaria crotonensis</i>	16
OHSIL	<i>Stephanocyclus meneghiniana</i>	18	<i>Stephanodiscus vestibulus</i>	15
NYFIN	<i>Cocconeis placentula</i>	12	<i>Fragilaria capucina</i>	10

**Table V. Dominant & subdominant diatom species present at each lake.**

Code	Diversity	Evenness	Richness
	H	J	S
NYFIN	3.3	0.79	64
OHAQU	3.0	0.76	52
OHBAS	2.9	0.73	52
OHBRA	2.9	0.76	45
OHCRY	1.3	0.43	18
OHEAS	2.9	0.76	46
OHMEY	1.7	0.52	24
OHMUZ	1.6	0.51	22
OHNES	3.2	0.80	52
OHNOR	2.7	0.71	48
OHPIP	2.9	0.82	36
OHPUN	2.7	0.76	33
OHSAN	2.8	0.78	34
OHSIL	2.9	0.76	43
OHSIP	2.2	0.62	33
OHSPR	2.5	0.67	41
OHSPV	2.5	0.73	32
OHSUM	3.0	0.81	39
OHSUN	2.0	0.55	41
OHWES	3.0	0.74	56
OHWYO	1.5	0.46	28
PACAN	2.7	0.74	36
PACON	3.0	0.81	41
PACRY	2.2	0.60	36
PAEDI	2.5	0.72	32
PALEB	3.6	0.82	77
PAPLE	2.5	0.79	25
PASAN	3.0	0.76	50
PASPR	2.4	0.63	46
PASUG	3.1	0.79	49

**Table VI. Shannon diversity (H), evenness (J), and richness (S, number of species) in 30 lakes. An evenness value close to 1.0 implies a diverse and equally divided population. J values close to 0.1 represent dominance by a small number of species.**

competition theory predicts that species diversity must be low. The lowest diversity score was found in hypereutrophic Crystal Lake (OH), which had the highest total phosphorus and ammonia, high total nitrogen and modest silica values supporting the theory of decreased diversity with increased productivity.

Maximum species richness is achieved when community structure is shifted or disturbed as a result of external or environmental forces. The process of allogenic forces opening niches for opportunistic organisms can be defined as ecological disturbance (Reynolds, 2006). The realization that diversity may be highest when disturbances are not very frequent yet not too rare gave rise to the intermediate disturbance hypothesis (Connell, 1978). In aquatic ecosystems, external events that lead to disturbance in community structure include rapid fluctuations of nutrients, and dramatic or persistent upset of established water column structure, as would occur in draining, flushing, or intense wind-mixing of a system (Reynolds, 2006). Biologist reports (Pennsylvania Fish and Boat Commission, 2007) indicate that gizzard shad *Dorosoma cepedianum*, was recently introduced into Lake Leboeuf (PA) ([http://www.fish.state.pa.us/water/lakes\\_nonpfbc/leboeuf/00index.htm](http://www.fish.state.pa.us/water/lakes_nonpfbc/leboeuf/00index.htm)). The young of this fish are known to feed on algae, and the infestation of this non-native species may have been a factor contributing to disruption of the established aquatic food web through selective grazing leading to disturbance in the diatom community.

High species diversity is also attained in lakes with a variety of available habitats (Hutchinson, 1961). A system with a large wetland buffer surrounding the lake, numerous stream inputs, large littoral zone with a multitude of substrates (sand, rocks, and mud), well established macrophytes, and a deep pelagic zone will inevitably support

a more diverse diatom assemblage than a shallow, muddy system with little depth variation and no macrophytes. Lake Leboeuf<sub>(PA)</sub> could be categorized as the former although lakes of both types were sampled in this research, resulting in a broad range of diatom diversity and species richness.

### 3.2.1 Planktonic taxa

The three planktonic species with highest total abundance across all lakes were *Aulacoseira granulata*, *A. ambigua*, and *Asterionella formosa*. Studies from European lakes find these taxa in turbid, wind-mixed, nutrient rich and slightly alkaline systems with low light penetration (Bennion, 1994; Bennion et al., 2001; Bradshaw et al., 2002). In this study, they are dominant in a large number of eutrophic systems. *Aulacoseira granulata* reached highest abundance in hypereutrophic systems while *A. ambigua* was dominant in less enriched systems like Sugar Lake<sub>(PA)</sub>, the site with the lowest TP. *Cyclostephanos dubius* was often the subdominant in lakes where *A. granulata* was dominant. Three other lakes dominated by *A. ambigua* are eutrophic systems. The *Aulacoseira* species present in this research dominated sites with higher nickel and silica measurements, and occurred over a wide total phosphorus range, occurring in lakes with TP between 12 ug/L and 153 ug/L.

*Asterionella formosa* peaked in abundance in deep lakes but showed wide tolerance range in TP. It comprised 71% of the diatoms from Crystal Lake<sub>(OH)</sub>, the site with the highest TP (338.9 ug/L), but was also dominant in Phippen Lake<sub>(OH)</sub>, where the TP was 24.4 ug/L. This finding is not entirely anomalous. These data suggest the *Aulacoseira* species and *Asterionella formosa* are generalists capable of exploiting

various concentrations of TP to achieve community dominance. These two genera are also considered disturbance-tolerant and indicative of high lake productivity (Grime, 1977). Lakes low in nutrients for most of the year may experience seasonal mixing, resulting in upwelling of water and sediments rich in phosphorus and nitrogen. This pulse may stimulate the growth of diatoms that would otherwise not succeed in the low nutrient open waters (Kuhn et al., 1981). This research sampled in fall, when increased wind and storm activity contributes to mixing and breakdown of summer thermal stratification in the water column. Mixing offers nutrients that were previously sequestered in colder bottom waters to phytoplankton, often causing diatom blooms. The apparent contradiction of a species thought to occur only at higher productivity but dominating a system with low TP supports the idea that certain species can exploit brief nutrient pulses.

Secchi depth measurements were low (< 4 m) in all cases, revealing that lakes sampled in this research are turbid. This may have been an artifact of mixing and sediment resuspension occurring in most lakes during fall turnover but may also reflect a large amount of particulates caused by erosion, sediment loading, organic fragments, and algal biomass. Light penetration in autumn decreases as the sun's rays become less direct. *Asterionella formosa* in particular is capable of higher growth rate than other diatoms under lower light conditions which may partially account for its dominance in a number of samples (Reynolds, 2006).

*Stephanodiscus* species are well-documented indicators of nutrient enrichment (Heinrichs, 2005; Heinsalu, 2007), and were major components in six of the eutrophic lakes studied. These forms include *Stephanodiscus parvus*, *S. niagarae*, *S. hantzschii*, *S.*

*medius*, and *S. vestibulus*. In eutrophic sites where *Aulacoseira granulata* or *A. ambigua* species dominated, *Stephanodiscus* species were present in very low numbers or absent from species counts, suggesting *Aulacoseira* may competitively exclude *Stephanodiscus* during nutrient assimilation and growth.

*Cyclostephanos dubius* is a reported eutrophic species (Reavie et al., 2006), and occurred in large numbers in relatively shallow systems with high TP (between 60 ug/L and 153 ug/L). This species was also present as a subdominant diatom in three hypereutrophic lakes dominated by *Aulacoseira granulata*. *Cyclostephanos invisitatus* is typically found in nutrient rich waters (Bradshaw et al., 2002; Bennion et al., 2004), and occurred in three lakes with TP values above 80 ug/L. *Cyclostephanos invisitatus*, *C. dubius*, and *Stephanodiscus parvus* were present in East Lake (OH) but each composed less than three percent of the total cell counts. *Cyclostephanos* species in this calibration set have TP optima of 108 ug/L and 111 ug/L and suggests their presence may be due to increased TP availability from seasonal mixing or eutrophication.

Many *Cyclotella* species are typically indicative of oligotrophic conditions (Fitzpatrick et al., 2007). *Cyclotella ocellata* is commonly found in low nutrient, deep, clear waters (Hall and Smol, 1996), and was a major taxa (13% to 39%) in six lakes with similar low TP, low TN, and relatively deep light penetration. *Stephanocyclus meneghiniana* was the dominant species in Silver Lake (OH), a site with moderately high TP, relatively low TN, and the lowest Si and Ni. *Puncticulata bodanica* var. *lemanica* achieved abundance greater than 10 percent in Pippen (OH) and Canadohta (PA), two deep lakes with similar large surface areas, low TP and moderate NH<sub>3</sub>. Both lakes were

considerably choppy when sampled, and consistent wind activity results in mixing capable of entraining the large, ornately silicified diatom.

### 3.2.2 Periphytic taxa

Benthic *Staurosira construens* occurs in slightly alkaline or circumneutral lakes (Hall and Smol, 1996), and dominated sample lakes with pH between 7.8 and 8.1, high light penetration, relatively low TP, and very comparable TN, NH<sub>3</sub>, and NO<sub>x</sub> values.

*Pseudostaurosira brevistriata* was the major taxon in systems with moderate light penetration, moderate depth, and TP between 38.9 ug/L and 91.5 ug/L, indicating it has a wide tolerance for TP, as reported by other calibration sets (Bennion et al., 1996).

*Pseudostaurosira brevistriata* is also reported to be tolerant of various conductivities (Patrick and Reimer, 1966) and was found in high numbers in lakes with Na concentrations between 17 ug/L and 19 ug/L, and Ca concentrations between 12 mg/L and 35 mg/L. *Tabellaria flocculosa* dominated two deeper PA lakes with similar TP, neutral pH, moderate NH<sub>3</sub> and relatively high light penetration. Benthic species *Nitzschia*, *Pinnularia*, and alkaliphilic *Cymbella* and *Encyonema* were present but in very low abundance.

Epiphytic species such as *Achnanthydium minutissimum* comprised around 5% of species present in lakes with moderate submerged and emergent macrophyte cover.

*Cocconeis placentula* is a well documented epiphytic diatom, and dominated Findley Lake (NY), a site with exceptionally dense submerged macrophytes. Other epiphytic and benthic species including *Gomphonema*, *Amphora*, and *Navicula* occurred, but only in very low abundances.



### 3.3 Data Exploration

#### 3.3.1 Detrended Correspondence Analysis

Detrended Correspondence Analysis (DCA) is one of the first analyses performed on a data set to determine whether species exhibit unimodal or linear responses to environmental variables measured. The choice of an appropriate species-response model is made by examining the standard deviation expressed as the length of the gradient displayed for each axis. If the first axis has a gradient length corresponding to a beta diversity value of greater than 3, unimodal methods of analysis are appropriate (Leps and Smilauer, 2003). Beta diversity measures community composition (the number of unique species in a sample) and increases as more unique species are included in analysis. A DCA of all 321 taxa encountered and 10 significant environmental variables yielded a gradient length of 3.79, indicating unimodal modeling of species response can be used.

It is necessary when constructing inference models to exclude rare species and include only those diatoms that achieve a minimum percent abundance and are present in a given number of lakes. This step ensures derivation of more realistic optima than indicator values calculated from a single occurrence of a species. For example, if *Aulacoseira granulata* occurred in a single site at 1.5% abundance with a TP value of 34 ug/L, this would be the optima. If however, *Aulacoseira granulata* occurred at an abundance of 1.5 at 4 sites with TP values of 34 ug/L, 64 ug/L, 77 ug/L, and 82 ug/L, its optima would be 64 ug/L. This illustrates the need to limit the species included in model development to those diatoms present in a large number of sample sites (DeNicola et al., 2004).

When the taxa list was reduced to 40 species ( $\geq 1\%$  abundance and present in at least 3 lakes) the DCA gradient length declined to 3.13, but unimodal methods were still appropriate (see Appendix 4 for response graphs of select abundant species and environmental variables.) Transformations of the species and environmental data resulted in no appreciable differences in further analyses; therefore only results based on non-transformed data are included.

### **3.3.2 Canonical Correspondence Analysis**

Canonical Correspondence Analysis (CCA) is a direct gradient analysis technique that defines the ordination axes as combinations of measured environmental variables that maximize dispersion of diatom species among sample locations (Leps and Smilauer, 2003). This method is contrasted with indirect gradient analysis, where theoretical variables that best fit a regression model of species response are predicted. In this research, CCA was used to determine those environmental variables most strongly affecting diatom species distribution through a process of multiple regression. The solution file generated by Canoco 4.54 for Windows can be used to determine significant effects of environmental variables.

Fifteen variables were initially assessed to determine their effect on the distribution of diatom species across sample sites. These variables were total phosphorus (TP), total nitrogen (TN), silica (Si), nitrite+nitrate ( $\text{NO}_x$ ), ammonia ( $\text{NH}_3$ ), Secchi disk depth (Secchi), maximum depth ( $Z_{\text{max}}$ ), pH (pH), surface area (SrfA), dissolved oxygen (DO), nickel (Ni), calcium (Ca), magnesium (Mg), sodium (Na), and potassium (K). Significant variables were determined by running each as the sole explanatory variable in

a CCA and examining its p-value (Table VII). Ten significant environmental variables were identified and their relationship to diatom species was explored.

Canonical correspondence analysis with variables TP, NH<sub>3</sub>, Secchi depth, Z<sub>max</sub>, Ni, Na, Mg, DO, and K explained 51% of the variation in distribution of diatom species across sample sites (sum of canonical eigenvalues/sum of unconstrained eigenvalues x 100; 1.647/3.237 x 100). This is a very large proportion of the variance in diatom species considering the multivariate and typically noisy nature of large ecological datasets. A slight arching effect indicated that some variables were correlated (Leps and Smilauer, 2003). In addition, some arrows corresponding to environmental variables were aligned in similar quadrants, indicating correlations between minerals Ca, Mg, Na and K (Davies, et al., 2002). To verify correlations that may result in redundancies, a Pearson product-moment correlation matrix was constructed (Enache and Prairie, 2002). Variables that were highly correlated ( $r > 0.6$ ) were removed based on their relative importance to diatom populations (Table VIII). From this analysis it was determined that Secchi depth and potassium were highly correlated with total phosphorus and therefore were removed from further analysis in development of the inference model. Dissolved oxygen had a high inverse correlation with ammonia, and was removed from the final subset of variables. Calcium was highly correlated with Na and Mg thus Ca and Na were removed from the final variable set. The set of predictive environmental variables was reduced to TP, NH<sub>3</sub>, Z<sub>max</sub>, Ni, and Mg. Two ordination diagrams were then produced to investigate the relationship of diatom species to the 5 environmental variables listed above (Figures 2-5)

To further identify those environmental variables most strongly determining the spread of diatom species, the 5-variable subset was reduced through a process of forward

Variable	p-value	Eigenvalue Axis	Eigenvalue Axis	Ratio
		1	2	
TP	0.002	0.365	0.443	0.82
K	0.001	0.352	0.452	0.78
Secchi	0.002	0.301	0.452	0.67
Ni	0.004	0.263	0.445	0.59
Zmax	0.001	0.265	0.510	0.52
Mg	0.006	0.245	0.554	0.44
Na	0.003	0.242	0.550	0.44
DO	0.019	0.217	0.507	0.43
NH <sub>3</sub>	0.025	0.229	0.551	0.42
Ca	0.026	0.208	0.542	0.38

**Table VII. Significance values for Monte Carlo permutations, eigenvalues for constrained CCA**

Variable	TP	Secchi	Zmax	DO	NH <sub>3</sub>	Ca	Mg	Na	K	Ni
TP	1.000									
Secchi	*-0.754	1.000								
Zmax	-0.419	0.354	1.000							
DO	-0.470	0.240	-0.117	1.000						
NH <sub>3</sub>	0.501	-0.165	0.169	*-0.812	1.000					
Ca	0.450	-0.419	-0.224	-0.212	0.249	1.000				
Mg	0.277	-0.163	-0.439	0.085	-0.015	*0.770	1.000			
Na	0.328	-0.419	-0.105	-0.119	0.005	*0.625	0.548	1.000		
K	*0.632	-0.537	-0.453	-0.251	0.154	0.483	0.4293	0.5541	1.000	
Ni	0.409	-0.423	-0.545	0.015	0.042	-0.014	-0.081	-0.219	0.406	1.000

**Table VIII. Pearson product-moment correlation matrix of 10 statistically significant variables executed in R statistical software. Total phosphorus (TP ug/L), Secchi disk depth (Secchi cm), maximum depth ( $Z_{max}$  m), dissolved oxygen (DO mg/L), ammonia (NH<sub>3</sub> ug/L), calcium (Ca mg/L), magnesium (Mg mg/L), sodium (Na mg/L), potassium (K mg/L), and nickel (Ni ug/L). Values close to 1.0 display a strong degree of colinearity. Significant correlation between two variables is indicated with an asterisk (\*), and one of the variables in a pair of correlated values was removed prior to further statistical analyses based on its known impact on diatom species distribution.**

<b>Lake Code</b>	<b>Number</b>	<b>Lake Code</b>	<b>Number</b>
OHAQU	1	PAPLE	15
OHBAS	2	PASAN	16
OHBRA	3	PASPR	17
OHEAS	4	PASUG	18
OHMEY	5	OHPIP	19
OHMUZ	6	OHSAN	20
OHNES	7	OHSIL	21
OHNOR	8	OHSIP	22
NYFIN	9	OHSPR	23
PACAN	10	OHSPV	24
PACON	11	OHSUM	25
PACRY	12	OHSUN	26
PAEDI	13	OHWES	27
PALEB	14	OHWYO	28

**Table IX. Numbers and lake codes used in CCA analysis and biplot diagrams.**

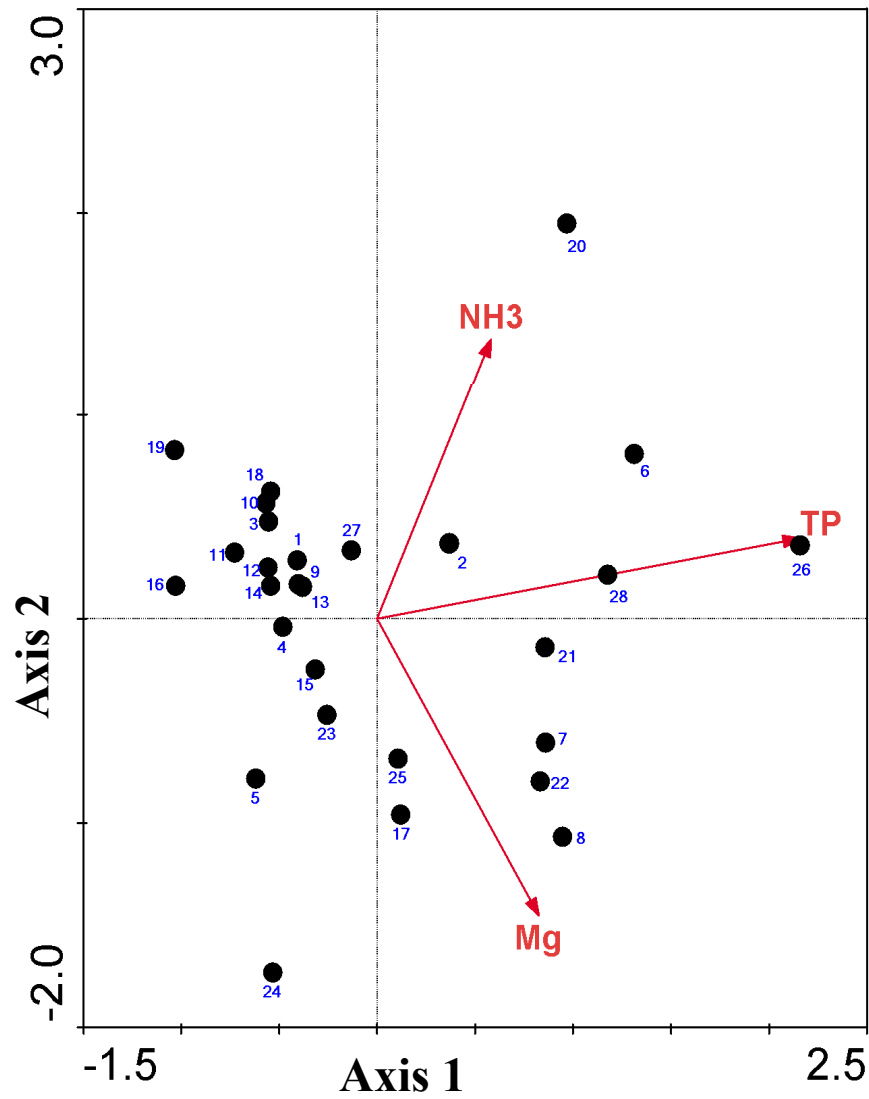


Figure 2. Ordination biplot of lake sites (black circles) and three statistically significant environmental variables (red arrows). Total phosphorus (TP ug/L), ammonia (NH3 ug/L), and magnesium (Mg mg/L). Numbers correspond to lakes in Table IX.

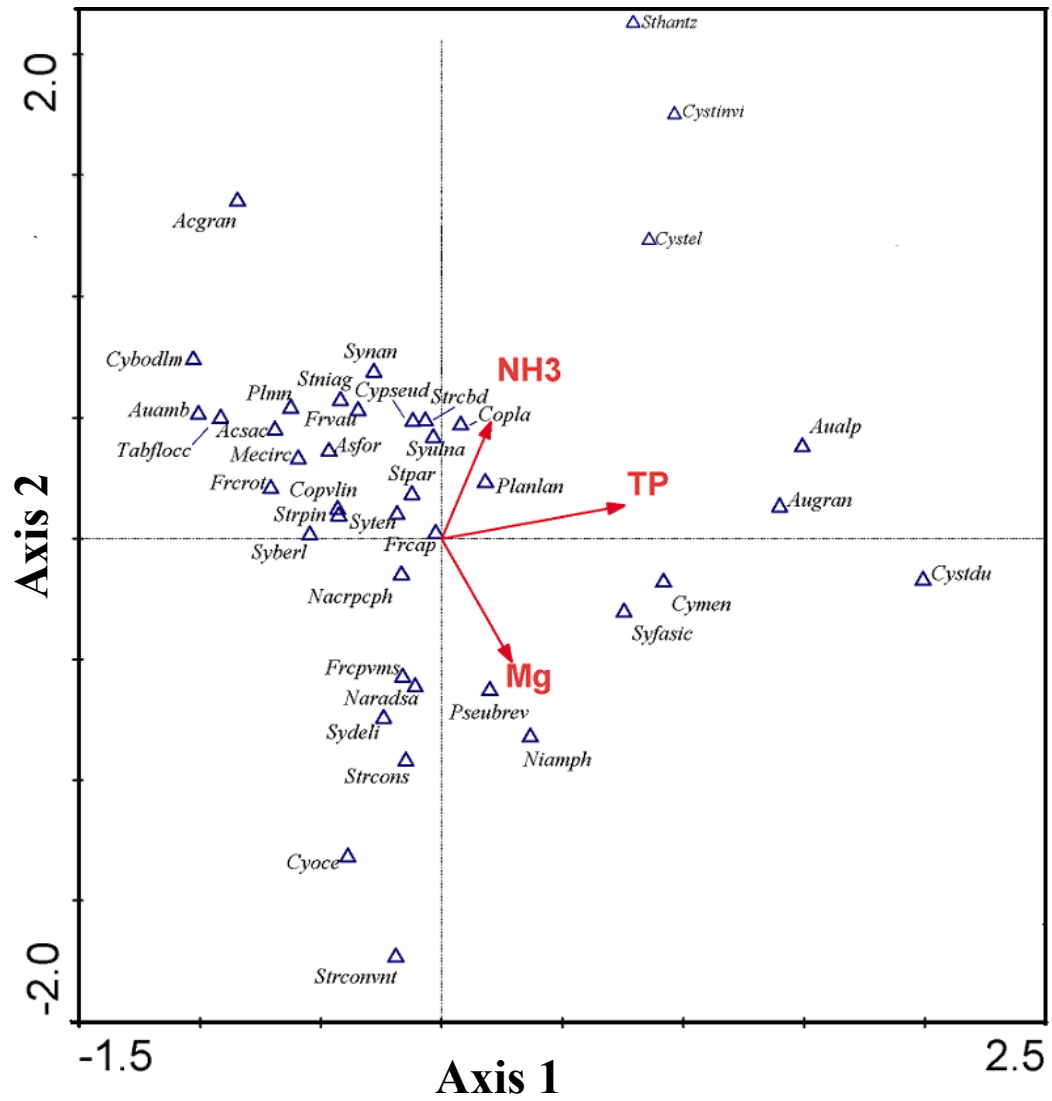


Figure 3. Ordination biplot showing species and 3 statistically significant environmental variables. Species codes listed in Appendix 3.



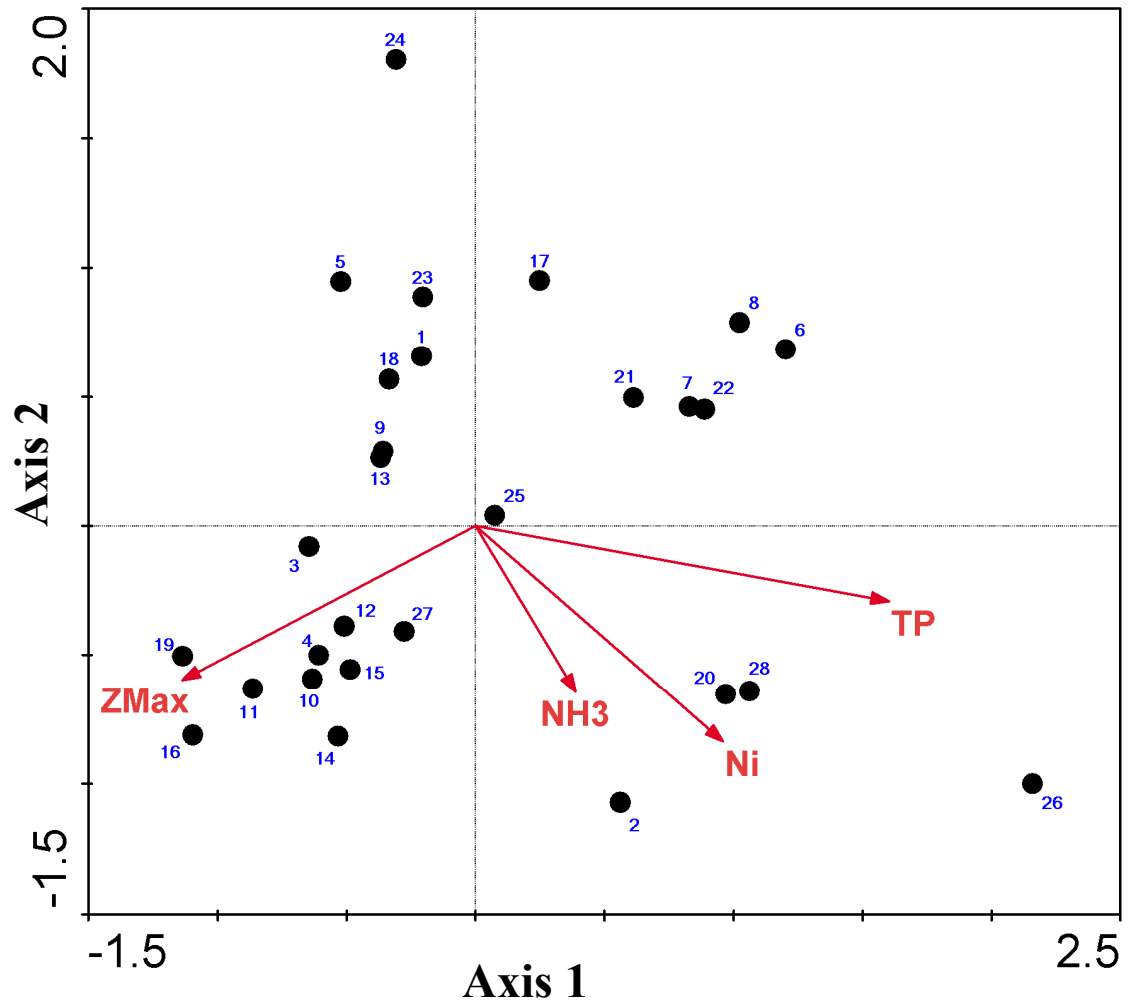


Figure 4. Biplot showing lake sites (black circles) and 4 environmental variables (red arrows). Maximum depth ( $Z_{\max}$  meters), nickel (Ni ug/L), total phosphorus (TP ug/L) and ammonia (NH3 ug/L). Numbers correspond to lakes in Table IX.

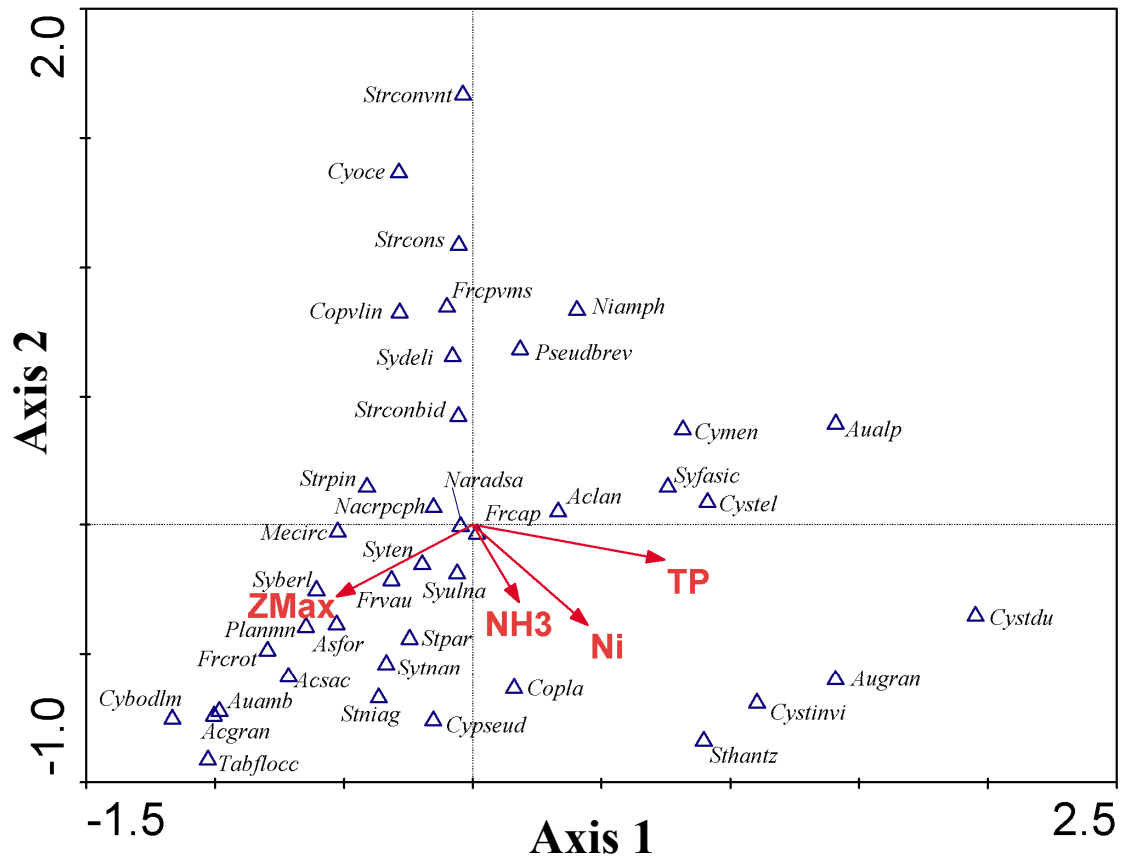


Figure 5. Ordination biplot showing species in relation to four exploratory environmental variables. Species codes listed in Appendix 3.

selection (Cooper et al., 1999). From the forward selection TP ( $p = 0.002$ ),  $\text{NH}_3$  ( $p = 0.002$ ), and Mg ( $p = 0.006$ ) were the environmental variables most strongly influencing distribution of 40 diatom species among the 28 study sites. These three variables account for 25% of the original variance (51%), illustrating their powerful and independent effect on diatom species. All three variables had variance inflation factors (VIF) of less than 1.6. A VIF less than 20.0 indicates the variable is not highly correlated with any other factor included in the analysis, and therefore explains a unique portion of the variance in diatom distribution (Hall and Smol, 1996).

Canonical coefficients and approximate t-tests of the regression coefficients can be used to determine environmental variables that account for a significant portion of variance in diatom species distribution along the first few ordination axes (ter Braak, 1986; Siver, 1999). The canonical or regression coefficient for TP and axis 1 was large (0.86), suggesting it exhibited the most powerful effect on diatom species distribution. From examination of t-values it was determined that TP was strongly correlated with the first axis, while  $\text{NH}_3$  and Mg were strongly associated with the second and third axes.  $\text{NH}_3$  was found to be correlated higher with the third axis than the second axis (Tables X & XI).

Eigenvalues reflect the significance of an ordination axis in CCA. Values close to 0.5 in ecological data reflect a large fraction of variability in species data explained by an axis. The axes are constrained to be combinations of environmental variables. The first three eigenvalues reported are canonical, the fourth is not since only three independent constraints were formed from environmental variables included in this analysis. The species responses are constrained to be a factor of the three variables used in analysis.

<b>Variable</b>	<b>AX 1</b>	<b>AX 2</b>	<b>AX 3</b>
TP	0.8619	0.0902	-0.3612
NH <sub>3</sub>	-0.228	0.4399	0.6284
Mg	0.0225	-0.5314	0.5045

**Table X. Regression/canonical coefficients for 3 statistically significant forward selected variables.**

<b>Variable</b>	<b>AX 1</b>	<b>AX 2</b>	<b>AX 3</b>
TP	7.769	0.7442	-2.9846
NH <sub>3</sub>	-2.1633	3.8189	5.4656
Mg	0.2333	-5.0463	4.7999

**Table XI. t-values of regression coefficients.**

Eigenvalues for axis 1 and axis 2 were 0.38 and 0.25, respectively and explained 11.6% and 7.6% of variation in the distribution of diatoms. The third axis had an eigenvalue of 0.19, and explained 5.9% of variation in diatom distribution. The first two axes accounted for 76.5% of the cumulative variance between diatoms and environmental variables. The relationship of species to their environment was highly correlated with axis 1 (0.87), axis 2 (0.83), and axis 3 (0.81), confirming variables that are important in determining diatom species distribution were selected for final analysis (Table XII).

Monte Carlo permutation (999 permutations) was used to determine the statistical significance of the first three axes in explaining the distribution of diatom species. All three axes were significant (axis 1,  $p = .004$ ; axis 2,  $p = .004$ ; axis 3  $p = .01$ ). The significance of axis 2 was determined by using sample scores from axis 1 as covariable data, thereby analyzing the data independent of the contribution made by the first axis. This process was repeated using sample scores from axis 1 and 2 as covariables to assess the significance of axis 3 (Leps and Smilauer, 2003).

### **3.4 Ordination**

Biplots showing the relationships between species, samples, and environmental variables were created in CanoDraw (ter Braak and Smilauer, 2002). A biplot of sample locations and environmental variables provides an easy way of grouping similar lake systems based on water chemistry parameters. In the ordination diagrams of the significant environmental variables and 28 sample sites, lakes with the highest TP values are shown along the right side of the first axis (Figure 2). These sites include Sandy<sub>(OH)</sub>, Muzzy<sub>(OH)</sub>, Sunny<sub>(OH)</sub>, and Wyoga<sub>(OH)</sub> lakes. These lakes are clearly separated along

<b>Axis</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>Total inertia</b>
Eigenvalues	0.375	0.247	0.191	0.339	3.237
Species-environment correlations	0.872	0.826	0.809	0	
Cumulative percentage variance					
of species data	11.6	19.2	25.1	35.6	
of species-environment relation	46.2	76.5	100	0	
Sum of all eigenvalues					3.237
Sum of all canonical eigenvalues					0.813

**Table XII. Results of CCA with 40 species constrained to environmental variables TP, NH<sub>3</sub>, and Mg.**

the horizontal axis from Pippen<sub>(OH)</sub> and Sandy<sub>(PA)</sub>, lakes with the lowest TP concentrations. Sites in the middle of this gradient have measured TP in the range of 12 - 153 ug/L. Sample lakes with the highest NH<sub>3</sub> concentrations are found in the right quadrant of the ordination diagram. Lakes Wyoga<sub>(OH)</sub>, Sandy<sub>(OH)</sub>, and Muzzy<sub>(OH)</sub> had the highest NH<sub>3</sub> concentrations and were clearly distinguished from Meyers<sub>(OH)</sub>, Springfield<sub>(OH)</sub>, Spring Valley<sub>(OH)</sub>, and Pleasant<sub>(PA)</sub>, which had the lowest NH<sub>3</sub> values. In this analysis hypereutrophic systems were grouped together on the right side of the first axis, while more mesotrophic or moderately eutrophic sites were clustered on the left. Lakes with Mg levels above 10 mg/L were situated along the bottom of axis 2, and were farthest away from sites with lower concentrations of Mg positioned at the middle of the second axis (Fig. 2).

Diatom species that respond to TP, NH<sub>3</sub>, and Mg were identified by constructing a biplot of species and environmental variables. Species present in lakes with high TP concentrations are shown on the right side of axis 1, while species able to exist in lakes with lower TP are shown on the left side. *Aulacoseira granulata*, *A. alpigena*, and *Cyclostephanos dubius* increase in abundance with increases in TP. Species more common in alkaline water are shown along the bottom of axis 2, while species found in systems with a neutral pH and lower Mg are shown in the top of the biplot. Diatoms that occur in sites with higher NH<sub>3</sub> are shown along the top of axis 2, and species present in sites with lower NH<sub>3</sub> concentrations are displayed towards the bottom of the second axis (Fig. 3).

### **Ordination with TP, NH<sub>3</sub>, Ni, Z<sub>max</sub>**

An alternative biplot was constructed using TP, NH<sub>3</sub>, Ni and Z<sub>max</sub> to examine the lakes according to depth and nickel content in addition to nutrient levels (Fig. 4) and is useful in determining major differences among lakes resulting from these variables. Lakes with the highest TP are present along the right side of axis 1, as are shallow lakes with elevated nickel content. Shallow hypereutrophic lakes like Bass<sub>(OH)</sub> and Sunny<sub>(OH)</sub>, with high total phosphorus, highest nickel values, and Z<sub>max</sub> depths of less than 2m, were along the right side of the first axis, while the deepest mesotrophic lakes such as Phippen<sub>(OH)</sub>, Conneaut<sub>(PA)</sub>, and Sandy<sub>(PA)</sub> are on the left side of this axis. This ordination may be useful in detecting systems especially vulnerable to rapid eutrophication because of their shallow depth. The effects of nutrient enrichment are readily apparent in systems with less water to dilute the nutrients. A shallow lake with phosphorus-laden sediments is sure to suffer from a long history of algal blooms, low dissolved oxygen, and other symptoms associated with eutrophication. These events are brought about primarily from internal loads of phosphorus re-suspended through mixing. Identification of systems vulnerable to these effects can help direct restoration or lake management efforts toward controlling autochthonous rather than allochthonous nutrient sources.

A biplot of species vs. Z<sub>max</sub>, TP, NH<sub>3</sub>, and Ni was constructed (Fig. 5). This plot effectively separated planktonic taxa from benthic. Water depth is important in determining differences in available diatom habitat (Rühland and Smol, 2002; Schmidt et al., 2004), and can account for dominance of periphytic or planktonic organisms (Hall and Smol, 1992). *Synedra* and *Asterionella* species are located along the maximum depth vector, and were dominant in plankton samples. Benthic species such as



*Pseudostaurosira brevistriata*, *Nitzschia amphibia*, and *Staurosira construens* and its associated varieties were clustered opposite the depth vector, and dominated shallow systems.

### **3.5 Inference Model Development**

Environmental variables that are statistically significant predictors of diatom species distribution are good candidates for inference model development. This relationship is examined by comparing the ratios of constrained CCA eigenvalues. Robust inference models can be constructed from environmental variables when ratios of the first and second eigenvalues in a constrained CCA are high (Hall and Smol, 1996; Bigler and Hall, 2002). The eigenvalue ratio of CCA constrained to TP was 0.82 in my research, indicating that a strong TP inference model can be developed. Statistically robust TP inference models have been developed for data sets with eigenvalue ratios as low as 0.36 and 0.45 (Schönfelder et al., 2002). Eigenvalue ratios from CCA constrained to Ni (0.59),  $Z_{\max}$  (0.52), Na (0.44), Mg (0.44), and  $\text{NH}_3$  (0.42), indicated that each variable could potentially be reconstructed from diatom fossil assemblages, however inference models were only developed for the three statistically significant variables TP,  $\text{NH}_3$ , and Mg using  $C^2$  (Juggins, 2007). Exploratory models were constructed for Ni,  $Z_{\max}$ , and Na, but were unreliable.

#### **3.5.1 Total phosphorus inference model**

Priority was given to development of the total phosphorus inference model. Total phosphorus is the variable most often targeted in cultural eutrophication management

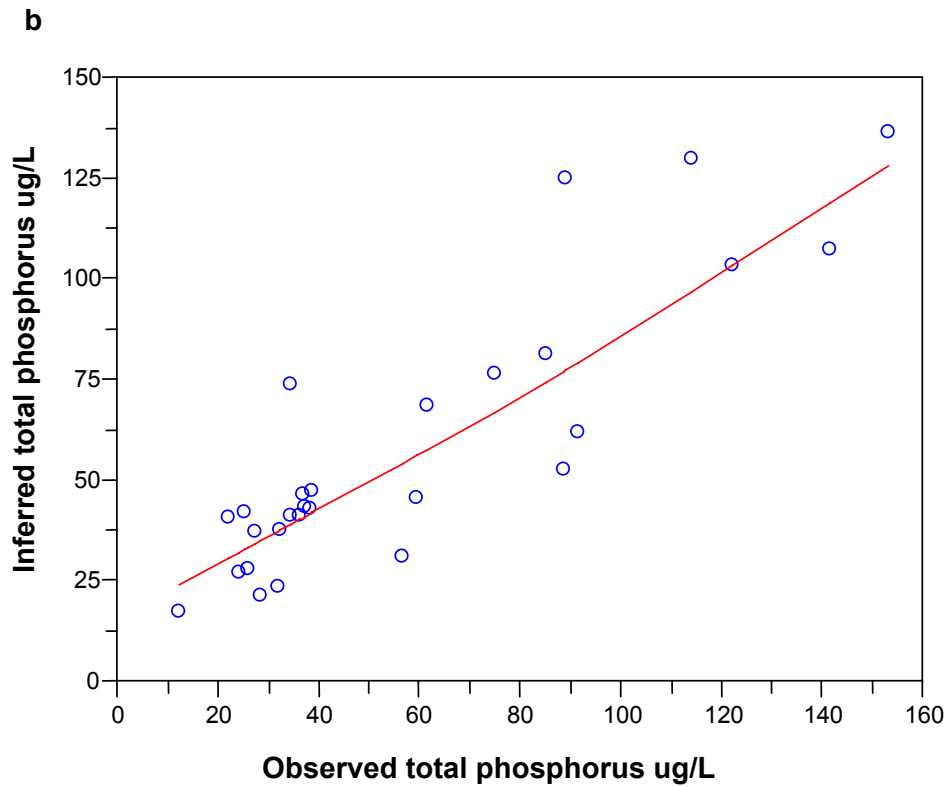
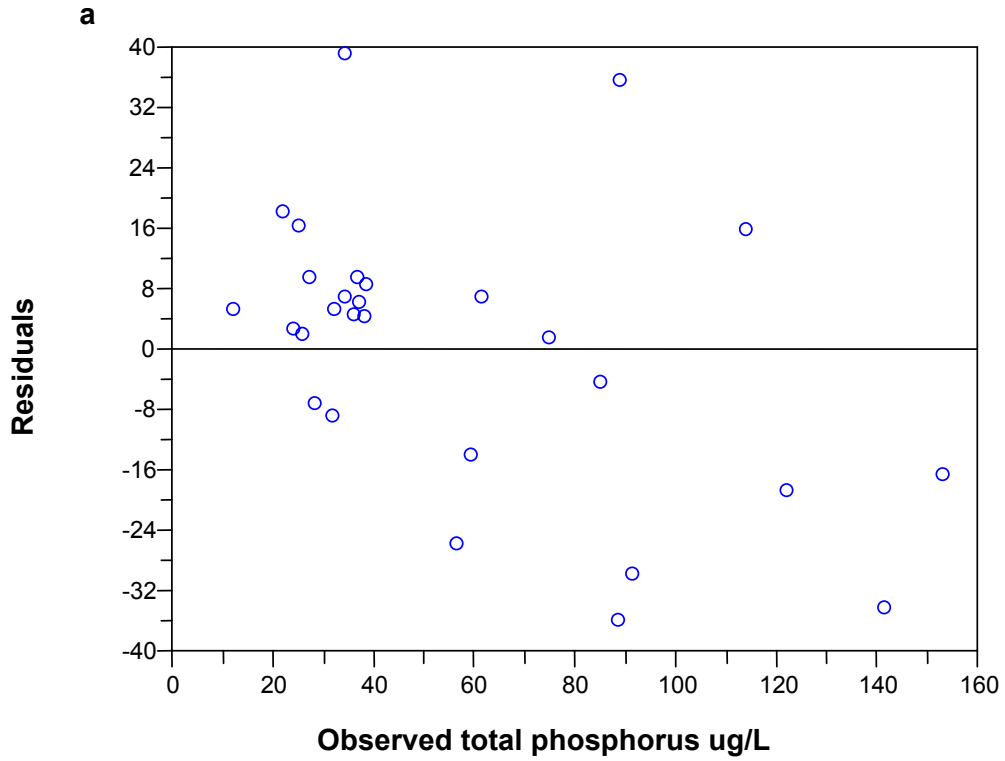
efforts, and has been the subject of numerous worldwide diatom calibration studies. This study covered a large range of TP concentrations when compared with other diatom calibration sets and is indicative of the nutrient enrichment experienced by lakes in NE Ohio. Mean TP from all 30 lakes was 72 ug/L, and 58 ug/L was the average TP excluding two outlier sites (Crystal<sub>(OH)</sub> and Punderson<sub>(OH)</sub>) which had extremely high TP values. Total phosphorus optima and tolerances for the 40 most abundant diatom species in this study are given in Table XIII. The root mean squared error (RMSE) of prediction is 17 ug/L and is an artifact of the long TP gradient and small number of lakes in the data set. Other calibration studies with shorter gradients report smaller RMSE. There is a strong linear correlation ( $R^2 = .77$ ) between inferred and observed TP, indicating good reconstructive capability of this TP inference model (Figures 6a, 6b).

The model could be slightly improved by dropping additional sites that had high TP values, but not without costs. Removing additional sample lakes would limit the range of reconstructed TP values within prediction error. As a result, only two extreme outlier sites were dropped from model development. Diatom species that occurred at only a few sites were also excluded from the model to improve the reconstructive capability. The low number of abundant species necessitates inclusion of some diatoms with larger TP tolerances and lower percent abundances.

In an effort to explore modifications to improve the TP inference model, species with an N2 lower than 4.0 (effective occurrence, calculated as the reciprocal of Simpson's diversity index following Hill, 1973) were removed from model development, reducing the training set from 40 taxa to 33 widely occurring species. This had the negative effect of reducing the number of species used in the reconstructive model, and

Name	# of lakes	Abundance	N2	Optimum	Tolerance	Bootstrap Optimum	Bootstrap Tolerance
Accons	3	3.1	2.5	30.2	7.8	31.3	12.6
Acthmn	16	6.9	12.0	36.8	14.6	37.0	13.6
Acsac	4	1.9	3.7	24.8	11.1	25.4	10.6
Asfor	20	21.9	12.4	46.2	27.4	46.9	26.8
Aualp	4	5.0	3.9	105.0	28.9	105.6	22.9
Auamb	12	40.6	5.6	28.3	13.8	29.4	13.6
Augran	13	72.1	5.5	95.3	43.4	93.7	37.8
Copla	15	15.1	7.5	58.9	42.2	57.7	38.9
Copvlin	4	11.4	2.7	40.9	19.7	46.0	19.9
Cyocce	10	39.0	6.0	42.5	17.7	42.4	16.7
Cypseud	4	4.8	3.1	43.9	26.7	49.3	21.8
Cystel	4	7.2	2.4	89.2	61.0	82.0	45.8
Cystdu	9	30.6	4.2	111.2	40.7	106.6	34.2
Cystinvis	4	5.8	2.6	120.5	46.1	118.2	39.9
Frcap	16	10.0	12.1	55.3	35.2	56.7	34.2
Frcpvms	10	32.4	3.6	46.8	23.3	47.3	19.0
Frcrot	19	16.5	13.7	37.9	17.4	38.3	16.8
Frvau	12	3.7	9.4	49.8	34.1	50.8	31.3
Mecirc	4	2.4	3.7	37.2	13.5	39.0	14.4
Nacrpcph	6	2.2	5.6	50.5	20.4	50.6	17.9
Naradsa	3	3.8	2.6	47.6	28.4	54.5	24.6
Niamph	3	2.0	2.7	58.7	39.7	59.3	29.0
Planlan	5	1.8	4.8	58.0	53.9	59.5	44.4
Pseurbrev	15	25.5	7.9	61.7	33.1	63.2	32.7
Puncbod	11	12.5	6.1	32.3	22.8	33.5	21.8
Staurcons	25	50.4	7.5	48.4	32.9	50.2	30.6
Staurconbi	10	10.4	7.1	58.7	40.2	60.0	36.3
Starconven	10	23.9	4.2	46.6	32.2	47.1	25.6
Starlpinn	16	13.3	10.6	43.8	27.0	44.4	25.6
Sthantz	6	26.3	2.2	109.2	60.9	96.7	44.3
Stniag	8	10.5	5.6	45.6	28.1	45.6	25.4
Stpar	7	14.0	4.8	52.4	29.7	54.5	26.4
Stpcymen	15	17.5	7.5	81.6	31.2	80.5	31.4
Syberlo	15	4.1	13.3	39.7	23.2	40.3	21.4
Sydeli	3	1.4	2.9	43.8	24.7	41.8	21.5
Syten	17	10.6	10.8	54.1	34.4	54.7	32.9
Synan	9	3.7	7.7	55.8	44.6	56.8	38.8
Syulna	21	7.6	16.5	57.3	39.4	57.6	37.9
Tabfasci	4	2.0	3.7	87.4	48.5	86.4	38.1
Tabflocc	6	29.2	3.8	30.6	9.5	30.6	9.9

**Table XIII. TP optima and tolerances, ug/L for 40 abundant diatom species.**



**Figure 6 a. Observed vs. WA residuals (observed – inferred) values for 28 lakes.**  
**Figure 6 b. Observed vs. WA diatom-inferred TP values for 28 lakes.  $R^2 = 0.77$**

decreased the  $R^2$  value from 0.77 to 0.58. The RMSE also increased, although differences between bootstrapped and raw  $R^2$  values were decreased. In order to reduce the error of prediction while simultaneously including as many commonly occurring diatoms as possible, the training set with 40 taxa and 28 lakes was selected.

The reliability of diatom species optima were determined by measures of fit with and without bootstrapping (Hall and Smol, 1996). For this TP inference model, the  $R_{2\text{ boot}}$  was 0.47, while the raw  $R_2$  was 0.77. The  $\text{RMSE}_{\text{boot}}$  was 28.9 ug/L, while the raw RMSE was 17 ug/L. These differences may seem large, but result from the low number of sites and species included in model development, and do not negate accurate TP inference from the diatom species. The RMSE of 17 ug/L represents 12.8% of the range in TP values. Cross validation using leave-one-out or jackknifing was also performed but did not produce significant differences when compared with bootstrapping, therefore bootstrapping was selected as the final method of cross validation. Bootstrapping in this analysis leaves out 9 lakes and reconstructs TP values using the remaining 19 measurements in each cycle. This is a demanding process and results in lower correlations between observed and inferred TP values.

### **3.5.2 Ammonia inference model exploration**

The  $R^2$  value of the relationship between observed and inferred  $\text{NH}_3$  measurements was 0.76 with a RMSE of prediction of 70 ug/L. When Sandy Lake <sub>(OH)</sub>, a system with extremely high  $\text{NH}_3$  was removed, the RMSE was reduced to 55 ug/L, but the  $R^2$  value was reduced to 0.52. Since Crystal <sub>(OH)</sub> and Punderson <sub>(OH)</sub> lakes had already been removed because due to their high  $\text{NH}_3$  values, Sandy Lake <sub>(OH)</sub> was retained in

model development at the cost of a slightly higher error of prediction. Comparison of  $R^2$  with  $R^2_{boot}$  showed that  $NH_3$  indicator values for diatoms may not be accurately captured. The raw  $R^2$  was 0.77, while the  $R^2_{boot}$  was .0127, representing strong disagreement between raw and cross-validated measures of fit in observed vs. weighted average (WA) - inferred values. The raw RMSE was 70.1 ug/L, while the  $RMSE_{boot}$  was 153.0 ug/L. The differences are most likely a result of the large range in ammonia concentrations (1800 ug/L). This indicates that optima and tolerance values for  $NH_3$  reconstruction may not be robust, and this inference model should not be blindly utilized. An increase in the number of sample sites and diatom taxa may increase the strength of the ammonia inference model.

### **3.5.3 Magnesium inference model exploration**

The model for magnesium inference had an  $R^2$  value of 0.67 with a RMSE of prediction of 2.6 mg/L. Exclusion of Summit Lake (OH) increased the  $R^2$  value to 0.72, but only slightly reduced the RMSE to 2.3 mg/L. In order to include as many systems as possible, Summit was retained in model development with a slightly lower  $R^2$  linear correlation coefficient between observed and inferred Mg values. Examination of cross-validated measures of model robustness indicated that the magnesium inference model cannot be used to accurately reconstruct historical levels in freshwater lakes of the ecoregions studied. The raw  $R^2$  was 0.67, while the  $R^2_{boot}$  was 0.31. The raw RMSE of prediction was 2.6 mg/L, while the  $RMSE_{boot}$  was 4.1 mg/L. These values are high, and would be expected to decline with inclusion of more sample locations and an increase in the number of species utilized in model development.

### **Maximum depth inference model exploration**

An exploratory model was also developed to determine if diatom species could be used to accurately predict the maximum depth of a lake using the same procedure as described above. The  $R^2$  raw value of the goodness of fit between observed and WA inferred  $Z_{\max}$  values was 0.63, while the  $R^2_{\text{boot}}$  was 0.31. The raw RMSE of prediction was 2.2 m and the  $\text{RMSE}_{\text{boot}}$  was 3.2 m. These differences indicate that the number of sample lakes and species encountered do not yield statistically reliable  $Z_{\max}$  optima and tolerances.

### **Nickel inference model exploration**

The reliability of diatoms as indicators of nickel concentrations in the sample lakes was also assessed. A comparison of the  $R^2$  with and without cross validation indicated optima for Ni concentrations developed in this research are not robust. The raw  $R^2$  value was 0.51, while the  $R^2_{\text{boot}}$  was 0.17. The raw RMSE was 69 ug/L, while the  $\text{RMSE}_{\text{boot}}$  was found to be 99 ug/L. These large errors between observed and diatom-inferred Ni values suggest realistic optima have not been achieved for the 40 species used in this analysis.

### **Sodium inference model exploration**

An exploratory model was constructed to determine whether diatoms from this research could be used to predict sodium concentration. The raw  $R^2$  was 0.68, while the  $R^2_{\text{boot}}$  was 0.29. The raw RMSE was 3.36 mg/L Na, while the  $\text{RMSE}_{\text{boot}}$  was 5.37 mg/L. The large differences in bootstrapped versus non-validated model performance indicate that reliable sodium optima and tolerance for these diatom species were not obtained.

### 3.6 Comparison of TP optima values

Calibration sets with dissimilar environmental gradients often show poor agreement in nutrient optima. Certain diatom species common in low nutrient regions may also exist under higher nutrient conditions or exhibit physiological plasticity. This point is illustrated in a number of species when my optima are compared with those of training sets from lakes with lower ranges of TP concentrations (Table XIV). The WA-TP optima for 40 abundant diatoms were compared to existing calibration sets to examine the model constructed.

The range of TP (0.8 ug/L to 49.2 ug/L) in a calibration data set from English lakes (King et al. 2000) was much smaller than my study and resulted in relatively low species optima when compared with my calibration set. In fact, TP optima from my data set were nearly two times higher than those calculated for English lakes. The lakes in King et al. (2000) are much deeper (average depth = 33m) and larger (average surface area = 1055 acres) than my sample lakes (average depth = 8m, average surface area = 134 acres). Larger lakes can sustain greater inputs of nutrients without showing symptoms of eutrophication. Deep lakes may sequester phosphorus in sediments that are easily resuspended in shallow lakes. The English lakes exist within slate, Silurian, and Volcanic rocks, which may supply nutrients at rates different than the shales, sandstones and conglomerates of my sample lakes. The English Lake District is subject to higher rainfall than the ecoregions in my study (approx. 2500 mm/year vs. approx. 889mm/year;



<b>Reference</b>	<b>TP (ug/L) min</b>	<b>TP (ug/L) max</b>	<b>Region</b>
Fritz et al 1993	1	51	Michigan
Reavie et al 1995	6	41	British Columbia
Hall & Smol 1992	5	28	British Columbia
Hall & Smol 1996	2.7	24.3	Ontario
Reavie & Smol 2001	14	54	Ontario
King et al 2002	0.8	49.2	English lakes
Enache & Prairie 2002	2.8	52	Quebec
Ramstack et al 2003	10	100	Minnesota
Reavie et al 2006	1	521	Great Lakes*
Scotese 2008	12	153	Ohio/Pennsylvania

**Table XIV. List of TP calibration studies, range of TP values in ug/L and regions examined. \*Reavie et al 2006 sampled nearshore areas, riverine wetlands, and tributaries of the Laurentian Great Lakes, USA.**

USGS, 2005), which may flush and recharge the lakes with fresh water. The landscape of the English Lake District is composed of prairies, forests, and open fields, contrasted with cropland, housing, and impervious surface surrounding lakes in OH and PA. Perhaps most importantly, land use in the English Lake District consists of some light agriculture and tourism rather than industrial, urban, and agricultural land use around my sample lakes. Conditions in the low-nutrient English Lakes support many diatoms that did not occur in my sample lakes, and contribute to differences between the two calibration sets.

A calibration data set developed from 46 North American lakes in British Columbia (B.C.) (Hall and Smol, 1992; Hall and Smol, 1996) was also constructed from a low TP gradient (2 ug/L to 28 ug/L) and optima were lower than those developed in the current study. Sample lakes in these B.C. studies exist within cool temperate, low human-impact, non-agricultural forested areas, contributing to the lack of eutrophic systems. The lowest TP optima in my calibration set was 28 ug/L for *Aulacoseira ambigua* compared with 16.9 ug/L in the B.C. set, and was higher than the highest optimum (23.3 ug/L) in the B.C. set. A second training set developed for lakes in British Columbia ranged between 6 ug/L and 41 ug/L TP (Reavie et al., 1995). These optima were also much lower when compared with optima for diatom species in OH and PA. This B.C. model included a total of 151 species, while my data set was limited to 40 taxa.

In a calibration set from Minnesota (Ramstack et al., 2003), TP concentrations from 55 lakes ranged between 10 ug/L and 100 ug/L, close to the range obtained in my set (12 ug/L to 153 ug/L). Similarities in TP between these two studies resulted in many comparable optima (*Aulacoseira ambigua* 28 ug/L this study, 29 ug/L Minnesota set;

*Fragilaria vaucheriae* 50 ug/L this study, 49 ug/L Minnesota set; *Fragilaria capucina* var. *mesolepta* 47 ug/L this study, 40 ug/L Minnesota set; *Stephanodiscus niagarae* 46 ug/L in both data sets). These similarities reveal that indicator values (optima) are often analogous when comparable ranges of environmental variables are captured. The study regions are also similar in terms of land use. Lakes in the Twin Cities metropolitan region receive urban runoff and storm water input and lakes in the Western Corn Belt Plains ecoregion of Minnesota are large, shallow, and subject to agricultural runoff (Ramstack et al., 2003). Lakes in Ohio and Pennsylvania are exposed to the same point and non-point sources of nutrients, contributing to agreement between these two calibration sets.

In contrast, TP optima for diatom species from 41 lakes in a Michigan study (Fritz et al., 1993) were much lower than those identified from lakes in Ohio and Pennsylvania (*Aulacoseira granulata* in the Michigan set was 18.1 ug/L versus 95 ug/L in my study). This taxon dominated hypereutrophic Ohio lakes but has a wide tolerance range. Sample lakes from north-western lower Michigan are alkaline, large, and deep, surrounded by forested watersheds with varying but comparatively low human impact. The naturally high dissolved carbonates in these lakes bind with phosphorus making it unavailable and contributing to their low phosphorus content. Total phosphorus concentrations in the Michigan set ranged from 1 ug/L to 51 ug/L, while the mean TP for my sample sites was 58 ug/L.

A similar pattern was seen in a calibration set based on 64 southern Ontario lakes developed by Reavie and Smol (2001). Total phosphorus concentrations were comparable to previously mentioned training sets from Canada and ranged from 4 ug/L to

54 ug/L. Total phosphorus optima were again lower than those developed from my calibration set even though it is located closer to the OH/PA region. Lakes in the Ontario data set exist in deciduous and coniferous forests on limestone and granite bedrock, and range in shoreline use from pristine to recreational development (Reavie and Smol, 2001). In contrast, nutrient inputs to the relatively small and shallow lakes of the OH/PA region are more heavily impacted by land use and development.

The Great Lakes Environmental Indicators (GLEI) project sampled 155 coastal wetlands, embayments and shoreline sites on the U.S. side of all 5 Laurentian Great Lakes (Reavie et al., 2006). Total phosphorus measurements for the GLEI project covered a very broad range from 1 ug/L to 521 ug/L, and included the extent of TP concentrations found in my research, although mean TP was only 27 ug/L for the GLEI project. Inland lakes are comparatively small with smaller volumes of water to dilute nutrients that run directly off the surrounding landscape. In comparison, the Great Lakes have enormous water volumes and are less directly influenced by landscape nutrient runoff except in the nearshore regions along populated river mouths. The optimum for *Cyclostephanos dubius* constructed from my data set was 111 ug/L, compared to 107.5 ug/L for the GLEI study; *Staurosira construens var. binodis* 59 ug/L this study, 55.7 ug/L GLEI; *Stephanocyclus meneghiniana* 82 ug/L this study, 89.6 ug/L GLEI; *Cyclotella pseudostelligera* 44 ug/L this study, 56.8 ug/L GLEI. The higher GLEI optimum is a function of the longer TP gradient captured in the GLEI project, although a number of other indicator values were comparable to my current study. The similarities are noteworthy considering the large variation in habitats sampled in the GLEI project.

Differences in the range of TP concentrations and the total number of species used, as well as size, geology and geography account for the disagreement between optima developed in the Michigan, B.C., southern Ontario and English training sets and my data set. On average, optima developed from Canadian data sets have been about 10 units lower than those developed from the northern OH/PA lakes. Lakes in my sample region are relatively shallow, small, and more impacted and developed than lakes in any of the aforementioned calibration sets, and the TP optima reflect this enrichment. Some diatoms are present in all of the data sets. The same generalist species of diatom can exist in different water conditions among regions. The large number of eutrophic and hypereutrophic systems sampled in Ohio results in species optima that are much higher than those derived from regions based on oligotrophic and mesotrophic systems. These examples reinforce the importance of developing region-specific calibration sets. Such regional data sets can later be combined with others to yield more realistic tolerance and optima for cosmopolitan diatom species.

### **3.7 Factors influencing TP inference models**

Studies sampling similar TP gradients reported many optima close to those found in this study, but there were still a number of differences. Other than the large TP gradient, a number of factors can contribute to the difficulty of constructing reliable inference models from eutrophic systems. Shallow, nutrient enriched systems may experience dramatic fluctuations in the level of TP in the water column. Internal loading of TP from sediments can cause inconsistent epilimnetic nutrient concentrations (Bennion et al., 2005). An inference model based on shallow eutrophic lakes would benefit from

repeated and seasonal measurements to ensure development of accurate nutrient optima (Gibson et al., 1996). Lakes in this project were sampled once during autumn, when mixing can quickly resuspend phosphorus-rich sediments. The TP measured during sampling could have been higher than is typical, and optima developed from the single water chemistry measurement may be larger than annual measurements would otherwise yield.

Another factor that may complicate model development is the dominance of generalist species with wide ecological tolerances. These species may convey ambiguous or broad ranging indicator values. Additionally, the inclusion of particular benthic taxa found to be unreliable indicators of nutrient conditions may reduce model reconstructive capability (Bennion et al., 2005). *Staurosira*, *Staurosirella* and *Pseudostaurosira* species are problematic in development of nutrient inference models because of wide ecological tolerances and their response to physical (habitat) characteristics rather than water column nutrient levels (Bradshaw and Anderson, 2001). These three genera comprised a significant portion of the abundant diatoms used in my model development (20% of all species enumerated). Due to the low number of abundant taxa, these species were included to insure a model applicable to systems in the ecoregions studied. When removed, the model suffered from a low number of species employed for reconstruction (Bennion et al., 2001). If more lakes with lower nutrient concentrations and less problematic taxa were sampled, a TP inference model more comparable to those published may have been constructed.

Increasing the number of lakes in a calibration study may substantially reduce errors in TP inference models (Reavie et al., 2002; Ginn et al., 2007), however increasing

the number of sediment samples does not improve species data (Anderson, 1998), although seasonal plankton measurements could contribute to a more comprehensive understanding of fluctuations in community parameters (Köster and Pienitz, 2006). This current study is the only existing diatom calibration set for lakes in NE Ohio and NW Pennsylvania, but should be used with consideration of the above potential sources of error and suggestions for improvements.

### **3.8 Conclusion**

The relationship between abundant diatom species and environmental variables in 30 lakes from NE Ohio and NW Pennsylvania was explored in order to develop a diatom-based calibration data set. Total phosphorus, ammonia, and magnesium were identified as significant environmental variables influencing the distribution of diatom species in this study. Dominant diatom types (planktonic or periphytic) are influenced by lake depth and heavy metal contamination and maybe useful in classifying impaired systems. The primary goal of this research was to construct a total phosphorus diatom-inference model from lakes in Ohio and Pennsylvania. This calibration set will be used to reconstruct TP values in a paleoecological study of Bass Lake (OH) (Znidarsic, unpublished data). Inference models were examined for TP, NH<sub>3</sub>, Mg, Z<sub>max</sub>, Ni, and Na. The total phosphorus inference model was the most reliable of the six models analyzed, with an error of prediction = 17 ug/L. This value is tolerable given the long gradient of TP concentrations encountered, the small number of sample sites, low number of abundant taxa used in model development, and potential sources of error discussed in the previous section. Including more lakes could improve the reliability of optima derived

from this calibration. A number of reservoirs in Ohio and Pennsylvania could be added to the data set and would increase the number of calibration sites by a factor of four or more. Repeated seasonal water chemistry measurements could provide a more accurate depiction of annual nutrient fluctuations and may also improve this current calibration set.

A comparison of optima with other values derived from previous calibration studies revealed that the same species can have different optima depending on the region examined. When the TP gradient is long the optima is usually high, and may overestimate reconstructed TP in a low nutrient lake. When the environmental gradients are short and near-pristine lakes are sampled, TP optima are low and are not useful in regions with enriched, productive lakes. This necessitates the development of calibration sets that span wide environmental gradients, and suggests combination of data sets to provide realistic indicator values for cosmopolitan diatom species.



## LITERATURE CITED

- Anderson, N. (1998) Variability of diatom-inferred phosphorus profiles in a small lake basin and its implications for histories of lake eutrophication. *Journal of Paleolimnology* 20:47-55.
- Anderson, P. and Davic, R. (2002) Biological and water quality study of Bass Lake, 2002. OEPA Technical Report: NEDO/2005-09-01 September 8, 2005.
- Axler, R., and Henneck, J. (2002) Field protocol for determination of turbidity and chlorophyll fluorescence using the Turner Designs Aquafluor for Great Lakes coastal zone waters. Center for Water & the Environment, Natural Resources Research Institute, University of Minnesota-Duluth, MN 55811 (unpublished addition to Central Analytical Laboratory Manual: Ameal, J., Rurzycki, E., Owen, C., and Axler, R. (1998 – revised 2005). Analytical chemistry and quality assurance procedures for natural water, wastewater, and sediment samples NRRI, University of Minnesota-Duluth, Technical Report NRRI/TR-98/28. Duluth, MN).
- Battarbee, R., Cameron, N., Golding, P., Brooks, S., Switsur, R., Harkness, D., Appleby, P., Oldfield, F., Thompson, R., Monteith, D., McGovern, A. (2001) Evidence for Holocene climate variability from the sediments of a Scottish remote mountain lake. *Journal of Quaternary Science* 15:339-346.

- Bennion, H. (1994) A diatom-phosphorus transfer function for shallow, eutrophic ponds in southeast England. *Hydrobiologia* 275/276:391-410.
- Bennion, H., Juggins, S. and Anderson, J. (1996) Predicting epilimnetic phosphorus concentrations using an improved diatom-based transfer function and its application to lake eutrophication management. *Environmental Science and Technology* 30:2004-2007.
- Bennion, H., Appleby, P., and Phillips, G. (2001) Reconstructing nutrient histories in the Norfolk Broads, UK: implications for the role of diatom-total phosphorus transfer functions in shallow lake management. *Journal of Paleolimnology* 26:181-204.
- Bennion, H., Fluin, J., and Simpson, G. (2004) Assessing eutrophication and reference conditions for Scottish freshwater lochs using subfossil diatoms. *Journal of Applied Ecology* 41:124-138.
- Bennion, H., Johnes, P., Ferrier, R., Phillips, G., and Haworth, E. (2005) A comparison of diatom total phosphorus transfer functions and export coefficient models as tools for reconstructing lake nutrient histories. *Freshwater Biology* 50:1651-1670.
- Berg, T., and Speck, S. (2006) Glacial Map of Ohio. State of Ohio Department of Natural Resources. Division of Geologic Survey, Columbus, OH. Accessed 5/21/2006 online at <http://dnr.state.oh.us/portals/10/pdf/glacial.pdf>.

- Bigler, C., and Hall, R. (2002) Diatoms as indicators of climatic and limnological change in Swedish Lapland: a 100-lake calibration set and its validation for paleoecological reconstructions. *Journal of Paleolimnology* 27:97-115.
- Björck, S., Hakansson, H., Zale, R., Karlen, R., and Jonsson, B. (1991) A late Holocene lake sediment sequence from Livingston-Island, South-Shetland-Islands, With Paleoclimatic implications. *Antarctic Science* 3:61-72.
- Bradshaw, E., and Anderson, N. (2001) Validation of a diatom-phosphorus calibration set for Sweden. *Freshwater Biology* 46:1035-1048.
- Bradshaw, E., Anderson, N., Jensen, J., and Jeppesen, E. (2002) Phosphorus dynamics in Danish lakes and the implications for diatom ecology and palaeoecology. *Freshwater Biology* 47:1963-1975.
- Brazner, J., Danz, N., Niemi, G., Regal, R., Trebitz, A., Howe, R., Hanowski, J., Johnson, L., Ciborowski, J., Johnston, C., Reavie, E., Brady, V., and Sgro, G. (2007) Evaluation of geographic, geomorphic and human influences on Great Lakes wetland indicators: A multi-assembly approach. *Ecological Indicators* 7:610-635.
- Brönmark, C., and Hansson, L. (2005) *The Biology of Lakes and Ponds*. 2<sup>nd</sup> ed. Oxford University Press, New York.

Carlson, R. (1977) A trophic state index for lakes. *Limnology and Oceanography* 22:361-369.

Charles, D. and Whitehead, D. (1986) The PIRLA project: Paleoecological Investigation of Recent Lake Acidification. *Hydrobiologia* 143:13-20.

Charles, D., Whitehead, D., Anderson, D., Bienert, R., Camburn, K., Cook R., Crisman, T., Davis, R., Ford, J., Fry, B., Hites, R., Kahl, J., Kingston, J., Kreis, R., Mitchell, M., Norton, S., Roll, L., Smol, J., Sweets, P., Uutala, A., White, J., Whiting, M., and Wise, R. (1986) The PIRLA project (Paleoecological investigations of recent lake acidification) – preliminary-results for the Adirondacks, New-England, Northern-Great Lakes States, and Northern-Florida. *Water and Soil Pollution* 30:355-365.

Charles, D., Binford, M., Furlong, E., Hites, R., Mitchell, M., Norton, S., Oldfield, F., Paterson, M., Smol, J., Uutala, A., White, J., Whitehead, D., and Wise, R. (1990) Paleoecological investigation of recent lake acidification in the Adirondack Mountains, N.Y. *Journal of Paleolimnology* 3:195-241.

Charles, D., and Smol, J. (1990) The PIRLA-II Project: Regional assessment of lake acidification trends. *Verhandlungen Internationale Vereinigung für Theoretische und Angewandte Limnologie* 24:474-480.

- Cholnoky, B. (1968) Die Ökologie der Diatomeen in Binnengewässern. Cramer, 699 pp.
- Christie, C., and Smol, J. (1993) Diatom assemblages as indicators of lake trophic status in southeastern Ontario lakes. *Journal of Phycology* 29:575-586.
- Cleve, P. (1894-95) Synopsis of the naviculoid diatoms. *Kgl. Sven. Vet. Akad. Handl.*, 26:1-194, 27:1-219.
- Connell, J. (1978) Diversity in tropical rainforests and coral reefs. *Science* 199:1304-1310.
- Cooper, S., Huvane, J., Vaithyanathan, P., and Richardson, C. (1999) Calibration of diatoms along a nutrient gradient in Florida Everglades Water Conservation Area-2A, USA. *Journal of Paleolimnology* 22:413-417.
- Corkan, J. (1993) Magnitude and temporal patterns of epiphyte nitrogenase activity in East twin lake, Portage County, OH and the neutral substratum theory. Kent State University, PhD thesis.
- Cubizolle, H., Serieyssol, K., Argant, J. (2005) Diatom and pollen evidence of Holocene vegetation dynamics and human impact on the Vireennes fenland (French Massif Central). *Diatom Research* 20:257-273.

- Cumming, B., Smol, J., and Kingston, J. (1992) How much acidification has occurred in Adirondack Region Lakes (New York, USA) since Preindustrial Times? *Canadian Journal of Fisheries and Aquatic Sciences* 49:128-141.
- Davic, R., Eicher, D., and DeShon, J. (1996) Ohio Water Resource Inventory. Volume 3: Ohio's Public Lakes, Ponds, & Reservoirs. Ohio EPA technical bulletin MAS/1997-10-2. State of Ohio Environmental Protection Agency, Division of surface water, Columbus, OH.
- Davies, S., Metcalfe, S., Caballero, M., and Juggins, S. (2002) Developing diatom-based transfer functions for Central Mexican lakes. *Hydrobiologia* 467:199-213.
- Davis, R., Anderson, D., Dixit, S., Appleby, P., and Schauffler, M. (2006) Responses of two New Hampshire (USA) lakes to human impacts in recent centuries. *Journal of Paleolimnology* 35:669-697.
- DeNicola, D., de Eyto, E., Wemaere, A., and Irvine, K. (2004) Using epilithic algal communities to assess trophic status in Irish lakes. *Journal of Phycology* 40:481-495.
- Dixit, S., Smol, J., and Kingston, J. (1992) Diatoms: Powerful indicators of environmental change. *Environmental Science and Technology* 26:23-33.

- Dixit, S., Smol, J., Charles, D., Hughes, R., Paulsen, S., and Collins, G. (1998) Assessing water quality changes in the lakes of the northeastern United States using sediment diatoms. *Canadian Journal of Fisheries and Aquatic Sciences* 56:131-147.
- Enache, M., and Prairie, Y. (2002) WA-PLS diatom-based pH, TP and DOC inference models from 42 lakes in the Abitibi clay belt area (Quebec, Canada). *Journal of Paleolimnology* 27:151-171.
- Fehlmann, W. and Minder, L. (1920) Beitrag zum Problem der Sedimentbildung und Besiedlung im Zürichsee. *Festschr. Zschokke* 11.
- Findlay, D., Kling, H., Ronicke, H., and Findlay W. (1998) A paleolimnological study of eutrophied lake Arendsee (Germany). *Journal of Paleolimnology* 19:41-54.
- Fitzpatrick, M., Munawar, M., Leach, J., and Haffner, G. (2007) Factors regulating primary production and phytoplankton dynamics in western Lake Erie. *Fundamental and Applied Limnology* 169:137-152.
- Fritsch, F. (1931) Some aspects of the ecology of fresh-water algae. *Journal of Ecology* 19:233-272.
- Fritz, S., Kingston, J., and Engstrom, D. (1993) Quantitative trophic reconstruction from

sedimentary diatom assemblages: a cautionary tale. *Freshwater Biology* 30:1-23.

Garcia-Rodriguez, F., Bate, G., Smailes, P., Adams, J., and Metzeltin, D. (2007)

Multivariate analysis of the dominant and sub-dominant epipelagic diatoms and water quality data from South African rivers. *Water SA* 33:653-658.

Gibson, C., Foy, R., and Bailey-Watts, A. (1996) An analysis of the total phosphorus cycle in some temperate lakes: the response to enrichment. *Freshwater Biology* 35:525-532.

Ginn, B., Cumming, B., and Smol, J. (2007) Diatom-based environmental inferences and model comparisons from 494 northeastern North American lakes. *Journal of Phycology* 43:647-661.

Grime, J. (1977) Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *American Naturalist* 111:1169-1194.

Haberyan, K., and Horn, S. (2005) Diatom paleoecology of Laguna Zonocho, Costa Rica. *Journal of Paleolimnology* 33:361-369.

Hall, R., and Smol, J. (1992) A weighted-averaging regression and calibration model for inferring total phosphorus concentration from diatoms in British Columbia



(Canada) lakes. *Freshwater Biology* 27:417-434.

Hall, R., and Smol, J. (1996) Paleolimnological assessment of long-term water-quality changes in south-central Ontario lakes affected by cottage development and acidification. *Canadian Journal of Fisheries and Aquatic Science* 53:1-17.

Heinrichs, M., Cumming, B., Laird, K., and Hart, J. (2005) Diatom-and-chironomid-inferred eutrophication of Bouchie Lake, British Columbia. *Journal of Water Quality Research* 40:418-430.

Heinsalu, A., Alliksaar, T., Leeben, A., and Noges, T. (2007) Sediment diatom assemblages and composition of pore-water dissolved organic matter reflect recent eutrophication history of Lake Peipsi (Estonia/Russia). *Hydrobiologia* 584:133-143.

Hill, M. (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology* 54:427-432.

Hustedt, F. 1937-39. Systematische und ökologische Untersuchungen über den Diatomeen-Flora von Java, Bali, Sumatra. *Arch. Hydrobiol.* 15 & 16.

Hustedt, F. (1956) *Kieselalgen (Diatomeen)*. Kosmos-Verlage Franckh, Stuttgart.

- Hustedt, F. (1957) Die Diatomeenflora des Fluss-systems der Weser im Gebiet der Hansetadt Bremen. *Abh. Naturw. Ver. Bremen.* 34: 181-440.
- Hutchinson, G. (1961) The paradox of the plankton. *American Naturalist.* 96:137-147.
- Interland, S., Kilham, S. (2001) Limiting resources and the regulation of diversity in phytoplankton communities. *Ecology* 82:1270–1282.
- Jongman, R., ter Braak, C., and Van Tongeren, O. (1995) Data analysis in community and landscape ecology. Cambridge University Press, Cambridge, UK.
- Juggins, S. (2007) C2 Version 1.5 User guide. Software for ecological and palaeoecological data analysis and visualization. Newcastle University, Newcastle upon Tyne, UK.
- Kireta, A., Reavie, E., Danz, N., Axler, R., Sgro, G., Kingston, J., Brown, T., Hollenhorst, T. (2008) Coastal Geomorphic and Lake Variability in the Laurentian Great Lakes: Implications for a Diatom-based Monitoring Tool. *Journal of Great Lakes Research* 33(S13):136-153.
- King, L., Barker, P., and Jones, R. (2000) Epilithic algal communities and their relationship to environmental variables in lakes of the English Lake District. *Freshwater Biology* 45:425-442.

- Kolkwitz, R., and Marsson, M. (1908) Ökologie der pflanzlichen Saprobien. Berichte der Deutschen Botanischen Gesellschaft. 26:505–519.
- Köster, D., and Reinhard, P. (2006) Seasonal variability and paleolimnological inferences - a case study. *Journal of Paleolimnology* 35:395-416.
- Krammer, K., and Lange-Bertalot, H. (1991, 1997, 2004) Bacillariophyceae. 1-4. Teil., in *Süßwasserflora von Mitteleuropa*. Gustav Fischer Verlag, Stuttgart, New York, NY.
- Kuhn, D., Plafkin, J., Cairns, J. Jr., and Lowe, R. (1981) Qualitative characterization of aquatic environments using diatom life-form strategies. *Transactions of the American Microscopical Society* 100:165-182.
- Laird, K., and Cumming, B. (2008) Reconstruction of Holocene lake level from diatoms, chrysophytes and organic matter in a drainage lake from the Experimental Lakes Area (northwestern Ontario, Canada). *Quaternary Research* 69:292-305.
- Leps, J., and Smilauer, P. (2003) *Multivariate analysis of ecological data using CANOCO*. Cambridge University Press, New York, NY.
- Meriläinen, J. (1967) The diatom flora and the hydrogen ion concentration of the water. *Ann. bot. fenn.* 4:51-58.

Merilainen, J., and Huttunen, P. (1983) Ecological interpretation of diatom assemblages by means of two-way indicator species analysis (Twinspan). Proceedings of the seventh international Diatom Symposium 385-391.

Merilainen, J., and Huttunen, P. (1990) Lake acidification in Finland. Philosophical Transactions of the Royal Society of London 327:423-425.

Moffett, M. (1979) Changes in microcrustacean communities of East and West twin lakes, Ohio following lake restoration. Kent State University Masters thesis.

Moser, K., MacDonald, G., and Smol, J. (1996) Applications of freshwater diatoms to Geographical research. Progress in Physical Geography 20:21-52.

Naumann, E. (1919) Nagra synpunkter angående limnoplanktons okologi med sarskild hansyn till fytoplankton. Sv. Bot. Tidskr. 13:129-163

Nygaard, G. (1956) Ancient and recent flora of diatoms and chrsophyceae in Lake Gribso. Studies on the humic acid lake Gribso. Folia limnol. scand. 8:32-94.

Ohio EPA. (2006). Manual of Ohio EPA Surveillance methods and quality assurance practices. State of Ohio Environmental Protection Agency, Division of Surface Water, Division of Environmental Services. Appendix B: Lake Sampling Procedures

- Ostrofsky, M., and Bradley, C. (2006) Reconstructing the historical trophic status of northwestern Pennsylvania lakes using GIS. *Hydrobiologia* 571:273-281.
- Patrick, R., and Reimer, C. (1966) The diatoms of the United States I. Monographs of The Academy of Natural Sciences of Philadelphia 13.
- Pennsylvania Fish and Boat Commission. (2000) Fisheries Management Biologist Field Report - Lake Leboeuf . accessed online 5/1/2008  
<http://www.fish.state.pa.us/images/fisheries/afm/2000/leboeuf.htm>  
Updated 5/5/2008.
- Pientiz, R., Smol, J., and Birks, H. (1995) Assessment of freshwater diatoms as quantitative indicators of past climate change in the Yukon and Northwestern Territories, Canada. *Journal of Paleolimnology* 13:21-49.
- R Foundation for Statistical Computing (2006) R-statistical software version 2.3.1.
- Ramstack, J., Fritz, S., Engstrom, D., and Heiskary, S. (2003) The application of a diatom-based transfer function to evaluate regional water-quality trends in Minnesota since 1970. *Journal of Paleolimnology* 29:79-94.
- Reavie, E., Hall, R., and Smol, J. (1995) An expanded weighted-averaging model for

inferring past total phosphorus concentrations from diatom assemblages in eutrophic British Columbia (Canada) lakes. *Journal of Paleolimnology* 14:49-67.

Reavie, E., and Smol, J. (2001) Diatom-environmental relationships in 64 alkaline southeastern Ontario (Canada) lakes: a diatom-based model for water quality reconstructions. *Journal of Paleolimnology* 25:25-42.

Reavie, E., Smol, J., and Dillon, P. (2002) Inferring long-term nutrient changes in southeastern Ontario lakes: comparing paleolimnological and mass-balance models. *Hydrobiologia* 481:61-74.

Reavie, E., Axler, R., Sgro, G., Danz, N., Kingston, J., Kireta, A., Brown, T., Hollenhorst T. and Ferguson, M. (2006) Diatom-based weighted-averaging transfer functions for Great Lakes coastal water quality: relationships to watershed characteristics. *Journal of Great Lakes Research* 32:321-347.

Reavie, E., Kireta, A., Kingston, J., Sgro, G., Danz, N., Axler, R., Hollenhorst, T. (2008) Comparison of simple and multimetric diatom-based indices for Great Lakes coastline disturbance. *Journal of Phycology* 44:787-802.

Reynolds, C. (2006) *Ecology of Phytoplankton*. Cambridge University Press, New York, NY.

- Round, F., Crawford, R., and Mann, D. (1990) *The Diatoms*. Cambridge University Press, New York, NY.
- Rühland, K., and Smol, J. (2002) Freshwater diatoms from the Canadian arctic treeline and development of paleolimnological inference models. *Journal of Phycology* 38:249-164.
- SAS Institute Inc., SAS software 9.1.3 (2002-2004) Help and Documentation, Cary, NC: SAS Institute Inc.
- Schindler, D, and Fee, L. (1974) Primary production in freshwater. *Proceedings of First International Congress of Ecology, Structure, Functioning and Management of Ecosystems*. 155-158.
- Schiner, G., and Gallaher, J. (1979) *Geology and groundwater resources of western Crawford County, Pennsylvania*. Water resource report 46. Commonwealth of Pennsylvania, Department of Environmental Resources. Bureau of topographic and geologic survey.
- Schmidt, R., Kamenik, C., Kaiblinger, C., and Hetzel, M. (2004) Tracking Holocene environmental changes in an alpine lake sediment core: application of regional diatom calibration, geochemistry, and pollen. *Journal of Paleolimnology* 32:177-196.

- Schönfelder, I., Gelbrecht, J., Schönfelder, J., and Steinberg, C. (2002) Relationships between littoral diatoms and their chemical environment in northeastern German lakes and rivers. *Journal of Phycology* 38:66-82.
- Sgro, G., Reavie, E., Kingston, J., Kireta, A., Ferguson, M., Danz, N., Johansen, J. (2007) Development of a habitat quality index from a diatom-based total phosphorus inference model. *Environmental Bioindicators* 2:1:15-34.
- Siver, P. (1999) Development of paleolimnological inference models for pH, total nitrogen and specific conductivity based on planktonic diatoms. *Journal of Paleolimnology* 21:45-59.
- Smol, J. 1988. Paleoclimate proxy data from freshwater arctic diatoms. *Verh. int. Ver. Limnol.* 23:837-844.
- ter Braak, C. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67:1167-1179.
- ter Braak, C., and Smilauer, P. (2002) CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination version 4.5. Microcomputer Power, Ithaca, NY.
- ter Braak, C. (2005) CANOCO for Windows Software for Canonical Community



Ordination version 4.54. Microcomputer Power, Ithaca, NY.

Tilman, D., Kilham, S., and Kilham, P. (1982) Phytoplankton community ecology: The role of limiting nutrients. *Annual Review Ecological Systems* 13:349-372.

Tilman, D. (1977) Resource competition between planktonic algae: an experimental and theoretical approach. *Ecology* 58:338-348.

United States Environmental Protection Agency. (2008) Great Lakes Monitoring, Limnology Program. Date last updated: 2/12/08. Date accessed: 8/15/08.  
URL: <http://www.epa.gov/glnpo/monitoring/limnology/index.htm>

United States Geological Survey. (2005) The National Atlas of the United States of America. Date last updated: April 29, 2008. Date accessed: 8/15/08. URL: <http://www.nationalatlas.gov/>

Van Dorn, W. (1956) Large-volume water samplers. *Transactions of the American Geophysical Union* 37:682-684.

Vasari, Y., Kiznetsov, O., Lavrova, N., Shelekhova, T., and Vasari, A. (2007). Ainlampi, a Late-Glacial site in the northern Karelian Republic. *Annales Botanici Fennici* 44:42-55.

- Wasylikowa, K., Witkowski, A., Walanus, A., Hutorowicz, A., Alexandrowicz, S., and Langer, J. (2006) Palaeolimnology of Lake Zeribar, Iran, and its climatic implications. *66*:477-493
- Wetzel, R., and Likens, G. (1991) *Limnological Analyses*. 2<sup>nd</sup> ed. Springer-Verlag, New York, NY.
- Williams, L. (1964) Possible relationships between plankton-diatom species numbers and water-quality estimates. *Ecology* 45:809-823.
- Woods, A., Omernik, J., Brockman, S., Gerber, T., Hosteter, W. and Azevedo, S. Ecoregions of Indiana and Ohio. Date accessed: 11/15/2006.  
url: [http://www.epa.gov/wed/pages/ecoregions/ohin\\_eco.htm](http://www.epa.gov/wed/pages/ecoregions/ohin_eco.htm).  
Updated 10/28/2008
- YSI Environmental Incorporated. (1999) YSI Model 63 Handheld pH, conductivity, salinity and temperature system Operations Manual.
- YSI Environmental Incorporated. (2002) YSI Model 550A Dissolved oxygen instrument Operations Manual.
- Zhang E., Jones, R., Bedford, A., Langdon, P., and Tang, H. (2007) A chironomid-based salinity inference model from lakes on the Tibetan Plateau. *Journal of*

Paleolimnology 38:477-491.

## **APPENDICES**

## APPENDIX 1

### WATER COLUMN DATA FOR THIRTY SAMPLE LAKES

<b>NYFIN</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Findley</b>	S	7.9	5.8	10.17	82.8
<b>11/11/06</b>	1m	7.93	5.9	10.5	80.7
	2m	7.99	6	10.09	80
	3m	8	6	10.09	80
	4m	7.96	5.9	8.51	76.9
	7m	7.99	5.5	10.12	81.4
	Average	8.0	5.9	9.9	80.3

<b>OHAQU</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Aquilla</b>	S	7.01	7.8	8.08	68.6
<b>11/12/06</b>	1m	7.01	7.8	8.18	68.5
	2m	7	7.8	8.18	69.6
	3m	6.98	7.8	8.09	68.2
	4m	7.02	7.6	7.31	61.5
	Average	7.0	7.8	8.0	67.3

<b>OHBAS</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Bass</b>	S	7.22	8.2	8.18	69.7
<b>11/12/06</b>	1m	7.23	8.2	8.11	68.3
	1.5m	7.22	8.2	8.19	70.1
	Average	7.2	8.2	8.2	69.4

<b>OHBRA</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Brady</b>	S	6.59	8.9	8.05	70.2
<b>10/29/06</b>	1m	6.5	9	8.08	70.2
	2m	6.48	9.1	8.03	69.8
	3m	6.52	9.2	8.08	70.1
	4m	7.37	8.3	7.96	68.4
	11m	7.39	8.4	8.21	70.4
	Average	6.8	8.8	8.1	69.9

**Appendix 1. Water column data at surface (S) and various depth intervals. Celsius temperature (temp °C), dissolved oxygen (DO) in mg/L and % saturation.**

<b>OHCRY</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Crystal</b>	S	7.51	11.7	7.76	71.6
<b>10/22/06</b>	1m	7.52	11.8	7.79	71.9
	2m	7.5	11.9	7.77	72.1
	3m	7.5	11.9	7.75	72
	4m	7.66	11	6.82	69.2
	13.5m	6.99	7.2	0.36	2.5
	Average	7.4	10.9	6.4	59.9

<b>OHEAS</b>	depth	pH	temp °C	DO mg/L	DO %
<b>East</b>	S	7.19	9.5	8.45	73.5
<b>10/29/06</b>	1m	7.24	9.5	8.29	72
	2m	7.31	9.5	8.32	72.8
	3m	7.31	9.5	8.3	72.8
	4m	7.27	9.5	8.03	69.2
	11m	7.46	8.9	6.92	60
	Average	7.3	9.4	8.1	70.1

<b>OHMEY</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Meyers</b>	S	7.84	13	9.51	90.2
<b>10/15/06</b>	1m	7.88	13	9.37	89.4
	2m	7.9	12.9	9.42	89.4
	3m	7.93	12.8	9.35	88.3
	4m	7.57	12.4	8.99	83.1
	5.75m	7.97	12.5	7.78	74.4
	Average	7.8	12.8	9.1	85.8

<b>OHMUZ</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Muzzy</b>	S	7.32	11.8	6.14	57.3
<b>10/22/06</b>	1m	7.21	11.4	6.18	57.1
	2m	7.19	11.9	5.96	55.4
	3m	7.17	11.9	5.95	56.1
	4m	7.37	11.3	5.5	49
	7m	7.4	11.5	4.74	44.6
	Average	7.3	11.6	5.7	53.3

**Water column data continued.**

<b>OHNES</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Nesmith 10/9/06</b>	S	8.79	19.5	13.69	150.4
	1m	8.61	17.3	12.34	125.5
	2m	8.2	16.6	8.12	80
	3m	7.82	15.9	5.29	53.9
	4.8m	8.15	18	6.16	62.8
	Average	8.3	17.5	9.1	94.5

<b>OHNOR</b>	depth	pH	temp °C	DO mg/L	DO %
<b>North Reservoir 10/15/06</b>	S	7.83	12	10.41	97.4
	1m	7.93	12	10.52	98
	2m	7.97	11.8	10.31	96.1
	Average	7.9	11.9	10.4	97.2

<b>OHPIP</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Pippen 10/29/06</b>	S	7.7	9.5	8.26	73
	1m	7.66	9.5	8.23	72.5
	2m	7.63	9.5	8.22	72
	3m	7.6	9.5	8.17	72
	4m	7.67	8.7	7.68	67.6
	13m	7.13	6.4	0.37	3.9
	Average	7.6	8.9	6.8	60.2

<b>OHPUN</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Punderson 11/12/06</b>	S	7.79	8.3	9.13	78
	1m	7.8	8.3	8.89	75.5
	2m	7.78	8.3	8.76	77
	3m	7.76	8.3	8.74	76.5
	4m	7.82	7.8	8.86	76.1
	5m	7.8	7.5	8.19	72.6
	10m	7.68	7.1	4.06	35.7
	13m	7.38	6.7	1.4	11.7
	Average	7.7	7.8	7.3	62.9

**Water column data continued.**

<b>OHSAN</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Sandy 10/22/06</b>	S	7.29	11.7	4.05	37.5
	1m	7.28	11.7	4	37.5
	2m	7.29	11.8	3.98	37.6
	3m	7.28	11.8	3.94	36.4
	4m	6.96	11	3.9	36
	6.8m	7.47	11.1	3.79	34.8
	Average	7.3	11.5	3.9	36.6

<b>OHSANR</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Sandy Rep 10/22/06</b>	S	7.34	11.7	4.1	38.3
	1m	7.32	11.8	4.13	38.2
	2m	7.31	11.8	4.01	36.9
	3m	7.3	11.8	3.96	36.1
	4m	7.21	11.1	3.84	36
	6.8m	7.47	11.1	3.79	34.8
	Average	7.3	11.6	4.0	36.7

<b>OHSIL</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Silver 10/29/06</b>	S	8.38	9.5	9.86	85.6
	1m	8.44	9.6	9.75	85.9
	2m	8.45	9.6	9.67	85.2
	3m	8.44	9.6	9.41	87.3
	4m	8.45	9.3	9.08	78.6
	10m	8.47	8.9	9.32	80.8
	Average	8.4	9.4	9.5	83.9

<b>OHSIP</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Sippo 10/15/06</b>	S	7.31	11.1	6.86	64.1
	1m	7.35	11.2	6.92	62.4
	2m	7.39	11.3	6.88	62.4
	3m	7.18	11.1	6.80	61.7
	Average	7.31	11.17	6.86	62.6

**Water column data continued.**



<b>OHS PF</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Springfield 10/9/06</b>	S	8.44	19.4	10.59	116.3
	1m	8.54	17.9	12.54	12.7
	2m	8.42	16.7	10.3	106
	3m	8.21	15.9	8.27	83.8
	4m	7.92	18.3	8.43	91
	6.1m	8.17	17.2	6.98	74.5
	Average	8.3	17.6	9.5	80.7

<b>OHS PV</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Spring Valley 10/15/06</b>	S	8.12	11.9	10.11	92.1
	1m	8.14	11.9	9.6	89.6
	2m	8.16	11.9	9.27	84.4
	3m	8.17	11.8	9.68	90.9
	Average	8.1	11.9	9.7	89.3

<b>OHS UM</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Summit 10/9/06</b>	S	8.45	16.8	11.26	116.9
	1m	8.46	16.9	11.22	116.9
	2m	8.26	16.4	10.84	110.1
	3m	8.05	16.7	7.5	77
	4m	7.99	16.6	8.39	87.4
	6.8m	7.79	16.2	5.3	53
	Average	8.2	16.6	9.1	93.6

<b>OHS UN</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Sunny 10/30/06</b>	S	7.42	8	9.52	81.5
	1m	7.44	7.9	9.67	82
	Average	7.4	8.0	9.6	81.8

<b>OHS WES</b>	depth	pH	temp °C	DO mg/L	DO %
<b>West 10/29/06</b>	S	7.09	9.3	8.4	72.7
	1m	7.15	9.3	8.43	73.6
	2m	7.19	9.3	8.54	74.5
	3m	7.23	9.3	8.45	73.4
	4m	7.53	8.7	7.34	64.6
	11m	7.84	9	7.45	64.5
	Average	7.3	9.2	8.1	70.6

**Water column data continued.**

<b>OHWYO</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Wyoga 10/22/06</b>	S	7.21	10.9	5.74	52.1
	1m	7.22	11	5.53	49.8
	2m	7.24	10.9	5.56	50.6
	3m	7.24	10.9	5.55	50.3
	4m	7.39	10.5	5.49	48.7
	6.3m	7.42	10.3	5.45	43.8
	Average	7.3	10.8	5.6	49.2

<b>PACAN</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Canadohta 11/11/06</b>	S	7.03	6.8	9.06	74.5
	1m	7.09	6.7	9.13	74.8
	2m	7.1	6.8	9.05	74.4
	3m	7.11	6.7	9.06	74.6
	4m	7.06	7	8.69	73.8
	9m	7.22	6.9	8.1	66.7
	Average	7.1	6.8	8.8	73.1

<b>PACANR</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Canadohta Rep 11/11/06</b>	S	7.46	6.8	9.02	74
	1m	7.28	6.8	8.96	74.4
	2m	7.28	6.8	9.05	74.2
	3m	7.28	6.7	9.05	73.8
	4m	7.45	6.9	8.28	69.3
	8.9m	7.48	7.3	8.8	73.6
	Average	7.4	6.9	8.9	73.2

<b>PACON</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Conneaut 11/19/06</b>	S	7.22	8.4	9.41	80.4
	1m	7.42	8.4	9.37	80.2
	2m	7.41	8.4	9.32	79.5
	3m	7.4	8.4	9.28	79.1
	12m	7.22	8.1	8.97	76.3
	Average	7.3	8.3	9.3	79.1

**Water column data continued.**

<b>PACRY</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Crystal</b> (Mud Lake) <b>11/19/06</b>	S	7.05	7.9	9.05	76.8
	1m	7.05	7.9	9.19	77.2
	2m	7.03	7.9	9.19	76.5
	3m	7.02	7.9	9.13	77.7
	4m	7.16	7.7	9.11	76.6
	7m	7.18	7.8	8.69	73.2
	Average	7.1	7.9	9.1	76.3

<b>PACRY R</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Crystal Rep</b> (Mud Lake Rep) <b>11/19/06</b>	S	7.08	7.8	9.19	77.1
	1m	7.08	7.9	9.29	78.1
	2m	7.08	7.9	9.22	77.6
	3m	7.07	7.9	9.22	77.6
	4m	7.16	7.8	9.16	77.3
	7m	7.19	7.8	9.24	77.8
	Average	7.1	7.9	9.2	77.6

<b>PAEDI</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Edinboro</b> <b>11/11/06</b>	S	7.69	5.9	10.17	81.7
	1m	7.7	5.9	10.03	80.4
	2m	7.69	5.9	9.86	79.3
	3m	7.69	5.9	10.19	81.4
	4m	7.74	6	10.35	83.2
	7m	7.78	5.9	9.49	77.8
	Average	7.7	5.9	10.0	80.6

<b>PALEB</b>	depth	pH	temp °C	DO mg/L	DO %
<b>LeBoeuf</b> <b>11/11/06</b>	S	7.26	5.1	9.36	75.8
	1m	7.3	5.1	9.76	77.2
	2m	7.32	5.1	9.76	77.3
	3m	7.33	5	9.83	77.2
	4m	7.45	5.4	9.87	78.1
	7m	7.47	5	9.63	77.4
	Average	7.4	5.1	9.7	77.2

**Water column data continued.**

<b>PAPLE</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Pleasant 11/11/06</b>	S	7.37	7	10.01	82.4
	1m	7.41	7	9.9	81.8
	2m	7.47	7	9.83	81.1
	3m	7.46	7	9.88	80.5
	4m	7.79	6.5	10.06	81.7
	11m	7.93	6.5	9.5	76.5
	Average	7.6	6.8	9.9	80.7

<b>PASAN</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Sandy 11/19/06</b>	S	7.44	8	9.12	77.2
	1m	7.5	8	8.86	75.2
	2m	7.52	8.1	9.06	76.9
	3m	7.52	8.1	9.09	76.8
	4m	7.59	8	9.29	78.2
	11m	7.58	8.1	7.76	65.8
	Average	7.5	8.1	8.9	75.0

<b>PASPR</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Spring 11/19/06</b>	S	7.38	7.7	10.22	85.8
	1m	7.39	7.7	10.31	86.5
	2m	7.38	7.7	10.25	86.3
	3m	7.37	7.7	10.29	86.9
	Average	7.4	7.7	10.3	86.4

<b>PASUG</b>	depth	pH	temp °C	DO mg/L	DO %
<b>Sugar 11/19/06</b>	S	6.88	7.3	8.3	69.3
	1m	6.87	7.3	8.29	68.9
	2m	6.85	7.3	8.33	69.5
	3m	6.84	7.3	8.31	69
	4m	6.74	7	7.69	63.9
	Average	6.8	7.2	8.2	68.1

**Water column data continued.**

## **APPENDIX 2**

### **DESCRIPTION OF AUTOMATED ANALYSES USED FOR METAL DETECTION IN WATER SAMPLES and DETAILED METAL ANALYSES METHODS WITH LITERATURE CITED**

#### **Flame Emission Atomic Spectroscopy**

In this analysis metal atoms are subjected to energy in the form of a flame. Energy is transferred to the electrons surrounding the nucleus of the metal atoms. These electrons become excited and change position to a higher energy shell or orbital. Electrons return to the neutral or ground state and emit the energy applied by flame as light. Unique wavelengths of light are detected by a monochromator for each metal in question and the intensity of this emission is proportional to the concentration of metal in the sample (Wetzel and Likens, 1991).

#### **Atomic Absorption Spectroscopy**

This analysis is founded on the ability of metal atoms to absorb precise resonant wavelengths of radiation energy. Light from a cathode lamp with a filament composed of the element being analyzed is shone through a flame and into the reader of a spectrometer, which isolates and detects a specified wavelength. A sample is atomized in the flame, and atoms in the neutral or ground state decrease the concentration of the specific light wavelength detected by absorbing some of its energy. How much light is absorbed is proportional to the concentration of the metal atoms present in the sample

(Wetzel and Likens, 1991). This method is extremely precise because atoms will only absorb energy within a very specific (resonant) wavelength.

### **Inductively Coupled Plasma Emission Spectroscopy**

This method involves aerosolizing water samples and transport of the aerosol by a carrier gas to an inductively coupled argon plasma (Shugar and Ballinger, 1996). The plasma (a gas with ionized atoms) is formed by a magnetic field produced from a radio frequency generator that heats the gas to a temperature of between 6,000 and 10,000 °K. This extreme temperature excites the atoms in the plasma, which discharge light energy at specific frequencies. A spectrometer with an incurrent filter separates these unique wavelengths which are detected by highly-sensitive photomultiplier tubes, and combinations of metals can be analyzed from a single water sample. The quantity of light emitted is proportional to the concentration of each metal species in the sample.

### **Preliminary Treatment of Samples**

#### **Microwave-Assisted Digestion**

Collected unfiltered water samples were digested in a Teflon® microwave digestion vessel using a CEM MDS 81D Microwave (SM 3030K). The digestion vessels were preliminarily washed with a 50% trace-metal grade concentrated nitric acid solution, followed by a thorough rinse with Type I polished DI water. A 50 mL portion of the collected water sample was placed into the digestion vessel and 5 mL of concentrated nitric acid (trace metal grade, Fisher Scientific) were added via pipette. Samples (and a digestion blank prepared using Type I DI water) were microwave

digested at 100% power for 30 minutes (Microwave Application Note EW-2). The samples were then cooled to room temperature and placed in a previously acid-washed sample cup for further analysis.

### **Lead and Cadmium Analyses**

The concentrations of lead and cadmium were determined in microwave-digested samples using graphite furnace atomic absorption (GFAA) spectroscopy (SM 3113). Measurements were made using a Perkin-Elmer Analyst 100 atomic absorption spectrometer equipped with a HG-800 graphite furnace and AS-72 autosampler, using Perkin-Elmer AA WinLab Software, version 3.2.

#### **Lead Concentrations**

The following Pb standard solutions were prepared from a 1000 ppm stock solution: 3.13, 6.25, 12.5, 25, and 50 ppb. The following GFAA conditions were used: 20  $\mu$ L sample size, 283.3 nm wavelength, 0.7 nm slit, D<sub>2</sub> background correction, peak height was measured, Argon purge. Temperature Program: Segment 1: 60 to 120°C in 10 s, 60 s hold time; Segment 2: 120 to 180°C in 5 s, 5 s hold time; Segment 3: 180 to 400°C in 10s, 20 s hold time; Segment 4: 400 to 20°C in 1 s, 10 s hold time; Segment 5: 20 to 1800°C in a fast ramp, 5 second hold (read absorbance); Segment 6: 1800 to 2600°C in 1 s, 5 s hold time. Samples or standards were pipetted into a sample cup (1 mL), the cup placed in the autosampler, 20  $\mu$ L of sample/standard was introduced into the graphite furnace tube via the autosampler, the temperature program run and a plot of absorbance versus time was obtained. The peak height was measured. A calibration curve was

prepared by plotting the peak height (y-axis) versus Pb concentration (in ppb). The detection limit was 2 ppb and the linear working region was 2 to 50 ppb. Water samples were analyzed and their peak heights measured and converted to concentration values using the linear regression line of the calibration curve. The calibration curve had a square of the linear correlation coefficient ( $R^2$ ) value of 0.9975.

### **Cadmium Concentrations**

The following Cd standard solutions were prepared from a 1000 ppm stock solution: 0.63, 1.25, 2.5, and 5 ppb. The following GFAA conditions were used: 20  $\mu$ L sample size, 228 nm wavelength, 0.7 nm slit, D<sub>2</sub> background correction, peak height was measured, Argon purge. Temperature Program: Segment 1: 60 to 120°C in 10 s, 50 s hold time; Segment 2: 120 to 300°C in 1 s, 15 s hold time; Segment 3: 300 to 50°C in 1s, 15 s hold time; Segment 4: 50 to 1700°C in a fast ramp, 5 second hold (read absorbance); Segment 5: 1700 to 2600°C in 1 s, 5 s hold time. Samples or standards were pipetted into a sample cup (1 mL), the cup placed in the autosampler, 20  $\mu$ L of sample/standard was introduced into the graphite furnace tube via the autosampler, the temperature program run and a plot of absorbance versus time was obtained. The peak height was measured. A calibration curve was prepared by plotting the peak height (y-axis) versus Cd concentration (in ppb). The detection limit was 0.63 ppb and the linear working region was 0.63 to 5 ppb. Water samples were analyzed and their peak heights measured and converted to concentration values using the linear regression line of the calibration curve. The calibration curve had a square of the linear regression correlation coefficient ( $R^2$ ) value of 0.9969.



### **Aluminum, Copper, Iron, Nickel and Zinc Analyses**

The concentrations of these elements in the digested water samples were determined using inductively-coupled plasma (ICP) emission spectroscopy (SM 3120B). Measurements were made using a Perkin-Elmer Plasma 400 Emission Spectrometer running Perkin-Elmer Plasma 400 Software, Color Version 3.30.

#### **Aluminum Analyses**

The following Al standard solutions were prepared from a 1000 ppm stock solution: 25, 50, 100, 250, 500, 750, 1000, and 10000 ppb. The emission wavelength of 309.271 nm was monitored. A standard or sample solution was aspirated into the plasma and the emission of the sample was measured. A calibration curve was prepared by plotting the emission (y-axis) versus Al concentration (in ppb). The detection limit was 25 ppb and the linear working region was 25 to 10000 ppb. Water samples were analyzed and their peak heights measured and converted to concentration values using the linear regression line of the calibration curve. The calibration curve had a square of the linear regression correlation coefficient ( $R^2$ ) value of 0.9999.

#### **Copper Analyses**

The following Cu standard solutions were prepared from a 1000 ppm stock solution: 12.5, 25, 50, 100, 250, 500, 750, and 1000 ppb. The emission wavelength of 324.754 nm was monitored. A standard or sample solution was aspirated into the plasma and the emission of the sample was measured. A calibration curve was prepared by plotting the emission (y-axis) versus Cu concentration (in ppb). The detection limit was

12.5 ppb and the linear working region was 12.5 to 1000 ppb. Water samples were analyzed and their peak heights measured and converted to concentration values using the linear regression line of the calibration curve. The calibration curve had a square of the linear regression correlation coefficient ( $R^2$ ) value of 0.9983.

### **Iron Analyses**

The following Fe standard solutions were prepared from a 1000 ppm stock solution: 12.5, 25 and 50 ppb. The emission wavelength of 238.204 nm was monitored. A standard or sample solution was aspirated into the plasma and the emission of the sample was measured. A calibration curve was prepared by plotting the emission (y-axis) versus Fe concentration (in ppb). The detection limit was 12.5 ppb and the linear working region was 12.5 to 1000 ppb. Water samples were analyzed and their peak heights measured and converted to concentration values using the linear regression line of the calibration curve. The calibration curve had a square of the linear regression correlation coefficient ( $R^2$ ) value of 0.9973.

### **Nickel Analyses**

The following Ni standard solutions were prepared from a 1000 ppm stock solution: 12.5, 25, 50, 100, 250, and 500 ppb. The emission wavelength of 221.647 nm was monitored. A standard or sample solution was aspirated into the plasma and the emission of the sample was measured. A calibration curve was prepared by plotting the emission (y-axis) versus Ni concentration (in ppb). The detection limit was 12.5 ppb and the linear working region was 12.5 to 500 ppb. Water samples were analyzed and their

peak heights measured and converted to concentration values using the linear regression line of the calibration curve. The calibration curve had a square of the linear regression correlation coefficient ( $R^2$ ) value of 0.9997.

### **Zinc Analyses**

The following Zn standard solutions were prepared from a 1000 ppm stock solution: 12.5, 25, 50, and 100 ppb. The emission wavelength of 213.856 nm was monitored. A standard or sample solution was aspirated into the plasma and the emission of the sample was measured. A calibration curve was prepared by plotting the emission (y-axis) versus Zn concentration (in ppb). The detection limit was 12.5 ppb and the linear working region was 12.5 to 100 ppb. Water samples were analyzed and their peak heights measured and converted to concentration values using the linear regression line of the calibration curve. The calibration curve had a square of the linear regression correlation coefficient ( $R^2$ ) value of 0.9945.

### **Sodium and Potassium Analyses**

The concentrations of these elements in the digested water samples were determined using flame atomic emission spectroscopy. Measurements were made using a Perkin-Elmer Analyst 200 with an air-acetylene flame. The wavelengths of emission for sodium and potassium were 589 and 766.5 nm, respectively. The slit width for both was 0.2/0.4 nm. The following sodium standards were prepared from a 1000 ppm stock solution: 1.56, 3.13, 6.25, 12.5, 25, 50 and 100 ppm. The following potassium standards were prepared from a 100 ppm stock solution: 0.63, 1.25, 2.5, 5, and 10 ppm. To

suppress the ionization of sodium in the flame, a 100,000 ppm K solution (as KCl, 0.1 mL per 10 mL of sample) was added to each sample and standard.(5) To suppress the ionization of potassium in the flame, a 100,000 ppm Cs solution (as CsCl, 0.1 mL per 10 mL of sample) was added to each sample/standard (*Analytical Methods for Atomic Absorption Spectroscopy*,1996) Calibration curves were generated by aspirating standard solutions into the air-acetylene flame and measuring the emission of either Na or K. Due to the large ranges of sodium and potassium in the water samples, calibration curves with non-linear responses were generated to speed the analyses of the samples. Calibration curves of Na and K concentrations (ppm, y-axis) versus emission intensities (x-axis) were prepared and fit the form  $\text{Na/K Conc} = K * (\text{emission intensity})$ . The square of the correlation coefficient for sodium and potassium calibration curves were 0.9742 and 0.9638, respectively.

### **Calcium and Magnesium Analyses**

The concentrations of these elements in the digested water samples were determined using flame atomic absorption spectroscopy (SM 3111A-B). Measurements were made using a Perkin-Elmer Analyst 200 with an air-acetylene flame. The wavelengths of absorption for calcium and magnesium were 422.7 nm and 285.2 nm, respectively. The slit width for both was 0.7 nm. The following calcium standards were prepared from a 1000 ppm stock solution: 1, 2.5, 5, and 10 ppm. The following magnesium standards were prepared from a 1000 ppm stock solution: 0.1, 0.25, 0.5, and 1.0 ppm. To suppress interferences of calcium and magnesium in the flame, a 10,000 ppm Sr solution (as  $\text{Sr}(\text{NO}_3)_2$ , 1 mL per 10 mL of sample) was added to each sample and

standard (*Analytical Methods for Atomic Absorption Spectroscopy*, 1996). Water samples were typically diluted 1/20 to bring absorbances within calibration range. The absorbance of a sample/standard was measured by aspirating the solution into the air-acetylene flame. The detection limit for Ca was 1.0 ppm and its linear working range was 1 to 10 ppm. The detection limit for Mg was 0.1 ppm and its linear working range was 0.1 to 1 ppm. The square of the linear regression correlation coefficients were 0.9995 and 0.9844 for calcium and magnesium, respectively.

## LITERATURE CITED FOR METAL ANALYSES METHODS

- Analytical Methods for Atomic Absorption Spectroscopy*, Perkin-Elmer Corporation, 1996, pp. 62 (Ca), 83(K), 87 (Mg), 90(Na).
- Microwave Application Note EW-2, Revision 6-88, CEM Corporation, Matthews, NC.
- Shugar, G., and Ballinger, J. 1996. *Chemical Technicians' Ready Reference Handbook*, 4<sup>th</sup> Ed. McGraw-Hill New York, NY.
- Standard Method 3030 K. American Public Health Association, American Water Works Association, Water Environment Federation, *Standard Methods for the Examination of Water and Wastewater*, 20<sup>th</sup> Ed.; Clesceri, L.S.; Greenberg, A.E.; Eaton, A.D.; Franson, A.H, Eds. American Public Health Association: Washington, D.C., 1998.
- Standard Method 3113. American Public Health Association, American Water Works Association, Water Environment Federation, *Standard Methods for the Examination of Water and Wastewater*, 20<sup>th</sup> Ed.; Clesceri, L.S.; Greenberg, A.E.; Eaton, A.D.; Franson, A.H, Eds. American Public Health Association: Washington, D.C., 1998.
- Standard Method 3120B. American Public Health Association, American Water Works Association, Water Environment Federation, *Standard Methods for the Examination of Water and Wastewater*, 20<sup>th</sup> Ed.; Clesceri, L.S.; Greenberg, A.E.; Eaton, A.D.; Franson, A.H, Eds. American Public Health Association: Washington, D.C., 1998.
- Standard Method 3111A-B. American Public Health Association, American Water Works Association, Water Environment Federation, *Standard Methods for the Examination of Water and Wastewater*, 20<sup>th</sup> Ed.; Clesceri, L.S.; Greenberg, A.E.; Eaton, A.D.; Franson, A.H, Eds. American Public Health Association: Washington, D.C., 1998.
- Wetzel, R., and Likens, G. (1991) *Limnological Analyses*. 2<sup>nd</sup> ed. Springer-Verlag, New York, NY.

## APPENDIX 3

### SPECIES LIST

Code	Taxon
Acsp1	<i>Achnanthes</i> sp. 1
Acsp2	<i>Achnanthes</i> sp. 2
Acexveli	<i>Achnanthes exigua</i> var. <i>elliptica</i> Hustedt
Acingrat	<i>Achnanthes ingratiiformis</i> Lange-Bertalot
Aclnrbab	<i>Achnanthes lanceolata</i> ssp. <i>robusta</i> var. <i>abbreviata</i> Reimer
Aclanvby	<i>Achnanthes lanceolata</i> var. <i>boyei</i> (Oestrup) Lange-Bertalot
Acrupest	<i>Achnanthes rupestoides</i> Hohn
Actherm	<i>Achnanthes thermalis</i> (Rabenhorst) Schönfeldt
Acaffine	<i>Achnantheidium affine</i> (Grunow) Czarnecki
Accat	<i>Achnantheidium catenatum</i> (Bily & Marvan) Lange-Bertalot
Acexi	<i>Achnantheidium exiguum</i> Grunow
Acgra	<i>Achnantheidium gracillimum</i> (Meister) Lange-Bertalot
Ajac	<i>Achnantheidium jackii</i> (Rabenhorst) Lange-Bertalot
Acmn	<i>Achnantheidium minutissimum</i> (Kützing) Czarnecki
Acsap	<i>Achnantheidium saprophilum</i> (Kobayasi & Mayama) Round & Bukht.
Actnormi	<i>Actinocyclus normanii</i> (Gregory) Hustedt
Amphiped	<i>Amphipleura pellucida</i> (Kützing) Kützing
Amphrcof	<i>Amphora coffeaeformis</i> (Agardh) Kützing
Aminar	<i>Amphora inariensis</i> Krammer
Amova	<i>Amphora ovalis</i> (Kützing) Kützing
Amped	<i>Amphora pediculus</i> (Kützing) Grunow
Amsubcap	<i>Amphora subcapitata</i> (Kisselev) Hustedt
Asfor	<i>Asterionella formosa</i> Hassall
Aualp	<i>Aulacoseira alpigena</i> (Grunow) Krammer
Auamb	<i>Aulacoseira ambigua</i> (Grunow) Simonsen
Aucren	<i>Aulacoseira crenulata</i> (Ehrenberg) Thwaites
Augran	<i>Aulacoseira granulata</i> (Ehrenberg) Simonsen
Aulaev	<i>Aulacoseira laevissima</i> (Grunow) Krammer
Calschum	<i>Caloneis schumanniana</i> (Grunow) Cleve
Catgaili	<i>Catacombus gaillonii</i> (Bory) Williams & Round
Cavscute	<i>Cavinula scutelloides</i> (W. Smith) Lange-Bertalot in L.-B. & Metzeltin
Copedicl	<i>Cocconeis pediculus</i> Ehrenberg
Copla	<i>Cocconeis placentula</i> Ehrenberg
Copvlin	<i>Cocconeis placentula</i> var. <i>lineata</i> (Ehrenberg) VanHeurck
Cospusula	<i>Cosmioneis pusilla</i> (W. Smith) Mann & Stickle in Round, Crawford, & Mann.
Cracuspud	<i>Craticula cuspidata</i> (Kützing) Mann in Round, Crawford & Mann
Crahaloph	<i>Craticula halophila</i> (Grunow) Mann in Round, Crawford & Mann
Ctpulchl	<i>Ctenophora pulchella</i> (Ralfs ex Kuetzing) Williams & Round
Cystodam	<i>Cyclostephanos damasii</i> (Hustedt) Stoermer & Hakansson
Cystdu	<i>Cyclostephanos dubius</i> (Fricke) Round
Cystinvis	<i>Cyclostephanos invisitatus</i> (Hohn & Hellerman) Theroit, Stoerm. & Hakn.

### Species list

Code	Taxon
Cycstsp1	<i>Cyclostephanos sp. 1</i>
Cystthol	<i>Cyclostephanos tholiformis</i> Stoermer, Hakansson & Theroit
Cyato	<i>Cyclotella atomus</i> Hustedt
Cycom	<i>Cyclotella comensis</i> Grunow
Cyglabrs	<i>Cyclotella glabriuscula</i> (Grunow) Hakansson
Cycliris	<i>Cyclotella iris</i> Brun & Héribaud
Cymichga	<i>Cyclotella michiganiana</i> Skvortzow
Cyocce	<i>Cyclotella ocellata</i> Pantocsek
Cyplankt	<i>Cyclotella planctonica</i> Brunthaler
Cyrosii	<i>Cyclotella rossii</i> Hakansson
Cycsp1	<i>Cyclotella sp. 1</i>
Cymatsol	<i>Cymatopleura solea</i> (Brébisson & Godey) W. Smith
Cymaffin	<i>Cymbella affinis</i> Kützing
Cymasper	<i>Cymbella aspera</i> (Ehrenberg) Cleve
Cymcistla	<i>Cymbella cistula</i> (Ehrenberg) Kirchner
Cymehren	<i>Cymbella ehrenbergii</i> Kützing
Cymheter	<i>Cymbella heteropleura</i> (Ehrenberg) Kützing
Cymlaevi	<i>Cymbella laevis</i> Naegeli
Cymlance	<i>Cymbella lanceolata</i> (Ehrenberg) Kirchner
Cympsaft	<i>Cymbella peraffinis</i> Tynni
Cymproxi	<i>Cymbella proxima</i> Reimer
Cymsp1	<i>Cymbella sp. 1</i>
Cymtumid	<i>Cymbella tumidula</i> Grunow
Cymturgd	<i>Cymbella turgida</i> Gregory
Cymturgl	<i>Cymbella turgidula</i> Grunow
Cymbamph	<i>Cymbopleura amphicephala</i> (Naegeli) Krammer
Cymbsubc	<i>Cymbopleura subcuspidata</i> (Krammer) Krammer
Diameso	<i>Diatoma mesodon</i> (Ehrenberg) Kützing
Diaspc1	<i>Diatoma sp. 1</i>
Diatenu	<i>Diatoma tenuis</i> Agardh
Diavulg	<i>Diatoma vulgare</i> Bory
Diavlpro	<i>Diatoma vulgare var. producta</i> Grunow
Dipparma	<i>Diploneis parma</i> Cleve
Dipsmvp	<i>Diploneis smithii var. pumila</i> (Grunow) Hustedt
Dipsp1	<i>Diploneis sp. 1</i>
Dispseud	<i>Discostella pseudostelligera</i> (Hustedt) Houk & Klee
Disstell	<i>Discostella stelligera</i> (Cleve & Grunow) Houk & Klee
Encyalpi	<i>Encyonema alpinum</i> (Grunow) D.G. Mann
Encycaes	<i>Encyonema caespitosum</i> Kützing
Encylate	<i>Encyonema latens</i> (Krasske) D.G. Mann
Encymin	<i>Encyonema minutum</i> (Hilse) D.G. Mann
Encymuel	<i>Encyonema muellerii</i> (Hust.) D.G. Mann
Encysile	<i>Encyonema silesiacum</i> (Bleisch) D.G. Mann
Epadnata	<i>Epithemia adnata</i> (Kützing) Brébisson
Eprick	<i>Epithemia frickei</i> Krammer

**Species list continued**



Code	Taxon
Epsore	<i>Epithemia sorex</i> Kützing
Eisp1	<i>Epithemia</i> sp. 1
Epturg	<i>Epithemia turgida</i> (Ehrenberg) Kützing
Eptuvgran	<i>Epithemia turgida</i> var. <i>granulata</i> (Ehrenberg) Brun
Epturvws	<i>Epithemia turgida</i> var. <i>westermannii</i> (Ehrenberg) Grunow
Euarcus	<i>Eunotia arcus</i> Ehrenberg
Eubilun	<i>Eunotia bilunaris</i> (Ehrenberg) Mills
Euclaevi	<i>Eucocconeis laevis</i> (Oestrup) Lange-Bertalot
Euflexuo	<i>Eunotia flexuosa</i> (Brébisson) Kützing
Euform	<i>Eunotia formica</i> Ehrenberg
Euimpl	<i>Eunotia implicate</i> Nörpel & Lange-Bertalot
Euinc	<i>Eunotia incise</i> W. Smith
Euinv2PR	<i>Eunotia incise</i> var. 2 <i>PIRLA</i>
Euminor	<i>Eunotia minor</i> (Kützing) Grunow
Eumonodn	<i>Eunotia monodon</i> Ehrenberg
Eupirla	<i>Eunotia pirla</i> Carter
Eurhy	<i>Eunotia rhyncephala</i> Hustedt
Eusilvah	<i>Eunotia silvahercynia</i> Nörpel et. al
Eusps1	<i>Eunotia</i> sp. 1
Eusp2	<i>Eunotia</i> sp. 2
Eusubarc	<i>Eunotia subarcuata</i> (Naegeli) Pantocsek
Euveners	<i>Eunotia veneris</i> (Kützing) De Toni
Fragbidn	<i>Fragilaria bidens</i> Heiberg
Frcap	<i>Fragilaria capucina</i> Desmazières
Frcpvamp	<i>Fragilaria capucina</i> var. <i>amphicephala</i> (Grunow) Lange-Bertalot
Frcpvcpi	<i>Fragilaria capucina</i> var. <i>capucina</i> Desmazières
Frcpvgra	<i>Fragilaria capucina</i> var. <i>gracilis</i> (Oestrup) Hustedt
Frcpvms	<i>Fragilaria capucina</i> var. <i>mesolepta</i> Rabenhorst
Frcpvprm	<i>Fragilaria capucina</i> var. <i>perminuta</i> (Grunow) Lange-Bertalot
Fracprum	<i>Fragilaria capucina</i> var. <i>rumpens</i> (Kützing) Lange-Bertalot
Frcrot	<i>Fragilaria crotonensis</i> Kitton
Frfameli	<i>Fragilaria famelica</i> (Kützing) Lange-Bertalot
Frflitt	<i>Fragilaria famelica</i> var. <i>littoralis</i> (Germain) Lange-Bertalot
Fraheide	<i>Fragilaria heidenii</i> Oestrup
Fraincog	<i>Fragilaria incognita</i> Reichardt
Frneopr	<i>Fragilaria neoproducta</i> Lange-Bertalot
Frantz	<i>Fragilaria nitzschioides</i> Grunow
Frreiche	<i>Fragilaria reicheltii</i> (Voigt) Lange-Bertalot
Frbusta	<i>Fragilaria robusta</i> Hustedt
Frshulzi	<i>Fragilaria schulzii</i> Brockmann
Frvau	<i>Fragilaria vaucheriae</i> (Kützing) J.B. Peterson
Frzeille	<i>Fragilaria zeilleri</i> Héribaud
Frzveli	<i>Fragilaria zeilleri</i> var. <i>elliptica</i> Gasse

**Species list continued**

Code	Taxon
Frabricapi	<i>Fragilariforma bicapitata</i> (Mayer) Williams & Round
Frcnstrc	<i>Fragilariforma constricta</i> (Ehrenberg) Williams & Round
Fransim	<i>Frankophila similioides</i> Lange-Bertalot
Fruhomb	<i>Frustulia rhomboides</i> (Ehrenberg) De Toni
Geidecuss	<i>Geissleria decussis</i> (Oestrup) Lange-Bertalot, & Metzeltin in Lange-Bertalot
Geschoen	<i>Geissleria schoenfeldii</i> (Hustedt) Lange-Bertalot & Metzeltin in Lange-Bert.
Goacumin	<i>Gomphonema acuminatum</i> Ehrenberg
Goangust	<i>Gomphonema angustatum</i> (Kützing) Rabenhorst
Goangrs	<i>Gomphonema angustatum</i> f. <i>sarcophagus</i> (W. Greg.) Grunow
Goaugur	<i>Gomphonema augur</i> Ehrenberg
Gogracl	<i>Gomphonema gracile</i> Ehrenberg
Gomcrps	<i>Gomphonema micropus</i> Kützing
Gominutu	<i>Gomphonema minutum</i> (Agardh) Agardh
Gooliva	<i>Gomphoneis olivacea</i> (Hornemann) P. A. Dawson ex R. Ross & P. A. Sims
Goolimin	<i>Gomphonema olivaceum</i> var. <i>minutissimum</i> Hustedt
Goparv	<i>Gomphonema parvulum</i> (Kützing) Kützing
Goprlag	<i>Gomphonema parvulum</i> var. <i>lagenula</i> (Kützing) Frenguelli
Goproce	<i>Gomphonema procerum</i> Reichardt & Lange-Bertalot
Gopum	<i>Gomphonema pumilum</i> (Grunow) Reichardt & Lange-Bertalot
Gosp1	<i>Gomphonema</i> sp. 1
Gotrunct	<i>Gomphonema truncatum</i> Ehrenberg
Gyraccum	<i>Gyrosigma accuminatum</i> (Kützing) Rabenhorst
Gyrnodif	<i>Gyrosigma nodiferum</i> (Grunow) Reimer
Hanamphi	<i>Hantzschia amphioxys</i> (Ehrenberg) Grunow
Hantzsp1	<i>Hantzschia</i> sp. 1
Hipcapit	<i>Hippodonta capitata</i> (Ehrenberg) Lange-Bertalot, Metzeltin & Witkowski
Hiphung	<i>Hippodonta hungarica</i> (Grunow) Lange-Bertalot, Metzeltin & Witkowski
Karaclevei	<i>Karayevia clevei</i> (Grunow) Round & Bukht.
Lemhunga	<i>Lemnicola hungarica</i> (Grunow) Round & Basson
Martysp1	<i>Martyanna</i> sp.1
Martysp2	<i>Martyanna</i> sp.2
Martymar	<i>Martyanna martyi</i> (Heribaud) Round
Melvar	<i>Melosira varians</i> Agardh
Mecirc	<i>Meridion circulare</i> Agardh
Mecrcvcn	<i>Meridion circulare</i> var. <i>constrictum</i> (Ralfs) Van Heurck
Nabryoph	<i>Navicula bryophila</i> J.B. Petersen
Nacptrad	<i>Navicula capitatoradiata</i> Germain
Nacari	<i>Navicula cari</i> Ehrenberg
Nacinct	<i>Navicula cincta</i> (Ehrenberg) Ralfs
Nacncmin	<i>Navicula cincta</i> f. <i>minuta</i> Grunow
Naclemen	<i>Navicula clementis</i> Grunow
Naconsym	<i>Navicula constans</i> var. <i>symmetrica</i> Hustedt
Nacrpcph	<i>Navicula cryptocephala</i> Kützing
Nacryptl	<i>Navicula cryptonella</i> Lange-Bertalot

### Species list continued

Code	Taxon
Nadigitr	<i>Navicula digitoradiata</i> (Gregory) Ralfs
Naeidrig	<i>Navicula eidrigiana</i> Carter
Naexpect	<i>Navicula expecta</i> VanLandingham
Nahambrg	<i>Navicula hambergii</i> Hustedt
Nakotsch	<i>Navicula kotschyi</i> Grunow
Nalanc	<i>Navicula lanceolata</i> (Agardh) Kützing
Naleptos	<i>Navicula leptostriata</i> Jorgensen
Nalngcph	<i>Navicula longicephala</i> Hustedt
Namenisc	<i>Navicula menisculus</i> Schumann
Nanamibc	<i>Navicula namibica</i> Rumrich & Lange-Bertalot
Naoblong	<i>Navicula oblongella</i> Naegeli
Naphyll	<i>Navicula phyllepta</i> Kützing
Naprotra	<i>Navicula protracta</i> (Grunow) Cleve
Napuslnd	<i>Navicula pusilla</i> var. <i>lundstroemii</i> (Cleve) Lange-Bertalot
Naradsa	<i>Navicula radiosa</i> Kützing
Nardofix	<i>Navicula radiosafallax</i> Lange-Bertalot
Narecen	<i>Navicula recens</i> Lange-Bertalot
Nareichn	<i>Navicula reichardtiana</i> Lange-Bertalot
Narhncht	<i>Navicula rhynhotella</i> Lange-Bertalot
Narhynce	<i>Navicula rhynchocephala</i> Kützing
Naseibgi	<i>Navicula seibigii</i> Lange-Bertalot
Nasp1	<i>Navicula</i> sp. 1
Nasubrhy	<i>Navicula subrhynchocephala</i> Hustedt
Nasuecor	<i>Navicula suecorum</i> Carlson
Natenell	<i>Navicula tenelloides</i> Hustedt
Natrivia	<i>Navicula trivialis</i> Lange-Bertalot
Naveneta	<i>Navicula veneta</i> Kützing
Naviridu	<i>Navicula viridula</i> (Kützing) Ehrenberg
Navirgrm	<i>Navicula viridula</i> var. <i>germainii</i> (Wallace) Lange-Bertalot
Navirlin	<i>Navicula viridula</i> var. <i>linearis</i> Hustedt
Neidsp1	<i>Neidium</i> sp. 1
Niamph	<i>Nitzschia amphibia</i> Grunow
Nibacill	<i>Nitzschia bacillum</i> Hustedt
Nitzcomu	<i>Nitzschia commutata</i> Grunow
Niconstr	<i>Nitzschia constricta</i> (Gregory) Ralfs
Nidis	<i>Nitzschia dissipata</i> (Kützing) Grunow
Nidsvmed	<i>Nitzschia dissipata</i> var. <i>media</i> (Hantzsch) Grunow
Nifnticl	<i>Nitzschia fonticola</i> Grunow
Niheuflr	<i>Nitzschia heufleriana</i> Grunow
Nihungar	<i>Nitzschia hungarica</i> Grunow
Niinterm	<i>Nitzschia intermedia</i> Hantzsch
Nileibet	<i>Nitzschia liebethuthii</i> Rabenhorst
Nimacile	<i>Nitzschia macilenta</i> Gregory
Nimcrocp	<i>Nitzschia microcephala</i> Grunow

**Species list continued**

Code	Taxon
Niobtusa	<i>Nitzschia obtusa</i> W. Smith
Niobtusa	<i>Nitzschia obtusa</i> W. Smith
Nipalea	<i>Nitzschia palea</i> (Kützing) W. Smith
Nipalc	<i>Nitzschia paleacea</i> Grunow
Nipermin	<i>Nitzschia perminuta</i> (Grunow) M. Peragallo
Nitprspc	<i>Nitzschia perspicua</i> Cholnoky
Nipura	<i>Nitzschia pura</i> Hustedt
Niradicu	<i>Nitzschia radricula</i> Hustedt
Nirect	<i>Nitzschia recta</i> Hantzsch
Nitsigma	<i>Nitzschia sigma</i> (Kützing) W. Smith
Nisigmde	<i>Nitzschia sigmoidea</i> (Nitzsch) W. Smith
Nisp1	<i>Nitzschia sp. 1</i>
Nisbtils	<i>Nitzschia subtilissima</i> Cleve
Niumbn	<i>Nitzschia umbonata</i> (Ehrenberg) Lange-Bertalot
Opole	<i>Opephora olsenii</i> Möller
Opepacif	<i>Opephora pacifica</i> (Grunow) Petit
Pinalpin	<i>Pinnularia alpine</i> W. Smith
Pinborrc	<i>Pinnularia borealis</i> var. <i>rectangularis</i> Carlson
Pinlundi	<i>Pinnularia lundi</i> Hustedt
Pinmajor	<i>Pinnularia major</i> (Kützing) Rabenhorst
Pinmicst	<i>Pinnularia microstauron</i> (Ehrenberg) Cleve
Pinsp 1	<i>Pinnularia sp. 1</i>
Pinsp2	<i>Pinnularia sp. 2</i>
Pinstrep	<i>Pinnularia streptoraphe</i> Cleve
Pinvirid	<i>Pinnularia viridis</i> (Nitzsch) Ehrenberg
Plaeligin	<i>Placoneis elginensis</i> (Gregory) E.J. Cox
Plagastrm	<i>Placoneis gastrum</i> (Ehrenberg) Mersechk.
Plapsdang	<i>Placoneis pseudanglica</i> (Lange-Bertalot) E.J. Cox
Plancons	<i>Planothidium conspicuum</i> (A. Mayer) M. Aboal
Plandeli	<i>Planothidium delicatulum</i> (Kütz) Round & Bukht.
Plandist	<i>Planothidium distinctum</i> (Messik.) Lange-Bert.
Planfrq	<i>Planothidium frequentissimum</i> (Lange-Bert.) Round & Bukht.
Plangran	<i>Planothidium granum</i> (Hohn & Hellermann) Lange-Bert.
Planlan	<i>Planothidium lanceolatum</i> (Breb) Round & Bukht.
Planros	<i>Planothidium rostratum</i> (Oestrup) Round & Bukht.
Plansep	<i>Planothidium septentrionalis</i> (Oestrup) Round & Bukht.
Planstw	<i>Planothidium stewartii</i> (Patrick) Lange-Bertalot
Pleuangu	<i>Pleurosigma angulatum</i> W. Smith
Pleurlae	<i>Pleurosira laevis</i> (Ehrenberg) Compère
Psamchil	<i>Psammothidium childanos</i> (Hohn & Hellermann)
Psamgris	<i>Psammothidium grischunum</i> Bukht. & Round
Psamhelv	<i>Psammothidium helveticum</i> (Hust.) Bukht. & Round
Psamobl	<i>Psammothidium oblongella</i> (Oestrup) Van de Vijver
Psamsac	<i>Psammothidium sacculum</i> (Carter) Bukht.

### Species list continued

Code	Taxon
Psamsct	<i>Psammothidium scoticum</i> (Flower et Jones) L.Bukhtiyarova et Round
Psamsemi	<i>Psammothidium semiapertum</i> (Hustedt) M. Aboal
Psamvent	<i>Psammothidium ventralis</i> (Krasske) Bukht. & Round
Psebv	<i>Pseudostaurosira brevistriata</i> (Grunow in Van Heurck) Williams & Round
Psebrvin	<i>Pseudostaurosira brevistriata</i> var. <i>inflata</i> (Pantocsek) M.B. Edlund
Pntbodan	<i>Puncticulata bodanica</i> (Grunow in Schneider) Håkansson
Rhcurv	<i>Rhoicosphenia curvata</i> (Kützing) Grunow
Rhopagib	<i>Rhopalodia gibba</i> (Ehrenberg) O. Müller
Sepupula	<i>Sellaphora pupula</i> (Kützing) Mereschk.
Stauranc	<i>Stauroneis anceps</i> Ehrenberg
Staurphn	<i>Stauroneis phoenicenteron</i> (Nitzsch) Ehrenberg
Staursmt	<i>Stauroneis smithii</i> Grunow
Staursp 1	<i>Stauroneis</i> sp. 1
Stauravn	<i>Staurosira aventralis</i> Lange-Bertalot & Rumrich
Strcons	<i>Staurosira construens</i> Ehrenberg
Strcnsbsa	<i>Staurosira construens</i> var. <i>subsalina</i> (Hustedt) Andresen, Stoermer & Kreis
Strconbid	<i>Staurosira construens</i> var. <i>bidonis</i> (Ehrenberg) P.B. Hamilton
Strconvnt	<i>Staurosira construens</i> var. <i>venter</i> (Ehrenberg) P.B. Hamilton et. al.
Strellip	<i>Staurosira elliptica</i> (Schumann) Williams & Round
Strpsucon	<i>Staurosira pseudoconstruens</i> (Ehrenberg) Williams & Round
Staurlap	<i>Staurosirella lapponica</i> (Grunow in Van Heurck) Williams & Round
Staurlept	<i>Staurosirella leptostauron</i> (Ehrenberg) Williams & Round
Staurmin	<i>Staurosirella minuta</i> Morales & M.B. Edlund
Staurpin	<i>Staurosirella pinnata</i> (Ehrenberg) Williams & Round
Staurpvi	<i>Staurosirella pinnata</i> var. <i>intercedens</i> (Grunow) P.B. Hamilton
Strpinvla	<i>Staurosirella pinnata</i> var. <i>lancettula</i> (Schumann) E.Y. Haw & M.G. Kelly
Stenocur	<i>Stenopterobia curvula</i> (W. Smith) Krammer
Stepcymen	<i>Stephanocyclus meneghiniana</i> (Kützing) Skabitshevsky
Stbinder	<i>Stephanodiscus binderanus</i> (Kützing) Kreiger
Statmosp	cf. <i>Stephanodiscus atmosphaericus</i> (Ehrenberg) Hakansson & Locker
Sthantz	<i>Stephanodiscus hantzschii</i> Grunow
Stmedius	<i>Stephanodiscus medius</i> Hakansson
Stminu	<i>Stephanodiscus minutulus</i> (Kützing) Cleve & Möller
Stniag	<i>Stephanodiscus niagarae</i> Ehrenberg
Stpar	<i>Stephanodiscus parvus</i> Stoermer & Hakansson
Stvest	<i>Stephanodiscus vestibulus</i> Hakansson, Stoermer & Theroit
Staurprd	<i>Stauroneis producta</i> Grunow
Subohem	<i>Surirella bohemica</i> Maly
Suminu	<i>Surirella minuta</i> Brébisson in Kützing
Suovalis	<i>Surirella ovalis</i> Brébisson
Supatel	<i>Surirella patella</i> Kützing
Surobust	<i>Surirella robusta</i> Ehrenberg
Sutenera	<i>Surirella tenera</i> Gregory

**Species list continued**

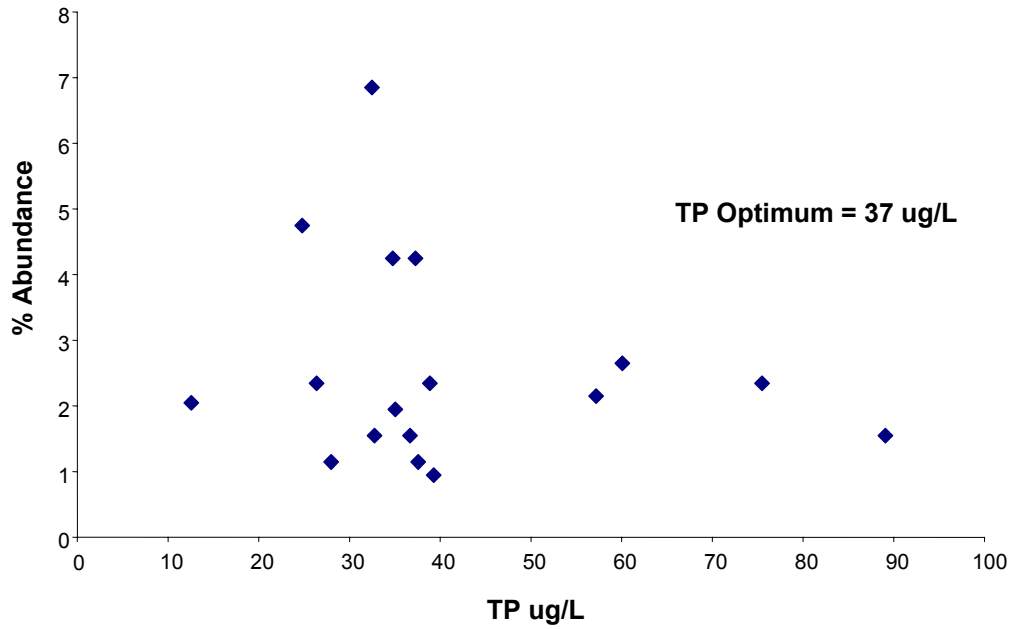
Code	Taxon
Syacus	<i>Synedra acus</i> Kützing
Syberl	<i>Synedra berlinensis</i> Lemmermann
Sydeli	<i>Synedra delicatissima</i> W. Smith
Synan	<i>Synedra nana</i> F. Meister
Syparasi	<i>Synedra parasitica</i> (W. Smith) Hustedt
Syparsub	<i>Synedra parasitica</i> var. <i>subconstricta</i> (Grunow) Hustedt
Synrump	<i>Synedra rumpens</i> Kützing
Syten	<i>Synedra tenera</i> W. Smith
Syulna	<i>Synedra ulna</i> (Nitzsch) Ehrenberg
Tabfenes	<i>Tabellaria fenestrata</i> (Lyngbye) Kützing
Tabflocc	<i>Tabellaria flocculosa</i> (Roth) Kützing
Taufasic	<i>Tabularia fasciculata</i> (Agardh) Williams & Round
Thalvisu	<i>Thalassiosira visurgis</i> Hustedt

**Species list continued**

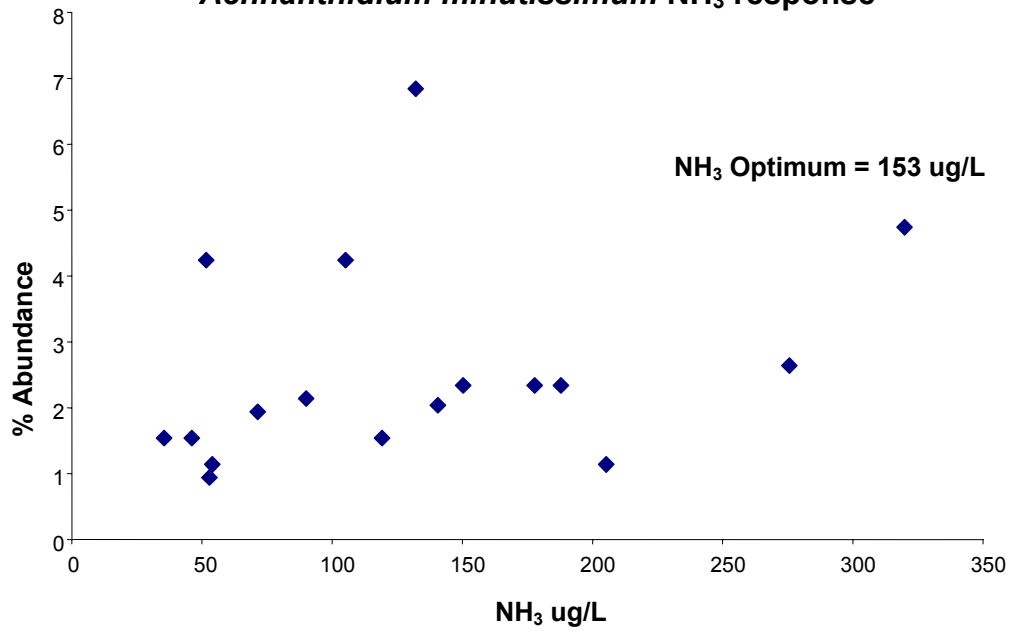
## APPENDIX 4

### RESPONSE OF SELECT SPECIES TO ENVIRONMENTAL VARIABLES

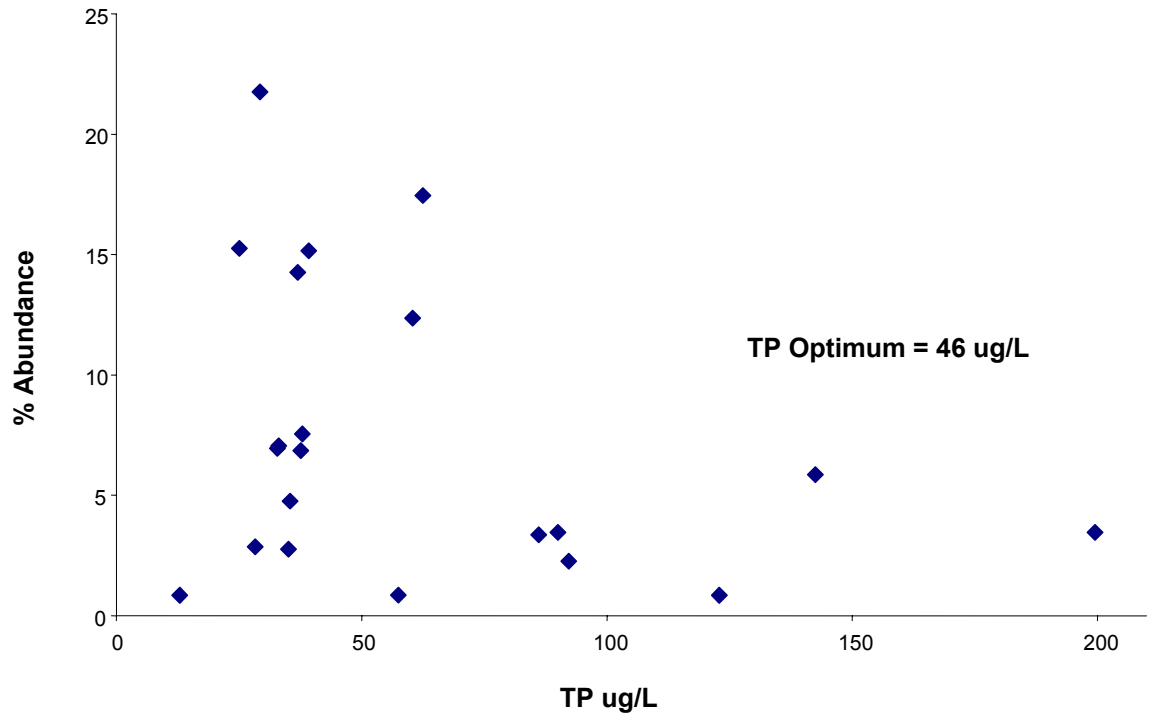
#### *Achnanthydium minutissimum* TP response



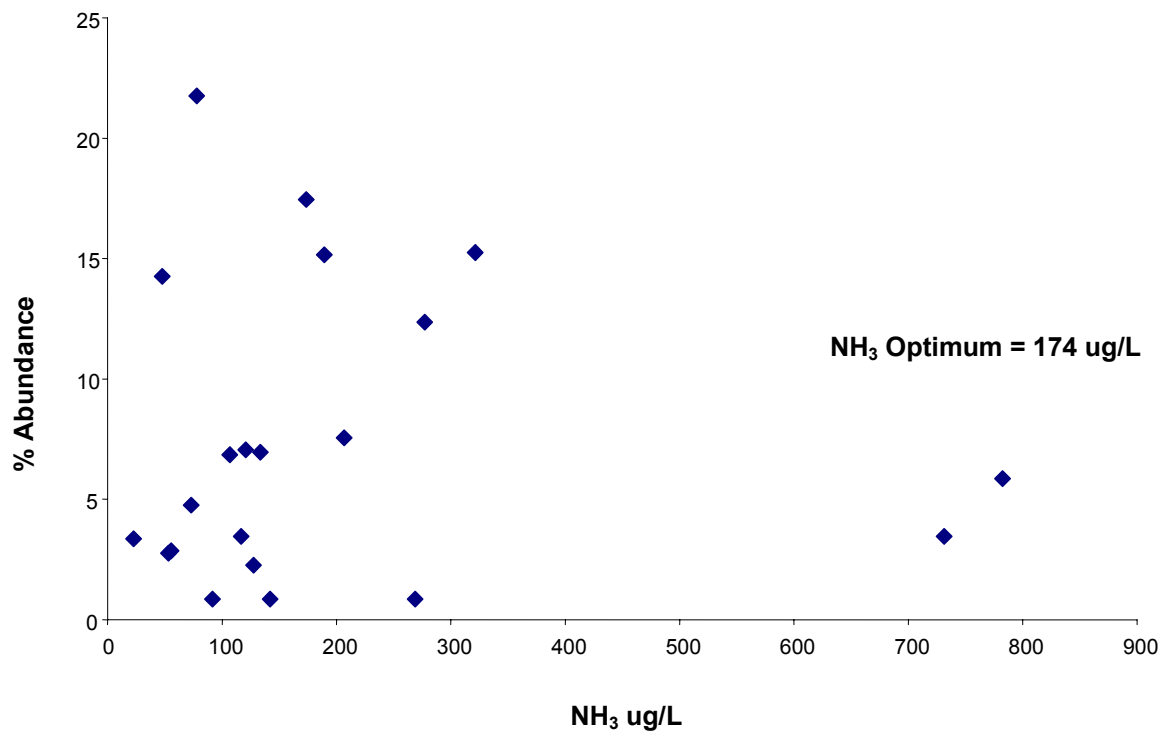
#### *Achnanthydium minutissimum* NH<sub>3</sub> response



***Asterionella formosa* TP response**

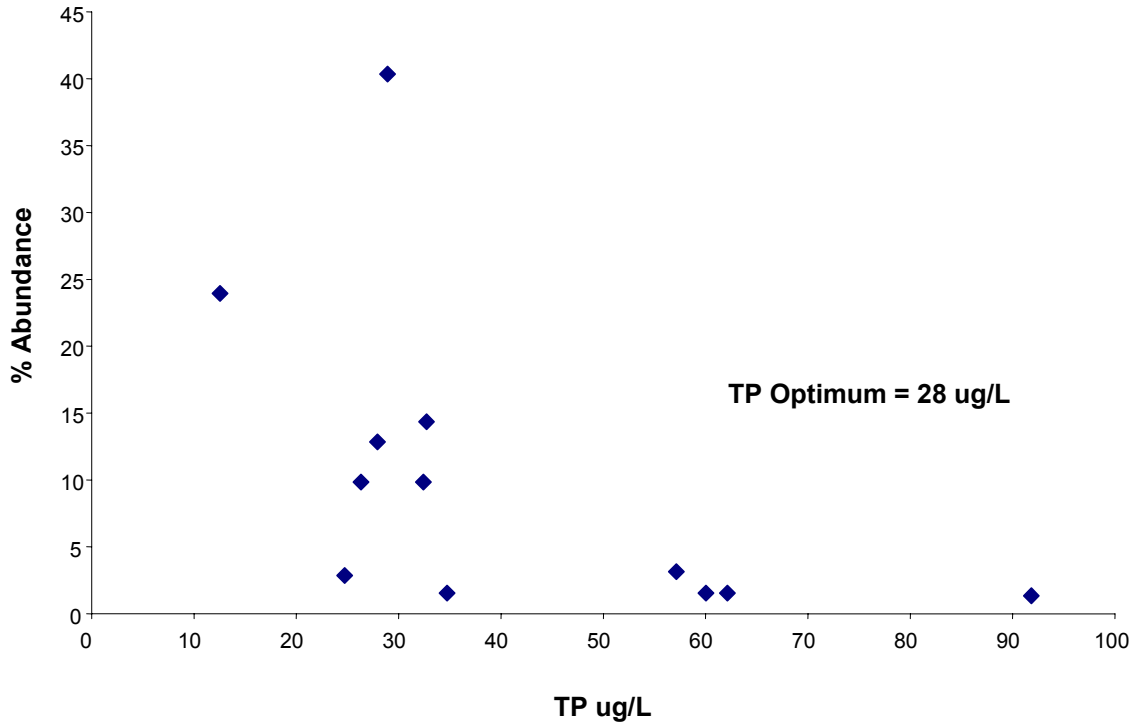


***Asterionella formosa* NH<sub>3</sub> response**

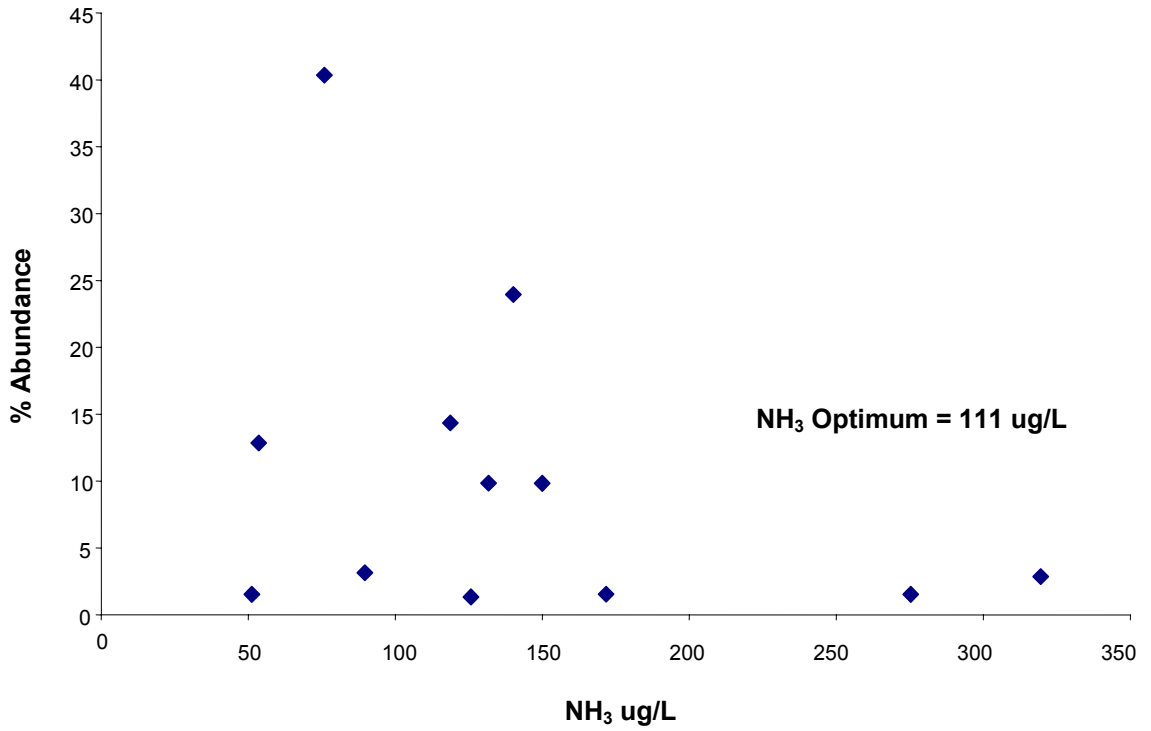




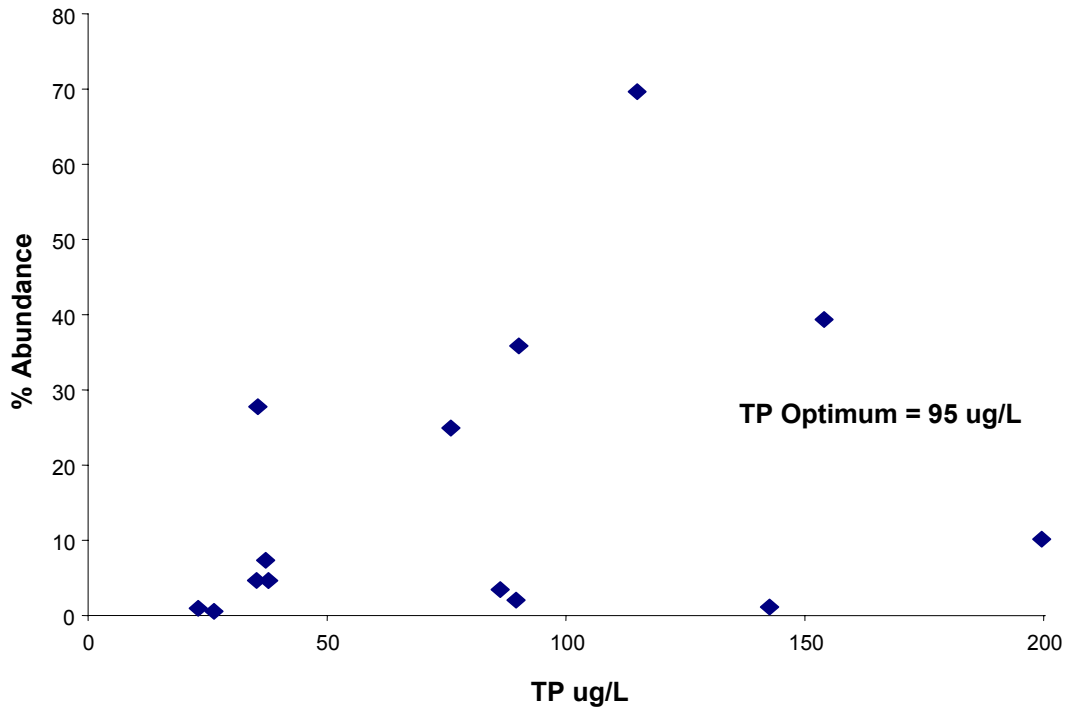
***Aulacoseira ambigua* TP response**



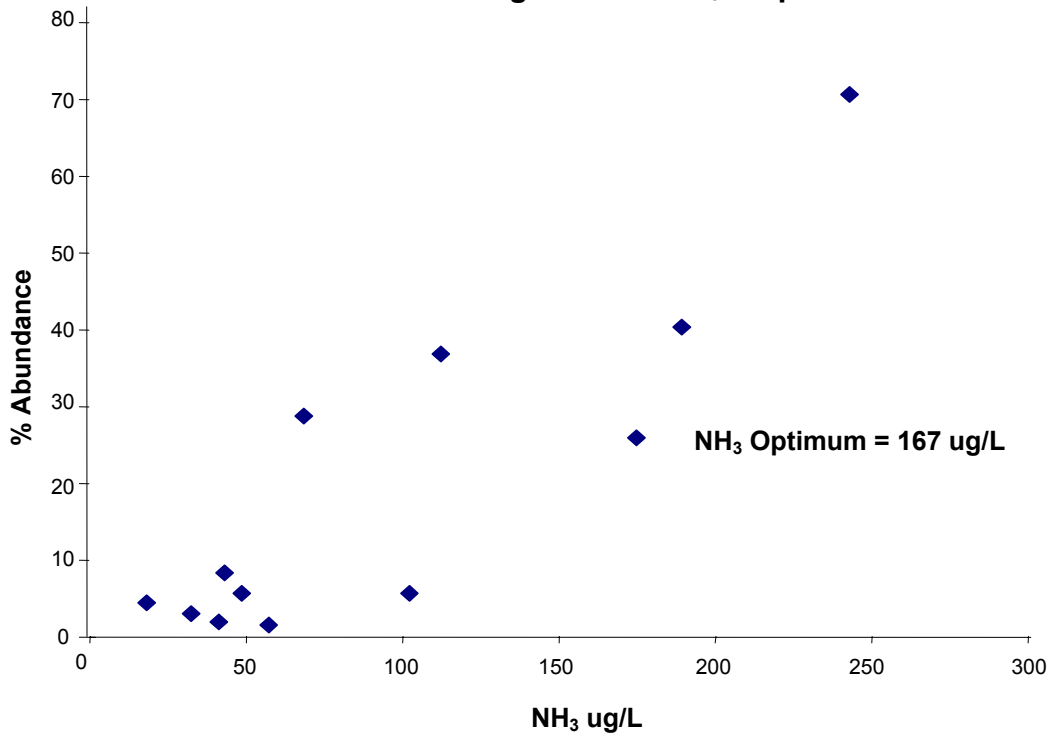
***Aulacoseira ambigua* NH<sub>3</sub> response**



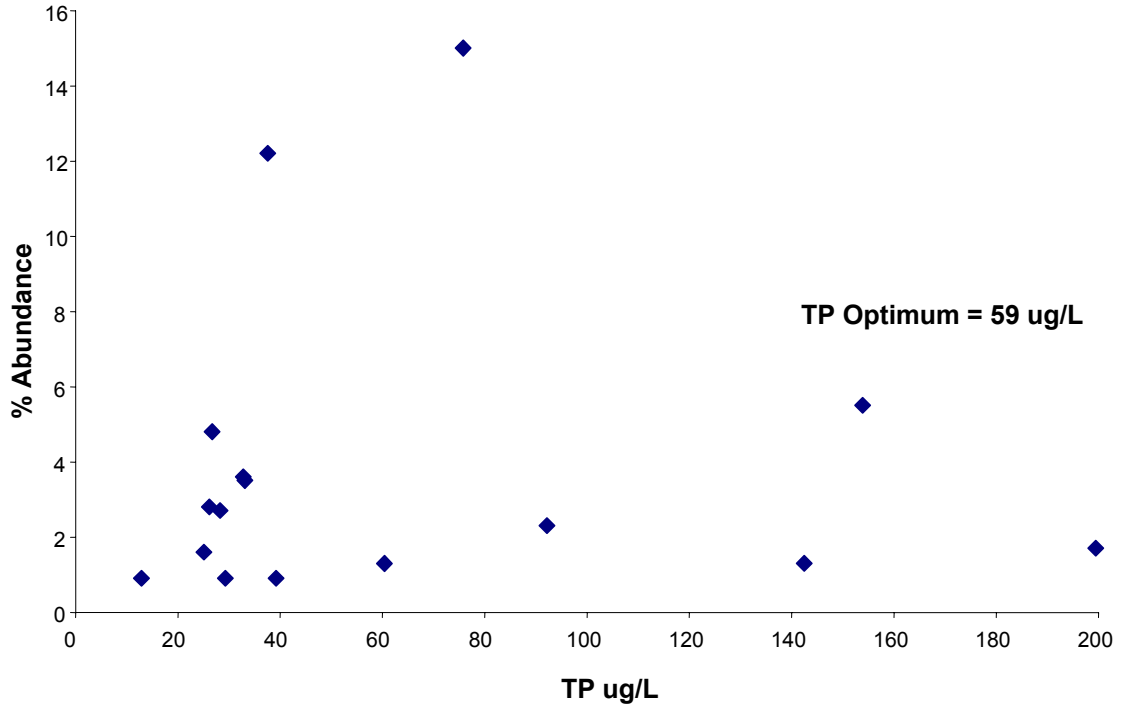
***Aulacoseira granulata* TP response**



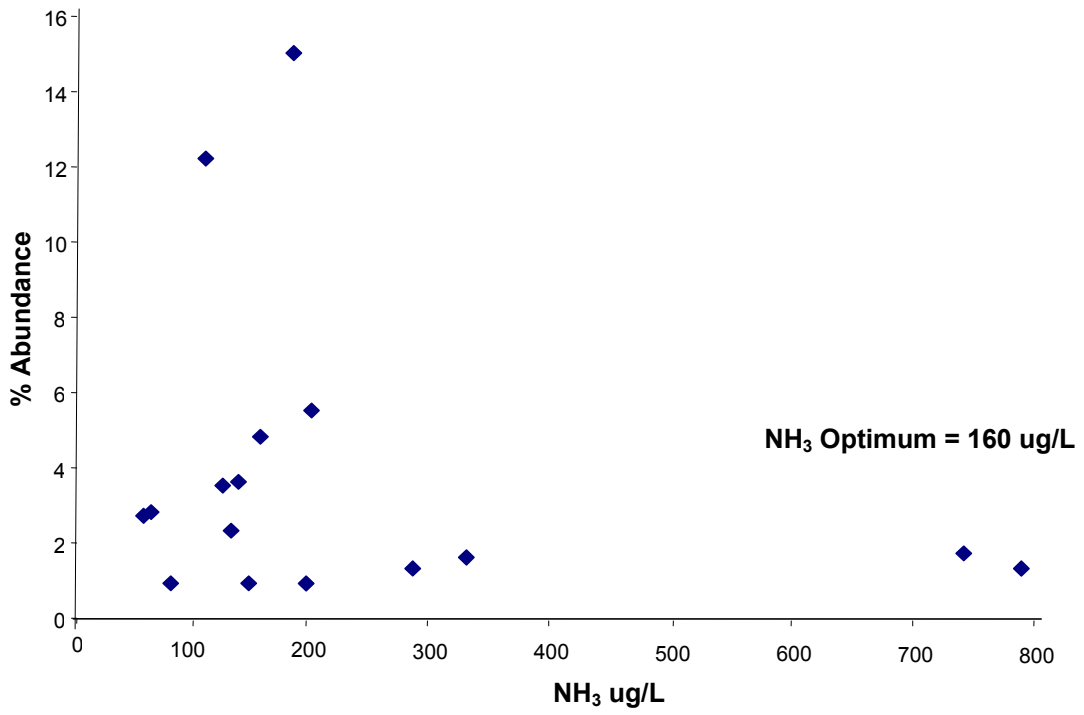
***Aulacoseira granulata*  $\text{NH}_3$  response**



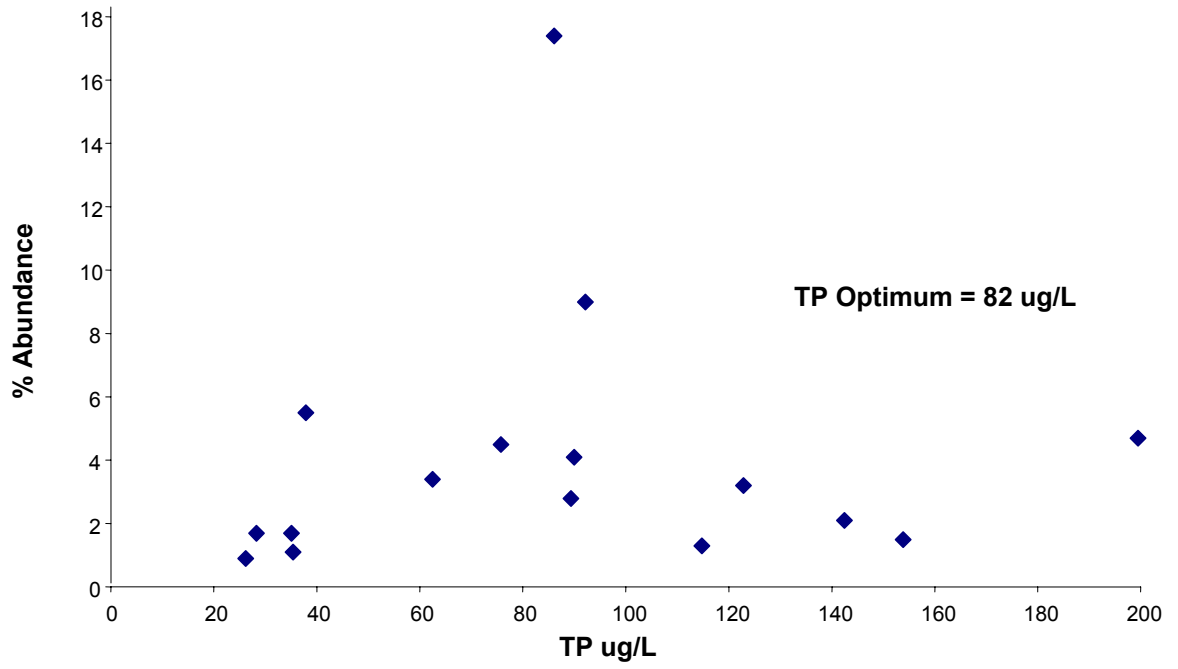
***Cocconeis placentula* TP response**



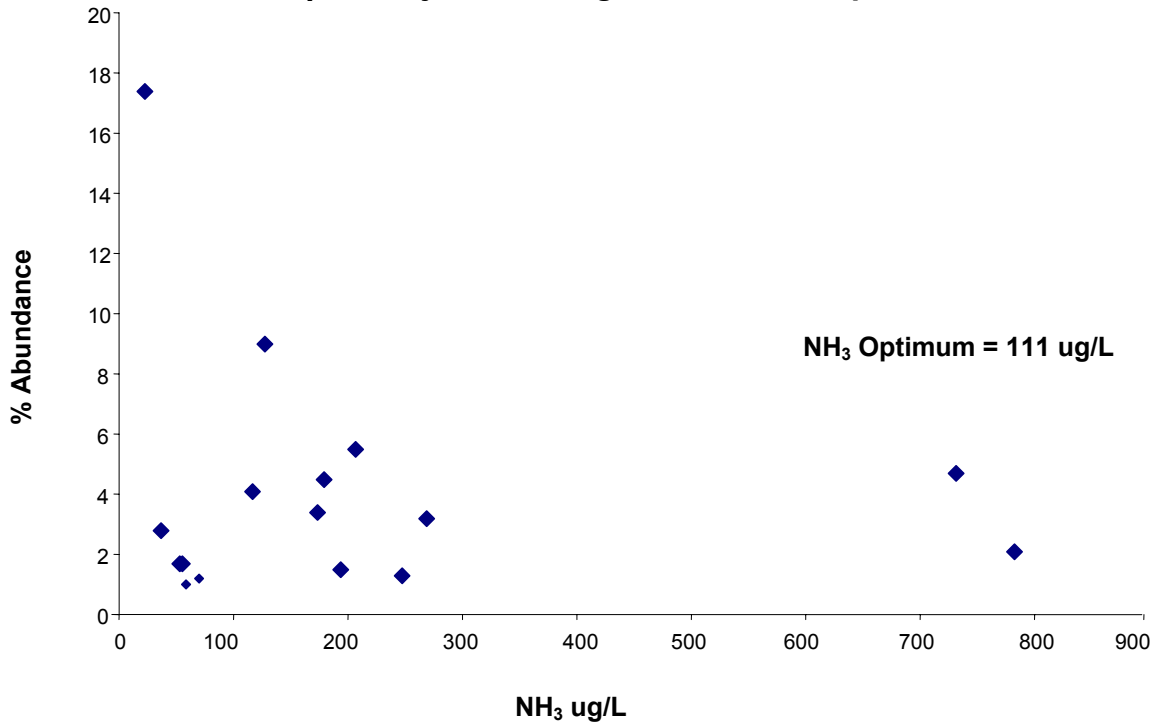
***Cocconeis placentula* NH<sub>3</sub> response**



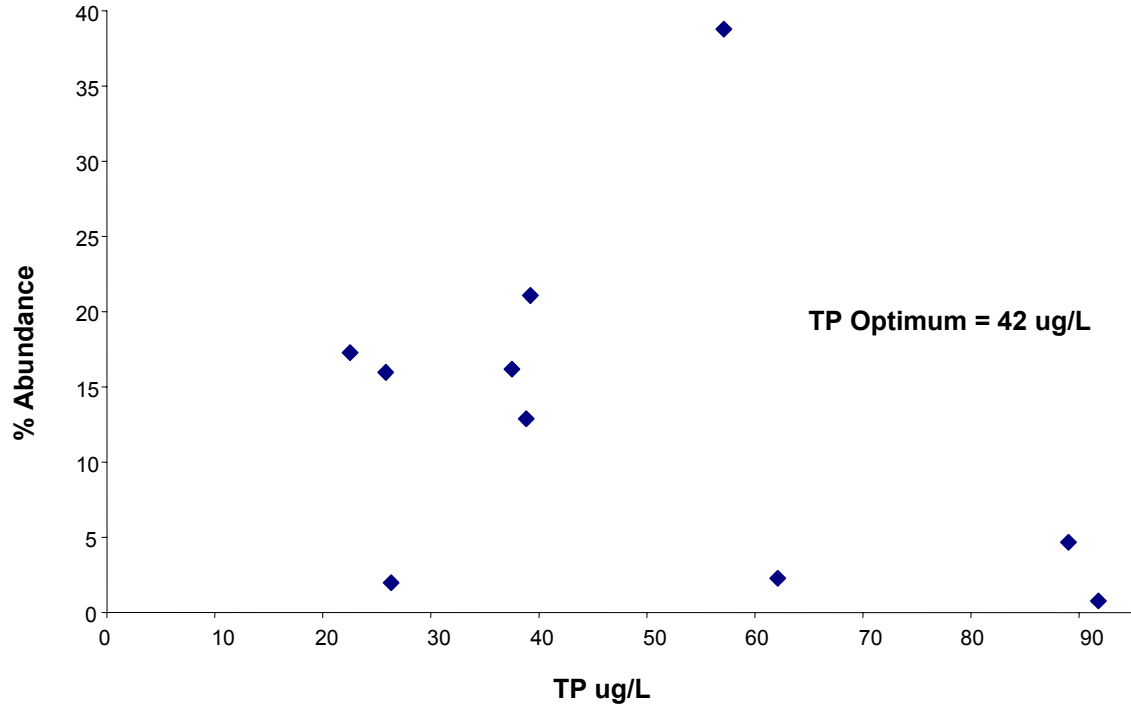
***Stephanocyclus meneghiniana* TP response**



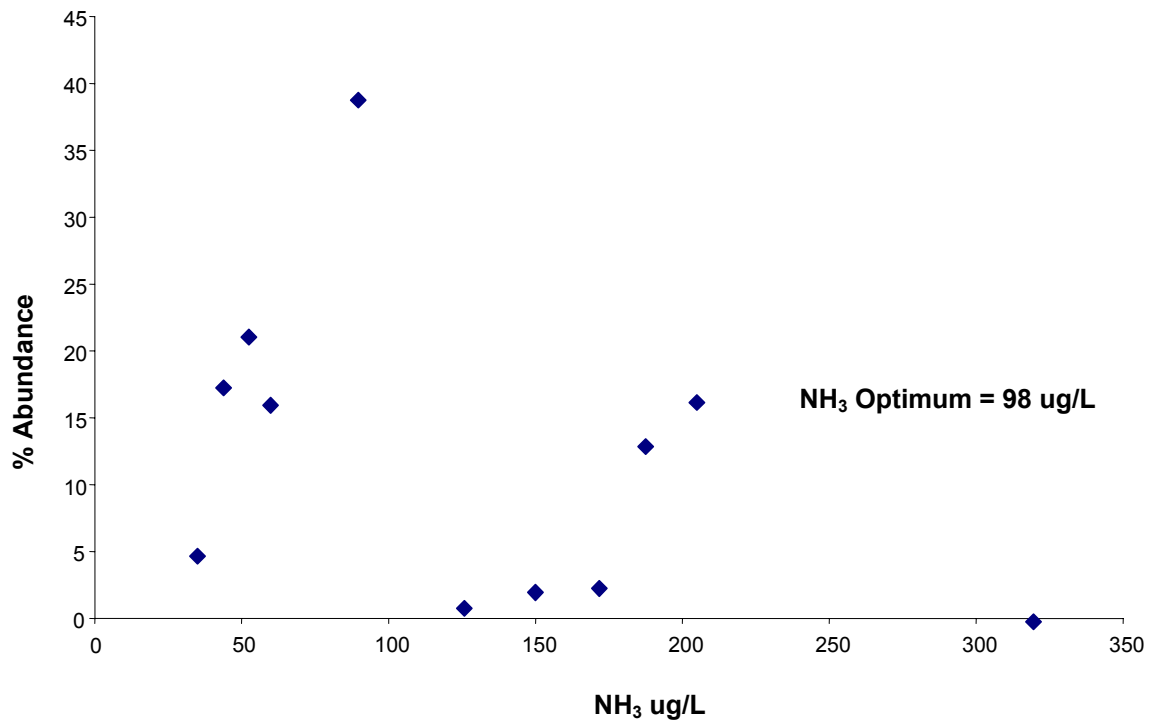
***Stephanocyclus meneghiniana* NH<sub>3</sub> response**



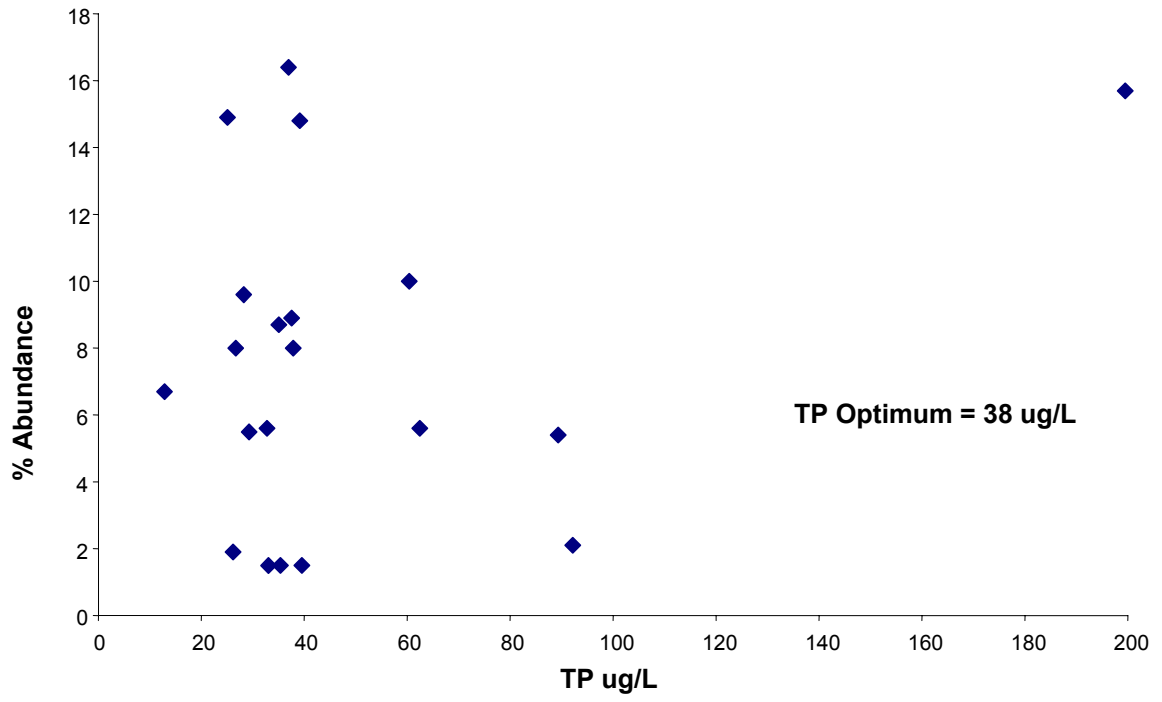
### *Cyclotella ocellata* TP response



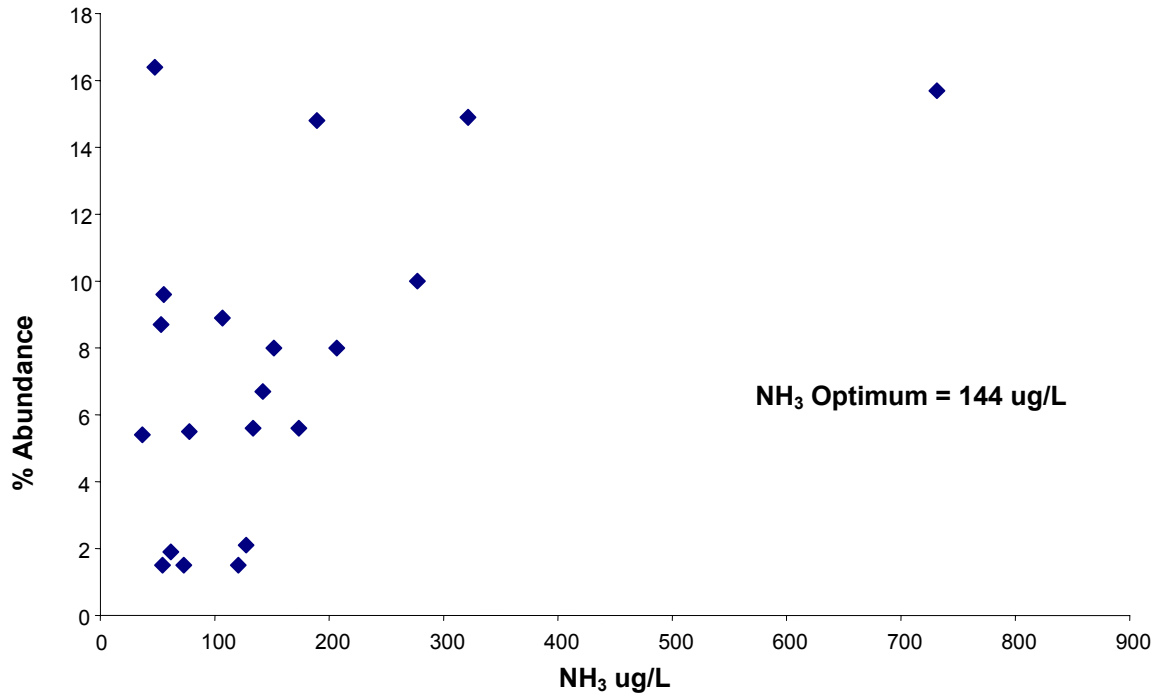
### *Cyclotella ocellata* NH<sub>3</sub> response



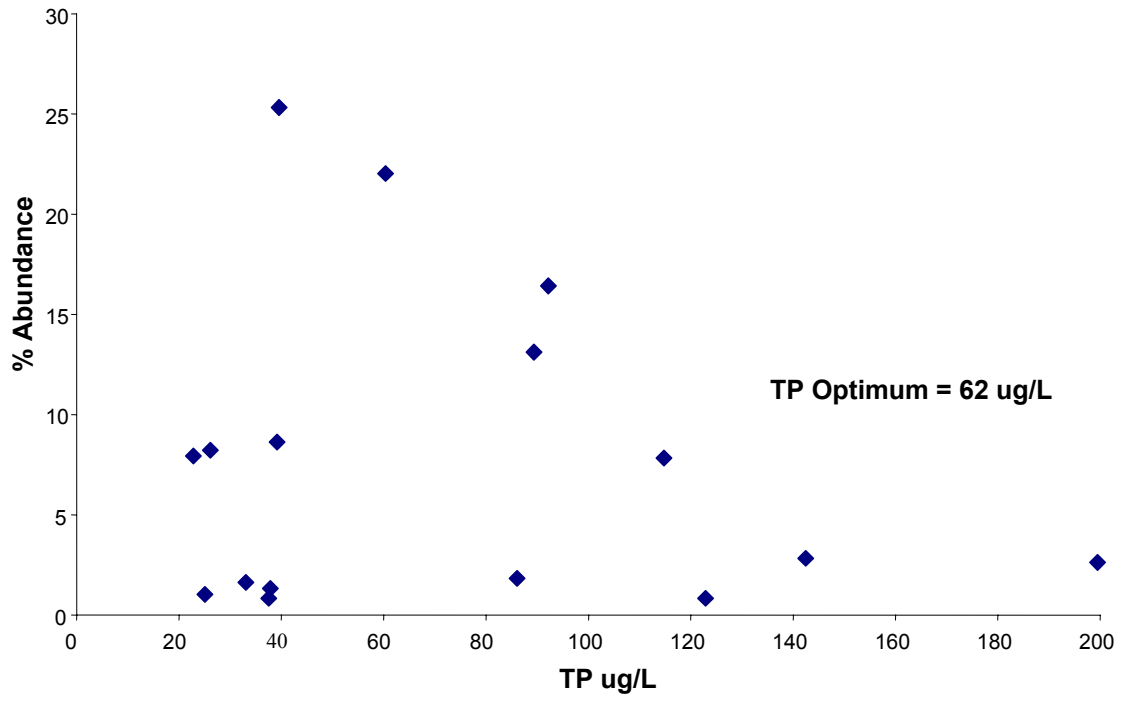
***Fragilaria crotonensis* TP response**



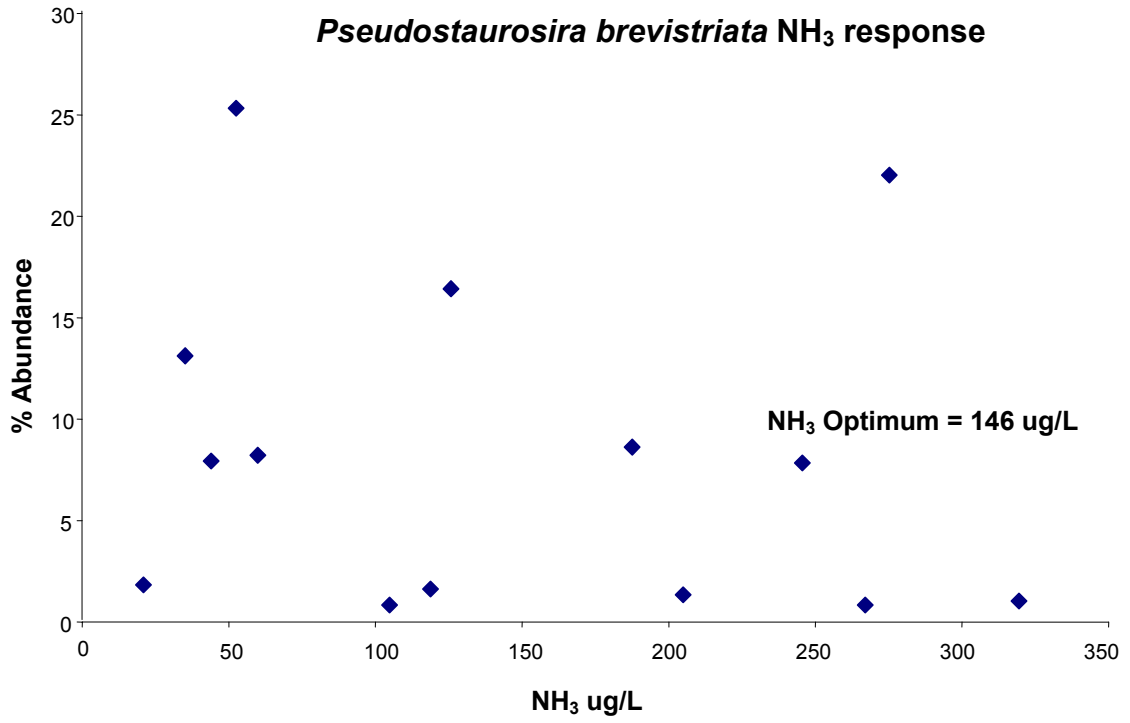
***Fragilaria crotonensis*  $\text{NH}_3$  response**



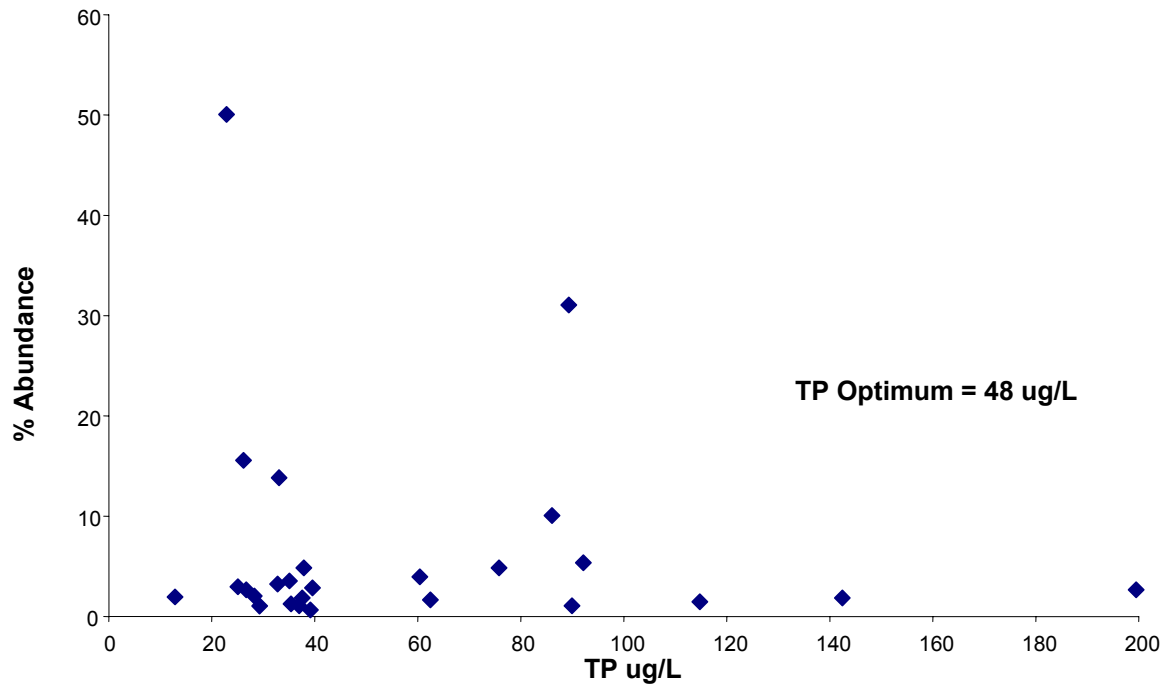
***Pseudostaurosira brevistriata* TP response**



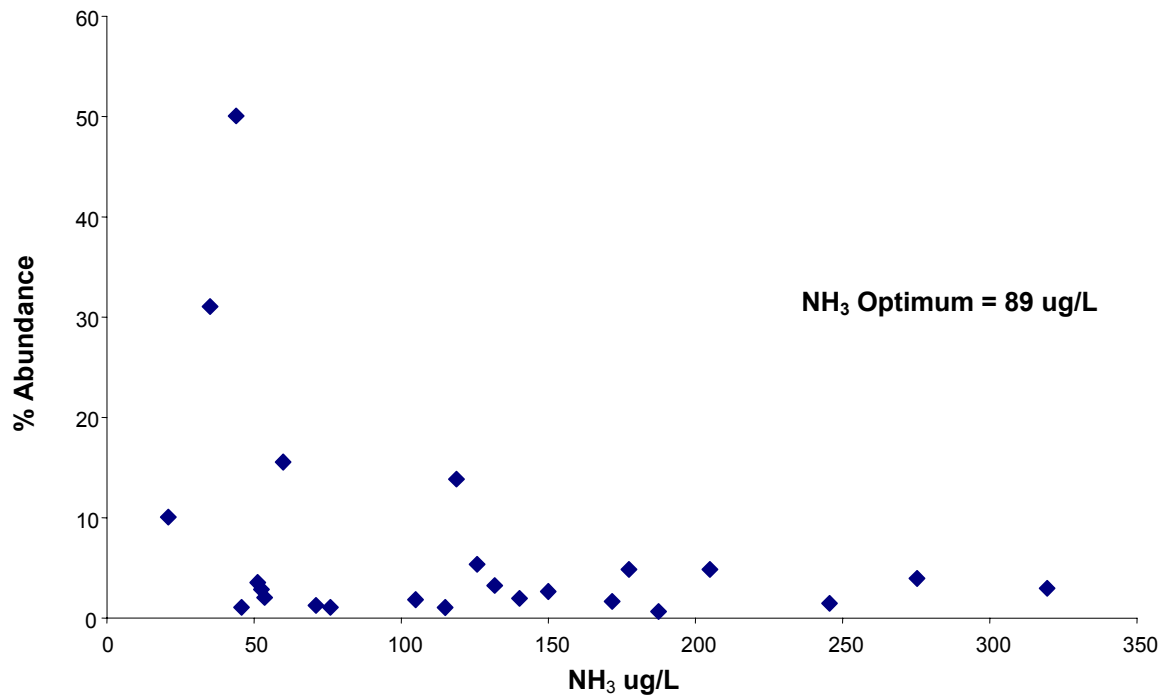
***Pseudostaurosira brevistriata* NH<sub>3</sub> response**



### *Staurosira construens* TP response

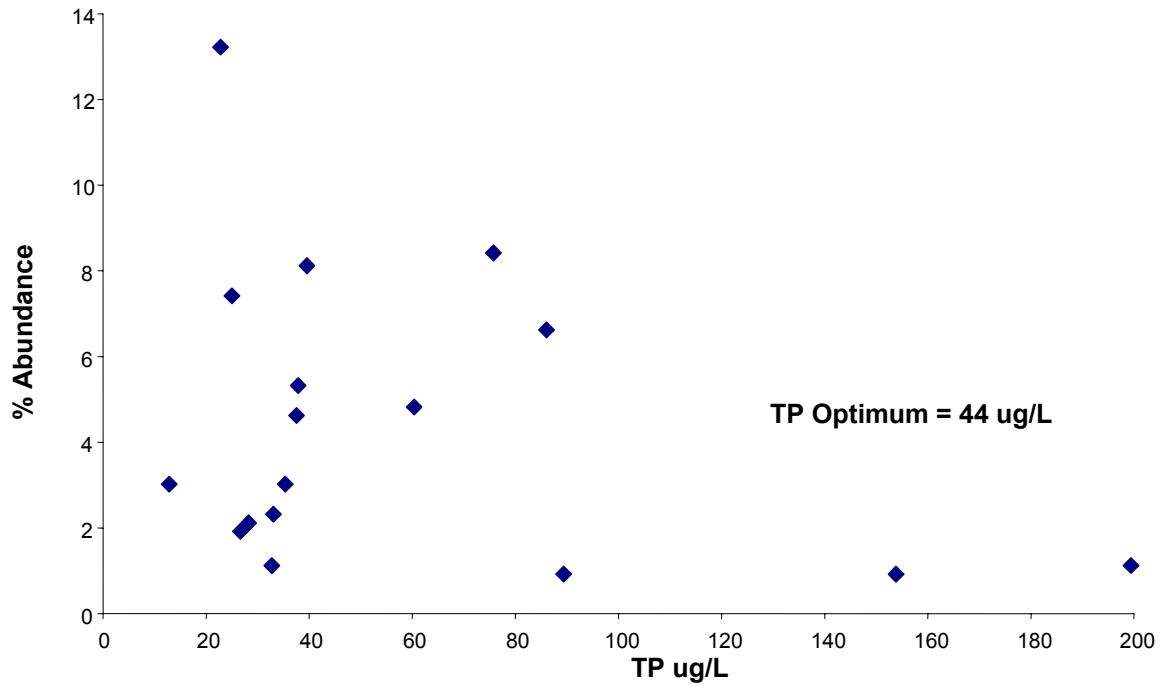


### *Staurosira construens* $\text{NH}_3$ response

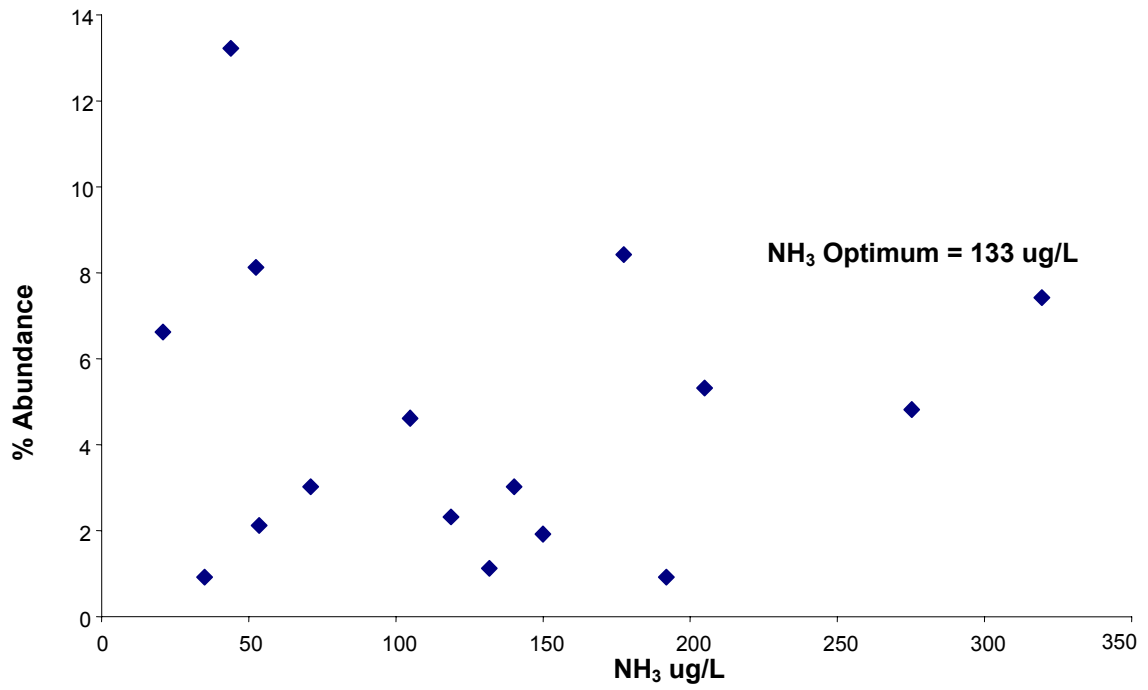




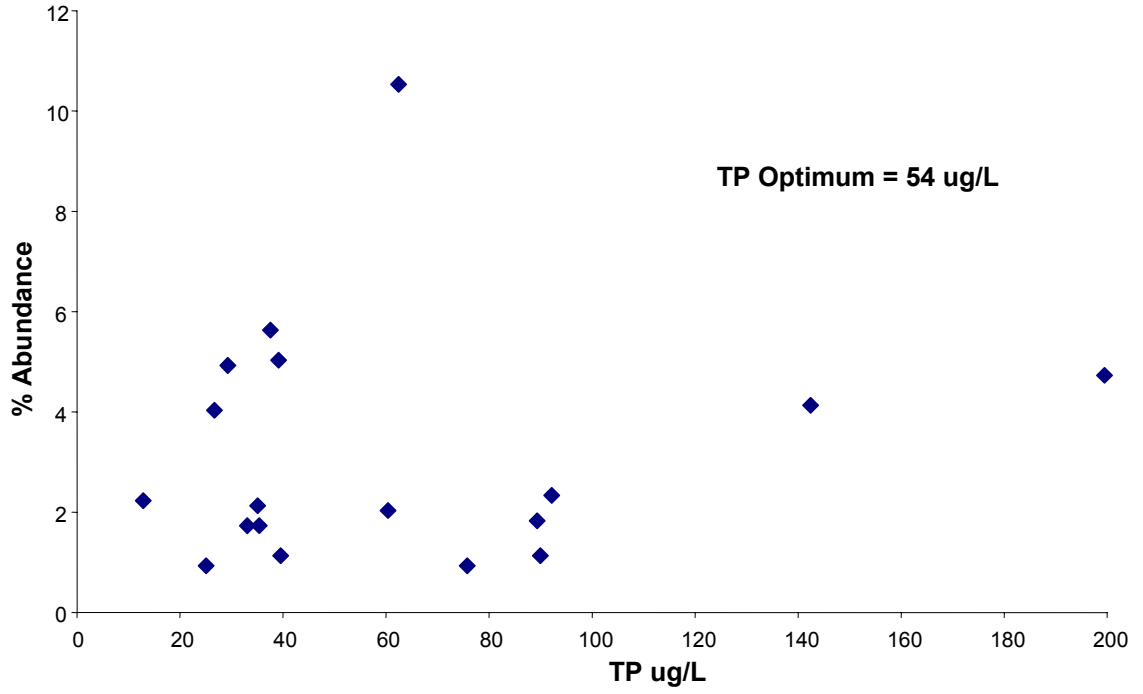
***Staurosirella pinnata* TP response**



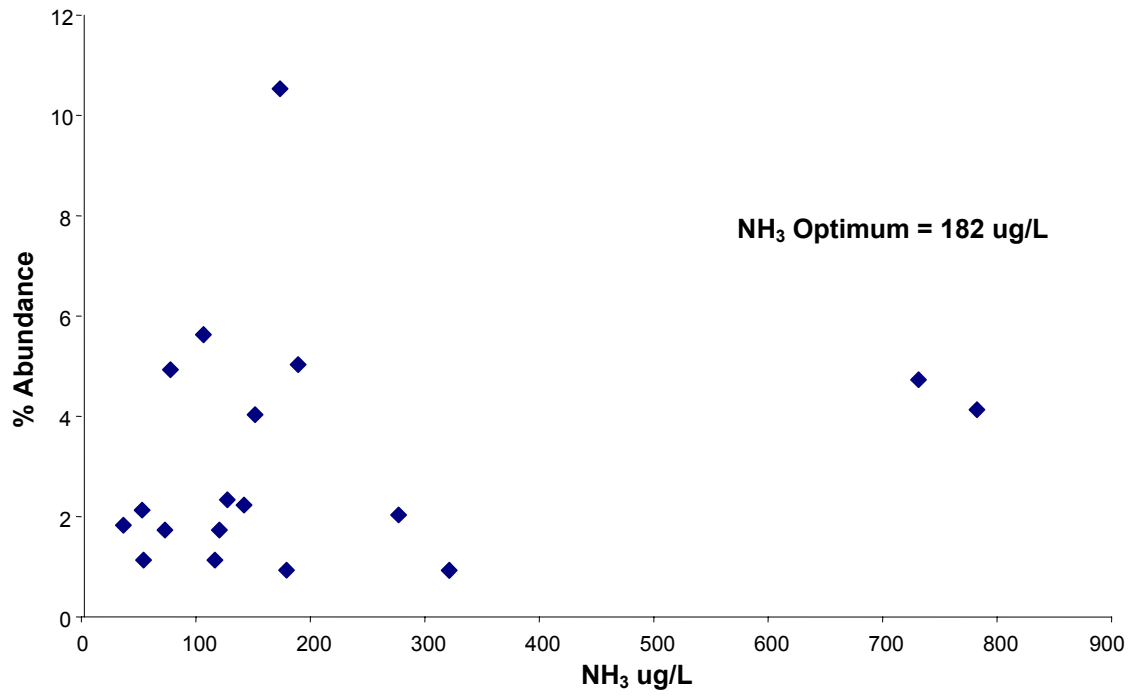
***Staurosirella pinnata*  $\text{NH}_3$  response**



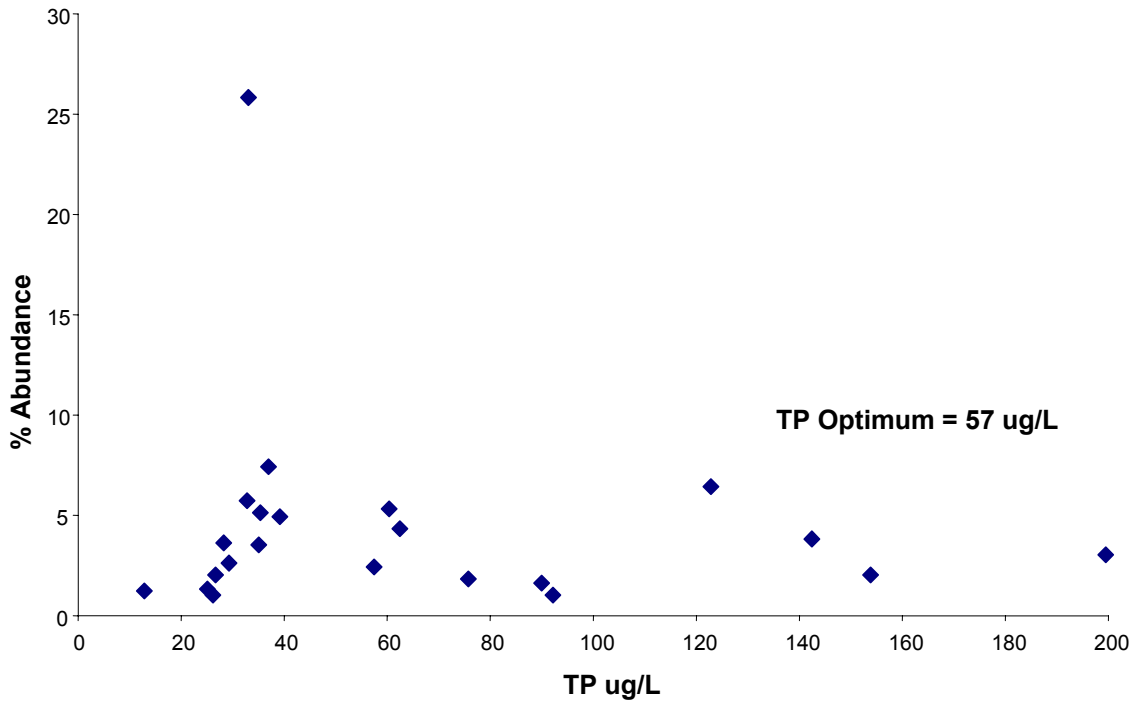
### *Synedra tenera* TP response



### *Synedra tenera* NH<sub>3</sub> response



### *Synedra ulna* TP response



### *Synedra ulna* NH<sub>3</sub> response

