Survey of the Nation's Lakes – Indiana Data Analysis

Vicky Meretsky & William W. Jones

SCHOOL OF PUBLIC AND ENVIRONMENTAL AFFAIRS

INDIANA UNIVERSITY

Bloomington, IN

Prepared for:
Water Assessment and Planning Branch
Office of Water Quality
Indiana Department of Environmental Management
Indianapolis, IN

April 2010

Survey of the Nation's Lakes – Indiana Data Analysis

Vicky Meretsky
William W. Jones

School of Public & Environmental Affairs
Indiana University
Bloomington, IN

Prepared for:

Water Assessment and Planning Branch
Office of Water Quality
Indiana Department of Environmental Management
Indianapolis, IN

April 2010

Table of Contents

Survey of the Nation's Lakes – Indiana Data Analysis

INTRODUCTION

In 2007, the U.S. Environmental Protection Agency (EPA), states, and tribes conducted a nationwide survey of the condition of the nation's lakes. The survey was designed to

a) Assess the condition of the Nation's Lakes
b) Establish a baseline to compare future surveys for trends assessment and evaluate trends since the 1970's National Eutrophication Survey Study
c) Help build State and Tribal capacity for monitoring and assessment and promote collaboration across jurisdictional boundaries (EPA 2007).

Designed to estimate the percentage of lakes that are in good, fair, or poor condition, the survey serves as a scientific report card on America's lakes. It examined ecological, water quality, and recreational indicators, and assessed how widespread key stressors (such as nitrogen, phosphorus, and acidification) are across the country.

The survey was a collaborative effort that involved dozens of state environmental and natural resource agencies, federal agencies, universities, and other organizations. In Indiana, the effort was coordinated through the Indiana Department of Environmental Management (IDEM). Staff and faculty at Indiana University's School of Public and Environmental Affairs (SPEA) conducted the field analysis and sampling.

This report provides a statistical analysis of the water chemistry data for Indiana lakes and reservoirs included in the National Lakes Assessment (NLA).

METHODS

Lake Selection

A total of 1,028 randomly-selected and 124 selected reference lakes – representing 49,546 target lakes in five size classes and distributed relatively across the lower 48 states –were included in the survey. EPA selected the lakes from the nation's natural and man-made freshwater lakes, ponds, and reservoirs (hereafter referred to as "lakes"). Both public and private lakes were selected and sampled. Lakes included in the draw had to be at least one meter deep and over ten acres in size. The survey did not include the Great Lakes or the Great Salt Lake. Lakes were selected randomly using a statistical survey design to represent the population of lakes in their ecological region, or Ecoregion – the geographic area in which climate, ecological features, and plant and animal communities are similar.

In Indiana, 21 lakes were selected in the initial draw. An additional 29 "overdraw" lakes were also sampled in Indiana by selecting the next 29 lakes from the master EPA list of eligible lakes. Permission of the landowner was required before sampling on private lakes, and several landowners denied us approval to sample. In such a situation, the next unsampled lake on the list was then elevated to the draw. The total of 50 Indiana lakes sampled was intended to be a large enough sample size to allow IDEM to statistically analyze lake conditions within Indiana. An additional lake, Olin Lake, was added as a reference lake (Figure 1).

Data Collection

SPEA students and faculty who would collect the Indiana field data attended a three-day training program in Madison, Wisconsin in May 2007, prior to the start of sampling. To insure data consistency, all field crews throughout the U.S. used the same sampling protocols. Field protocols were documented in the "Survey of the Nation's Lakes Field Operations Manual" (EPA 2007).

At each lake, the point of maximum water depth (call the Index Site) was found using available bathymetry and boat-mounted depth meters. Temperature, dissolved oxygen and pH profiles were made using a multi-parameter sonde (HydroLab Quanta HQ series). Secchi disk transparency was determined using a standard Secchi disk. Four 2-liter water samples were collected from the upper 2-meters of the water column using a 2-meter integrated sampler. A 4-liter cubitainer was filled for determination of water chemistry, a 2-liter sample was retained for phytoplankton determination, and the final 2-liters were available to filter water for chlorophyll *a* determination. The water chemistry samples were stored on ice in a cooler, the chlorophyll *a* filters were stored frozen with dry ice and the phytoplankton sample was preserved with Lugol's solution. At the end of each day samples were shipped to EPA-approved laboratories for analysis.

Additional field samples were collected for zooplankton, Microcystin toxin, sediment diatoms, sediment mercury, and benthic macroinvertebrates, however, these results were not available in time to be included in this report.

Watershed Features

An **O**nline **W**atershed De**L**ineation **S**ystem (OWLS), available as part of a suite of watershed characterization and management tools available on-line through Purdue University at: http://cobweb.ecn.purdue.edu/~watergen/, was used to delineate the watersheds and estimate watershed land use of the NLA lakes. The program uses data from a 2007 USGS 10-meter digital elevation model to delineate watersheds and uses a layer from the National Landuse Cover Dataset (2001) for land use designation. The program could not delineate watersheds for 6 of the 51 lakes due to very flat land around these lakes.

Water Sample Analysis

Data flags

Water quality data were flagged for long handling time on measurements of pH (5-8 d), conductivity (8-11 d), gran acid-neutralizing capacity (8 d), turbidity (4 d), total organic carbon (15-244 d), dissolved organic carbon (15-153 d), total phosphorus (29-91 d), total nitrogen (29-

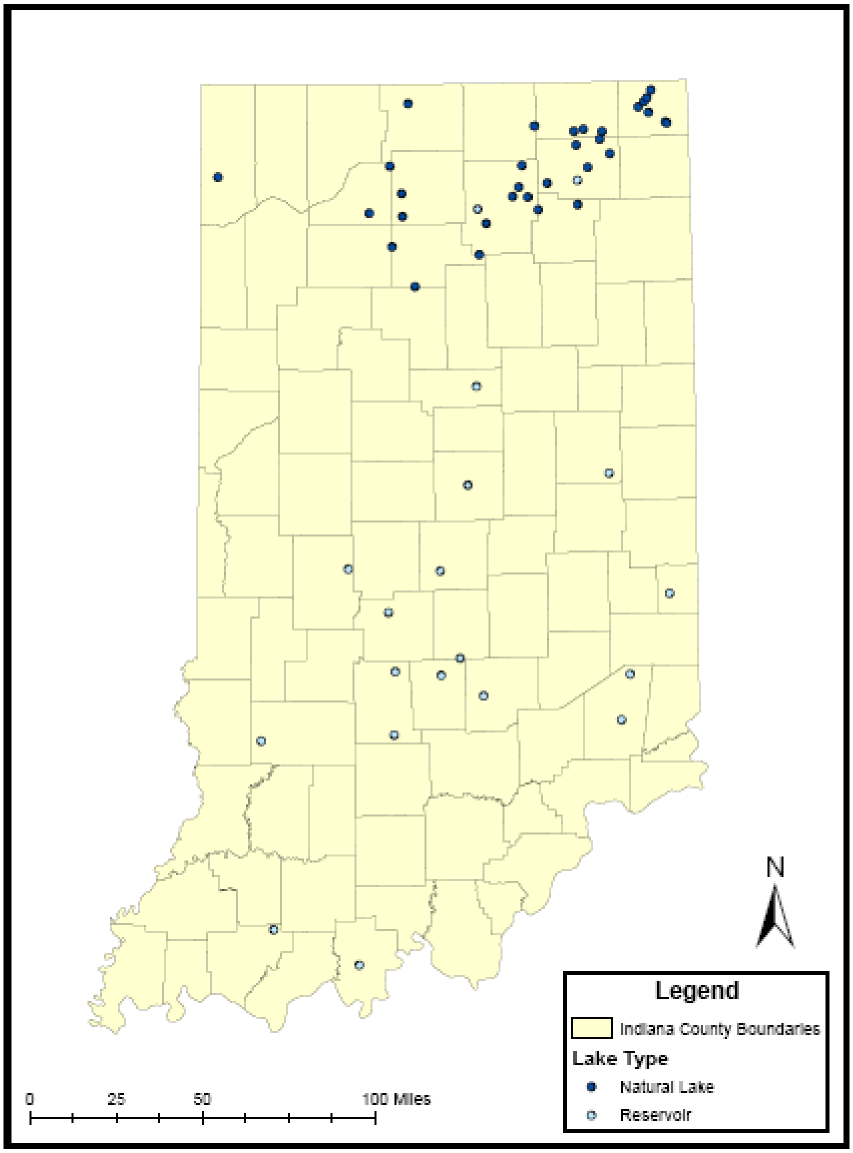


Figure 1. Location of the 51 lakes sampled in Indiana during the National Lakes Assessment.

95 d), ammonium (34-37 d), nitrate plus nitrite (8-9 d), nitrate (94-165 d), chlorine, (97-170 d), SiO_2 (8-9 d), sodium (181-323 d), sulfate (94-169 d), calcium (181-302 d), magnesium (181-315 d), and potassium (181-323 d). Eighteen turbidity values were below detection limits, as were 2 phosphorus values, 9 ammonium values, and 12 nitrate values (Appendix 1).

<u>Data analysis</u>

All flagged data were included in the analysis. Values below detection limits were not modified. Of the variables with measurements below detection limit, only P (with 2 measurements below detection levels) was used extensively in analysis.

To determine how the NLA data compared to the larger sample of the Clean Lakes Program, we used the Indiana Clean Lakes data from 2000-2005, a period which covers all the lakes generally sampled in the CLP. Only first visits to each lake during that period were used. In comparing NLA and CLP data, we use Mann-Whitney-Wilcoxon tests because, although the smallest sample size (n = 51) is above the rule of thumb used to suggest parametric tests (n = 30), some distributions were strong non-normal, suggesting, instead, a more conservative approach. In comparing NLA data for natural lakes and impoundments, we again used Mann-Whitney-Wilcoxon tests, in this instance because the smaller sample size was below the rule-of-thumb value for parametric tests.

Correlation and location tests were run using the SPSS statistical package. Comparative tests were run using nonparametric statistics due to the relatively small sample size of the NLA data set and the obviously non-normal distribution of most of the variables. Spearman rank correlations were used to detect monotonic, but not necessarily linear relationships among variables. Like the parametric Pearson correlation coefficients, Spearman rank correlation coefficients vary between -1 and 1; 0 values suggest no relationship between the analyzed variables, -1 shows a perfect, monotonic, negative relationship and 1 shows a perfect, monotonic, positive relationship.

Cluster analysis and ordinations were used to investigate multivariate relationships in the data; cluster analyses were run in SPSS; ordinations were run using PC-Ord (McCune and Mefford 2006). Multivariate normality of the data was unlikely given the lack of normality in the individual variables, so Bray-Curtis and nonmetric multidimensional scaling (NMDS or NMS) ordinations were used, which do not require multivariate normality. We used hierarchical clustering with Euclidean distance and average linkage as an exploratory approach. For Bray-Curtis ordination, we used the PC-Ord default settings of Sørenson distance (also called Bray-Curtis distance) with endpoint selection using variance-regression. Sørenson distance minimizes impact of outliers relative to Euclidean distance, and variance-regression minimizes impact of outliers relative to the original method of endpoint selection, resulting in higher likelihood of finding endpoints at the end of the true major axis of the data. For NMDS, we used relative Euclidean distance, which standardizes units, and 500 iterations, to maximize likelihood of the best global solution. The chosen techniques are standard for their ordination type; as our work was an initial exploration and the data set had no particular characteristics that suggested other choices, we used the common techniques.

Cluster analysis looks for points that are close to each other in the data cloud, as judged by whatever distance measure is used. As points are joined together into clusters, distances among clusters are judged using the linkage technique (between the center of the clusters, the edges of the clusters, etc).

Bray-Curtis ordination looks for the longest axis in the data cloud – the axis of greatest variation in the data – as judged by whatever distance measure is used and whatever endpoint selection method is used – and arrays the sites along that axis. Then it looks for the longest axis perpendicular to the first, and again arrays sites along that axis, etc. Researchers interpret an axis by examining the correlation of measured variables to the synthetic “variable” of the axis.

Particularly when used with the Sørenson distance measure, it has a reputation for producing readily-interpreted axes.

Non-metric multidimensional scaling is a constrained ordination technique – ordinations constructed using one data set (here, a subset of water-quality variables) must also explain variation in a related data set (here, a subset of the landscape variables). NMDS is almost aggressively nonparametric; McCune and Grace (2002, p 125) describe it as “well suited to data that are nonnormal or are on arbitrary, discontinuous, or otherwise questionable scales.” NMDS is an ordination technique that finds new axes so that the data cloud as graphed on the new axes bears the strongest possible resemblance to the original data cloud in the original variable space.

Of the three techniques, only NMDS includes a measure of whether the ordination performs better than would be expected by chance. Cluster analysis is a purely exploratory technique that offers no explanatory mechanisms. Bray-Curtis creates synthetic axes that can be interpreted by correlation to the original variables, but no measure of performance. NMDS produces synthetic axes that are also interpreted by correlation to the original variables, and also provides a measure of performance.

RESULTS

Descriptive Statistics

Descriptive statistics for the water quality variables (Table 1) and landscape variables (Table 2) show a number of skewed variables with medians displaced from the means. Variables with distributions skewed to the right tended to have one or a few unusually large values (e.g., area, turbidity) as can be seen in the histograms (Figure 2).

The HYMAPS-OWL program was unable to correctly define the watershed around 6 lakes due to excessively flat topography. Agriculture was the most common land cover in the watersheds of the measured lakes for which watershed coverage could be determined (n = 45; Table 2). Forest was the next most common land cover, and only these land covers exceeded the average % cover of the lakes themselves. Commercial and industrial land covers were less than 1% of watershed area for more than half the measured lakes.

Table 1. Descriptive statistics for lake area and water quality variables for the 50 Indiana lakes selected in the NLA sample and Olin Lake, a reference lake added to the sample.

	Mean	**Median**	**25-ile**	**75-ile**	**SD**	**Min**	**Max**
Area (ha)	159.71	37.50	15.19	119.58	406.67	5.10	2761.79
pH	8.45	8.48	8.37	8.57	0.26	7.57	9.07
Conductivity (µS/cm at 25°C)	384.3	383.9	308	465	140.0	130.6	864.6
Gran Acid-Neutralizing Cap (µeq/L)	2640	2797	1983	3218	872	555	4682
Turbidity (NTU)	4.68	2.70	1.50	4.99	5.89	0.42	35.5
Water Color (PCU)	11.7	10.0	6.0	16.0	9.9	0	64.0
Total Organic Carbon (mg/L)	7.56	7.16	4.25	8.58	4.29	2.69	27.59
Dissolved Organic Carbon (mg/L)	6.87	6.51	4.08	8.11	3.66	2.62	25.76
Total Phosphorus (µg/L)	26.29	17.0	8.0	34.0	29.79	2.0	170.0
Total Nitrogen (µg/L)	826	713	501	1101	493	88	2091
Ammonium (mg N/L)	0.030	0.021	0.015	0.028	0.044	0.003	0.315
Nitrate + Nitrite (mg N/L)	0.078	0	0	0.007	0.273	0	1.720
Nitrate (mg N/L)	0.087	0	0	0	0.278	0	1.739
Chloride (mg/L)	25.38	21.85	12.19	29.12	20.76	1.64	124.4
Sulfate (mg/L)	28.41	22.53	12.64	34.49	27.99	2.82	183.1
Calcium (mg/L)	37.60	39.58	27.44	46.79	13.59	11.7	72.29
Magnesium (mg/L)	15.67	16.30	11.34	19.63	6.81	3.99	42.20
Sodium (mg/L)	13.45	9.04	5.85	15.94	12.60	1.12	72.94
Potassium (mg/L)	2.42	2.13	1.69	2.79	1.43	0.343	9.47
Silica (mg/L SiO_2)	4.02	2.97	1.70	4.73	3.90	0.050	23.46
Chlorophyll *a* (µg/L)	15.74	5.65	3.00	20.52	21.52	0.944	118.8

Table 2. Land cover in watersheds of measured lakes, shown as absolute (ac) and proportional area. Areas were determined using the HYMAPS-OWL hydrologic model.

Area (ha) and % area	Mean	Median	25-ile	75-ile	SD	Min	Max
Water	927	224	81	725	2,039	0	10,901
Commercial	122	2	0	38	570	0	3,812
Agricultural	6,334	885	72	2,985	19,027	0	114,364
High density urban	331	37	3	250	1,128	0	7,464
Low density urban	681	193	33	508	1,467	0	6,935
Grass/pasture	711	201	53	725	1,900	0	12,421
Forest	2,802	408	96	1,147	9,114	0	46,329
Industrial	33	0	0	13	86	0	310
Undefined	0	0	0	0	1	0	7
Total watershed area	11,941	2,965	509	7,467	27,957	0.7	135,474
% water	16.8	11.6	4.9	23.9	15.3	0	61.3
% commercial	1.0	0.1	0	0.5	3.1	0	20.4
% agricultural	37.5	38.1	14.4	57.3	25.0	0	84.4
% high density urban	3.2	1.3	0.3	4.0	4.6	0	20.6
% low density urban	6.6	5.6	4.4	8.5	3.9	0	17.5
% grass/pasture	9.9	8.2	2.6	13.2	9.7	0	45.5
% forest	22.8	15.2	6.5	24.4	26.0	0	95.5
% industrial	1.2	0	0	0.2	7.2	0	48.0

Extreme Values

The nature of the water quality measurements was such that unusual values tend to occur only as large numbers. For some variables (e.g., turbidity, ammonium sulfate, chlorophyll *a*), lakes with low values were common. For others, the measurements were bounded by zero and measures near zero, if not common, were not sufficiently uncommon as to make low values obviously unusual. The values described here are not limited by any particular statistical cut-off; the varied distributions of the data make such cut-offs inconsistent. Rather, examination of histograms, and results from multivariate analysis were used to identify extreme values (see Figure 2).

Strakis and Hert Lakes were both affected by industrial activity – Strakis by a landfill and limestone mining (as well as agriculture), Hert by coal mining (Table 3). Rock, Fish, Cedar, Versailles, Bischoff, Harper, Palestine, Skinner, South Chain, and Whitewater Lakes all had some values associated with eutrophication (high nutrients, turbidity). Higher pH values in some lakes were also associated with high algal productivity (Figure 3). Little Otter Lake is in a more residential setting, with some natural land. Saddle and Monroe Lakes had values associated with the more acidic soils of the unglaciated southern part of the state; Monroe Reservoir is also the largest body of water in the sample, at almost 4 times the size of the next largest lake. Olin Lake, added to the NLA sample to serve as a reference lake (closer to pristine conditions than most in the northern part of the state), has an undeveloped shoreline, but its surrounding mature forest is not extensive and the landscape beyond is heavily agricultural.

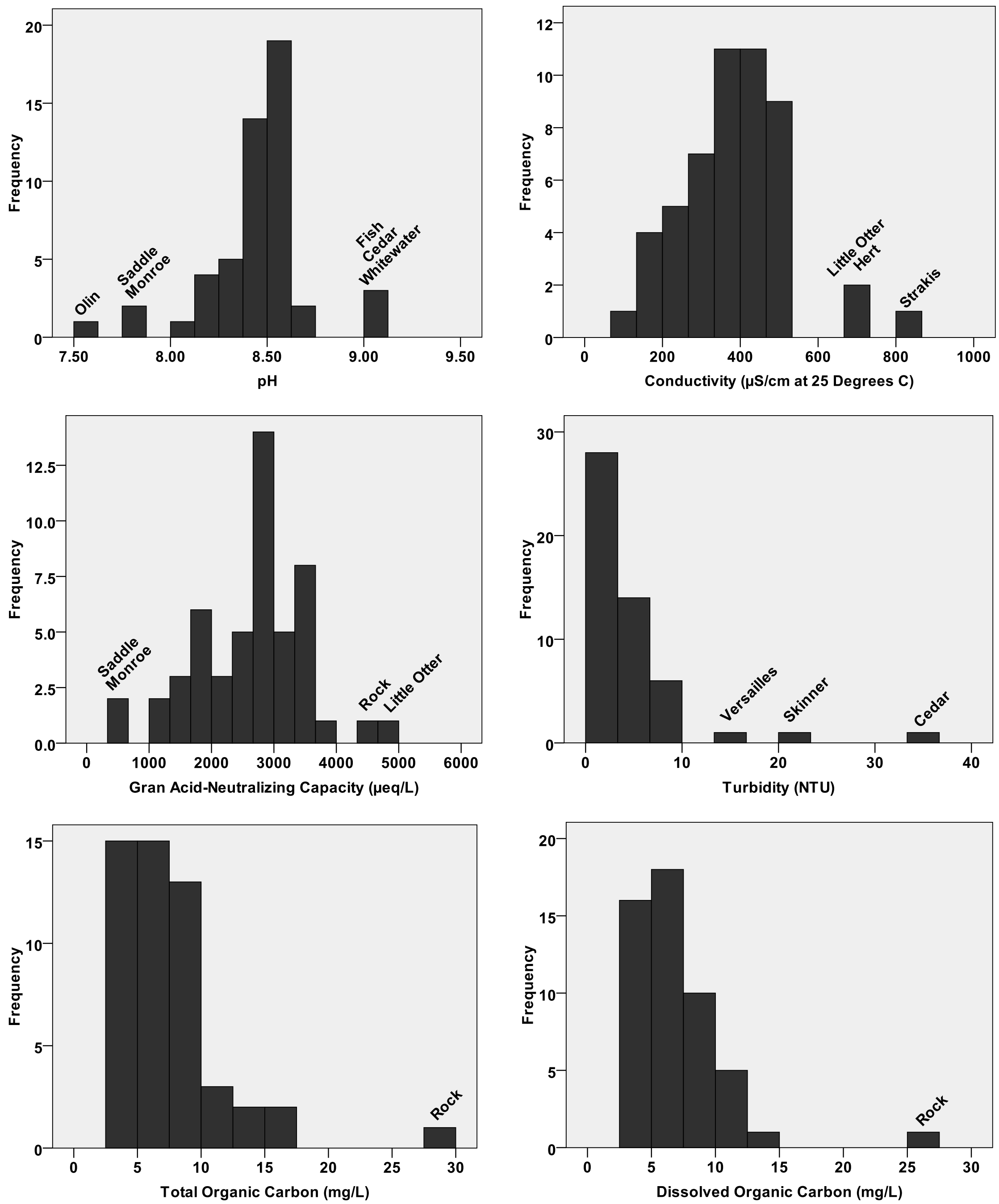


Figure 2. Distributions of water-quality values for NLA lakes. Lake names are shown for unusual values.

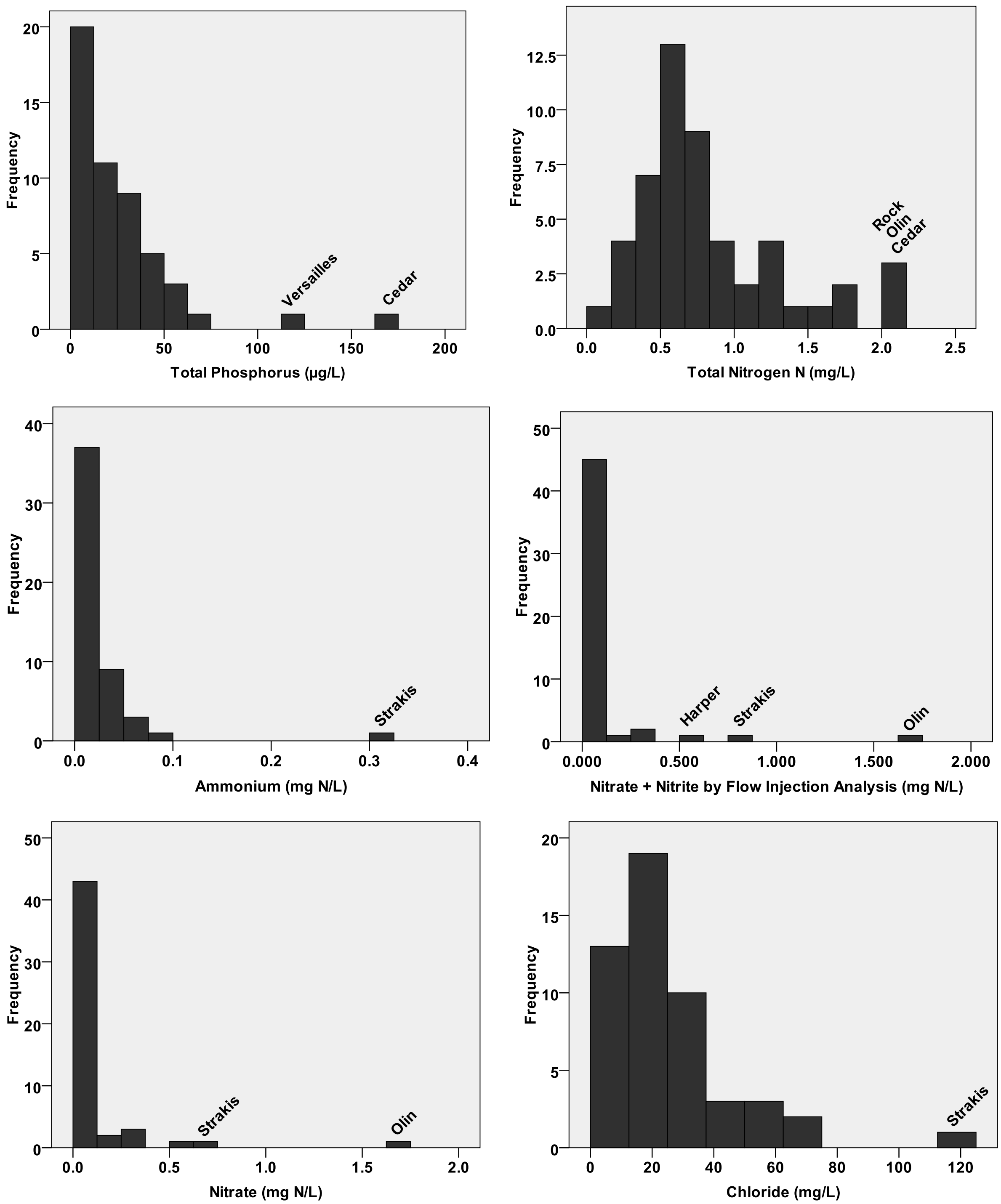
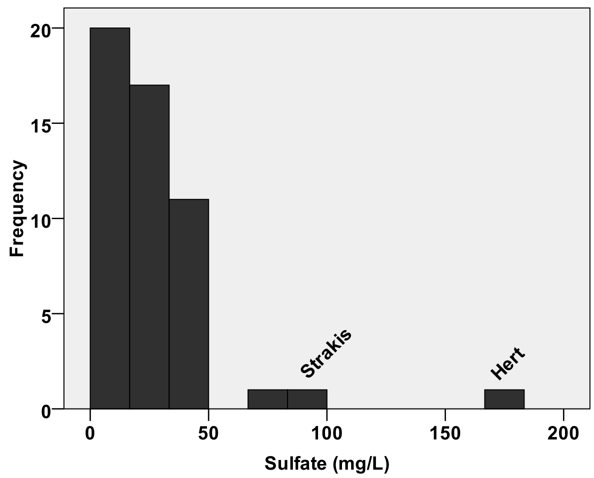
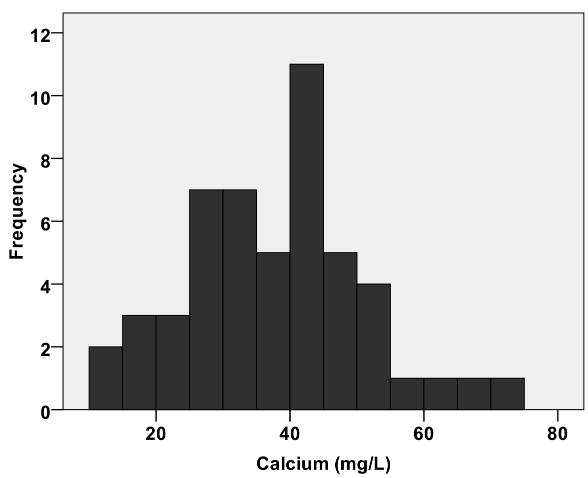
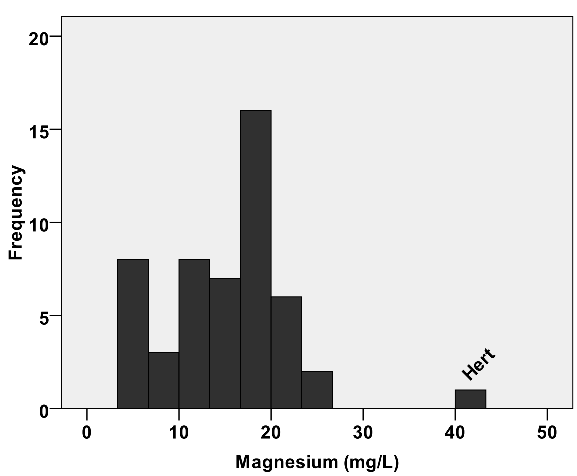
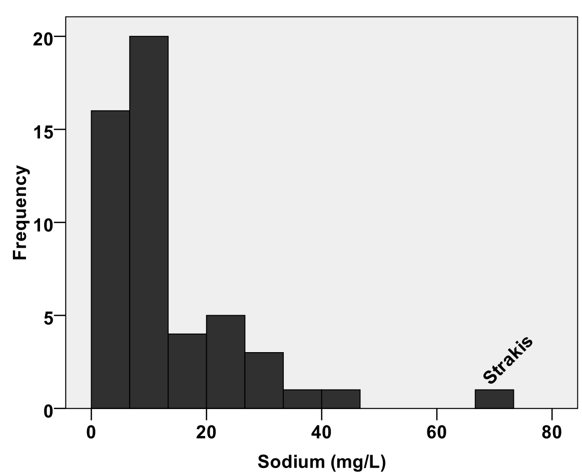


Figure 2 (continued). Distributions of water-quality values for NLA lakes. Lake names are shown for unusual values.


Frequency
Strakis
Hert
Sulfate (mg/L)


Frequency
Calcium (mg/L)


Frequency
Hert
Magnesium (mg/L)


Frequency
Strakis
Sodium (mg/L)

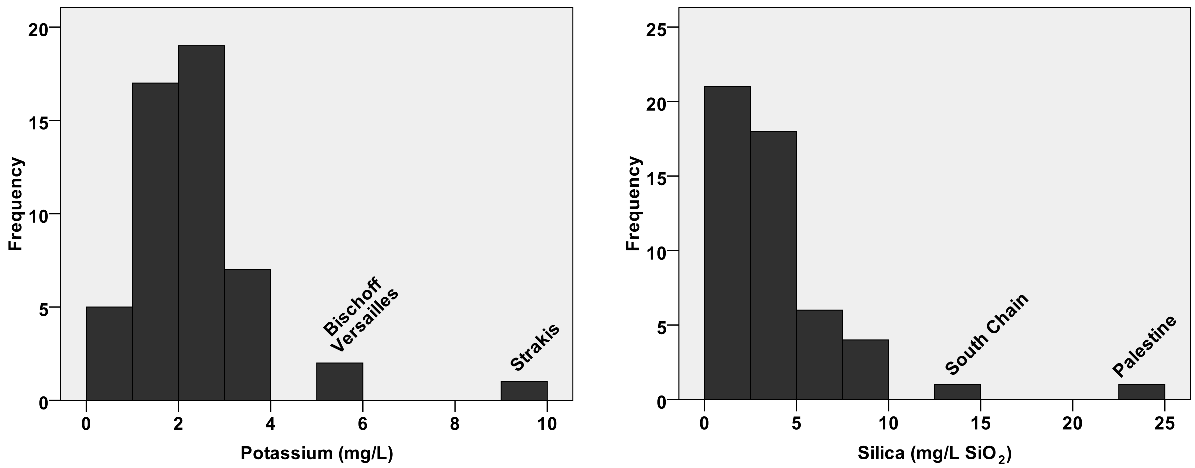


Figure 2 (continued). Distributions of water-quality values for NLA lakes. Lake names are shown for unusual values.

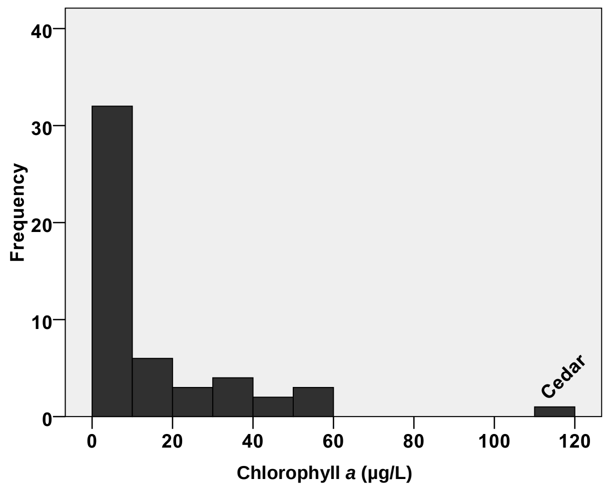


Figure 2 (continued). Distributions of water-quality values for NLA lakes. Lake names are shown for unusual values.

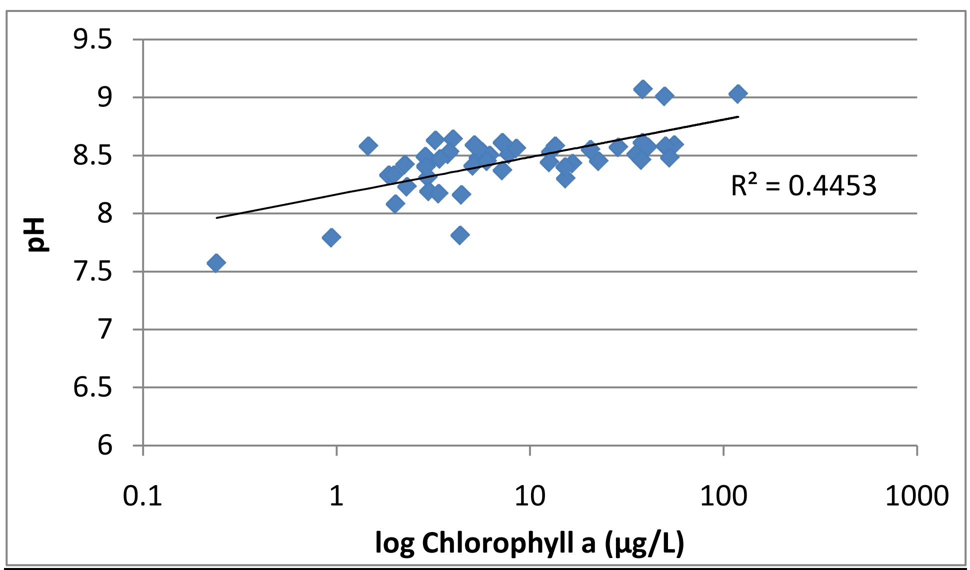


Figure 3. Higher chlorophyll *a* production results in increased pH in the Indiana NLA lakes.

Table 3. Extreme water quality values in the NLA lakes.

Lake	County	Extreme	Value	Units	Landscape features
Strakis	Marion	high conductivity	864.6	µS/cm	Landfill, limestone mine, and agricultural activity all in immediate vicinity
		high NH_4	22.5	mg N/L	
		high NO_3+NO_2	0.75	mg N/L	
		high NO_3	0.68	mg N/L	
		high Cl	124	mg/L	
		high SO_4	96.3	mg/L	
		high Na	72.9	mg/L	
		high K	9.5	mg/L	
Hert	Greene	high conductivity	242	µS/cm	Final cut lake left from coal mining
		high SO_4	676	mg/L	
		high Mg	42.2	mg/L	
Rock	Kosciusko	high color	64	PCU	Agriculture
		high gran acid-neutral cap	4,563	µeq/L	
		high TOC	27.59	mg/L	
		high DOC	25.76	mg/L	
		high ttl N	2	mg/L	
Fish	Elkhart	high pH	9.07		Homes, forest, surrounded by agriculture
Cedar	Lake	high pH	9.03		Homes, agriculture
		high turbidity	35.5	NTU	
		high ttl P	170	µg/L	
		high ttl N	2.0	mg/L	
		high chlorophyll *a*	119	µg/L	
Versailles	Ripley	high turbidity	15.4	NTU	Agriculture
		high ttl P	121	µg/L	
		high K	5.1	mg/L	
Bischoff	Ripley	high K	5.2	mg/L	Agriculture
Harper	Noble	high NO_3+NO_2	0.54	mg N/L	Campground, forest surrounded by agric
Palestine	Kosciusko	high silica	23.5	mg/L SiO2	Homes, forest, surrounded by agriculture
Skinner	Noble	high turbidity	22.1	NTU	Homes, agriculture
South Chain	St. Joseph	high silica	12.6	mg/L SiO2	Golf course, homes, forest, ag, auto salvage
Whitewater	Union	high pH	9.01		Forest-park, agriculture
Little Otter	Steuben	high conductivity	682	µS/cm	Parks and homes
		high gran acid-neutral cap	4,682	µeq/L	
Saddle	Perry	low pH	7.79		Forest on acid soils
		low gran acid-neutral cap	555	µeq/L	
Monroe	Monroe/ Brown/ Jackson	large area	2,762	ha	Primarily forest on acid soils
		low pH	7.81		
		low gran acid-neutral cap	602	µeq/L	
Olin	LaGrange	low pH	7.57		Forest, surrounded by agriculture, on acid soils
		high ttl N	2.1	mg/L	
		high NO_3+NO_2	1.72	mg N/L	
		high NO_3	1.74	mg N/L	

EPA Water Quality Standards

Values of variables measured during the NLA survey did not exceed the maximum values provided as guidelines for aquatic life use by IDEM (DO values were not provided and could not be checked). No lake exceeded the total phosphorus value of 0.3 mg/l used for Aquatic Life Use (see extreme value section). Similarly, no lake exceeded the NO_3+NO_2 value of 10 mg N/l used for Aquatic Life Use.

Three lakes exceeded the pH value of 9 used for Aquatic Life Use (Fish and Cedar Lakes in Lake County, and Whitewater Lake in Union County: 9.07, 9.03, and 9.01). However, the Aquatic Life Use Support water quality standards given in the Indiana Department of Environmental Management Integrated Water Monitoring and Assessment Report (2008, Figure 8) allow pH values to exceed 9 if the pH values “are correlated with photosynthetic activity.” All three of the exceeding lakes are highly eutrophic, and the pH values are likely in response to photosynthetic activity and its impact on the carbonate chemistry. Thus, no lakes exceeded the pH guidance, so long as eutrophication can be considered.

The ammonia guidelines are given for unionized NH_4, which is a function of temperature and pH. Temperature is not in the database. The highest total ammonia reading was 0.315 mg/l at a pH of 8.41. Assuming a water temperature of 22°C, about 10% of the ammonia would be unionized, or about 0.032 mg/l. The maximum value permitted at that temperature and pH is 0.214 mg/l, almost an order of magnitude higher. The next lowest reading is only approximately 1/3 as high, and at a similar pH. The unionized ammonia would be similarly lower. None of the measurements exceeds the maximum unionized ammonia values.

Comparison to CLP

Comparisons between the NLA and CLP data were possible for pH, conductivity, ammonium, nitrate, phosphorus, and chlorophyll *a*. Although the histograms indicated a general similarity (Figure 4), some distributional differences were noticeable in the descriptive statistics (Table 4) and the sample size, particularly for the CLP data (n = 370-380 values, depending on the variable) was high. All variables tested had significantly different medians and distributions in the two data sets (Mann-Whitney-Wilcoxon tests and Kolmogorov Smirnov, all $p < 0.01$). Sampling dates for the two programs (Figure 5) and selected weather data for the sampling periods (Figure 6) are also shown.

We also used epilimnetic total phosphorus data to calculate Carlson’s Trophic State Index (TSI) (Carlson 1977). Distributional differences are apparent as the NLA data result in a greater proportion of oligotrophic lakes and a lower proportion of eutrophic and hypereutrophic lakes than do the CLP data (Figure 7).

The ratio of total nitrogen to total phosphorus (TN:TP) is often used to evaluate nutrient limitation in phytoplankton growth. While the TN:TP ratio varies depending on the phytoplankton community present, Smith (1982) and others have generally used a 12:1 ratio as the break between nitrogen limitation (<12:1) and phosphorus limitation (>12:1). Using these criteria, 48 NLA lakes (94%) and 334 CLP lakes (87%) were phosphorus limited.

Table 4. Comparison of water quality values for lakes sampled in the NLA (n = 51) and CLP (n = 270-280) programs. NLA data were sampled in 2007, CLP data over 2000-2005.

	NLACLP									
	CLP					NLA				
	Mean	**Median**	**25-ile**	**75-ile**	**St Dev**	**Mean**	**Median**	**25-ile**	**75-ile**	**St Dev**
pH	8.31	8.30	8.10	8.60	0.48	8.45	8.48	8.37	8.57	0.26
Conductivity (µS/cm 25°C)	571	421	329	549	550	384	384	308	465	140
Total Phosphorus (µg/L)	51.1	39.0	24.0	65.0	48.6	26.3	17.0	8.0	34.0	29.8
Ammonium (mg N/L)	0.056	0.031	0.018	0.066	0.064	0.030	0.021	0.015	0.028	0.044
Nitrate+nitrite (mg N/L)	0.415	0.022	0.013	0.294	0.883	0.078	0.000	0.000	0.007	0.273
Chlorophyll *a* (µg/L)	10	3	1	11	16	16	6	3	21	22

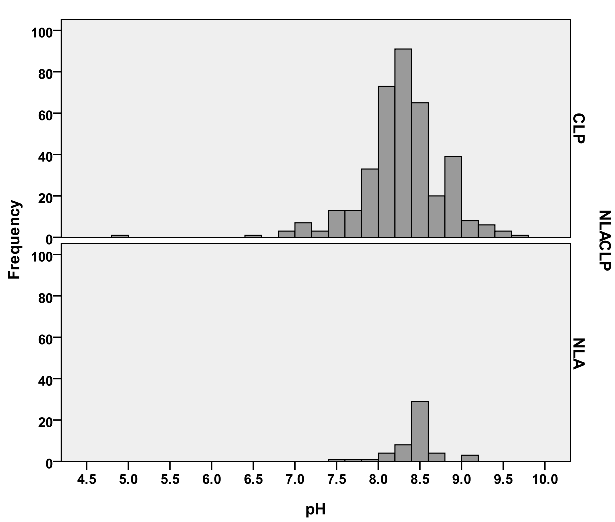
Natural Lakes Compared to Others (impoundments, quarry and mine lakes, etc.)

Natural lakes were more northern in their distribution (Figure 1), and had less forested area and more agricultural area than impoundments and other created lakes. In terms of water chemistry, natural lakes had higher conductivity, ANC, TOC, TN, and Ca. Other ions and nutrient-related measurements were also higher in natural lakes, but not significantly so (Table 5). Watersheds around natural lakes had significantly more agricultural area and significantly less forested area than watersheds around impoundments.

Glacial history and geology are the primary drivers for land uses and land features within Indiana. The most recent glacial era some 10,000 to 14,000 years ago covered the northern third of the state and left behind numerous ice-block or kettle lakes. Earlier glaciers extended to central Indiana. Glacial till in these regions is more suitable for agriculture than the thin soils in unglaciated southern Indiana.

Table 5. Mean and median water quality and watershed composition values for impoundments (n = 19) and natural lakes (n = 32) sampled in the NLA program. A check mark denotes a statistically significant difference as indicated by a Mann-Whitney-Wilcoxon result of $p < 0.05$. Differences with $0.05 < p < 0.10$ are marked with a 'b;' all other results had $p > 0.10$.

		Impound		Natural	
	Sgn	Mean	Median	Mean	Median
pH		8.4	8.4	8.5	8.5
Conductivity (uS/cm at 25 °C)	ü	346.2	330.8	413.7	429.1
Gran Acid-Neutralizing Capacity	ü	2164.5	2223.3	2971.7	2925.5
Turbidity (NTU)		4.2	2.7	5.0	3.0
Total Organic Carbon (mg/L)	ü	4.9	4.2	9.2	8.1
Total Phosphorus (ug/L)		28.7	16.5	25.5	17.0
Total Nitrogen N (mg/L)	ü	0.65	0.48	0.95	0.76
Chloride (mg/L)		23.9	15.0	26.9	23.5
Sulfate (mg/L)		31.5	15.0	26.9	27.1
Calcium (mg/L)	ü	30.1	28.2	42.6	41.1
Magnesium (mg/L)	b	14.5	13.5	16.7	18.4
Potassium (mg/L)		2.7	1.8	2.3	2.3
Silica (mg/L SiO2)		2.9	2.5	4.6	3.5
CHLA		14	5	17	7
pwaterar	ü	0.11	0.07	0.20	0.20
pcommar		0.02	0.00	0.01	0.00
pagarea	ü	0.29	0.24	0.44	0.47
phidensar		0.03	0.01	0.03	0.02
plodensar		0.07	0.05	0.07	0.06
pgraspastar	b	0.07	0.06	0.12	0.10
pforestar	ü	0.38	0.23	0.12	0.09
pindustar		0.03	0.00	0.00	0.00



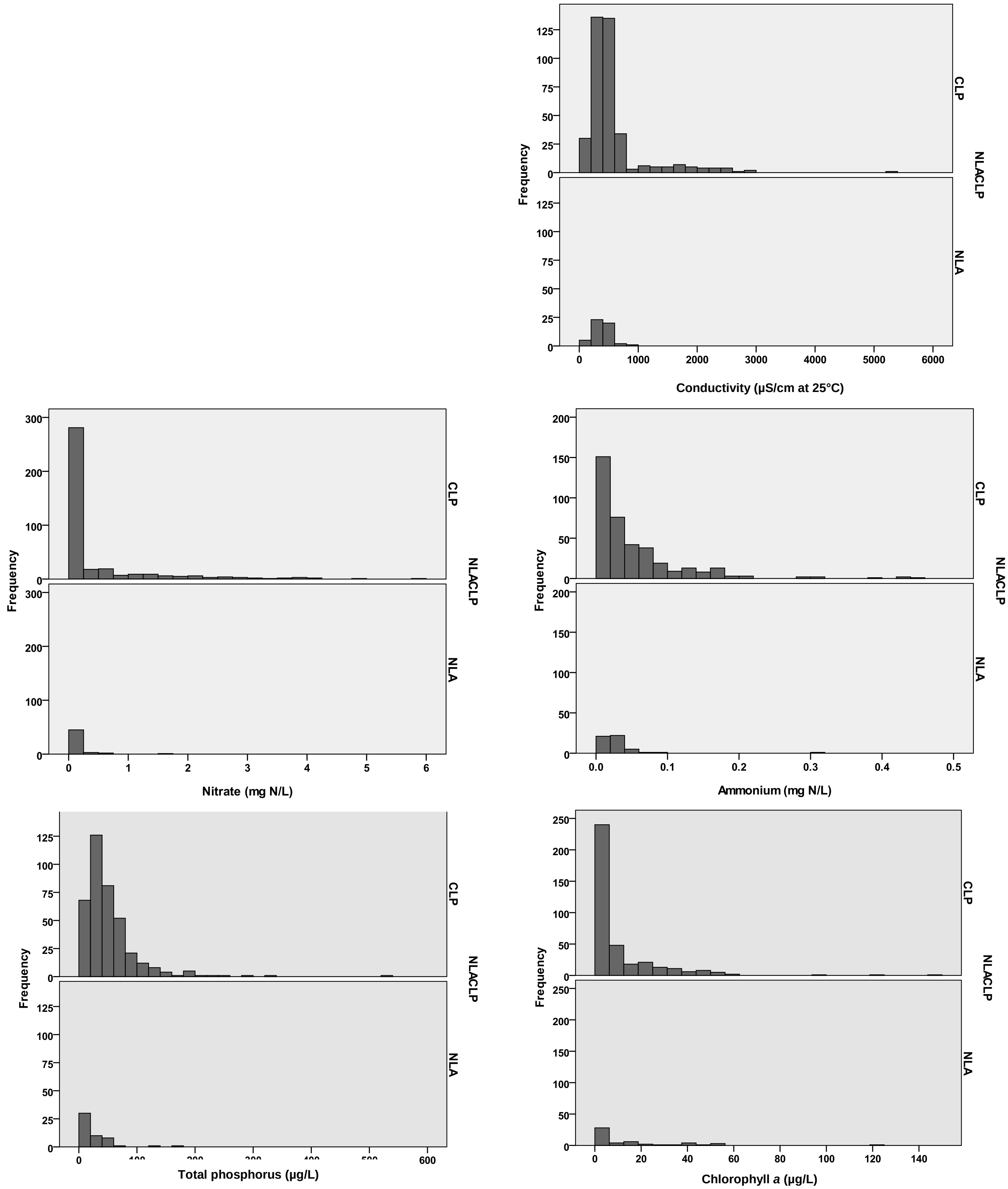


Figure 4. Frequencies of water-quality values for lakes sampled in the NLA (*n* = 51) and CLP (*n* = 270-280) programs. NLA data were sampled in 2007, CLP data over 2000-2005. In each pair, the Y axes are forced to be identical.

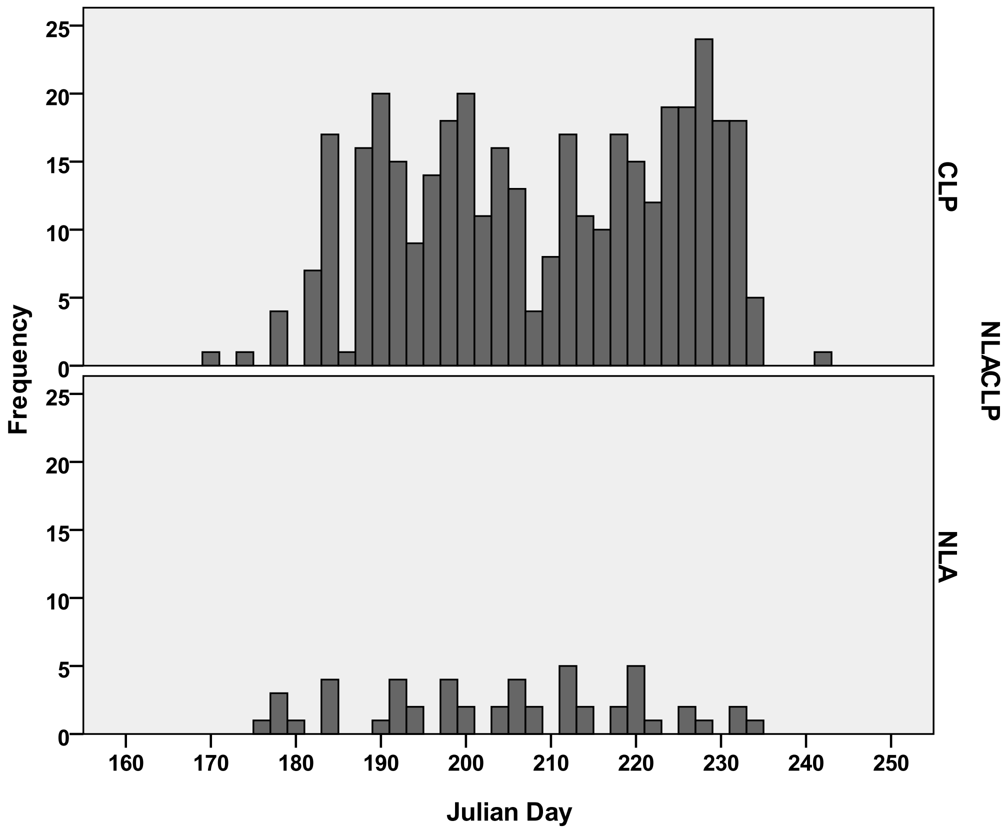


Figure 5. Sampling dates for 2007 NLA data and CLP data collected during 2000-2005. Julian date gives the day of the year (from 1 to 365) sampling occurred.

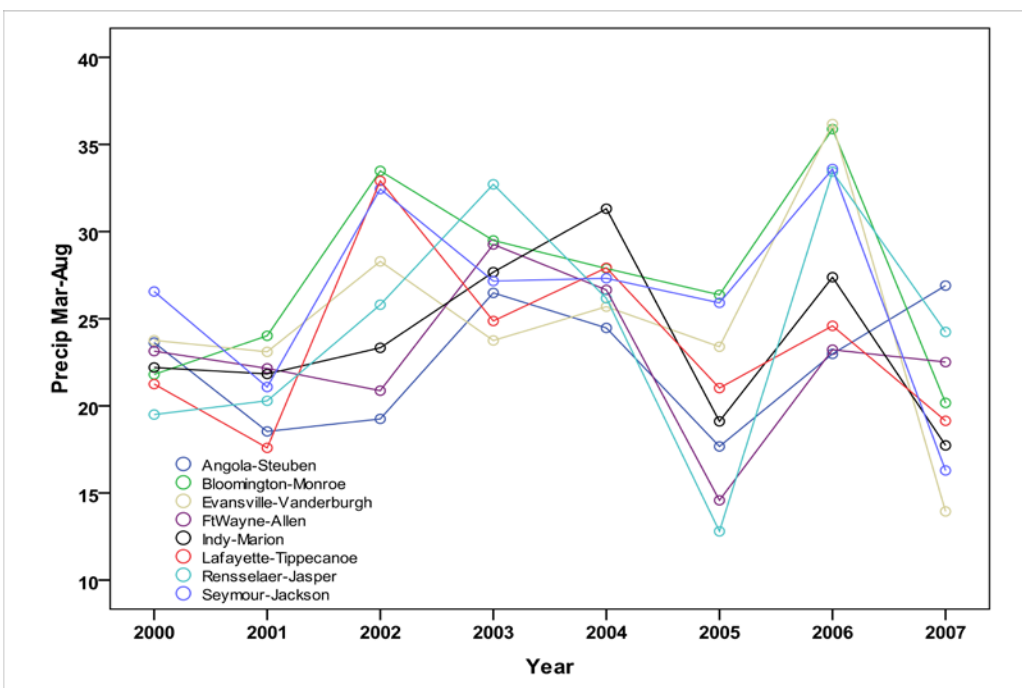


Figure 6. Precipitation patterns during CLP and NLA sampling periods.

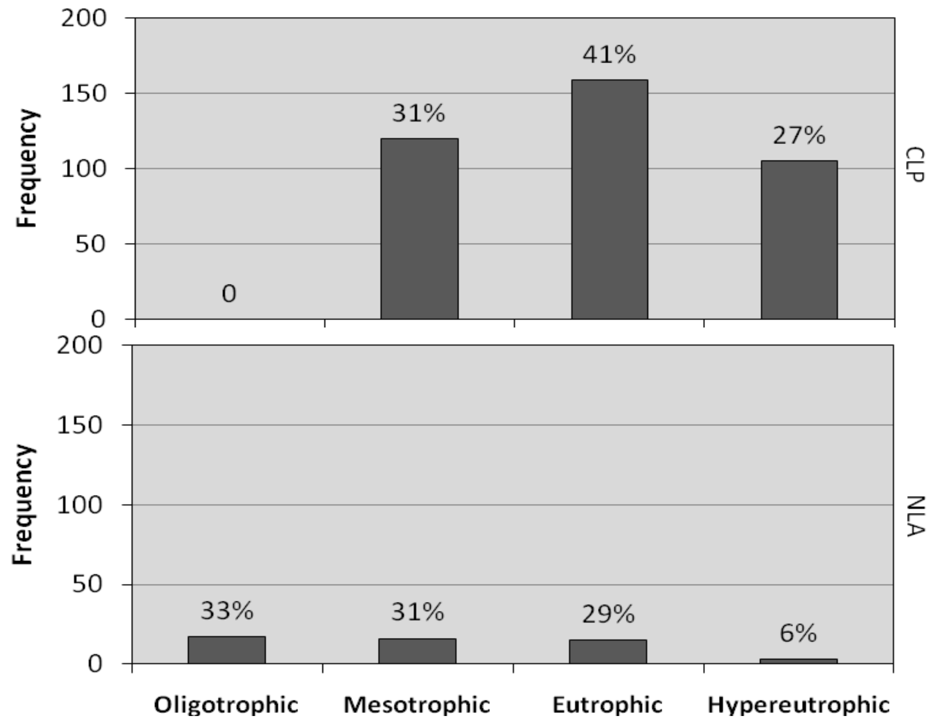


Figure 7. Frequency of trophic state for 2007 NLA data and CLP data collected during 2000-2005. The Y-axes are forced to be identical. Percentages for each treatment are shown.

Correlational Analysis

Lake depth was strongly correlated with lake area (r_s = 0.407, p = 0.003), correlated with chloride (r_s = 0.31, p = 0.029), and suggestively related ($0.05 < p < 0.10$) to total phosphorus (r_s = -0.275), % forested area (r_s = -0.279), % ag area (r_s = 0.270), and latitude (r_s = -0.274).

<u>Water quality variables</u>

Most of the water quality variables can be placed into one of three categories: nutrient variables (measurements of C, N, P, K), salts (chloride, sulfate, calcium, magnesium, sodium), and summary variables (pH, conductivity, gran acid neutralizing capacity, turbidity, color). Silica and chlorophyll *a*, associated with diatoms and algae, respectively, complete the list. Secchi disk depths were available, but were not used as several lakes could not be measured due to high water clarity (the disk was still visible at the bottom).

The nutrient variables associated with C, N, P, and K were often strongly intercorrelated (Table 6; e.g., total organic carbon and nitrogen with a Spearman rank correlation coefficient of 0.7). Although total N was in this group, as was ammonium, although to a lesser extent, nitrate, and nitrate plus nitrite were not strongly related to the other nutrient variables. Nitrate and nitrate plus nitrate were weakly correlated with ammonium and even more weakly correlated with total N. K was weakly correlated with several of the other nutrient variables. The summary variables associated with algal blooms – turbidity, water color, chlorophyll *a* were generally highly correlated with the limiting nutrient variables C, N, and P, but not with K. The nutrient variables had a variety of intermediate correlations with the salts, but there were no obvious patterns.

Table 6. Spearman rank correlations among water quality variables for the 50 NLA lakes and the reference lake. Light yellow shows relationships with correlation coefficients between 0.3 and 0.5. Bright yellow shows relationships with correlation coefficients between 0.5 and 0.7. Orange shows relationships with correlation coefficients $\geq$ 0.7 (these breakpoints are arbitrary).

		Total Organic Carbon (mg/L)	Dissolved Organic Carbon (mg/L)	Total P (µg/L)	Total N (mg/L)	NH_4 (mg N/L)	Nitrate + Nitrite (mg N/L)	Nitrate (mg N/L)	Potassium (mg/L)
pH	Corr Coeff	0.471	0.461	0.317	0.402	0.157	-0.054	-0.022	0.27
	P	<0.0005	0.001	0.024	0.003	0.27	0.709	0.879	0.056
Conductivity (µS/cm)	Corr Coeff	0.297	0.358	0.019	0.431	0.394	0.289	0.241	0.798
	P	0.034	0.01	0.894	0.002	0.004	0.04	0.089	<0.0005
Gran Acid-Neutral Capacity (µeq/L)	Corr Coeff	0.383	0.457	0.138	0.395	0.344	0.161	0.126	0.695
	P	0.006	0.001	0.333	0.004	0.013	0.261	0.38	<0.0005
Turbidity (NTU)	Corr Coeff	0.485	0.401	0.691	0.579	0.296	-0.15	-0.048	0.14
	P	<0.0005	0.004	<0.0005	<0.0005	0.035	0.292	0.739	0.328
Water Color (PCU)	Corr Coeff	0.624	0.633	0.637	0.57	0.237	0.043	0.072	0.138
	P	<0.0005	<0.0005	<0.0005	<0.0005	0.094	0.763	0.615	0.334
Total Organic Carbon (mg/L)	Corr Coeff	1	0.969	0.416	0.7	0.46	-0.057	-0.088	0.338
	P	.	<0.0005	0.002	<0.0005	0.001	0.692	0.537	0.015
Dissolved Organic Carbon	Corr Coeff		1	0.344	0.653	0.463	-0.007	-0.03	0.329
	P		.	0.013	<0.0005	0.001	0.961	0.832	0.019
Total P (µg/L)	Corr Coeff			1	0.571	0.022	-0.036	-0.052	0.471
	P			.	0	0.877	0.8	0.715	<0.0005
Total N (mg/L)	Corr Coeff				1	0.519	0.255	0.288	0.494
	P				.	<0.0005	0.071	0.04	<0.0005
NH_4 (mg N/L)	Corr Coeff					1	0.359	0.393	0.27
	P					.	0.01	0.004	0.056
Nitrate + Nitrite (mg N/L)	Corr Coeff						1	0.836	-0.02
	P						.	<0.0005	0.889
Nitrate (mg N/L)	Corr Coeff								0.116
	P								0.419
Chloride (mg/L)	Corr Coeff	0.275	0.26	-0.02	0.365	0.335	0.167	0.235	0.402
	P	0.051	0.066	0.889	0.009	0.016	0.24	0.096	0.003
Sulfate (mg/L)	Corr Coeff	-0.028	0.033	-0.128	0.188	0.241	0.503	0.34	0.061
	P	0.847	0.816	0.372	0.186	0.089	<0.0005	0.015	0.671
Calcium (mg/L)	Corr Coeff	0.368	0.457	0.045	0.407	0.487	0.212	0.238	0.276
	P	0.008	0.001	0.755	0.003	<0.0005	0.135	0.092	0.05
Magnesium (mg/L)	Corr Coeff	0.103	0.161	0.063	0.388	0.305	0.426	0.345	0.303
	P	0.471	0.26	0.66	0.005	0.029	0.002	0.013	0.031
Sodium (mg/L)	Corr Coeff	0.119	0.103	-0.035	0.21	0.185	0.173	0.145	0.448
	P	0.404	0.472	0.808	0.14	0.195	0.226	0.31	0.001
Silica (mg/L SiO_2)	Corr Coeff	-0.077	-0.004	0.005	-0.071	0.009	0.264	0.137	-0.114
	P	0.589	0.976	0.972	0.62	0.952	0.062	0.339	0.427
Chlorophyll *a* (µg/L)	Corr Coeff	0.527	0.46	0.742	0.532	0.184	-0.065	-0.078	0.481
	P	<0.0005	0.001	<0.0005	<0.0005	0.195	0.651	0.584	<0.0005

Table 6 (continued). Correlations among water quality variables for the 50 NLA lakes and the reference lake. Light yellow shows relationships with correlation coefficients between 0.3 and 0.5. Bright yellow shows relationships with correlation coefficients between 0.5 and 0.7. Orange shows relationships with correlation coefficients $\geq$ 0.7 (these breakpoints are arbitrary).

		Conductivity (µS/cm)	Gran Acid-Neutralizing Capacity (µeq/L)	Turbidity (NTU)	Water Color (PCU)
pH	Corr Coeff	0.162	0.078	0.453	0.34
	P	0.257	0.587	0.001	0.015
Conductivity (µS/cm)	Corr Coeff	1	0.835	0.19	0.244
	P		0	0.182	0.084
Gran Acid-Neutralizing Capacity (µeq/L)	Corr Coeff		1	0.184	0.36
	P			0.195	0.009
Turbidity (NTU)	Corr Coeff			1	0.519
	P			.	0

		Cl (mg/L)	SO_4 (mg/L)	Ca (mg/L)	Mg (mg/L)	Na (mg/L)	SiO_2 (mg/L)	Chlorophyll *a* (µg/L)
pH	Corr Coeff	0.375	0.01	-0.058	0.27	0.406	-0.207	0.556
	P	0.007	0.945	0.685	0.056	0.003	0.145	<0.0005
Conductivity (µS/cm)	Corr Coeff	0.625	0.594	0.803	0.798	0.61	0.265	0.092
	P	<0.0005	<0.0005	<0.0005	<0.0005	<0.0005	0.06	0.521
Gran Acid-Neutralizing Capacity (ueq/L)	Corr Coeff	0.344	0.362	0.897	0.695	0.257	0.313	0.16
	P	0.014	0.009	<0.0005	<0.0005	0.069	0.026	0.261
Turbidity (NTU)	Corr Coeff	0.154	0.013	0.191	0.14	0.107	-0.049	0.86
	P	0.28	0.925	0.179	0.328	0.454	0.734	<0.0005
Water Color (PCU)	Corr Coeff	0.126	0.044	0.343	0.138	0.034	0.16	0.606
	P	0.38	0.758	0.014	0.334	0.81	0.263	<0.0005
Cl (mg/L)	Corr Coeff	1	0.196	0.31	0.493	0.886	-0.112	0.062
	P	.	0.167	0.027	<0.0005	<0.0005	0.435	0.664
SO_4 (mg/L)	Corr Coeff		1	0.47	0.684	0.238	0.443	0.007
	P		.	<0.0005	<0.0005	0.092	0.001	0.963
Ca (mg/L)	Corr Coeff			1	0.6	0.172	0.297	0.112
	P			.	<0.0005	0.226	0.034	0.432
Mg (mg/L)	Corr Coeff				1	0.48	0.308	0.172
	P				.	<0.0005	0.028	0.228
Na (mg/L)	Corr Coeff					1	-0.04	0.018
	P					.	0.78	0.901
SiO_2 (mg/L)	Corr Coeff						1	-0.009
	P						.	0.952

The two summary variables associated with acidity – pH and gran acid neutralizing capacity were highly correlated; the two summary variables associated with water clarity – turbidity and color – were also fairly highly correlated (Table 6). Conductivity was strongly correlated with all the salts. The salts showed a number of correlations, the strongest between anion-cation pairs that regularly co-occur, such as sodium with chloride.

Landscape Analysis

The 7 major proportional terrestrial landscape categories showed several intercorrelations, particularly among the highly-developed land covers (Table 7). In this group, the highest correlations were between commercial landcover and industrial, and high-density urban landcover. Lower correlations occurred between low-density urban landcover and high-density urban and commercial landcovers and between high-density urban and industrial landcovers. The two open landcover types, agriculture and forest, were negatively correlated.

Most of the correlations between proportional landcover and water quality variables were of lesser significance (Table 8). Among the stronger correlations were sodium and chloride with commercial and high-density landcovers, magnesium with agricultural landcover, and conductivity with low-density area. Among the landcover types, agricultural landcover was correlated with the highest number of water quality variables, including nutrient variables DOC, total N, and K, but not total TOC or P. Forest cover was only negatively correlated with water quality variables. Among the water quality variables, only total phosphorus and sulfate showed no significant correlations with any landscape variable. Silica showed only negative correlations with upland landcover types.

Table 7. Spearman rank correlational analyses of proportional landscape cover variables and water-quality variables (n = 45 for all correlations). Light yellow shows relationships with correlation coefficients between 0.3 and 0.5. Bright yellow shows relationships with correlation coefficients between 0.5 and 0.7. Orange shows relationships with correlation coefficients $\geq$ 0.7.

		% water	% comm	% ag	% hi-d urb	% lo-d urb	% gr-past	% forest	% indust
% water	Corr Coeff	1	0.042	-0.284	0.264	0.211	.403**	-0.264	-0.092
	Sig (2-t)		0.783	0.059	0.08	0.164	0.006	0.08	0.549
% com-mercial	Corr Coeff		1	-0.109	.765**	.445**	0.29	-0.066	.719**
	Sig (2-t)			0.477	0	0.002	0.053	0.667	0
% ag	Corr Coeff			1	-0.011	0.16	-0.247	-.424**	-0.189
	Sig (2-t)				0.943	0.294	0.101	0.004	0.215
% hi-dens urban	Corr Coeff				1	.475**	0.234	-0.28	.456**
	Sig (2-t)				.	0.001	0.121	0.062	0.002
% lo-dens urban	Corr Coeff					1	0.057	-0.129	0.23
	Sig (2-t)					.	0.712	0.399	0.129
% grass-pasture	Corr Coeff						1	0.011	0.13
	Sig (2-t)						.	0.94	0.393
% forest	Corr Coeff							1	0.039
	Sig (2-t)							.	0.801
% indus-trial	Corr Coeff								1
	Sig (2-t)								.

Table 8. Spearman rank correlational analyses of proportional landscape cover variables and water-quality variables (n = 45 for all correlations). Light yellow shows relationships with correlation coefficients between 0.3 and 0.5. Bright yellow shows relationships with correlation coefficients between 0.5 and 0.7. Orange shows relationships with correlation coefficients $\geq$ 0.7.

		% water	% commercial	% ag	% hi-density urban	% low-density urban	% grass pasture	% forested	% industrial
pH	Corr Coeff	.145	.172	.340*	.216	.204	.323*	-.295*	.018
	Sig (2-t)	.340	.257	.022	.154	.179	.030	.049	.906
Conductivity	Corr Coeff	.082	.292	.372*	.294*	.547**	.139	-.374*	.196
	Sig (2-t)	.592	.051	.012	.050	.000	.362	.011	.196
Gran Acid-Neut. Cap.	Corr Coeff	.107	.063	.465**	.061	.382**	.253	-.329*	.014
	Sig (2-t)	.483	.680	.001	.692	.010	.093	.027	.925
Turbidity (NTU)	Corr Coeff	-.087	.397**	.151	.239	.301*	.304*	.057	.330*
	Sig (2-t)	.569	.007	.322	.113	.044	.042	.710	.027
Water Color (PCU)	Corr Coeff	-.107	.086	.317*	.043	.178	.362*	.038	-.074
	Sig (2-t)	.483	.573	.034	.778	.243	.015	.806	.630
Total Organic Carbon (mg/L)	Corr Coeff	.292	.103	.276	.129	.167	.470**	-.302*	-.062
	Sig (2-t)	.052	.502	.067	.400	.272	.001	.044	.686
Dissolved Organic C	Corr Coeff	.294	.000	.319*	.039	.198	.425**	-.311*	-.129
	Sig (2-t)	.050	1.000	.033	.798	.191	.004	.038	.397
Total Phosphorus	Corr Coeff	-.213	.142	.253	-.002	.073	.277	.187	-.003
	Sig (2-t)	.160	.353	.093	.990	.633	.065	.220	.986
Total Nitrogen N (mg/L)	Corr Coeff	-.036	.250	.353*	.252	.095	.341*	-.324*	.139
	Sig (2-t)	.816	.098	.017	.095	.536	.022	.030	.362
Chloride (mg/L)	Corr Coeff	.221	.535**	.257	.563**	.491**	.130	-.440**	.236
	Sig (2-t)	.145	.000	.088	.000	.001	.395	.002	.119
Sulfate (mg/L)	Corr Coeff	-.076	.063	.248	.209	.257	-.239	-.134	.242
	Sig (2-t)	.619	.683	.101	.168	.088	.114	.379	.110
Calcium (mg/L)	Corr Coeff	.070	.035	.374*	.125	.352*	.147	-.383**	.026
	Sig (2-t)	.649	.817	.011	.412	.018	.334	.009	.863
Magnesium (mg/L)	Corr Coeff	-.061	.149	.604**	.253	.433**	-.096	-.391**	.129
	Sig (2-t)	.690	.328	.000	.094	.003	.532	.008	.398
Sodium (mg/L)	Corr Coeff	.196	.554**	.129	.554**	.466**	.137	-.266	.288
	Sig (2-t)	.196	.000	.399	.000	.001	.369	.078	.055
Potassium (mg/L)	Corr Coeff	-.006	.411**	.360*	.356*	.382**	.255	-.211	.210
	Sig (2-t)	.971	.005	.015	.016	.010	.090	.164	.166
Silica (mg/L SiO2)	Corr Coeff	-.326*	-.324*	.030	-.335*	-.122	-.050	.296*	-.090
	Sig (2-t)	.029	.030	.846	.025	.425	.745	.048	.557
CHLA (µg/L)	Corr Coeff	-.029	.223	.339*	.172	.206	.264	.000	.067
	Sig (2-t)	.852	.140	.023	.257	.174	.080	.999	.660
Watershed area	Corr Coeff	-.261	.398**	.452**	.255	.209	-.021	.109	.517**
	Sig (2-t)	.084	.007	.002	.091	.169	.893	.477	.000

Multivariate Analyses

Multivariate analyses of the NLA data were run to explore relationships among the variables and among the variable sets. We avoided entering correlated variables into the analyses, as this distorts results. As a result, the nutrient data set was reduced to total phosphorus, total nitrogen and potassium; the salts variables were reduced to chloride, calcium, and sulfate; the summary variables were reduced to gran acid-neutralizing capacity; the proportional landscape variables were reduced to forest, commercial, and ag/pasture.

Cluster analysis

Cluster analyses are entirely exploratory. Here, we show two examples of outcomes, one from clustering of the water-quality variables, and one from clustering of the landcover variables. Data were clustered, then major clusters were analyzed to determine how they differed on the major variables.

Cluster analysis of water-quality variables. The sample water-quality cluster used squared Euclidean distance and average linkage algorithms and produced 3 major clusters of lakes (Figure 8). Tests of group differences in water quality variables showed greatest differences among the cluster groups in calcium, gran acid-neutralizing capacity (Kruskal-Wallis and Welch's Robust ANOVA both producing $p < 0.0005$). Sulfate also varied significantly using both parametric and nonparametric tests, whereas total N and chloride only differed significantly using the Kruskal-Wallis test. Given the strongly nonnormal distributions of some variables, aspects of assumptions for both Kruskal-Wallis and Welch's Robust ANOVA were likely unmet for some tests.

Cluster 1 was associated with high gran acid-neutralizing capacity, N, Ca, and SO_4; lakes in this cluster were in the northern part of the state with the exception of one lake in Ripley County, in the southeast. Cluster 2 was associated with high N and intermediate gran acid-neutralizing capacity (ANC), Ca, and SO_4; lakes in this cluster were from the northern and central parts of the state. Cluster 3 was associated with low alkalinity, N, Ca, and SO_4; lakes in this cluster were widely scattered throughout the state.

Cluster analysis of landcover variables. The sample landcover cluster used squared Euclidean distance and average linkage algorithms and produced 4 major clusters of lakes (Figure 9). Of these, cluster 3 was most easily described, having high forest cover, and low ANC, P, N, and K (significantly so in all cases against groups 1 and 3, not consistently significantly so against group 4, due to sample size (3, for group 4)). Group 4 held the 3 sites with highest cover in grass/pasture (Figure 10), and Groups 1 and 2 had the points with 2nd highest and highest agricultural land cover, respectively. The highest N and P values were in group 4, rather than in the high-agriculture groups, but due in part to the small size of group 4, differences were not significant. Although the separation of clusters among landscape variables was good (Figure 10), the relationships of the clusters to water quality were not entirely clear.

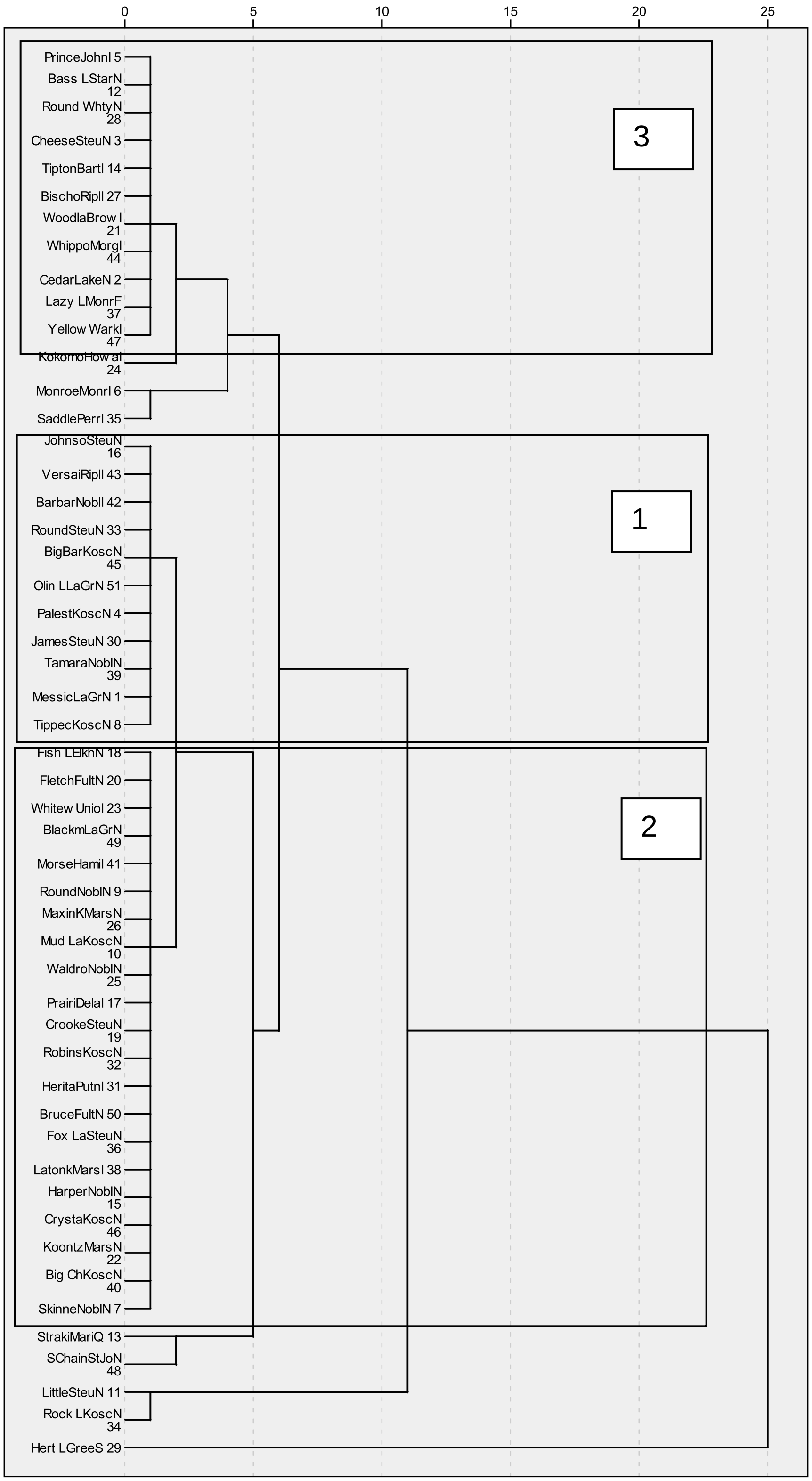


Figure 8. Cluster analysis of the NLA and reference lakes (n = 51). Cluster analysis used squared Euclidean distance and average linkage to assess clustering of water-quality variables. The major clusters showing greatest similarity are outlined and numbered (see text). Lakes are identified by partial name and county, and type of water body (primarily *natural* (N) and *impoundment* (I)).

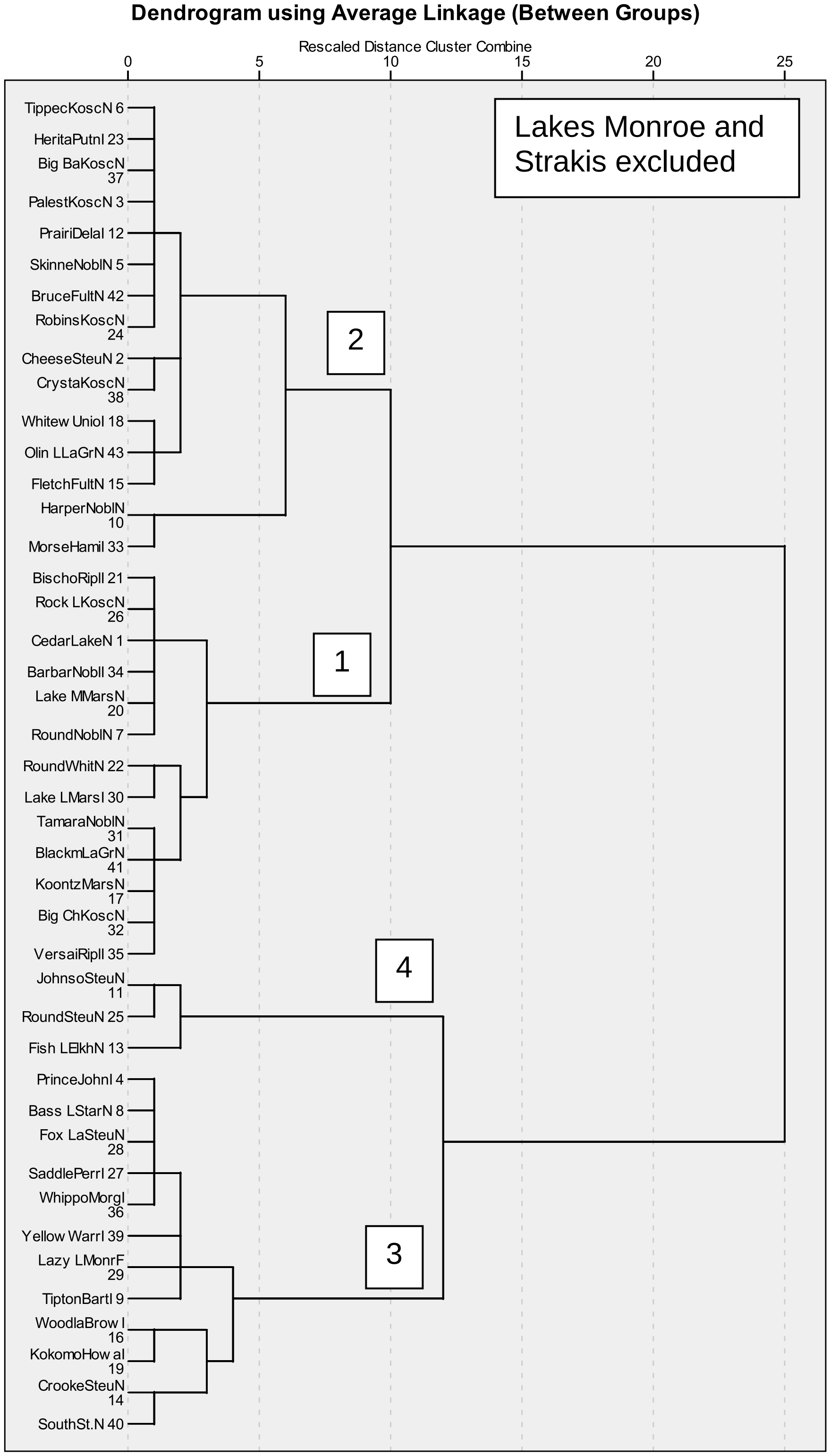


Figure 9. Cluster analysis (squared Euclidean distance and average linkage) using proportional landcover variables for agriculture, grass/pasture and commercial landcovers. Lakes are identified by partial name and county, and type of water body. The four major clusters are identified at their branch points. Two outlier lakes (Monroe and Strakis) were excluded.

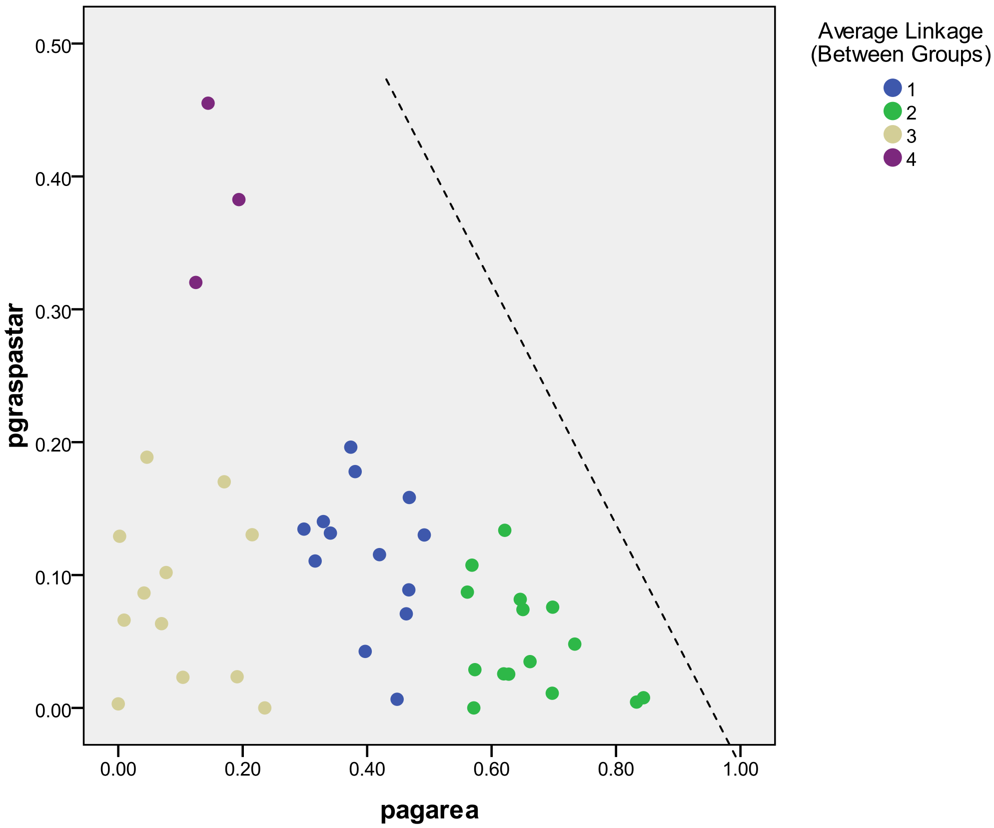


Figure 10. Proportion of agriculture and grass/pasture landcovers for NLA lakes, shown by membership in clusters derived from cluster-analysis of landscape variables. Note that landcover proportions must sum to one, so that as cover in grass/pasture increases, cover in agricultural area must decrease. Points cannot fall above the dashed line, and can only approach the line if the landscape has no additional cover types.

Bray-Curtis ordination

The first axis of the Bray-Curtis ordination including all 51 NLA and reference lakes was strongly correlated with gran ANC and calcium (Kendall's tau, τ (a nonparametric correlation coefficient similar to Spearman rank) = 0.738, 0.682) and to a lesser extent, with total N (τ = 0.539). A Kendall's tau value greater than 0.19 would be considered statistically significant for our sample size, but useful explanatory power (as with Pearson's correlation coefficient) is generally at higher levels – on the order of 0.30 or higher.

The analysis was rerun without calcium. Ca was highly correlated with ANC, and the inclusion of both variables could have contributed to the overriding effect of the pair on the analysis. The first axis of the new analysis extracted 65. 3% of the original matrix; Kendall's tau values were -0.735 for ANC and -0.539 for total N. Lake Monroe and Saddle Lake, both with very low ANC, increased the length of the axis considerably (Figure 11 – see the two points in the upper right); the ranking of lakes on this axis are shown in Appendix 2. The second, much weaker, axis was also highly correlated with ANC. The possibility of strong explanatory variables being correlated with more than one axis is relatively common in ordination. By analogy, in a linear regression, we would expect to model variables that are strongly correlated with multiple axes using squared or higher terms. Because the strong variables dominate multiple axes, they hide weaker relationships that might otherwise show up on axes after the first axis. In this case, however, given the proportion of the original matrix extracted, there may be relatively little explanatory power in the remaining variables.

Following up on this first ordination, we removed Lake Monroe and Saddle Lake from the analysis, to see if removing their lengthening effect on the first axis would yield a secondary axis with different correlations. The result was a first ordination axis with a strong correlation with ANC ($\tau = 0.679$) and with total N ($\tau = 0.539$) that extracted 71% of the original matrix. In the overall dataset, ANC and total N were moderately correlated ($r_s = 0.395$).

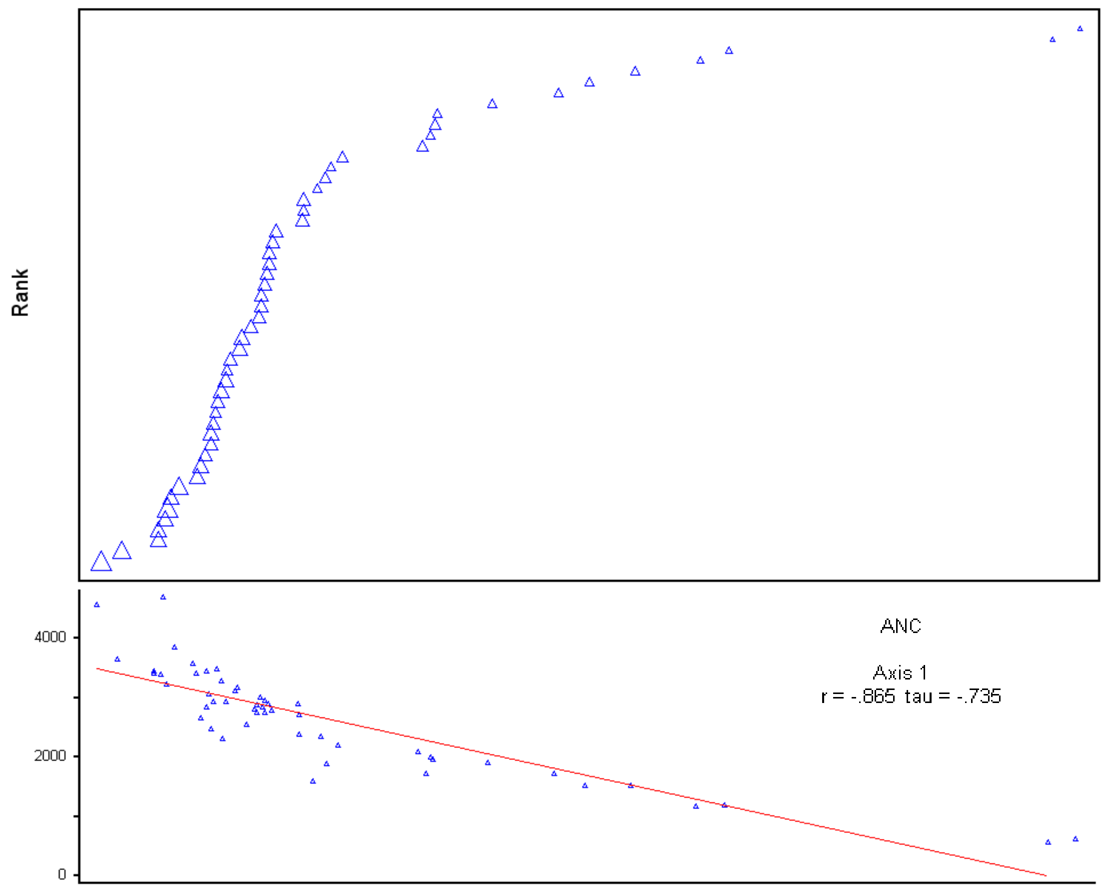


Figure 11. _First Bray-Curtis axis scores (top) and regression of axis-1 scores with acid-neutralizing capacity (ANC; bottom). The correlation coefficients are shown in an inset. Symbol size in the top graph is related to fit; small symbols indicate good fit to the axis.

To investigate effects from other variables, we removed ANC from the explanatory variables. The next run had a first axis that extracted only 31.5% of the original matrix. Axis 1 was strongly correlated with sulfate and chloride ($\tau = 0.579$ and 0.550). Strakis and South Chain Lakes were outliers (see Table 3), but the axis fit the remaining data cloud well. Because this ordination had so much less explanatory power than earlier ordinations, we did not pursue additional removal of variables.

Acid-neutralizing capacity dominated the correlations of the first axis from the Bray-Curtis ordinations, even when the two lakes with unusually low values were eliminated. Nitrogen values were moderately strongly correlated with ANC, but when ANC was eliminated from the explanatory values, nitrogen did not dominate the resulting analysis, and no single axis dominated the resulting ordination.

Nonmetric multidimensional ordination

Ordination constrained by the landscape variables was less successful in describing the data than the unconstrained Bray-Curtis ordination. Forty-five lakes had landscape compositions that HYMAP-OWLS could quantify, and these were used for nonmetric multidimensional ordination. The full data set used water quality variables ANC, total P, total N, K, and Cl, SO_4, and landscape variables total watershed area, proportion of commercial area, proportion of agricultural area, and proportion of grass/pasture area.

The best ordination of the full data set used only a single axis, which had strong correlation with total N ($\tau = -0.624$) and a weaker correlation with total P ($\tau = -0.349$). Cedar Lake, with an extremely high N value, was at the low end of the axis (the correlation was negative, so the low end corresponds to high N values; Figure 12). However, the correlations with the landcover variables were quite low (-0.188 for proportion of commercial area and -0.143 for proportion of grass and pasture; neither value was statistically significant). In addition, the overall variability explained by the model was only 9%. We experimented with modifications of the NMS ordination, but the constraint that water quality axes must be meaningful in the space of the landscape variables seemed unproductive overall.

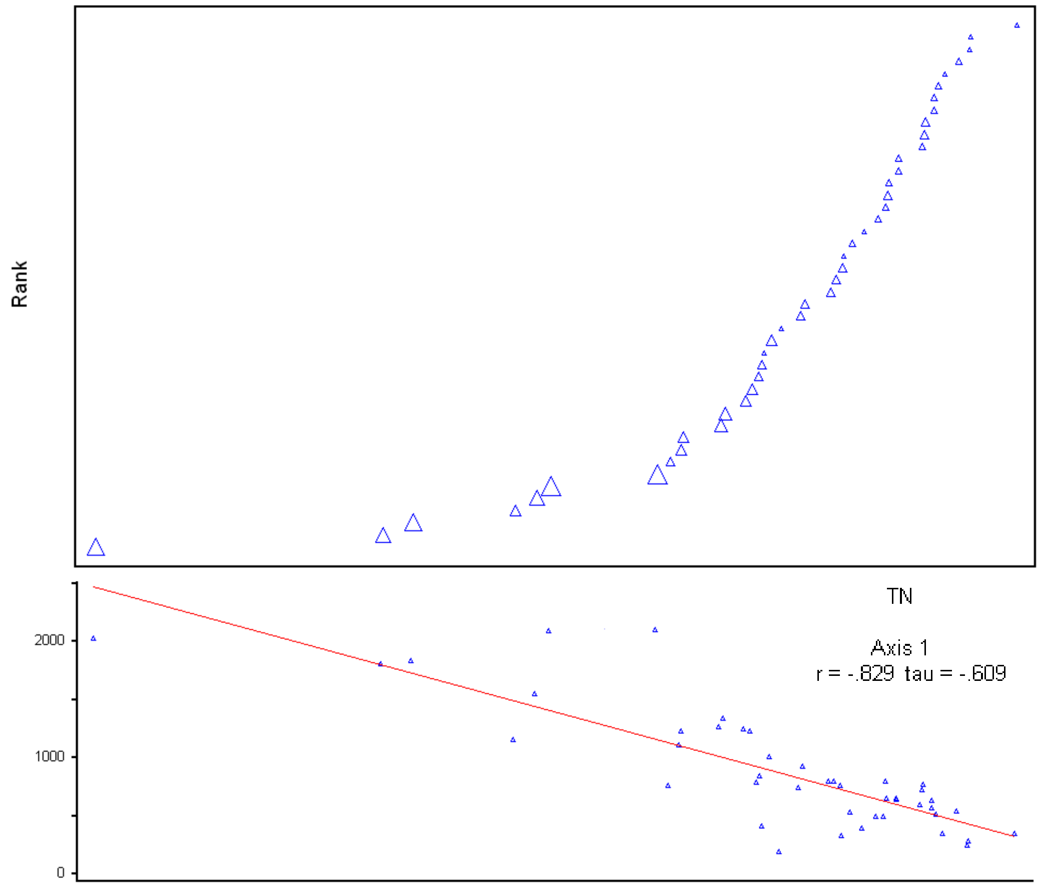


Figure 12. First axis from nonmetric multidimensional ordination (top) and regression with total N (bottom). The correlation coefficients are shown in an inset. Symbol size in the top graph is related to fit; small symbols indicate good fit to the axis.

DISCUSSION

Overview of Patterns in Results

Our results demonstrated many commonly observed patterns among water-quality variables, and between water-quality and land-cover variables. Not surprisingly, patterns related to agricultural run-off dominated several kinds of analyses and many locations around the state; Indiana lakes are denser in the unglaciated northern portions of the state, but lakes and agriculture occur throughout the state. Lakes showing signs of industrial pollution were less common than lakes showing agricultural pollution. Other findings were related to geological and watershed differences in the state. As indicated in methods, some of the variables measured during sampling were unavailable for analysis here. Data external to this study, related to water quality, particularly local precipitation information and lake management practices that undoubtedly help to explain some water-quality patterns also were not available for this analysis.

Extreme Data

The frequency distribution plots for the sample parameters illustrate the central tendency of the results. While this is useful in characterizing a population of lakes, examining the instances of extreme values can provide insight into specific lakes.

Three lakes (Cedar, Fish and Whitewater) had pH values above 9.0. In biologically productive lakes, the photosynthetic demand of algae for CO_2 (a weak acid) drives pH to the alkaline side of neutrality (Wetzel 2001). The NLA data show this relationship (Figure 3) where pH increases as the chlorophyll *a* concentration increases.

Cedar Lake, one of the more eutrophic lakes in Indiana, had extreme values on the high end of the turbidity, total phosphorus and chlorophyll *a* distributions, all characteristic of eutrophication. Versailles Lake had high levels of turbidity and total phosphorus but turbidity in this lake is more likely due to watershed soil erosion rather than due to algal production.

Strakis Lake had the highest number of extreme values for measured parameters: conductivity, ammonia, chloride, sulfate, sodium and potassium. This lake lies in a former limestone quarry and is also near a landfill and agricultural land, all of which likely contribute to its high ionic content. Hert Lake fills a coal surface mine end-cut and had extreme values for conductivity, sulfate and magnesium, all of which are characteristic of mine spoils.

Rock Lake also had extreme values in multiple parameters: acid-neutralizing capacity (ANC), total organic carbon (TOC) and dissolved organic carbon (DOC). None of these parameters are routinely measured in the Indiana Clean Lakes Program or in the Lake and River Enhancement (LARE) Program. A 2008 LARE study of Rock Lake (Richardson et al. 2009) concluded that the lake was well-oxygenated and mesotrophic, with high epilimnetic total phosphorus (128 µg/L) and chlorophyll *a* concentrations (57.8 µg/L). No findings from that study explain why the lake had extreme values in ANC, TOC, and DOC in the present NLA Study.

Olin Lake had the highest nitrate concentration (>1.5 mg/L) of all the NLA lakes sampled in Indiana. Indiana Clean Lake Program data show that for four assessments between 1993 and 2008, Olin Lake had a mean epilimnetic nitrate concentration of 1.208 mg/L. The NLA concentration is consistent with the CLP data. Olin Lake is surrounded by wetlands and has no homes along its shore. Its watershed drains primarily agricultural land but 87% of Olin Lake's watershed drains through Martin Lake, which would act as a settling basin (Pranckus et al. 2009). Martin Lake's epilimnetic nitrate concentration in a 2008 LARE study was 0.928 mg/L. We have no definitive explanation for Olin Lake's high nitrate concentration.

Comparison to CLP Data

The population of 51 NLA lakes sampled in Indiana had lower mean conductivity, total phosphorus, ammonium, and nitrate, lower TSI scores (Figure 4), and higher mean chlorophyll *a* than did the population of 370-380 lakes sampled between 2000 and 2005 by the Indiana Clean Lakes Program (CLP) (Table 4). The NLA data also categorized a higher percentage of lakes as oligotrophic (33%) using Carlson's TP TSI than did the same analysis of the CLP lake data (0%), and a lower proportion of eutrophic and hypereutrophic lakes when compared to the CLP data (Figure 7). These differences could be due to differences in lake selection, in sampling or analytical techniques, or in weather influencing the lakes' condition.

The CLP samples public lakes throughout Indiana that have boat launches. No probabilistic sampling design is used to select CLP lakes but rather, all accessible public lakes are sampled on a rotating schedule. The NLA draw included private and public lakes so the two sample populations are not exactly equivalent.

The differences between the two populations of lakes suggests that private lakes may have lower trophic states and mean values for conductivity, total phosphorus, ammonium and nitrate concentrations and thus their inclusion in the NLA may have driven down the means for these parameters. The lower means for total phosphorus (26.3 vs. 51.1 µg/L) and for nitrate (0.087 vs. 0.416 mg/L) are particularly striking in the NLA lakes. We investigated the possibility that ownership (private vs. public) might be related to water quality, but found no statistical evidence to support such a link.

Another factor that could cause the lower concentrations of TP and nitrate in the NLA samples is the long sample holding time prior to analysis. Many NLA samples, particularly TP and nitrate, had holding time alerts attached to the results to inform users that sample holding times were exceeded. When holding times are exceeded, the accuracy of the data could be compromised. In particular, phosphorus can be adsorbed onto the low-density polyethylene (LDPE) cubitainers used in the NLA study and nitrate can be lost through denitrification, a microbially mediated reaction whereby nitrate is reduced to molecular nitrogen (N_2). That analytical problems affected the results is suggested in the nitrate analyses where the nitrate + nitrate concentrations are lower than those of nitrate alone.

Precipitation during the 2007 NLA field sampling period was generally lower than during the years that the CLP data were sampled. Precipitation during the summer of 2007 was lower than summer precipitation in 2002, 2003, and 2004, and similar to patterns in 2000, 2001 and 2005

(Figure 6). Our observational data on Indiana lakes suggests that, for lakes with disturbed watersheds, drought results in less watershed runoff and better lake quality. However, we have no long-term climate and lake-quality data to evaluate statistically.

Correlation Analyses

The results of the correlational analyses showed that lake depth was strongly correlated with lake area and suggested an inverse relationship with total phosphorus, % forested area, and latitude, and suggested a positive relationship with % agricultural area. That deep lakes tend to have larger surface areas is not surprising. Glacial lakes in Indiana tend to be deeper than reservoir and the glacial lakes occur in the northern third of the state where forests and wetlands have been largely converted to agricultural land. Deep lakes also have more volume, which allows for more dilution of nutrients such as phosphorus. In an analysis of CLP data collected between 1994 and 1998, lake depth was inversely proportional to total Kjeldahl nitrogen, total phosphorus, and chlorophyll *a* with $p < 0.001$ (Jones and Barnes 2005) .

Our results also showed that summary variables associated with algal blooms – turbidity, water color, chlorophyll *a* – were generally highly correlated with the limiting nutrient variables C, N, and P. Phosphorus is most often the limiting nutrient in Indiana lakes and eutrophication has been associated with high concentrations of nitrogen and phosphorus (Montgrain and Jones, 2009). Indiana's eutrophic lakes also tend to have lower transparency due to increased turbidity caused by excessive algal densities and suspended sediments.

In the landscape analysis, most of the correlations between landcover and water quality variables were of lesser significance. Among the stronger correlations were sodium and chloride with commercial and high-density landcovers. The use of road salt during winter months to melt snow and ice has been associated with increased sodium and chloride concentrations in surface waters (Sassan and Kahl 2007). With greater density of roads and traffic in urbanized areas, it is not surprising that sodium and chloride concentrations in urbanized areas were found to be higher.

Agricultural landcover was positively correlated with DOC, total N and potassium but not with TOC or total phosphorus. While all of these nutrients are associated with agriculture and soil erosion, loss rates depend upon cropping practice, soil type, topography, and other variables. A recent study of nutrient losses from row-crop agriculture in northern Indiana showed a stronger relationship between nitrogen losses and land use class ($r^2 = 0.56 - 0.82$; $p < 0.05$) than with total phosphorus ($r^2 = 0.42$; $p < 0.05$) (Smith et al. 2008).

Multivariate Analysis

Ordination results revealed that acid neutralizing capacity is a major source of variability among Indiana lakes. Similarly, in the results of the cluster analysis, cluster 1, the most significant cluster, was associated with high gran acid-neutralizing capacity. Lakes within cluster 1 were primarily in the northern part of the state and this coincides with the part of Indiana affected by the most recent period of glaciation. The bedrock underlying the Northern Lake and Moraine Region is composed of siltstone, limestone and shale. The land surface is covered with thick deposits of glacial till composed of a diverse mix of sediments (Fenelon and Bobay 1994). The glacial till is layered with crushed calcium carbonate, a primary source of anions responsible for acid-neutralizing capacity. In fact, the amount of leaching of this calcium carbonate has been used to date the outwash sediments according to J. Robert Dodd (pers. comm. 2010).

Cluster 2 generally includes lakes in Central Indiana and is associated with high total nitrogen and intermediate gran-acid-neutralizing capacity. The predominant land use in the Central Indiana Corn Belt is row-crop agriculture, especially corn, which depends upon significant nitrogen fertilization.

Cluster analysis of land cover variables produced four clusters. Cluster 3 was most easily described, having high forest cover and low ANC, P, N, and K. These results are consistent with the low P, N and K losses associated with forest land cover.

<u>Sampling Considerations for Following up on Patterns Observed in this Study</u>

Acid-neutralizing capacity

Acid-neutralizing capacity, which defined the primary axis of ordination results, is related to soil chemistry. In Indiana, major regional variation in soil chemistry is linked to bedrock characteristics and glacial history. The state's major watersheds cross, rather than define, these regions. Omernik's ecoregions (1987) coincide well with some of the regions – the karst-rich south-central and southeast areas, the glacial lakes of the northeast, and the deep soils of the till plain. However, lakes are not evenly divided among the ecoregions, and the NLA sampling regime within Indiana did not involve sufficient lakes to provide good coverage in the areas with lower lake density. As a result, effective statistical comparison of ecoregions was prevented by small sample sizes.

In view of the importance of soil- and bedrock-related characteristics, future studies might use stratified sampling to ensure a minimum sample size in the strata of interest. The Omernik ecoregions or Homoya's natural regions of Indiana (Homoya 1985) provide divisions based in part on topographic and glacial characteristics that may be relevant. Similarly, soil maps that provide information on pH and alkalinity could be used to characterize soils in the immediate vicinity of lakes. Understanding soil variation within ecoregions may best explain some of the similarities among lakes from disparate parts of the state.

Nutrient relationships

Nutrient relationships were apparent in correlational analyses and cluster analysis, but did not define strong axes during ordination. Many nutrient variables are too confounded by variation in algal growth cycles at a variety of scales to be easily measured by one or two visits in a season. Daily variation from photosynthetic activity, seasonal variation in algal growth, and variation related to storm-caused changes in run-off and sediment all combine to make it difficult to properly characterize the nutrient status of lakes. Unfortunately, storm events need not be large to create spikes in run-off, and few lakes have sufficiently local weather stations to accurately record weather immediately preceding sampling. Possibly, discussion with local residents could be used for a relatively informal measure such as occurrence of strong rains in the week preceding sampling or onsite evidence of elevated turbidity, etc. Variables chosen for ordination and other multivariate analyses should be those least affected by flood pulses – total N, rather than individual forms of N, and in preference to total P, which is affected by sediment inflow. Droughts, more likely under future climate change, can result in reduced nutrient and sediment loading in those lakes dominated by watershed runoff.

Landscape analysis

The HYMAP-OWLS analytical tool used to provide landscape information in this study summarizes land cover within a single watershed. However, due to marl mining in the late 1800s and early 1900s (Blatchley and Ashley 1900), lake level manipulations by farmers in the early 1900s, the Indiana Drainage Code of 1965, and the creation of artificial channels to connect residences to lakes, many Indiana lakes are now chained together by channels. The importance of canal inflow in determining water quality obviously varies with flow volume and difference in water quality among lakes in such a chain. As a result, landscape impacts are smeared, with upstream basins potentially affecting downstream lakes. A more sophisticated landscape analysis would reflect the total watershed area contributing to lakes, factoring in the relative importance of upper basins to the downstream waters. Most between-lake connecting channels are not instrumented, and flow may be a function not only of weather but also of management (e.g., by water-control structures), thus, determining when to consider upstream watersheds is difficult, but, where possible, visual inspection may provide at least basic guidance.

During this study, landscape analysis of 6 lakes could not be completed due to extreme lack of topography in the surrounding landscape, which prevented the program from delineating a relevant watershed. In such landscapes, ditches may more strongly affect lake water quality than overland flow. In less hydrologically manipulated landscapes, sheet flow may be important, but in the primary agricultural landscapes, drainage ditches minimize such flow and substitute for it.

Any analysis of landscape features affecting lakes in agricultural regions must account for subsurface drainage tiles. However, the location of these tiles is not well-document and even fewer have been evaluated regarding water quality. In one comprehensive study of the Leary-Weber Ditch agricultural watershed west of Indianapolis (Baker et al. 2006), drainage tiles accounted for substantial nutrient loadings to the receiving stream.

Limitations of statistics and sampling

Statistical analysis is predicated on the use of a finite sample to characterize an infinite universe. Indiana lakes are numerous, and as a single group, more than sufficiently numerous to adequately approximate infinity. However, some subsets are distinctly less numerous. Large reservoirs and lakes strongly affected by certain kinds of industrial pollutants represent two such subsets. Large reservoirs are of sufficient interest that lake-by-lake individual description may be most useful. Water quality of any large single reservoir may be a result of such a large watershed and of so many separate inputs that the reservoirs do not fit typologies or statistical methods designed to describe or categorize smaller lakes. Some highly polluted lakes (e.g., Strakis) are sufficiently idiosyncratic that they constitute outliers for most kind of analyses, and potentially confound analysis that seeks to categorize or ordinate less exceptional lakes.

Where obviously finite and distinct subgroups can be identified, separate analyses, or lake-by-lake description will likely be more appropriate than inclusion in a larger analysis. Random sampling that seeks to characterize more usual lakes should probably avoid lakes in these subgroups. In studying a collection as heterogeneous as Indiana lakes, a single, unstratified sampling scheme is unlikely to satisfy all analytical aims. When possible, research questions should be identified that can guide the sampling scheme or schemes, or, at least, subsets of lakes such as those described above might be identified and characterized so that they can be excluded from a simple sampling scheme.

REFERENCES CITED

Baker, N.T., W.W. Stone, J.T. Wilson and M.T. Meyer. 2006. Occurrence and Transport of Agricultural Chemicals in Leary Weber Ditch Basin, Hancock County, Indiana, 2003–04. Scientific Investigations Report 2006–5251, U.S. Department of the Interior, U.S. Geological Survey, Indianapolis.

Blatchley, W.S. and G.H. Ashley. 1900. The Lakes of Northern Indiana and their Associated Marl Deposits. Indiana Department of Geology and natural Resources, Twenty-Fifth Annual Report, Indianapolis.

Choi, Jin-Yong and Bernard A. Engel. 2009. Watershed Delineation Watershed Generation. Agricultural and Biological Engineering Department, Purdue University. http://cobweb.ecn.purdue.edu/~watergen/owls/

Dodd, J.R. 2010. Conversation with author, 5 March 2010.

Fenelon, J.M., K.E. Bobay and others. 1994. Hydrogeologic Atlas of Aquifers in Indiana. Water-Resources Investigations Report 92-4142. U.S. Geological Survey, Indianapolis.

Homoya, M.A., D.B. Abrell, J.R. Aldrich, and T.W. Post. 1985. The natural regions of Indiana. Proceedings of the Indiana Academy of Sciences 94:245-268.

Jones, W.W. and B. Barnes. 2005. Indiana Lake Water Quality Assessment Report for 1994-1998. School of Public & Environmental Affairs, Indiana University, Bloomington.

McCune, B. and J.B. Grace. 2002. Ecological analysis of ecological communities. MjM Software, Gelneden Beach, Oregon, USA.

McCune, B. and M.J. Mefford. 2006.PC-ORD. Multivariate analysis of ecological data. Version 5. MjM Software, Gleneden Beach, Oregon, USA.

Montgrain, L.A. and W.W. Jones. 2009. Indiana Lake Water Quality Assessment Report for 2004-2008. School of Public & Environmental Affairs, Indiana University, Bloomington.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). Annals of the Association of American Geographers 77(1):118-125.

Pranckus, M., W.W. Jones, T. Parr and others. 2009. Olin, Oliver and Martin Lakes Diagnostic Study. Prepared for Oliver and Martin Lakes Improvement Association and Indiana Department of Natural Resources. J.F. New & Associates, Inc., Walkerton, IN. 280 pp.

Richardson, J., W.W. Jones, and others. 2009. Rock Lake Diagnostic Study. Prepared for Rock Lake Conservation and Improvement Association and Indiana Department of Natural Resources. J.F. New & Associates, Inc., Walkerton, IN. 150 pp.

Sassan, D. and S. Kahl. 2007. Beaver Brook/ Policy Brook I-93 Chloride TMDL: Salt Loading Due to Private Winter Maintenance Practices. Plymouth State University, Center for the Environment, Plymouth, NH.

Smith, D.R., S.J. Livingston, B.W. Zuercher, M. Larose, G.C. Heathman and C. Huang. 2008. Nutrient losses from row crop agriculture in Indiana. J. Soil Water Cons., 63(6): 396-409.

Smith, V.H. 1983. The nitrogen and phosphorus dependence of algal biomass in lakes: An empirical and theoretical analysis. *Limnol. Oceanogr.*, 27(6): 1982:1101-1112.

USEPA. 2007. Survey of the Nation's Lakes. Field Operations Manual. EPA 841-B-07-004. U.S. Environmental Protection Agency, Washington, DC. http://cobweb.ecn.purdue.edu/~watergen/.

Appendix 1:

Data flags for all flagged variables used in analysis. Data flags indicate conditions of potential concern, usually with regard to handling time.

Flag for pH_Lab

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		40	78.4	78.4	78.4
	H5	2	3.9	3.9	82.4
	H6	1	2.0	2.0	84.3
	H7	1	2.0	2.0	86.3
	H8	7	13.7	13.7	100.0
	Total	51	100.0	100.0	

Flag for Conductivity

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		46	90.2	90.2	90.2
	H10	1	2.0	2.0	92.2
	H11	1	2.0	2.0	94.1
	H8	3	5.9	5.9	100.0
	Total	51	100.0	100.0	

Flag for Gran Acid-Neutralizing Capacity

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		44	86.3	86.3	86.3
	H8	7	13.7	13.7	100.0
	Total	51	100.0	100.0	

Flag for Turbidity

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		31	60.8	60.8	60.8
	< RL (2)	17	33.3	33.3	94.1
	H4	2	3.9	3.9	98.0
	H4; < RL (2)	1	2.0	2.0	100.0
	Total	51	100.0	100.0	

Flag for Total Organic Carbon

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		10	19.6	19.6	19.6
	H15	9	17.6	17.6	37.3
	H16	5	9.8	9.8	47.1
	H17	4	7.8	7.8	54.9
	H18	4	7.8	7.8	62.7
	H19	5	9.8	9.8	72.5
	H20	5	9.8	9.8	82.4
	H21	3	5.9	5.9	88.2
	H23	1	2.0	2.0	90.2
	H244	1	2.0	2.0	92.2
	H26	1	2.0	2.0	94.1
	H71	2	3.9	3.9	98.0
	H72	1	2.0	2.0	100.0
	Total	51	100.0	100.0	

Flag for Dissolved Organic Carbon

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		46	90.2	90.2	90.2
	H15	2	3.9	3.9	94.1
	H153	1	2.0	2.0	96.1
	H16	2	3.9	3.9	100.0
	Total	51	100.0	100.0	

Flag for Total Phosphorus

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		7	13.7	13.7	13.7
	H29	4	7.8	7.8	21.6
	H30	3	5.9	5.9	27.5
	H34	2	3.9	3.9	31.4
	H35	4	7.8	7.8	39.2
	H36	7	13.7	13.7	52.9
	H37	3	5.9	5.9	58.8
	H38	2	3.9	3.9	62.7
	H39	1	2.0	2.0	64.7
	H40	1	2.0	2.0	66.7
	H44	2	3.9	3.9	70.6
	H45	2	3.9	3.9	74.5
	H46	2	3.9	3.9	78.4
	H47	2	3.9	3.9	82.4
	H47; < RL (4)	1	2.0	2.0	84.3
	H85	2	3.9	3.9	88.2
	H86; < RL (4)	1	2.0	2.0	90.2
	H87	1	2.0	2.0	92.2
	H90	2	3.9	3.9	96.1

	H91	2	3.9	3.9	100.0
	Total	51	100.0	100.0	

Flag for TN

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		7	13.7	13.7	13.7
	H29	4	7.8	7.8	21.6
	H30	3	5.9	5.9	27.5
	H34	2	3.9	3.9	31.4
	H35	4	7.8	7.8	39.2
	H36	6	11.8	11.8	51.0
	H37	3	5.9	5.9	56.9
	H38	2	3.9	3.9	60.8
	H39	1	2.0	2.0	62.7
	H40	1	2.0	2.0	64.7
	H44	2	3.9	3.9	68.6
	H45	2	3.9	3.9	72.5
	H46	2	3.9	3.9	76.5
	H47	2	3.9	3.9	80.4
	H84	1	2.0	2.0	82.4
	H85	2	3.9	3.9	86.3
	H86	1	2.0	2.0	88.2
	H87	1	2.0	2.0	90.2
	H90	2	3.9	3.9	94.1
	H91	2	3.9	3.9	98.0
	H95	1	2.0	2.0	100.0
	Total	51	100.0	100.0	

Flag for NH4

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		38	74.5	74.5	74.5
	< RL (0.02)	8	15.7	15.7	90.2
	H34	1	2.0	2.0	92.2
	H35	1	2.0	2.0	94.1
	H35; < RL (0.02)	1	2.0	2.0	96.1
	H36	1	2.0	2.0	98.0
	H37	1	2.0	2.0	100.0
	Total	51	100.0	100.0	

Flag for N03_N02

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		46	90.2	90.2	90.2
	H8	4	7.8	7.8	98.0
	H9	1	2.0	2.0	100.0
	Total	51	100.0	100.0	

Flag for NO3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		2	3.9	3.9	3.9
	H101	1	2.0	2.0	5.9
	H104	1	2.0	2.0	7.8
	H105	1	2.0	2.0	9.8
	H107	2	3.9	3.9	13.7
	H108	1	2.0	2.0	15.7
	H109	1	2.0	2.0	17.6
	H110	1	2.0	2.0	19.6
	H111 ; < RL (0.0)	1	2.0	2.0	21.6

H112	1	2.0	2.0	23.5
H113	1	2.0	2.0	25.5
H119	1	2.0	2.0	27.5
H119 ; < RL (0.0)	1	2.0	2.0	29.4
H126	2	3.9	3.9	33.3
H131	2	3.9	3.9	37.3
H132	2	3.9	3.9	41.2
H140	1	2.0	2.0	43.1
H140 ; < RL (0.0)	1	2.0	2.0	45.1
H141 ; < RL (0.0)	1	2.0	2.0	47.1
H142	2	3.9	3.9	51.0
H150	1	2.0	2.0	52.9
H151	2	3.9	3.9	56.9
H152	2	3.9	3.9	60.8
H153	4	7.8	7.8	68.6
H154	2	3.9	3.9	72.5
H155	1	2.0	2.0	74.5
H161	1	2.0	2.0	76.5
H165	1	2.0	2.0	78.4
H165 ; < RL (0.0)	2	3.9	3.9	82.4
H94 ; < RL (0.02)	2	3.9	3.9	86.3
H96 ; < RL (0.02)	1	2.0	2.0	88.2
H97	1	2.0	2.0	90.2
H97 ; < RL (0.02)	1	2.0	2.0	92.2
H99	2	3.9	3.9	96.1
H99 ; < RL (0.02)	2	3.9	3.9	100.0
Total	51	100.0	100.0	

Flag for CL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		1	2.0	2.0	2.0
	H100	3	5.9	5.9	7.8
	H101	1	2.0	2.0	9.8
	H102	2	3.9	3.9	13.7
	H104	2	3.9	3.9	17.6
	H106	1	2.0	2.0	19.6
	H107	3	5.9	5.9	25.5
	H109	1	2.0	2.0	27.5
	H110	1	2.0	2.0	29.4
	H111	1	2.0	2.0	31.4
	H112	2	3.9	3.9	35.3
	H113	2	3.9	3.9	39.2
	H119	1	2.0	2.0	41.2
	H126	1	2.0	2.0	43.1
	H129	2	3.9	3.9	47.1
	H131	2	3.9	3.9	51.0
	H133	2	3.9	3.9	54.9
	H140	1	2.0	2.0	56.9
	H141	1	2.0	2.0	58.8
	H142	2	3.9	3.9	62.7
	H151	2	3.9	3.9	66.7
	H152	1	2.0	2.0	68.6
	H153	3	5.9	5.9	74.5
	H154	1	2.0	2.0	76.5
	H155	4	7.8	7.8	84.3
	H156	1	2.0	2.0	86.3
	H165	1	2.0	2.0	88.2
	H168	1	2.0	2.0	90.2

H169	2	3.9	3.9	94.1
H170	1	2.0	2.0	96.1
H97	2	3.9	3.9	100.0
Total	51	100.0	100.0	

Flag for SIO2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		46	90.2	90.2	90.2
	H8	4	7.8	7.8	98.0
	H9	1	2.0	2.0	100.0
	Total	51	100.0	100.0	

Flag for Na

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		1	2.0	2.0	2.0
	H181	1	2.0	2.0	3.9
	H182	2	3.9	3.9	7.8
	H183	2	3.9	3.9	11.8
	H280	1	2.0	2.0	13.7
	H281	2	3.9	3.9	17.6
	H282	4	7.8	7.8	25.5
	H283	1	2.0	2.0	27.5
	H286	4	7.8	7.8	35.3
	H287	7	13.7	13.7	49.0
	H288	3	5.9	5.9	54.9
	H289	3	5.9	5.9	60.8
	H290	1	2.0	2.0	62.7
	H291	2	3.9	3.9	66.7

H292	1	2.0	2.0	68.6
H293	4	7.8	7.8	76.5
H295	1	2.0	2.0	78.4
H296	2	3.9	3.9	82.4
H297	3	5.9	5.9	88.2
H298	3	5.9	5.9	94.1
H302	1	2.0	2.0	96.1
H315	1	2.0	2.0	98.0
H323	1	2.0	2.0	100.0
Total	51	100.0	100.0	

Flag for SO4

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		2	3.9	3.9	3.9
	H100	2	3.9	3.9	7.8
	H101	1	2.0	2.0	9.8
	H102	1	2.0	2.0	11.8
	H104	1	2.0	2.0	13.7
	H106	1	2.0	2.0	15.7
	H107	4	7.8	7.8	23.5
	H108	1	2.0	2.0	25.5
	H109	1	2.0	2.0	27.5
	H110	1	2.0	2.0	29.4
	H111	1	2.0	2.0	31.4
	H112	1	2.0	2.0	33.3
	H113	1	2.0	2.0	35.3
	H119	2	3.9	3.9	39.2
	H126	2	3.9	3.9	43.1
	H131	2	3.9	3.9	47.1
	H132	2	3.9	3.9	51.0

	H140	2	3.9	3.9	54.9
	H142	1	2.0	2.0	56.9
	H151	2	3.9	3.9	60.8
	H152	2	3.9	3.9	64.7
	H153	3	5.9	5.9	70.6
	H154	2	3.9	3.9	74.5
	H155	2	3.9	3.9	78.4
	H156	1	2.0	2.0	80.4
	H157	1	2.0	2.0	82.4
	H161	1	2.0	2.0	84.3
	H165	3	5.9	5.9	90.2
	H169	1	2.0	2.0	92.2
	H94	1	2.0	2.0	94.1
	H96	1	2.0	2.0	96.1
	H97	1	2.0	2.0	98.0
	H99	1	2.0	2.0	100.0
	Total	51	100.0	100.0	

Flag for Ca

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		1	2.0	2.0	2.0
	H181	1	2.0	2.0	3.9
	H182	2	3.9	3.9	7.8
	H183	2	3.9	3.9	11.8
	H280	1	2.0	2.0	13.7
	H281	2	3.9	3.9	17.6
	H282	4	7.8	7.8	25.5
	H283	1	2.0	2.0	27.5
	H286	4	7.8	7.8	35.3
	H287	7	13.7	13.7	49.0

H288	3	5.9	5.9	54.9
H289	4	7.8	7.8	62.7
H290	1	2.0	2.0	64.7
H291	2	3.9	3.9	68.6
H292	1	2.0	2.0	70.6
H293	5	9.8	9.8	80.4
H295	1	2.0	2.0	82.4
H296	2	3.9	3.9	86.3
H297	3	5.9	5.9	92.2
H298	3	5.9	5.9	98.0
H302	1	2.0	2.0	100.0
Total	51	100.0	100.0	

Flag for Mg

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		1	2.0	2.0	2.0
	H181	1	2.0	2.0	3.9
	H182	2	3.9	3.9	7.8
	H183	2	3.9	3.9	11.8
	H280	1	2.0	2.0	13.7
	H281	2	3.9	3.9	17.6
	H282	4	7.8	7.8	25.5
	H283	1	2.0	2.0	27.5
	H286	4	7.8	7.8	35.3
	H287	7	13.7	13.7	49.0
	H288	3	5.9	5.9	54.9
	H289	3	5.9	5.9	60.8
	H290	1	2.0	2.0	62.7
	H291	2	3.9	3.9	66.7
	H292	1	2.0	2.0	68.6

H293	5	9.8	9.8	78.4
H295	1	2.0	2.0	80.4
H296	2	3.9	3.9	84.3
H297	3	5.9	5.9	90.2
H298	3	5.9	5.9	96.1
H302	1	2.0	2.0	98.0
H315	1	2.0	2.0	100.0
Total	51	100.0	100.0	

Flag for K

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid		1	2.0	2.0	2.0
	H181	1	2.0	2.0	3.9
	H182	2	3.9	3.9	7.8
	H183	2	3.9	3.9	11.8
	H280	1	2.0	2.0	13.7
	H281	2	3.9	3.9	17.6
	H282	4	7.8	7.8	25.5
	H283	1	2.0	2.0	27.5
	H286	4	7.8	7.8	35.3
	H287	7	13.7	13.7	49.0
	H288	3	5.9	5.9	54.9
	H289	4	7.8	7.8	62.7
	H290	1	2.0	2.0	64.7
	H291	2	3.9	3.9	68.6
	H292	1	2.0	2.0	70.6
	H293	4	7.8	7.8	78.4
	H295	1	2.0	2.0	80.4
	H296	2	3.9	3.9	84.3
	H297	3	5.9	5.9	90.2

H298	3	5.9	5.9	96.1
H302	1	2.0	2.0	98.0
H323	1	2.0	2.0	100.0
Total	51	100.0	100.0	

Appendix 2:

Lake scores on Axis 1 of a Bray-Curtis Ordination using acid-neutralizing capacity, total phosphorus, total nitrogen, Cl, SO_4, and K as lake characteristics. The axis was highly negatively correlated with acid-neutralizing capacity and moderately highly negatively correlated with total nitrogen.

Score on Axis 1	Lake ID	Lake Name
-0.051	NLA06608-1867	Rock Lake
-0.035	NLA06608-IN:646	Olin Lake
-0.006	NLA06608-2644	Versailles Lake
-0.005	NLA06608-2987	South Chain Lake
0	NLA06608-0811	Johnson Lake
0.002	NLA06608-0555	Little Otter
0.005	NLA06608-2523	Barbara Lake
0.011	NLA06608-0139	Palestine Lake
0.026	NLA06608-2779	Big Barbee Lake
0.029	NLA06608-2267	Tamarack Lake
0.032	NLA06608-0660	Strakis Lake
0.037	NLA06608-0043	Messick Lake
0.037	NLA06608-0731	Harper Lake
0.039	NLA06608-1243	Waldron Lake
0.04	NLA06608-0219	Skinner Lake
0.043	NLA06608-3147	Bruce Lake
0.045	NLA06608-1579	James Lake
0.049	NLA06608-0235	Tippecanoe Lake
0.05	NLA06608-0987	Fish Lake
0.053	NLA06608-0299	Round Lake
0.06	NLA06608-0491	Mud Lake
0.061	NLA06608-1835	Round Lake
0.069	NLA06608-2507	Morse Reservoir
0.076	NLA06608-2891	Crystal Lake
0.077	NLA06608-0971	Prairie Creek
0.077	NLA06608-2155	Fox Lake
0.08	NLA06608-1556	Hert Lake
0.081	NLA06608-2219	Lake Latonka
0.084	NLA06608-1195	Koontz Lake
0.084	NLA06608-2283	Big Chapman
0.086	NLA06608-1355	Lake Maxinkuckee
0.09	NLA06608-1755	Robinson Lake
0.11	NLA06608-1684	Heritage Lake
0.111	NLA06608-1199	Whitewater Lake
0.112	NLA06608-1131	Crooked Lake

0.123	NLA06608-0091	Cedar Lake
0.129	NLA06608-3035	Blackman Lake
0.133	NLA06608-0587	Bass Lake
0.142	NLA06608-1163	Fletcher Lake
0.207	NLA06608-1227	Kokomo Reservoir No. 1
0.213	NLA06608-1492	Bischoff Reservoir
0.217	NLA06608-1499	Round Lake
0.219	NLA06608-0107	Cheeseboro Lake
0.263	NLA06608-0148	Princes East
0.316	NLA06608-0724	Tipton Lakes
0.342	NLA06608-2708	Whippoorwill Lake
0.378	NLA06608-1172	Woodland Lake
0.43	NLA06608-2196	Lazy Lake
0.453	NLA06608-2916	Yellow Banks
0.713	NLA06608-1908	Saddle Lake
0.736	NLA06608-0212	Monroe Lake