U.S. Environmental Protection Agency
Office of Water/Office of Research and Development
Washington, D.C. 20460
EPA 841-R-09-001a

National Lakes Assessment: Technical Appendix

Data Analysis Approach

January 2010

National Lakes Assessment: Technical Appendix

Data Analysis Approach

U.S. Environmental Protection Agency
Office of Water
Office of Research and Development

January 2010

EPA 841-R009-001a

National Lakes Assessment: Technical Appendix

Data Analysis Approach

Table of Contents

Overview

This document provides additional information to supplement the results and discussion presented in the 2007 National Lakes Assessment (NLA). It is intended to serve as a more technical reference than the report itself on the conceptual basis and the methods and procedures used for the NLA. Although it is intended to provide a comprehensive summary of these procedures, it is not intended to present additional data analysis results or an in-depth report of the design, sampling, or analysis protocol. For additional details, citations are provided.

Objectives of the NLA

The objective of the NLA is to characterize the ecological condition of the nation's lakes throughout the conterminous United States. The NLA is an ecological assessment of lakes based on chemical, physical, and biological data. It employs a statistically-valid probability design stratified to allow estimates of the condition of streams on a national and regional scale. The two key questions the NLA addresses are

- To what degree are the Nation's lakes in good, fair, and poor condition?
- What is the relative importance of the different stressors evaluated in the NLA?

The NLA is a collaboration among the U.S. Environmental Protection Agency (EPA), states, tribal nations, U.S. Geological Survey (USGS), and other partners. It is intended as a document for the public and Congress. It is not a technical document, but rather a report geared towards a broad audience. This Technical Addendum is a supplemental document used to support the results in the NLA report. It describes the process used to collect, evaluate, and analyze data for the NLA. It outlines steps taken to assess the biological condition of the nation's freshwater resources and identify the relative impact of stressors on this condition. Results from the analysis are included in this 2007 NLA Report; the data collected and methods described will continue to be studied and used for future analyses.

The NLA data analysis procedures described in this technical report were developed from the input and experience of the participating cooperators and technical experts. A small workgroup was held in the spring of 2009 to consider approaches for data analysis. Findings from this workshop were presented at the larger group of cooperators and lake managers at the National Lakes Meeting in April 2009. Here, state agencies, universities, non-profits, and EPA participated in a one day workshop where they discussed topics such as analysis options, data presentation, and reference sites. Discussions from this meeting were used to help define the steps taken for the data analysis presented in the final report.

NLA analysts used two processes for establishing the good/fair/poor findings in the NLA report. For trophic status and recreational indicators, alysts used fixed, nationally consistent thresholds. This approach is not covered in detail in this Techincal Addendum. The second approach was to establish regionally consistent reference-based thresholds. Detailed information on this approach is presented below.

Reference Condition

To assess current ecological condition, it is necessary to compare measurements today to an estimate of “good” quality. Setting reasonable expectations for each indicator was one of the greatest challenges for the NLA analysts. Because of the difficulty in estimating historical conditions for many NLA indicators, the 2009 NLA used “least-disturbed condition” as the reference condition. Least-disturbed condition can be defined as the best available chemical, physical, and biological habitat conditions given the current state of the landscape – or “the best of what’s left” (Stoddard et al. 2006). Data from reference sites were used to develop seven regional specific reference conditions against which test results could be compared.

Sources of Reference Sites

Sites sampled during the NLA index period using consistent sampling protocols and analytical methods were screened to meet regional specific physical and chemical criteria. These included both sites selected from the probability sample sites and an additional 124 hand-picked sites thought to be reference by best professional judgment. Like the probability sample sites, the hand-picked sites were sampled using the NLA methods. These sites were obtained from a number of sources. Some states submitted their best reference sites to be sampled as part of the NLA while other sites from the west and northeast were selected in a prescreening analysis utilizing landuse to find least-disturbed lake watersheds. Regardless of whether sites were probability-based or hand-selected, only those that met the final screening criteria were used in developing the reference condition.

Screening NLA Site Data for Biological Reference Condition

Prior to identifying reference lakes, all lakes from the NLA were grouped into distinct regional clusters based on nine environmental variables. This clustering was undertaken in order to identify regional reference lakes. These variables took into account geographic and geologic differences such as elevation, precipitation, air temperature, longitude, latitude, and calcium concentrations. In addition to these geographic/geologic variables, other variables such as lake area, depth, and shoreline development were also used to segregate lakes.

Seven regional clusters were identified during this process, and these seven regions were grouped into one of three larger regions, eastern highlands (EHIGH; which constituted the Appalachians and the Northeast), plains and lowlands (PLNLOW; which constituted the coastal plains, northern and southern plains, and Midwest), and western mountains (WMNTS; which constituted the western mountains and xeric region of the west). The PLNLOW region, which was the largest of the three combined regions, was stratified along 40 degree latitude to insure that reference sites south of the upper Midwest would be included in the analysis. It is important to keep in mind that the seven regional clusters were identified to group like lakes for purposes of identifying regional reference lakes, but are different from the NLA reporting regions. Lakes from more than one lake cluster can and does exist within the reporting regions.

To identify biological reference sites for purposes of the NLA, analysts used the chemical and physical data collected at each site to determine whether any given site is in least-disturbed condition for its region. In the NLA, screening values were established for ten chemical and physical parameters to screen for reference sites. These parameters included total nitrogen, total phosphorus, chloride, sulfate, turbidity, euphotic zone dissolved oxygen, acid-neutralizing capacity, shoreline disturbance by agriculture, shoreline disturbance by non-agriculture, and shoreline disturbance intensity and extent. If a site exceeded the screening value for any one stressor, it was dropped from reference consideration.

Given that expectations of least-disturbed condition vary across regions, the criteria values for exclusion varied by region. The seven aggregate reference clusters developed for the NLA used regionalize biological reference condition thresholds (Table A-1). The first threshold value in each of the 10 screening variables, from Table A-1, is the reference threshold. All sites in the NLA (both probability and hand-picked) that passed all criteria were considered to be biological reference sites for the NLA (Table A-2, Figure A-1). However, if any site exceeded one or more threshold, then it was not considered a reference lake.

In addition to selecting biological reference sites, analysts also determine poor quality (highly disturbed) sites that would be used in the biological assessment of the nation's lakes. Similar to the reference selection process, thresholds were used to determine which lakes were to be considered poor, in each of the seven cluster regions. The second threshold value in each of the 10 screening variables, from Table A-1, is the poor threshold. If any site exceeded the threshold for any one of these screening criteria, then the site was considered to be in poor condition. However, in regional clusters C, D, and E, a site had to exceed two or more of these thresholds to be considered in poor condition. Analysts incorporated this rule due to the high number of highly disturbed condition sites in these regions, when we applied a single failure as the screening variable threshold.

Note that the NLA did not use data on landuse in the watersheds for the final reference site screening—sites in agricultural areas (for example) may well be considered least disturbed, provided that their chemical and physical conditions are among the best for the region. Additionally, the NLA did not use data from the biological assemblages themselves because these are the primary components of the lake ecosystems being evaluated and to use them would constitute circular reasoning.

Screening NLA Site Data for Nutrient Reference Condition

Setting reference condition for nutrients requires a different process then the one used for biological reference condition evaluation. Because nutrients (TN, TP) were used to select biological reference sites, the biological reference sites could not be used as nutrient reference lakes due to circularity. During the development of nutrient reference sites, 11 nutrient ecoregions were utilized to categorized different portions of the conterminous United States (USEPA 2000). These included Coastal Plain, Temperate Plains, Southeastern Plains and Piedmont, Grass Plains, Cultivated Great Plains, Southern Glaciated, Northern Glaciated, Southern Appalachian Mountains, Xeric West, and Western Mountains. The Grass Plains was separated into to categories, natural and man-made, due to the Sand Hills high natural nutrient levels (Table A-3).

As with selection of biological reference lakes, chemical and lake riparian and littoral condition thresholds were used to select nutrient reference lakes. An initial screening of all ecoregions for inorganic acidity excluded all lakes with an ANC ≤ 50 μeq/L and DOC < 5 mg/L. Once these lakes were excluded, selection of reference conditions by nutrient ecoregion was conducted, using chloride, sulfate, shoreline disturbance by agriculture, shoreline disturbance by non-agriculture, and shoreline disturbance intensity and extent, and in field assessment of agricultural (Assess ag), residential (Assess resid.), and industrial (Assess ind.) landuse from field data form (Table A-3). Similar to biological reference selection, if a lake exceeded any one of these eight selection criteria then the lake was not considered a reference lake. However, chloride was not used to select reference lakes in two Omernik level III ecoregions of the Western Mountains (ecoregion 1) and Northern Glaciated (ecoregion 82) nutrient ecoregions due to ocean influence.

Once the nutrient reference lakes were selected, nutrient levels for separating Good, Fair, and Poor were determine from the distribution of reference lake nutrient concentrations from the 11 nutrient ecoregions. Nutrient levels were determined for both total phosphorus (TP) and total nitrogen (TN). The cutoff between Good and Fair lakes was set at the 75^{th} percentile (Q3) of reference lakes, and the cutoff between Fair and Poor lakes was set at the 95^{th} percentile (P95) of reference lakes (Table A-4). If a nutrient ecoregion had < 20 lakes, then the cutoff between the Fair and Poor lakes was the maximum nutrient concentration (P95 = maximum) for reference lakes in that nutrient ecoregion.

In addition to developing thresholds for nutrients, we determined thresholds from population percentiles in the reference lakes in each of the nutrient ecoregion for chlorophyll-*a* and turbidity (Table A-5). Like the nutrient thresholds, these percentile-based thresholds were used to determine Good, Fair, and Poor lake conditions for the NLA. With the cutoff between Good and Fair lakes set at the 75^{th} percentile (Q3), and the cutoff between Fair and Poor lakes set at 95^{th} percentile (P95).

Table A-1. Regional biological reference thresholds for the reference/highly disturbed lakes

	Phosphorus	Nitrogen	Chloride	Sulfate	Turbidity	ANC[4]	Dissolved oxygen	Agriculture disturbance	Nonagricultural disturbance	Disturbance intensity
	µg/L	µg/L	ueq/L	ueq/L	NTU	µeq/L	mg/L	RDISINAG	RDISINNONAG	RDISINEX1A
A	12/ 100	300 / 1500	200 / 10,000	400 / 1000	5 / 50	≤50	>4 / ≤3	0 / 0.5	0.6 / 0.80	0.5 / 0.85
B	10 / 100	300 / 1500	250 / 10,000	250 / 1000	2 / 50	≤50	>4 / ≤3	0 / 0.5	0.5 / 0.75	0.4 / 0.85
C[1, 2]	15 / 125	500 / 1500	250 / 10,000	250 /1000	5 / 50	≤50	>4 / ≤3	0 / 0.3	0.6 / 0.8	0.5 / 0.85
C[1, 3]	50 / 125	750 / 1500	250 / 10,000	NA / 1000	10 / 50	≤50	>4 / ≤3	0 / 0.3	0.6 / 0.8	0.5 / 0.85
D[1]	75 / 250	750 / 1500	NA / 2000	250 / 1000	10 / 50	≤50	>4 / ≤2	0 / 0.5	0.6 / 0.75	0.6 / 0.85
E[1]	100 / 500	1500 / 5000	600 / 10,000	1500/ 10,000	10 / 50	≤50	>4 / ≤3	0.1 / 0.5	0.6 / 0.75	0.6 / 0.85
F	10 / 100	300 / 1500	250 / 10,000	250 / 1000	2 / 50	≤50	>4 / ≤3	0 / 0.5	0.5 / 0.75	0.4 / 0.85
G	50 / 250	750 / 1500	500 / 10,000	500 / 4000	10 / 50	≤50	>4 / ≤3	0.1 / 0.5	0.5 / 0.75	0.5 / 0.85

[1] Because of the number of highly disturbed sites in these four clusters, a site had to exceed two thresholds to be categorized as highly disturbed, unlike the other cluster where a site only had to exceed one threshold to be considered highly disturbed.

[2] Lakes at latitude greater than 40 degrees – the lake classification number is the sum of reference or highly disturbed sites within the cluster.

[3] Lakes at latitude less than or equal to 40 degrees.

[4] ANC thresholds were only used to determine if a lake would be considered highly disturbed (i.e. adversely affected by atmospheric deposition). If the ANC value was >50 µeq/L, or if dissolve organic carbon is ≥ 5mg/L (even with an ANC value < 50 µeq/L) then the lake was assumed to be an acceptable candidate for as a reference lake with regard to this parameter.

Table A-2. Biological Reference Sites

Reference Clusters	Data Source		Total
	Hand-pick	Random	
Eastern highlands			
A	6	11	17
B	16	14	30
Plains and lowlands			
C	6	24	30
D	5	14	19
E	2	22	24
Mountain West			
F	8	32	40
G	0	10	10
Total	43	127	170

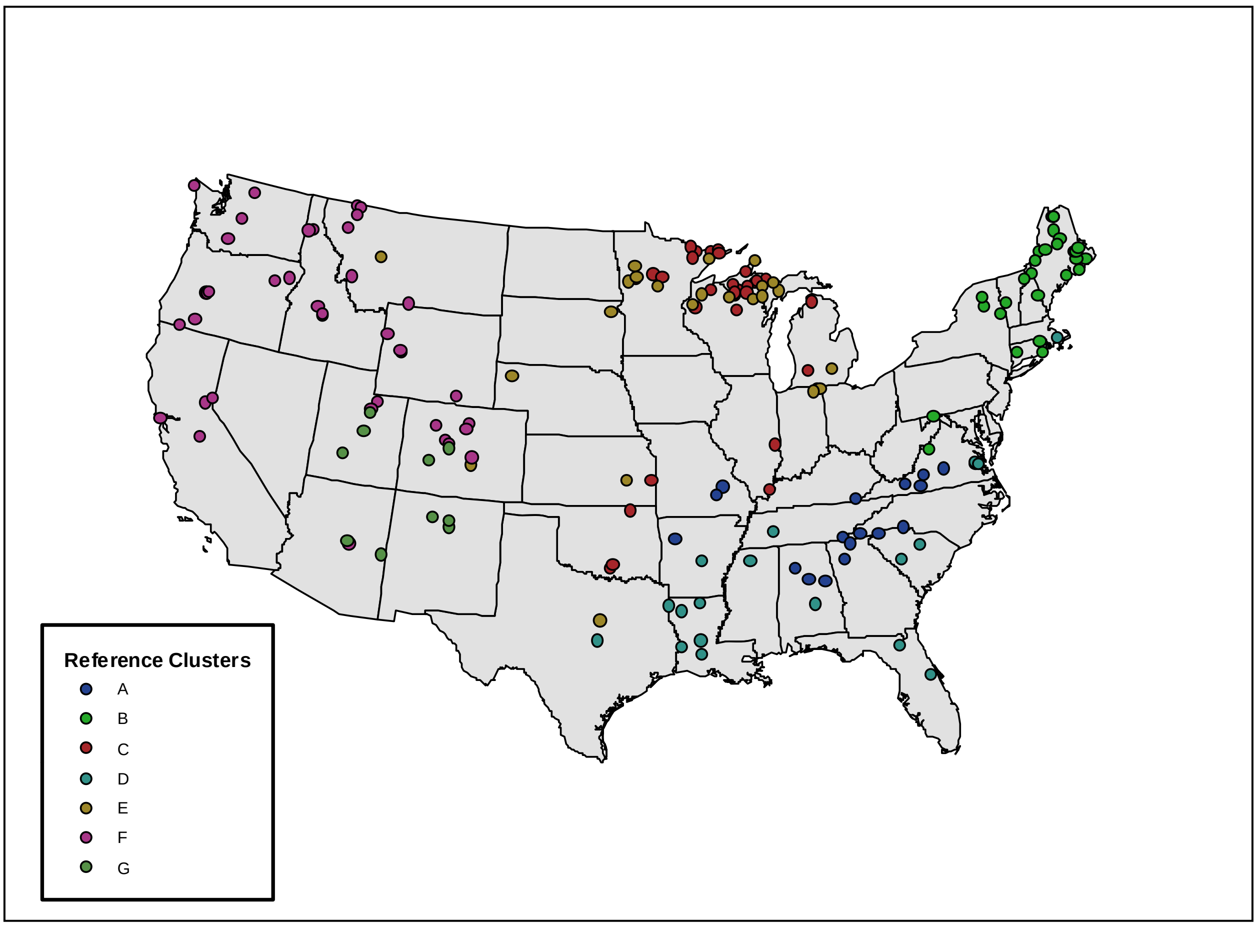


Figure A-1. Biological reference sites by seven regional clusters

Table A-3. Nutrient reference site screening criteria by nutrient ecoregion

Nutrient Ecoregion	Chloride (ueq/L)	Sulfate (ueq/L)	Habitat ag disturb	Habitat non-ag disturb	Habitat Ex1a disturb	Assess ag	Assess resid.	Assess ind.
Coastal Plain	>1000	>400	>0	>0.6	>0.6	>4	>9	>4
II. Western Mts.	>20	>50	>0	>0.2	>0.2	>4	>4	>4
III. Xeric West	>500	>10000	>0.1	>0.6	>0.6	>6	>6	>6
IV. Grass Plains-Man-made	>1000	>10000	>0.2	>0.6	>0.6	>9	>9	>9
IV. Grass Plains-Natural	>400	>400	>0	>0.1	>0.1	>5	>5	>5
IX. SE Plains/Piedmont	>200	>400	>0	>0.4	>0.4	>4	>9	>4
V. Cultivated Great Plains	>1000	>10000	>0.2	>0.6	>0.6	>9	>9	>9
VI. Temperate Plains	>1000	>10000	>0	>0.6	>0.6	>9	>9	>9
VII. Southern Glaciated	>400	>400	>0	>0.6	>0.6	>9	>9	>9
VIII. Northern Glaciated	>20	>200	>0	>0	>0	>4	>9	>4
XI. S. Appalachian Mts.	>500	>500	>0.1	>0.5	>0.5	>9	>9	>9

Table A-4. Good/Fair/Poor Condition Class thresholds for total phosphorus (TP) and total nitrogen (TN)

Nutrient Ecoregion	# Ref Lakes	TP (ug/L) Good-Fair	TP (ug/L) Fair-Poor	TN (ug/L) Good-Fair	TN (ug/L) Fair-Poor
Coastal Plain	14	26	75	629	2311
II. Western Mts.	23	15	19	278	380
III. Xeric West	14	48	130	514	2286
IV. Grass Plains-Man-made	9	37	56	513	824
IV. Grass Plains-Natural	6	839	1719	8647	9359
IX. SE Plains/Piedmont	30	62	176	680	1531
V. Cultivated Great Plains	16	117	159	1106	1355
VI. Temperate Plains	10	108	193	1240	2447
VII. Southern Glaciated	13	24	102	828	1410
VIII. Northern Glaciated	24	16.5	36	674	1174
XI. S. Appalachian Mts.	21	10	29	311	665

Table A-5. Condition Class thresholds for chlorophyll-*a* and turbidity

Nutrient Ecoregion	Total # of Lakes in Data	Chl-a (ug/L) Good-Fair	Chla-a (ug/L) Fair-Poor	Turb. (NTU) Good-Fair	Turb (NTU) Fair-Poor
Coastal Plain Ecos	89	29.1	75.6	6.30	19.9
II. Western Mts.	165	1.81	2.74	1.44	5.47
III. Xeric West	88	7.79	29.5	3.69	24.9
IV. Grass Plains-Man-made	40	13.9	25.1	4.49	14.4
IV. Grass Plains-Natural	24	118	144	47.5	75.9
IX. SE Plains/Piedmont	186	31.7	84.0	11.4	37.3
V. Cultivated Great Plains	122	49.9	76.3	26.5	50.3
VI. Temperate Plains	106	37.8	49.6	10.7	19.7
VII. Southern Glaciated	125	8.56	46.4	5.19	102
VIII. Northern Glaciated	140	7.56	12.5	2.75	5.41
XI. S. Appalachian Mts.	72	5.34	23.8	1.91	2.38

Biological Data

Data Preparation: Standardizing Counts

NLA analysts standardized the number of individuals in a sample to a constant number to provide an adequate number of individuals (i.e. diatom valves, natural algal units, microcrustaceans and rotifers) that was the same for nearly all samples and that could be used for both multimetric index development and/or O/E predictive modeling index. For sediment diatoms, a subsample was place on a microscope slide to be enumerated. All samples were scribed with transect lines, which were used for counting a know field of the subsample. Taxonomists were to enumerate no more than 600 valves per sample. For phytoplankton, a subsample was place on a microscope slide and the taxonomists were to enumerate up to 300 natural algal units. Finally, for zooplankton, two subsamples were enumerate for microcrustaceans and rotifers. For each taxonomic group, taxonomist were supposed to count a minimum of 200 individuals and not more than a maximum of 400 individuals.

Samples that did not contain the minimum number of units/individuals were reviewed and retained for further analysis when appropriate (i.e. if the sampling effort was determined to be sufficient) because low counts can indicate a response to one or more stressors. For example, samples from sites classified as least disturbed were retained if zooplankton counts were 100 or more individuals.

Operational Taxonomic Units

To provide a nationally consistent database for the diatoms, phytoplankton and zooplankton, taxonomic lists were reviewed for discrepancies. In some cases it was necessary to combine taxa to a coarser level of common taxonomy. This new combination of taxa is called the "Operational Taxonomic Unit" or OTU and improves the level of confidence in an overall assessment.

Sediment Diatom Metric Development

The taxonomic composition and relative abundance of different taxa that compose the sediment diatom assemblages in lake sediments were used to develop a diatom Index of Biological Integrity (IBI) or a Lake Diatom Condition Index (LDCI). IBIs for fish and benthic macroinvertebrates have been used extensively in North America, Europe, and Australia to assess how human activities affect ecological condition (Barbour et al., 1995, 1999; Karr and Chu 1999). IBIs usually contain multiple measures of a given assemblage, such as structural, functional, and/or tolerance metrics, that respond positively or negatively to anthropogenic stressors (Barbour et al. 1999). The purpose of these indicators is to present the complex data represented within an assemblage in a way that is understandable and informative to resource managers and the public. This approach has been recommended for use in previous EPA surveys such as the Wadeable Streams Assessment.

While diatoms have been used extensively in North America, Europe, and Australia to monitor water quality, development of diatom IBIs has been much more limited compared to other biological assemblages (Bahls 1993, Hill et al. 2000, Wang et al. 2005). Additionally, most of these IBIs have been developed for lotic ecosystems with a few exceptions for wetland ecosystems (Wang et al. 2006). This study contains the first known published IBI for lentic ecosystems.

The following sections provide a general overview of the approach used to develop ecological indicators based on sediment diatoms, followed by details regarding data preparation and the process used for each approach to arrive at a final indicator.

Metric Evaluation and Selection

Candidate metrics were derived from the sediment diatom count data and traits of each taxon. Morphological and growth form traits were obtained from literature or best professional judgment. Indicator species analysis was used to determine diatom taxa that were characteristically found in reference or impaired lakes, and to determine diatom taxa characteristically found in either high TN and/or TP, or low TN and/or TP (Dufrene and Legendre 1997). In most cases, three variants of each candidate metric were calculated: one based on taxa richness, one based on the proportion of individuals, and one based on the proportion of taxa. All candidate metrics were assigned to one of the following five categories representing different aspects of biotic integrity (Barbour et al., 1999; Karr, 1993; Karr et al., 1986; Stoddard et al., 2005).

- **Similarity to Reference Condition:** Proportions of individuals and taxa characteristically found in reference or impacted sites
- **Diversity:** e.g. taxa number observed in samples and evenness of the distribution of individuals across taxa
- **Composition:** e.g. the relative abundance of different genera
- **Morphological and growth forms:** e.g. Benthic, planktonic, motile, epiphytic, colonial, chainforming
- **Tolerance:** e.g. low and high nutrient

Three performance evaluations were conducted to identify the best metric from each metric category. Candidate metrics that failed a test were eliminated from additional consideration and testing.

- **Signal to noise (S:N) test:** "Signal to noise" is the ratio of variance among sites and the variance within a site (based on repeated visits to the same site). A low S:N value indicates a metric that cannot distinguish among sites very well. S:N ratios were calculated for each assessment region. Generally, candidate metrics having S:N values ≤ 1 were eliminated.
- **Mann-Whitney *U* test:** Metrics were selected using this method when these tests showed significant differences ($\alpha=0.05$) between reference and highly disturbed sites (see the description of how reference and poor sites were identified under Setting Expectations). Additionally, analysts evaluated the separation power of each significant metric using deviation in median ranks of metrics in reference and impaired sites and the Z-statistic. Separation power has defined as the amount of overlap (i.e. 25th and 75th percentile) in box plots of values of metrics for reference and impaired sites (Barbour et al. 1996, 1999). The Z-statistic accounts for the separate variation in ranks within reference and impaired groups as well as the difference in magnitude of ranks among groups.
- Independence among metrics in different metric categories was evaluated using correlations among metrics. Metrics within categories were often highly correlated. Independence among metrics was maintained by calculating averages of metrics within categories before calculating an overall average LDCI.

Metrics with the highest S:N and Z-statistics (either positive or negative) and lowest correlations with other categories were selected for inclusion in the LDCI.

Metric Selection, Scaling, Transformation and Calculation of Lake Diatom

Condition Indices

Multiple versions of the Lake Diatom Condition Index were calculated to evaluate their relative performances for distinguished reference and highly disturbed sites. The same metrics were used in calculating all versions of the LCDIs. All LCDIs were based on metrics that were scaled to the same range.

Structural and Tolerance Metrics

Most structural and tolerance metrics were scaled to a 0-1 range using the "Blocksom – 5th-95th percentile" method (Blocksom 2003):

- determine 5th and 95th percentiles of metrics;
- subtract the 5th percentile of the metric from metric value at a site; and
- divide that quantity by difference between the 5th and 95th percentiles (Table A-6).

If metrics were positively correlated with the chemical principle components analysis (PCA) factor score, they were subtracted from 1.0 to reverse the scale such that a 0.0 and 1.0 metric scores indicated low and high biological condition, respectively. If metrics were positively correlated with the chemical PCA factor, they were not subtracted from 1.0 to reverse the scale.

Genus Level Composition Metrics and Percent Epiphytic Individuals

Genus level species composition metrics and percent epiphytic individuals were normalized using a 0, 0.5, 1.0 values that were assigned to metric values based on quartile separations in the ranges of the metrics (Table A-7). This scale was used because many sites had 0.0 relative abundances of genera and percent epiphytic individuals, which would skew the distribution of a metric normalized using the 5-95th percentile method described above.

- If metrics were negatively correlated to the chemical PCA factor score, high values of metrics indicated high biological condition. In this case 0 was assigned to metric values if they were less than the 25th percentile of the metric range, 0.5 was assigned to metric values if greater than or equal to the 25th percentile and less than the 75th percentile, and 1.0 was assigned to metric values if greater than or equal to the 75th percentile.
- If metrics were positively correlated to the chemical PCA factor score, high values of metrics indicated low biological condition. In this case 10 was assigned to metric values if they were less than the 25th percentile of the metric range, 0.5 was assigned to metric values if greater than or equal to the 25th percentile and less than the 75th percentile, and 0.0 was assigned to metric values if greater than or equal to the 75th percentile.

Table A-6. The 5th and 95th percentile of final evaluated structural and tolerance metrics.

Metric	5th percentile	95th percentile
Prop. Impacted spp	0.000	0.349
Prop. Reference spp	0.026	0.050
Shannon diversity *H'*	1.414	3.486
Richness	19	71.45
Prop. Colonial Individuals	0.034	0.825
Prop. Low TP Taxa	0.034	0.606
Prop. High TP Taxa	0.042	0.651
Prop.Low TN Taxa	0.045	0.611
Prop. High TN Taxa	0.029	0.618

If more than 25% of the sites had zero relative abundance at a site, then greater than 25% of the values would be assigned either a 0.0 or 1.0, depending upon the relationship between the metric and the PCA score. For example, the % Cocconeis individuals at 41% of the sites were 0.0 and this metric was positively correlated to the chemical PCA factor score of the site; accordingly 41% of the sites with relative abundances equal to 0.0 were assigned a 1.0, 34% of site were assigned with relative abundances between the 41st and 75th percentiles were assigned a 0.5, and the remaining 25% of values was assigned a 0.0.

Table A-7. The 25th and 75th percentile of example genus/growth form metrics.

Metric	25th percentile	75th percentile
Prop. Achnanthidium Individuals	0.002	0.045
Prop. Cocconeis Individuals	0.000	0.012
Prop. Cyclotella and Stephanodiscus Individuals	0.003	0.234
Prop. Epiphytic	0.002	0.024

The LDCI was calculated in three steps to produce a multimetric index ranging between 0 and 100. First the averages of selected metrics within the five metric categories were determined. Then these averages of each metric category were weighted evenly by multiplying by 20 and these products were summed to calculate the final value of the observed LDCI. Finally, LDCI was calculated as the deviation in observed LDCI values and the expected LDCI value for a specific lake, where the latter accounted for variation in LDCI values due to natural lake and lake watershed features.

Performances of different models predicting expected LDCI (using the 10 selected and scaled metrics in Tables A-6, A-7) were evaluated by calculating variation in expected LDCI among reference sites. Models with low variation in expected condition at reference sites will more precisely distinguish between reference and impaired condition. Models of expected LDCI differed as a result of two calculation methods, whether individual metrics or the multimetric LDCI were predicted by models, and in the variables used to account for natural variation in expected LDCI.

Expected LDCIs models were calculated using the following (Table A-8):

1. no natural variables;
2. natural versus man-made lakes;
3. seven lake type clusters;
4. nine ecoregions
5. Classification and regression tree (CART) model predictions of natural variation in metrics at reference sites using all “GIS” variables as in Cao et al. (2007);
6. adjusted using CART model predictions of natural variation in LDCI at reference sites using all “GIS” variables;

7. regression predictions of natural variation in LDCI at reference sites using all "GIS" variables and forward stepwise regression;
8. regression predictions of natural variation in LDCI at reference sites using "GIS" variables selected for first principles as important causal variables and forward stepwise regression;
9. regression predictions of natural variation in LDCI at reference sites using "GIS" variables selected for first principles as important causal variables and using all subset regression; and
10. regression predictions of natural variation in LDC at reference sites using "GIS" variables selected for first principles as important causal variables, minus maximum lake depth because it was negatively related to LDCI, and using all subset regression.

Table A-8. The amount of variation in expected LDCI among reference sites using the models above.

Model	LDCI model	Ref_Var
1	None	226.6
2	lake type clusters	157.9
3	lake vs reservoir	210.6
4	ecoregions level 9	140.2
5	all natural factors for metrics by CART model	93.2
6	all natural factors for LDCI by CART model	116.6
7	all natural factors for LDCI by GLM	102.8
8	1st P natural factors for LDCI by GLM	97.2
9	all subset NF for LDCI by GLM	97.2
10	all subset NF - depth for LDCI by GLM	107.2

Option 10 was chosen, with reference variance of 107.2. This model had lower variance than any of the a priori categorical classifications of lakes. Model 5 (14 metrics independently adjusted by CART model) had lower variance among reference site LDCI values, but requires more evaluation before finalizing. Of the last four models, we decided that all subset regression models (models 9 and 10) were better than models using forward stepwise regression (models 7 and 8). Model 10 was chosen because maximum lake depth, a predictor variable in the LDCI was negatively related to LDCI in model 9, which does not make sense based on first principles, i.e. we predicted that deep lakes should naturally have higher LDCIs than shallow lakes. Exploration of regional patterns indicted maximum lake depth was positively and negatively correlated to LDCI at reference sites, depending upon region. Model 10 was:

$$Expected LDC = (73.363 - 79.948 \times KFCT\_AVE + 0.008167 \times LAT\_DD^2 - 1.367 \times Log(BASIN\_LAKE\_RATIO) + 2.79 \times Log([ELEV\_PT + 1] + 0.581 \times LON\_DD + 1.116 \times \sqrt{SUMP\_PT}))$$

In the model, KFCT_AVE is watershed mean soil erodibility factor of soils (no units) from State Soil Geographic (STATSGO) Database, LAT_DD is latitude in decimal degrees, BASIN_LAKE_RATIO is ratio of basin area to lake area, ELEV_PT is site elevation (meters) from the National Elevation Dataset, LON_DD is longitude in decimal degrees, and SUMP_PT is annual sum of the predicted mean monthly precipitation (mm) derived from the PRISM data.

Plankton O/E: Predictive (RIVPACS) Models

Observed over Expected (O/E) indices provide a quantitative measure of biological condition by measuring the agreement between the taxonomic composition expected under reference conditions and that observed at individual sites. For the NLA, we developed a combined phytoplankton-zooplankton O/E index based on the 259 plankton taxa observed across reference-quality lakes (351 total plankton taxa were identified across all of the 1157 NLA lakes).

Because taxonomic composition can vary markedly with natural environmental factors, application of the O/E index depends substantially on development of models that predict how taxonomic composition varies with natural environmental setting. These models are calibrated with data collected at reference sites.

One hundred and seventy lakes were initially identified as candidate reference lakes for use in calibrating 3 regional (western mountains and xeric (WMTNS), plains and lowlands (PLNLOW), and eastern highlands (EHIGH)) models. As described, these reference lakes were selected from the 1157 lakes sampled for the NLA based on application of regional screening criteria (see pages A-2). Of these 170 lakes, 14 large PLNLOW lakes lacked littoral predictor variable data and were dropped from model development resulting in156 calibration lakes for purposes of the O/E model (Fig. A-2).

The 50 (WMTNS), 59 (PLNLOW), and 47 (EHIGH) lakes were then classified into 4 (WMTNS = groups 1-4), 5 (PLNLOW = groups 5-9), and 3 (EHIGH = groups 10-12) groups based on their combined phytoplankton and zooplankton taxa composition (Fig A-2). Following reference site classification, discriminant functions models were developed to predict the probability of group membership from naturally occurring landscape attributes. For the WMTNS model, $\log_{10}$ average available water holding capacity (proportion), $\log_{10}$ soil permeability (inches/hr), average depth to the water table (feet), longitude (decimal degrees), and $\log_{10}$ average calcium oxide content (%) of the watershed's bedrock were identified as predictors. For the PLNLOW model, mean long-term maximum monthly air temperature (^{o}C), mean long-term maximum monthly precipitation (mm), $\log_{10}$ lake surface area (km^2), $\log_{10}$ watershed area (km^2), $\log_{10}$ lake perimeter (km), dummy variable specifying natural lake (1) or reservoir (0), square root of the percent littoral cover as aquatic macrophyte, and square root of the percent littoral cover as organic matter were identified as predictors of plankton composition. For the EHIGH model $\log_{10}$ average available water holding capacity (proportion), average bulk soil density (g/cm^3), and average depth to the water table (feet) were identified as predictors.

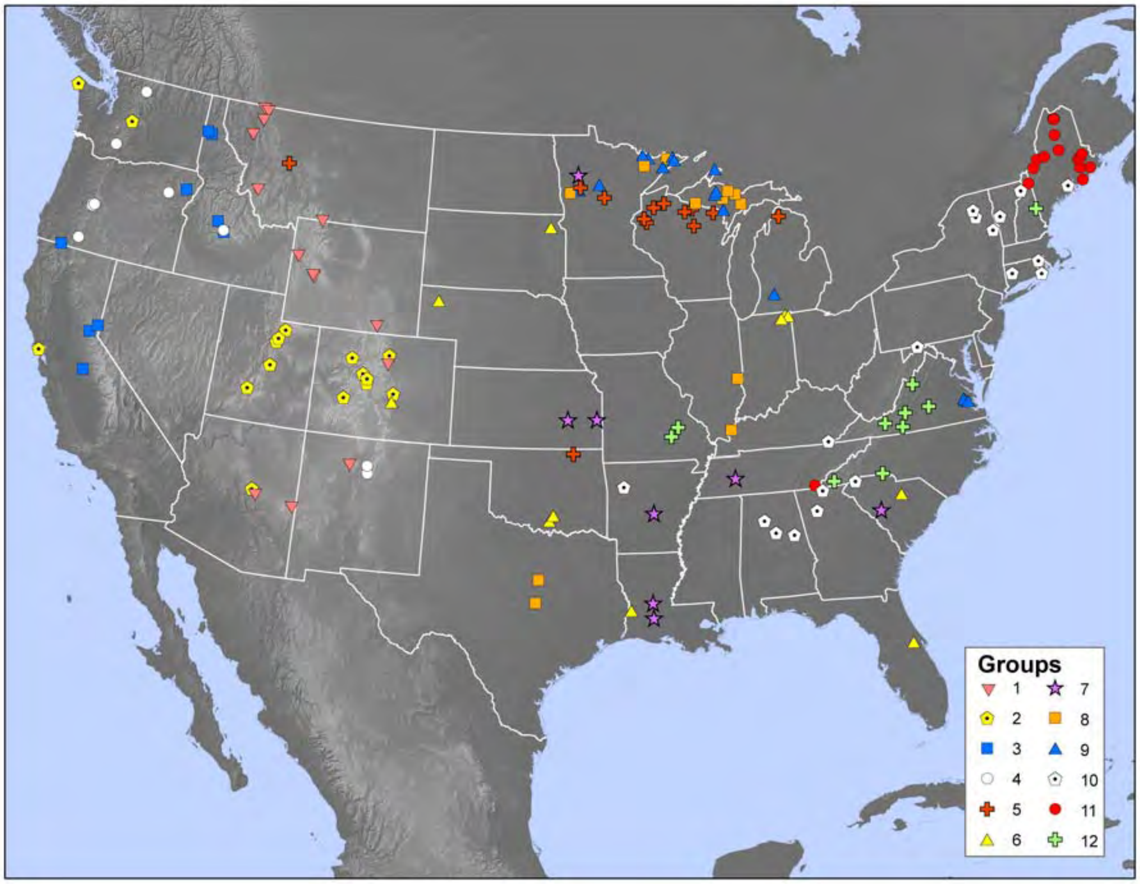


Figure A-2. Location of the 156 reference lakes used to develop the phytoplankton-zooplankton O/E indices. Individual lakes are symbol and color coded by the groups they were assigned to based on similarity in taxa composition.

These discriminant functions models were used to predict the probabilities of each of the 1157 NLA lakes belonging to each of the biologicacally-defined groups. These probabilities, together with the observed frequencies of occurrence of each taxon among the lakes in each group, were used to predict the probabilities of observing each of the 259 reference lake taxa at each of the 1157 NLA lakes. For each lake, probabilities of detection > 0.5 were then summed to estimate the number of reference-lake taxa expected (E) at each lake. The O/E index is the proportion of those reference-lake taxa predicted at a specific lake that were observed (O) in a sample.

The distribution of O/E values observed at reference-quality lakes is used to evaluate the condition of assessed lakes. Because the reference-site O/E distributions for the 3 regions had 1 or 2 statistical outliers, we dropped these values when estimating the 5^{th} and 25^{th} percentile of reference site values. Lakes with O/E values > 25^{th} percentile of reference site values were considered to be in good condition. Lakes with O/E values between the 5^{th} and 25^{th} percentiles of reference site values were considered to be in fair condition. Lakes with O/E values < the 5^{th} percentile of reference site values were considered to be in poor condition. The 5^{th} and 25^{th} percentile reference site O/E values, were 0.60 and 0.80.

Relative Risk, Attributable Risk, and Relative Extent

In the NLA, each targeted and sampled lake was classified as being in either "Good," "Fair," or "Poor" condition, separately for each stressor variable and each biological response variable. From this data, we estimated the *relative extent* (prevalence) of lakes in Poor condition for a specified stressor or response variable. We also estimated the *relative risk* (*RR*) of each stressor for a biological response. *RR* measures the severity of a stressor's effect on that response in an individual lake, when that stressor is in Poor condition (Van Sickle, et al. 2006). Finally, we estimated the population *attributable risk* (*AR*) of each stressor for a biological response. *AR* combines *RR* and relative extent into a single measure of the overall impact of a stressor on a biological response, over the entire population of lakes (Van Sickle and Paulsen 2008).

To estimate *RR* and *AR*, we first combined the "Good" and "Fair" classes of condition into a single class designated "Not Poor". Thus, each sampled lake was designated as being in either Poor (*P*) or Not-Poor(*NP*) condition, separately for each stressor and response variable.

To estimate the relative extents, *RR,* and *AR* for one stressor (S) and one response (Y) variable, we compiled a 2x2 table (Table A-9), based on data from all lakes that were included in the probability sample. A separate table must be compiled for each pair of stressor and response variables:

Table A-9.

	Stressor (*S*)	
Response (*Y*)	Not-Poor (*NP*)	Poor (*P*)
Not-Poor (*NP*)	*a*	*b*
Poor (*P*)	*c*	*d*

Table entries (*a,b,c,d*) are the sums of the sampling weights of all sampled lakes that were found to have each combination of *P* or *NP* condition for *S* and *Y*. Thus, $RE_{S.est} = (b+d)/(a+b+c+d)$ is the estimated relative extent of Poor stressor condition, calculated by estimating the number of lakes in the population that have Poor stressor condition (totaled over both classes of response condition), divided by the total estimated number of lakes. Similarly, $RE_{Y,est} = (c+d)/(a+b+c+d)$ estimates the relative extent of lakes in Poor condition for the biological response variable.

RE_S can also be interpreted as the probability that a lake chosen at random from the lake population will have Poor stressor condition. In shorthand, this probability can be written as $RE_S = \Pr(S = P)$. We use similar concepts and language from probability theory to define and interpret *RR* and *AR*.

[Note: *RE* estimates made from Table A.1 may differ slightly, due to sampling variation, from our extent estimates that were made using separate data for each *S* and *Y* variable. This is because the weight sums in Table A.1 include only those lakes that have nonmissing class assignments for both *S* and *Y*.]

Relative Risk (RR)

Relative risk measures the likelihood (that is, the “risk”, or probability) of finding Poor biological response condition in a lake when the condition of a specific stressor is also Poor. This likelihood is expressed relative to the likelihood of Poor response condition in lakes that have Not-Poor stressor condition. That is,

$$RR = \frac{\Pr(Y = P \mid S = P)}{\Pr(Y = P \mid S = NP)} \qquad \text{(A1)}$$

Using Table A.1, *RR* is estimated by:

$$RR_{est} = [d/(b+d)]/[c/(a+c)] \qquad \text{(A2)}$$

RR = 1.0 indicates “No association” between stressor and response, that is, Poor biological condition in a lake is equally likely to occur whether or not the stressor condition is Poor. RR < 1.0 indicates that Poor response condition is actually less likely to occur when the stressor is Poor.

Further details of RR and its interpretation, including estimation of a confidence interval for RR_{est}, can be found in Van Sickle et al. (2006).

Attributable Risk (AR)

Attributable risk (AR) measures how much of the extent of Poor condition for a biological response variable can be attributed to the Poor condition of a specific stressor. AR is based on a scenario in which the stressor would be entirely eliminated from the lake population, by means of restoration activities. (By “eliminated”, we mean that all lakes in Poor condition for the stressor are restored to the Not-Poor condition.) Under this scenario, AR is defined *as the proportional decrease in the extent of Poor biological response condition that would occur if the stressor were eliminated from the lake population.* Mathematically, AR is defined as (Van Sickle and Paulsen 2008)

$$AR = \frac{\Pr(Y = P) - \Pr(Y = P \mid S = NP)}{\Pr(Y = P)} \qquad \text{(A3)}$$

We estimated AR using Equation A3. We first calculated $RE_{Y,est}$ (see above), which is an estimate of Pr(Y = P). Then, using the weight sums in Table A-9,

$$AR_{est} = [RE_{Y,est} - c/(a+c)] \,/\, RE_{Y,est} \qquad \text{(A4)}$$

We calculated a confidence interval for AR_{est} following Van Sickle and Paulsen (2008). AR can take a value between 0 and 1. An AR value of 0 indicates either “No association” between stressor and response, or else a stressor having zero extent.

A strict interpretation of AR in terms of stressor elimination, as described above, requires one to assume that the stressor-response relation is strongly causal and that stressor effects are reversible. Van Sickle and Paulsen (2008) discuss the reality of these assumptions, along with other issues such as interpreting the *AR’s* of multiple, correlated stressors, and using AR to express the joint effects of multiple stressors.

However, AR can also be interpreted more informally, as a measure that combines RR and relative extent into a single index of the overall, population-level impact of a stressor on a response. After some algebra, AR can be written as (Van Sickle and Paulsen 2008)

$$AR = \frac{RE_S(RR-1)}{1 + RE_S(RR-1)} \qquad \text{(A5)}$$

Equation A5 shows that the numerator of *AR* is the product of the relative extent of Poor stressor condition and the “excess” relative risk (*RR*-1) of that stressor. The denominator standardizes this product to yield *AR* values between 0 and 1. Thus, a high *AR* for a stressor indicates that the stressor is widely prevalent (has a high relative extent of Poor condition), and the stressor also has a large effect (high *RR*) in those lakes where it does have Poor condition.

Water Chemistry Analysis

Six chemical stressors are summarized in the NLA report: total nitrogen, total phosphorus, dissolved oxygen, turbidity, acidity and salinity. For total nitrogen and total phosphorus, threshold values were determined during the reference nutrient process. For setting nutrient class boundaries, reference sites from the screened NLA dataset were used (see page 3). Because nutrients were the focus, the two nutrient screening levels used in defining biological reference sites were dropped and the other screening factors were used by themselves to identify a set of “nutrient reference sites.” Before calculating percentiles from this set of sites, outliers (values outside 1.5 times the interquartile range) were removed (Herlihy and Sifneos 2008). For acidity, threshold values were determined based on values derived during the NAPAP program. Sites with acid neutralizing capacity (ANC) less than zero were considered acidic. Those with dissolved organic carbon (DOC) greater than 10 mg/L were classified as organically acidic (natural). Acidic sites with DOC less than 10 and sulfate less than 300 μeq/L were classified as acidic deposition impacted, those with sulfate above 300 were acid mine drainage impacted. Sites with ANC between 0 and 25 μeq/L were considered acidic deposition influenced, but not currently acidic.

Salinity classes and dissolved oxygen were divided into low, medium, or high classes. Salinity classes were defined by specific conductance using ecoregional specific values (Table A-10). Dissolved oxygen classes were defined by oxygen concentration for each ecoregion (Table A-10).

Table A-10. Select Water Chemistry Criteria for NLA Assessment

Ecoregion	Salinity as Conductivity (μS/cm) Low-Medium	Salinity as Conductivity (μS/cm) Medium-High	Dissolved Oxygen (mg/L) Low-Medium	Dissolved Oxygen (mg/L) Medium-High
CPL	500	1000	≥ 3	5 ≤
NAP	500	1000	≥ 3	5 ≤
SAP	500	1000	≥ 3	5 ≤
UMW	500	1000	≥ 3	5 ≤
TPL	1000	2000	≥ 3	5 ≤
NPL	1000	2000	≥ 3	5 ≤
SPL	1000	2000	≥ 3	5 ≤
WMT	500	1000	≥ 3	5 ≤
XER	500	1000	≥ 3	5 ≤

Trends Studies

Trend Analysis of National Eutrophication Survey (NES)

Monitoring and surveillance programs have, in the past, often dealt with site-specific questions of ecosystem condition, thus concentrating on single lakes or small groups of lakes. For example, sites are often monitored for nutrient levels, frequency of algal blooms, fisheries, fecal coliform counts at swimming beaches, etc. However, present day pressures on aquatic systems across large geographic areas have driven the need to assess lakes over far wider regions. This kind of need produced the National Eutrophication Survey (NES) in 1972-1976. While this survey was National in scope, the design was not of a rigorus scientific nature. The purpose of the NES was to assess the trophic condition (defined as the nutrient enrichment) of lakes influenced by domestic waste water treatment plants (WWTP) with a sprinkling of special purpose lakes that were not necessarily influenced by WWTP's. The specific purpose of the survey was to measure nutrient inputs from all sources in the watershed relative to those of the WWTP source to determine if WWTP upgrades might be successful in modifying the lake or reservoir trophic state.

The trophic state definition and assessment approach to water quality employed by the NES did not necessarily result in the identification of degraded water quality, but was used simply to characterize the water quality based on the nutrient levels. The perception of good or poor water quality based on the nutrient level depends on the intended beneficial use of the water body. For example a reservoir managed for a warm water fishery can tolerate a much greater degree of nutrient enrichment than can a lake intended for a cold water trout fishery. Therefore, nutrient enrichment, or degree of eutrophication was employed as a tool to gauge the water quality, while the interpretation of good or poor condition depended on the intended use of the lake or reservoir. While the NES survey assessed the condition of selected lakes across the United States, the focus of the survey was to assess the condition of individual lakes in detail rather than to extrapolate results to the condition of all lakes. Trophic state condition based on nutrient levels with some characterization of correlative Secchi disk transparency, phytoplankton, and chlorophyll-*a* became the focus for individual NES lake reports. Because of the correlative nature between nutrients and chlorophyll-a concentration, NLA analysts decided to utilize chlorophyll-*a* concentrations as the trophic condition indicator for the current comparison between studies. Additionally, because of the significance of nutrients in inland waters, NLA analysts also made a comparison of nutrient concentrations between studies.

Since 1972 there has been increased interest in characterization of lake and reservoir quality on a regional basis relative to biogeochemical and land use characteristics. This approach seeks to quantify those characteristics relative to lake water quality, such that regionalized management techniques might be utilized to minimize adverse effects on water quality. Relatively new developments in statistical sampling designs, ecoregional landscape classification and TM and AVHRR technologies coupled with GIS capabilities and other techniques make it possible to do regional resource analyses in ways that did not exist at the time of the NES. We now have tools to conduct a regional census, survey, modeling effort or other techniques designed to describe, infer, or extrapolate lake, reservoir, and wetland conditions across temporal, spatial and biological scales (boundaries).

Because the design of the NLA selected lakes on a probability basis by number and size we can infer analytical results of the sampling to the population of lakes from which the sample was drawn. As part of the NLA design process, the team also had the opportunity to include a subset of lakes from the NES survey of 1972-1976 and make a comparison of trophic state, i.e. how have these lakes changed in trophic state over 35 years. This afforded NLA analysts the opportunity to use the rigorous statistical design of the regional NLA to provide a benchmark of condition change for lakes originally selected and evaluated on a site-specific basis. As a result, the subset of NES lakes sampled for NLA, was used to estimate the current condition of all 800 lakes in the original NES survey.

The goal of this trend analysis was to compare the 1972-1976 trophic state of a subset of the NES lakes with the trophic state of those same lakes today (2008). As described above, the NLA sampling and analysis provided the opportunity to accomplish this goal. While sampling and analysis techniques differed somewhat for the two surveys, NLA analysts determined that the differences were insignificant relative to comparing trophic states using chlorophyll-*a* and nutrient concentrations. The NLA sampling consisted of a single, mid-summer integrated water sample at the deepest spot in the lake and from just below the surface to a depth of about 1.5m (a sampling tube). The NES sampling consisted of sampling several sites on a lake as well as the inlets and outlets. However, this sampling also included a site at the perceived deepest spot in the lake. Sampling was done with a Van Dorn Bottle at just below the surface and at 1 -2 m depth intervals. Therefore, the current comparison used the integrated sample NLA chlorophyll-*a* and nutrient concentrations and compared them to NES samples taken at the site nearest the NLA site and from depth(s) that most nearly mimicked the depth of the NLA integrated depth sample. The accuracy and precision of analytical results are considered comparable to each other based on the methods and the QA of both the NES (USEPA 1974; USEPA 1975a; USEPA 1975b) and the NLA.

We used a 4 X 4 factorial design to assess the trophic state of the original NES survey and the NLA. We determined how many lakes were in each of the following trophic classes hypereutrohic (HE), to eutrophic (E), to mesotrophic (M), to oligotrophic (O), as well as the reverse series from O, to M, to E, to HE, during 1972 and 2007. The results are summarized in Figure A-3.

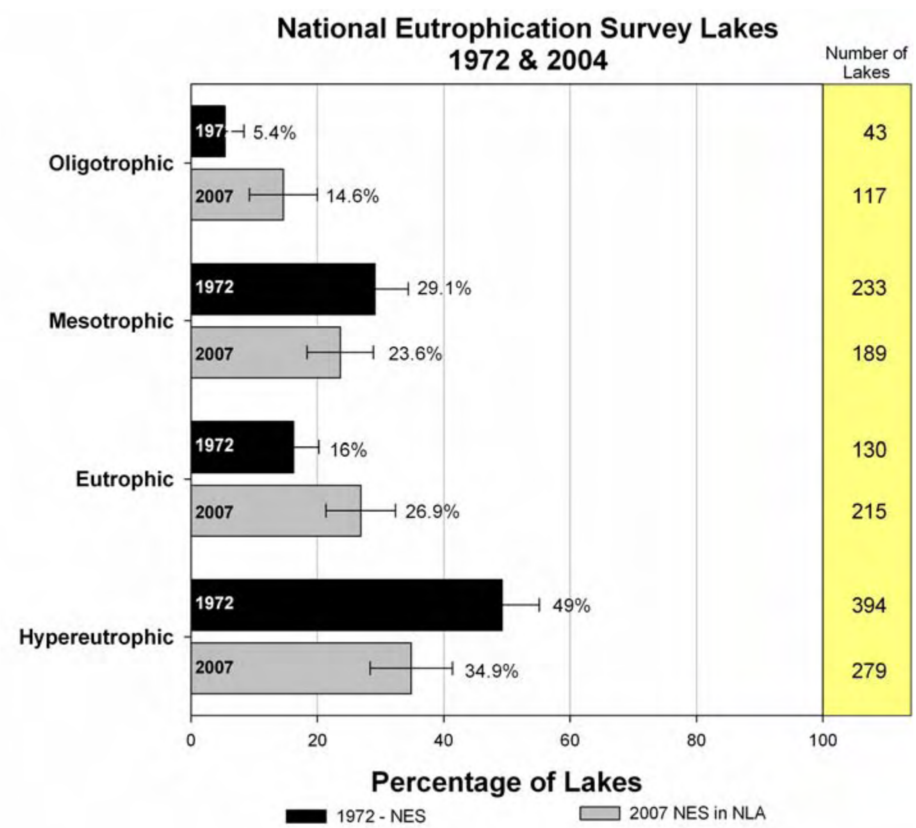


Figure A-3. Percentage and number of NES lakes estimated in each of four trophic classes in 1972 and in 2007 based on chlorophyll-*a* concentrations.

The 2007 NLA indicates that 51.1% of the NES lakes remain in the same trophic state category as they were in 1972 (Figure A-4). Another 22.6 degraded to lower trophic state categories. Only 26.3 percent of the lakes actually improved in their trophic state category. While at first glance this seems a rather bleak picture, the results must be put into proper perspective. First, most of the original NES lakes were eutrophic or hypereutrophic to begin with because they were selected for their proximity to domestic waste treatment plants. Second, there likely has been an increasing population density associated with these lakes, again most likely since they were in the vicinity of waste treatment plants where populations usually grow.

While the 2007 NLA indicates a modest improvement in trophic state, as assessed as chlorophyll-*a*, there does appear to be a much more substantial improvement in total phosphorus concentrations between 1972 and 2007. The 2007 NLA indicates that 50.4% of lakes decreased in there total phosphorus concentrations, while only 25.9% increased, with another 23.6% showing no change. The most likely reason for this out come is the improvements in waste treatment plants between 1972 and 2007. However, this inconsistency between chlorophyll-*a* and total phosphorus concentrations is difficult to resolve.

While we have land use cover for the 2007 NLA we do not have similar land cover data for the original NES, that we might be able to make land use change associations with the trophic state changes (chlorophyll-*a*).

NLA analysts did not identify the causes for the improvements or the declines, and other factors may be influencing chlorophyll-*a* that are not influencing total phosphorus concentrations. Additional analyses are recommended to delve into these results in greater detail.

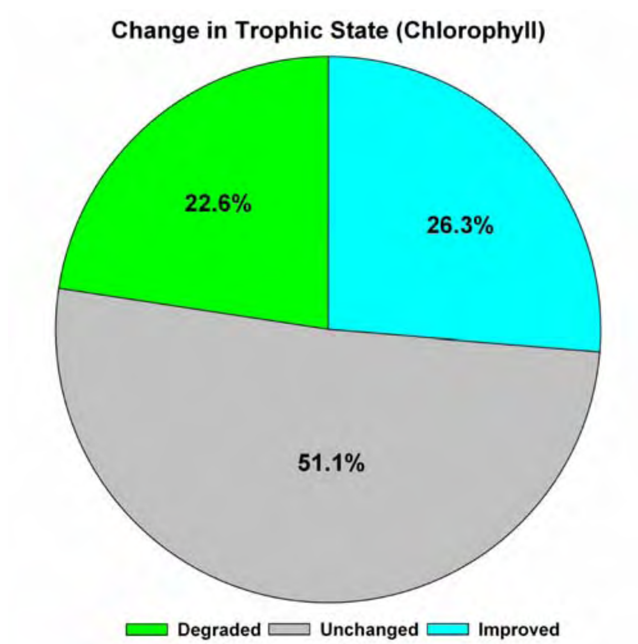


Figure A-4. Change in trophic state of lakes between the 1972 NES and 2007 NLA studies.

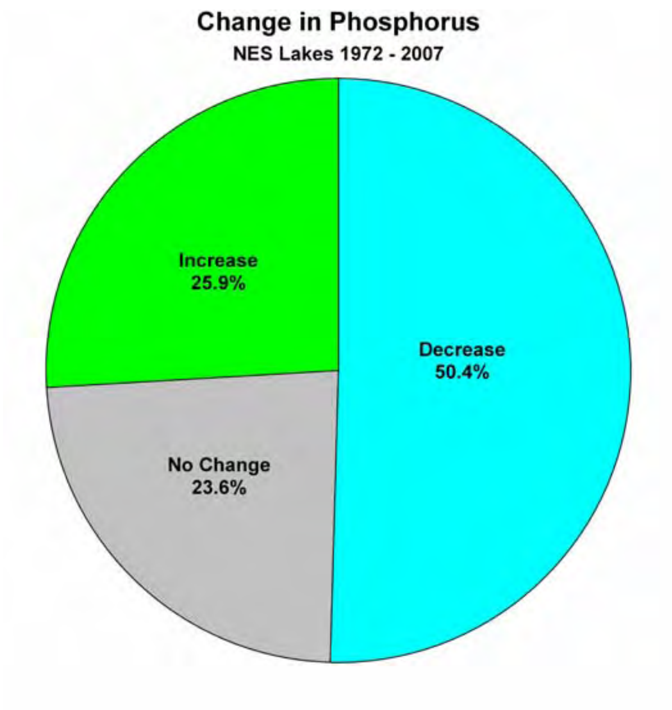


Figure A-5. Change in total phosphorus of lakes between the 1972 NES and 2007 NLA studies.

Diatom Sediment Core Analysis

Sediment Diatom Transfer Function

Variations in fossil diatom species composition were used to assess the amount of change that has occurred in lake systems cored within the NLA since the European settlement. Diatoms are one of the most powerful water quality indicators used in paleolimnological studies. They colonize virtually every freshwater microhabitat and many diatom species have well-defined optima and tolerances for environmental variables such as lake pH, nutrient concentration, water salinity or color (Stoermer and Smol 1999). Thus, they constitute a powerful approach to allow lake managers characterize natural background or reference conditions and to track past changes in lake systems (Smol 1992; Charles et al. 1994).

In order to reconstruct changes in study lakes, the following main steps were taken: 1) Calibration and development of transfer functions; 2) Reconstruction and assessment of the magnitude of change in lake characteristics (see summary of method in Figure A-6).

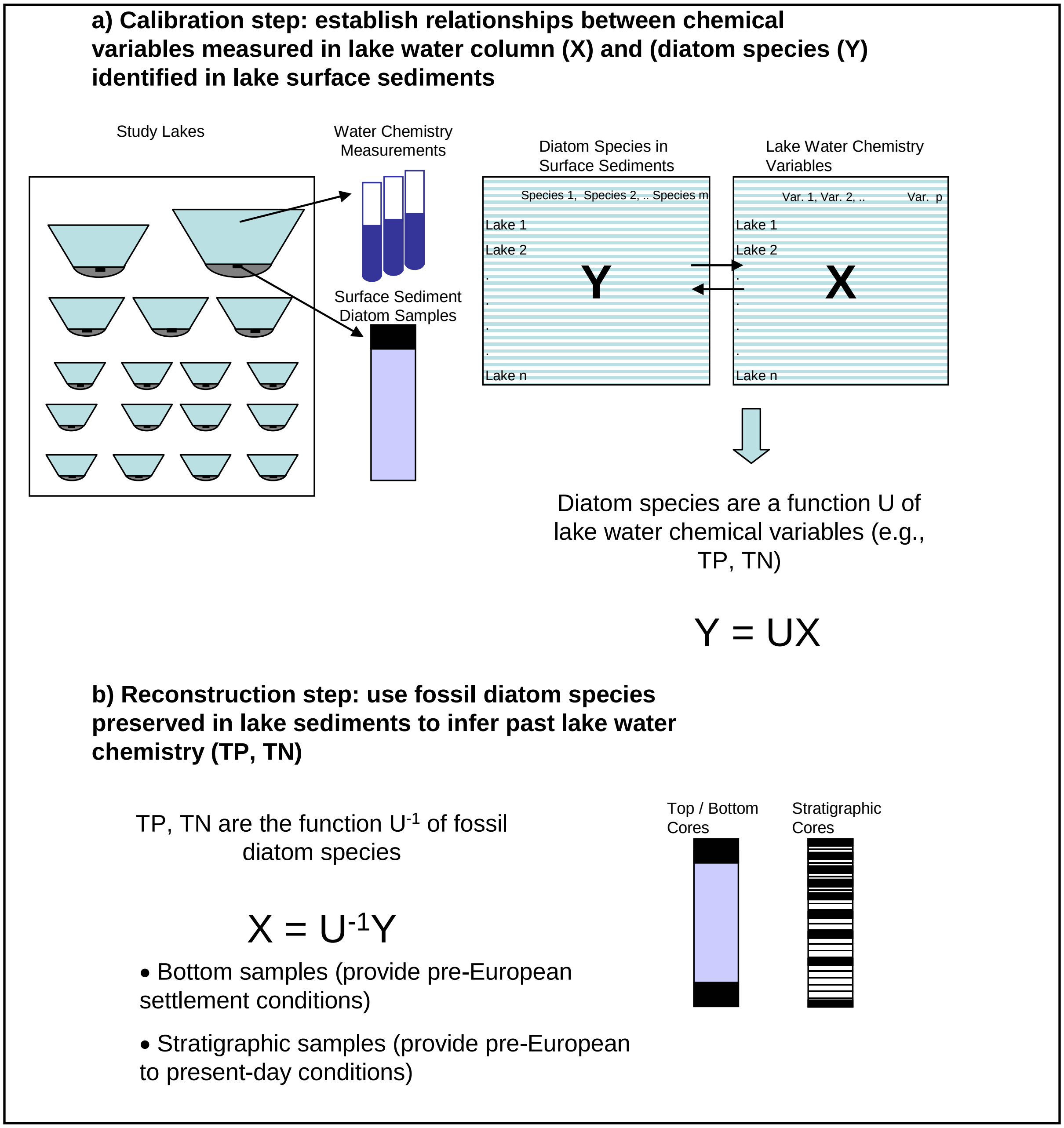


Figure A-6. Quantification of relationships between diatom species in lake surface sediment samples (calibration set) and measured water chemistry ($Y = UX$) and development of transfer functions ($X = U^{-1}Y$) to infer water characteristics (e.g., TP, TN) based on the composition and number of diatom species in top and bottom (or stratigraphic) sediment samples.

Calibration and Development of Transfer Functions

1 Ordinations in reduced space. Prior to the development of diatom-based transfer functions, diatom distributions and relationships with ecological data were explored. This step is necessary to examine whether linear or unimodal methods are appropriate for the available training set in relation to the environmental variables of interest. Detrended Correspondence Analysis (DCA) of diatom species data was performed to determine if the length of environmental gradient is >2.5 SD in order to allow the use of unimodal statistical modeling. Since the environmental gradient was >2.5 SD, unimodal methods were considered appropriate to explore variation in surface sediment diatom assemblages among the lakes and to determine which chemical and other environmental factors explain statistically significant proportions of their variance.

A CCA (and associated Monte Carlo permutation test) using all measured environmental variables was carried out to find a minimal set of variables that explain, in a statistical sense, some variance within the diatom species data. Species relative abundances were log-transformed and down weighted for rare taxa. The statistical significance of each variable was assessed using Monte Carlo unrestricted permutation tests involving 999 permutations (ter Braak, 1989). CCA allowed us to determine how strongly diatom species composition is related to ambient lake water measured parameters. Within the measured variables included in the NLA data set, PTL, COND, NTL, and pH were found to explain highly significant proportions of the species variance ($p=0.001$). CCA also allowed for the identification of outlier samples (i.e., samples with unusual diatom assemblages, unusual combination of environmental variables or a diatom assemblage with poor relationship to environmental variables) that have extreme (more than 5 times) influence and very high squared residual chisquared distance. Other analyses performed within this step were Principal Component Analysis to assess variations in lake physico-chemical caracteristics, and a Pearson correlation matrix with Bonferroni adjusted probabilities to idenrify groups of significantly ($p \leq 0.05$) correlated environmental variables. Multivariate analyses were performed using the computer program CANOCO (ter Braak and Šmilauer 2002) and correlations using the program JMP 5.01 (SAS Institute Inc.).

2. Development of transfer functions. In this step, diatom species from surface sediments (‘modern samples’, constituting the calibration or ‘training’ data set) that are significantly influenced by select variable (i.e., they have well defined optima and narrow tolerances with regard to TP, TN, pH or CON) can be used to develop inference models for these variables. Models to infer TP, TN, conductivity and pH were developed using weighted-averaging partial-least-squares (WA-PLS) models. This reconstruction procedure has the advantage of taking into account the residual correlations that remain after fitting the environmental variable of interest (ter Braak, 1993; Birks 1998). In WA-PLS, the first component is selected to maximize the covariance between the vector of species weighted averages and the environmental variable of interest. Subsequent components are chosen in the same way but with the restriction that they be orthogonal and hence uncorrelated to earlier components (ter Braak, 1993). The number of components to be retained is determined by cross-validation (leaveone-out-jackknifing) on the basis of prediction error sum-of–squares (PRESS), to minimize RMSEP and bias. WA-PLS calibration functions were developed on log-transformed species data. The species data set was based on species names corrected for misspelling and synonyms. The data set includes 220 diatom species that occur with at least 2% relative abundance in at least 6 samples. Transfer Functions computed including species < 2% had lower predictive power except for conductivity. The chemistry data was log transformed except for pH; if chemistry data were available for more than 1 visit, values were averaged.

Calculations were done using the computer program C^2 (Juggins 2003). Error estimates based on cross-validation were provided for each model. The models were subsequently used to infer lake water TP, TN, CON and pH in top and bottom samples and assess the amount of change that occurred in sediment cores.

Table A-11. Transfer functions developed for the national calibration data set

Transfer function	R^2	Jack_R^2	RMSEP_Jack	Equation
TP	0.74	0.67	0.36	Inferred Log TP = 0.4502122 + 0.7038829*Obsrved Log TP
TN	0.67	0.58	0.29	Inferred Log TN = 1.0462799 + 0.6275195*Observed Log TN
pH	0.59	0.52	0.63	Inferred pH = 3.6986058 + 0.542196*Observed pH
Cond	0.78	0.75	0.28	Inferred Log Cond = 0.5794113 + 0.7501362*Observed Log Cond

Models were developed using the whole data set and by splitting data for individual clusters defined within this project. The predictive power was also explored for the original data set (fixed names) and for data sets with lumped diatom species.

Top-Bottom Changes

A fast way to quantify the changes that have affected lake systems, beginning with the European settlement, is to examine how much the composition and relative abundance of sedimentary diatom assemblages changed between the top (representing present-day conditions) and at the bottom of sediment cores, representing the reference, undisturbed conditions (i.e., top-bottom approach). The top-bottom approach provides two ‘snap-shots’ of environmental conditions, before and after human impacts, and has proven successful in addressing diverse environmental questions such as the impact of acid rain, eutrophication or global warming (Cumming et al. 1992; Dixit et al. 1999; Smol et al. 2007).

The transfer functions were used to reconstruct pH, TP, TN, and conductivity in top and bottom samples and assess the amount of change since pre-industrial (pre-European) times in cores that were considered long enough to reach pre-disturbance conditions. Since Pb-210 dating of bottom cores was not performed, we used an alternative approach to evaluate whether or not bottom cores may represent reference conditions (see below).

Determination of whether sediment cores represent pre-disturbance conditions

We assigned each sediment core to one of three categories based on our confidence that the bottom interval represented time prior to European-settlement disturbance typical for the region. "Yes" indicates confidence that the bottom represents a pre-disturbance time period. Usually these are from longer cores and / or from lakes with lower sedimentation rate. "No" means it is unlikely that the interval is sufficiently deep to represent pre-disturbance time. These are usually from shorter cores and/ or lakes with presumed high sedimentation rates. "Uncertain" means that it is difficult to make a determination. This category was used for lakes officially designated "man-made" (reservoirs), oxbow lakes, and others that were borderline in terms of core length, presumed sedimentation rate, and disturbance history.

Category assignments were based on several factors, including sediment core dates from previous studies and evaluation of lake and watershed characteristics that can have a strong influence on sedimentation rates. Key variables considered were nutrient ecoregion, lake cluster (A-G), total percent watershed disturbance, total P, depth, surface area and watershed area. As general principles, lake watersheds with highly erodible soils and high watershed disturbance (especially Ag) tend to have greater input of inorganic particles due to erosion. Watersheds with high percent agriculture and urban tend to have higher algal growth stimulated by increased nutrient inputs. Sediments in shallower lakes might be mixed to a greater depth than deeper lakes. In all the cases above, a longer core would be required to reliably represent pre-disturbance times.

Many final decisions were based on viewing the lake and its watershed using Google Earth. This was a very valuable source of information and showed many important characteristics that otherwise would not have been taken into account (e.g., land-use disturbance patterns, location of shoreline riparian vegetation, local hydrology). Over half the lakes were viewed with GE.

There were 501 lakes with both top and bottom sediment core intervals, and that had sufficient number of diatoms counted (about 30+ lakes had cores for which one or more intervals had a low number of diatoms counted). Of these, 294 were categorized "yes", 106 as "no", and 101 as "uncertain." There were also data for 30 duplicate cores that were not included in the analysis. In most cases, cores from visit 1 were used; visit 2 cores were sometimes used if they were longer.

Even though a sediment core bottom-sample may not be deep enough to represent a pre-settlement time horizon, it can still represent reference conditions if the lake has not been disturbed. We made no attempt to make these types of determinations.

General criteria for categories

Yes

Lakes were accepted for analysis if: they occurred in nutrient ecoregions 2 and 8 where sedimentation rates are known to be relatively low, based on previous studies; lakes with undisturbed, or relatively undisturbed watersheds, and at least moderately long cores for the region; lakes in the Northeast US greater than 25 cm in length were generally considered sufficiently long based on results from the EMAP Surface Water study (NE lakes; ragweed pollen was analyzed in the bottom sediment samples to help confirm pre-disturbance time period).

No

Lakes were rejected if: cores less than 20 cm in length, except a few reference lakes that seemed clearly undisturbed and to have low sedimentation rate; all lakes in nutrient ecoregion 6 with percent watershed disturbance (usually Ag) greater than 50%, bottom sediments in this region with high % Ag would need to be at least 60 cm depth to be in pre-settlement time (Dan Engstrom, Pers. comm.); and all cores in this ecoregion, regardless of percent watershed disturbance, that were less than 30 cm long were not considered for analysis.

Uncertain

Lakes that were considered uncertain were as follows: man-made lakes (reservoirs); date of formation (e.g., dam building) was not known; sedimentation rates were also unknown, and could potentially vary over a wide range; and all oxbow lakes (as determined using Google Earth).

Development of Indices for Lakeshore and Littoral Habitat Condition

Introduction

The physical habitat of a lake includes the environment at the bottom of the lake (*substrate*), the vegetation and substrate along its shoreline (*riparian zone*), and the biotic and abiotic structure of the near shore water (*littoral zone*). Physical habitat condition is critically important to benthic communities, fish and other aquatic organisms.

The NLA and other lake survey and monitoring efforts increasingly rely upon biological assemblage data to define lake condition. Information concerning the multiple dimensions of physical and chemical habitat is necessary to interpret this biological information and comprehensively assess ecological condition. The controlling influence of littoral structure and complexity on lake biota has been long recognized, and recent research highlights the roles of littoral woody debris in providing refuges from predation and affecting nutrient cycling and littoral production. National Lake Assessment field crews characterized lake depth, water surface characteristics, bank morphology and evidence of lake level fluctuations, littoral and shoreline substrate, fish concealment features, aquatic macrophytes, riparian vegetation cover and structure, and human land use activities. These littoral and riparian physical habitat measurements and visual observations were made in a randomized array of 10 littoral plots (10m x 15m) with adjoining riparian plots (15m x 15m) systematically spaced along the shoreline of each sample lake. Metrics describing a rich variety of lake characteristics were calculated from this raw data, and many of these were determined with moderate precision in the national dataset.

For the NLA, we summarized the shoreline and littoral physical habitat information with four integrative measures of lake condition: *RipDist*, incorporating measures of the extent and intensity of human land use activities; *RipVeg,* incorporating the structure and cover in three layers of riparian vegetation, including inundated vegetation; *LitCvr*, a combined biotic cover complexity measure including large woody snags, brush, overhanging vegetation, aquatic macrophytes, boulders, and rock ledges; and *LitRipCvr*, which combines *RipVeg* and *LitCvr* in an index of the cover and complexity of the land-water interface of lakes.

Riparian and littoral habitat structure serves as both an indicator of ecological condition and a context for interpreting biological information. These habitat components are important to lake biological assemblages, providing refuge from predation, living and egg-laying substrates, and food. Shoreline structure also affects nutrient cycling, littoral production, and sedimentation rates. Human activities along lakeshores often adversely affect these ecosystem functions by reducing habitat complexity. Compared with riparian and littoral conditions in lesser disturbed reference lakes throughout the U.S.A, lakes with moderate or high human disturbances in the same region have reduced cover and extent of multi-layered riparian vegetation or natural wetlands. Those with moderate or high disturbance generally also have reduced snag, brush and emergent aquatic macrophyte cover. Our general expectation is that wetland and multi-layered riparian vegetation and abundant, complex fish concealment features foster native fish, macroinvertebrate, and avian assemblage diversity, whereas extensive and intensive shoreline human activities that reduce natural riparian vegetation and reduce littoral cover complexity are probably detrimental to native biota.

Our physical habitat assessment approaches and expectations are based on previous research. In Midwestern lakes, Christensen et al. (1996) reported negative associations between lakeshore cabin development and the density of riparian trees and littoral coarse woody debris, and Jennings et al. (1999) reported cumulative negative effects on fish assemblages as riparian alteration increased. More recent Wisconsin lake studies have found reductions in the quantity of woody debris, and the cover of emergent and floating aquatic macrophytes with increases in cumulative lakeshore human development (Jennings et al. 2003; Hatzenbeler et al. 2004). Radomski and Geoman (2001) also reported loss of emergent and floating-leaf vegetation as a result of human lakeshore development in upper midwest lakes. In the Northeast, shoreline disturbance has been associated with the decline of species richness of native minnows and with an increase in nonnative predator fish species (Whittier et al. 1997). Halliwell (2007,2008) described a number of recent and ongoing studies showing detrimental effects on fish and their habitat resulting from human development of shorelines. Among these, an ongoing study by Merrell et al. (2008, 2009) showed that human development of shorelines resulted in a decrease of woody debris (snag habitat), an increase in sandy shorelines, and an increase in submerged aquatic macrophyte cover in Vermont lakes. Wagner et al. (2006) reported negative effects of residential lakeshore development on littoral fishes and habitat, citing their use of near-shore habitat for nesting, foraging and refuge from predators and adverse conditions. Ness (M.S. 2006 Univ. of Maine) reported that both riparian (shore) and littoral habitat complexity was simplified (at the site scale), with lower densities of trees and shrubs, aquatic macrophytes, and in-lake coarse woody debris.

In an early probability survey of Northeastern lakes, Whittier et al. (2002) reported population estimates of the number of lakes with no direct evidence of human activities in 27% (± 9%). The methods used in this survey (Kaufmann and Whittier, 1997) were modified by Rowan et al. (2006) for use in surveys serving the needs of the European Union's Water Framework Directive, and by the USEPA (2007) for use in the present NLA. Based on these methods, Whittier reported that 67% of NE lakes had relatively undisturbed shorelines, and at these lakes the median canopy-layer tree cover was 67%, with a median combined canopy-layer, mid-layer, and ground-layer woody cover of 170% (of a possible 300%), indicating substantial structural complexity and the potential for sustaining that complexity over time. At the other end of the spectrum, 23% (± 10%) of lakes had at least one type of human structure or activity at half or more of the shoreline stations. This human activity was associated with reduced canopy-layer cover (median = 35%) and three-layer woody cover (median = 82%). Half of these lakes had buildings at more than a third of the shoreline stations. Habitat complexity, in the form of woody snags, overhanging trees, and aquatic plants, was markedly reduced at lakes with higher levels of human activity along the shoreline. The national findings of the NLA reinforced and expanded the geographic extent of these earlier findings for Northeastern U.S. reported by Whittier et al. (2002), who concluded that although stressors such as non-native fish introductions, mercury contamination, and shoreline alteration were not generally considered subjects for environmental management, they were as widespread as eutrophication, and more extensive than acidification, in the lakes of their survey.

Lakeshore disturbances caused by human activity are direct stressors to littoral and riparian habitat, and range in impact from minor effects (such as removal of small areas of riparian tree cover to develop a picnic area) to major alterations (such as construction of a large year-round lakeshore home complete with retaining walls, unstable and erosive landscaping, and concrete shoreline walls and docks). Publications of the effects of lakeshore development on the integrity of littoral habitat present a compelling story. Effects like sedimentation, loss of native plant growth, alteration of native plant communities, loss of habitat structure, and modifications to substrate types are all commonly associated with shoreline human development (Christensen et al., 1996; Whittier et al., 1992, Engel and Pederson, 1998, Merrell et al., Human activities along lakeshores often adversely affect shoreline and littoral structure and ecosystem function by reducing habitat complexity. For example, in the presence of human activity, habitat complexity, in the form of woody snags, overhanging trees, and aquatic plants, becomes markedly reduced, resulting in impacts to macrophytes, fish, and other aquatic biota (e.g., Wagner et al., 2006, Taillon and Fox, 2004, Engel and Pederson, 1998; Whittier et al., 2002). In the following sections, we describe an index of shoreline disturbance and three measures of shoreline habitat: shoreline (or riparian) vegetation cover and complexity, littoral habitat structure and cover for biota, and a combined index of lakeshore and littoral habitat cover and structural complexity.

Lakeshore Disturbance Index

The Lakeshore Distrubance Index *RDis_IX* was based on field observations tallying the presence and proximity of 12 types of human activities or disturbances at 10 systematically located shoreline positions. *RDis_IX* incorporates both the extent of human activities and the intensity of those activities. The extent was expressed simply as the proportion of the shoreline stations that have at least one type of human activity recorded within their 15 x 15 m shore plot and adjacent 10 x 15 m littoral plot (*hifpAnyCirca*). The intensity of human disturbances was expressed by the mean proximity-weighted tally of the number of types of human land-use activities per observation station, both agricultural (*hiiAg*) and non-agricultural (*hiiNonAg*), where disturbances observed outside of the plots were given half the weight of those within the shoreline-littoral plots.

The field procedures tallied nine designated non-agricultural human disturbances: buildings, commercial developments, parks/man-made beaches, docks/boats, walls/dikes/revetments, trash/landfill, road/railroad, power lines, and lawns. Similarly field crews tallied three designated types of agricultural disturbances: row crops, pasture/range/hayfield, and orchard. The field procedures classified Agricultural disturbances into one-third as many categories as for non-Agricultural types. Consequently *hiiAg* ranges from 0 to 1.1, whereas *hiiNonAg* has a range 5.6 times as great (0 to 6.4). To avoid over-representing non-agricultural disturbances and under-representing agricultural disturbances in RDis_IX, the disturbance intensity tallys for agricultural land use were weighted by 5x. This weighting effectively scales agricultural land-uses equal in disturbance potential to those for non-agricultural land use. The index is scaled from 0 to 1, where 0 indicates absence of any human disturbances and 1 indicates extremely high disturbance.

$$RDis\_IX = \frac{1 - \{ 1 / [1 + hiiNonAg + (5 \times hiiAg)]\} + hifpAnyCirca}{2}$$

The same formulation of *RDis_IX* was applied to all ecoregions and all multivariate lake type classes (“CLUS”) in the NLA.

Condition Criteria for RDis_IX:

We applied uniform condition criteria within the NLA:

Low Disturbance *RDis_IX* <u><</u>0.20
Medium Disturbance *RDis_IX* > 20 but <u><</u> 0.75
High Disturbance *RDis_IX* >0.75

Whittier et al (2002) defined thresholds of 25 and 50% of shoreline with human activities as "low" and "high" in an assessment of Northeastern U.S. lakes. The NLA metric *hifpAnyCirca*, a measure of the proportion of lakeshore with one or more disturbance types, is directly comparable to Whittier et al's measure and uses the same field methods. *RDis_IX* is not directly comparable, as it incorporates both the extent and intensity of human activities in its calculation. However, disturbance extent and intensity are correlated, and Table A-12 below shows the regression association of *RDis_IX* with *hifpAnyCirca* in the NLA dataset. Whittier et al's 25 and 50% thresholds roughly correspond with *RDis_IX* thresholds set at 0.34 and 0.54.

Table A-12. Regression association of *RDis_IX* with *hifpAnyCirca*

hifpAnyCirca	*RDIS_IX*
0.00	0.13
0.25	0.34
0.50	0.54
0.75	0.74
0.80	0.78
1.00	0.94

Thresholds of *RDis_IX* (0.20 and 0.75) used in the NLA correspond to *hifpAnyCirca* values of 0.08 and 1.00. The NLA thresholds are more stringent at both ends of the disturbance spectrum than those of Whittier et al. (2002), in the sense that they identify lakes with both lesser and greater amounts of shoreline disturbance. Lakes with *RDis_IX* <u><</u>0.20 have very low levels of lake and near-lake disturbance, and those with *RDis_IX* >0.75 have very high levels of disturbance.

Lakeshore Habitat Index

Indices of riparian cover and complexity were based on visual estimates of vegetation cover and structure in three vegetation layers at 10 evenly-spaced 15 x 15 m plots adjacent to the lake shore. Field data used to calculate indices of riparian cover and complexity included cover-class estimates of large (<u>></u>0.3m dbh) and small diameter (<0.3m dbh) tree cover in the >5m high vegetation layer, woody and non-woody vegetation in the mid-layer (0.5 to 5 m), and woody, non-woody, inundated, and barren classes in the ground cover layer (<0.5 m) of the 10 lakeshore plots. These vegetation classes are virtually identical to those described by Kaufmann et al. (1999) for streams, as is the procedure for converting cover class data to estimates of mean cover for all the types and combinations of vegetation. For each vegetation cover type, field crews estimated areal cover in five classes: absent (0), sparse (0-10%), moderate (10-40%), heavy (40-75%), and very heavy (>75%). Based on cover estimation techniques described by Daubenmire (1968), lakeshore vegetation metrics were calculated by assigning cover class midpoint values (i.e., 0%, 5%, 25%., 57%, and 87.5%) to each plot's observations and then averaging those cover values across all 10 stations.

Three *RVegQ* index formulations were used in the NLA, and were assigned by aggregated ecoregion (ECOWSA9). To calculate lake Riparian Vegetation Cover-Complexity (*RVegQ*) indices, sub-metrics were scaled from 0-1, using dataset maximum values. The summary indices were calculated as the mean of their weighted subcomponents, so also vary from 0 to 1.

RVegQ_2 sums the woody cover in three lakeside vegetation layers and includes inundated vegetation as a positive characteristic. This is appropriate for moist ecoregions (NAP,SAP,UMW,CPL) where tree vegetation can be expected in relatively undisturbed locations.

$$RVegQ\_2 = \frac{\{(rviwoody / 2.5) + rvfcGndInundated\}}{2}$$

RVegQ_7 accommodates lack of tree canopy in ref sites by summing only lower two layers of woody vegetation, where *rviLowWood = rvfcGndWoody + rvfcUndWoody*. It also includes inundated vegetation as a positive characteristic. This index is appropriate for plains ecoregions (NPL, SPL) where tree canopy may not be expected in the absence of human activities, or where presence of tree canopy or enhanced tree canopy cover around lakes may be associated with human activities (e.g. TPL,NPL,SPL).

$$RVegQ\_7 = \frac{\{(rviLowWood / 1.75) + rvfcGndInundated\}}{2}$$

RVegQ_8 sums the woody cover in three lakeside vegetation layers and includes inundated vegetation as a positive characteristic. In contrast *to RVegQ_2*, this index also includes the presence of large diameter trees and accommodates bedrock and boulders as natural shoreline. Sub-metric *ssiNATBedBld* is an index of natural rock shoreline that precludes vegetation. It is calculated as *ssiNATBedBld = ssfcbedrock + sfcboulders*, but the value of *ssiNATBedBld* is set to 0 in lakes that have a substantial amount of anthropogenic seawalls and revetment (i.e., *if hipwWalls* $\geq$0.10). This index is appropriate for ecoregions that have potential to grow large diameter trees, are relatively arid, or lack vegetated lake shorelines at high elevations (WMT, XER).

$$RVegQ\_8 = \frac{(rviwoody/2.5) + rvfpCanBig + rvfcGndInundated + ssiNATBedBld}{4}$$

The formulation of *RvegQ* is assigned by aggregate ecoregion (ECOWSA9):

For ECOWSA9 = NAP, SAP, UMW, CPL:	*RVegQ = RVegQ_2*
For ECOWSA9 = NPL, SPL, TPL:	*RVegQ = RVegQ_7*
For ECOWSA9 = WMT, XER:	*RVegQ = RVegQ_8*

Shallow Water Habitat Index

Indices of littoral cover and complexity were based on field visual estimates of the areal cover of 10 types of littoral cover features within each of ten littoral plots (15 x 10 m) adjacent to the shoreline, and spaced evenly around the periphery of each lake. Field data included cover-class estimates of Woody snags >0.3m diameter, Wood and brush <0.3m diameter, inundated live trees >0.3m diameter, inundated aquatic and herbaceous vegetation, overhanging vegetation <1m above water surface, rock ledges, boulders, and human structures, plus a separate estimation of floating, emergent, and submergent aquatic macrophytes. The cover classes used by the field crews are identical to those described above for riparian vegetation and applied by Kaufmann et al. (1999) to streams, as are the procedures for converting cover class data to estimates of mean cover for all the types and combinations of fish concealment and aquatic macrophyte cover. For each vegetation cover type, field crews estimated areal cover in five classes: absent (0), sparse (0-10%), moderate (10-40%), heavy (40-75%), and very heavy (>75%). Littoral cover metrics were calculated by assigning cover class midpoint values (i.e., 0%, 5%, 25%., 57%, and 87.5%) to each plot's observations and then averaging those cover values across all 10 stations. Mean cover values for various types of cover were summed to yield combined cover metrics.

Three Shallow Water Habitat (*LitCvrQ*) index formulations were used in the NLA, and were assigned by lake type class ("CLUS"). To calculate *LitCvrQ* indices, sub-metrics were scaled from 0-1 using dataset maximum values. The summary indices were calculated as the mean of their weighted subcomponents, so also vary from 0 to 1.

LitCvr_B includes all types of aquatic macrophytes along with the other natural biotic and abiotic cover structure elements included in *fciNatural.* Its added emphasis on snag cover, but lack of additional emphasis on floating and emergent aquatic macrophytes is appropriate for Clus D (warm, low conductivity lakes, mostly in CPL).

$$LitCvr\_B = \frac{fciNatural + (fcfcSnag / 0.2875)}{2}$$

LitCvr_C includes submerged and other types of aquatic macrophytes in *fciNatural*, but its emphasis on floating and emergent forms in addition to snags is appropriate for CLUS A (SAP reservoirs), where presence of submerged aquatic macrophytes indicates water clear enough (low turbidity) for submergent vegetation.

$$LitCvr\_C = \frac{fciNatural + (fcfcSnag / 0.2875) + \{(amfcEmergent + amfcFloating) / 1.515\}}{3}$$

LitCvr_D is the appropriate littoral cover-complexity index formulation for most lake types in the NLA, where increases in submerged aquatic macrophytes is typically associated with nutrient inputs from human disturbances. It excludes submerged aquatic macrophytes, but increases the weighting of floating and emergent macrophytes in addition to snags.

$$LitCvr\_D = \frac{(SomeNatCvr / 1.5) + (fcfcSnag / 0.2875) + \{(amfcFltEmg) / 1.515\}}{3}$$

Where: *SomeNatCvr* = (*fcfcBoulders* + *fcfcBrush* + *fcfcLedges* + *fcfcLivetrees* + *fcfcOverhang*)

amfcFltEmg = (*amfcEmergent* + *amfcFloating*)

The formulation of regional *LitCvrQ* used for each lake was assigned according to lake type ("CLUS"), a geographically-constrained multivariate classification. Within each of three broad geographic areas (Eastern Highlands, Plains and Lowlands, Xeric and Mountain West), lakes were grouped by their size, depth, morphology, depth, elevation, temperature, precipitation, calcium concentration, latitude, and longitude. Assignments were as follows:

For CLUS = B, C, E, F, G: *LitCvrQ* = *LitCvr_D*.
For CLUS = A: *LitCvrQ* = *LitCvr_C*.
For CLUS = D: *LitCvrQ* = *LitCvr_B*.

Physical Habitat Complexity Index

The Physical Habitat Complexity index (*LitRipCvQ*) is simply the arithmetic mean of the respective values for the Riparian and Littoral Cover Complexity indices *RVegQ* and *LitCvrQ*:

$$LitRipCvQ = \frac{(RVegQ + LitCvrQ)}{2}$$

For example, lakes in the NPL and in the SPL lakes that are within lake type "CLUS B" use the Riparian formulation *RVegQ_2* and the Littoral formulation *LitCvrQ_D*. Their formulation for combined Littoral-Riparian Cover Complexity

LitRipCvQ is *LRCvQ_2D* = (*RVegQ_2 + LitCvrQ_D*)/2. Formulations were assigned by Lake Class (CLUS) and Ecoregion (ECOWSA9) as follows:

For CLUS A: *LitRipCVQ = LRCVQ_2C*
For CLUS B: *LitRipCVQ = LRCVQ_2D*
For CLUS C:
  in UMW, CPL: *LitRipCVQ = LRCVQ_2D*
  in NPL, SPL, TPL: *LitripCVQ = LRCVQ_7D*
For CLUS D:
  in CPL: *LitRipCVQ = LRCVQ_2B*
  in SPL, TPL: *LitripCVQ = LRCVQ_7B*
For CLUS E:
  in UMW: *LitRipCVQ = LRCVQ_2D*
  in NPL, SPL, TPL: *LitripCVQ = LRCVQ_7D*
For CLUS F: *LitRipCVQ = LRCVQ_8D*
For CLUS G: *LitRipCVQ = LRCVQ_8D*

Physical Habitat Index Precision and its Interpretation

Physical habitat measurements were repeated at a stratified random subset of 91 NLA sample lakes during the summer 2007 index sampling period. These repeat samples allow an assessment of the within-season repeatability of lake habitat metrics. Table A-13 shows the precision of the four physical habitat condition indicators used in the NLA. The basic measure of repeatability is RMS_{rep}, the Root Mean Square of repeat visits. The RMS_{rep} is a measure of the absolute (unscaled) precision of measurement, and incorporates both measurement and short-term temporal variability.

Table A-13. Precision and distribution characteristics of the four Physical Habitat condition indicators applied in the National Lakes Assessment --- Calculated for the survey as a whole, and separately for three combined ecoregions: the Eastern Highlands, Plains and Lowlands, and the West.

Metric	RMS_{rep}	S/N	Mean/Med	Rg_{pot}	Rg_{obs}	Rg_{obs}/RMS_{rep}
NLA: (df=91)						
RDis_IX	0.115	4.8	0.48 / 0.49	0-1	0 – 0.947	8.2
RVegQ	0.058	2.9	0.17 / 0.16	0-1	0 – 0.558	9.6
LitCvrQ	0.059	2.7	0.12 / 0.09	0-1	0 – 1.0 (0.79)	16.9 (13.4)
LitRipCvQ	0.043	3.9	0.15 / 0.13	0-1	0 – 0.588	11.6
EHigh:(df= 21)						
RDis_IX	0.096	7.0	0.42 / 0.42	0-1	0 – 0.932	9.7
RVegQ	0.052	2.6	0.21 / 0.20	0-1	0 – 0.489	9.4
LitCvrQ	0.060	1.6	0.14 / 0.12	0-1	0.002-0.630	10.5
LitRipCvQ	0.042	2.4	0.18 / 0.17	0-1	0.011-0.457	10.6

PlnNLow (df=49)						
RDis_IX	0.129	3.5	0.52 / 0.54	0-1	0 – 0.936	7.3
RVegQ	0.064	1.9	0.16 / 0.14	0-1	0 – 0.558	8.7
LitCvrQ	0.061	3.4	0.13 / 0.09	0-1	0 – 1.0 (0.79)	16.4 (13.0)
LitRipCvQ	0.043	4.3	0.14 / 0.12	0-1	0 – 0.588	13.7
West: (df=23)						
RDis_IX	0.096	7.1	0.42 / 0.41	0-1	0 – 0.947	9.9
RVegQ	0.046	6.5	0.16 / 0.14	0-1	0 – 0.491	10.7
LitCvrQ	0.054	0.7	0.079 / 0.062	0-1	0 – 0.423	7.8
LitRipCvQ	0.041	3.3	0.12 / 0.11	0-1	0 – 0.421	10.3

The RMS_{rep} for a metric (column 1 of Table A-13) is an estimate of the average standard deviation of that metric if measurements were repeated at all lakes, and standard deviations for each lake were averaged across lakes. It is often scaled by comparing it to some magnitude of variation that is of interest. Alternative scalers might be the magnitude of expected change or the magnitude of an ecologically important difference. It is often difficult to define such a change for a broad survey region. Useful and relevant alternatives are to compare RMS_{rep} to the potential (theoretical) range (Rg_{pot} in Table A-13) or the observed range (Rg_{obs} in Table A-13) of the metric in a survey such as the NLA.

The ratio of Rg_{obs}/RMS_{rep} for metric is an expression of its potential for discerning differences among lakes. The last column of Table A-13, shows that the ratio Rg_{obs}/RMS_{rep} ranged from 7.3 to 18.8 the four Physical Habitat metrics used in the NLA. These results show good potential for these variables to discern lake differences over the ranges observed nationally and in the major subregions.

Another way of scaling the precision of habitat metrics to the "job at hand" is to examine their components of variance. The ratio of variance among lakes to that due to measurement (or temporal) variation within individual lakes has been termed a "Signal-to-noise" ratio, (S/N shown in column 2 of Table A-13). One can think of S/N as the ability of the metric to discern differences among lakes in this survey context. If the among-lake variance in the region or nation is a meaningful variation in lake condition, then the S/N is a measure of the ability of a metric to discern lake condition. This variance-partitioning approach is explained in Kaufmann et al. (1999) and Faustini and Kaufmann (2007), where the authors referred to RMS_{rep} as RMSE and evaluated S/N in stream physical habitat variables. In those publications, the authors generally interpreted precision to be high relative to regional variation if S/N >10, low if S/N <2.0, and moderate if in-between. The NLA physical habitat metrics have mostly moderate precision in this set of data (S/N 2.7 to 4.8 in the national dataset), which means that there can be a substantial, but not crippling influence of measurement "noise" in our classification, regression, plots, and distributions. Larsen et al. (2004) examined the effects of measurement imprecision on the ability of stream physical habitat metrics and sampling designs to detect temporal trends. Kaufmann et al. (1999) and Faustini and Kaufmann (2007) discuss the effect of various levels of S/N on classification, regression and population estimates.

Within subregions of the NLA, one or another of the metrics showed low S/N. Although measured with approximately the same RMS_{rep} in all three regions, the S/N ratio of *RVegQ* ranged from 1.9 in the Plains-Lowlands to 6.5 in the West. This finding suggests that low S/N in the Plains-Lowlands is largely an indication of low variation in riparian vegetation as measured by *RVegQ*, rather than poor measurement precision. However, we may want to explore more sensitive indicators of riparian vegetation condition in the Plains-Lowlands to reformulate *RVegQ* for these regions. Similarly, *LitCvrQ* was measured with approximately the same absolute precision (RMS_{rep}) in all three regions, but S/N ranged from 0.7 in the West to 3.4 in the Plains-Lowlands – this time indicating lack of variation in littoral complexity in the West compared with the other regions, rather than low absolute precision in the West. Again, however, we may want to explore more sensitive indicators of littoral complexity in lakes of the West.

Reference Site Screening

The general NLA strategy was to base most indicator metric expectations on the distribution of indicator values observed in least-disturbed reference sites. These expectations were either site-specific predictions derived by modeling the influence of important non-anthropogenic environmental factors, or were simple statistics concerning the central tendency and distribution of metric values in reference sites within an appropriate region or class of lakes. For biotic and habitat indicators, lakes were first classified into one of seven types, CLUS A through G, based on a geographically-constrained multivariate classification. Within each of three broad geographic areas (Eastern Highlands, Plains and Lowlands, Xeric and Mountain West), lakes were grouped by their similarity in size, depth, morphology, depth, elevation, temperature, precipitation, calcium concentration, latitude, and longitude (see previous sections of this appendix for details). Within each lake type (CLUS), we attempted to obtain approximately 20 or more reference lakes by choosing the least disturbed lakes on the basis of chemical variables and direct observations of agricultural and non-agricultural human disturbances along the lake margin. For each group (CLUS), a series of reference threshold concentrations were established. These varied by group to account for regional variations in water chemistry and littoral-riparian disturbances. Any lake sampled in the survey was considered to be reference if it met every threshold established for the relevant group. Screening parameters were: total phosphorus; total nitrogen; chloride; total sulfate; acid neutralizing capacity, dissolved organic carbon; dissolved oxygen in the epilimnion; proportion of lakeshore with non-agricultural disturbances; proportion of lakeshore with agricultural disturbances; and the relative extent and intensity of human influences of all types together. Following this process, 170 reference lakes (CLASS = Yes) were identified for the entire survey, of which 160 had useable field physical habitat data.

For use in setting lake physical habitat expectations, we modified reference CLASS (used for biological indicators throughout the NLA), by excluding lakes with level control structures <u>in addition to evidence of large lake level fluctuations visible in aerial</u>

photos and field measurements (*bfxhorizdist* $\geq$10m). This lake level fluctuation screen (CLASSP) dropped 11 reference sites from the set used by biological indicators, including 6 in WMT, 1 in XER, 1 in SPL, 3 in SAP; but none in NAP, NPL, TPL, and CPL, and UMW. With one exception in the UMW, all reference sites with the *bfxhorizdist* $\geq$10m also had level-control structures and extensive "bathtub rings" visible in aerial photographs (e.g. Cle Elum Lake in Washington). The one site in the UMW ecoregion that met the *bfxhorizdist* $\geq$10m criterion was not dropped as a reference site because it was a natural lake on very flat topography, and had no level control structures. The revised reference screen yielded 149 least-disturbed reference lakes (CLASSP = Yes), of which 107 were in the UMW, WMT, and NAP (Table A-14). It was difficult to find sites of minimal reference quality in the NPL, SPL, TPL, and XER ecoregions.

Habitat Indicator Expectations

The paucity of reference sites necessitated some analytical strategies for deriving expectations for the various physical habitat indices. First was to coalesce reference sites from relatively similar regions (NPL, SPL, TPL combined into "CENPLN"), or regions where sites varied as much within regions as between regions (WMT and XER combined into "WEST"). The second strategy was to attempt to normalize reference distributions and estimate their percentiles rather than use simple non-parametric statistics (e.g., 5^{th}, 25^{th} percentiles); we did this by log-transforming to approximate normal distributions and to calculate the logarithmic (geometric mean) and logarithmic standard deviations to estimate the percentiles of the reference distributions. The third strategy was to alter the habitat metrics to take into account subregional differences (see above discussion of revisions in *RVegQ, LitRipCvrQ* to accommodate different regional expectations concerning tree canopy vegetation in the Central Plains regions: NPL, SPL and TPL). Lastly, we calculated site-specific expectations for all lakes in the WEST (MTN and XER), based on their latitude, elevation, and subregion.

For the Western United States (ECOWSA9= WMT or XER), expected values of *RVegQ, LitCvrQ*, and *LitRipCvrQ* were modeled by multiple linear regression predictions based on lake elevation, latitude, and subregion (WMT versus XER) in all sites of the WEST excluding highly disturbed lakes (Table A-15). It was necessary to define the controlling effects of these natural factors on a larger set of lakes than only the relatively small number of reference lakes. However, calculating ratio of observed/expected values of the metrics based on moderately disturbed lakes in addition to reference lakes necessitated refining the O/E expectations. We did this by examining the distribution of O/E values in Western reference sites --- the expectation was the geometric mean O/E in reference sites. In all the other regions of the NLA, the geometric mean of the distribution of *RVegQ, LitCvrQ*, and *LitRipCvrQ* values in regional reference sites were set as the expected values for sites within their respective regions. Table A-16 summarizes the reference expectations for *RVegQ, LitCvrQ*, and *LitRipCvrQ* in the nine ecoregions of the NLA.

Table A-14. Number of regional reference sites used for setting Physical Habitat Expectations and Condition

Ecoregion (ECOWSA 9)	Reference sites (CLASSP = Yes)	
NAP	28	
SAP	14	
UMW	41	
CPL	13	
CENPLN:	12	
TPL		(5)
NPL		(1)
SPL		(6)
WEST:	41	
MTN		(38)
XER		(3)
	149	

Setting Condition Criteria

Recall that the high (>0.75) and low (<u><</u>0.20) disturbance criteria for the direct field quantification of lakeshore human disturbances (*RDis_IX*) were set independently of reference site distributions. By contrast, reference site distributions were used to set condition criteria for *RVegQ, LitCvrQ*, and *LitRipCvrQ.* The metrics we examined were observed/expected values (O/E) for each metric, calculated by dividing observed values for each lake by its appropriate expected value. In the WMT and XER, the expectations were lake-specific values calculated from the regression models in Table A-15. For the other regions, the expectations were the region-specific geometric mean values of the respective regional reference lake distributions shown in Table A-16.

We recognize that there is natural variability in expectations that is not captured by our modeling in the WEST, and certainly not by the single regional geometric mean values that serve as expected values in the other regions. Consequently, we compared each lakes O/E value with the statistical distribution of O/E values in the set of regional reference sites. Lakes with O/E values lass than the 5^{th} percentile of the reference distribution were scored in “poor” condition, whereas those above the 25^{th} percentile were scored in “good condition. Those with O/E values between the 5^{th} and 25^{th}

Table A-15. Multiple Linear Regression Models used for calculating reference expectations for Physical Habitat in Western Region (MTN, XER) Lakes

----

<u>Riparian Vegetation:</u>

Expected Log(RVegQ_8) = -1.2108 - (0.000037*ELEV_PT) + (0.0126*FLD_LAT) + (0.1112*WMT)

Rsq=0.2364 RMSE=0.1715 p<0.0001 Model=3 Error=122 CTotal=125 ---- excludes CLASSP=Trash, bfxhorizdist>=10m, and 16 low outliers outside of bounds of ref residuals;

Littoral Cover:

Expected Log(LitCvrQ_D) = -0.9738 - (0.000073*ELEV_PT);

Rsq=0.0500 RMSE=0.3145 p=00075, Model=1df Error=140 CTotal=141 ---- excludes CLASSP=Trash, bfxhorizdist>=10m;

Littoral-Riparian Cover:

Expected Log(LitRipCvQ_8D =
-1.0751- (0.000038*ELEV_PT)+(0.0083*FLD_LAT)-(0.000079*XER_X_ELEV);

Rsq=0.1927 RMSE=0.1869 p<0.0001 Model=3 Error=129 CTotal=132 ---- excludes CLASSP=Trash, bfxhorizdist>=10m, and 9 low outliers outside of bounds of ref residuals;
----

percentiles of the reference distribution were rated in "fair" condition (Table A-17). Note that because the reference lake sample sizes were small, the Log(O/E) mean (i.e., geometric mean) and its standard deviation were used to estimate these percentiles.

PHAB Metric Performance

Examination of the lakeshore human disturbance metric *RDis_IX* across all nine NLA ecoregions (unweighted sample statistics) shows a wide distribution in all ecoregions, but notably higher disturbance in the NPL and lowest disturbance in the WMT and NAP (Figure A-7). Similarly, Riparian Vegetation (*RVegQ*) metric values were highest in the WMT and NAP and lowest in the NPL (Figure A-8A). Although less pronounced, this pattern persisted after *RVegQ* was transformed into O/E values (Figure A-8B).

Littoral Habitat Cover and Complexity (*LitCvrQ*) was notably higher in the Coastal Plain (CPL) than in any of the other regions (Figure A-9A). Lakes in the Northern Plains, by contrast had the lowest *LitCvrQ*, although sample distributions of *LitCvrQ* were relatively low throughout the other inland plains ecoregions (SPL, TPL). After scaling *LitCvrQ* as the O/E variable *LitCvr_OE*, many ecoregions (e.g., WMT, CPL, NAP) had sample median O/E near 1.0 (Figure A-9B). The Northern Plains still had the lowest sample median.

Table A-16. Reference expectations for *RVegQ, LitCvrQ*, and *LitRipCvrQ* in the nine ecoregions of the NLA.

Ecoregion	Metric	Metric Basis	Expected value
NAP	*RVegQ*	Log *RVegQ_2*	0.268
SAP	*RVegQ*	Log *RVegQ_2*	0.235
UMW	*RVegQ*	Log *RVegQ_2*	0.252
CPL	*RVegQ*	Log *RVegQ_2*	0.290

TPL	*RVegQ*	Log *RVegQ_7*	0.176
NPL	*RVegQ*	Log *RVegQ_7*	0.176
SPL	*RVegQ*	Log *RVegQ_7*	0.176
MTN	*RVegQ*	Log *RVegQ_8*	*Lake-specific f(Elevation, Latitude, ECOWSA9)*
XER	*RVegQ*	Log *RVegQ_8*	*Lake-specific f(Elevation, Latitude, ECOWSA9)*
NAP	*LitCvrQ*	Log *LitCvrQ_D*	0.147
SAP	*LitCvrQ*	Log *LitCvrQ*_C(*D)*	0.191
UMW	*LitCvrQ*	Log *LitCvrQ_D*	0.169
CPL	*LitCvrQ*	Log *LitCvrQ_B(D)*	0.299
TPL	*LitCvrQ*	Log *LitCvrQ_D(B)*	0.114
NPL	*LitCvrQ*	Log *LitCvrQ_D*	0.114
SPL	*LitCvrQ*	Log *LitCvrQ_D(B)*	0.114
MTN	*LitCvrQ*	Log *LitCvrQ_D*	*Lake-specific f(Elevation)*
XER	*LitCvrQ*	Log *LitCvrQ_D*	*Lake-specific f(Elevation)*
NAP	*LitRipCvrQ*	Log *LitRipCvrQ_2D*	0.214
SAP	*LitRipCvrQ*	Log *LitRipCvrQ_2C(D)*	0.231
UMW	*LitRipCvrQ*	Log *LitRipCvrQ_2D*	0.220
CPL	*LitRipCvrQ*	Log *LitRipCvrQ_2B(D)*	0.305
TPL	*LitRipCvrQ*	Log *LitRipCvrQ_7D(B)*	0.153
NPL	*LitRipCvrQ*	Log *LitRipCvrQ_7D*	0.153
SPL	*LitRipCvrQ*	Log *LitRipCvrQ_7D(B)*	0.153
MTN	*LitRipCvrQ*	Log *LitRipCvrQ_8D*	*Lake-specific f(Elevation, Latitude, ECOWSA9 x Elev)*
XER	*LitRipCvrQ*	Log *LitRipCvrQ_8D*	*Lake-specific f(Elevation, Latitude, ECOWSA9 x Elev)*

The Coastal Plain (CPL) had the highest combined Littoral-Riparian Cover (*LitRipCvrQ*), followed by the NAP, UMW, SAP and WMT (Figure A-10). The NPL had the lowest median *LitRipCvrQ*, with low medians also in the XER and the other inland plains ecoregions. Once scaled as the O/E variable (*LitRipCvr_OE*), the WMT, NAP, CPL and UMW had median values approaching 1, but lakes in the NPL remained substantially below 0.5 (Figure A-10B).

There were clear, usually progressive, and often substantial declines in habitat cover and complexity (*RVegQ, LitCvrQ*, and *LitRipCvrQ* and their O/E transforms) from minimally disturbed to highly disturbed lakes in most of the nine NLA ecoregions (Figures A-11). While scaling these variables as O/E values masked natural differences in cover and complexity among regions, it facilitated comparisons of condition and impairment across regions (see especially Figure 6). The weakest relationships to disturbance were generally seen in the littoral cover index (*LitCvrQ* and *LitCvr_OE*), especially in the CPL and the two western ecoregions (Figures A-12 and -14). Very strong contrasts in *RVegQ* (and its O/E transform) were seen in many regions, especially CPL, UMW, WMT (Figures A-12 and -14). *LitRipCvrQ* and its O/E transform showed strongest declines with disturbance in the UMW, SPL and WMT (Figures A-12, -13, -14).

Table A-17. Condition criteria for rating lake condition as good, fair and poor. The 5th and 25th percentiles of the reference O/E distributions, respectively, were set as the upper bounds for poor and fair condition. These percentiles were estimated, respectively, the log mean minus 1.65 and 0.67 times the log standard deviation of the reference distribution of the habitat metric shown. They are expressed as antilogs of those values, i.e., as O/E fractions

Ecoregion	Metric	O/E_5th	O/E_25th
CENPLN	*Rveg_OE*	0.548864321	0.783803142
WEST	*Rveg_OE*	0.573040082	0.866209615
NAP	*Rveg_OE*	0.616062821	0.821438398
SAP	*Rveg_OE*	0.548655832	0.783682231
CPL	*Rveg_OE*	0.52679907	0.770851994
UMW	*Rveg_OE*	0.590622517	0.807491575
CENPLN	*LitCvr_OE*	0.277411845	0.594121132
WEST	*LitCvr_OE*	0.271963111	0.591858651
NAP	*LitCvr_OE*	0.469158944	0.735420954
SAP	*LitCvr_OE*	0.338649158	0.644243433
CPL	*LitCvr_OE*	0.46826856	0.734853894
UMW	*LitCvr_OE*	0.415609054	0.700104711
CENPLN	*LitRipCvr_OE*	0.510252334	0.76092701
WEST	*LitRipCvr_OE*	0.578149294	0.860819309
NAP	*LitRipCvr_OE*	0.668243895	0.84901038
SAP	*LitRipCvr_OE*	0.64137939	0.834981782
CPL	*LitRipCvr_OE*	0.586374398	0.805128129
UMW	*LitRipCvr_OE*	0.634351542	0.831254482

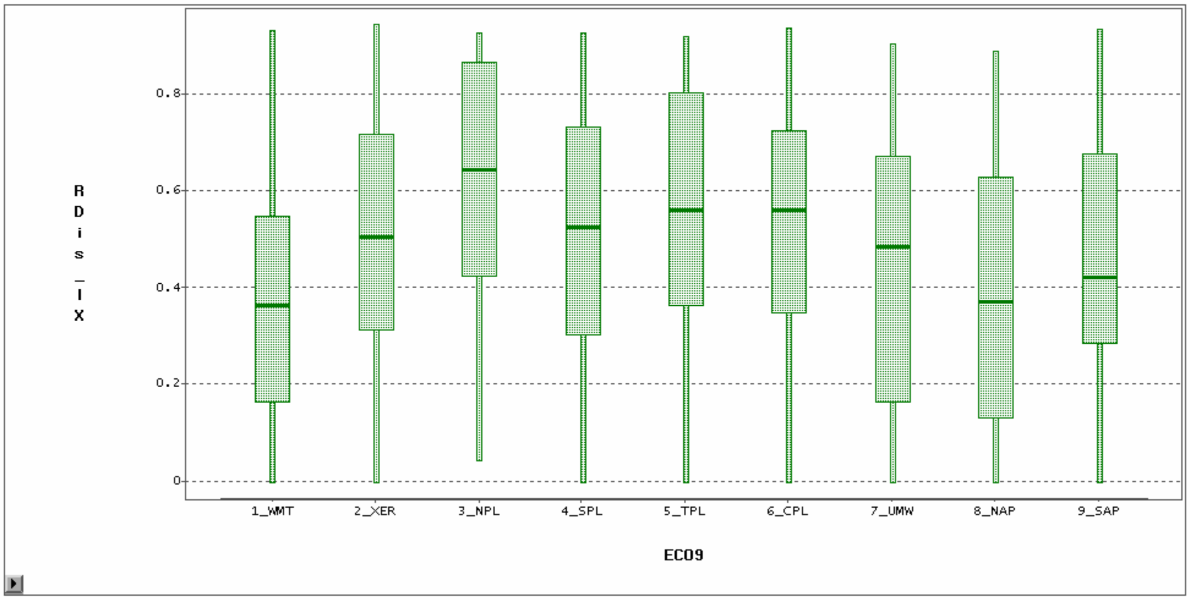
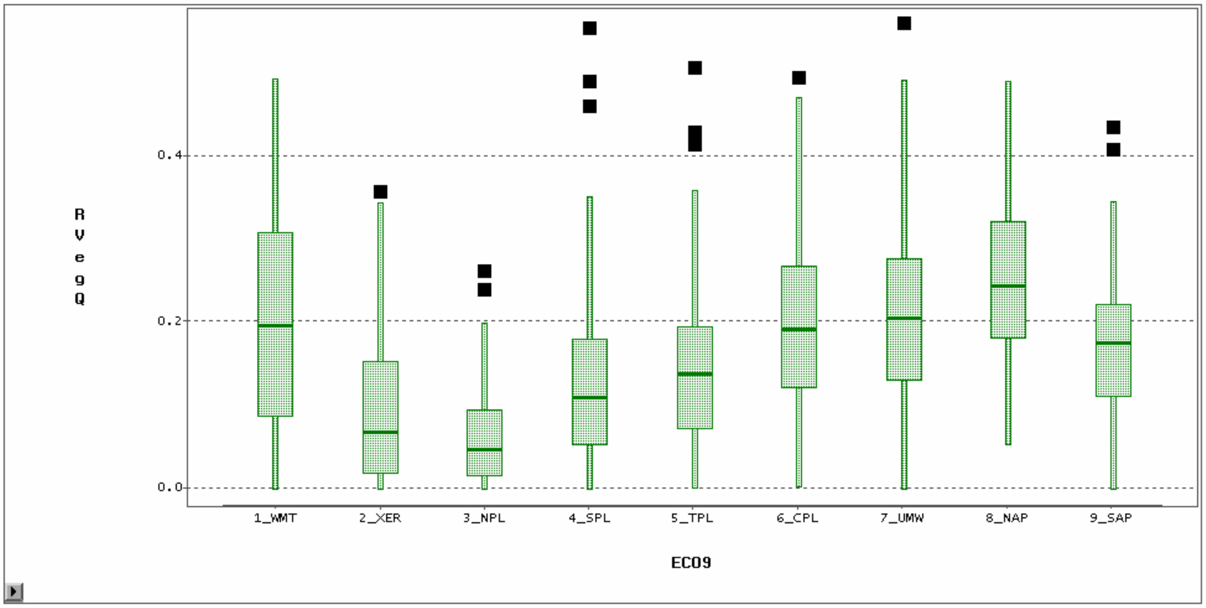


Figure A-7. Comparison of Lakeshore Disturbance (RDis_IX) across nine NLA ecoregions



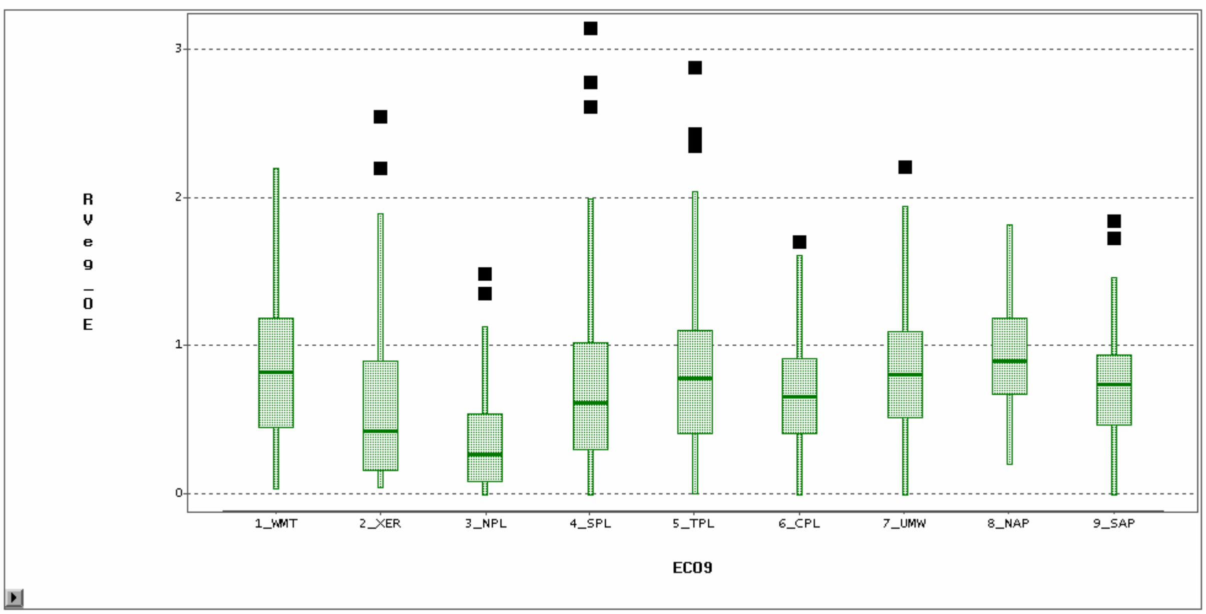
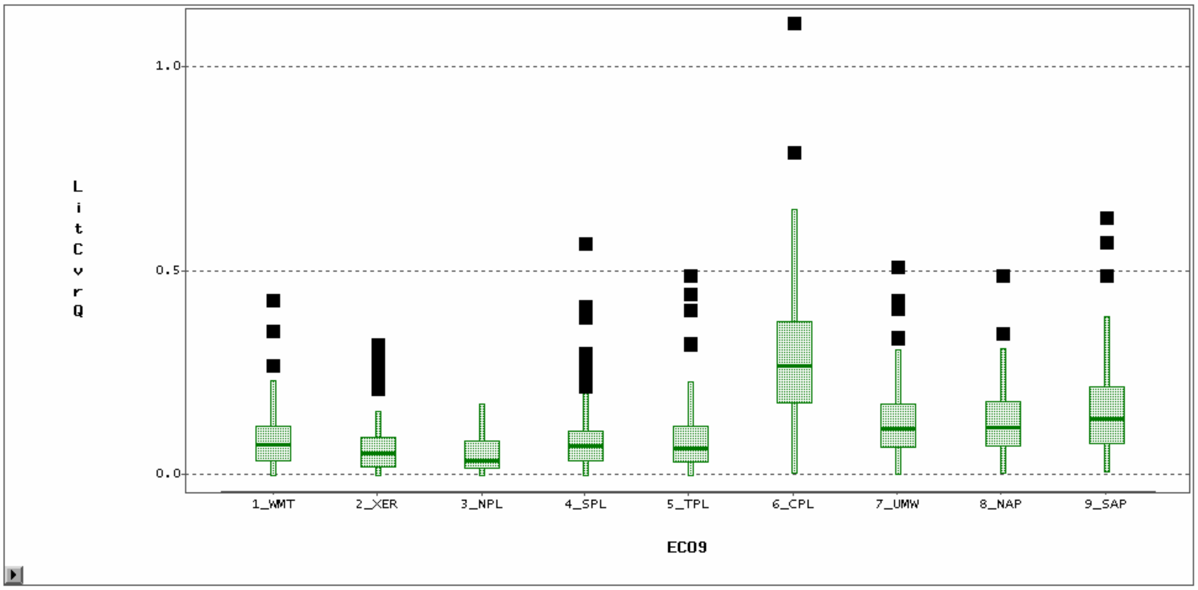


Figure A-8. Comparison of A) Observed, and B) Expected , Lakeshore Riparian Vegetation Structure and Cover (RVegQ) across nine NLA ecoregions



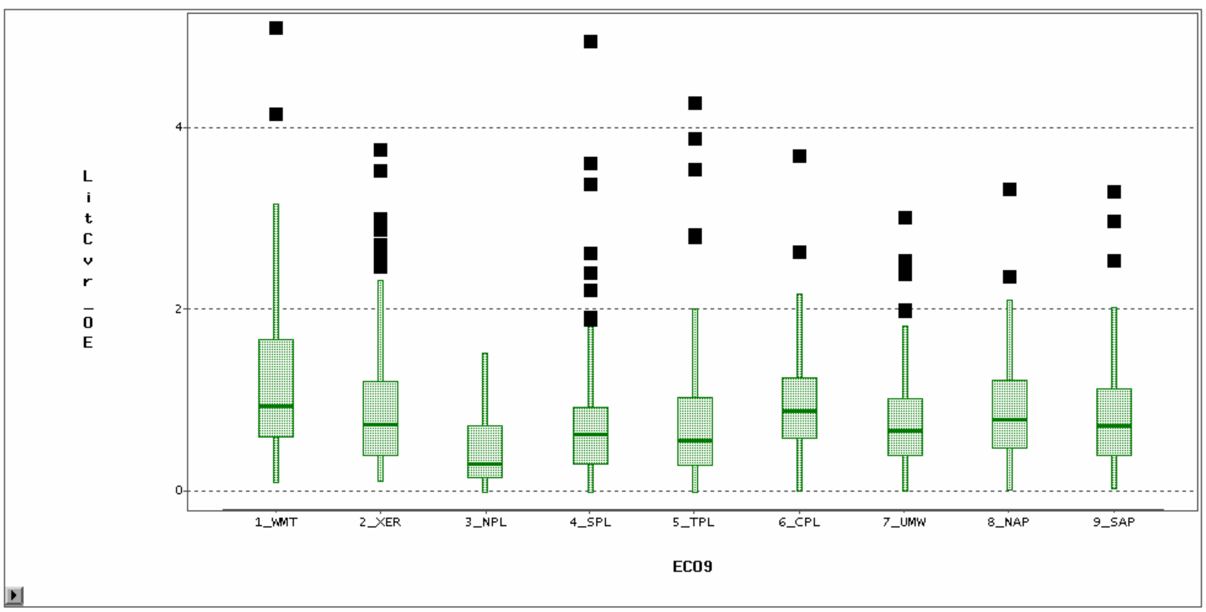
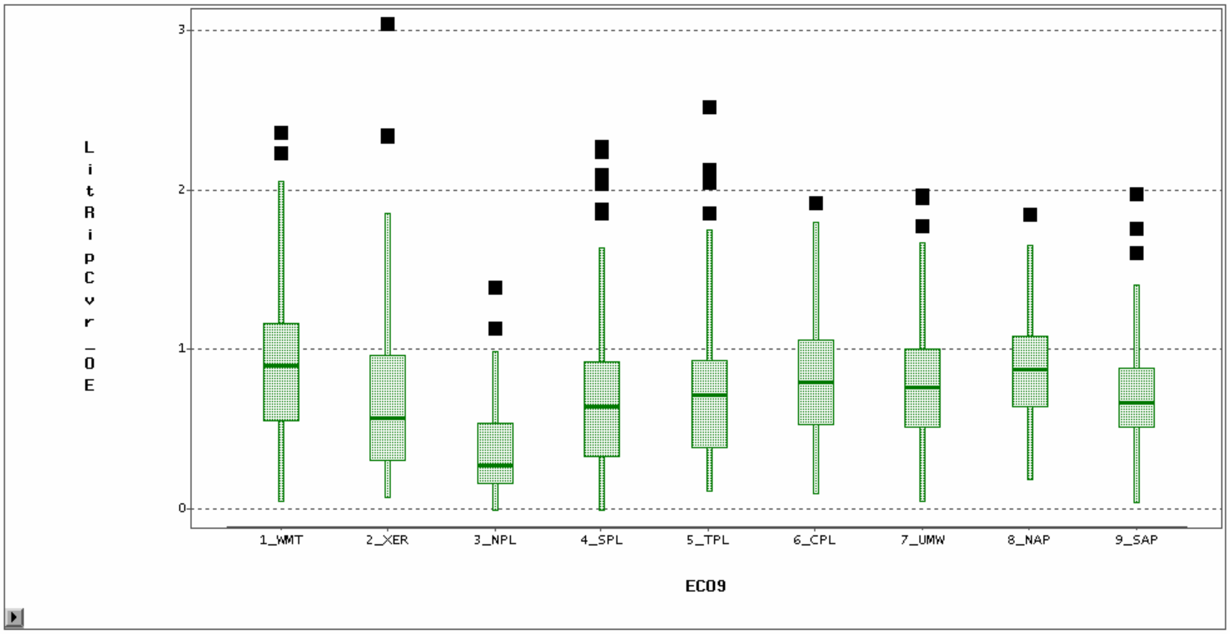


Figure A-9. Comparison of A) Observed, and B) Expected, Littoral Habitat Structure and Cover (LitCvrQ) across nine NLA ecoregions



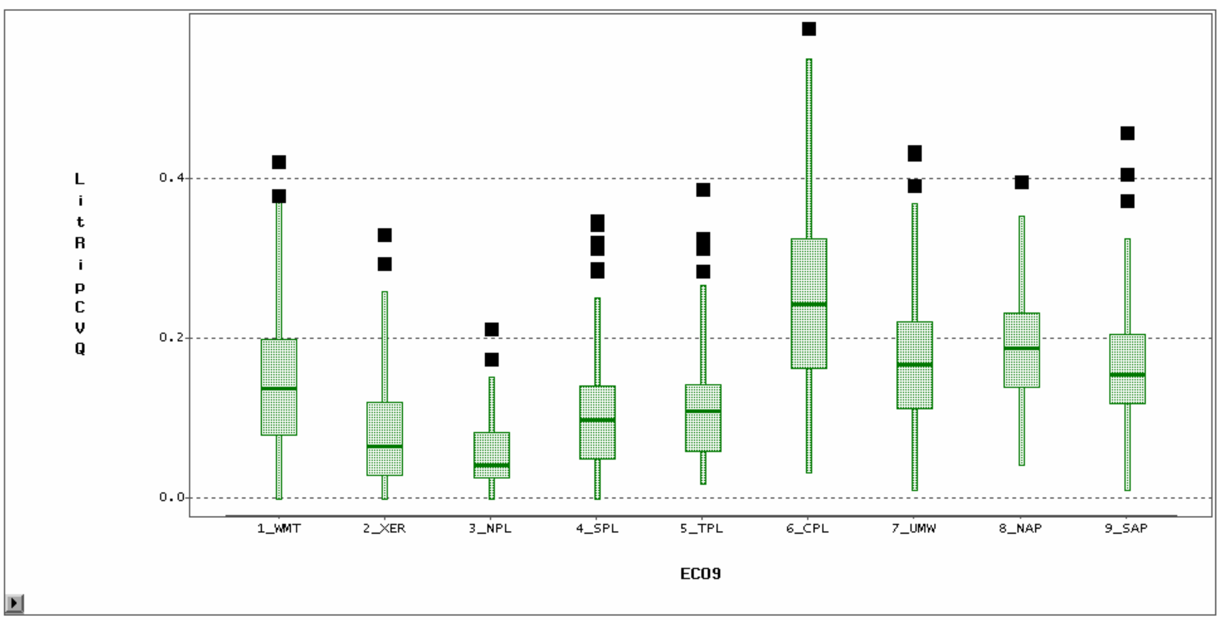


Figure A-10. Comparison of A) Observed, and B) Expected, Littoral and Riparian Habitat Structure and Cover (LitRipCvrQ) across nine NLA ecoregions

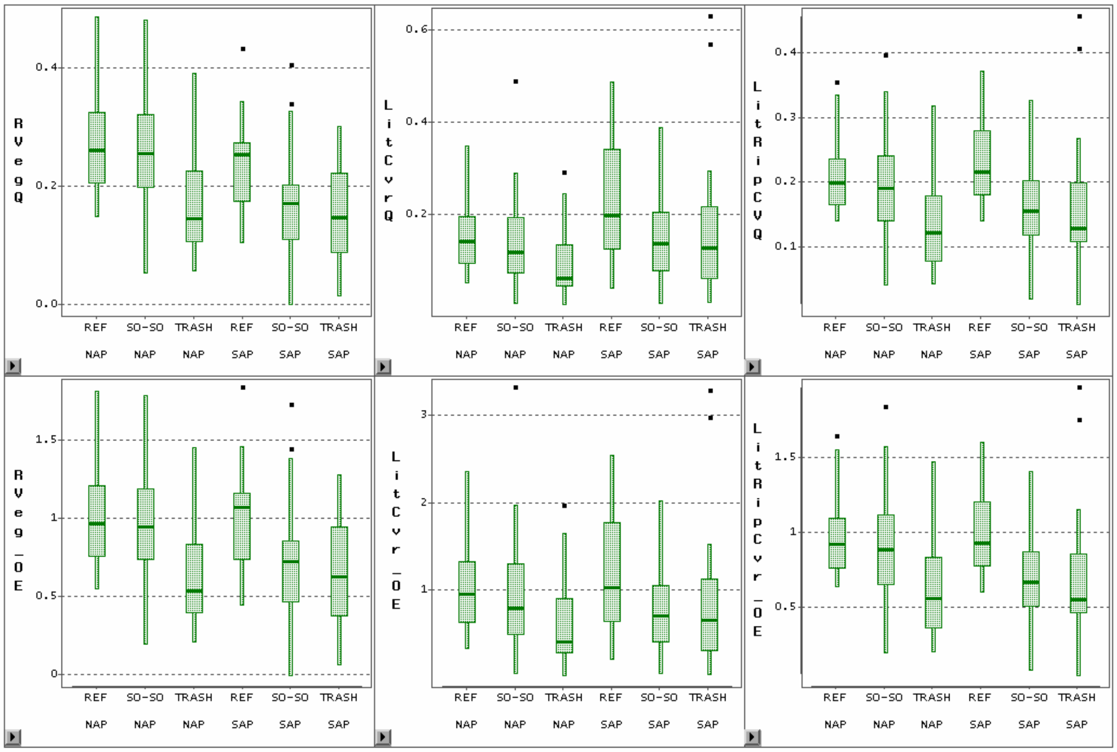


Figure A-11. Contrasts in scaled and unscaled Physical Habitat metrics among reference (Ref), moderately disturbed (So-So), and highly disturbed (Trash) lakes in the Eastern Highlands (Northern and Southern Appalachians)

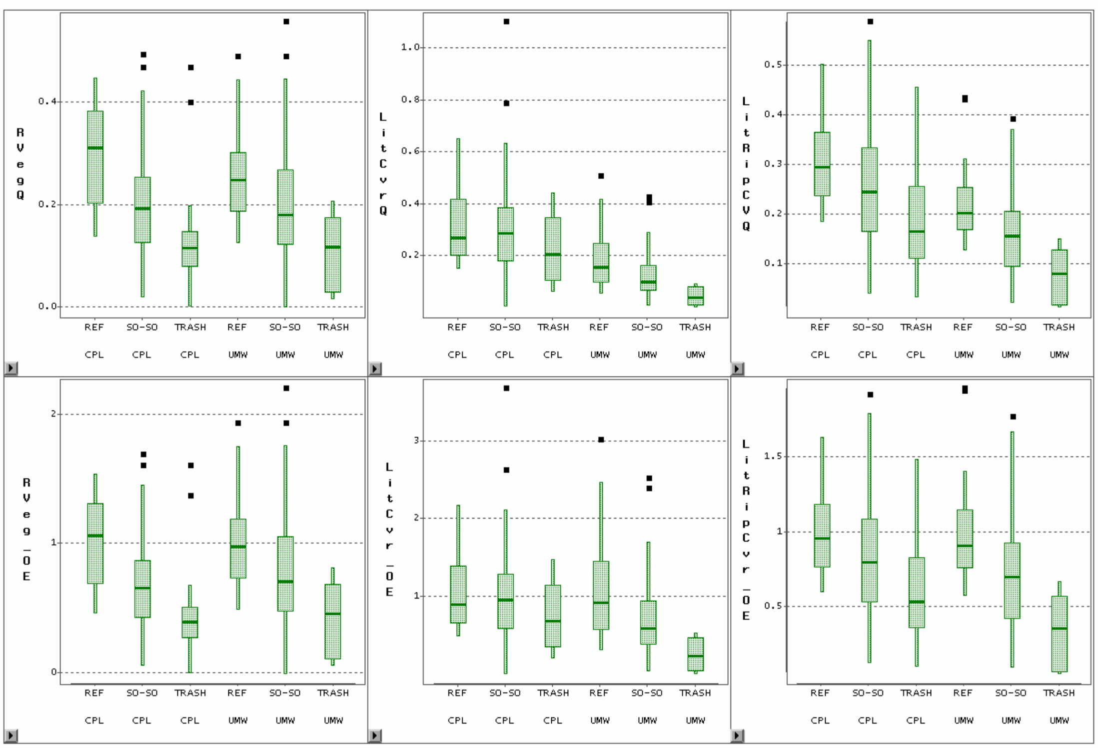


Figure A-12. Contrasts in scaled and unscaled Physical Habitat metrics among reference (Ref), moderately disturbed (So-So), and highly disturbed (Trash) lakes in the Coastal Plain and the Upper Midwest

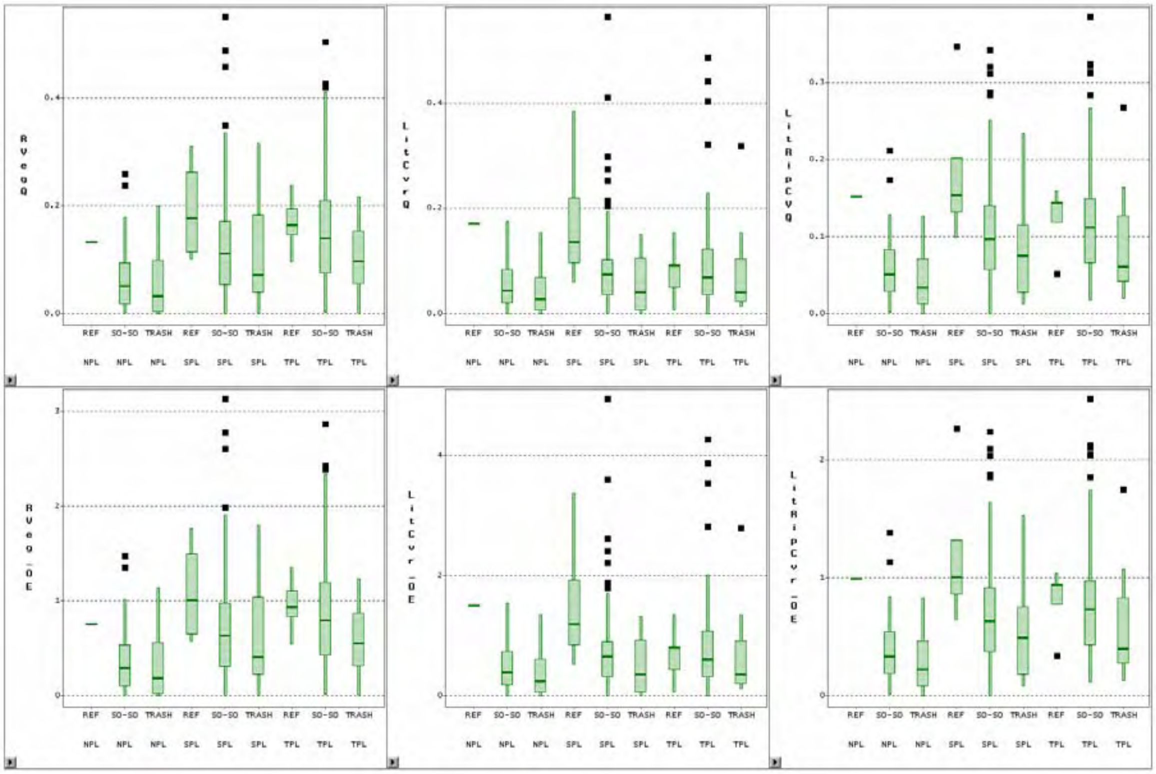


Figure A-13. Contrasts in scaled and unscaled Physical Habitat metrics among reference (Ref), moderately disturbed (So-So), and highly disturbed (Trash) lakes in the Central Plains (Northern, Southern, and Temperate

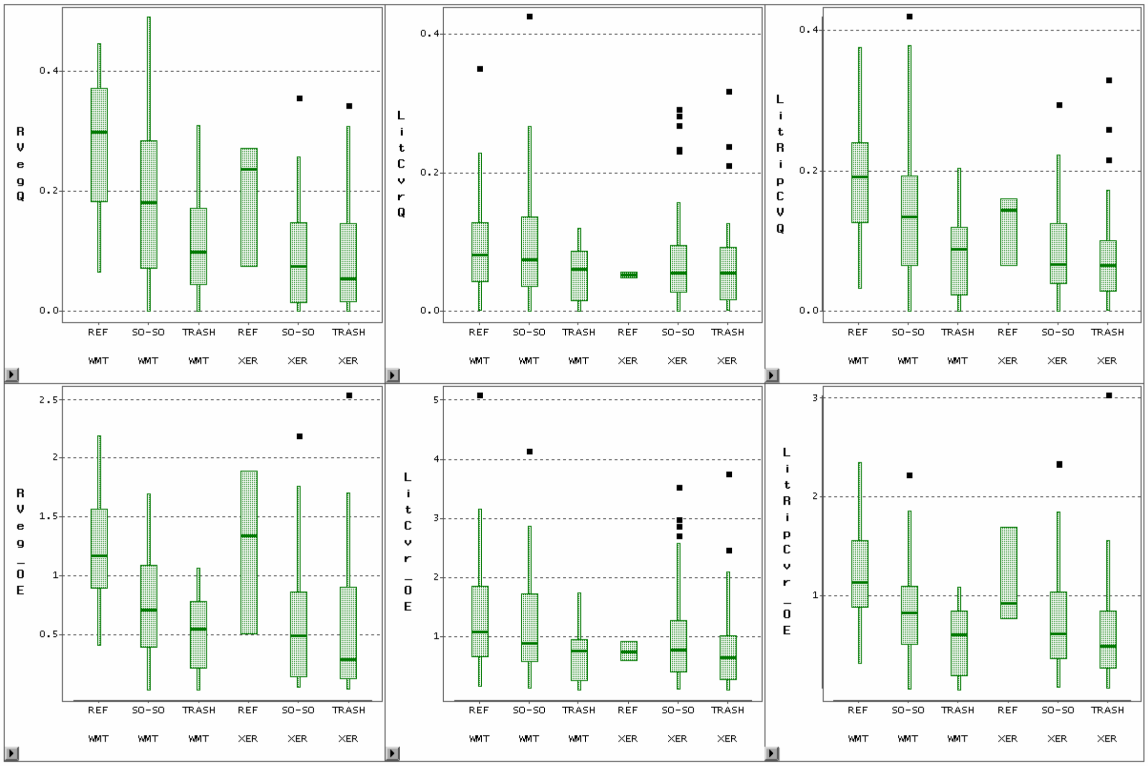


Figure A-14. Contrasts in scaled and unscaled Physical Habitat metrics among reference (Ref), moderately disturbed (So-So), and highly disturbed (Trash) lakes in the Central Plains (Northern, Southern, and Temperate Plains)

NLA Index Precision and Interpretation

Lake condition indicators were repeated at a stratified random subset of 36 to 96 NLA sample lakes during the summer 2007 index sampling period (df_{rep} +1 in Table A-15). These repeat samples allow an assessment of the within-season repeatability of these metrics. Table A-15 shows the precision of a selection of lake condition indicators used in the NLA. The basic measure of repeatability is RMS_{rep}, the Root Mean Square of repeat visits. The RMS_{rep} is a measure of the absolute (unscaled) precision of the whole measurement and analytical process, incorporating also short-term temporal variability within the summer sampling period. One can envision RMS_{rep} for a metric is an estimate of its average standard deviation if measured repeatedly at all lakes, and standard deviations for each lake were averaged across lakes. For Log transformed variables, one can view the antilog of the RMS_{rep} as a proportional standard deviation; something like a . The antilog of 0.179 is 1.51. Then, for example the, RMS_{rep} of 0.179 for Log10(PTL+1) means that the +/- error bound on a measurement in a lake is the measured value times 1.51 and divided by 1.51. So, the +/- 1 StdDev error bounds on a PTL measurement of 10 ug/L during the index period is (10 $\div$1.51) to (10 $\times$1.51) or 6.6 to 15.1.

RMS_{rep} is often scaled by comparing it to some magnitude of variation that is of interest. Alternative scalers might be the magnitude of expected change or the magnitude of an ecologically important difference. It is often difficult to define such a change for a broad survey region. Useful and relevant alternatives are to compare RMS_{rep} to the potential (theoretical) range or the observed range (Rg_{obs} in Table A-15) of the metric in a survey such as the NLA. The ratio of Rg_{obs}/RMS_{rep} for metric is an expression of its potential for discerning differences among lakes. The last column of Table A-15 shows that the ratio Rg_{obs}/RMS_{rep} ranged from 7.6 to 25.9 for the 11 selected NLA metrics. These results show good potential for these metrics to discern lake differences over the ranges observed nationally.

Another way of scaling the precision of metrics to the "job at hand" is to examine their components of variance. The ratio of variance among lakes to that due to measurement (or temporal) variation within individual lakes has been termed a "Signal-to-noise" ratio, (S/N shown Table A-15). One can think of S/N as the ability of the metric to discern differences among lakes in this survey context. If the among-lake variance in the region or nation is a meaningful variation in lake condition, then the S/N is a measure of the ability of a metric to discern lake condition. This variance-partitioning approach is explained in Kaufmann et al. (1999) and Faustini and Kaufmann (2007), where the authors referred to RMS_{rep} as RMSE and evaluated S/N in stream physical habitat variables. In those publications, the authors generally interpreted precision to be high relative to regional variation if S/N >10, low if S/N <2.0, and moderate if in-between. When S/N is over about 10, the effect of measurement error on most interpretations is nearly insignificant within the national context; between 6 and 10 these effects are minor. Between S/N of 2 and 5, the effects of imprecision should be acknowledged, examined and evaluated. From 2 to 4 they are usually adequate to make good-fair-poor classifications, but there is some distortion of CDFs and a significant limitation of the amount of variance that can be explained by approaches such as multiple linear regression (The magnitude of the within-lake variance component limits on the amount of among-lake variance that can be explained by multiple linear regression using single visit data).

S/N for TPL and NPL had high precision (S/N>10) in the national survey. Secchi depth, turbidity, riparian disturbances, and the diatom IBI determined from the top of the sediment cores all had moderately high precision (S/N 4.8-7.1). Chlorophll-a, Riparian and Littoral habitat cover and complexity, and the sediment core bottom diatom IBI had moderate precision in this set of data (S/N 2.0 to 3.9 in the national dataset), which means that there can be a substantial, but not crippling influence of measurement "noise" in classification, regression, plots, and distributions based on those variables. Larsen et al. (2004) examined the effects of measurement imprecision on the ability of stream physical habitat metrics and sampling designs to detect temporal trends. Kaufmann et al. (1999) and Faustini and Kaufmann (2007) discuss the effect of various levels of S/N on classification, regression and population estimates.

Table A-15 Precision and distribution characteristics of diatom IBI and indices of diatom assemblage integrity, nutrient concentrations, trophic status, water clarity, shoreline human disturbance, and lakeshore physical habitat applied in the National Lakes Assessment

Metric	df_{rep}	RMS_{rep}	S/N	Mean/Median	Rg_{obs}	Rg_{obs}/RMS_{rep}
DiatomIBI						
LDC_ADJ(Top)	93	5.37	5.8	-8.8 / -8.5	-50 – +38	16.4
LDC_ADJ(Bottom)	35	8.30	2.0	-9.7 / -9.0	-47 – +45	11.1
Chem-Nutrients						
Log(1+PTL)	95	0.179	12.1	1.51 / 1.41	0.00 – 3.69	20.6
Log(NTL)	95	0.132	11.3	2.79/ 2.76	1.00 – 4.42	25.9
Log(1+CHLA)	95	0.389	3.6	1.04/0.93	0.028 – 2.97	7.6
Log(SECMEAN)	91	0.164	7.1	0.111/0.130	-1.40 – 1.56	18.0
Log(TURB)	95	0.216	7.1	0.619 / 0.563	-0.833 – 2.76	16.6
PHab:						
RDis_IX	90	0.115	4.8	0.48 / 0.49	0 – 0.947	8.2
RVegQ	89	0.058	2.9	0.17 / 0.16	0 – 0.558	9.6
LitCvrQ	88	0.059	2.7	0.12 / 0.09	0 – 1.0	16.9
LitRipCvQ	87	0.043	3.9	0.15 / 0.13	0 – 0.588	11.6

Quality Assurance Summary

The NLA has been designed as a statistically valid report on the condition of the Nation’s lakes at multiple scales, i.e., ecoregion (Level II), and national, employing a randomized site selection process. The NLA is an extension of the EMAP methods for assessing lakes, similar to the 1997 Northeastern Lakes Assessment; therefore, it uses similar EMAP-documented and tested field methods for site assessment and sample collection as the Northeast Lakes Assessment. The NLA collected data on phytoplankton, zooplankton, sediment diatoms, water chemistry and physical habitat.

Key elements of the Quality Assurance (QA) program include:

- **Quality Assurance Project Plan** – A Quality Assurance Project Plan (QAPP) was developed and approved by a QA team consisting of staff from EPA’s Office and Wetlands Oceans and Watersheds (OWOW) and Office of

Environmental Information (OEI) and a Project QA Officer. All participants in the program signed an agreement to follow the QAPP standards. Compliance with the QAPP was assessed through standardized field training, site visits, and audits. The QAPP addresses all levels of the program, from collection of field data and samples and the laboratory processing of samples to standardized/centralized data management.

- **Field training and sample collection** – EPA provided training sessions throughout the study area (with at least one instructor in each session) for all field crew members of each field crew team. All field teams were audited on site within the first few weeks of fieldwork. Adjustments and corrections were made on the spot for any field team problems. To assure consistency, EPA supplied standard sample/data collection equipment and site container packages. 1,028 random site, reference site, and repeat site samples were collected.
- **Water chemistry laboratory QA procedures** – NLA used the same single lab all water chemistry samples. The Western Ecology Division (WED) was responsible for QA oversight in implementing the NLA QAPP and lab standard operating procedures (SOPs) for sample processing.
- **Zooplankton laboratory QA procedures** – NLA used four labs, all four were audited for adherence to the NLA QAPP/SOP for benthic sample processing. This included internal quality control (QC) checks on sorting and identification of zooplankton and the use of the Integrated Taxonomic Information System for correctly naming species collected, as well as the use of a standardized data management system. Independent taxonomists were contracted to perform QC analysis of 10% of each labs samples (audit samples).
- **Phytoplankton laboratory QA procedures** – NLA used one lab, this lab was audited for adherence to the NLA QAPP/SOP for benthic sample processing. This included internal quality control (QC) checks on sorting and identification of phytoplankton and the use of the Integrated Taxonomic Information System for correctly naming species collected, as well as the use of a standardized data management system. Independent taxonomists were contracted to perform QC analysis of 10% of each labs samples (audit samples).
- **Sediment Diatom laboratory QA procedures** – NLA used four labs, all four were audited for adherence to the NLA QAPP/SOP for benthic sample processing. This included internal quality control (QC) checks on sorting and identification of sediment diatoms. Independent taxonomists were contracted to perform QC analysis of 10% of each labs samples (audit samples).
- **Entry of field data** – NLA used a standardized data management structure, i.e., the same standard field forms for data collected in the field, with centralized data entry through scanning in to electronic data files. Internal error checks were used to confirm data sheets were filled out properly.

- **Records management** – These records include (1) planning documents, such as the QAPP, SOPs, and assistance agreements and (2) field and laboratory documents, such as data sheets, lab notebooks, and audit records. These documents are ultimately to be maintained at EPA. All data are archived in the STORET data warehouse at www.epa.gov/STORET.

References

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1997. DRAFT Revision to Rapid Bioassessment Protocols for Use in Streams and Rivers. 841-D-97-002. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

Barbour, M.T., J. Gerritsen, B.D. Snyder, J.B. Stribling. 1999. Rapid Bioassessment Protocols: For Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish. EPA 841-B-99-002. U.S. Environmental Protection Agency, Washington, D.C.

Bauer, S.B., and T.A. Burton. 1993. Monitoring Protocols to Evaluate Water Quality Effects of Grazing Management on Western Rangeland Streams. EPA 910/9-91-001. U.S. Environmental Protection Agency, Region X, Seattle, WA. 166 p.

Benson, B.J. and J.J. Magnuson. 1992. Spatial heterogeneity of littoral fish assemblages in lakes: relation to species diversity and habitat structure. Canadian Journal of Fisheries and Aquatic Sciences 49:1493-1500.

Blocksom, K. A. 2003. A Performance Comparison of Metric Scoring Methods for a Multimetric Index for Mid-Atlantic Highlands Streams.B.A. Markert, A.M. Breure and H.G. Zechmeister (ed.), Bioindicators and Biomonitors, Chapter14. Environmental Management. Springer-Verlag, New York, NY, 31(5):670-682.

Birks, H.J.B. 1998. Numerical tools in paleolimnology – Progress, potentialities, and problems. Journal of Paleolimnology. 20: 307-332.

Charles, D.F., J.P. Smol, and D.R. Engstrom. 1994. Paleolimnological approaches to biological monitoring. In: Biological monitoring of aquatic systems, edited by L. L. Loeb and A. Spacie, CRC Press:Boca Raton, p. 233-293.

Christensen DL,Herwig BR, Schindler DE, Carpenter, SR. 1996. Impacts of lakeshore residential development on coarse woody debris in north temperate lakes. Ecological Applications 6: 1143–1149.

Cumming, B.F., Smol, J.P., Kingston, J.C., Charles, A.F., Birks, H.J.B., Camburn, K.E., Dixit, S.S., Uutala, A.J., and Selle, A.R. 1992. How much acidification has occurred in Adirondack region (New York, USA) lakes since pre-industrial times? Canadian Journal of Fisheries and Aquatic Sciences 49: 128-141.

Daubenmire, R. 1968. Plant Communities: A Textbook of Plant Synecology. Harper & Row, New York. 300 p.

Dixit, S.S., J.P. Smol, D.F. Charles, R.M. Hughes, S.G. Paulsen, and G.B. Collins. 1999. Assessing lake water quality changes in the lakes of the northeastern United States using sediment diatoms. Canadian Journal of Fisheries and Aquatic Sciences 56: 131-152.

Dunne, T., and L.B. Leopold. 1978. Water in Environmental Planning. W.H. Freeman, New York. NY. 818 p.

Dufrene, M., and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs 67:345-366

Faustini, J.M. and P.R. Kaufmann. 2007. Adequacy of Visually Classified Particle Count Statistics from Regional Stream Habitat Surveys. Journal of the American Water Resources Association 43(5):1293-1315.

Halliwell, D. 2007. Lake habitat measures. New England Lake News -- New England Chapter of the North American Lake Management Society 2(2):1-4.

Halliwell, D. 2008. Lake habitat measures. Out of the Blue --A Newsletter of the Water Quality Division,Vermont Agency of Natural Resources Department of Environmental Conservation. 33:8-9.

Hatzenbeler, G.R., J.M. Kampa, M.J. Jennings, and E.E. Emmons. 2004. A comparison of fish and aquatic plant assemblages to assess ecological health of small Wisconsin lakes. Lake and Reservoir Management 20:211-218.

Hill, A. B. (1965). "The environment and disease: association or causation." Proceedings Royal Society Medicine 58: 295-300.

Hynes, H.B.N. 1972. Ecology of Running Waters. Univ. of Toronto Press, Canada. 555p.

Jennings, M.J., M.A. Bozek, G.R Hatzenbeler, E.E. Emmons, and M.D. Staggs. 1999. Cumulative effects of incremental shoreline habitat modification on fish assemblages in north temperate lakes. North American Journal of Fisheries Management 19:18-27.

Jennings, M.J., E.E. Emmons, G.R. Hatzenbeler, C. Edwards, and M.A. Bozek. 2003. Is littoral habitat affected by residential development and land use in watersheds of Wisconsin lakes? Lake and Reservoir Management 19:272-279.

Juggins, S. 2003. C^2 User guide. Software for ecological and paleoecological data analysis and visualization. University of Newcastle, Newcastle upon Tyne, UK. 69pp.

Karr, J.R. 1993. Defining and assessing ecological integrity beyond water quality. Environmental Toxicology and Chemistry 12:1521-1531.

Kaufmann, P.R., P. Levine, E.G. Robison, C. Seeliger, and D.V. Peck. 1999. Quantifying Physical Habitat in Wadeable Streams. EPA 620/R-99/003. Environmental Monitoring and Assessment Program, U.S. Environmental Protection Agency, Corvallis, OR.

Kaufmann, P.R. and T.R. Whittier. 1997. Habitat Assessment. Pages 5-1 to 5-26 In: J.R. Baker, D.V. Peck, and D.W. Sutton (Eds.). Environmental Monitoring and Assessment Program -Surface Waters: Field Operations Manual for Lakes.EPA/620/R-97/001. U.S. Environmental Protection Agency, Washington, D.C.

Larsen, D.P, P.R. Kaufmann, T.M. Kincaid, and N.S. Urquhart . 2004. Detecting persistent change in the habitat of salmon-bearing streams in the Pacific Northwest. Canadian Journal of Fisheries and Aquatic Sciences 61:283-291.

Merrell, K. et al. (2008 in Capen et al. 2008) Merrell, K., E. Howe and S. Warren. 2009. Examining Shorelines, Littorally. Lakeline 29:1.

Radomski, P. and T.J. Geoman. 2001. Consequences of human lakeshore development on emergent and floating-leaf vegetation abundance. North American Journal of Fisheries Management 21:41-46.

Rowan, J.S., J.Carwardine, R.W.Duck, O.M.Bragg, A.R. Black, J.E.J. Cutler, I. Soutar, and P.J. Boon. 2006. Development of a technique for lake Habitat Survey (LHS) with applications for the European Union Water Framework Directive. Aquatic Conservation 16(6):637-657. Doi: 10.1002/aqc.786 .

Smol, J.P. 1992. Paleolimnology: An important tool for effective ecosystem management. Journal of Aquatic Ecosystem Health 1:49-58.

Smol J.P., Wolfe A.P., Birks H.J.B., et al. 2007. Climate-driven regime shifts in the biological communities of arctic lakes. Proceedings of the National Academy Sciences 102: 4397-4402.

Stoermer E.F. and J.P. Smol. 1999. Applications and uses of diatoms: prologue. In Stoermer E.F. and Smol J.P. (eds) The diatoms: Applications for the Environmental and Earth Sciences, Cambridge, Cambridge University Press, pp 3-8.

Taillon, D. and M. Fox. 2004. The influence of residential and cottage development on littoral zone fish communities in a mesotrophic north temperate lake. Environmental Biology of Fishes 71: 275-285 ter Braak, C.J.F. and van Dam, H. 1989. Inferring pH from diatoms: a comparison of old and new calibration methods. Hydrobiologia 178: 209-223.

ter Braak, C. J. F., and S. Juggins. 1993. Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. Hydrobiolgia: 269/270: 485-502.

ter Braak, C.J.F., and P. Šmilauer. 2002. CANOCO reference manual and CanoDraw for Windows. User's guide: software for canonical community ordination. Version 4.5. Microcoputer Power, Ithaca, NY, USA.

USEPA 2007 NLA Field Methods USEPA. 2007. Survey of the Nation's Lakes. Field Operations Manual. EPA 841-B-07-004. U.S. Environmental Protection Agency, Washington, DC.

Wagner, T., A. Jubar, M. Bremigan. 2006. Transactions of the American Fisheries Society 135:843–852.

Whittier TR, Halliwell DB, Paulsen SG. 1997. Cyprinid distributions in northeast USA lakes: Evidence of regional-scale minnow biodiversity losses. Canadian Journal of Fisheries and Aquatic Sciences 54: 1593–1607.

Whittier, T.R., S.G. Paulsen, D.P. Larsen, S.A. Peterson, A.T. Herlihy, and P.R. Kaufmann. 2002. Indicators of ecological stress and their extent in the population of Northeastern lakes: a regional-scale assessment. Bioscience 52(3):235-247.

Van Sickle, J. J.L. Stoddard, S.G. Paulsen and A.R. Olsen. 2006. Using relative risk to compare the effects of aquatic stressors at a regional scale. Environmental Management 38:1020-1030.

Van Sickle, J. and S.G. Paulsen. 2008. Assessing the attributable risks, relative risks, and regional extents of aquatic stressors. Journal of the North American Benthological Society 27:920-931.