

Standard Operating Procedure #9: Data Analysis

Version 1.2 (2021)

In Water Quality Monitoring Protocol for Inland Lakes: Great Lakes Inventory and Monitoring Network

Prepared by

Joan Elias and David VanderMeulen*

National Park Service

Great Lakes Inventory and Monitoring Network

2800 Lake Shore Drive East

Ashland, Wisconsin 54806

* *Contact author: david_vandermeulen@nps.gov*

SOP9 – Contents

	Page
Revision History Log.....	iii
Acknowledgements.....	iii
9.0 Introduction.....	2
9.1 Temporal and Spatial Domains.....	2
9.2 Initial QA/QC Checks for Outliers in Data.....	3
9.3 Annual Data Summaries	3
9.4 Analyses of Long-Term Trends	4
9.5 Literature Cited	7

Revision History Log

The following table lists all edits and amendments to this document since the original publication date. Information entered in the log must be complete and concise. Users of this standard operating procedure will promptly notify the project manager and/or the Great Lakes Network (GLKN) data manager about recommended and required changes. The project manager must review and incorporate all changes, complete the revision history log, and change the date and version number on the title page and in the header of the document file. For complete instructions, please refer to SOP #13: Revising the Protocol.

Previous Version #	Revision Date	Author (with title and affiliation)	Location in Document and Concise Description of Revision	Reason for Change	New Version #
1.0	5/01/2015	David VanderMeulen, GLKN	Updated the NPS contact author, distinguished discrete from continuous data, fixed a handful of grammatical errors throughout	To clarify and update SOP based on changes that have occurred since Version 1.0 was published in 2008	1.1
1.1	3/2021	Alexander Egan, Aquatic Ecologist, GLKN	Minor text changes.	To clarify and update SOP based on changes that have occurred since Version 1.1 was published in 2015	1.2
Add rows as needed for each change or set of changes tied to an updated version number					

Acknowledgements

I would like to acknowledge Brenda Moraska Lafrancois and Richard Axler for helpful input and discussions relative to analysis of water quality data and general limnological principles. Thanks also to Sue Magdalene and Erik Beever for providing verbiage in earlier drafts of this document.

9.0 Introduction

Data analysis is the process by which measurements of the environment are interpreted meaningfully. It begins with evaluations of data after the data have been collected and entered into an electronic file or database, and have undergone a check of the data entry to ensure quality. Data analysis includes quality assurance/quality control (QA/QC) checks for statistical outliers prior to data summarization, may include exploratory data analysis, and concludes with analyses that lead to summary and interpretations of the data.

Well-conceived and -developed monitoring strategies have clear connections between questions of interest, appropriate sampling designs, and resulting analytical approaches (Noon 2003). Accordingly, the utility and robustness of the Great Lakes Network's (GLKN) analyses for monitoring water quality are based on ecologically meaningful questions and relationships, which prescribe the monitoring design, which in turn prescribes an analytical approach. Increasingly, biometricians are advocating that ecologists seek to elucidate and quantify ecologically important phenomena, rather than exclusively pursue statistical significance (Yoccoz 1991, Johnson 1999, Anderson et al. 2001). Efforts of GLKN to monitor water quality seek to provide a quantitative understanding of the effect size (e.g., temporal trend, difference in indicator values among pre-defined strata) as well as the repeatability of that result (i.e., a measure of precision or uncertainty associated with the estimate). We are striving to address directed monitoring questions that reflect our prior knowledge of the system and may provide useful information for management decisions, rather than test myriad hypotheses about ecosystem change.

In addition to quantitatively describing the status and temporal trends of water quality indicators, a secondary goal is to begin to understand the dynamics and drivers of our indicators, following our conceptual models (Gucciardo et al. 2004). Although not every trend is a product of local management action, tests of association that address the underlying 'why' questions behind the 'what' questions in trend analysis will be explored for at least a subset of our metrics of water quality. Although the validity of these associations can be strengthened with focused research, these types of questions increase the likelihood that our monitoring can lead to an early correction of trends before the cascading ramifications become irreversible.

9.1 Temporal and Spatial Domains

As described in the protocol narrative for monitoring water quality in GLKN inland lakes (Elias et al. 2008), we have defined our temporal domain for discretely collected data as daytime periods during the ice-free months, ideally, when lakes are thermally stratified. Selected lakes will be sampled three times annually, at the same location within each lake. We selected lakes using several criteria, including lake type based on ordinations of past chemistry data to determine groupings of similar lakes within a given park, amount of past data, spatial balance within each park, and particular interest by park staff. We also selected lakes to span various gradients within a park, such as gradients across intensity of visitor use, surface area, maximum depth, and watershed area. Continuous water temperature data will be collected from a subset of inland lakes at Isle Royale and

Voyageurs National Parks, and Pictured Rocks and Sleeping Bear Dunes National Lakeshores, which are the four parks in the Great Lakes I&M Network that contain a significant number of inland lakes.

Both discrete and continuous monitoring of lakes will take place at the deepest location of a given lake unless a question specific to the littoral zone is being addressed. It has been shown that spatial variation within a basin is small relative to temporal variation (Hanna and Peters 1991, Marshall and Peters 1989) for highly variable parameters such as chlorophyll and phosphorus. Therefore, for discrete monitoring we will maximize the number of sampling visits per lake, within budget constraints, and sample at a single location (the deepest part of the basin). For continuous monitoring, autonomous temperature sensors record water temperature from a lake's surface to the bottom at discrete intervals, once per hour, year-round, with site visits typically coincident with discrete monitoring visits at the beginning and end of a field season.

9.2 Initial QA/QC Checks for Outliers in Data

All water quality data undergo several quality assurance/quality control (QA/QC) procedures (e.g., duplicate sampling, data-entry filters, removal of logically inconsistent entries; see SOP #8 for data management details) to ensure that the data accurately represent the natural environment at the time of sampling. In addition to these procedures, several analytical and data sorting techniques (e.g., scatter plots, box-and-whisker plots, stem-and-leaf plots, sorting values of a given indicator in ascending or descending order, "COUNTIF" statements in Excel) are available to identify potentially erroneous values and statistical outliers. During this process, data points that do not meet limitations for precision or bias may be flagged or eliminated from the database. Some statistical measures, such as population mean, are sensitive to extreme atypical values, or outliers. Therefore, to reveal the central tendency of the population, the project manager may elect to remove outliers from the data pool. Such removals should be performed with great caution, and only when it is clear that the outlier truly did not reflect system properties (e.g., the outlier resulted from instrument error, transcriptional error, contamination of the sample).

9.3 Annual Data Summaries

Brief characterizations of the data from each lake, each park, and the network as a whole will be performed each year, after all QA/QC procedures have been completed. For each lake sampled, and for each parameter measured, descriptive statistics may include mean, median, maximum and minimum values, as well as skew, kurtosis, and measures of variability, when appropriate (e.g., coefficient of variation, standard error, variance), and provided as hard copy or web based summary reports, with the primary audience being the parks

Given the relevant legislation (e.g., Clean Water Act of 1972), it may be of interest to individual NPS units and to other entities to assess the proportion of measurements during a time period or across a domain (at a single point in time) that exceed pre-determined thresholds, such as State water quality

standards or ecoregional nutrient criteria. As with nearly all percentage data, arcsine transformations must be performed on those data before statistical analyses can be conducted (Sokal and Rohlf 1995).

9.4 Analyses of Long-Term Trends

In addition to these descriptive statistics, analytical approaches may also include estimation of interannual change, graphic approaches (e.g., comparison of mean and variability of a parameter in the current year versus during past years), and occasionally qualitative analysis (Guthery et al. 2001), as well as modeling, correlational analyses, and various parametric and nonparametric analyses. Results of such analyses will be distributed via synthesis reports and/or articles in peer-reviewed journals.

Because lakes were not selected randomly, we will not make inferences about trends in lakes other than those we sample. We will analyze data from each lake independently of all other lakes. All lakes sampled within a park may be grouped for analytical comparison with lakes sampled in another park. Whenever lakes are pooled for analysis, we will ensure inferences are not made beyond those lakes that were sampled. Characteristics of the lakes that are known or suspected to affect water quality (e.g., lake size, maximum depth, underlying geology, watershed characteristics) can be included as covariates.

9.4.1 Recommended Methods for Long-Term Trend Analysis

Synthesis reports will include more intensive analyses of change after at least three years of sampling in a given lake has occurred. In addition to repeated-measures, time-series, regression, and non-parametric equivalents of various methods (such as regression, paired-t tests, and ANOVA), monitoring data may also be evaluated through nonparametric trend tests (e.g., Mann-Kendall or Seasonal Kendall), Monte Carlo simulation analyses, Bayesian analyses, and comparisons of period means. For the latter-most approach, one is often interested in comparing values before and after an important event (e.g., change in management policy, remarkable anthropogenic disturbance, natural catastrophe, drought), and considers years within each of the two periods as replicates. We may also examine trends for breakpoints, or changes in slope, which may indicate the timing of an important event, and hence a potential cause. Trends in parameters that are analyzed with respect to biotic and abiotic covariates will be included in the synthesis reports, although cause-effect relationships may be investigated more thoroughly by NPS partners and collaborators.

In addition to analyzing each variable separately, water quality variables could be analyzed collectively through multivariate ordinations (e.g., nonmetric multi-dimensional scaling) of resource conditions through time, following West and Yorks (2002). This approach effectively integrates information across many indicators, and can suggest whether individual stations are all moving in the same direction in multidimensional ordination space. Furthermore, joint plots can be overlaid on the ordination, and can suggest which variables correlate most strongly to the direction of changes. Similarly, if specific comparisons are desired across a suite of uncorrelated variables, two or more pre-defined groups of samples could be compared using MANOVA or NPMANOVA, depending on whether parametric assumptions are met.

9.4.2 Approaches to Analyze Frequently Collected Data

In contrast to many other parameters that the network is monitoring, discrete water quality measurements are collected several times per year. Not only does the long-term nature of this data set allow for robust retrospective analyses of trend, but the multiple sampling sessions within each year allow for various analytical approaches to analyze long-term trends. For example, if the goal is to monitor trend in the *average* value of a particular water-quality parameter, then all measurements collected within the year (or during the ice-free season) at each sampling location would be considered temporal subsamples and averaged. Trend analysis (using repeated-measures, time-series, or other regression analyses) would thus be performed simply on the annual means at each sampling location. This approach seems particularly viable for parameters that do not exhibit strong intra-annual variability (e.g., pH). However, this approach may be vulnerable to imprecision or bias if a parameter exhibits significant, predictable variability in its value throughout the year and data are not collected (due to equipment failure or logistical constraint) at a high number of intended sampling occasions.

Alternatively, if a particular parameter is known to exhibit significant intra-annual variability (e.g., chlorophyll-*a*), samples within the year can be partitioned into one of several periods (ideally, defined by relevant phenological or biological phenomena). Thus, for example, if interannual trends in nutrient concentrations during algal blooms are of interest, the temporal domain can be accordingly defined, and all samples within that window averaged within each year and the means analyzed for trend across years. Finally, if there is a strong desire to incorporate intra-annual variability into interannual trend analyses, data within each year can be analyzed through a smoothing algorithm, and interannual variability is thus analyzed on the smoothed data. However, it may be the case that process variation is larger than the sampling variability (Burnham et al. 1987).

9.4.3 Duplicate Sampling

To ensure that understanding of water quality trends within lakes of the GLKN is not confounded by biased results, for discrete monitoring we will collect duplicate samples and field measurements at the rate of approximately once every 10 samples. We will assess the relative percent difference between duplicate samples or measurements and flag data that do not meet the QA/QC guidelines detailed in SOP #12. Those data that do not meet the QA/QC stipulations may or may not be used in analyses, on a case-by-case basis. For example, parameters measured at low concentrations (e.g., less than five times the method detection limit) may be accepted, and both duplicates used in analysis. If duplicate samples meet the QA/QC guidelines, we will use the mean value of the duplicate measurements in data analyses. Explanations of how duplicates are handled in data analyses will be included in reports.

9.4.4 Other Analytical Considerations

For trend analyses performed using regression, trend will be investigated using a linear relationship. If analyses suggest a non-linear temporal pattern, serial autocorrelation, or lagged response to stressors, appropriate analytical modifications will occur. When discussing the desire to be able to detect a trend of 20%, for example, it must be stated *during what time period* that level of change is to be detected. This allows one to calculate the minimum level of change that is important to detect

between successive sampling periods. A 20% change occurring over 1 year is obviously much different than that same level of change (20%) occurring over 10 years (i.e., an average change of 1.84% per year).

Adopting the philosophies of the precautionary principle (United Nations 1992) and the safe minimum standard of conservation (Ciriacy-Wantrup 1952, Berrens et al. 1998), the GLKN monitoring program seeks to identify potential natural resource problems early, before deleterious or irreversible ecosystem changes occur (e.g., crossing of ecological thresholds; Laycock 1991, van de Koppel et al. 1997, Laurance and Williamson 2001). Consequently, for most analyses that use the null hypothesis of no change through time, we will adopt $\alpha = 0.10$. Furthermore, for parameters that exhibit particularly high variability, magnitudes of change that seem biologically meaningful (Johnson 1999) yet have $0.10 < \alpha < 0.20$ may merit more intensive or extensive monitoring or experimental study.

If water quality monitoring tracks a relatively large suite of variables, Bonferroni or modified Bonferroni corrections (e.g., Hohm's method) may be performed, to maintain the 'family-wise' alpha at 0.10. Roback and Askins (2005) argue that if the main goal is simply initial screening for conservation problems or generation of hypotheses to be tested further, but not detection of real differences or trends, then alpha can be left at its standard comparison-wise significance level, to avoid Type II errors.

When the dataset contains non-detects, we will follow Helsel (2005) in conducting statistical analyses. Several methods exist for handling non-detects, each with advantages and drawbacks. Early in our program we will likely use a substitution method, where a value between zero and the detection limit will be substituted for the non-detect. The main drawback of this method is that estimates of the true variability are not possible. When the program is more mature and we have adequate data to determine the distribution of data, we will likely use maximum likelihood estimation, which works well for large sample sizes. Other methods, such as "regression on order statistics" and the non-parametric Kaplan-Meier method will be considered, and in consultation with a statistician, the most appropriate method for the data will be employed.

9.5 Literature Cited

- Anderson, D. R., K. P. Burnham, W. R. Gould, and S. Cherry. 2001. Concerns about finding effects that are actually spurious. *Wildlife Society Bulletin* 29:311–316.
- Berrens R. P., D. S. Brookshire, M. McKee, and C. Schmidt. 1998. Implementing the safe minimum standard approach: Two case studies from the U.S. Endangered Species Act. *Land Economics* 74:147–161.
- Burnham, K. P., D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock. 1987. Design and analysis methods for fish survival experiments based on release-recapture. *American Fisheries Society Monograph* 5:1–437.
- Ciriacy-Wantrup S. V. 1952. *Resource Conservation: Economics and Policies*. University of California Press, Berkeley, California.
- Elias, J. E, R. Axler, and E. Ruzycki. 2008. Water quality monitoring protocol narrative for inland lakes, version 1.0. National Park Service, Great Lakes Inventory and Monitoring Network. Natural Resource Technical Report NPS/MWR/GLKN/NRTR—2008/109. National Park Service, Fort Collins, Colorado.
- Gucciardo, S., B. Route, and J. Elias (editors). 2004. Conceptual ecosystem models for long-term monitoring in the Great Lakes Network. Great Lakes Network Technical Report GLKN/2004/04. National Park Service, Ashland, Wisconsin.
- Guthery, F. S., J. J. Lusk, and M. J. Peterson. 2001. The fall of the null hypothesis: Liabilities and opportunities. *Journal of Wildlife Management* 65:379–384.
- Hanna, M., and R. H. Peters. 1991. Effect of sampling protocol on estimates of phosphorus and chlorophyll concentrations in lakes of low to moderate trophic status. *Canadian Journal of Fisheries and Aquatic Science* 48:1979–1986.
- Helsel, D. R. 2005. *Nondetects and Data Analysis: Statistics for Censored Environmental Data*. John Wiley and Sons, Inc., New York.
- Johnson, D. H. 1999. The insignificance of significance testing. *Journal of Wildlife Management* 63:763–772.
- Laurance, W. F., and G. B. Williamson. 2001. Positive feedbacks among forest fragmentation, drought, and climate change in the Amazon. *Conservation Biology* 15:1529–1535.
- Laycock, W. A. 1991. Stable states and thresholds of range condition on North American rangelands: A viewpoint. *Journal of Range Management* 44:427–433.
- Marshall, C. T., and R. H. Peters. 1989. General patterns in the seasonal development of chlorophyll-*a* for temperate lakes. *Limnology and Oceanography* 34:856–867.

- Noon, B. R. 2003. Conceptual issues in monitoring ecological resources. Pages 27–72 *in* D. E. Busch, and J.C. Trexler, editors. *Monitoring Ecosystems: Interdisciplinary Approaches for Evaluating Ecoregional Initiatives*. Island Press, Covelo, California.
- Roback, P. L., and R. A. Askins. 2005. Judicious use of multiple hypothesis tests. *Conservation Biology* 19:261–267.
- Sokal, R. R., and F. J. Rohlf. 1995. *Biometry: The Principles and Practice of Statistics in Biological Research*, third edition. W. H. Freeman and Company, New York.
- United Nations. 1992. Agenda 21. Chapter 17 *in* Report of the United Nations conference on environment and development. United Nations, New York.
- van de Koppel, J., M. Reiterker, and F. J. Weissing. 1997. Catastrophic vegetation shifts and soil degradation in terrestrial grazing systems. *Trends in Ecology and Evolution* 12:352–356.
- West, N. E., and T. P. Yorks. 2002. Vegetation responses following wildfire on grazed and ungrazed sagebrush semi-desert. *Journal of Range Management* 55:171–181.
- Yoccoz, N. G. 1991. Use, overuse, and misuse of significance tests in evolutionary biology and ecology. *Bulletin of the Ecological Society of America* 72:106–111.