

Monitoring Design

Anthony R. Olsen
USEPA NHEERL Western Ecology Division
Corvallis, OR 97333

And

Dale M. Robertson
U.S. Geological Survey
Middleton, WI 53562

The monitoring design component of the overall Monitoring Framework provides answers to the questions: What site or environment will be monitored? What parameters will be measured at the site? Where will the measurements be taken? What methods will be used to collect the data? When and how frequently will the measurements be made? Since answers to these questions encompass most of the activities people think about in a monitoring program, some consider monitoring design as the entire process of implementing a monitoring program, i.e., the entire Monitoring Framework. Here, we take a more limited view for monitoring design activities. A natural tension exists between monitoring design, the monitoring objectives, scientific capabilities, and institutional capabilities (e.g., budgets, personnel). Monitoring objectives guide the development of the monitoring design. At the same time, the development of the monitoring design almost always requires clarification and prioritization. The available budget or personnel capabilities may limit available options for a design to meet the objectives and, therefore, might require elimination of one or more of the monitoring objectives. In addition, the monitoring design must be sufficient to enable the planned data analysis component of the study.

What will be monitored?

Generally, the monitoring objectives specify what site or environment will be monitored. The objectives may identify a specific waterbody, or location, to be monitored. Examples are the mixing zone associated with a point source discharge, drinking water intakes for a community water system, a single lake or reservoir, a specific river reach, ground water associated with a potentially contaminated site, or a wetland near a new development. Other monitoring programs may focus on determining the status and trends for a general class of aquatic resource for a region, state, or the entire nation. General classes may include streams and rivers, lakes and reservoirs, estuaries and coastal waters, recreational beach waters, ground water, drinking water sources, and drinking water intakes. Typically, a monitoring program focuses on only one of the general classes and may further restrict that class to some subset of interest. For example, a state may desire a monitoring program for a list of special type of lake, or all perennial, wadeable streams, rather than all lakes and all streams and rivers within a state. A community drinking water organization would restrict monitoring to all sources that

contribute to their drinking water system. A national survey of lakes may define a lake to be any waterbody greater than 1 hectare in area, a depth greater than 1 meter, at least 50% open water, and include natural as well as man-made ponds and reservoirs.

The development of a monitoring design typically requires that the initial monitoring objective be refined to clarify exactly what will be monitored. The refined definition must be sufficiently clear to enable any person in the monitoring program to determine if a specific waterbody is included or excluded from the study. Is the boundary of the mixing zone clearly defined? What lakes are included in a national lake monitoring program? Sometimes this is more difficult than we expect. The statistical survey design literature uses the phrase “target population” to identify explicitly what is to be included in the study. As an example, Alley (1993) discusses the target populations in designing regional ground water quality monitoring programs.

What parameters will be measured?

The monitoring objectives are the basis for determining what parameters will be measured. Many times this is straightforward. For example, the objectives may specify that a specific chemical, e.g., dioxin, is to be measured within the mixing zone of a point source discharge. A monitoring objective for a beach may specify that specific bacteria, such as *Escherichia* (E-coli) or enterococcus bacteria, must be monitored. Or a lake will be monitored to determine its trophic state. In this case, several alternative methods are available to determine the trophic state of a lake. Trophic classifications based on total phosphorus, continue to be most common; however, the monitoring objectives would need to be clarified before selecting specifically what needs to be measured to determine the trophic state.

Section 305(b) of the Clean Water Act requires biennial reports to Congress on the status and trends of navigable waters that meet water quality standards. The water quality standards in a state enable a specific set of parameters to be identified as necessary to be measured. In other situations, the objectives may only say that the condition of the aquatic resource be determined. Aquatic resources are complex ecosystems and can be viewed from many alternative perspectives. Assessing condition can depend on measurement of chemical contaminants, water temperature, in-stream physical habitat, riparian habitat, sediment contamination, benthic macro-invertebrate community, fish community, periphyton community, as well as other factors. Certainly, agreement will be needed on which measurements will constitute the monitoring program’s concept of “the condition of the aquatic resource.”

The discussion so far has emphasized what is to be measured and did not take the next step defining how the information is to be quantitatively reported. In some cases, this is straightforward, as when pH will be reported in pH units or total phosphorus as $\mu\text{g/l}$. Nitrate in water may be reported as $\text{NO}_3\text{-N}$ (nitrate as nitrogen) in mg/l . Fish community data are initially reported as species abundance counts. To relate these data to monitoring objectives requires specifying one or more quantitative metrics or indices be derived. One such index is the fish index of biotic integrity (IBI), developed by Karr (1981), although even this may not be sufficient to define how the IBI is constructed.

Simply defining what parameters are to be measured is not sufficient for developing a monitoring program to be used for environmental decision making because most parameters can be monitored using many different techniques. Therefore, we must also know how the data will be used or in other words know the objectives for the data. Data quality objectives (DQO's) define how accurate we must be when collecting the data and therefore, what methods can be used to collect the data in the field, and what laboratory procedures may be required. These topics are covered in the Collect Field and Lab Data cog.

When and how frequently will the measurements be taken?

Knowing what to measure at a site and how to measure it are essential in designing a monitoring program. Another important question is "When and how frequently should and will the measurements be taken?" There are many dimensions that one must consider when answering this question. When to sample is often determined by the objectives of the monitoring program. For example, one objective of a program may be to define phosphorus concentration during fall turnover in a lake or during summer base flow in a stream. The frequency of data collection should depend on the specific question to be answered and several factors specific to the parameter being monitored, such as the expected variability of the parameter, response time of the parameter and the system, how the parameter fluctuates with season and flow (if we are monitoring streams), variability and timing of inputs of the parameter, and the magnitude of response one expects or needs to detect in the parameter (USEPA, 1997).

States are required by USEPA to report on the waters of their state that meet designated uses, such as supporting aquatic life use. A state standard may specify that pollutant concentrations may not exceed a limit during an entire monitoring period or more than 10 percent of the time. As stated, the frequency of monitoring is not easily determined. In practice, states typically monitor monthly or quarterly for pollutants. When biological measurements are used, it is common to monitor only once during a year. If only a single measurement is made during a year, when should it be measured? One approach is to define an index period, for example, during summer lowflow in streams or after fall overturn in lakes. The index period is not intended to represent annual average conditions but instead result in a measurement that provides an index on the quality of the aquatic resource.

Various temporal strategies have been used to collect samples to describe changes in water quality. These strategies range from basic fixed-period sampling strategies, such as monthly, to extensive automated sampling techniques. The U.S. Geological Survey's national stream quality accounting network (NASQAN) used basic monthly sampling of streams across the nation for more than 20 years to monitor the water quality at a fixed set of locations (Smith et. al., 1987). In some cases, even less frequent sampling such as quarterly or annual sampling has been used by many state agencies to describe annual conditions. The U.S. Geological Survey's National Water-Quality Assessment

(NAWQA) Program (Hirsch et al., 1988) typically collects fixed-period monthly samples supplemented by a few manually collected high-flow samples each year (Gilliom et al., 1995). Although sampling strategies vary, monitoring programs often have the same common goals, that being to estimate mean, median, and maximum concentrations and, if sampling streams, loads of specific water-quality parameters during a specified time period. With intense sampling, estimating summary concentrations and loads are straight forward and accurate. However, with less intense sampling, the summary concentrations may be less accurate and different assumptions and approaches must be used that may also reduce the accuracy in these estimates.

The approach suggested by the USEPA (1997) to determine the sampling frequency required to measure a mean concentration to within a specified range relies on the assumption that there are no major trends, cycles, or patterns in the data. The number of samples (n) needed to estimate the mean concentration of a constituent within a range E , with a confidence of α is estimated with Eq. 1 once the variability (s^2) in concentrations is known. The t statistic requires that n is known; therefore, the $Z_{1-\alpha/2}$ statistic is used in Eq. 1 instead of the t statistic to obtain an initial estimate of n .

$$n = (t_{1-\alpha/2, n-1} \times s / E)^2 \quad (1)$$

For example, the number of samples to determine a mean phosphorus concentration within a range ± 0.1 mg/L with confidence of 95 % ($P < 0.05$, $t_{0.025}$ is approximately 2.1), with standard deviation of $s = 0.2$ mg/L would be about 18 samples. This approach, although commonly used, has been found to underestimate the number of samples actually needed to obtain a defined confidence level (Kupper and Hafner, 1989). Zar (1999) suggests a more appropriate approach is to decide the number of samples needed to determine if a concentration exceeds a given value by as little as δ given a specific level of significance (α), and a probability of detecting the difference (B) using Eq. 2. An initial value of n can again be determined using the Z statistic.

$$n = s^2 / \delta^2 (t_{\alpha, n-1} + t_{1-B, n-1})^2 \quad (2)$$

In the same example above, the number of samples required to determine with a 90 % probability at the 0.05 level of significance that a mean total phosphorus concentration exceeded a criterion by $+ 0.1$ mg/L would be about 47 samples. The above formulae assume that the data follow a normal distribution, unless the sample size is large (30 to 50 samples). For skewed data, a prior transformation such as the logarithm may be necessary to approximate a normal distribution.

These approaches rely on the assumption that water quality varies randomly throughout the year and throughout specific flow regimes; however, water quality in streams does not follow this behavior. A few studies have been conducted to determine how various temporal sampling strategies affect the estimated summary concentrations and loads in streams. Results from these studies can be used to define the appropriate frequency for a monitoring program. This type of analysis often requires oversampling at the onset of study to determine how various sampling designs would affect the summary statistics.

Walling and Webb (1981) used hourly suspended sediment data to compare results based on different fixed period sampling strategies. Robertson (2003) used data from several extensively sampled streams to examine how various temporal sampling strategies typically used by agencies affect estimated mean, median, and maximum concentrations, and annual loads in small streams.

The frequency of monitoring of some studies is based on evaluating compliance with an issued permit. Some facilities are required to obtain a state point source discharge permit. The permit may include requirements on the frequency of monitoring. For example, the permit may state that the concentration of total suspended solids shall not exceed a monthly average of 30 mg/L, a weekly average of 45 mg/L, or a daily maximum of 90 mg/L. To compute these summary statistics, the specified monitoring frequency is usually daily to satisfy the requirements. In some situations, a point source discharge permittee may be required to measure the condition of the receiving waters and sediments within a mixing zone. The frequency in most of these cases depend on the permit requirements. Beach monitoring programs can require daily bacteria measurements when the beach is open to the public.

So far frequency of measurement has focused on the number of measurements within a year. Frequency across years must also be considered. The frequency of sampling in a given year determines not only how accurate a mean annual concentration can be estimated, but it also determines how long it takes to detect a change in water quality. Parameters with high variability require more frequent sampling to describe mean annual concentrations and more years of sampling to detect changes in concentration. A typical long-term monitoring program continues taking measurements at the same frequency each year; therefore, it is important to consider the variability in water quality and how long it will take to detect a change in water quality of a specified magnitude prior to defining the sampling frequency of a monitoring program. Urquhart et. al. (1993, 1998) illustrate the impact of alternative designs in detecting trends. Financial constraints often limit the number of samples that can be collected in a given year and therefore often extend the number of years of data required to detect a change in concentration. An alternative is to conduct monitoring every several years, e.g., in sediments in an estuary where the expectation is that concentration changes will occur slowly. Data from the two different periods are then statistically compared or when sufficient data become available annual summaries from specified periods are compared. When collecting data every several years, it is important to consider possible factors that are periodic and may bias data collected every several years, such as periodic El Nino events.

Where will measurements be taken?

Site selection is a critical part of the monitoring design. Seldom is it possible to measure at all locations in a study area, i.e., at all elements/locations in the target population. What options are available for selecting sites and what governs which option should be used? The latter depends on a quantitative statement of the monitoring objectives.

In general, a goal in site selection is to obtain a “representative” sample. If the sample is not representative, then the information produced will not address the monitoring objectives. The difficulty is that representative means different things to different groups. In a review of non-scientific, scientific, and statistical literature, Kruskal and Mosteller (1979a,b,c) found seven alternative definitions for representative that appear to be in common use. Consequently, it is necessary to be very specific about what is meant by a representative sample in a monitoring program before an assessment can be made whether the site selection process is consistent with the objectives.

Sites are often selected to represent large geographic areas. Various approaches have been used to classify large areas into smaller regions of similar water quality. These approaches can be subdivided into geographically dependent and geographically independent approaches. In geographically dependent classification schemes, broad areas or regions are defined that reflect the geographic distribution of various explanatory characteristics. Bailey (1976; 1996), Omernik (1987), and Albert (1995) have used combinations of several environmental characteristics thought to influence ecosystem functioning (soil type, climate, natural vegetation, land use, land form, etc.) to subdivide large geographical areas into ecoregions. Geographically independent frameworks are usually determined by watershed attributes that can be defined independently of a geographic region. This approach was used to subdivide the United States into hydrologic landscape regions on the basis of land-surface form, geology, and climate (D.M. Wolock, U.S. Geological Survey, written commun. 2001) and used by the Army Corps of Engineers to develop a hydrogeomorphic classification scheme for wetlands (Brinson 1993). Which ever method is used, it is important that the regionalization scheme be based on the distribution of the most strongly related environmental factors. Robertson and Saad (2003) used SPATial Regression-Tree Analysis (SPARTA) to determine which environmental characteristics were most statistically important in describing the distribution of specific water-quality constituents and to delineate areas with relatively similar potential water quality.

One approach to site selection is to rely on professional judgment to determine where to sample. For example, it may be desirable to judgmentally locate a site in the receiving waters at the end of the pipe with the expectation that the maximum concentration will be located there. If an estimate of the average pollutant concentration within the mixing zone is required by the objectives, then that site may not be usable in computing that average. It would be useful in determining if concentrations within the mixing zone are likely to be above a detection limit. A probability survey design, such as a simple random sample or systematic sample, would be more appropriate to determine the average within the mixing zone. The chosen regionalization scheme may be the basis for defining strata in a probability survey design approach to site selection or for professional judgment site selection within regions.

A version of professional judgment is a gradient study that investigates the relationship between a response indicator and one or more stressor indicators. For example, a study may investigate the influence of agriculture and urban development on pesticide

concentrations in streams in headwater areas. Selecting watersheds to insure a gradient of percent watershed area developed would be essential for this objective. Particular watersheds could be selected using existing knowledge on their environmental characteristics such that those selected are typical watersheds. Alternatively, all watersheds may be placed in five categories based on percent agricultural use within each watershed. A stratified random sampling design could then be used to select watersheds across the gradient. The U.S. Geological Survey's National Water Quality Assessment Program (NAWQA) incorporates local knowledge about streams in selecting sites to meet an objective of understanding relationships between stressors and water quality (Gilliom et. al. 1995; Helsel 1995).

When selecting sites using existing knowledge it is important to understand the most important environmental factors that affect the variability in specific water-quality characteristics. With this information, one can determine an appropriate stratification or regionalization scheme, such as ecoregions or areas with similar soil types. It is also important to determine how well the data collected at the selected sites represent the water quality of the larger areas they were chosen to represent. Robertson (1998) evaluated the surface-water sampling design used by one study unit of the NAWQA program by sampling several sites in each of the areas individual sites were selected to represent and comparing the results from the various sites.

Reference conditions of specific environmental characteristics may be required to aid in the interpretation of data collected within a study region. A regionalization approach may first be applied to define sub-regions of similar expected water quality. Reference conditions at sites that are not subject to disturbance, minimally impacted sites, or pre-European settlement conditions are determined within each sub-region. In other cases, conditions that are typical for sites that represent a certain type of impact, such as water quality in areas with agriculture on clay soils may be of interest

The monitoring objectives may require overall quantitative estimates, such as the percent of all streams and rivers within a state that meet their designated uses or an average pollutant concentration in sediments within an estuary. Incorporating some form of "random selection" to locate sites is necessary to obtain a representative sample. For example, section 305(b) of the Clean Water Act requires states to estimate the stream length (or number of lakes or estuarine area) in their state that are impaired. A probability sample, or survey design, is the only site selection approach available that will produce quantitative estimates with an accompanying statement of precision (e.g., 95% confidence intervals). Simple random sample, stratified random sample, and systematic sample are common survey designs that have been used in monitoring programs. A simple random sample of all lakes in a state can be achieved by creating a list of all the lakes and using a random sampling algorithm to select a sample of lakes. The monitoring objectives may specify that estimates must be made for four subregions of a state or for a set of special interest lakes as well as all other lakes. The list of lakes is then separated into multiple lists (i.e., strata) and a simple random sample selected within each stratum to obtain a stratified random sample. For example, the U.S. EPA's Environmental Monitoring and Assessment Program (EMAP) in a study of lakes in the northeastern

United States selected lakes using a probability survey design that stratified by lake area (Larsen et. al. 1994).

States are increasing their use of probability survey designs as integral components of their monitoring programs, particularly for streams and rivers. The Consolidated Assessment and Listing Methodology prepared by the US EPA Office of Water (<http://www.epa.gov/owow/monitoring/calm.html>) discusses the use of probability survey designs for monitoring the quality of water in states. Further information on probability sampling applied to aquatic resources is available at <http://www.epa.gov/nheerl/arm/>.

Aquatic resources are integral parts of the spatial environment in which they exist. As such, geographic information systems (GIS) should be utilized in the monitoring design process, particularly site selection. GIS coverages, such as contained in the National Hydrologic Database (NHD) (<http://nhd.usgs.gov/>), can be used as a representation of the target population to be sampled. Computer programs are available that use NHD coverages, monitoring design specifications, and complete site selection using a probability survey design.

Implementing a monitoring design

Knowing what will be monitored, what will be measured, when and how frequently it will be measured, and where it will be measured, are all essential elements of a monitoring design. However, knowing this information may not be adequate for others to implement the monitoring program. Documenting the design, including rationale for decisions, is also critical when the design is implemented and when the data that are collected are used by others. The Develop Monitoring Objective cog and the Design Monitoring Program cog information must be consistent and detailed. When documenting the objectives, a two-tier approach is likely to be useful. The first tier gives the objectives as they might appear in an overview presentation of the monitoring program, being as specific as possible and being constrained by length. The second tier gives very specific, detailed, quantitative restatements of the objectives that are necessary for the monitoring design process.

Once the monitoring design is used to select sites, a plan must be developed to visit the sites and collect the agreed upon measurements. A typical situation is that all that is known about a selected site is its geographic location and environmental characteristics. To access the site and enable sampling to be conducted may require landownership be determined and permission granted before a visit to the site. Some states have a legal requirement for landowners to give written permission before access is granted. A well thought out protocol for how to contact landowners, what information to provide them, and how to follow-up with landowners can significantly increase the likelihood of a landowner granting access. A logistical plan may be necessary to increase the efficiency of visiting sites.

Before field work begins, some type of quality assurance project plan (QAPP) should be developed. The content of the plan will reflect all components of the Monitoring Framework. Details on field crew training, field measurement protocols, shipping and sample handling, laboratory analysis protocols, and data management are essential topics to be covered in the plans. Most important is that the quality assurance plan become an active, living document that is a routine part of the monitoring program. Many projects funded by the USEPA require and approved QAPP prior to starting the project. Guidelines for such plans are available from US EPA's Quality System web site (<http://www.epa.gov/quality/index.html>). The content of the plans will reflect all components of the Monitoring Framework.

Summary

The monitoring design cog can not be considered in isolation of other cogs in the monitoring framework. For example, the objectives of a monitoring program govern the monitoring design and at the same time the monitoring design forces clarification of the monitoring objectives. Assessing and interpreting the resulting data must be matched to the monitoring design. Designing a monitoring program is a team and an iterative process that involves all the cogs of the monitoring framework.

References

Albert, D.A. 1995. Regional landscape ecoregions of Michigan, Minnesota, and Wisconsin: a working map and classification (fourth revision: July 1994). General Technical Report NC-178. U.S. Department of Agriculture, Forest Service, St. Paul, MN. 250 pp.

Alley, William M. 1993. Regional Ground-Water Quality. Van Nostrand Reinhold, New York, 634 pp.

Bailey, R.G. 1976. Ecoregions of the United States Map (scale 1:7,500,000). Ogden Utah: U.S. Department of Agriculture. Forest Service. Intermountain Region.

Bailey, R.G. 1996. Ecosystem geography. Springer-Verlag, New York, 204 pp.

Cochran, W. G., 1977. Sampling Techniques, 3rd Edition. New York: John Wiley and Sons.

Brinson, M.M., 1993. A hydrogeomorphic classification for wetlands. U.S. Army Corps Engineer Waterways Experiment Station Technical Report WRP-DE-4.

Gilbert, R. O. (1987). Statistical Methods for Environmental Pollution. New York, Van Nostrand Reinhold.

Gilliom, R. J., W. M. Alley, et al. (1995). Design of the National Water-Quality Assessment Program: Occurrence and Distribution of Water-Quality Conditions. Denver, CO, U.S. Geological Survey: 33.

- Helsel, D. R. (1995). Design of a relational water-quality assessment program. 1995 Proceedings of the Biometrics Section. Washington, DC, American Statistical Association: 60-67.
- Karr, J. R. (1981). "Assessment of biotic integrity using fish communities." Fisheries **6**: 21-26.
- Kruskal, W. and F. Mosteller (1979). "Representative sampling, I: Non-scientific literature." International Statistical Review **47**: 13-24.
- Kruskal, W. and F. Mosteller (1979). "Representative sampling, II: Scientific literature, excluding statistics." International Statistical Review **47**: 111-127.
- Kruskal, W. and F. Mosteller (1979). "Representative sampling, III: The current statistical literature." International Statistical Review **47**: 245-265.
- Kupper, L.L., and K.B., Hafner 1989. "How appropriate are popular sample size formulas?" The American Statistician **43**: 101-105.
- Kwiatkowski, R. E. (1991). "Statistical needs in national water quality monitoring programs." Environmental Monitoring and Assessment **17**(2-3): 253-271.
- Larsen, D. P., K. W. Thornton, et al. (1994). "The role of sample surveys for monitoring the condition of the Nation's lakes." Environmental Monitoring and Assessment **32**(2): 101-134.
- MacDonald, L. H. and A. Smart (1993). "Beyond the guidelines: practical lessons for monitoring." Environmental Monitoring and Assessment **26**(2-3): 203-218.
- Maher, W. A., P. W. Cullen, et al. (1994). "Framework for designing sampling programs." Environmental Monitoring and Assessment **30**(2): 139-162.
- Olsen, A. R., J. Sedransk, et al. (1999). "Statistical issues for monitoring ecological and natural resources in the United States." Environmental Monitoring and Assessment **54**(1): 1-45.
- Omernik, J.M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers **77**:118-125.
- Overton, W. S. and S. V. Stehman (1995). "Design implications of anticipated data uses for comprehensive environmental monitoring programs." Environmental and Ecological Statistics **2**(4): 287-303.

Overton, W. S. and S. V. Stehman (1996). "Desireable design characteristics for long-term monitoring of ecological variables." Environmental and Ecological Statistics **3**(4): 349-361.

Peterson, S. A., N. S. Urquhart, et al. (1999). "Sample representativeness: a must for reliable regional lake condition estimates." Environmental Science & Technology **33**(10): 1559-1565.

Robertson, D.M., 1998, Evaluation of the surface-water sampling design in the Western Lake Michigan Drainages in relation to environmental factors affecting water quality, U.S. Geological Survey Water Resources Investigations Report 98-4072, 53 p.

Robertson, D.M., and D.A. Saad, 2003, "Environmental water-quality zones for streams: a regional classification scheme." Environmental Management 31(5): 581-602.

Robertson, D. M., 2003 (in press), Influence of different temporal sampling strategies on estimating water quality in small streams. Journal of the American Water Resources Association.

Sanders, T. G., R. C. Ward, et al. (1983). Design of Networks for Monitoring Water Quality. Highlands Ranch, Colorado, Water Resource Publication.

Smith, R. A., R. B. Alexander, et al. (1987). "Water-quality trends in the nation's rivers." Science 235(27 March): 1607-1615.

Stout, B. B. (1993). "The good, the bad, and the ugly of monitoring programs: defining questions and establishing objectives." Environmental Monitoring and Assessment 26: 91-98.

Urquhart, N. S., W. S. Overton, et al. (1993). Comparing sampling designs for monitoring ecological status and trends: impact of temporal patterns. Statistics for the Environment. V. Barnett and K. F. Turkman. New York, John Wiley & Sons: 71-86.

Urquhart, N. S., S. G. Paulsen, et al. (1998). "Monitoring for policy-relevant regional trends over time." Ecological Applications 8(2): 246-257.

U.S. Environmental Protection Agency, 1997. Monitoring guidance for determining the effectiveness of nonpoint source controls EPA/841-B-96-004. U.S. EPA Office of Water Nonpoint Source Control Branch. Washington D.C.

USEPA (U.S. Environmental Protection Agency). 1998. National strategy for the development of regional nutrient criteria. Office of Water, ERP-822-R-98-002.

Walling, D. E., and B. W. Webb, 1981. The reliability of suspended sediment load data. In *Erosion and Sediment Transport Measurement. Proc. Florence Symp.*, June 1981, IAHS Publ. no. 133, pp. 177-194.

Zar, J.H., 1999. *Biostatistical Analysis*, 4th Edition, New Jersey, Prentice Hall.