United States Environmental Protection Agency
Office of Water
Office of Science and Technology
Washington, DC 20460
EPA-822-R-07-004
September 2007
www.epa.gov

Nutrient Criteria Technical Guidance Manual

Wetlands

Full Guidance Document available in full at:
http://www.epa.gov/waterscience/criteria/nutrient/guidance/

Chapter 4 Sampling Design for Wetland Monitoring

4.1 INTRODUCTION

This chapter provides technical guidance on designing effective sampling programs for State wetland water quality monitoring programs. EPA recommends that States begin wetland monitoring programs to collect water quality and biological data in order to characterize the condition of existing wetlands as they develop nutrient criteria that will protect their wetlands. The best monitoring programs are designed to assess wetland conditions with statistical rigor while maximizing available resources.

At the broadest level, monitoring data should:

1. Detect and characterize the condition of existing wetlands.

2. Describe whether wetland conditions are improving, degrading, or staying the same.

3. Define seasonal patterns, impairments, and deviations in status of wetland conditions.

Water quality monitoring programs should collect a sufficient number of samples over time and space to identify changes in system condition or estimate average conditions with statistical rigor. Three approaches to study design for assessing water quality and biological and ecological condition, and identifying degradation in wetlands are described in this chapter. Specific issues to consider in designing monitoring programs for wetland systems are also discussed in this chapter. The study designs presented here can be tailored to fit the goals of specific monitoring programs.

The three approaches described below (Section 4.3) (probabilistic sampling, targeted/tiered, and Before/After-Control/Impact [BACI]), present study designs that allow one to obtain a significant amount of information with relatively minimal effort. Probabilistic sampling begins with a large-scale, random monitoring design that is reduced as the wetland system conditions are characterized. This approach is used to find the average condition of each wetland class in a specific region. Probabilistic sampling design is frequently used for new large-scale monitoring programs at the State and Federal level (e.g., Environmental Monitoring and Assessment Program (EMAP), Regional Environmental Monitoring and Assessment Program (REMAP), State programs [e.g., Maine, Montana, Wisconsin]). The tiered or targeted approach to monitoring begins with coarse screening and proceeds to more detailed monitoring protocols as impaired and high-risk systems are identified and targeted for further investigation. Targeted sampling design provides a triage approach to more thoroughly assess condition and diagnose stressors in wetland systems in need of restoration, protection, and intensive management. Several State pilot projects use this method or a modification of this method for wetland assessment (e.g., Florida, Ohio, Oregon, and Minnesota). The synoptic approach described in Kentula et.al., (1993) uses a modified targeted sampling design. The BACI design and its modifications are frequently used to assess the success of restoration efforts or other management experiments. BACI design allows for comparisons in similar systems over time to determine the rate of change in relation to the management activity, e.g., to assess the success of a wetland hydrologic restoration. The BACI design, in particular, is included to assist States in evaluating ongoing management actions, and may provide less statistical rigor if adopted as a general monitoring program design. This design, however, is of considerable value in assessing restoration success and has been included at the request of States with ongoing wetland restoration. Detenbeck et.al., (1996) used BACI design for monitoring water quality of wetlands in the Minneapolis/St. Paul, Minnesota metro area.

Monitoring programs should be designed to describe what the current conditions are and to answer under what conditions impairment may occur. A well-designed monitoring program can contribute to determining those conditions.

Sampling design is dependent on the management question being asked. Sampling efforts should be designed to collect information that will answer the management question. For example, probabilistic sampling might be good for ambient (synoptic) monitoring programs, BACI for evaluating management actions such as restoration, and targeted sampling/stratified and random sampling for developing index of biotic integrity (IBIs) or nutrient criteria thresholds. In practice, some State programs likely will need to use a combination of approaches.

4.2 CONSIDERATIONS FOR SAMPLING DESIGN

DESCRIBING THE MANAGEMENT QUESTION

Clearly defining the question being asked (identifying the hypothesis) encourages the use of appropriate statistical analyses, reduces the occurrence of Type I (false positive) errors, and increases the efficient use of management resources (Suter 1993; Leibowitz et al., 1992; Kentula et al., 1993). Beginning a study or monitoring program with carefully defined questions and objectives helps to identify the statistical analyses most appropriate for the study and reduces the chance that statistical assumptions will be violated. Management resources are optimized because resources are directed at monitoring that which is most likely to answer management questions. In addition, defining the specific hypotheses to be tested, carefully selecting reference sites, and identifying the most useful sampling interval can help reduce the uncertainty associated with the results of any sampling design and further conserve management resources (Kentula et al., 1993). Protecting or improving the quality of a wetland system often depends on the ability of the monitoring program to identify cause-response relationships, for example, the relationship of nutrient concentration (causal variable) to nutrient content of vegetation or vegetation biomass (response variable). Cause-response relationships can be identified using large sample sizes and systems that span the gradient (low to high) of wetland quality. All ranges of response should be observed along the causal gradient from minimally disturbed to high levels of human disturbance.

Monitoring efforts often are prioritized to best utilize limited resources. For example, the Oregon case study chose not to monitor depressional wetlands due to funding constraints. They further tested the degree of independence of selected sites (and thus the need to monitor all of those sites) using cluster analysis and other statistical tests (http://www.epa.gov/owow/wetlands/bawwg/case/or.html). Frequency of monitoring should be determined by the management question being asked and the intensity of monitoring necessary to collect enough information to answer the question. In addition, monitoring should identify the watershed level activities that are likely to result in ecological degradation of wetland systems (Suter et al., 1993).

SITE SELECTION

Site selection is one of many important tasks in developing a monitoring program (Kentula et.al., 1993). Site selection for a monitoring program is based on the need to sample a sufficiently large number of wetlands to establish the range of wetland quality in a specific regional setting. Wetland monitoring frequently includes an analysis of both watershed/landscape characteristics and wetland specific characteristics (Kentula et al., 1993; Leibowitz et al., 1992). Therefore, wetland sampling sites should be selected based on land use in the region so that watersheds range from minimally impaired with few expected stressors to high levels of development (e.g., agriculture, forestry, or urban) with multiple expected stressors (see the Land-Use Characterization for Nutrient and Sediment Risk Assessment, Wetland module #17). There is often a lag in time between the causal stress and the response in the wetland system. This time lag between stress and response and the duration of this lag depends on many factors, including the type of stressor, climate, and system hydrology; these factors should be considered when selecting sites to establish the range of wetland quality within a region.

LANDSCAPE CHARACTERIZATION

The synoptic approach described in Liebowitz et.al., (1992) provides a method of rapid assessment of wetlands at the regional and watershed levels that can help identify the range of wetland quality within a region. Liebowitz et.al., (1992) recommend an initial assessment for site selection based on current knowledge of watershed and landscape level features; modification of such an assessment can be made as more data are collected. Assessing watershed characteristics through aerial photography and the use of geographical information systems (GIS) linked to natural resource and land-use databases can aid in identifying reference and degraded systems (see the Land-Use Characterization for Nutrient and Sediment Risk Assessment, Wetland module #17); Johnston et al., 1988, 1990; Gwin et al., 1999; Palik et al., 2000; Brown and Vivas 2004). Some examples of watershed characteristics which can be evaluated using GIS and aerial photography include land use, land cover (including riparian vegetation), soils, bedrock, hydrography, and infrastructure (e.g., roads or railroads). Changes in point sources can be monitored through the NPDES permit program (USEPA 2000). Changes in nonpoint sources can be evaluated through the identification and tracking of wetland loss and/or degradation, increased residential development, urbanization, increased tree harvesting, shifts to more intensive agriculture with greater fertilizer use or increases in livestock numbers, and other land use changes. Local planning agencies should be informed of the risk of increased anthropogenic stress and encouraged to guide development accordingly.

IDENTIFYING AND CHARACTERIZING REFERENCE WETLANDS

The term "reference" in this document refers to those systems that are least impaired by anthropogenic effects. The use of the term reference is confusing because of the different meanings that are currently in use in different classification methods, particularly its use in hydrogeomorphic [HGM] wetland classification. A discussion of the term reference and its multiple meanings is provided in Chapter 3.

Watersheds with little or no development that receive minimal anthropogenic inputs could potentially contain wetlands that may serve as minimally impaired reference sites. Watersheds with a high percentage of the drainage basin occupied by urban areas, agricultural land, and altered hydrology are likely to contain wetlands that are impaired or could potentially be considered "at risk" for developing problems. Wetland loss in the landscape also should be considered when assessing watershed characteristics for reference wetland identification. Biodiversity can become impoverished due to wetland fragmentation or decreases in regional wetland density even in the absence of site-specific land-use activities. Reference wetlands may be more difficult to locate if fragmentation of wetland habitats is significant and may no longer represent the biodiversity of minimally disturbed wetlands in the region. The continued high rate of wetland loss in most States dictates that multiple reference sites be selected to ensure some consistency in reference sites for multiple year sampling programs (Liebowitz et al., 1992; Kentula et al., 1993). Once the watershed level has been considered, a more site-specific investigation can be initiated to better assess wetland condition.

The ideal reference site will have similar soils, vegetation, hydrologic regime, and landscape setting to other wetlands in the region (Adamus 1992; Liebowitz et al., 1992; Kentula et al., 1993; Detenbeck et al., 1996). Classification of wetlands, as discussed in Chapter 3, may aid in identifying appropriate reference wetlands for specific regions and wetland types. Wetland classification should be supplemented with information on wetland hydroperiod to assure that the selected reference wetlands are truly representative of wetlands in the region, class, or subclass of interest. Reference wetlands may not be available for all wetland classes. In that case, data from systems that are as close as possible to the assumed unimpaired state of wetlands in the wetland class of interest should be sought from States within the same geologic province. Development of a conceptual reference may be important if appropriate reference sites cannot be found in the local region or geologic province. Techniques for defining a conceptual reference are discussed at some length in Harris et.al., (1995), Trexler (1995), and Toth et.al., (1995).

Reference wetlands should be selected based on low levels of human alteration in their watersheds (Liebowitz et.al., 1992; Kentula et.al., 1993; USEPA 2000). Selecting reference wetlands usually involves assessment of land-use within watersheds and visits to individual wetland systems to ground-truth expected land-use and check for unsuspected impacts. Ground-truthing visits to reference wetlands are crucial for identification of ecological impairment that may not be apparent from land-use and local habitat conditions. Again, sufficient sample size is important to characterize the range of conditions that can be expected in the least impacted systems of the region (Detenbeck et.al., 1996). Reference wetlands should be identified for each ecoregion or geological province in the State lands and then characterized with respect to ecological integrity. A minimum of three low impact reference systems is recommended for each wetland class for statistical analyses. However, power analysis can be performed to determine the degree of replication necessary to detect an impact to the systems being investigated (Detenbeck et.al., 1996; Urquhart et.al., 1998). Highest priority should be given to identifying reference systems for those wetland types considered to be at the greatest risk from anthropogenic stress.

WHEN TO SAMPLE

Sampling may be targeted to the periods when effects are most likely to be detected – the index period. The appropriate index period should be defined by what the investigator is trying to investigate and what taxonomic assemblage or parameters are being used for that investigation (Barbour et.al., 1999). For example, increased nutrient concentrations and sedimentation from non-point sources may occur following periods of high runoff during spring and fall, while point sources of nutrient pollutants may cause plankton blooms and/or increased water and soil nutrient concentrations in wetland pools during times of low rainfall. Hence, different index periods may be needed to detect effects from point source and nonpoint source nutrients, respectively. Each taxonomic assemblage studied also should have an appropriate index period—usually in the growing season (see assemblage methods in the Maine case study: http://www.epa.gov/waterscience/criteria/wetlands/index.html).

The index period window may be early in the growing season for amphibians and algae. Other assemblages, such as vegetation and birds, may benefit from a different sampling window for the index period; see the assemblage specific modules for recommendations. Once wetland condition has been characterized, one-time annual sampling during the appropriate index period may be adequate for multiple year monitoring of indicators of nutrient status, designated use, and biotic integrity. However, criteria and ecological indicator development may benefit from more frequent sampling to define conditions that relate to the stressor or perturbor of interest (Karr and Chu 1999; Stevenson 1996; Stevenson 1997). Regardless of the frequency of sampling, selection of index periods and critical review of the data gathered and analyzed should be done to scientifically validate the site characterization and index periods for data collection.

Ideally, water quality monitoring programs produce long-term data sets compiled over multiple years to capture the natural, seasonal, and year-to-year variations in biological communities and constituent concentrations (e.g., Tate 1990; Dodds et.al., 1997; McCormick et.al., 1999; Craft 2001; Craft et al., 2003; Zheng et al., 2004). Multiple-year data sets can be analyzed with statistical rigor to identify the effects of seasonality and variable hydrology. Once the pattern of natural variation has been described, the data can be analyzed to determine the ecological state of the wetland. Long-term data sets have also been important in influencing management decisions about wetlands, most notably in the Everglades, where long-term data sets have induced Federal, State, and Authorized Tribal actions for conservation and restoration of the largest wetland system in the U.S. (see Davis and Ogden 1994; Everglades Interim Report, South Florida Water Management District [SFWMD, 1999]; Everglades Consolidated Report [SFWMD, 2000, 2001]; 1994 Everglades Forever Act, Florida Statute § 373.4592).

In spite of the documented value of long-term data sets, there is a tendency to intensively study a wetland for one year before and one year after treatment. A more cost-effective approach may be to measure only the indices most directly related to the stressor of interest (i.e., those parameters or indicators that provide the best information to answer the specific management question), but to double or triple the monitoring period. Multiple years (two or more) of data are often needed to identify the effects of years with extreme climatic or hydrologic conditions. Comparisons over time between reference and at risk or degraded systems can help describe biological response and annual patterns in the presence of changing climatic conditions. Multi-year data sets also can help describe regional trends. Flooding or drought may significantly affect wetland biological communities and the concentrations of water column and soil constituents. Effects of uncommon climatic events can be characterized to discern the overall effect of management actions (e.g., nutrient reduction, water diversion) if several years of data are available to identify the long-term trends.

At the very minimum, two years of data before and after specific management actions, but preferably three or more each, are recommended to evaluate the cost-effectiveness of management actions with some degree of certainty (USEPA 2000). If funds are limited, restricting sampling frequency and/or numbers of indices analyzed should be considered to preserve a longer-term data set. Reducing sampling frequency or numbers of parameters measured will allow for effectiveness of management approaches to be assessed against the high annual variability that is common in most wetland systems. Wetlands with high hydrological variation from year to year may benefit from more years of sampling both before and after specific management activities to identify the effects of the natural hydrologic variability (Kadlec and Knight 1996).

CHARACTERIZING PRECISION OF ESTIMATES

Estimates of cause-response relationships, nutrient and biological conditions in reference systems, and wetland conditions in a region are based on sampling; hence, precision should be assessed. Precision is defined as the "measure of the degree of agreement among the replicate analyses of a sample, usually expressed as the standard deviation" (APHA 1999). Determining precision of measurements for one-time assessments from single samples in a wetland is often important. The variation associated with one-time assessments from single samples can be determined by re-sampling a specific number of wetlands during the survey. Measurement variation among replicate samples then can be used to establish the expected variation for one-time assessment of single samples. Re-sampling does not establish the precision of the assessment process, but rather identifies the precision of an individual measurement (Kentula et.al., 1993).

Re-sampling frequency is often conducted for one wetland site in every block of 10 sites. However, investigators should adhere to the objectives of re-sampling (often considered an essential element of QA/QC) to establish an assessment of the variation in a one-time/sample assessment. Often, more than one in 10 samples should be replicated in monitoring programs to provide a reliable estimate of measurement precision (Barbour et.al., 1999). The reader should understand that this is a very brief description of the concerns about precision, and that any monitoring program or study involving monitoring should include consultation with a professional statistician before the program begins and regularly during the course of the monitoring program to assure statistical rigor.

4.3 SAMPLING PROTOCOL

APPROACHES TO SAMPLING DESIGN

The following sections discuss three different approaches to sampling design, probabilistic, targeted, and BACI. These approaches have advantages and disadvantages that under different circumstances warrant the choice of one approach over the other (Table 4). The decision as to the best approach for sample design in a new monitoring program should be made by the water quality resource manager or management team after careful consideration of the different approaches. For example, justification of a dose-response relationship is confounded by lack of randomization and replication and should be considered in choosing a sampling design for a monitoring program.

PROBABILISTIC SAMPLING DESIGN FOR ASSESSING CONDITION

Probabilistic sampling – a sampling process wherein randomness is requisite (Hayek 1994) – can be used to characterize the status of water quality conditions and biotic integrity in a region's wetland system. This type of sampling design is used to describe the average conditions of a wetland population, identify the variability among sampled wetlands, and help determine the range of wetland system conditions in a region. Data collected from a probabilistic random sample design generally will be characteristic of the dominant class or type of wetland in the region, but rare wetlands may be under-represented or absent from the probabilistically sampled wetlands. Additional sampling sites may need to be added to precisely characterize the complete range of wetland conditions and types in the region.

Probabilistic designs are often modified by stratification (such as classification). Stratified random sampling is a type of probabilistic sampling where a target population is divided into relatively homogenous groups or classes (strata) prior to sampling based on factors that influence variability in that population (Hayek 1994). Stratification by wetland size and class or types ensures more complete information about different types of wetlands within a region. Sample statistics from random selection alone would be most characteristic of the dominant wetland type in a region if the population of wetlands is not stratified.

Many State 305(b) and watershed monitoring programs utilize stratified random sampling designs, and we will further discuss this type of probabilistic sampling. Pilot projects in Maine, Montana, and Wisconsin all use stratified random sampling design. Details of these monitoring designs can be found in the Case Studies module #14 and on the Web at http://www.epa.gov/waterscience/criteria/wetlands/index.html.

Stratification is based on identifying wetland systems in a region (or watershed) and then selecting an appropriate sample of systems from the defined population. The determination of an appropriate sample population usually is dependent on the management questions being asked. A sample population of isolated depressional wetlands could be identified as a single stratum, but investigations of these wetlands would not provide any information on riparian wetlands in the same region. If the goal of the monitoring program is to identify wetland condition for all wetland classes within a region, then a sample population of wetlands should be randomly selected from all wetlands within each class. In practice, most State programs stratify random populations by size, wetland class (see Chapter 3), and landscape characteristics or location (see http://www.epa.gov/waterscience/criteria/wetlands/index.html, module #14).

Once the wetlands for each stratum have been identified, the list of wetlands can be used to select a spatially-balanced stratified random sample. Spatial-balance will ensure spatial coverage over the assessment region, usually increase the types of wetlands sampled (assuming classes of wetlands vary spatially), and reduce spatial autocorrelation among the sampled wetlands. For example, EMAP implements spatially-balanced samples using Generalized Random Tessellation Stratified (GRTS) designs applied to GIS coverages of wetlands within the assessment region. GRTS using a hierarchical grid randomization process to ensure the sites are spatially distributed (Paulsen et.al., 1991; Stevens and Olsen 2004). Estimates of ecological conditions from these kinds of modified probabilistic sampling designs can be used to characterize the water quality conditions and biological integrity of wetland systems in a region, and over time, to distinguish trends in ecological condition within a region. (See http://www.epa.gov/owow/wetlands/bawwg/case/mtdev.html and http://www.epa.gov/owow/wetlands/bawwg/case/fl1.html).

TARGETED DESIGN

A targeted approach to sampling design may be more appropriate when resources are limited (Stern 2004). The example of targeted sampling described here involves defining a gradient of impairment. Once the gradient has been defined and systems have been placed in categories of impairment, investigators focus the greatest efforts on identifying and characterizing wetland systems or sites likely to be impacted by anthropogenic stressors, and on relatively undisturbed wetland systems or sites (see Identifying and Characterizing Reference Systems, Chapter 3), that can serve as regional, sub-regional, or watershed examples of natural biological integrity. Florida Department of Environmental Protection (FL DEP) uses a targeted sampling design for developing thresholds of impairment with macroinvertebrates (http://www.epa.gov/owow/wetlands/bawwg/case/fl2.html). Choosing sampling stations that best allow comparison of ecological integrity at reference wetland sites of known condition can conserve financial resources. A sampling design that tests specific hypotheses (e.g., the FL DEP study tested the effect of elevated water column phosphorus on macroinvertebrate species richness) generally can be analyzed with statistical rigor and can conserve resources by answering specific questions. Furthermore, identification of systems with problems and reference conditions eliminates the need for selecting a random sample of the population for monitoring.

Targeted sampling assumes some knowledge of the systems sampled (Stern 2004; Kentula et.al., 1993). Systems based on independent variables with evidence of degradation are compared to reference systems that are similar in their physical structure (i.e., in the same class of wetlands). Wetland systems should be viewed along a continuum from reference to degraded. An impaired or degraded wetland is a system in which anthropogenic impacts exceed acceptable levels or interfere with beneficial uses. Comparison of the monitoring data to that collected from reference wetlands will allow characterization of the sampled systems. Wetlands identified as “at risk” should be evaluated through a sampling program to characterize the degree of degradation. Once characterized, the wetlands should be placed in one of the following categories:

1. Degraded wetlands—wetlands in which the level of anthropogenic perturbance interferes with designated uses.

2. High-risk wetlands—wetlands where anthropogenic stress is high but does not significantly impair designated uses. In high-risk systems, impairment is prevented by one or a few factors that could be changed by human actions, though characteristics of ecological integrity are already marginal.

3. Low-risk wetlands—wetlands where many factors prevent impairment, stressors are maintained below problem levels, and/or no development is contemplated that would change these conditions.

4. Reference wetlands—wetlands where the ecological characteristics most closely represent the pristine or minimally impaired condition.

Once wetland systems have been classified based on their physical structure (see Chapter 3) and placed into the above categories, specific wetlands need to be selected for monitoring. At this point, randomness is introduced; wetlands should be randomly selected within each class and risk category for monitoring. An excellent example of categorizing wetlands in this manner is given in the Ohio Environmental Protection Agency's (OH EPA) case study, available at: http://www.epa.gov/owow/wetlands/bawwg/case/oh1.html. They used the Ohio Rapid Assessment Method to categorize wetlands by degree of impairment. The Minnesota Pollution Control Agency (MPCA) also used a targeted design for monitoring wetlands (http://www.epa.gov/owow/wetlands/bawwg/case/mn1.html). They used the best professional judgment of local resource managers to identify reference sites and those with known impairment from identified stressors (agriculture and stormwater runoff).

Targeted sampling design involves monitoring identified degraded systems and comparable reference systems most intensively. Low risk systems are monitored less frequently (after initial identification) unless changes in the watershed indicate an increased risk of degradation.

Activities surrounding impaired wetland systems may be used to help identify which actions negatively affect wetlands, and therefore may initiate more intensive monitoring of at-risk wetlands. Monitoring should focus on factors likely to identify ecological degradation and anthropogenic stress and on any actions that might alter those factors. State water quality agencies should encourage adoption of local watershed protection plans to minimize ecological degradation of natural wetland systems. Development plans in the watershed should be evaluated to identify potential future stressors. Ecological degradation often gradually increases due to many growing sources of anthropogenic stress. Hence, frequent monitoring may be warranted for high-risk wetlands if sufficient resources remain after meeting the needs of degraded wetlands. Whenever development plans appear likely to alter factors that maintain ecological integrity in a high-risk wetland (e.g., vegetated buffer zones), monitoring should be initiated at a higher sampling frequency in order to enhance the understanding of baseline conditions (USEPA 2000).

BEFORE/AFTER, CONTROL/IMPACT (BACI) DESIGN

An ideal before/after impact survey has several features: 1) the type of impact, time of impact, and place of occurrence should be known in advance; 2) the impact should not have occurred yet; and, 3) control areas should be available (Green 1979). The first feature allows the surveys to be efficiently planned to account for the probable change in the environment. The second feature allows a baseline study to be established and extended as needed. The last feature allows the surveyor to distinguish between temporal effects unrelated to the impact and changes related to the impact. In practice however, advance knowledge of specific impacts is rare, and the ideal impact survey is rarely conducted. BACI designs modified to monitor impacts during or after their occurrence still can provide information, but there is an increase in the uncertainty associated with the results and the likelihood of finding a statistically significant change due to the impact is less probable. In addition, other aspects of survey design are dependent on the study objectives, e.g., the sampling interval, the length of time the survey is conducted (i.e., sampling for acute versus chronic effects), and the statistical analyses appropriate for analyzing the data (Suter 1993).

The best interval for sampling is determined by the objectives of the study (Kentula et.al., 1993). If the objective is to detect changes in trends (e.g., regular monitoring for detection of changes in water quality or biotic integrity), regularly spaced intervals are preferred because the analysis is easier. On the other hand, if the objective is to assess differences before and after impact, then samples at random time points are advantageous. Random sample intervals reduce the likelihood that cyclic differences unforeseen by the sampler will influence the size of the difference before and after the impact. For example, surveys taken every summer for a number of years before and after a clear-cut may show little difference in system quality; however, differences may exist that can only be detected in the winter and therefore may go undetected if sampling occurs only during summer.

The simplest impact survey design involves taking a single survey before and after the impact event (Green 1979). This type of design has the obvious pitfall that there may be no relationship between the observed event and the changes in the response variable—the change may be entirely coincidental. This pitfall is addressed in BACI design by comparing before and after impact data to data collected from a similar control system nearby. Data are collected before and after a potential disturbance in two areas (treatment and a control), with measurements on biological and environmental variables in all combinations of time and area (Green 1979). We will use a clear-cut adjacent to a wetland as an example to illustrate the BACI design. The sampling design is developed to identify the effects of clear-cutting on adjacent wetland systems. In the simplest BACI design, two wetlands would be sampled. One wetland would be adjacent to the clear-cut (the treatment wetland); the second wetland would be adjacent to a control site that is not clear-cut. The control site should have characteristics (soil, vegetation, structure, functions) similar to the treatment wetland and is exposed to climate and weather similar to the first wetland. Both wetlands are sampled at the same time points before the clear-cut occurs and at the same time point after the clear-cut takes place. This design is technically known as an area-by-time factorial design. Evidence of an impact is found by comparing the control site samples (before and after) with the treatment site before and after samples. Area-by-time factorial design allows for both natural wetland-to-wetland variation and coincidental time effects. If there is no effect of the clear-cut, then change in system quality between the two time points should be the same. If there is an effect of the clear-cut, the change in system quality between the two time points should be different.

CONSIDERATIONS FOR BACI DESIGN

There are some potential problems with BACI design. First, because the control and impact sites are not randomly assigned, observed differences between sites may be related solely to some other factor that differs between the two sites. One could argue that it is unfair to ascribe the effect to the impact (Hurlbert 1984; Underwood 1991). However, as pointed out by Stewart-Oaten et.al., (1986), the survey is concerned about a particular impact in a particular place, not in the average of the impact when replicated in many different locations. Consequently, it may be possible to detect a difference between these two specific sites. Even so, if there are no randomized replicate treatments, the results of the study cannot be generalized to similar events at different wetlands. In any case, the likelihood that the differences between sites are due to factors other than the impact can be reduced by monitoring several control sites (Underwood 1991) because multiple control sites provide some information about potential effects of other factors.

The second and more serious concern with the simple Before-After design with a single sampling point before and after the impact is that it fails to recognize that there may be natural fluctuations in the characteristic of interest that are unrelated to any impact (Hurlbert 1984; Stewart-Oaten 1986). Single samples before and after impact would be sufficient to detect the effects of the impact if there were no natural fluctuations over time. However, if the population also has natural fluctuations over and above the long-term average, then it is impossible to distinguish between cases where there is no effect from cases where there is an impact. Consequently, measured differences in system quality may be artifacts of the sampling dates and natural fluctuations may obscure differences or lead one to believe differences are present when they are not.

The simple BACI design was extended by Stewart-Oaten et.al., (1986) by pairing surveys at several selected time points before and after the impact to help resolve the issue of psuedoreplication (Hulbert 1984). This modification of the BACI design is referred to as BACI-PS (Before-After, Control-Impact Paired Series design). The selected sites are measured at the same time points. The rationale behind this paired design is that repeated sampling before the impact gives an indication of the pattern of differences of potential change between the two sites. BACI-PS study design provides information both on the mean difference in the wetland system quality before and after impact and on the natural variability of the system quality measurements. The resource manager has detected an effect if the changes in the mean difference are large relative to natural variability. Considerations for sampling at either random or regularly spaced intervals also apply here. Replication of samples should also be included if resources allow in order to improve certainty of analytical results.

Violation of the BACI assumptions may invalidate conclusions drawn from the data. Enough data should be collected before the impact to identify the trends in the communities of each sampling site if the BACI assumptions are to be met. Clearly defining the objectives of the study and identifying a statistically testable model of the relationships the investigator is studying can help resolve these issues (Suter 1993).

The designs described above are suitable for detecting longer-term chronic effects in the mean level of the variable of interest. However, the impact may have an acute effect (i.e., effects only last for a short while) or may change the variability in response (e.g., seasonal changes become more pronounced) in some cases. The sampling schedule can be modified so that it occurs at two temporal scales (enhanced BACI-PS design) that encompass both acute and chronic effects (Underwood 1991). The modified temporal design introduces randomization by randomly choosing sampling occasions in two periods (Before and After) in the control or impacted sites. The two temporal scales (sampling periods vs. sampling occasions) allow the detection of a change in mean and of a change in variability after impact. For example, groups of surveys could be conducted every year with five surveys one week apart randomly located within each group. The analysis of such a design is presented in Underwood (1991). Again, multiple control sites should be used to counter the argument that detected differences are specific to the sampled site. The September 2000 issue of the *Journal of Agricultural, Biological, and Environmental Statistics* discusses many of the advantages and disadvantages of the BACI design and provides several examples of appropriate statistical analyses for evaluation of BACI studies.

4.4 SUMMARY

State monitoring programs should be designed to assess wetland condition with statistical rigor while maximizing available management resources. The three approaches described in this module—probabilistic sampling, targeted/tiered approach, and BACI (Before/After, Control/Impact)—present study designs that allow one to obtain a significant amount of information for statistical analyses. The sampling design selected for a monitoring program should depend on the management question being asked. Sampling efforts should be designed to collect information that will answer management questions in a way that will allow robust statistical analysis. In addition, site selection, characterization of reference sites or systems, and identification of appropriate index periods are all of particular concern when selecting an appropriate sampling design. Careful selection of sampling design will allow the best use of financial resources and will result in the collection of high quality data for evaluation of the wetland resources of a State. Examples of different sampling designs currently in use for State wetland monitoring are described in the Case Study module #14 on the Web site: http://www.epa.gov/waterscience/criteria/wetlands/index.html. Well-designed monitoring programs tend to produce data that managers can use in nutrient criteria development, such as in developing reference networks or utilizing distribution-based approaches.

Table 4. Comparison of Probabilistic, Targeted, and BACI Sampling Designs

Probabilistic	Targeted	BACI
Random selection of wetland systems from entire population within a region.	Targeted selection of wetlands based on problematic (wetland systems known to have problems) and reference wetlands.	Selection of wetlands based on a known impact.
This design requires minimal prior knowledge of wetlands within the sample population for stratification.	This design requires prior knowledge of wetlands within the sample population.	This design requires knowledge of a specific impact to be analyzed.
This design may use more resources (time and money) to randomly sample wetland classes because more wetlands may need to be sampled.	This design utilizes fewer resources because only targeted systems are sampled.	This design may use fewer resources because only wetlands with known impacts and associated control systems are sampled.
System characterization for a class of wetlands is more statistically robust.	System characterization for a class of wetlands is less statistically robust, although characterization of a targeted wetland may be statistically robust.	Characterization of the investigated systems is statistically robust.
Rare wetlands may be under-represented or absent from the sampled wetlands.	This design may miss important wetland systems if they are not selected for the targeted investigation.	The information gained in this type of investigation is not transferable to wetland systems not included in the study.
This design is potentially best for regional characterization of wetland classes, especially if water quality conditions are not known.	This design is potentially best for site-specific and watershed-specific criteria development when water quality conditions for the wetland of interest are known.	This design is potentially best for monitoring restoration or creation of wetlands and systems that have specific known stressors.