



# **Chapter I**

# **Statistical Design**

## **Chapter Authors**

Donald G. Uzarski, Central Michigan University  
Sango Otieno, Grand Valley State University

# Sampling design

Sampling design is a description of the sample collection plan that specifies the number, type, and location (spatial and/or temporal) of sampling units to be selected for assessment. The paramount purpose of sampling designs -- within the context of the Consortium's recommendations for monitoring -- is to ensure the collection of data that are representative of an area and of adequate scope to permit one to draw logical conclusions about a population of interest. But deciding how to sample is often difficult, because one must consider trade-offs between the costs and benefits of the amount and type of sampling undertaken. Thus, any sampling design represents a balance between the study objectives and the constraints of cost, time, logistics, safety and existing technology.

One must also make numerous practical and statistical decisions to be confident that a sampling design and indicator measurements provide the necessary "vital sign" information (Busch and Trexler 2003). The following questions can help make those decisions:

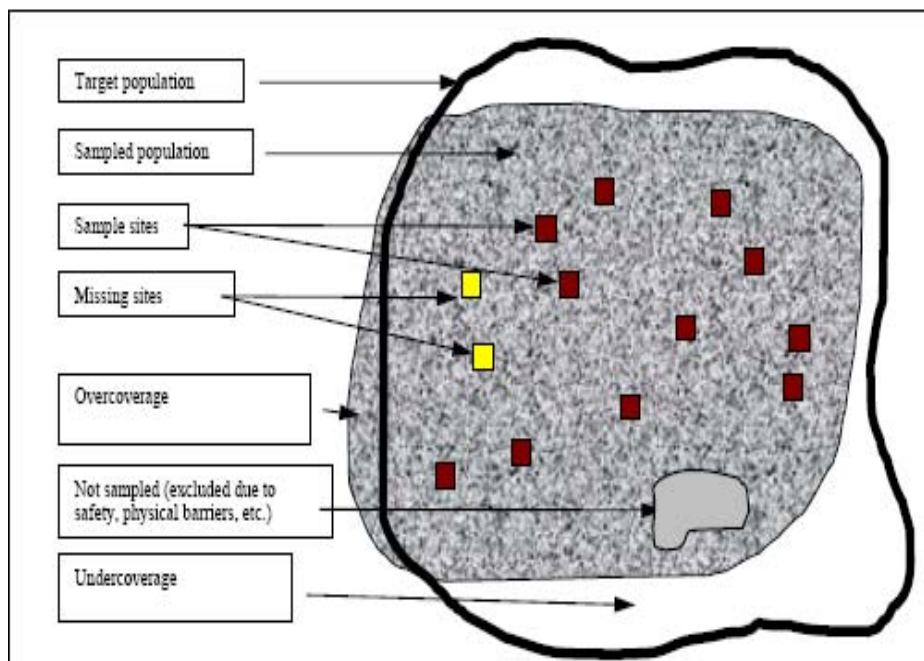
- What are the defining spatial boundaries of the ecological system?
- What is the appropriate temporal frame (time of year) for sampling?
- What is the appropriate time interval between samples?
- What sample size is necessary to estimate the value of the indicator?
- What survey design is most efficient (random, systematic, stratified random)?
- What is the appropriate unit of measure for the indicator variable?
- Is there an optimal sample unit size and shape for estimating the value of the indicator?
- What are the trade-offs between gains in precision and statistical power versus the additional costs per sample?
- How can the implementation plan be designed so that uncertainty about the true state of the ecological system is minimized?

Many of these questions are addressed in this chapter, including those involving target populations and sampling frames (Figure 1-1), allocation and arrangement of samples (*membership design*), frequency of sampling occasions (*revisit design*), measurements to be taken at sampling locations (*response design*), and the number of samples required to meet stated objectives (*sample size*). Italicized terms are described later in this chapter. Sampling designs, within the context of the Consortium's recommendations for the intended implementation plan, encapsulate the series of decisions that dictate where, when, and how to sample a "vital sign" indicator (e.g., the indicator nitrate as a measure of wetland water chemistry) (Elzinga et al. 2001). A sound sampling design requires clear and concise monitoring objectives and must be flexible.

Because the intent of this document is to propose a robust implementation plan that can meet the needs of the Great Lakes coastal wetland science community and policymakers well into the future, the designs must be able to accommodate changes in management and funding priorities, as well as environmental changes. Likewise, the description of a good sampling design should be appropriately concise, understandable and manageable. Overly complex designs can be confusing and may make an implementation plan less accessible to its key audience, few of whom are likely to be familiar with statistics and sampling design theory. Therefore, a description of an implementation plan should begin simply and add complexity conservatively and only when needed to explain how to achieve specific objectives. Of course, to monitor the health and integrity of a coastal wetland, some level of complexity cannot be avoided, particularly when the region of interest is large, remote, spatially complex and difficult to access (McDonald and Geissler 2004).

Since monitoring objectives call for estimating the *status* of wetlands in a region, *trends* in their condition, or both, these two terms are used explicitly in this chapter and follow definitions given by Urquhart et al. (1998) and McDonald(2003). **Status** is a measure of a current attribute, condition or state, and is typically summarized as a

population mean. In the case of assessing wetland **status**, the population consists of all wetlands in the region of interest. So “wetland status” describes the average condition of all wetlands in an area at a specific time. **Trend** is a measure of change through time. This change can refer to a population parameter such as a mean (*net trend*; i.e., change in the average condition of wetlands through time), or of an individual member or unit of a population (*gross trend*; i.e., change in the condition of one specific wetland of interest through time). Status is typically estimated by sampling many different units (different wetlands) throughout the area of interest during a single time interval. In contrast, the study of trends requires repeated sampling, sometimes of the same wetlands, and sometimes of different wetlands. The question of whether the program goal is to estimate status, trends or both is one of the first and most important things that must be addressed.



**Figure 1-1.** Conceptual illustration of terms used to describe various units associated with sampling a population of interest. Each square represents a different wetland within the sampling region of interest.

The first step in developing an implementation plan is to define the geographic bounds of the region of interest and perform an inventory of all of the study units (wetlands) within the region. The complete set of study units (study units = wetlands) in the region is the **target population** (Figure 1.1). Some portion of the region may be inaccessible or otherwise unsuitable for evaluation. Consequently, we compile a subset of units that can be sampled, and these together make up the **sample population**.

The next important step when developing a sampling design is to define the environmental units of interest within a specified study area (Figure 1-1). A population consists of **elements**, i.e., the objects on which a measurement is taken (Scheaffer et al. 1990). This is the basic “unit” of observation. In our case, the elements are the individual wetlands. The implementation plan becomes defined by selecting a sufficiently representative subset of units (wetlands) for sampling (indicated by squares in Figure 1-1).

**Sampling unit** refers to the place that is actually sampled. We quantify our target population by using a **sampling frame**, defined as the collection of sampling units. A **sample** is a subset of units chosen to evaluate the condition of each unit through counts, observation, or other form of measurement. If the sampling units are selected using some type of random draw, the sample is said to constitute a **probability-based sample** (because it is equally likely that any unit could have been chosen). Whenever possible, a probability-based sampling design should be used. This lets

us argue that the “average” value calculated for the wetlands sampled truly represents the sample population as a whole.

If estimates of average condition are biased because the locations were not completely randomly chosen, they are subject to “nonsampling error” (Lessler and Kalsbeek 1992). Nonsampling error reduces the precision and accuracy of estimates. Lessler and Kalsbeek (1992) identified three components of nonsampling error, some of which may be unavoidable.

1. Frame error results when the sampled population is very different than the target population (Figure 1-1). The two types of frame error are overcoverage and undercoverage. Overcoverage occurs when the sampled population contains elements that were not part of the target population. Undercoverage occurs when elements of the target population are omitted from the sampled population.
2. Nonresponse error results from the failure to obtain measurements for all of the samples originally selected (yellow squares in Figure 1-1). When the missing measurements are very different from the values obtained from the wetlands that could be sampled, the estimates calculated from the available data may be biased.
3. Measurement error is defined as the difference between the measurements obtained during sampling and the true value of the measure. This can result from observers’ detection errors or from using inaccurate instruments.

Once the target population and sampling frame have been determined, a strategy must be developed for determining how many samples should be collected, allocating the sampling effort appropriately across the sampling frame, determining (randomly) which specific wetlands in a subregion should be sampled, and timing the visits for sampling. Most sample designs selected for use with Consortium-developed metrics will involve rotating field sampling efforts through various sets of sample units over time. In this situation, it is useful to define a **panel** of sample units. A panel is a group of wetlands whose members are always sampled together according to a schedule of repeating “revisit” time periods (Urquhart and Kincaid 1999, McDonald 2003). See Figure 1-2 for a schematic representation of different revisit designs.

The rules by which units (wetlands) in the population become members of a panel are called the **membership design** (McDonald 2003). Membership design specifies the spatial allocation procedure. One familiar membership design strategy is simple random sampling. This procedure involves drawing units from a population at random (i.e., with equal probability). Unfortunately, this often fails to produce an ideal spatial pattern of samples across the study region because the habitat itself may be spatially uneven. In particular, simple-randomly selected samples can often be patchily distributed or clustered, leaving large areas of the frame unsampled. An alternative is to draw a spatially balanced random sample following the methods described by Stevens and Olsen (2004). This method involves splitting the sampling frame into a number of zones (strata) and randomly selecting the required number of sampling units from within each zone. This “stratified-random” approach allows for a spatially balanced, random draw of samples with variable inclusion probabilities. Often, a designer generates an ordered list of sample units for each stratum that can support additions and or omissions of sample collections while retaining spatial balance. These features provide considerable flexibility and efficiency to a sampling design.

**Index** sites – also known as *sentinel* or *intensive* sites – are sampling locations that are (i) visited repeatedly and regularly, (ii) sites where more detailed measures are made, or (iii) both. Conversely, a **survey** (or *extensive*) site is a sampling location that is visited once or on an irregular basis, or where less detailed measures are obtained. Generally, “always” visiting a sampling site provides data that are most useful for detecting temporal variation (trends). The data array from this type of design would look like Table 1-1 (from Urquhart et al. 1998), except that the rows would represent the selected set of wetlands, rather than all of the wetlands. This balanced data structure has substantial appeal, and is the design most monitoring personnel and ecologists seem to favor. In fact, even some statisticians (Skalski 1990) think this is by far the best temporal design for trend detection.

However, repeated sampling violates the equal-likelihood-of-sampling assumption of probability based designs and introduces bias into status estimates. A trend-detection design is most powerful for determining changes through time, if the same sites are sampled on every occasion and samples are collected at regular intervals. However, this design cannot be used to determine overall status because the sites are not randomly selected.

In contrast, the most statistically powerful design to summarize status in a region involves randomly selecting the complete set of sampling sites on every occasion. Some sites may be resampled, but only if they are selected by chance. This design is less powerful for illustrating trends.

What kind of sampling design should be used to monitor a spatially dispersed (regional) ecological resource of interest? Often, a monitoring program requires assessing both overall status and temporal trends in the region of interest. We want it to have good power for detecting temporal trends in a regional population while simultaneously providing precise estimates of that population's status. This is where compromise strategies such as panel sampling are most appropriate (Urquhart et al. 1998; see below). A mixed design incorporates some pattern of revisits to sites (wetlands), but that also involves collecting some new samples from the regional population for each "revisit."

**Table 1-1.** Tabular organization of response values, averages and slopes (Urquhart et al. 1998). Note that in the context of Consortium: sampling unit = wetland.

Sampling unit (= lake)	Time period (= Year)				Averages	Slope (= trend)
	1	2	...	t		
1	$Y_{11}$	$Y_{12}$	...	$Y_{1t}$	$\bar{Y}_1$	$\beta_1$
2	$Y_{21}$	$Y_{22}$	...	$Y_{2t}$	$\bar{Y}_2$	$\beta_2$
...	...	...	...	...	...	...
$l$	$Y_{l1}$	$Y_{l2}$	...	$Y_{lt}$	$\bar{Y}_l$	$\beta_l$
Averages	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$	...	$\bar{Y}_{.t}$	$\bar{Y}_{..}$	$\bar{\beta}_{.}$

The Consortium adopted the notation of Urquhart et al. (1998) to describe revisit designs for brevity and consistency. Fig. 1-2 schematically summarizes four designs for a monitoring program that assesses a study region over a period of 12 years. Figure 1-2 relates to Table 1-1 as follows:

PANEL	SIZE	TIME PERIODS (=YEARS)												
		1	2	3	4	5	6	7	8	9	10	11	12	...
<b>DESIGN 1 = SAME SITES (= LAKES)</b>														
1	60	X	X	X	X	X	X	X	X	X	X	X	X	...
<b>DESIGN 2 = NEW SITES (= LAKES)</b>														
1	60	X												
2	60		X											
3	60			X										
4	60				X									
5	60					X								
6	60						X							
7	60							X						
8	60								X					
9	60									X				
10	60										X			
11	60											X		
12	60												X	
...	...													...
<b>DESIGN 3 = AUGMENTED SERIALLY ALTERNATING</b>														
1	50	X				X				X				...
2	50		X				X				X			...
3	50			X				X				X		...
4	50				X				X				X	...
COMMON	10	X	X	X	X	X	X	X	X	X	X	X	X	...
<b>DESIGN 4 = PARTIALLY AUGMENTED SERIALLY ALTERNATING</b>														
1	35	X				X				X				...
2	35		X				X				X			...
3	35			X				X				X		...
4	40				X				X				X	...
1A	5	X	X			X				X				...
2A	5		X	X			X				X			...
...	...													...
1B	5	X				X	X			X				...
...	...													...
1C	5	X				X				X	X			...
...	...													...

**Figure 1-2.** Examples of four different revisit designs, beginning with the simplest, in which a single panel or set of sampling units are visited on every sampling occasion, and ending with a complex partially augmented serially alternating design (Urquhart et al. 1998).

Each panel in Fig.1-2 represents a selection of rows from Table 1-1, subject to the restriction that no wetland ('lake' reference from Urquhart et al. 1998) occurs in more than one panel; the columns in Fig.1-2, except for the first two, are the same as those in Table 1-1. The X's in Fig. 1-2 identify the year(s) in which wetlands from a particular panel will be visited. The first two columns identify the panels and give the numbers of wetlands in the respective panels. Each of the four designs was developed for the same number of wetland visits each year (60). The designs therefore operate under the same fixed budget for field work, except that Design 4 must have fewer wetlands (lakes) (55) during the first year of sampling in order to have 60 in subsequent years. Each of the designs continues for more years

than are displayed in the table. The pattern of repetition (revisits) in the first three designs should be obvious, but the pattern in Design 4 deserves a bit more explanation. Fig 1-3 below (Urquhart, et al., 1999) illustrates a detailed form of Design 4.

**Design 4 - Partially Augmented Serially Alternating**  
 1-4 = Panels  
 A-J = Sub-Panels

	# OF SITES	YEAR 1	YEAR 2	YEAR 3	YEAR 4	YEAR 5	YEAR 6	YEAR 7	YEAR 8	YEAR 9	YEAR 10	YEAR 11	YEAR 12
1	35	X				X				X			
2	35		X				X				X		
3	35			X				X				X	
4	40				X				X				X
1A	5	X	X			X				X			
2A	5		X	X			X				X		
3A	5			X	X			X				X	
4A	5				X	X			X				X
1B	5	X				X	X			X			
2B	5		X				X				X		
3B	5			X				X				X	
4B	5				X				X				X
1C	5	X				X				X	X		
2C	5		X				X				X	X	
3C	5			X				X				X	X
4C	5				X				X				X
1D	5	X				X				X			
2D	5		X				X				X		
3D	5			X				X				X	

**Figure 1-3.** Partitioning of each panel into subpanels for Design 4 (Urquhart, N.S., and Kincaid, T.M. 1999).

Randomly select (without replacement) four panels of 55 wetlands each from the sample population of wetlands (lakes) (see Fig 1-3). Schedule each panel to be visited every four years, with a different panel starting in each of the first four years. Randomly divide each panel into 11 subpanels of five sites each (these are labeled A-K above, but only A-D are shown). Each time a panel is visited, randomly select a subpanel of five wetlands (lakes) to visit for two consecutive years - that year and the following one. The first four lines of Design 4 in Fig.1-3 collect the remaining subpanels that would not be visited in two consecutive years during the first 12 years representing 35, 35, 35, and 40 sites that will not be revisited respectively. The next two lines, labeled 1A and 2A, display the visit pattern of the first subpanels from the first two panels (panel 1 subpanel A and panel 2 subpanel A); the lines labeled 1B and 1C display the visit pattern of the second and third subpanels of panel 1.

As mentioned above the desired design must make a compromise between competing concerns of status and trend. Designs 1, 3, and 4 have similar power to detect trends because they include revisits to sites and incorporate augmentation to achieve connectedness. However, Designs 3 and 4 will give the most precise estimates of overall status because they include visits to the most sites (after 20 years Design 2 is superior for estimating status but still does not provide a good estimate of trend). There are two reasons why Design 4 is superior to Design 3. First, it provides a better estimate of the amount of “Wetland x Year” variation (i.e., to what extent does the estimate for a wetland depend on the year in which it was sampled); and second, it also causes less impact of physical sampling on the wetland due to repeat visits (i.e., taking the same samples from the same locations annually for 20 or more years; Urquhart et al. 1998).

## Sample Size Considerations and Ability to Detect Change

Populations in the real world are dynamic; change over time is to be expected. However, what is important is whether or not there has been *meaningful* change (meaningful to the ecosystem, or public), what has caused the observed change, and whether or not the resource is expected to change further.

To understand what constitutes a meaningful and significant change, one must differentiate between statistical significance and biological significance. Statistical significance relies on probability and is influenced by sample size. Thus, even trivial changes (from a biological perspective) can be judged to be statistically significant if the sample size is large enough. So, regardless of statistical significance, one should consider something biologically significant if it represents a major shift in ecosystem structure or function (e.g., loss of one or more species, addition of non-native species, changes in ecosystem processes). The term “effect size” is operationally defined to be the smallest difference that represents a biologically meaningful change in a variable of interest. It is typically expressed as a percentage of the average existing value. Effect size is a value judgment that can be decided on the basis of prior scientific evidence, best professional judgment, public consensus, or legislated regulations. However, the effect size is also an attribute that must be defined as part of the implementation plan so that the sample design can be developed to maximize the power of detecting a change of that magnitude, should it occur. The power of a statistical comparison is the ability to detect a biologically meaningful change when it occurs. A statistically significant change is not always biologically significant, but a biologically significant change must also be statistically significant; if the latter is not true, then the power of the statistical comparison must be increased.

Thus, from a monitoring standpoint, one should be concerned with both statistical significance and the power to detect a biologically significant effect. To answer this, it must be decided what level of statistical significance to attain (i.e., what is our Type I error rate or  $\alpha$ , discussed below), what level of change to consider biologically meaningful (what is the effect size), how certain one wishes to be to detect that change (what is the power), and how variable the indicator measure is that we are trying to estimate (what is the variance). The relationships among power, sample size, effect size and variation are summarized in equation 1 below.

$$\text{Power } \alpha \quad (\text{sample size}) \times (\text{effect size}) \frac{\quad}{(\text{variance}) \times (\text{number of groups being compared})} \quad (1)$$

In addition to implementation objectives, a **sampling objective** must be defined. Sampling objectives establish a desired level of statistical power, the capacity to detect a “real” change or trend, a minimum detectable change or effect size, and acceptable levels of both a false-change ( $\alpha$  or the probability of a Type I error) and a missed-change ( $\beta$  or the probability of a Type II error) (Elzinga et al. 2001). Sample size affects each of these components. The larger the sample size, the lower error and the greater the power to detect a change (Eqn 1). Reducing sample size, which is desirable for cost-effectiveness, leads to reduced power and higher error rates. These tradeoffs are mitigated by reducing variance estimates (in the denominator of equation 1), either through modifications in response design, another component (e.g., revisit design), or by accepting a higher minimum effect size (in the numerator of equation 1) (Steidl et al. 1997).

In general, sample size should be large enough to give a high probability of detecting any changes that are of management, conservation or biological importance, but not unnecessarily large (Manly 2001). Scientists traditionally seek to reduce Type I errors, and accordingly prefer small alpha levels. In a regional implementation plan with a strong resource-conservation mandate, however, it is preferable to employ an early warning philosophy by tolerating a higher alpha, but consequently increasing the power to detect differences or trends (Sokal and Rohlf 1995, Mapstone 1995, Roback and Askins 2005).

Accordingly, the Consortium has initially set a very high target of an alpha of 0.10 and power of 0.80. The magnitudes of change that could be detected given these standards will depend on sampling effort, a given indicator, and on the wetland-to-wetland variability. Table 1-2 shows how sample size has to change to provide a given degree of power to detect various effect sizes of >20%, in agreement with other monitoring approaches. For some indicators and measures, it is possible to significantly increase power with acceptable increases in cost. For the initial set of protocols, *a priori* power analyses will be used to determine the approximate sample size needed to an effect size of 20%. Given the specification of alpha, desired power, and effect size, combined with information on the variance of the response variable in question (obtained from available data or comparable analogous data, where available), it is possible to calculate the sample size required to achieve these results. Statistical power analysis (Gerrodette 1993), is the typical approach to estimating sampling sizes for monitoring population trends.

From Table 1-2 below, Minimum sample size necessary to be 80% certain (i.e., power = 0.8) that a specified true difference (=effect size) between two groups will be found to be significant (p<0.05). **V** is the coefficient of variation of the 'reference' group ( $V = ([\text{standard deviation}/\text{mean}] \times 100)$ ). The variances of the two groups are assumed to be equal. The dotted line transecting the table indicates the minimum differences likely to be designated significant with triplicate sampling. The heavy line transecting the table indicates the minimum differences likely to be designated significant if 12 wetlands are sampled per group. Entries for which  $n > 100$  may be overestimated by approximately 2%. Three replicates are considered to be an absolute minimum sample size if outliers are to be identified. Entries were determined from power formulae given by Sokal and Rohlf (1981).

**Table 1-2.** Tabular organization of response values, averages and slopes (Sokal , et al. 1981).

V (%)	Effect Size (True difference between two means (%))									
	100	90	80	70	60	50	40	30	20	10
100	14	18	22	29	39	54	67	99	201	801
90	12	14	18	23	32	41	55	73	163	649
80	9	11	14	18	25	36	47	67	129	513
70	7	9	11	14	19	28	40	55	99	393
60	5	7	8	11	14	20	32	45	81	289
50	4	5	6	7	10	14	22	36	67	201
40	3	3	4	5	7	9	14	25	52	129
30	3	3	3	3	4	5	8	14	32	73
20	3	3	3	3	3	3	4	7	14	33
10	3	3	3	3	3	3	3	3	4	14

One may use existing software programs (e.g. Gerrodette 1993) and simple equations (Elzinga et al. 2001, Manly 2001) for approximating sample sizes. In the context of Consortium, the development of a monitoring and sampling program is best accomplished in a workshop setting involving key stakeholders and technical personnel and including the input of researchers and statisticians who can explain the theory and trade-offs to participants in straightforward terms. Such individuals could operate statistical software at the workshops to demonstrate the power and effectiveness of various sampling scenarios through simulations.

## Recommendations

For now, the Consortium will recommend a target number of wetlands stratified by Great Lake and ecoregion. Ideally, a minimum of 12 wetlands from each ecoregion on each Great Lake should be randomly selected for study. The recommendation of 12 is simply a starting point until power analyses using means and variances of IBI scores from the GLCWC pilot studies can be performed. The number of sites may then be adjusted accordingly. Twelve sites were chosen based on best professional judgment and experience using the Burton et al. (1999) IBI and protocol

on many Lake Huron and Lake Michigan wetlands over the past 10 years. If less than 12 sites are present within a given ecoregion, then all of the sites should be sampled. Whenever possible, more than 12 sites per ecoregion per Great Lake and connecting channel should be sampled.

Four panels of 12 wetlands were randomly selected (without replacement) from the population in each ecoregion. Each panel should be visited every four years, with a different panel starting in each of the first four years. From each panel is a randomly selected subset of three wetlands (subpanel) that will be visited two consecutive years. These subpanels were selected so they do not overlap, which effectively partitions each panel into four subpanels. Design 4 in Fig. 1-3 above shows the pattern of sampling.

## References

- Busch, D. E., and D. T. Trexler, editors. 2003. *Monitoring Ecosystems*. Island Press, Washington.
- Elzinga, C. L., D. W. Salzer, J. W. Willoughby, and J. P. Gibbs. 2001. *Monitoring plant and animal populations*. Blackwell Science, Inc.
- Gerrodette, T. 1993. TRENDS: Software for a power analysis of linear regression. *The Wildlife Society Bulletin* **21**:515-516.
- Lessler, J. T., and W. D. Kalsbeek. 1992. *Nonsampling Errors in Surveys*. John Wiley and Sons, Inc., New York.
- Manly, B. F. J. 2001. *Statistics for Environmental Science and Management*. Chapman & Hall/CRC, Boca Raton, Florida.
- Mapstone, B. 1995. Scalable decision rules for environmental impact studies: Effect size, type I & type II errors. *Ecological Applications* **5**:401-410.
- McDonald, T. L. 2003. Review of Environmental Monitoring Methods: Survey Designs. *Environmental Monitoring and Assessment* **85**:277-292.
- McDonald, T. L., and P. H. Geissler. 2004. Systematic and stratified sampling designs in long-term ecological monitoring studies. <http://science.nature.nps.gov/im/monitor/docs/SampleDesigns.doc>.
- Roback, P. J., and R. A. Askins. 2005. Judicious use of multiple hypothesis tests. *Conservation Biology* **19**:261-267.
- Scheaffer, R. L., W. Mendenhall, and L. Ott. 1990. *Elementary Survey Sampling*, 4<sup>th</sup> edition. PWS-Kent, Boston.
- Skalski, J. R. 1990. A design for long-term monitoring. *Journal of Environmental Management* **30**:139-144.
- Sokal, R. R., and F. J. Rohlf. 1995. *Biometry*, 3rd edition. W.H. Freeman and Company, New York.
- Sokal, R.R. and F.J. Rohlf. 1981. *Biometry*, 2<sup>nd</sup> Ed.. Freeman Publ., San Francisco, CA
- Steidl, R. J., J. P. Hayes, and E. Schaubert. 1997. Statistical power analysis in wildlife research. *Journal of Wildlife Management* **61**:270-279.
- Stevens Jr., D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. *Journal of American Statistical Association* **99**:262-278.
- Urquhart, N. S., and T. M. Kincaid. 1999. Trend detection in repeated surveys of ecological responses. *Journal of Agricultural, Biological, and Environmental Statistics* **4**:404-414.
- Urquhart, N. S., S. G. Paulsen, and D. P. Larsen. 1998. Monitoring for policy relevant regional trends over time. *Ecological Applications* **8**:246-257.

*This chapter is adapted from Sierra Nevada Network (SIEN) monitoring plan:  
Draft SIEN Phase III Report, December 2006 [http://www.nps.gov/archive/seki/snn/snn\\_index.htm](http://www.nps.gov/archive/seki/snn/snn_index.htm).*