Guidance for the Design of Sampling Schemes for Inventory and Monitoring of Biological Resources in National Parks

Steven G. Fancy
National Park Service Inventory and Monitoring Program
March 24, 2000 version

Purpose:

The National Park Service (NPS) Inventory and Monitoring Program provides funding and technical support to approximately 269 national park units to conduct basic inventories and long-term monitoring of park natural resources. In FY2000, parks were organized into 32 biome-based networks that share similar natural resources and management issues, and funding was provided to each network to develop a detailed study plan for conducting inventories of vertebrates and vascular plants in parks. Beginning in FY2001, networks of parks will begin receiving funding to implement long-term monitoring of "vital signs" of park ecosystem health. These inventory and monitoring efforts are part of the Natural Resource Challenge, an initiative to revitalize and expand the natural resource program within the National Park Service and improve park management through greater reliance on scientific knowledge.

The NPS recognizes the importance of collecting data in a scientifically credible manner so that they can be used to address current and future management issues. All parks and their contractors and cooperators should use certain "good sampling practices" so that data meet the purpose for which they were collected and withstand scrutiny by critics. Sample sizes will almost always be limited by shortages of funding and personnel, and it is critical to be able to make inferences to larger areas from data collected at relatively few sampling locations.

Beginning in Spring 2000, most of the 32 networks will hold workshops to develop a detailed study plan for inventories of vertebrates and vascular plants, and a few networks will also begin developing plans for long-term monitoring of vital signs. All of the plans for inventorying and monitoring in national parks will contain a statistical sampling design developed by the network parks and their cooperators and contractors. To provide some guidance to these 32 or more separate design efforts, and to help reduce duplication of effort and inconsistencies among networks, the Servicewide I&M Program convened a panel of sampling experts on February 23-24, 2000 to develop some recommendations for designing a sampling program. The six panel members were Drs. Paul Geissler, Douglas Johnson, and John Sauer of USGS/BRD; Drs. Lyman McDonald and Trent McDonald of West, Inc.; and Dr. Anthony Olsen of the US Environmental Protection Agency. All six panelists have considerable experience and have published numerous scientific papers on sampling design, and all six are known for their practical, real-world understanding of problems such as difficult access and high costs of sampling in many field situations. To provide focal points for discussion of various sampling approaches, four parks were chosen to present specific objectives and management questions, and to present some of the logistical problems for their park. The four park scenarios that were discussed were for Gates of the Arctic NP in Alaska, Olympic and North Cascades NP in Washington, and Great Smoky Mountains NP in Tennessee and North Carolina. A list of workshop participants and a summary of the park scenarios that were discussed is presented in Appendix A.

What is the best sampling design?

The "best" sampling design depends on the question being asked and the attribute being measured, and several equally valid designs may be possible even in cases where a very specific question with only one variable of interest is being asked. If the purpose of the survey is to determine the distribution or abundance of some rare species, the most efficient design may be very different from one intended to provide the same information for another species with a different life history or spatial distribution. Sampling design issues for monitoring continue to be an area of active research (e.g., see special issues on measuring change over time in Ecological Applications [1998, Vol. 8, No. 2] and the Journal of Agricultural, Biological and Environmental Statistics [1999, Vol. 4, No. 4]). Any statistician who is asked to design a survey will begin by asking for specific information regarding the goals of the survey. For the purposes of this workshop we asked the panel to recommend designs for an omnibus inventory or monitoring program with multiple goals because most parks are interested in knowing the status or trend of a number of various natural resources or "target populations" rather than answering a single, focused question. When discussing inventories, we focused on the biological inventories of vertebrates and vascular plants that the 32 networks are developing this year. However, it was recognized that even if the objective is to conduct a one-time inventory of the status of various species (e.g., presence/absence, distribution or abundance), someone may later decide to revisit some of the inventory sampling sites to look at long-term trends in the variable of interest, and therefore a sampling design that allows managers and scientists more flexibility in the future is better than one that is restrictive.

The panel was repeatedly reminded that access and cost have a major effect on how surveys are conducted, and some park areas are simply impossible to sample because of terrain, impenetrable vegetation, safety issues, or logistic considerations. After two days of discussing a number of different park sampling scenarios, and two weeks of trading email messages discussing the pros and cons of various approaches, it was decided to present several different designs and a discussion of some good sampling practices for park managers to keep in mind as they develop their sampling program. The companion document examples.doc presents details for the designs summarized below.

Working Definitions (for the purposes of this report):

- An **Inventory** is an extensive point-in-time effort to determine location or condition of a resource. In the case of biological inventories, the interest is in documenting at least 90% of the vertebrates and vascular plants that occur in the park, and to obtain distribution and abundance information for a subset of those species that are of special interest to park managers. It should be recognized that even if the inventory is designed as a one-time "snapshot" of the distribution or abundance of a resource, someone may later decide to repeat the inventory to look at change, and thus the inventory may become the initial set of observations for a monitoring program. Inventories focus on determining spatial occurrence and distribution, whereas monitoring focuses on detecting temporal change and usually involves a subset of the sites that were included in the inventory.
- **Monitoring** differs from inventory in adding the dimension of time, and the general

purpose of monitoring is to detect changes or trends in a resource. Elzinga et al. (1998) defined monitoring as "The collection and analysis of repeated observations or measurements to evaluate changes in condition and progress toward meeting a management objective". Detection of a change or trend may trigger a management action, or it may generate a new line of inquiry. Monitoring is often done by sampling the same sites over time, and these sites may be a subset of the sites sampled for the initial inventory.

- **Research** has the objective of understanding ecological processes and in some cases determining the cause of changes observed by monitoring. That understanding is needed for determining the appropriate management response to threats. The NPS monitoring program also includes a research component to design sampling protocols for various types of park resources at different locations and spatial scales. Differences between research and monitoring are listed on the web site "Monitoring Natural Resources in National Parks" at http://science.nature.nps.gov/im/monitor.

<u>Things to keep in mind as you develop a sampling design for inventory and monitoring</u>:

1. Some sort of probability sample should always be taken. Conceptually, the target population (usually the entire park) is divided into sampling units such that every point in the park is included in a sampling unit, but not in more than one. The sampling design is used to select a probabilistic sample of the sampling units. As a result, statistical estimates of population attributes can be produced with an estimate of their reliability. Probability samples occur when each unit in the target population has a known, non-zero probability of being included in the sample, and always include a random component (such as a systematic sample with a random start). The credibility of data that are not collected using these principles is easily undermined.

2. Statistical, design based inferences can only be made to areas that have a chance of being included in the sample. If study plots are chosen to be close to roads, design-based inferences can only be made to areas near roads. Since the NPS's mission is to protect resources in the entire park, sampling should be designed so that robust inferences can be made to the entire park and not some easily accessible portion of it. Model based inferences and professional judgement can be used to infer values in portions of a park that had no probability of being included in the sample. However, accuracy of model based inferences and professional judgement is only as good as the model and the decision making process of the individual providing the professional judgement, and models and judgement-based information can often be easily discredited by critics. Areas of the park that are too inaccessible or unsafe to sample can be simply excluded from the program, but then no inference can be made about resources in these areas.

3. Judgement sampling, using "representative" sites selected by experts, should be avoided. If there is no controversy, judgement sampling sounds good at the beginning, but "representative" sites may come back to haunt you in the future because they are easily discredited by critics and may produce biased, unreliable information.

4. Panel members supported a general framework of first spreading samples out over the entire park or target population, and then increasing the sampling intensity in areas of special interest. Simple random sampling is not recommended because you may select a sample that is not spatially balanced, and because we are often interested in species or other park resources that occur in limited areas and we want to make sure we include adequate samples in those areas. Samples can be spread out over the area of interest by using some sort of grid or cell design or a tesselation procedure (e.g., Stevens 1997). Within this overall design, areas of special interest such as rare habitats can be sampled with higher frequencies using either stratification or the more general approach of defining the cells corresponding to the areas of interest and varying their selection probabilities (the unequal selection probability approach). In either case, the areas are then sampled disproportionate to their availability so that adequate samples are taken from each. This unequal sampling probability approach accomplishes most of the advantages of stratification, but avoids some of the problems of stratification that are mentioned below. An overall framework based on this design that allows for including site-specific studies and legacy data is presented below.

5. The watershed approach, where a park first randomly selects one or more watersheds as a primary unit, does more intensive sampling within that watershed, and then makes inferences to the rest of the watersheds, was not recommended by any of the panel members. This approach does not allow for inference to the entire park.

6. A design based on stratification of the park by "habitats" derived from vegetation maps is not recommended because stratum boundaries will change over time, and unless you fix the stratum boundaries forever there will be problems in the future with data analysis and incorporating new information into the design. A vegetation map is a model based on remote sensing data and data collected on the ground at a series of plots; the map boundaries will change as the classification models change or as additional ground-truthing data becomes available. Using these units to define strata will limit (and greatly complicate) long-term uses of the data by restricting future park managers' abilities to include new information into the sampling framework.

7. It is legitimate, and better, to delineate areas of special interest such as riparian or alpine areas based on physical characteristics such as terrain, and use these to judiciously define either strata or areas to sample with higher probability.

8. Permanent plots that are revisited over time are recommended for monitoring, because the objective is to detect changes over time. Revisiting the same plots removes plot to plot differences from the change estimates, increasing the precision.

9. An important step in developing a sampling design for a park is determining the sample size needed to significantly reduce the uncertainty of guessing about the status or trend of a resource and consequently reduce the costs of stewardship. Taking too few samples may increase the costs of stewardship and put resources at risk because

important changes are missed or detected too late for management to be effective, whereas taking too many samples will waste time and resources. The sample size that is needed to meet a sampling objective is largely a function of the effect size, which is the amount of change in the resource from one point in time to the next that the manager seeks to detect, and the variability of the resource across space and time. For a statistician to be able to estimate the sample size needed for a particular program, the park manager needs to be able to specify how much change they need to be able to detect, and with what certainty, to affect their management strategies and practices or to confront and mitigate threats to the park in legal and political arenas. For planning purposes, sample size calculations such as those in Chapter 4 of Cochran (1977) or other statistical texts can be used to obtain a rough idea of the magnitude of the sample needed to produce a confidence interval of a specified width for a particular variable. If a statistical comparison is to be made between two samples, a "rule of thumb" minimum sample size is 6 measurements in each sample. It is useful to think about sampling over space when allocating samples.

10. Be sensitive to spatial integrity of the sample! These data will be used for many purposes, and an initial view of the sample on a map will help to clarify the use and limitations of the sample. When a sample is allocated, it is probably a good idea to display the sample on a GIS to ensure that adequate coverage occurs for areas of interest.

11. When repeated measurements of the same site are made to determine trend, remember that the precision will increase as the number of years of sampling increases. [The sample size of comparisons is usually the number of plots, which will not increase. However, the precision will increase because the means for each plot become less variable ($var=s^2/n.$)]. There may be considerable intra-year variability in a measure because of small sample sizes, sampling errors and spatial variation, all of which increase needed sample sizes, and yet you may still be able to identify a trend as you increase the number of years of data.

12. When designing a monitoring program, remember that it is not necessary to visit all of the selected sites every year. Sampling designs exist that allow for increased spatial coverage though "rotating panel" designs, where each site is sampled every five years, for example, but five times as many sites can be sampled because only 1/5 of them are visited each year. Data from a complex rotating panel design with multiple strata can be difficult, so data analysis needs to be considered when the design is put together.

13. Collocation of samples is recommended to allow comparisons among components. For example, in the same stream segments you might sample water quality, aquatic macroinvertebrates, amphibians, and fish. Another example would be to monitor changes in vegetation, birds, mammals, and certain invertebrates at sites that are close to each other.

Important Note: The actual process of selecting sampling sites as described in the framework below and the examples in examples.doc is not easy, and in most cases will require people with advanced GIS, statistical, and ecological expertise. The good news is that you only have to do it once, and the national and regional offices in conjunction with experts from other agencies will provide technical support – you don't have to figure this all out yourself! The important step for park managers and resource specialists is to determine the goals and objectives of the sampling efforts, because without clear objectives, little progress can be made in designing efficient inventory or monitoring programs.

An overall framework for sampling in a park:

We recommend that parks develop an overall sampling design that is simple and yet provides flexibility for incorporating studies with different levels of sampling intensity both now and in the future. Even if current objectives and funding levels allow only a small number of sites to be visited, we recommend at the outset that you develop a design that includes the entire park (or could be expanded beyond park boundaries if desired). The following framework is flexible enough to incorporate different levels of sampling now and in the future.

Begin with a **Base grid** that spreads samples out evenly over the entire park or target population. One way to do this is to define a very dense grid and either use the center of each grid cell as the sampling point or randomly select a point within each cell, such that points are spaced 50 to 150 meters apart over the entire park or area of interest. The starting point for this grid should be randomly selected. For a large park, this dense grid could potentially contain hundreds of thousands of points. From this large number of potential sampling points, the desired number of actual sampling points for different purposes can be selected now and in the future. If certain areas of the park cannot be sampled for some reason (e.g., unsafe or inaccessible terrain, or areas beyond the range of a helicopter flight), those areas should be blocked out of the target population and design-based inferences cannot say anything about those areas. **Note that this grid is not a systematic sample. Instead, it is a sampling framework, or the list of all possible samples from which the actual sample will be selected.**

Another possible approach is to use one of the randomly-placed grids developed by other agencies to obtain regional and national data as the basis for the park's base grid. For example, the USDA Forest Service administers the Forest Inventory and Analysis (FIA) program, where remote sensing data is taken on a 1-km grid and ground plots are sampled at 5-km intervals. Recent legislation requires the FIA program to now include forested NPS lands and many parks will be sampled by the FIA program during the next few years, with much of the cost covered by the Forest Service. The FIA base grid was randomly located, and parks could use this grid as a means of collocating other sampling in the park, or could intensify this grid in areas of special interest as described below.

None of the panel members knew of a single sampling design that would work for both terrestrial and aquatic systems, so we recommend using different approaches to selecting stream segments or lakes or ponds to sample than for selecting other sites. An example of how to select sites for aquatic studies is given in examples.doc.

Within the base grid, some areas will require more intensive sampling (areas of special interest). These areas can be accommodated in the sampling framework by changing the chance of selecting the areas. **Additional Samples** are added using unequal probability sampling as needed to incorporate different sampling objectives and to assure adequate sample sizes for all communities of interest. There are many different approaches to this intensified sampling (stratification or variable selection probabilities are the two discussed here), and examples are presented in examples.doc.

Specialized studies that only allow inferences to be made to specific areas can also be incorporated into the overall sampling framework. For example, certain sites may need to be near a source of electricity because of the equipment used, or a study requiring the use of bulky traps or equipment may restrict the study to areas near roads. In these cases, inferences can only be made to the areas that could have potentially been selected to sample. Another special case is for legacy data: plots that were selected in the past, perhaps as judgment samples that were supposed to be representative of some area, could be defined as separate strata within the overall base grid. It may be possible to include these data in certain analyses, but sampling at these sites should not define the monitoring sample design. If sampling at the sites is discontinued it should not affect the overall sampling design.

Examples of some different designs

Detailed examples of some different designs are available in the companion document examples.doc, "Examples of Park Sampling Designs". The detailed examples were put in a separate document because some people might be intimidated and perhaps not read these sampling guidelines if it seemed too difficult or lengthy. For those that will be involved in developing the actual sampling design for a park, you can download a copy of the file examples.doc from http://science.nature.nps.gov/im/monitor/vsm.htm#design. The following includes a brief discussion of several designs that might be used to sample a park.

Design #1 is based on the framework above and consists of three levels of sampling: (1) an overall extensive framework is established using some sort of grid-cell approach to spread samples out over the entire park or target population; (2) in areas of special interest (e.g., riparian areas, rare habitats, or areas with special management interest) the grid is intensified so that certain areas are sampled disproportionate to their availability; and (3) special studies occur at specific locations that are selected separately from the overall framework (e.g., sampling that requires electricity to run equipment, or requires hauling in lots of traps or special equipment and therefore must be near a road). If certain areas of the park cannot be sampled (e.g., steep slopes that are unsafe to traverse), these areas are simply blocked out and no inferences can be made to them. Details of this design with an example for aquatic and terrestrial systems are presented in examples.doc.

Design #2 also begins with an extensive grid of potential sampling locations that covers the entire park or area of interest. The grid is then intensified in areas of special interest to park managers, such as rare habitats or areas with high visitor use. To increase the efficiency (precision) of the design, the park is stratified into two zones based on the cost or difficulty of sampling within them, and optimum allocation of samples among strata is done based on the cost of sampling a point in each stratum. A detailed example of this approach, including calculations for allocating samples, is presented in examples.doc.

Design #3 defines two or more strata based on the difficulty or cost of access. This design is not as good as the first two and is not recommended because it isn't as flexible. In this design, a buffer of some fixed distance is placed around roads and trails, and most of the samples are taken from this "easy access" stratum. At least 10% of the samples are placed in the "difficult access" stratum to test whether there is a near-trail bias in the measures of interest and to see whether you've missed anything (new species, etc.). This design can be approached as an iterative process whereby additional samples could be taken if the first set of "difficult access" samples is too small to adequately test for bias based on variability within the data sets. It is important that the original stratification based on the difficulty or cost of access must be fixed forever or else analysis and interpretation of the data will be compromised. For example, even though additional trails are added to the park, or it becomes easier or cheaper in the future to sample further away from trails, the original stratification should always be used.

A major problem with this approach is that areas near roads and trails may offer a very biased sample of the entire park (e.g., they often follow ridges or streams), and there may be considerable variation within the "difficult access" sample such that large numbers of samples must be taken to adequately test whether a bias exists. Many parks may find by doing a cost analysis that this design does not offer as great an advantage as it might first seem in terms of number of samples that can be taken. Nevertheless, there may be some park situations where the cost or difficulty of sampling in certain inaccessible areas is so overwhelming that this design must be used.

Design #4 is a stratified random design that is often used, but which has some pitfalls you need to be aware of. In this design, the park is first stratified based on vegetation boundaries or physical attributes such as elevation, slope and aspect, and a certain number of samples are randomly chosen (or laid out systematically with a random start) within each stratum. A stratified design is a special case of an unequal probability design, where the probabilities of selection are constant within strata, and it is possible to analyze a stratified design using unequal probability methods. For that reason, Paul Geissler suggests starting out with the more familiar stratified design, unless there is a compelling reason to use a more flexible design. However, Lyman McDonald pointed to a number of potential problems with stratification that must be taken into account:

- Strata boundaries may change over time (e.g., a pool may become a run after a flood, or vegetation map boundaries change as classification models change or as additional ground-truthing data becomes available).
- Stratification can be optimized for only one variable at a time. If multiple measurements are taken at each site, or multiple species are sampled, stratification for one measure may do a very poor job for another measure.
- If everyone stratifies separately, collocation of sites is not possible, and correlations over space cannot be easily made or design based.
- Sites are often misclassified in the office and must be reclassified into the proper stratum after site visits.
- Stratification and misclassification of sites leads to unequally weighted data which

must be released to the public and critics who may not understand or properly use the weights when considering the data.

If stratification is used to develop a design for a long-term program, it is important to recognize that the strata boundaries must be fixed forever, which argues that attributes such as terrain should be used to develop the strata rather than vegetation or some other attribute that will change. However, strata boundaries can be changed if one is willing to switch to the unequal probability approach to select new samples and to analyze the data.

Literature Cited

Cochran, W. G. 1977. Sampling techniques. 3rd ed. John Wiley, New York. 428 pp.

Elzinga, Caryl L., D. W. Salzer, and J. W. Willoughby. 1998. Measuring and monitoring plant populations. BLM Tech. Reference 1730-1. BLM/RS/ST-98/005+1730. Available from Bureau of Land Management, National Business Center BC-650B, PO Box 25047, Denver, CO 80225-0047.

Stevens, D. L., Jr. 1997. Variable density grid-based sampling designs for continuous spatial populations. Environmetrics 8:167-195.

Thompson, S. K., G. A. F. Seber, and G. F. Seber. 1996. Adaptive sampling. John Wiley & Sons, NY. 288 pp.

Appendix A
Contact Information for Workshop Participants

Panel Members:

Dr. Paul Geissler
USGS Biological Resources Div.
12100 Beech Forest Road
Laurel, MD 20708
(301) 497-5780
paul_geissler@usgs.gov

Dr. Douglas H. Johnson
USGS Biological Resources Div.
8711 37th Street SE
Jamestown, ND 58401-7317
(701) 253-5539
douglas_h_johnson@usgs.gov

Dr. Lyman MacDonald
West, Inc.
2003 Central Avenue
Cheyenne, WY 82001
(307) 634-1756
lmcdonald@west-inc.com

Dr. Trent MacDonald
West, Inc.
2003 Central Avenue
Cheyenne, WY 82001
(307) 634-1756
tmcdonald@west-inc.com

Dr. Anthony R. Olsen
US EPA Western Ecology Division
200 SW 35th Street
Corvallis, OR 97333
(541) 754-4790
tolsen@mail.cor.epa.gov

Dr. John R. Sauer
USGS Patuxent Wildlife Res. Center
11510 American Holly Drive
Laurel, MD 20708
(301) 497-5662
john_r_sauer@usgs.gov

Participants providing examples from National Parks:

Name	Park Example
Reed Glesne, NPS	North Cascades NPS Complex
Dr. Kurt Jenkins, USGS-BRD	Olympic NP
Dr. Mike Jenkins, NPS	Great Smoky Mountains NP
Dr. Tom Stohlgren, USGS-BRD	Rocky Mountain NP
Sara Wesser, NPS	Gates of the Arctic NP
Dr. Andrea Woodward, USGS-BRD	Olympic NP

Organizers:

Dr. Steven Fancy	NPS Inventory and Monitoring Program	(970) 225-3571
Mark Wotawa	NPS Inventory and Monitoring Program	