WETLANDS, Vol. 27, No. 3, September 2007, pp. 515–523

SAMPLE DESIGN, EXECUTION, AND ANALYSIS FOR WETLAND ASSESSMENT

Don L. Stevens, Jr. and Susan F. Jensen
Department of Statistics
Oregon State University
Corvallis, Oregon, USA 97331
E-mail: stevens@stat.oregonstate.edu

Abstract: Probability-based sample surveys are increasingly being used to assess natural resource condition, yet many of the techniques commonly used in survey sampling are not well known to environmental scientists. Here we discuss several techniques from survey sample methodology that can substantially increase the efficiency of a sample (lower the variance for the same sample size), make the survey more cost-effective, or adjust the analysis to accommodate difficulties in implementation, such as lack of access to all sample sites. The techniques are illustrated with two surveys that were designed to assess wetland condition: one in the Juniata watershed in Pennsylvania, USA, and one in the Nanticoke watershed in Maryland and Delaware, USA.

Key Words: access denial, missing data, multiple frames, non-response, post-stratification, probabilistic sampling, spatially balanced environmental samples

INTRODUCTION

Probability-based sample surveys are increasingly being used to assess natural resource condition. Wardrop et al. (2007) set out the case for using probability surveys to assess wetland condition. Much of existing survey methodology has been developed in the context of describing the characteristics of finite population, especially human populations. For the most part, this methodology can be transferred in a straightforward manner to sampling natural resources. Even though the easily transferable methodology is immediately applicable, very useful, and generally known among survey samplers, it may not be familiar to most environmental scientists (and in fact, to many statisticians!). Also, sampling natural resources, particularly wetlands, presents some challenges that are not often encountered in the usual sample survey application.

In this article, we discuss several techniques of survey sampling that have potential for wide application in sampling wetlands. The first, sampling an infinite continuum, presents a perspective that differs from the finite-population approach adopted in most sampling textbooks. The second, spatially balanced sampling, describes some techniques for selecting samples in continuous, two-dimensional space that capitalize on the spatial correlation that is likely present in the response. The third technique, the use of multiple frames, provides a partial answer to the difficulty of developing a comprehensive frame from which to select a sample. The fourth, two-phase sampling provides an additional technique to use when developing frame information is difficult, time-consuming, or expensive. The final technique, post-stratification, provides some remedy to the problem of non-response resulting from lack of access to all of the sites in the sample draw. In some instances, post-stratification can also be used to correct for a poor random selection by ‘‘re-balancing’’ the sample after the fact.

We illustrate the application of these techniques using the samples drawn to assess the Juniata and Nanticoke watersheds, which are discussed in other papers in this issue. (Wardrop et al. 2007, Whigham et al. 2007). Generally, the wetland resource was described and its condition evaluated in the Upper Juniata watershed in Pennsylvania, USA, using landscape-level, rapid assessment, and intensive (hydrogeomorphic) methodologies; the wetland resource in the Nanticoke River watershed in Maryland and Pennsylvania was also evaluated using a hydrogeomorphic approach. Important resources for these studies included the Penn State Cooperative Wetlands Center (CWC) and the National Wetland Inventory (NWI). These case studies assessed the functional performance of the wetlands at the watershed scale, while ensuring a sample without geographic bias. They also demonstrated tools that can be used in development of strategies for wetland restoration at the watershed scale.

SAMPLING AN INFINITE CONTINUUM WITH A MAP FRAME

Probability surveys have several characteristics that distinguish them from other sample-selection techniques. The most important characteristic is that sample units are selected using a technique that explicitly incorporates randomization. The randomization has two functions: it guards against selection bias, intentional or otherwise, and it provides an objective, inferential basis for extrapolating from the sample to the population level. Because the explicit randomization requires that the inclusion probability be specified for every population unit, a comprehensive definition of the population is required. The process of defining the population is often a valuable exercise that helps focus the survey. The population definition needs to be expressed as a *population frame*, which is a representation of the target population in a form that enables sample selection. Common frames for finite populations are simply lists of every population unit. For environmental resources, the frame is often a map or a GIS representation of a map. For map frames, the population can be conceptualized as an infinite continuum. In this case, a sample can be selected by picking points, (e.g., map coordinates).

The analysis of a probability survey is often called *design-based* because the validity of the population inference rests on the design rather than on an assumed statistical model. The randomness is explicitly included in the sample-selection process and forms the basis for estimating population characteristics. The key quantity in the estimation is the inclusion probability for a population unit, which is the probability that that unit is included in the sample. It must be positive for every unit. In the case of a continuum, the inclusion probability is defined by an inclusion density, usually denoted by $\pi(s)$. In contrast to a probability density, the inclusion density has units. For example, an inclusion density for a point sample from a map might have units of (number of sample points)/ km^2. In the case of a finite population, the inclusion probability sums to the sample size; in the continuous case, the integral of the inclusion density over the target domain gives the sample size. The importance of the inclusion probability for a sample element is that its reciprocal is a measure of the portion of the population represented by that element. Thus, for example, in a simple random sample of size n from a finite population with N total elements, the inclusion probability for each sample element is n/N, and each sample element represents N/n population elements. If a simple random sample of n sites were selected in a wetland with area A km^2, then the inclusion density would be $\pi(s) = n/A$, and each site would represent A/n km^2 of wetland.

The basic tool used in our analyses is the Horvitz-Thompson or π-weighted estimator (Horvitz and Thompson 1952, Thompson 2002). The continuous version of this estimator is given in Cordy (1993) or Stevens (1997). The concept of the π-weighted estimator is that estimates of totals are obtained by weighting individual observations with a weight inversely proportional to their inclusion probability. In the Juniata and Nanticoke examples described following, the samples were point samples from the target wetland area, so that the inclusion probability density is the point density with units of (number of sample points) / (*ha* of target wetland area). The weight attached to each observation is then the wetland area represented by that observation.

Let n be the number of sample plots, let z_i be the response for the i^{th} sample plot, and let π_i be the inclusion probability (or density) evaluated at i^{th} sample point. Note that z_i could be a numeric score (e.g., per cent forested land cover) or a binary classification, for example,

$$z_i = \begin{cases} 1, \textit{ if } i^{th} \textit{ plot} \text{ is } \textit{in degraded condition} \\ 0, \textit{ if } i^{th} \textit{ plot} \text{ is not } \textit{in degraded condition} \end{cases}.$$

The Horvitz-Thompson estimate of the total of z is given by $\hat{z}_T = \sum_i^n \frac{z_i}{\pi_i}$ and the estimate of the mean value by $\bar{z} = \frac{\hat{z}_T}{A}$, where A, the population size, is the total area of the target population. These formulas are the same for both finite and infinite populations. Note that in the case of z_i being a binary classification, $\bar{z}$ estimates the proportion of the resource in the condition class, e.g., the proportion of the watershed in degraded condition.

An alternative estimator of the mean value uses the estimated population size $\hat{A} = \sum_1^n \frac{1}{\pi_i}$ as a divisor in place of A. In some circumstances, use of the estimated population size in place of a known population size can lead to a more precise estimate of the mean because of positive covariance between $\bar{z}$ and $\hat{A}$. If the size of the target population is not known, for example, the imperfect frame case described following, then the alternative estimator must be used. Also, if some plots were not accessible, say because access permission was not obtained, then an estimate of the average condition of the *accessible* wetlands is $\bar{z} = \frac{\hat{z}_T}{\hat{A}}$, where both $\bar{z}$ and $\hat{A}$ are computed using only those sites for which a response was obtained.

SPATIALLY BALANCED SAMPLING

In sample survey terminology, an auxiliary variable is one that is known (or knowable) for every element in the population. It is often possible to make use of an auxiliary variable to improve the precision of sample, especially if the auxiliary variable, say X, is correlated with the unknown target response variable Z. A sample of Z is *balanced* over an auxiliary variable X if the x-values (which are known beforehand) are chosen so that the sample mean of the x-values is exactly equal to the true population mean of X (Yates 1981). A stricter version of balance was suggested by Royall and Herson (1973), who required that the first several sample moments of the x-sample match the population moments. The intuition behind balancing is that the auxiliary variable is correlated with the unknown response to be assessed. By balancing over the auxiliary variable, we hope to get approximate balance over the unknown response (and hence a more precise estimate than simple random sampling (SRS) would give, as discussed following).

Kott (1986) noted that an option intermediate between random sampling and strict balancing can be obtained by splitting the range of X into n quantiles and picking one sample element in each quantile. Although this option does not achieve balance in the strict sense of having sample moments match population moments, it does guarantee that the sample distribution function of X will be close to the true distribution function for every sample draw.

For a response variable distributed over some region R, the distribution function of the response gives the proportion of the area of R, where the response is less than some specified value. For example, if the response is the wetland condition score, then the distribution function evaluated at z is the proportion of the area of R where the wetland condition score takes on a value of z or less. Thus, the distribution function for Z is $F_z(z) = \dfrac{\text{Area of region with Z} \leq z}{\text{Total area}}$.

We can use space itself as an auxiliary variable, and balance our sample over space in the sense of Kott (1986), if the response is believed to have spatial structure. A spatially balanced sample, then, is one that is more or less evenly and regularly spread out over the spatial extent of the target population. A criterion for spatial balance is the closeness of the sample spatial moments (the center of gravity and the second moment about the center of gravity) to the domain spatial moments. If the target domain is a two-dimensional continuum, (e.g., a large lake or an estuary), then the extreme spatially balanced sample locates points at the nodes of a regular grid. A less extreme, balanced design might use some form of spatial stratification to spread out the sample.

Techniques for Obtaining Spatially Balanced Samples

The precept that a spatially balanced sample is more efficient than SRS for sampling for environmental populations is well established. Accordingly, there are numerous methods for inducing some spatial regularity in a random sample. One strategy for sampling ecological resources is to develop an area sampling design based upon a single area sampling frame. In such a frame, the entire region to be monitored (e.g., the conterminous United States) is partitioned into a set of mutually exclusive and exhaustive areas. These areas are frequently designated primary sampling units (PSUs). The partition can be based on arbitrary geometric figures or on some characteristic of the landscape, such as the USGS hydrologic cataloging units. Commonly, PSUs are chosen with boundaries that are easily discernible in the field, such as permanent roads, railroads, or rivers. A sample is selected from these PSUs according to a probability-based protocol, such as selecting a PSU with a probability proportional to its size. Usually, some restriction is imposed on the sample selection to ensure spatial dispersion of the sample. The ecological resources occurring in each sample PSU are identified, characterized, and measured.

The Soil Conservation Service (SCS) has used an area frame in several national resource surveys (Goebel and Schmude 1982). The 1958 Conservation Needs Inventory (CNI) used a frame based on 100-acre ~40 ha) squares of land in the northeastern states and partitions of public land survey sections (approximately 640 acres, 160 ha) in the rest of the country. The 1967 CNI treated the 1958 sample areas as PSUs and subsampled within them at specific points. The 1977 National Resource Inventory (NRI) used the 1958 CNI area frame and a two-stage sample. The 1982 NRI also used an area frame based on public land-survey sections in most cases. Louisiana was sampled using a design based on the Universal Transverse Mercator (UTM) grid system. PSUs for the northeastern states were defined by areas of 20 seconds of latitude by 30 seconds of longitude. The sample was stratified to control spatial coverage, to achieve increased resource homogeneity within strata, and to allow variation of sampling intensity.

A variation on the basic area frame design is a multi-stage cluster design. In this scheme, an area selected at a given stage is further split into subareas, and a sample is selected from the subareas. Complete characterization and measurement take place only at the lowest order set of areas. This is essentially the design used by the National Agricultural Statistics Service (NASS) (Cotter and Nealon 1987) in their June Enumerative Survey of national agricultural production, where each sampled PSU is split into secondary sampling units called segments. Field visits are made to a sample of segments.

Systematic sampling is often used (Bickford et al. 1963, Messer et al. 1986, Hazard and Law 1989). Several variants that perturb the strict alignment are available (Olea 1984). Spatial stratification can use regular polygons, natural boundaries, political boundaries, or arbitrary tessellations as strata. Maximal stratification (i.e., one or two points per stratum) has been viewed as most efficient. To this end, Munholland and Borkowski (1996) have used a Latin square with a single additional independent sample to achieve a spatially balanced sample. Breidt (1995) used a Markov process to generate a one-unit-per-stratum spatially distributed sample. Both of these techniques select cells in a regular grid. Another approach is to use space to order a list frame of the (finite) population and then use the order of the list to structure the sample, say by defining strata as successive segments of the ordered list, or by systematic random sampling. For example, Saalfeld (1991) drew on graph theory to define a tree that leads to a spatially articulated list frame, and the National Agricultural Statistics Service has used serpentine strips (Cotter and Nealon 1987) to order their primary sample units within a state. A related idea that originated in geography is the General Balanced Ternary (GBT) spatial addressing scheme (Gibson and Lucas 1982). The concept behind a GBT address is related to the concept of space-filling curves, such as those first constructed by Peano (1890), or the Hilbert curve (Simmons 1963). Wolter and Harter (1990) have used a construction similar to Peano's to construct a "Peano key" to maintain the spatial dispersion of a sample as the underlying population experiences births or deaths. Saalfeld (1991) also has used the Peano key to maintain spatial dispersion of a sample. Stevens and Olsen (1999, 2004) used a similar concept, recursive partitioning, together with hierarchical randomization, to distribute sample points through space and time.

The Generalized Random Tessellation Stratified (GRTS) (Stevens and Olsen 2004) technique is another possible approach for spatially balanced samples. The tessellation used in the standard GRTS design is a square grid. The basic concept of the GRTS design is to construct a random spatial stratification using equal-sized tessellation cells and then to select one sample point at random within each cell. The tessellation is constructed beginning with a single large square cell randomly placed over the population domain. The cell is split into four quadrants, which are in turn split into four quadrants, and so on, until the required resolution is obtained. The usual practice is to continue the partitioning until five to ten times as many points are available as the target sample size. A spatial address, which can be used to order the points linearly, is constructed using the pattern of subdivision. The address is used to arrange the points in an order, called *reverse hierarchical order*, which has the property that any consecutive sequence of points will be a spatially balanced sample. See Stevens and Olsen (2004) for details. This property allows dynamic adjustment of sample sizes while maintaining spatial balance. It can be especially useful when we don't know beforehand whether a selected point will in fact be in our target population (imperfect frame information) or when a substantial portion of the initial sample sites may fall on private property where access permission may not be available. We select more samples than we expect to need, divide the sample into equal-sized, spatially distributed groups, and proceed with the assessment. We then add groups to the initial set if access is limited, while maintaining a sample that is well dispersed over the accessible portion.

A spatially balanced sample will normally be more precise than a simple random sample of the same size because its spatial balance capitalizes on the spatial structure of the response. However, because of the restricted randomization inherent in the spatial balance, variance estimation can be an issue. Technically, the variance depends on pair-wise or joint inclusion probabilities (the probability that a pair of points are both included in the sample). The restricted randomization implicit in spatial balance makes some of those joint probabilities very small or zero. The joint probabilities appear in the denominator of the usual variance estimators, so the estimators are undefined if joint probabilities are zero and unstable if small. A commonly used approach is to ignore the spatial constraint in the design and apply the simple random sample variance estimator. The resulting estimator will almost always be biased high. Horvitz and Thompson (1952) derived an unbiased variance estimator to accompany their estimator of the total, but the joint inclusion probability appears as a divisor in the estimator, so it is unsuitable for spatially balanced designs.

One general purpose technique that provides reasonably good results is to apply a post-selection spatial stratification with at least two points per stratum. The strata can be selected arbitrarily, but the points in a stratum should be close together. The usual stratified sample variance estimator is then applied. Stevens and Olsen (2003) developed a variance estimator specifically for GRTS designs that is based on a similar concept. Instead of explicitly forming strata, a local variance is computed at each sample point. The local neighborhood of a point is defined as a region containing the point's four nearest neighbors and then expanded to satisfy a symmetry constraint (if *a* is in the neighborhood of *b*, then *b* must be in the neighborhood of *a*). The overall variance estimate is a weighted average of the local estimates. Our analyses use the Stevens and Olsen local variance estimator.

Advantages of Spatially Balanced Sampling

Spatially balanced sampling is advantageous for wetland surveys, since most responses of ecological systems show some spatial pattern, at least to the extent that sites that are close together tend to have more similar responses than widely separated sites. Proximate sites tend to be in the same physiographic region, have the same geologic substrate, and be subjected to the same natural and anthropogenic disturbances and stressors. That structure may be patchy rather than smoothly changing because of influences such as localized management practices, localized contamination, localized development, or natural discontinuities. Nevertheless, there will be spatial patterns in the condition of wetlands, so it is appropriate to account for spatial patterns when designing samples for wetland condition.

Spatially balanced sampling has an advantage over SRS, as it is more efficient. Intuitively, the efficiency gain occurs because sites that are close together tend to contain redundant information. In statistical terms, the sites show positive spatial correlation that decreases with separation distance. A sample with a regular distribution of sites will tend to have less between sites correlation and, hence, more information and lower variance than a completely random site pattern. So, if we expect that the wetland population has a spatial pattern in its condition, then a spatially balanced sample helps us avoid a biased sample or one with redundant information from multiple, proximate sites, which we might get with SRS.

Implementation of Spatially Balanced Sampling on Juniata and Nanticoke

The GRTS technique was used to draw both the Juniata and Nanticoke samples. The target domains were the wetland area in the two watersheds, and samples were point samples from the target wetland area, so that the inclusion probability is the point density with units of (number of sample points) / (*ha* of target wetland area). This number will be adjusted based on the actual disposition of the initial points.

Two benefits of the GRTS technique were mentioned prior: it allows for 1) imperfect frame information and 2) augmentation of the sample based on field experience. In both the Juniata and the Nanticoke studies, both of these properties came in handy. In the Juniata study, one of the maps (CWC) used as a sampling frame was a map of areas where wetlands were only likely to occur. The sample using the CWC map was formed by first selecting a point sample from the map, examining an aerial photo of a 5-ha rectangular plot centered on the points for wetlands, and then selecting a subsample from all identified wetlands. Because wetlands were only likely to occur, not all of the rectangular plots contained wetlands, making the CWC map an imperfect frame. Only the points with wetlands from the CWC sample were used in the final sample.

The GRTS design property of allowing augmentation of the sample also was beneficial. In both the Juniata and the Nanticoke studies, a sizeable number of sites had access permission denied or were non-response, but by using the GRTS design, the exclusion of points where sampling was not allowed did not interfere with the spatial balance of the sample. Also, in the Nanticoke study, an excess of points was drawn initially, approximately ten times the target sample size. Points were arranged in replicate groups of 25 points each, where each group was a spatially balanced sample, and groups could be evaluated sequentially until a sufficient sample size was obtained.

TWO-PHASE SAMPLING

An essential first step in probability sampling is to develop a representation of the target population from which to select a sample (a frame, in statistical terminology). NWI is the most complete catalog of wetland location, type, and extent that is nationally available. As such, it is typically the default frame from which to draw a wetland sample. Our experience suggests that NWI may miss over half of the smaller wetlands in forested portions of the watershed, resulting in significant undercoverage. To solve this problem, CWC used geologic structural and stratigraphic information to generate a map (CWC map) of areas with high probability of wetland occurrence (Wardrop et al. 2007). One way to proceed would be to select sample points from the CWC map, use aerial photos to locate all wetlands in a rectangular plot around the sample points, and then select a second sample from the identified wetlands. In the survey-sampling field, this operation of taking a sample of a sample is known as a two-phase sample (Särndal et al. 1992). It is frequently used in circumstances where a sampling frame is too difficult or expensive to construct for the entire population. Another application is when a desired target variable is expensive to measure, but a correlated surrogate variable is relatively easy or cheap to obtain (Thompson 2002). A large primary sample is taken with the surrogate variable measured, and a smaller secondary sample is taken for the target variable. The analysis draws on the correlation between the two variables. Two-phase sampling can be a very cost-effective technique.

For example, an initial sample might be selected to determine whether wetlands are present at selected sites, where presence/absence would be the auxiliary information. A second sample might be taken from those sites with wetlands present, and wetland condition could be estimated from that second sample. This avoids the possibility of selecting sites with no wetlands while attempting to sample wetland condition.

Analysis of Two-Phase Samples

The probability of a site being selected for the second-phase sample is a function of its probability of being selected for the first-phase sample and must be adjusted appropriately. In the general case, the inclusion probability at the second phase of sampling depends on the results of the first phase. Because of this dependence, the true inclusion probability is unknown. However, we can calculate the inclusion probability conditional on the achieved first-stage sample. Given our achieved first-phase sample, we control the second-phase selection, so we know the conditional probability of selecting a particular element. We then use the quantity $\pi^*(s) = \pi_1(s)\pi_2(s \mid s \in S_1)$ where $\pi_1(s)$ is the Phase 1 inclusion probability and $\pi_2(s \mid s \in S_1)$ is the conditional probability of including the element at s in the Phase 2 sample given the Phase 1 sample S_1 in the Horvitz-Thompson estimator. The resulting estimator is unbiased. See Särndal et al. (1992) for details.

Juniata Example of Two-Phase Sampling

In the Juniata study, a two-phase sample was used to select sites from the CWC map, which showed areas of high probability of wetland presence. The phase-one CWC sample (n = 79) was selected from the CWC map and was combined with NWI sample points for the landscape assessment. Aerial photos were used to locate all wetlands in a rectangular plot around the sample points of the phase-one CWC sample, and then a second sample was selected from the identified wetlands (the phase-two CWC sample). The phase-two CWC sample (n = 30) was combined with NWI sample points for the rapid and intensive assessments. By using a two-phase sampling design for the CWC map, the final samples for the rapid and intensive assessments did not include sites that didn't contained wetlands.

MULTIPLE FRAME SURVEY DESIGNS

One of the difficulties in designing a probability sample for wetlands is obtaining a complete frame (i.e., a comprehensive catalog of wetlands in the target region). The NWI is the default source of location information on wetlands, but as noted in Brooks et al. (1999), the NWI can miss a substantial portion of wetlands. Thus, some method is needed to augment the NWI catalog. One way to do this is to use two or more sources of frame information, each of which may be incomplete but, collectively, span the entire population.

Implementation of a multiple frame design is straightforward. One simply draws independent samples from each frame. Sample sizes may be equal or proportional to the anticipated population coverage of each frame. For example, if one frame is thought to cover 80% of the population and the other 30% with 10% overlap, one might split the total sample size 80:30. Frequently, one of the frames is difficult or expensive to construct or contains a rare or difficult to locate subpopulation. In these instances, one may sample that frame disproportionately to its population coverage.

Sample Analysis with Multiple Frames

Again, the Horvitz-Thompson theorem is the appropriate analysis tool. We need only determine how the inclusion probability is affected by the use of multiple frames, which is normally a straightforward application of basic probability rules. The usual case is that the samples from the different frames were selected independently of one another. In this case, the inclusion probability for any element is the sum of the inclusion probabilities from every frame that covers that particular element. For example, for two frames with respective inclusion probabilities π_1 and π_2, a sample element is in frame 1 only, frame 2 only, or both frames. If the element is only in frame 1, its inclusion probability is π_1; if it is only in frame 2, its inclusion probability is π_2; and if it is in both frames, its inclusion probability is $\pi_1 + \pi_2$. Note that the overlap need only be determined for the sample elements. We don't need to know what the frame overlap is before the sample is drawn; we only need to be able to determine it for the elements selected as sample elements. This can be a substantial advantage if the frame development is time-consuming or expensive, as in the example that follows.

Juniata Example of Multiple Frames

As mentioned prior, in the Juniata study, there was a need to augment the NWI frame, as it does not include all target wetlands. This led to the use of a second frame, which was a map from the CWC showing areas with a high probability of wetland existence. By using the CWC map, it was judged that all wetlands could be sampled. However, the CWC frame was substantially more difficult and time-consuming to use. The use of multiple sampling frames was a decision to balance completeness and cost-effectiveness in the sample.

The samples from the NWI frame and the CWC frame can be merged and analyzed as a single sample. We assume that the NWI frame is a proper subset of the CWC frame (i.e., all NWI wetlands could be sampled from the CWC frame, but some wetlands in the CWC frame do not appear in the NWI frame). The samples from the two frames were drawn independently, so the inclusion densities are additive. The composite inclusion density is therefore

$$\pi_{Total} = \begin{cases} \pi_{NWI} + \pi_{CWC}, \text{ if wetland is on NWI map} \\ \pi_{CWC}, \text{ otherwise} \end{cases}.$$

All the NWI sample points have inclusion density $\pi_{NWI} + \pi_{CWC}$, as do those CWC points that are also in the NWI frame. The only points that get density π_{CWC} are those in the CWC frame that do not appear on NWI maps.

POST-STRATIFICATION

Post-stratification (Holt and Smith 1979) introduces the effect of stratification into the analysis, and the resulting population estimates when the actual sampling was not stratified but random. This is accomplished by sorting the data into the appropriate strata after sampling is completed, calculating the desired statistic for each stratum, and then weighting the statistics from each stratum. Each weight is the proportion of the population that falls into that stratum. In this way, post-stratification is a technique for adjusting statistics to account for strata existing in the population but not acknowledged until after collecting the sample.

Post-stratification is useful in several different situations in which random sampling was used, but it is thought that accounting for population strata after the fact will improve the resulting population estimates. For example, the stratifying variable may be obvious but unavailable for use during sampling, or it may be available during sampling but not noticed as important until after sampling is completed.

Post-Stratification to Adjust for Non-Response

An interesting application of post-stratification is in the case of very different non-response rates on privately and publicly managed lands, when attempting to sample sites on both. Permission to sample is explicitly or implicitly denied (a non-response) more often for privately managed sites than for publicly managed sites, leading to achieved samples with a disproportionate number of public sites. This phenomenon is worrisome because straightforward extension of results from the sampled subpopulation to the entire target population relies on the assumption that the missing responses are missing completely at random (MCAR). This assumption is violated when management is shown to be associated with site-accessibility for the study. Furthermore, management practices are sometimes found to be associated with wetland condition. These circumstances, commonly encountered in large-scale studies of ecological condition, are strong reasons for post-stratifying on ownership (private versus public).

Post-Stratification to Correct for Poor Random Selection

Another situation that might prompt post-stratification is getting a bad draw, just by chance, for a random sample. In this situation, it is realized after the fact that a disproportionate amount of the sample falls in a particular population stratum. This can only be recognized if the proportion of the population that falls into that stratum is known. If it is suspected that the stratifying variable influences the response, post-stratifying can help correct the imbalance that occurred purely by chance.

To use post-stratification, the proportion of the population that falls into each stratum must be known. The stratified statistic for stratum *i* is then the proportion of the population that falls into stratum *i* multiplied by the original statistic for stratum *i*.

Juniata Example of Post-Stratification

Both the Nanticoke and the Juniata watershed assessments (Wardrop et al. 2007) used post-stratification to adjust for non-response, as described preceding, since the MCAR assumption was not reasonable.

The post-stratification technique is described here in detail, as it was used in the Juniata study, which is very similar to how it was used in the Nanticoke study. First, the researchers suspected that the achieved sample had a disproportionate number of publicly owned sites. This was verified by use of Pearson's χ^2 test, which showed highly significant lack of independence of ownership and sampling access for both the CWC and NWI samples (CWC: $\chi^2 = 15.5$, df = 1, p = 0; NWI: $\chi^2 = 23.3$, df = 1, p = 0). That non-response was not independent of ownership raised the concern that the MCAR assumption was not tenable, and differing landscape positions and management practices between publicly and privately owned wetlands could affect wetland condition.

Recognizing that public versus private ownership does not always indicate management practices, sites were reassigned to the appropriate category based on known management. For example, a number of the points occurred on privately owned hunting clubs located near state forests. These properties were managed similarly to the public land and were treated the same in the analysis.

A t-test showed a significant difference in the mean landscape-level scores (for which there was complete information) for access granted versus denied within each entire sample (CWC: t = 2.66, df = 74, p = 0.009; NWI: t = 2.11, df = 247, p = 0.036). Also, even though there was no direct comparison of condition between responders and non-responders for either rapid or intensive site-assessment scores, landscape-level scores were highly correlated with both site-level scores (rapid assessment: r = 0.929, 95% c.i. = (0.891, 0.953); intensive assessment: r = 0.580, 95% c.i. = (0.416, 0.707). These results suggested that the MCAR assumption was probably not true for the site-level samples.

To investigate if there was some additional bias resulting from non-response (e.g., landowners may have been less willing to give access permission to wetlands in poor condition), two-way analysis of variance was used to evaluate how much of the non-response differential could be attributable to differential sampling rates by ownership class. It was found that, after adjusting for the effect of ownership, neither access nor ownership-access interaction effects were close to being significant ($p > 0.4$) for both the NWI and CWC samples. This suggested that condition of the wetland did not influence access permission.

Once it was decided to post-stratify publicly and privately managed sites, we needed to modify the original inclusion-probability-based weights of each sampled site, resulting in a weight for each sampled site. The modifiers should have been the proportion of the population in each stratum (either the proportion of publicly or privately managed sites in the population), but those proportions were unknown, so they were estimated using information from the sample. For example, the modifier for the privately managed sampled sites was the ratio of the number of privately managed sampled sites and the total number of privately managed sites in the sample draw; this represents, in the population of privately managed sites, proportion that are accessible. Once these modifiers were calculated, the weight for each sampled site was calculated by multiplying the appropriate modifier (depending on whether the site was publicly or privately managed) by the inclusion probability. Finally, the weight, not the inclusion probability, for each sampled site was used when evaluating wetland condition, post-stratification.

SUMMARY

Several survey sampling techniques have been discussed with wide application for sampling wetlands. Most survey sampling methodology is rooted in a finite population perspective, but the infinite, continuous population described here is often more appropriate for wetland surveys. Spatially balanced sampling capitalizes on the spatial correlation that is present in ecological data; sites in closer proximity to each other tend to be more similar. The use of multiple sampling frames, such as more than one map, helps to create the most comprehensive from which to select a sample. Two-phase sampling, or taking a sample from a sample, provides an additional technique to use when developing frame information is difficult, time-consuming, or expensive. Finally, post-stratification can help correct a biased sample resulting from non-response, or lack of access to particular sites in the sample draw (e.g., private land).

ACKNOWLEDGMENTS

We thank all of the contributors to the Juniata and Nanticoke studies, especially Denice Wardrop and Dennis Whigham for generously allowing their data to be used as examples, and Mary Kentula for her guidance and encouragement. We also thank Gretchen Moisen and Jean Opsomer for their kind reviews and helpful comments. Preparation of this publication was partially supported under STAR Research Assistance Agreement No. CR82-9096-01 awarded by the U.S. Environmental Protection Agency to Oregon State University. It has not been formally reviewed by EPA. The views expressed in this document are solely those of the authors and EPA does not endorse any products or commercial services mentioned in this publication.

LITERATURE CITED

Bickford, C. A., C. E. Mayer, and K. D. Ware. 1963. An efficient sampling design for forest inventory: the Northeast Forest Resurvey. Journal of Forestry 61:826–33.

Breidt, F. J. 1995. Markov chain designs for one-per-stratum sampling. Survey Methodology 21:63–70.

Cordy, C. 1993. An extension of the Horvitz-Thompson theorem to point sampling from a continuous universe. Probability and Statistics Letters 18:353–62.

Cotter, J. and J. Nealon. 1987. Area frame design for agricultural surveys. U.S. Department of Agriculture, National Agricultural Statistics Service, Research and Applications Division, Area Frame Section, Washington, DC, USA.

Gibson, L. and D. Lucas. 1982. Spatial data processing using balanced ternary. Proceedings of the IEEE Computer Society Conference on Pattern Recognition and Image Processing. IEEE Computer Society Press, Silver Springs, MD, USA.

Goebel, J. J. and K. O. Schmude. 1982. Planning the SCS National Resources Inventory. Arid Land Resource Inventories' Workshop. p. 148–53. USDA, Forest Service General Technical Report, WO-28., Washington, DC, USA.

Hazard, J. W. and B. E. Law. 1989. Forest Survey Methods Used in the USDA Forest Service. U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory, Corvallis, OR, USA. EPA/600/3-89/065.

Holt, D. and T. M. F. Smith. 1979. Post stratification. Journal of the Royal Statistical Society, Series A: General 142:33–46.

Horvitz, D. G. and D. J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47:663–85.

Kott, P. S. 1986. Some asymptotic results for the systematic and stratified sampling of a finite population. Biometrika 73:485–91.

Messer, J. J., C. W. Arsiss, J. R. Baker, S. K. Drousé, K. N. Eshleman, P. R. Kaufmann, R. A. Linthurst, J. M. Omernik, W. S. Overton, M. J. Sale, R. D. Schonbrod, S. M. Stambaugh, and J. R. Tuschall Jr. 1986. National Surface Water Survey: National Stream Survey, Phase I-Pilot Survey. U.S. Environmental Protection Agency, Washington, DC, USA. EPA-600/4-86/026.

Munholland, P. L. and J. J. Borkowski. 1996. Simple Latin square sampling +1: a spatial design using quadrats. Biometrics 52:125–36.

Olea, R. A. 1984. Sampling design optimization for spatial functions. Mathematical Geology 16:369–92.

Peano, G. 1890. Sur Une Courbe, Qui Remplit Toute Une Aire Plane. Mathematische Annalen 36:157–60.

Royall, R. M. and J. Herson. 1973. Robust estimation in finite populations. II: Stratification on a size variable. Journal of the American Statistical Association 68:890–93.

Saalfeld, A. 1991. Construction of spatially articulated list frames for household surveys. p. 41–53. *In* Proceedings of Statistics Canada Symposium 91, Spatial Issues in Statistics. Ottawa, Ontario, Canada.

Särndal, C., B. Swensen, and J. Wretman. 1992. Model Assisted Survey Sampling. Springer-Verlag, New York, NY, USA.

Simmons, G. F. 1963. Introduction to Topology and Modern Analysis. McGraw-Hill, New York, NY, USA.

Stevens, D. L., Jr. 1997. Variable density grid-based sampling designs for continuous spatial populations. Environmetrics, 8:167–95.

Stevens, D. L., Jr. and A. R. Olsen. 1999. Spatially restricted surveys over time for aquatic resources. Journal of Agricultural, Biological, and Environmental Statistics 4:415–28.

Stevens, D. L., Jr. and A. R. Olsen. 2003. Variance estimation for spatially balanced samples of environmental resources. Environmetrics 14:593–610.

Stevens, D. L., Jr. and A. R. Olsen. 2004. Spatially-balanced sampling of natural resources. Journal of the American Statistical Association 99:262–77.

Thompson, S. K. 2002. Sampling. Wiley, New York, NY, USA.

Wardrop, D. H., M. E. Kentula, D. L. Stevens, Jr., S. F. Hornsby, and R. P. Brooks. 2007. Assessment of wetland condition: an example from the Upper Juniata watershed in Pennsylvania, USA. Wetlands 27:416–31.

Whigham, D. F., A. D. Jacobs, D. E. Weller, T. E. Jordan, M. E. Kentula, S. F. Jensen, and D. L. Stevens, Jr. 2007. Combining HGM and EMAP procedures to assess wetlands at the watershed scale: status of flats and non-tidal riverine wetlands in the Nanticoke River Watershed, Delaware and Maryland (USA). Wetlands 27:462–78.

Wolter, K. M. and R. M. Harter. 1990. Sample maintenance based on Peano Keys. p. 21–31. *In* Proceedings of the 1989 International Symposium: Analysis of Data in Time, Ottawa, ON, Canada.

Yates, F. 1981. Sampling Methods for Censuses and Surveys, fourth edition. Griffin, London, UK.

Manuscript received 11 August 2006; accepted 29 January 2007.