NONSAMPLING ERROR CONSIDERATIONS IN ENVIRONMENTAL SURVEYS

Virginia M. Lesser, Oregon St. Univ, and William D. Kalsbeek, Univ. of N.C.
Virginia M. Lesser, Statistics Dept., 44 Kidder Hall, OSU, Corvallis, OR 97331-4606

KEY WORDS: Nonsampling error, NRI, NASS, EMAP.

ABSTRACT

A number of probability surveys are in place in the U.S. to estimate and monitor changing conditions of ecological resources. In addition to sampling errors, nonsampling errors also need to be investigated to assess total survey error. Examples of nonsampling errors encountered in environmental surveys are discussed. More specifically, the impact of frame, measurement, and nonresponse errors encountered when monitoring various ecological resources within the Environmental Protection Agency's Environmental Monitoring and Assessment (EPA-EMAP), the U.S. Department of Agriculture's (USDA) probability surveys, and USDA's National Resources Inventory (NRI) program are examined and remedies to adjust for nonsampling errors in environmental surveys are discussed.

INTRODUCTION

There is an increasing level of interest in assessing environmental conditions around the world. A number of ongoing ecological sampling programs have been established in the United States (U.S.) in order to monitor ecological condition over time. The purpose of these sample surveys is to describe current conditions of certain ecosystem health as well as assess change through comparison of the repeated measurements. The use of probability sampling to draw statistical inferences in these surveys assumes ideally that we can: (1) construct a perfect frame for the target population; (2) select a sample from that frame that follows the probability sampling design; (3) collect information for each member of the chosen sample; (4) observe the true value for each study variable for each member of the sample; (5) process the data without introducing errors; and (6) use the processed, correct data to make valid inferences about the finite population (Sarndall, Cassel and Wretman; 1992) In most surveys, these ideal conditions are not met and nonsampling errors, as well as sampling errors introduced by studying a subset rather than the population, exist.

Nonsampling errors include errors of nonresponse, measurement, and frame errors. The objective of this paper is to discuss the prevalence of nonsampling errors in some probability environmental surveys and some remedies to deal with them, either preventative or compensatory. It is hoped that this paper might be a first attempt in development of error profiles for some major recurrent environmental surveys, similar to the one prepared by Brooks and Bailar (1978), for the Current Population Survey. Each of the three types of nonsampling errors will be briefly introduced followed by a short description of three national-scale probability-based environmental surveys, emphasizing aspects of the survey design that may lead to nonsampling error. Finally, each of the three types of nonsampling errors will be discussed in terms of the their prevalence in each of the three environmental surveys. Discussion of approaches to deal with nonsampling errors in these environmental surveys are addressed.

TYPES OF NONSAMPLING ERRORS

Frame Error

The mechanism that provides access to the population is called the sampling frame. For surveys of human populations, lists of households, employees, employers, special interest groups, etc., are often gathered at the design stage of the survey in order to identify members of the population from which to draw the sample. Similarly, in environmental surveys, a list of the population members is required to draw a sample. However, the population members are not individuals but ecological resources such as prairie wetlands, lakes, or streams. List frames that define a list of sampling units such as lakes for ecological resources are not as prevalent. Therefore, many environmental surveys are using area frames which define land areas as sampling units (Cotter and Nealon, 1987; Nusser and Goebel, 1997; Overton, et.al, 1990; Powell, 1994).

To minimize errors attributed to the frame it is critical that prior to drawing the sample selection, a perfect frame has the one-to-one correspondence between the listing on the frame and the sampling unit in the population. These perfect frames are rare, and hence Sarndal, Swensson, and Wretman (1992), Lessler and Kalsbeek (1992), and Kish (1965) describe the various types of frame error, which include undercoverage, overcoverage, and multiplicity. Lessler and Kalsbeek (1992) define undercoverage as a failure to include all units belonging to the defined population in the frame, overcoverage as a failure of the frame resulting from inclusion of elements that are not part of the target population, and multiplicity as multiple frame units linking to the same sampling unit in the population. Each of these references also discuss treatments to these problems. The discussion below will discuss frame errors, as well as treatments, with respect to the environmental surveys under study in this paper.

Nonresponse Error

Nonresponse errors refer to the errors associated with the failure to obtain information from a selected sampling unit in the survey. Most surveys of human populations have some form of nonresponse error due to either refusal, person unavailable for providing information, incorrect address or telephone number for the selected sampling unit, death, etc. (Sarndall, et. al., 1992; Groves, 1989). The effect of nonresponse is usually bias, although increase variance can also occur when nonresponse is viewed as a stochastic event (Lessler and Kalsbeek, 1992).

Nonresponse can arise in environmental surveys. If information is required from the landowner such as the amount of herbicide applied to the selected areal unit, nonresponse can occur due to all the same reasons as just mentioned above. Nonresponse can also be attributed to instrument failure, for example the camera used in aerial photography or pH meter. In addition, if physical descriptions of the habitat and/or water or soil samples about the selected site are needed, the enumerator must still request site access from the landowner to obtain this information. The landowner may still deny access and thus create nonresponse.

Measurement Error

The third source of error that may be present in surveys is measurement error (Cochran, 1977; Lessler and Kalsbeek, 1992). Measurement errors on the sampling unit arise in the data collection process, either due to mistakes or limitations tied to the measuring device, interviewer, or survey respondent. Cochran (1977) summarizes measurement errors as either a constant bias introduced over all sampling units, a variable component of bias, or a fluctuating component of error.

In large scale environmental surveys, many measurements by a large number of individuals are recorded. Instruments need to be consistently calibrated and all enumerators equally and thoroughly trained on using the instruments and recording survey data. In order to assess the magnitude of measurement error, repeated measurements may be taken for some sample members, or a "gold standard" measurement may be obtained to compare against the measurements made on survey respondents. Cochran (1977) and Lessler and Kalsbeek (1992) provide models applied to these data to quantify the extent of the measurement error.

In order to deal with nonsampling errors, Lessler and Kalsbeek (1992) propose three remedies. Although these remedies and examples are presented in terms of nonresponse error, the general approach can be applied to other nonsampling errors. The first remedy includes methods to prevent or reduce the prevalence of nonsampling error and assessments to identify the potential for nonsampling error, such as documented training manuals. Identification studies are assessments to identify the potential for nonsampling error effect on the survey results. The last approach includes a number of adjustment procedures that can be introduced at the analysis stage to adjust for nonsampling error bias. Within the context of the discussion of nonsampling errors in the three environmental programs, the type of remedy currently adopted to address the nonsampling errors is also presented.

EXISTING NONSAMPLING ERRORS EVALUATED IN THREE NATIONAL ENVIRONMENTAL SURVEYS

Aspects of the survey design that may lead to nonsampling error for each of the three environmental surveys were examined. Due to page constraints, details of the survey designs are not provided here but can be obtained from the references listed in the following section. However, areas of the survey design that were carefully reviewed include details of frame construction, questionnaire wording and format, data collection procedures, training and quality control, and data processing including imputation. The prevalence of frame, nonresponse, and measurement errors within these three national surveys is discussed below. A number of different ecological resources are studied in the EMAP survey. This involves separate frame development and data collection procedures for each ecological resource sampled in EMAP. Nonsampling errors were examined for all the different resources, but our discussion of nonsampling error will be limited to a review of just one of the ecological resources for each type of nonsampling error.

Frame Error

A critical component to the success of some of these nationally based sampling programs is obtaining and maintaining a suitable, up-to-date frame to select sampling units, obtaining and maintaining access to the selected sampling sites across the U.S., and obtaining consistent unbiased measurements throughout the U.S. Up-to-date frames for these environmental surveys involve use of recent remote sensing and aerial photography to develop current listings of these ecological resources.

NASS

Quite extensive development of the frames used by NASS to sample agricultural resources have been made since the mid-1930's (Fecso, et.al., 1986). Although the area frame is complete in that it covers the U.S., using just an area frame for sampling rare resources would not be efficient due to the amount of sampling necessary to obtain an adequate sample size to estimate these rare resources, such as rice production. With the inclusion of a list frame, specifically a list defining all rice farmers in the U.S., these rare resources can be sampled more efficiently. However, using the dual frame as a remedy to sample more efficiently also introduces an additional source of error to consider, duplication of farm operations. Duplication can occur both within the list frame and between the list and area frames.

NASS has employed a model by Fellegi and Sunter (1969) to determine and quantify whether within list duplication exists. As Fecso et.al., (1986) point out, the manual resolution process is a time-consuming one. Also, the chance of duplication still exists due to human error, and thus may surface in the selected sample. The manual comparison of names between the list and area frames to identify cross-frame duplication can cause similar problems. NASS has built other ways to catch duplicate farm operations by adding questions on the survey instrument to determine whether or not any other individuals might also be associated with this farm operation. This would increase the chance for this farm operation to be selected in the sample due to multiple listings. Musser and Mergerson (1994) have also proposed a methodology to measure the percent of duplication present in NASS list frames.

Once area frames for a state are developed, NASS uses this frame for sample selection until the area frame deteriorates and a new frame is developed. The area frames do not become out of date due to population coverage, but efficiency is reduced due the attenuation of the land use stratification resulting from changes in land utilization (Fecso, et.al, 1986). However, list frames of farm operators become out of date due to constant changes in the population of farms. About 20% of the records in the U.S. agricultural list frame will change from year to year. These changes result in problems with the list in terms of both out of scope records (overcoverage) and new operations not covered (undercoverage). NASS uses the area frame to estimate for the incompleteness of the list frame.

NASS uses both a list and area frame as the basis for selecting its sample in order to obtain the advantages of completeness in the area frame and efficiency of sampling rare and variable items in the list frame. However, it is clear that the potential for frame error, such as multiplicity, exists and the need to update frames is necessary to improve frame efficiency. NASS has employed quality control procedures to determine frame construction errors and help guide them to determine when States should have a new frame, which are typically every 15 years. For example, as the frame ages and land use changes over time, boundaries delineating areal sampling units may change which may lead to nonsampling error.

NRI

For each NRI survey, sample points are selected from the nonfederal land area of the U.S. (Nusser and Goebel, 1997). In order to assure all nonfederal lands are located, the sampling frame is developed that extends beyond the target population, nonfederal lands. The location and extent of the target population is identified at each NRI survey. Only one frame is used to select a separate two-stage stratified sample within each county. Strata are formed by geographical units defined by the Public Land Survey. Obtaining estimates of landuse for primary sampling units may be complicated due to non-permanent boundaries.

EMAP STREAMS

A systematic sample of hexagons was selected across specific regions in the U.S. to represent an areal sample of approximately 1/16th of the land area in the U.S. (Overton et. al., 1990). Descriptions of the areal extent of ecological resources, such as streams, within these hexagons are then determined. To accomplish this to assess extent of streams in the U.S., hydrographic maps (1:100,000) provided by the United States Geological Service (USGS) were used. These maps were then modified by USEPA to create the River Reach File, Version 3 (RF3). This file was chosen as the frame to represent stream resources in the U.S. (Larson, et.al, 1997). The stream resources cover large rivers to the smallest tributaries, including both perennial and intermittent streams. This frame was expected to be a complete frame of all stream resources within the sampled hexagons to be used for sample selection within the EMAP survey.

When reviewing this RF3 file, it was apparent that inconsistencies with regard to stream density existed between regions of the country and at state or county borders. For example, there is quite a sharp contrast in stream density between Virginia and West Virginia. The numbers that appear on these quadrants are the number of digitized lines that border the edge of the quadrant. These inconsistencies are not due to differences in stream density, but can be attributed to frame error (undercoverage), most likely caused by the cartographers working on this project using different guidelines in producing their statewide or regional maps. It is expected that some cartographers only digitized listed streams on this file, while others digitized both unlisted and listed streams.

In order to obtain a complete frame, EPA's River Reach File was updated prior to sample selection in order to remove the undercoverage problem. The maps used to create the digital line graphs were used to supplement the RF3 file. Finally, since no permanent delineated boundaries exist for the sample hexagons, nonsampling error may arise in determining the extent of stream resources within these hexagons.

Nonresponse Error

An objective for all three environmental surveys is to assess changing condition. This results in obtaining some repeated measurements on selected areal sampling units over time. The area frame portions of USDA's June Agricultural Survey incorporates a rotating panel design, while NRI and EMAP have adopted a longitudinal design. Another form of nonsampling error that arises in repeated surveys is due to respondent burden and subsequent denied access, i.e., nonresponse. The following discussion examines nonresponse error within these three environmental surveys.

NASS

NASS has experienced decreasing response rates over the past 10 years (Dale Atkinson, personal communication). Some of this can be attributed to respondent burden. For example, large farm operators representing a large farm output have a higher probability of selection, not only for the Quarterly Agricultural Survey, but for other specific surveys also run by NASS.

In order to prevent nonresponse bias and improve response rates, NASS has formed focus groups with farmers to determine how to reduce the burden of obtaining response. In addition, public relations activities have increased to better explain to farmers the importance of responding to the survey. Another approach to compensate for nonresponse is analytical. NASS uses nonresponse adjustment models in order to minimize nonresponse bias (Cox, 1993).

NRI

Nonresponse for NRI has been nearly nonexistent until very recently (Sarah Nusser, personal communication). In 1995, a nonresponse rate of less than 1%, attributed to denied access, was obtained in a special soil erosion study of 8000 sampling points. Since nonresponse rates are so small with NRI, no procedures have yet been implemented to adjust for nonresponse. However, with the introduction of remote sensing in the 1987 NRI survey, an increasing number of sampling points have been assessed without the need to obtain a site visit. Thus, the consequence of lower response rates due to respondent burden have been reduced with the introduction of this remote sensing technology. Only one-fourth of the sample sites were actually field visited in the 1992 NRI study since required data on the remaining sample points was available on maps.

EMAP PRAIRIE WETLANDS

In 1995 and 1996, field studies were conducted jointly with EMAP and the National Biological Survey (NBS) scientists to assess the status and condition of prairie wetlands in Eastern North Dakota. A systematic sample of hexagons was selected across Eastern North Dakota to represent an areal sample of approximately 1/16th of the land area. Descriptions of the extent of the prairie wetlands (approximately 28,000 prairie wetlands) using aerial photography within these hexagons were then determined. A stratified sample of 260 and 253 sites were selected for the 1995 and 1996 samples, respectively. The strata were defined by region and prairie wetland type. Site visits were necessary on all selected sites in order to obtain information on biological condition. Some procedures were used to prevent nonresponse, such as mailing introductory letters explaining the importance of cooperation from landowners and making personal telephone calls to again emphasize the importance of the study. Each landowner was asked to sign a consent form to grant NBS scientists permission to obtain site access. Even with these preventative measures, access rates of 42% and 41% were obtained for each year.

In order to make some adjustments for nonresponse, design-based and model-based weighting class adjustments are being investigated in order to adjust for the potential bias introduced by the high nonresponse rate. Weighting class adjustments were applied to adjust for the variable response rates across geographic regions. Another approach currently under investigation is a model-based approach combined field data with remote sensing data. Although field data were only collected from a subset of the initial sample, remote sensing data were collected on all wetlands selected in the sample.

Measurement Error

Measurement error for a particular measurement is defined as the difference between the true value or "gold standard" of the measurement and the actual value obtained during the measurement process (Lessler and Kalsbeek, 1992). Measurement taken in these three environmental surveys require both handling of special instruments and asking questions. To assess the quality of the collected statistical data, it is important to assess the magnitude of errors of measurement in the data collection process. The following programs have described some approaches adopted to assess the impact and remedy the problems of measurement error.

NASS

NASS attempts to identify all sources of nonsampling error, including measurement error (USDA, 1983). Critical reviews of the the details of the survey procedures, feedback from enumerators and field statisticians, reinterviews of respondents, and comparisons of survey data with previous survey data, censuses or administrative information aids in quantifying the prevalence of measurement error (Atkinson, 1997). After identifying the source of the measurement error, such as interviewer or questionnaire, experiments are designed to measure the magnitude of the measurement error. These experiments commonly use replicated sampling to assess the reliability of measurements. Once this is accomplished, alternative survey procedures are evaluated to assess their effectiveness in reducing or removing the various sources of measurement error.

In order to prevent measurement error at the onset of a survey, supervisors for each state are trained at national schools. These supervisors are then responsible for training field enumerators in the state just prior to going out to the field. Each enumerator also receives additional supervision while in the field and can be monitored by field supervisors. Other areas of nonsampling error that have been studied by NASS include assessing interviewer effects, impact on the questionnaire design, respondent biases, data editing, and mathematical biases associated with the estimators and variance formulae.

NRI

The NRI program invests considerable effort in developing instructions and training personnel so that protocols are followed as uniformly as possible across the U.S. (USDA-SCS, 1992). Guidelines on the definition of the various land use categories are defined in the instruction manuals, as well as guidelines on selecting the sample point in the PSUs. Quality control procedures have been applied to all data to detect unlikely and inconsistent values. This includes the use of edit routines and developing summary tables that can help detect suspicious data points. Computer programs have been developed by independent staff as part of this process (Nusser and Goebel, 1997).

NRI periodically does studies to assess measurement error (Sarah Nusser, personal communication). Independent data gatherers recollect data on a subsample of the selected PSUs. A special study was conducted to assess measurement error in the 1982 NRI survey (NRI, 1987). Results from this study indicated no adjustments for measurement error were necessary to NRI data.

EMAP LAKES

Sampling of lakes within the EMAP program has been conducted since the early 1990's. Measurements of biological, physical or chemical condition have been collected on selected lakes in the Northeastern U.S in order to monitor condition of lakes across the United States.

Prior to sending field crews in the field, all field crews attend training sessions and receive training manuals (USEPA, 1989; USEPA, 1997). Currently, teams of field crews are assigned to obtain data from the sampled lakes. Measurements are collected to quantify the spatial, temporal, and crew variability (Steve Paulsen, personal communication). An extensive quality assurance program exists within EPA to add another level of quality control on data collected in EMAP.

Repeated measurements are part of the survey design for monitoring lakes in order to make some assessment of measurement error attributed to crew, measurement and temporal sampling over the index period. Estimates are computed that use procedures outlined by Stefanski and Bay (1996) which eliminates much of the bias induced through measurement error.

DISCUSSION

Nonsampling errors refer to all errors other than sampling error that can impact final survey estimates. The problems associated with nonsampling errors is defining them, measuring them, and finally controlling the errors in the survey process (Lessler and Kalsbeek, 1992). A number of measures are presently being implemented within these programs to address the impact of nonsampling errors and employing preventative or analytical measures to address these forms of error.

The two longer established surveys, both NASS and NRI, appear to have less frame error problem than EMAP. Since these studies have developed over the past 50 years, that is not surprising since much time has passed to work out a clean frame and make quality assurance comparisons with previous year's data. Since NRI has adopted a single frame with fixed geographic strata multiplicity concerns expected with dual frames are minimized. With new technology becoming more readily available on remote sensing imagery, frame error might be further reduced for all these surveys by updating frames more quickly and cheaply.

There was also quite a difference in response or access rates among the three programs. Response rates for the EMAP prairie wetland study showed much lower response rates than found by studies conducted by NRI or NASS (Sarah Nusser, Dale Atkinson, personal communication). This may be largely attributed to how each of these three agencies are viewed by the respondent. Nearly all respondents for NASS, and a majority of respondents for NRI, are farmers. Historically, both NASS and NRCS could be considered a 'friend' of the farmer since they provide useful information for the farmer. Farmers are told information provided to these agencies is confidential, while this was not done in the prairie wetland study. The EMAP study is a monitoring program of the EPA, a regulatory and enforcement agency, which may be a factor in lower response rates.

Another reason that may explain the variable response rates across the three programs is the type of data that is collected. NASS collects mostly inventory information on crop acreage and livestock in Agricultural Surveys and NRI collects information on soil type, land use and vegetative cover. EMAP is interested in assessing biological condition, and therefore collects water or soil samples to assess deterioration of the selected sample site. This information collected by EMAP may be viewed as more sensitive data as compared to the other two programs, which is reflected in EMAP encountering the lowest response rates.

All three programs expresses concern over increasing nonresponse issues. Approaches need to be developed to prevent decreasing respondent participation. NASS is currently initiating new approaches in the field. Assessments of the impact of imputed data on survey estimates should also be incorporated into the survey process. Finally, new biological measures need to be developed that can be used to meet program objectives but can be obtained by means other than site field visits, such as the use of remote sensing.

As studies expand nationally, it is important training sessions and instruction manuals are continually evolving and scheduled as part of the survey process. Instruments need to be adopted that are easily calibrated and that can maintain stable readings over the range of environmental conditions they will be used.

Some sources of nonsampling error appear to have a greater degree of impact in some programs than others. The error profile discussed in this document needs to be expanded in greater detail for each major survey, such as those discussed in this paper, and be made more accessible in the open literature. It would be fruitful for these programs to share what has been learned in terms of prevention, quantification, and analytical remedies to gain insight on major sources of nonsampling error which need to be addressed in an environmental program.

REFERENCES

Atkinson, D. *Assessing Nonsampling Errors in Survey Data Through Random Intercept Models.* Proceedings of the 1997 American Statistical Meetings, Anaheim, CA.

Brooks, C.A. and B.A. Bailar. 1978. *An Error Profile: Employment as Measured by the Current Population Survey.* U.S. Department of Commerce, Washington, D.C.

Cochran, W.G. (1977). *Sampling Techniques.* New York: Wiley.

Cox, B.G. (1993). *Weighting Class Adjustments for Nonresponse in Integrated Surveys: Framework for Hog Estimation.* National Agricultural Statistics Service, Washington, D.C.

Fecso, R., Tortora R., and Vogel, F. (1986). Sampling frames for agriculture in the United States. *Journal of Official Statistics*, 2, 279-292.

Groves, R.M. (1989). *Survey Errors and Survey Costs.* New York: Wiley.

Kish, L. (1965). *Survey Sampling.* New York: Wiley.

Larson, P., Urquhart, S., Herlihy, A., Rosenbaum, B., Paulsen, S. *A spatially balanced, random stream site selection process for regional scale stream surveys.* In preparation.

Lessler, J.T., Kalsbeek, W.D. (1992). *Nonsampling Error in Surveys.* New York: Wiley.

Musser, O. and Mergerson, J.W. (1994) *Estimating List Frame Duplication.* National Agricultural Statistics Service, Washington, D.C.

Nusser, S.M., Goebel, J.J. (1997) *The National Resources Inventory: a long-term multi-resource monitoring programme.* Accepted in Environmental and Ecological Statistics.

Overton, W.S., White, D., Stevens, D.L. (1990). *Design Report for EMAP, Environmental Monitoring and Assessment Program.* EPA/600/3-91/053. U.S. Environmental Protection Agency, Washington, D.C.

Powell, D.S., McWilliams, W.H., Birdsey, R.A. 1994. *History, change, and the U.S. Forest Inventory.* Journal of Forestry 92(12): 6-11;

Salant, P., Dillman, D. 1994. *How to Conduct Your Own Survey.* New York: Wiley.

Stefanski, L.A., Bay, J.M. 1996. *Simulation Extrapolation Deconvolution of Fitie Population Cumulative Distribution Function Estimators*, Biometrika, 83,2, 407-417.

United States Department of Agriculture, Statistical Reporting Service, *Scope and Methods of the Statistical Reporting Service,* September 1983.

United States Environmental Protection Agency. 1989. *Handbook of Methods for Acid Deposition Studies: Field Operations for Surface Water Chemistry.* EPA 600/4-89/020. USEPA, Office of Research and Development, Washington, DC. 139 pp.

United States Environmental Protection Agency. 1997. *Field Operations and Methods for Measuring the Ecological Conditions of Wadeable and Nonwadeable Streams.* EPA 600/4-89/020. USEPA, National Health and Environmental Effects Research Laboratory, Western Ecology Branch, Corvallis, OR.