Journal of Environmental Management 78 (2006) 353–361
Journal of Environmental Management
www.elsevier.com/locate/jenvman

Multi-stage sampling for large scale natural resources surveys: A case study of rice and waterfowl

Joshua D. Stafford[a,*], Kenneth J. Reinecke[b], Richard M. Kaminski[a], Patrick D. Gerard[c]

[a]Department of Wildlife and Fisheries, Box 9690, Mississippi State University, Mississippi State, MS 39762, USA
[b]USGS Patuxent Wildlife Research Center, 2524 South Frontage Road, Suite C, Vicksburg, MS 39180, USA
[c]Experimental Statistics Unit, Box 9653, Mississippi State University, Mississippi State, MS 39762, USA

Received 10 December 2004; revised 5 April 2005; accepted 28 April 2005
Available online 9 September 2005

Abstract

Large-scale sample surveys to estimate abundance and distribution of organisms and their habitats are increasingly important in ecological studies. Multi-stage sampling (MSS) is especially suited to large-scale surveys because of the natural clustering of resources. To illustrate an application, we: (1) designed a stratified MSS to estimate late autumn abundance (kg/ha) of rice seeds in harvested fields as food for waterfowl wintering in the Mississippi Alluvial Valley (MAV); (2) investigated options for improving the MSS design; and (3) compared statistical and cost efficiency of MSS to simulated simple random sampling (SRS). During 2000–2002, we sampled 25–35 landowners per year, 1 or 2 fields per landowner per year, and measured seed mass in 10 soil cores collected within each field. Analysis of variance components and costs for each stage of the survey design indicated that collecting 10 soil cores per field was near the optimum of 11–15, whereas sampling >1 field per landowner provided few benefits because data from fields within landowners were highly correlated. Coefficients of variation (CV) of annual estimates of rice abundance ranged from 0.23 to 0.31 and were limited by variation among landowners and the number of landowners sampled. Design effects representing the statistical efficiency of MSS relative to SRS ranged from 3.2 to 9.0, and simulations indicated SRS would cost, on average, 1.4 times more than MSS because clustering of sample units in MSS decreased travel costs. We recommend MSS as a potential sampling strategy for large-scale natural resource surveys and specifically for future surveys of the availability of rice as food for waterfowl in the MAV and similar areas.

Keywords: Design efficiency; Multi-stage sampling; Rice fields; Sample survey; Waterfowl habitat

1. Introduction

Sample surveys are important for monitoring the status and trends of natural resources because complete censuses rarely are possible (Olsen et al., 1999). Examples of large-scale North American surveys include monitoring breeding and wintering populations of waterfowl (Conroy et al., 1988; Smith, 1995), grassland and forest bird populations (Link and Sauer, 1999), water quality (Gilliom et al., 1995), and land use (Nusser et al., 1998). Effective surveys provide reliable data on resource status and trends and are linked to processes for making management decisions (Nichols et al., 1995).

Sampling theory offers many strategies for designing and conducting surveys that yield precise estimates (Cochran, 1977; Särndahl et al., 1992; Lohr, 1999). Yet, natural resource scientists often apply simple and less robust methods, such as convenience sampling, haphazard sampling, and indices (Eberhardt and Thomas, 1991; Morrison et al., 2001:59–60). Employing inappropriate sampling strategies may result in imprecise or biased estimates, leading to inconclusive or incorrect management decisions and inefficient use of resources (Buckland, 1994; Anderson, 2001). Articulating an explicit objective is necessary for designing an efficient survey and often requires consensus among statisticians, ecologists, and managers responsible for applying results.

Here, we focus on the mean mass (kg/ha) of rice seed available in harvested rice fields because it has been defined

* Corresponding author. Address: Illinois Natural History Survey, Bellrose Waterfowl Research Laboratory, Forbes Biological Station, P.O. Box 590, Havana, IL 62644, USA. Tel.: +1 309 543 3950; fax: +1 309 543 4999.
E-mail address: stafford@inhs.uiuc.edu (J.D. Stafford).

0301-4797/$ - see front matter
doi:10.1016/j.jenvman.2005.04.029

as an important resource for waterfowl and linked to decisions about foraging habitat conservation (Reinecke et al., 1989; Reinecke and Loesch, 1996). Current plans for conserving North American bird populations are implemented by joint venture partnerships (e.g. Loesch et al., 1995) seeking to create sustainable habitat landscapes within ecoregions (Council for Environmental Cooperation, 1997). In the Mississippi Alluvial Valley (MAV) ecoregion, the Lower Mississippi Valley Joint Venture (LMVJV) uses biological models of waterfowl populations and habitats to guide conservation (Reinecke and Loesch, 1996). Because the biological models predict populations will benefit most from increases in foraging habitat, one of the LMVJV conservation strategies is to encourage farmers to flood harvested rice fields in winter to make waste seed available. Thus, estimating rice seed abundance in fields is critical to evaluating conservation strategies because management objectives are directly related to mass of rice seed per unit area. Current management objectives are based on estimates of seed mass from the 1980s (Reinecke et al., 1989) and now may be inaccurate because of recent changes in agricultural practices (Manley et al., 2004; Stafford et al., 2005).

Multi-stage sampling (MSS) has been used for surveys in forestry (Gove et al., 2002), fisheries (Allen et al., 2002), and land-use (Nusser et al., 1998) and is especially appropriate for large-scale natural resources surveys because of their inherent spatial structure (e.g. hierarchically nested habitats; Conroy and Smith, 1994). For example, each ricefield may be considered as a cluster of plots in which seed mass potentially can be measured, each farmer manages several fields, and the landscape contains a population of rice farmers. MSS involves selecting samples sequentially within each stage of sampled populations. MSS often is used when a list of first-stage elements (e.g. landowners) is available, but it is impossible or cost prohibitive to construct a complete list of sample units (e.g. plots within fields) for simple random sampling (SRS). The statistical efficiency (i.e. precision of estimates for a fixed n) of MSS always is less than that of SRS because elements within clusters (e.g. plots within fields) are partially correlated and not assumed independent as in SRS (Thompson, 1992). Nevertheless, MSS can provide more statistical precision per dollar invested if decreases in the cost of measuring plots in clusters compensate for decreases in statistical efficiency. Methods are available to design MSS that minimize costs for estimates of specified precision or maximize precision for fixed costs, provided estimates of variance and cost are available for each stage in the design (e.g. Cochran, 1977:285–288).

Additional research is needed on the design and cost efficiency (e.g. Sargeant et al., 2003) of large-scale surveys as natural resource conservation moves toward landscape and ecosystem management (Boyce and Haney, 1997). Our general objective was to design and evaluate a 3-year (2000–2002) MSS survey to estimate mass of rice seed lost during harvest (i.e. waste rice) and potentially available in late autumn (i.e. late November-early December) as food for waterfowl in the MAV. Our specific objectives were to: (1) describe the design and implementation of a large-scale MSS; (2) compare the statistical and cost efficiency of MSS and SRS; and (3) investigate options for increasing efficiency of the MSS design and consider implications for similar surveys.

2. Materials and methods

2.1. Study area

The MAV or Mississippi Delta is a 10-million ha floodplain of the lower Mississippi River and extends south from Cape Girardeau, Missouri to the Louisiana gulf coast (Fig. 1). Originally the largest forested wetland in North America, the MAV now is a matrix of croplands containing fragments of forested wetlands (Twedt and Loesch, 1999). Rice is an important crop in the MAV and, during 1998–2002, farmers in the MAV produced 63% of the U.S. rice harvest (National Agricultural Statistics Service, 1998–2003). Approximately 10% (Uihlein, 2000) of rice farmers in the MAV flood harvested fields in winter to attract waterfowl, reduce soil erosion, suppress winter weeds, or enhance rice straw decomposition (Manley et al., 2004,2005; Anders et al., 2005).

2.2. Survey design and data collection

We used a stratified 3-stage MSS to estimate mean mass (kg/ha) of rice seed in harvested fields in the MAV. Primary sample units were landowners who farmed rice and flooded fields in winter to provide waterfowl habitat; secondary sample units were rice fields; and tertiary sample units were soil cores collected from fields to measure rice seed mass.

The target population for our inferences was the 80,000 ha (Uihlein, 2000) of harvested rice fields flooded by landowners in winter. Because no complete list of landowners existed, we sampled from a database (i.e. the sampled population) of landowners who enrolled in cooperative conservation programs with Ducks Unlimited, Inc., (Stafford et al., 2005). Each year, we used PROC SURVEYSELECT in SAS v8.2 (SAS Institute, 1999) to select a pre-determined number of landowners (primary sample units) from the database. We selected landowners randomly and with replacement and, to ensure geographic representation, we stratified by state (Arkansas, Louisiana, Mississippi, and Missouri) and allocated samples to states proportional to area of harvested rice (e.g. Cochran, 1977:89).

Annually, we visited each randomly selected landowner to obtain permission to collect samples and select fields (secondary units) from those available. We randomly sampled 1 or 2 fields per landowner in 2000 depending on availability, whereas we sampled 2 fields per landowner

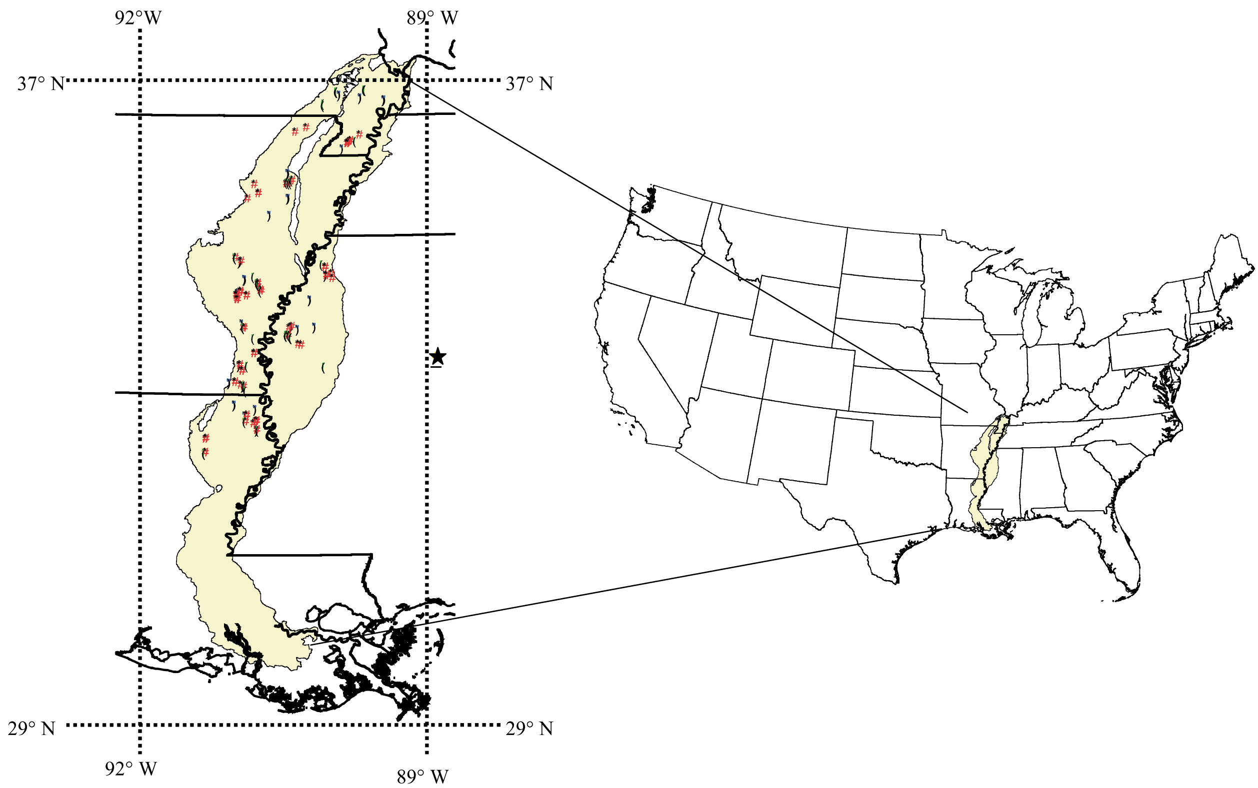

Fig. 1. The lower Mississippi Alluvial Valley, a major rice growing area of the United States, where a multi-stage sample survey was done to estimate abundance of rice remaining after harvest as food for wintering waterfowl. Symbols indicate the starting point for surveys (black star) and locations of sites sampled in early winter 2000 (green circles), 2001, (red triangles), and 2002 (blue squares).

in 2001 and 2002. Within selected fields, we collected 10 soil cores (tertiary sample units) with a depth and diameter of 10 cm each year during 29 November–7 December. We washed soil samples through sieves (mesh sizes 4 [4.75 mm], 10 [2.0 mm], and 18 [1.0 mm]), and separated rice seeds from soil and organic matter. We dried seeds at 87 °C and measured sample mass to the nearest 0.001 g (Manley et al., 2004).

2.3. Parameter estimation

We used PROC SURVEYMEANS (SAS Institute, 1999) to estimate annual means and variances for the mass (kg/ha) of rice seed for waterfowl. For MSS, PROC SURVEYMEANS uses Taylor series linearization (the Delta method; Seber, 1982:7) to estimate variances (SAS Institute, 1999:3200) and, because the proportion of primary sample units we selected was small, variances of the annual means depended primarily on variation among primary unit means (e.g. Cochran, 1977:288). PROC SURVEYMEANS uses sample weights derived from sample selection probabilities to estimate parameters for survey data (SAS Institute, 1999:3275). For our design, the sample weight for each measure of seed mass (tertiary sample unit) was the inverse of the product of the separate probabilities of selecting a given landowner, field, and soil core. The probability of selecting a landowner was n_i/N_i, where n_i and N_i were the numbers of landowners selected and enrolled each year in stratum (state) i. Similarly, the conditional probability of selecting a field from a selected landowner was m_j/M_j, where m_j was the number of fields (1 or 2) selected among M_j fields enrolled by landowner j. Finally, the probability of selecting a soil core within a selected field was $10/(K_{ij}/8.107\times10^{-7}$ ha), where the number of cores collected in each field was 10 and the potential number of cores was the area (K_{ij}) of rice field j within landowner i in hectares divided by the area of a core sample in hectares.

2.4. Statistical efficiency

Efficiency of sample designs is typically measured relative to the performance of SRS. For each year of sampling, we calculated a design effect ($\hat{deff}$), which was the ratio of the variance of a statistic (e.g. estimated mean) obtained with a candidate design to the variance expected with SRS (Cochran, 1977:85). We calculated ratios using variances of estimated means from PROC SURVEYMEANS for MSS and, for SRS, estimated variances of estimated means using PROC MEANS but weighted observations to account for effects of stratification and clustering (Verma et al., 1980:444). We also considered estimating variances of means for SRS by resampling data with replacement, but results were similar (J. D. Stafford, unpublished data). We then estimated the effective sample size (*EFFn*) for MSS each year; *EFFn* is the size of a SRS that provides precision equal to an alternative design (e.g. MSS). We calculated *EFFn* as $n/\hat{deff}$ where n was the number of tertiary sample units (i.e. core samples) used for MSS. Estimates of means and variances of rice abundance were not corrected for bias associated with seeds missed during sample processing (*sensu* Stafford et al., 2005) because coefficients of variation (and, thus, cost ratios; see below) did not differ between corrected and uncorrected estimates (Stafford, 2004:70).

2.5. Cost Efficiency

We developed functions to estimate relative costs of MSS and SRS following the logic of Cochran (1977:244). Typically, cost functions include variable and fixed costs, with variable costs depending on sampling effort and fixed costs assumed equal between candidate designs (Skalski and Millspaugh, 2002). We expressed total variable costs (derived from 2000 to 2002 U.S. currency values), C_t, of surveys as

$$C_t = C_l + C_e + C_v + C_p.$$

The cost of visiting landowners, C_l, to obtain permission to sample and select fields was

$$150.00(n/5) + 0.224d,$$

where \$150.00 was the average daily cost of salary, lodging, and meals for 1 employee, n the number of landowners sampled annually, d distance driven (km) to visit landowners, 5 was the mean number of landowners visited per day, and \$0.244 the cost per km for vehicle rental. We estimated costs for employees, C_e, during field sampling as \$275.00t, where \$275.00 was the daily cost of salaries, lodging, and meals for two workers and t was the time (days) required to collect samples. Costs for vehicle rental during actual sampling, C_v, were \$0.224d (variables previously defined). Finally, costs for processing samples (C_p; separating rice seeds from soil, drying, and weighing) were \$1.375m, where \$1.375 was the cost to process a sample (calculated from an hourly wage of \$5.50 US and an average processing time of 4 samples per hour; J. D. Stafford, unpublished data), and m the total number of samples processed. After simplifying and rearranging expressions, the total cost for completing a survey was

$$C_t = 30n + 275t + 0.448d + 1.375m.$$

We had 3 years of observed costs for MSS to determine if estimates from the cost function were reasonable but no observed costs for SRS. Therefore, we simulated a SRS and derived optimal (i.e. minimal time) travel routes among sample sites to standardize cost comparisons between MSS and SRS. To compare costs between survey designs of equal precision, we used PROC SURVEYSELECT (SAS Institute, 1999) to draw SRS of size *EFFn* each year from the landowner database by selecting fields proportional to size (i.e., area in ha) and with replacement so that samples drawn under SRS might include >1 soil core per field.

We used the Network Analyst extension of ArcView v3.2 (Environmental Systems Research Institute, 1996) to estimate travel times and distances as input to the cost function. To accomplish this estimation, we overlaid data layers representing locations (Zone 15 UTM coordinates) of sampled fields on a road network created from 2002 TIGER/Line Files (U.S. Census Bureau, 2003). Network Analyst minimized travel time among road segments using attribute data indicating speed of travel. Simulated travel routes began at Mississippi State University, Starkville, Mississippi, USA and followed an optimal route among landowners and fields before returning to the point of origin. We added times required for travel and sampling sequentially to determine the number of 8-h days needed to complete surveys.

2.6. Options for improving the design

We investigated two options for improving our MSS design. First, we used estimates of variance among landowners (primary units) to determine relations between number of landowners sampled and the precision of annual estimated means. Second, we estimated cost and variance components for each stage of the design to investigate allocation of sample effort to the second and third stages.

To understand the relationship between the number of landowners sampled and precision expressed as the coefficient of variation (CV), we first estimated annual mean rice abundance (kg/ha) using PROC SURVEYMEANS (SAS Institute, 1999). Then, because variance of annual estimates was determined by variation among landowner means, we used PROC MEANS with the WEIGHT variable representing sample weights to estimate landowner means and variances. Finally, we predicted precision of annual estimated means for samples of 10–100 landowners by calculating CVs as $(\hat{var}/n)^{1/2}/\bar{x}$, where $\hat{var}$ represented annual variances among landowner means, n the numbers of landowners, and $\bar{x}$ the annual means.

To investigate allocation of sample effort to the second and third stages of the design (i.e. fields within landowners and soil cores within fields, respectively) we calculated measures of homogeneity for fields within landowners and soil cores within fields and then estimated variance components for all 3 stages of the design. Because clusters were unequal in size, we assessed homogeneity using analysis of variance (ANOVA) and an adjusted R^2 statistic (R_a^2; Lohr, 1999:140) calculated as $R_a^2 = 1-(MS_{\text{within}}/MS_{\text{total}})$. We conducted 2 ANOVAs for each year. In the first, MS_{within} was the pooled mean square for soil cores within fields and MS_{total} the mean square among all cores. In the second, MS_{within} was the pooled mean square for fields within landowners and MS_{total} the mean square among all fields. To estimate variance components for each stage of the sample design and year, we used PROC VARCOMP and a method-of-moments analysis with Type I sums of squares (Milliken and Johnson, 1992:419; SAS Institute, 1999:3631) to analyze data for all landowners where 2 fields were sampled. This analysis produced negative or low measures of variance for fields within landowners and precluded joint estimation of optimal sample size for the second and third stages. As an alternative, we assumed the data were collected from a 2-stage design (soil cores within landowners) and used variance components from the reduced model (Milliken and Johnson, 1992:262) to determine the optimal number (m_{opt}) of soil samples to collect per field. We calculated m_{opt} as (Cochran, 1977:281)

$$m_{opt} = \frac{S_2}{\sqrt{S_1^2 - S_2^2/M}}\sqrt{\frac{c_1}{c_2}}$$

where S_1^2 was the variance among landowner means, S_2^2 the pooled variance within landowners, $c_1 = 120$ the cost (minutes) of sampling an additional landowner, $c_2 = 3$ (minutes) the cost of collecting an additional soil core within landowners, and M the mean number of potential soil cores per field (i.e., mean field size divided by area of a soil core).

3. Results

3.1. Sampling effort

We completed 3 surveys using the MSS design to estimate mean mass of rice seed available to waterfowl during 27 November–7 December 2000–2002. The number of landowners selected varied among years from 25 to 35, the number of fields sampled from 40 to 69, and the number of soil cores collected from 400 to 690 (Table 1). Estimates of mean mass of rice seed available to waterfowl were 104.6 kg/ha (SE = 24.7) in 2000, 49.1 kg/ha (11.1) in 2001, and 59.2 kg/ha (18.5) in 2002.

3.2. Statistical efficiency

As expected, the lack of independence among measurements in MSS decreased statistical efficiency relative to SRS and estimates of design effect ($\hat{deff}$) ranged from 3.2 to 9.0 (Table 1). The effective sample size (*EFFn*), or number of soil cores required for a SRS equal in precision to MSS, varied inversely with $\hat{deff}$ and ranged from 44 when $\hat{deff}$ was greatest to 157 when it was least. Generally, simulated SRS required collecting only a single soil core from most landowners because the ratio of *EFFn* to the number of landowners in SRS increased only from 1.01 to 1.04 with increases in *EFFn* of 44 to 157 (Tables 1, 2).

3.3. Cost efficiency

Total costs (C_t) of simulated SRS increased with the number of landowners in the sample and varied among years from \$3784 to \$9334, whereas C_t of MSS remained relatively stable (Table 2). Total costs of SRS and MSS were nearly equal in 2000 when the $\hat{deff}$ was 9.0 (Table 1), whereas costs of SRS were 1.26–2.08 times the cost of MSS in 2001 and 2002 when $\hat{deff}$ decreased (Table 2). On average, the cost of SRS was 1.40 times the cost of MSS. The cost of processing samples (C_p) was the smallest cost component (Table 2). Costs of contacting landowners (C_l), employee salaries (C_e), and vehicle rental (C_v) all increased with the number of landowners in SRS and accounted for most differences between SRS and MSS (Table 2).

3.4. Options for improving the design

Initially, we hoped to achieve a CV ≤ 0.15 for annual estimates of mean rice seed mass. However, estimates of variation among landowner means (primary units) indicated the number of landowners needed to achieve a CV = 0.15 was approximately 55 in years 2000–2001 and >100 in 2002 (Fig. 2). Generally, gains in precision were large when the number of landowners was less than the 25–35 we sampled, but decreased rapidly thereafter (Fig. 2).

Estimation of variance components indicated variation among landowners (13.8–24.0%) and among soil cores within fields (76.0–84.4%) accounted for most of the total variance (Table 3). Measures of homogeneity among soil cores within fields were positive and relatively low (0.13–0.22; Table 3).

Table 1
Variances, coefficients of variation (CV), design effects ($\hat{deff}$), and effective sample sizes (*EFFn*) for estimates of mean rice seed mass available as food for waterfowl from a 3-year (2000–2002) multi-stage survey (MSS) in the Mississippi Alluvial Valley, USA, and from hypothetical simple random samples (SRS)

Year	MSS					SRS			$\hat{deff}$[d]	*EFFn*[e]
	n_l[a]	n_f[b]	n_c[c]	$v\hat{a}r(\bar{x})$MSS	CV	n_l[a]	$v\hat{a}r(\bar{x})$SRS	CV		
2000	27	40	400	610.1	0.24	42	67.7	0.08	9.0	44
2001	35	69	690	123.2	0.23	100	18.4	0.09	6.7	103
2002	25	50	500	342.3	0.31	151	107.7	0.18	3.2	157

[a] n_l = number of primary sample units (landowners).
[b] n_f = number of secondary sample units (fields within landowners).
[c] n_c = number of tertiary sample units (soil cores within fields).
[d] $\hat{deff} = v\hat{a}r(\bar{x})_{MSS}/v\hat{a}r_{SRS}$
[e] $EFFn = n_c/\hat{deff}$.

Table 2
Estimated partial and total costs for using multi-stage sampling (MSS) and simple random sampling (SRS) to estimate mean rice seed mass available as food for waterfowl with equal precision in the Mississippi Alluvial Valley, USA, 2000–2002

Year	MSS					SRS					Cost Ratio
	n[a] (C_l)	t[b] (C_e)	m[c] (C_p)	d[d] (C_v)	C_t[e]	n[a] (C_l)	t[b] (C_e)	m[c] (C_p)	d[d] (C_v)	C_t[e]	SRS:MSS
2000	27	8	390	2007		42	5	44	2429		0.85
	$810	$2200	$536	$899	$4445	$1260	$1375	$61	$1088	$3784	
2001	35	10	690	2155		100	9	103	3536		1.26
	$1050	$2750	$949	$965	$5714	$3000	$2475	$142	$1584	$7201	
2002	25	8	500	1881		151	10	157	4103		2.08
	$750	$2200	$688	$843	$4480	$4530	$2750	$216	$1838	$9334	
Mean					$4880					$6772	1.40

[a] n = number of primary sample units (landowners); C_l = cost (\$30n) of contacting landowners.
[b] t = time in days required to collect samples; C_e = cost (\$275t) for employees to collect samples.
[c] m = number of tertiary sample units (soil cores); C_p = cost (\$1.375m) for processing samples.
[d] d = estimated total driving distance in km; C_v = cost (\$0.448d) for vehicle rental.
[e] Total cost $C_t = 30n + 275t + 0.448d + 1.375m$.

In contrast, we were unable to obtain estimates of the variance among fields within landowners, and high measures of homogeneity among fields (0.56–0.80; Table 3) indicated sampling 2 fields per landowner may not have been an effective strategy. When we analyzed the data as a 2-stage design (soil cores within landowners), estimates of variance and costs for sampling landowners and soil cores indicated the optimal number of soil cores to collect per field (m_{opt}) varied among years from 11 to 15.

4. Discussion

Data from the survey we conducted may benefit waterfowl conservation planning in the MAV and elsewhere in North America. Food abundance in this region is believed to limit waterfowl populations in winter (Reinecke and Loesch, 1996), and rice fields are a critical source of foraging habitat (Uihlein, 2000). Previous studies of the abundance of rice as food for waterfowl in the MAV have been limited in spatial scope (Reinecke et al., 1989; Manley et al., 2004). Stafford et al. (2005) used our survey design to obtain data representative of the entire MAV and showed that decreases in rice available to waterfowl in recent decades will necessitate modification of objectives for managing waterfowl habitat (e.g. increase area of foraging habitats other than rice). Our sample design also can be useful in other areas of North America where rice fields provide critical winter habitat for waterfowl. The Central Valley of California (Heitmeyer et al., 1989) and rice prairies in southwest Louisiana and southeast Texas (Hobaugh et al., 1989) are areas where current and representative data on rice available to waterfowl are needed to understand adequacy of waterfowl food resources in winter (Miller and Newton, 1999).

Designing successful surveys requires planning, evaluation, and iterative improvements (Buckland, 1994:149). In MSS, the ideal scenario is that measurements in clusters at the second and subsequent stages of the design are independent rather than correlated. In the context of our study, measurements from soil cores in each field ideally would exhibit the full range of variation in the population, and means for fields within landowners would be no more alike than fields selected at random. A potential shortcoming of our design was that means for fields within landowners were strongly correlated ($R_a^2 = 0.56$–0.80; Table 3). The R_a^2 statistic is similar to an intraclass correlation coefficient (ρ), and values of $\rho > 0.5$ are unusually high (Cochran, 1977:282). An option for improving the design was to sample only 1 field per landowner, thereby increasing the number of landowners and decreasing the standard error of estimates. We opted not to do this for 2 reasons. First, in 2000 we obtained an unusually high estimate of rice abundance for a landowner where only 1 field was sampled (J. D. Stafford, unpublished data). This may have been real variation or sampling error, but we decided to continue sampling 2 fields per landowner to prevent an extreme field mean from determining a landowner mean and inflating the variance of annual

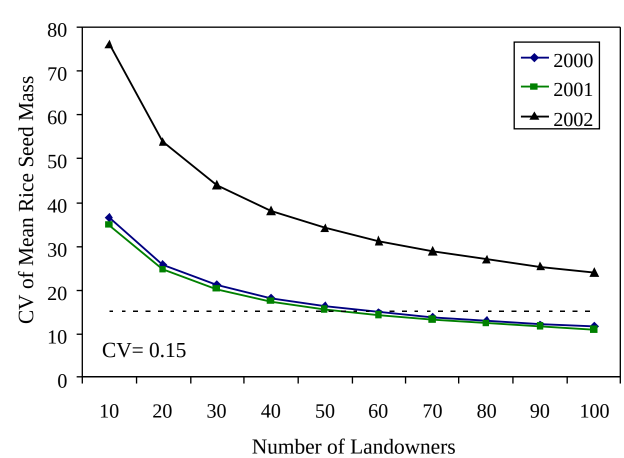

Fig. 2. Predicted relations between coefficients of variation (CV) of mean rice seed mass available as food for waterfowl and number of primary sample units (landowners) selected in a multi-stage sample survey in the Mississippi Alluvial Valley, USA, 2000–2002.

Table 3
Components of variance and estimates of homogeneity R_a^2 for measurements of rice seed available as food for waterfowl from a multi-stage sample survey in which landowners were primary sample units, rice fields within landowners were secondary units, and soil cores within rice fields were tertiary units, Mississippi Alluvial Valley, USA, 2000–2002

Year	Landowners		Fields			Soil Cores		
	Variance	%	Variance	%	R_a^2	Variance	%	R_a^2
2000	9468	24.0	−590	()	0.80	29,907	76.0	0.22
2001	2213	13.8	297	1.8	0.56	13,584	84.4	0.15
2002	17,271	15.9	−3402	()	0.75	91,484	84.1	0.13

estimates. Second, increasing the number of landowners by sampling only 1 field per landowner would provide limited overall benefits because the square-root of the number of landowners is inversely proportional to the CV of estimates (Fig. 2) but directly proportional to travel costs (Cochran, 1977:283; J. D. Stafford, unpublished data).

In contrast to the high correlations among fields within landowners, correlations among soil cores within fields were consistently low (R_a^2 = 0.13–0.22; Table 3). The optimal number of soil cores to collect per field (m_{opt} = 11–15) was slightly greater than the 10 we collected. However, small deviations from optima generally have little effect on precision (Cochran, 1977:281), and precision of estimates depends primarily on variation among primary sample units when the proportion of primary units selected is small (often the case in MSS; Cochran, 1977: 288). The only other design variable we could manipulate to increase precision of annual estimates was the number of landowners (primary units) selected. Inherent variation among landowner means limited precision of estimates for samples of 25–35 landowners to CVs = 0.23–0.31 (Table 1). Factors responsible for variation among landowners and lack of variation among fields within landowners probably were similar. Stafford (2004) attributed variation among landowners to such local effects as storms damaging crops prior to harvest, landowner harvest equipment or operator efficiency, and seed consumption by birds and rodents. In situations such as ours, where budget constraints prevented achieving desired precision within years, an option worth noting was that precision of estimates sometimes may be increased by treating surveys from multiple years as replicates (Neter et al., 1985:4). However, pooling over years may not result in significant precision gains if estimates vary greatly among years (Williams et al., 2002:45). In sum, we recommend future surveys use a MSS design to estimate abundance of rice as food for waterfowl, collect ≥10 soil cores from 1 field per landowner, and determine the number of landowners to sample from requirements for precision or budget constraints. Sampling a second field per landowner is an option that incurs relatively low cost but provides marginal gains in precision.

Despite high correlation among fields within landowners and some correlation among soil cores within fields, MSS generally provided more precision per dollar invested than simulated SRS. The cost of SRS was competitive with MSS only in 2000 when values for $\hat{deff}$ (Table 1) and homogeneity within and among fields (Table 3) were greatest. Otherwise, the cost of SRS was 1.3–2.1 times the cost for MSS of equal precision (Table 2). However, we recommend caution about generalizing the advantages of MSS over SRS because their relative efficiencies are sensitive to changes in cost and variance components. In our study, costs for laboratory processing of samples were relatively low (C_p; Table 2) and, in surveys where final data are recorded in the field (e.g. tree height or diameter), no processing costs are incurred. These scenarios favor MSS where more measurements are needed because they are not independent. In contrast, high costs for analyzing samples (e.g. chemical extractions) favor SRS, because decreases in the number and total cost of measurements resulting from statistical independence in SRS can exceed travel costs saved with MSS. Reducing travel costs by collecting multiple measurements at selected locations is primarily responsible for the cost effectiveness of MSS. Travel costs were especially high for our study because we visited landowners (C_l; Table 2) to negotiate land access and select fields prior to sampling. Survey designs that account for a cost component related to land access may be more common in the future as interest increases in the role of private lands in ecosystem management (Boyce and Haney, 1997; Hilty and Merenlender, 2003).

5. Conclusions

Data from sample surveys are critical to large-scale management of natural resources (Gilliom et al., 1995; Smith 1995; Nusser et al., 1998). Effective surveys require clear objectives, relevance to management decisions, and efficient designs (Nichols et al., 1995). In turn, efficient survey designs require estimates of cost and variance components, evaluation, and iterative improvement. Too often, researchers use sample designs that are familiar but inappropriate (Anderson, 2001; Cassell and Rousey, 2003), despite availability of designs that provide unbiased estimates while minimizing cost or maximizing precision (Cochran, 1977; Särndahl et al., 1992; Lohr, 1999). MSS often is an appropriate design for large-scale surveys because it exploits the natural clustering of objects (e.g. plots in fields in farms; Conroy and Smith, 1994). Regardless of the sampling design chosen, researchers should acquaint themselves with sampling theory, seek biometric expertise to aid in selecting an appropriate design, and conduct pilot sampling when possible prior to survey implementation (Buckland, 1994; Conroy and Smith, 1994).

The MSS survey we designed has provided important information about waterfowl food resources and affected habitat conservation planning in the MAV. If future surveys are needed in the MAV or other regions where rice in harvested fields is important as waterfowl food, we recommend using a MSS design where landowners are primary sample units, 1 (or at most 2) fields are sampled per landowner, and ≥ 10 soil cores are collected per field to measure seed abundance. The number of landowners to sample can be determined from variation in mean rice abundance among landowners and the desired precision of estimates.

Acknowledgements

Principal and initial funding for our research was provided by Ducks Unlimited, Inc., through the Institute for Wetland and Waterfowl Research, and the North American Wetlands Conservation Council. Support was also provided by Federal Aid in Wildlife Restoration through the Mississippi Department of Wildlife, Fisheries and Parks (Project W-48, Study 44), the Arkansas Game and Fish Commission, Mississippi State University (MSU) Forest and Wildlife Research Center (FWRC) and Delta Research and Extension Center, and the U.S. Geological Survey, Patuxent Wildlife Research Center. We are indebted to the rice farmers in the MAV who allowed us to sample their fields. We could not have completed this project without help from many individuals, including B. Batt, M. Checkett, B. Dugger, D. Godwin, K. Hartke, D. Jackson, B. Leopold, M. Kurtz, M. Petrie, T. Moorman, D. Rutka, B. Uihlein, and R. Wilson. Finally, we thank the following people who assisted with field, lab, and/or analytical work, especially J. Arnold, B. Bowers, C. Galle, B. Holder, N. Jerstad, J. Kross, A. Pearse, E. Penny, E. Seckinger, M. Stafford, J. Vest, T. Whyte, and C. Woods. Our manuscript was approved for publication as MSU-FWRC Journal Article WF-208.

References

Allen, M., Kilpatrick, D., Armstrong, M., Briggs, R., Course, G., Pérez, N., 2002. Multistage cluster sampling design and optimal sample sizes for estimation of fish discards from commercial trawlers. Fisheries Research 55, 11–24.

Anders, M.M., Windham, T.E., Reinecke, K.J., McNew, R.W., Robinson, J.F., 2005. The effects of fall rice stubble management and winter flooding on subsequent conventional and no-till irrigated and dryland soybeans. Journal of Sustainable Agriculture.

Anderson, D.R., 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29, 1294–1297.

Boyce, M.S., Haney, A., 1997. Ecosystem Management: Applications for Sustainable Forest and Wildlife Resources. Yale University Press, New Haven, CT.

Buckland, S.T., 1994. The principles and practice of large-scale wildlife surveys. Transactions of the North American Wildlife and Natural Resources Conference 59, 149–158.

Cassell, D.L., Rousey, A., 2003. Complex sampling designs meet the flaming turkey of glory, SAS User's Group International Paper 262–28 2003.

Cochran, W.G., 1977. Sampling Techniques. Wiley, New York.

Conroy, M.J., Smith, D.R., 1994. Designing large-scale surveys of wildlife abundance and diversity using statistical sampling principles. Transactions of the North American Wildlife and Natural Resources Conference 59, 159–169.

Conroy, M.J., Goldsberry, J.R., Hines, J.E., Stotts, D.B., 1988. Evaluation of aerial transect surveys for wintering American black ducks. Journal of Wildlife Management 52, 694–703.

Council for Environmental Cooperation, 1997. Ecological Regions of North America: Toward a Common Perspective. Council for Environmental Cooperation, Montreal, Que..

Eberhardt, L.L., Thomas, J.M., 1991. Designing environmental field studies. Ecological Monographs 61, 53–73.

Environmental Systems Research Institute, 1996. Using ArcView GIS. Environmental Systems Research Institute, Redlands, CA.

Gilliom, R.J., Alley, W.M., Gurtz, M.E., 1995. Design of the National Water-quality Assessment Program: Occurrence and Distribution of Water-quality Conditions. U.S. Geological Survey Circular 1112, Denver, CO.

Gove, J.H., Ducey, M.J., Valentine, H.T., 2002. Multistage point relascope and randomized branch sampling for downed coarse woody debris estimate. Forest Ecology and Management 155, 153–162.

Heitmeyer, M.E., Connelly, D.P., Pederson, R.L., 1989. The central, imperial, and Coachella valleys of California. In: Smith, L.M., Pederson, R.L., Kaminski, R.M. (Eds.), Habitat Management for Migrating and Wintering Waterfowl in North America. Texas Tech University Press, Lubbock, TX, pp. 475–505.

Hilty, J., Merenlender, A.M., 2003. Studying biodiversity on private lands. Conservation Biology 17, 132–137.

Hobaugh, W.C., Stutzenbaker, C.D., Flickinger, E.L., 1989. The rice prairies. In: Smith, L.M., Pederson, R.L., Kaminski, R.M. (Eds.), Habitat management for migrating and wintering waterfowl in North America. Texas Tech University Press, Lubbock, TX, pp. 367–383.

Link, W.A., Sauer, J.R., 1999. Estimating population change from count data: application to the North American Breeding Bird Survey. Ecological Applications 8, 258–268.

Loesch, C.R., Twedt, D.J., Reinecke, K.J., 1995. Conservation partnerships in the lower Mississippi Alluvial Valley. Wildlife Society Bulletin 23, 791–795.

Lohr, S.L., 1999. Sampling: Design and Analysis. Duxbury Press, New York.

Manley, S.W., Kaminski, R.M., Reinecke, K.J., Gerard, P.D., 2004. Waterbird food values of winter-managed rice fields in Mississippi. Journal of Wildlife Management 68, 74–83.

Manley, S.W., Kaminski, R.M., Reinecke, K.J., Gerard P.D., 2005. Agronomic implications of waterfowl management in Mississippi ricefields. Wildlife Society Bulletin 33, in press.

Miller, M.R., Newton, W.E., 1999. Population energetics of northern pintails wintering in the Sacramento Valley, California. Journal of Wildlife Management 63, 1222–1238.

Milliken, G.A., Johnson, D.E., 1992. Analysis of messy data, Volume I: Designed Experiments. Chapman & Hall, New York.

Morrison, M.L., Block, W.M., Strickland, M.D., Kendall, W.L., 2001. Wildlife Study Design. Springer, New York.

National Agricultural Statistics Service, 1998–2003. Crop Production Summaries, 1998–2003. Department of Agriculture, National Agricultural Statistics Service, Washington, DC.

Neter, J.W., Wasserman, W., Kutner, M.H., 1985. Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Designs. R.D. Irwin, Homewood, IL.

Nichols, J.D., Johnson, F.A., Williams, B.K., 1995. Managing North American waterfowl in the face of uncertainty. Annual Review of Ecology and Systematics 26, 177–199.

Nusser, S.M., Breidt, F.J., Fuller, W.A., 1998. Design and estimation for investigating the dynamics of natural resources. Ecological Applications 8, 234–245.

Olsen, A.R., Sedransk, J., Edwards, D., Gotway, C.A., Liggett, W., Rathbun, S., Reckhow, K.H., Young, L.J., 1999. Statistical issues for monitoring ecological and natural resources in the United States. Environmental Monitoring and Assessment 54, 1–45.

Reinecke, K.J., Loesch, C.R., 1996. Integrating research and management to conserve wildfowl (Anatidae) and wetlands in the Mississippi Alluvial Valley, USA. Gibier Faune Sauvage, Game and Wildlife 13, 927–940.

Reinecke, K.J., Kaminski, R.M., Moorehead, K.J., Hodges, J.D., Nassar, J.R., 1989. Mississippi Alluvial Valley. In: Smith, L.M., Pederson, R.L., Kaminski, R.M. (Eds.), Habitat Management for Migrating and Wintering Waterfowl in North America. Texas Tech University Press, Lubbock, TX, pp. 203–247.

Sargeant, G.A., Johnson, D.H., Berg, W.E., 2003. Sampling designs for carnivore scent-station surveys. Journal of Wildlife Management 62, 289–298.

Särndahl, C.-E., Swensson, B., Wretman, J., 1992. Model Assisted Survey Sampling. Springer, New York.

SAS Institute, 1999. SAS/STAT User's Guide. SAS Institute, Cary, NC.

Seber, G.A.F., 1982. The Estimation of Animal Abundance and Related Parameters. MacMillan, New York.

Skalski, J.R., Millspaugh, J.J., 2002. Generic variance expressions, precision, and sampling optimization for the sex-age-kill model of population reconstruction. Journal of Wildlife Management 66, 1308–1316.

Smith, G.W., 1995. A critical review of the aerial and ground surveys of breeding waterfowl in North America. U.S. Department of the Interior, National Biological Service Biological Science Report 5.

Stafford, J.D., 2004. Abundance and conservation of waste rice for wintering waterfowl in the Mississippi Alluvial Valley. Dissertation, Mississippi State University, Mississippi State, MS, USA.

Stafford, J.D., Kaminski, R.M., Reinecke, K.J., Manley, S.W., 2005. Waste rice for waterfowl in the Mississippi Alluvial Valley. Journal of Wildlife Management 69, in press.

Thompson, S.K., 1992. Sampling. Wiley, New York.

Twedt, D.J., Loesch, C.R., 1999. Forest area and distribution in the Mississippi Alluvial Valley: implications for breeding bird conservation. Journal of Biogeography 26, 1215–1224.

Uihlein, W.B., 2000. Extent and distribution of waterfowl habitat managed on private lands in the Mississippi Alluvial Valley. Dissertation, Mississippi State University, Mississippi State, MS, USA.

U.S. Census Bureau, 2003. TIGER/Line Files Technical Documentation. U.S. Department of Commerce Census Bureau, Washington, DC.

Verma, V., Scott, C., O'Muircheartaigh, C., 1980. Sample designs and sampling errors for the world fertility survey. Journal of the Royal Statistical Society, Series A 143, 431–473.

Williams, B.K., Nichols, J.D., Conroy, M.J., 2002. Analysis and Management of Animal Populations. Academic Press, San Diego, CA.