J. N. Am. Benthol. Soc., 2005, 24(3):525–537
© 2005 by The North American Benthological Society

Optimal sampling design for estimating spatial distribution and abundance of a freshwater mussel population

PENELOPE S. POOLER[1] AND DAVID R. SMITH[2]

Aquatic Ecology Laboratory, Leetown Science Center, US Geological Survey, 11649 Leetown Road, Kearneysville, West Virginia 25430 USA

Abstract. We compared the ability of simple random sampling (SRS) and a variety of systematic sampling (SYS) designs to estimate abundance, quantify spatial clustering, and predict spatial distribution of freshwater mussels. Sampling simulations were conducted using data obtained from a census of freshwater mussels in a 40 × 33 m section of the Cacapon River near Capon Bridge, West Virginia, and from a simulated spatially random population generated to have the same abundance as the real population. Sampling units that were 0.25 m^2 gave more accurate and precise abundance estimates and generally better spatial predictions than 1-m^2 sampling units. Systematic sampling with ≥2 random starts was more efficient than SRS. Estimates of abundance based on SYS were more accurate when the distance between sampling units across the stream was less than or equal to the distance between sampling units along the stream. Three measures for quantifying spatial clustering were examined: Hopkins Statistic, the Clumping Index, and Morisita's Index. Morisita's Index was the most reliable, and the Hopkins Statistic was prone to false rejection of complete spatial randomness. SYS designs with units spaced equally across and up stream provided the most accurate predictions when estimating the spatial distribution by kriging. Our research indicates that SYS designs with sampling units equally spaced both across and along the stream would be appropriate for sampling freshwater mussels even if no information about the true underlying spatial distribution of the population were available to guide the design choice.

Key words: systematic sampling, quadrat sampling, freshwater mussels, spatial distribution, kriging, estimating abundance.

Successful conservation of a biological population often requires a reliable description of its distribution and abundance. This information is needed to address questions regarding the number of individuals present, the locations of the individuals within a site, and how the abundance or spatial distribution has changed as a result of an impact or management action. Sampling methods for describing spatial distribution and abundance vary widely. Simple random sampling (SRS) is a standard method that provides unbiased estimates of abundance or density regardless of the population's spatial distribution (Thompson 1992). Implementing SRS is cumbersome in a biological survey. However, systematic sampling (SYS) is relatively easy to implement in the field. Moreover, if SYS is implemented using multiple random starts, error can be estimated accurately, regardless of the population's spatial distribution. McArthur (1987) compared the ability of SRS, stratified random sampling, SYS with and without stratification, and unequal probability sampling to predict mean concentration of pollution over a region. His results indicated that stratified SYS was the most efficient of these designs. However, SYS without stratification was the next most efficient method when information on how best to stratify was not available. SYS is preferred for sampling rare, spatially clustered populations in the absence of prior information on distributions (Christman 2000).

Freshwater mussels are an imperiled taxon with projected extinction rates in the same range as tropical species (Ricciardi and Rasmussen 1999). Accurate estimates of their spatial distribution and abundance have obvious benefits to their conservation. Smith et al. (2001) demonstrated the value of estimates of abundance and predictions of spatial distribution for an assessment of endangered species of freshwater mussels at a bridge construction site on the Allegheny River in Pennsylvania. A relatively large sampling effort is required to obtain accurate estimates and spatial predictions because

[1] Present address: Department of Statistics, Mail Code 0439, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0439 USA. E-mail: ppooler@vt.edu

[2] E-mail address: drsmith@usgs.gov

freshwater mussels are often rare and spatially clustered (Downing and Downing 1992, Cawley 1993, Strayer and Smith 2003). Evaluations of simulated sampling designs based on field-collected data can be useful for identifying efficient (less costly) and reliable sampling methods and for promoting application of rigorous methods for sampling freshwater mussel populations.

Designing an efficient survey for freshwater mussels depends on characteristics of the site and population of interest, survey objectives, and the level of precision that is required. If the objective is to estimate abundance accurately and precisely or to predict spatial distribution, then one has a wide array of sampling designs from which to choose. A number of studies have indicated that adaptive sampling or a combination of systematic and adaptive sampling is effective when estimating density of rare or clustered populations (Smith et al. 1995, Thompson and Seber 1996, Christman 2000). However, recent applications to freshwater mussels have indicated that adaptive sampling increases detection of rare species and collection of individual mussels, but does not decrease sampling error compared to conventional sampling (Strayer et al. 1997, Smith et al. 2003).

We compared a variety of SRS and SYS designs using data from a real freshwater mussel population and a simulated spatially random population generated to have the same abundance as the real population to determine how well the designs estimated abundance, quantified the degree of spatial clustering, and predicted spatial distribution. We also considered the effect of sampling-unit size on the performance of each sampling design. Both 1-m^2 and 0.25-m^2 sampling units are used commonly for sampling freshwater mussels (Strayer and Smith 2003). Downing and Downing (1992) recommended the use of 1-m^2 sampling units for freshwater mussel surveys, but clustered or contagious populations may be more efficiently sampled using smaller sampling units (Beall 1939, Finney 1946, Taylor 1953). We make specific recommendations for each of the research goals based on comparisons of the designs.

Methods

Study populations

The data used in our study have been used in other sampling simulation studies (Strayer and Smith 2003, Salehi and Smith 2005). Mussels were censused over 11 d in June and July 1994 on a section of pool-run habitat in the Cacapon River near Capon Bridge, West Virginia. The census required 35.5 h of search time (5 people averaging 7.1 h/person search time). The river section was approximately rectangular and was 33 m wide (bank to bank) by 40 m long. Water depth ranged from 12 to 75 cm (median depth = 52 cm). Dominant substrate types were large gravel and small cobble with interstitial sand. Water clarity was high, so visual observations were possible. A grid of 0.25-m^2 cells was superimposed on the entire surface area to provide spatial references for each mussel. Each cell was searched thoroughly, and all mussels found were measured lengthwise and returned to the substrate.

The population consisted of 1205 mussels belonging to 3 species: 901 *Elliptio complanata* (eastern elliptio), 277 *Elliptio fisheriana* (northern lance), and 27 *Lampsilis cariosa* (yellow lampmussel). In ecological terms, 3 species is a mussel community rather than a population, but the term *real population* is used throughout this manuscript to refer to all mussels at the census location. The real population was spatially aggregated into 2 prominent bands, a larger band in the center of the stream and a smaller off-center band (Fig. 1). Observed mussel locations were recorded to the nearest 0.25 m^2.

A pseudo-simple random sampling algorithm (S-Plus, version 4.5, MathSoft, Seattle, Washington) was used to generate 1205 random locations in an area the same size as the census area (33×40 m). This procedure was done to create a spatially random simulated population that could be used to understand how sampling designs performed when studying a non-clustered population. The simulated population was generated to have the same abundance as the real population, and individuals were tallied within 0.25-m^2 cells.

Survey designs

We compared SRS and SYS designs that included 1 to 6 random starting points. The sampling effort (% of the site sampled) ranged from 0.5% to 10.4% using both 0.25-m^2 and 1-m^2 sampling units. We simulated 20 SRS designs and 198 different SYS designs. Sampling effort (1%, 2%, . . . , 10%) differed for each SRS design for

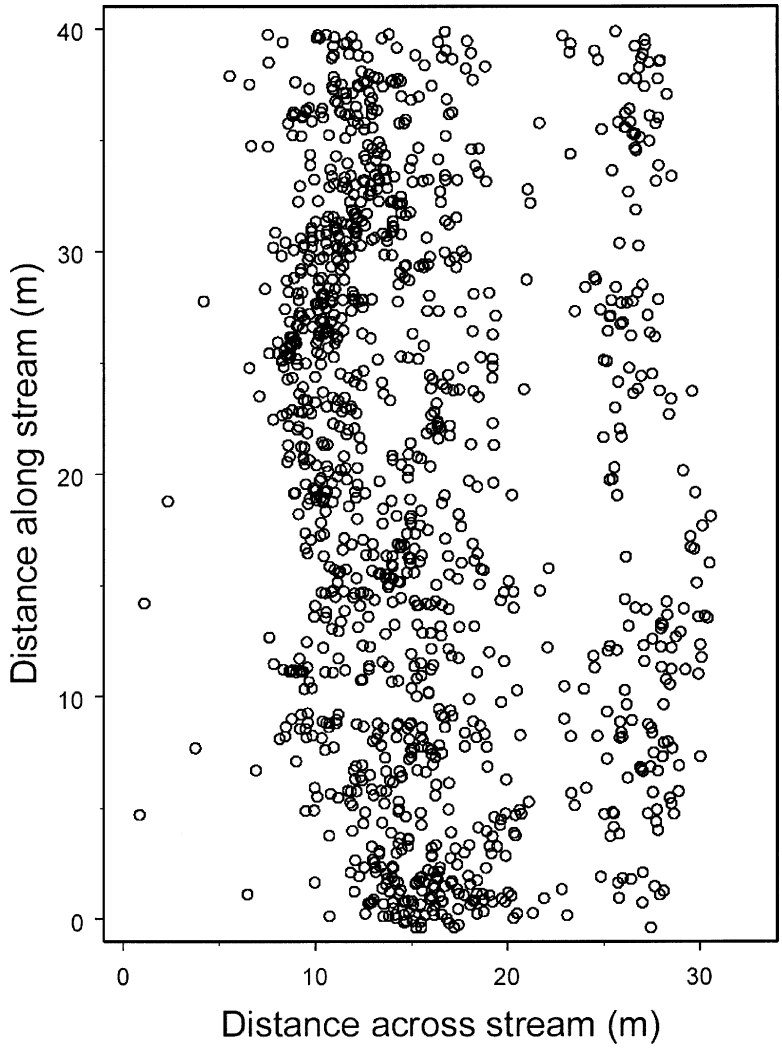

FIG. 1. Distribution of a freshwater mussel population in a pool–run on the Cacapon River near Capon Bridge, West Virginia. Each circle represents the location of an individual mussel.

the 2 sampling-unit sizes (0.25 m^2 and 1 m^2). SYS designs were created by varying the number of random starts and the distance between sampling units across (D_{AC}) and along (D_{AL}) the stream (Table 1). The 198 SYS designs consisted of 125 designs using 0.25-m^2 sampling units and 73 designs using 1-m^2 sampling units. Each of the 218 design options (20 SRS + 198 SYS) was simulated 1000 times. Mean, variance, and skewness were calculated for estimates of abundance and spatial clustering based on the 1000 sampling realizations. Relative efficiency (RE) was used to compare the performances of 0.25-m^2 and 1-m^2 sampling units for the 83 designs that used both sampling-unit sizes. We also used RE to compare SYS and SRS designs with the same level of sampling effort. For both comparisons, RE was calculated as

$$\mathrm{RE} = \frac{\widehat{\mathrm{var}}(\hat{T}_{\mathrm{Design1}})}{\widehat{\mathrm{var}}(\hat{T}_{\mathrm{Design2}})} \qquad [1]$$

where $\widehat{\mathrm{var}}$ is the estimated sampling variance and $\hat{T}$ is the estimated abundance. The SRS design variance was in the numerator when SRS designs were compared to SYS designs. Thus, RE

TABLE 1. Systematic sampling (SYS) designs examined using simulations. Each SYS design was specified by the number of random starts (RS) for a given distance across the stream between sampling units (D_{AC}) and distance along the stream between sampling units (D_{AL}), and by the sampling-unit size (0.25 m^2 or 1 m^2). Percent sampled is the range of values for the % of the site sampled in association with each design. Each design was used to simulate sampling from a real population and a simulated population 1000 times (see text for details). Blank spaces = design not included.

		0.25 m^2		1 m^2	
$D_{AC} \times D_{AL}$	D_{AC}/D_{AL}	RS	% of site sampled	RS	% of site sampled
2 × 8	0.25	1	6.3	1	6.4
8 × 2	4.00	1	6.3	1	7.6
3 × 12	0.25	1, 2, 3	2.8–8.3	1, 2, 3	3.3–8.3
12 × 3	4.00	1, 2, 3	2.8–8.3	1, 2, 3	3.2–8.5
5 × 20	0.25	1, 2, 3, 4, 5	1.0–5.0		
20 × 5	4.00	1, 2, 3, 4, 5	1.0–5.0		
7 × 28	0.25	1, 2, 3, 4, 5, 6	0.5–3.1		
28 × 7	4.00	1, 2, 3, 4, 5, 6	0.5–3.1		
8 × 25	0.32	1, 2, 3, 4, 5, 6	0.5–3.0		
25 × 8	3.13	1, 2, 3, 4, 5, 6	0.5–3.0		
3 × 9	0.33	1	3.7	1	4.2
9 × 3	3.00	1	3.7	1	4.2
4 × 12	0.33	1, 2, 3, 4, 5	2.1–10.4	1, 2, 3, 4, 5	2.7–10.4
12 × 4	3.00	1, 2, 3, 4, 5	2.1–10.4	1, 2, 3, 4, 5	2.3–10.4
6 × 17	0.35	1, 2, 3, 4, 5	1.0–4.9	1, 2, 3, 4, 5	1.4–4.9
17 × 6	2.83	1, 2, 3, 4, 5	1.0–4.9	1, 2, 3, 4, 5	1.1–4.9
4 × 9	0.44	1, 2, 3	2.8–8.3	1, 2, 3	3.4–8.3
9 × 4	2.25	1, 2, 3	2.8–8.3	1, 2, 3	3.0–8.3
5 × 10	0.50	1, 2, 3, 4, 5	2.0–10.0	1, 2, 3, 4, 5	2.1–10.0
10 × 5	2.00	1, 2, 3, 4, 5	2.0–10.0	1, 2, 3, 4, 5	2.4–10.0
7 × 14	0.50	1, 2, 3, 4, 5	1.0–5.1	1, 2, 3, 4, 5	1.1–5.1
14 × 7	2.00	1, 2, 3, 4, 5	1.0–5.1	1, 2, 3, 4, 5	1.4–5.1
10 × 20	0.50	1, 2, 3, 4, 5, 6	0.5–3.0		
20 × 10	2.00	1, 2, 3, 4, 5, 6	0.5–3.0		
3 × 5	0.60	1	6.7	1	6.7
5 × 3	1.67	1	6.7	1	7.4
4 × 4	1.00	1	6.3	1	6.8
5 × 5	1.00	1	4.0	1	4.2
6 × 6	1.00	1, 2, 3	2.8–8.3	1, 2, 3	3.2–8.3
7 × 7	1.00	1, 2, 3, 4, 5	2.0–10.2	1, 2, 3, 4, 5	2.3–10.2
10 × 10	1.00	1, 2, 3, 4, 5	1.0–5.0	1, 2, 3, 4, 5	1.2–5.0
14 × 14	1.00	1, 2, 3, 4, 5, 6	0.5–3.1		

> 1 if the SYS design was more efficient than the SRS design (i.e., if SYS sampling variance was smaller). Variances of designs using 1-m^2 sampling units were in the numerator when identical designs with the 2 different sampling-unit sizes were compared. Thus, RE > 1 if the design using 0.25-m^2 sampling units was more efficient than the same design using 1-m^2 sampling units.

Estimating abundance

Abundance estimates from SRS and SYS designs and sampling variances of those estimates were calculated 2 ways. Method 1 assumed simple random selection of sampling units. Strictly speaking, this method is not appropriate for SYS and will lead to unbiased variance estimates only if the population of sampling units can be described as being in random order. The sampling units contained counts from the real population, so units that were close to each other tended to be similar. Therefore, method 1 should tend to overestimate the variance of the estimator (Thompson 1992). The abundance estimate by method 1 is

$$\hat{T}_{\text{SRS}} = \frac{A}{am}\sum_{j=1}^{m} y_j \qquad [2]$$

and the sample variance of this abundance estimator is

$$\widehat{\text{var}}(\hat{T}_{\text{SRS}}) = A^2\left(\frac{M-m}{Mma^2}\right)\frac{1}{m-1}\sum_{j=1}^{m}(y_j-\bar{y})^2 \qquad [3]$$

where M is the number of sampling units in the study site, m is the number of sampling units in the sample, A is the area of the study site, a is the sampling-unit size, y_j is the number of individuals found in each sampling unit, and $\bar{y}$ is the mean number of individuals/sampling unit.

Method 2 for estimating abundance was used only for SYS designs that included ≥2 random starts. Mean abundance and its sampling variance were calculated using the unbiased estimators available when cluster sampling is used with ≥2 primary units that are randomly selected without replacement (Thompson 1992). The abundance estimate by method 2 is

$$\hat{T}_{\text{SYS}} = \frac{N}{n}\sum_{i=1}^{n}\bar{y}_i, \qquad \bar{y}_i = \sum_{j=1}^{m_i} y_{ij} \qquad [4]$$

where N is the number of primary units in the study site, n is the number of primary units in the sample, m_i is the number of secondary units (i.e., sampling units in the ith primary unit), $\bar{y}_i$ is the mean number of individuals/primary unit, and y_{ij} is the number of mussels found in the jth sampling unit of the ith primary unit. The sample variance of this abundance estimator is

$$\widehat{\text{var}}(\hat{T}_{\text{SYS}}) = \left[\frac{N(N-n)}{n}\right]\frac{1}{n-1}\sum_{i=1}^{n}(\bar{y}_i-\bar{y})^2,$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n}\bar{y}_i. \qquad [5]$$

Confidence interval (CI) coverage for both abundance estimates was found by calculating the 95% CI for each of the 1000 realizations of each of the 218 sample designs. CIs for the abundance estimate and sampling variance were calculated as

$$95\%\ \text{CI} = \hat{T} \pm (t_{1-\alpha/2,df})\sqrt{\widehat{\text{var}}(\hat{T})} \qquad [6]$$

where $\alpha = 0.05$ (probability of a Type I error) and df (degrees of freedom) were determined by the formulae used. For CIs based on equations 2 and 3, which treat the secondary units as if they were selected randomly, $df = m - 1$, the number of secondary units in the sample minus 1. For CIs based on equations 4 and 5, which required ≥2 primary units, $df = n - 1$, the number of primary units in the sample minus 1. The proportion of the 1000 CIs calculated for a given design that contained the true population abundance was the CI coverage. Ideally, the empirical coverage should be equal to nominal coverage, $1 - \alpha$.

Accuracy of abundance estimates from each survey design was compared by calculating relative bias as

$$\text{RB} = (E[\hat{T}] - \tau)/\tau \qquad [7]$$

where $E[\hat{T}]$ is the expected value of the abundance estimate attained from all 1000 sample realizations and τ is the true abundance calculated from the census. Accuracy of variance estimates also was examined for all SYS designs with ≥2 random starts. We compared the variance estimate found using equation 5 to the true variance of abundance found in the census, σ_τ^2 using the following measure of relative bias of variance:

$$\text{RBV} = [\widehat{\text{var}}(\hat{T}_{\text{SYS}}) - \sigma_\tau^2]/\sigma_\tau^2. \qquad [8]$$

Bias of variances was examined because previous research (Hedayat and Sinha 1991) indicated that increasing the number of random starts in a SYS design decreased bias in estimates of variability.

Quantifying degree of spatial clustering

Three statistics that measure spatial clustering, the Hopkins Statistic, Morisita's Index, and the Clumping Index, were calculated. The Hopkins Statistic (Cressie 1993) is

$$H_F = \sum_{i=1}^{p} U_i^2 \Big/ \sum_{i=1}^{p} W_i^2 \qquad [9]$$

where U_i is the distance between the ith randomly chosen point in the survey area and the closest location of an observed mussel and W_i is the distance between the location of the ith randomly chosen observed mussel and the nearest mussel. The p randomly chosen locations of observed mussels are those mussel locations that fall within the bounds of the sampling units of a specific sample. Previous studies have indicated that the Hopkins Statistic is the most effective method of detecting spatial clustering (see Diggle et al. 1976, Byth and Ripley 1980, Upton and

Fingleton 1985, Cressie 1993). Holgate (1965) suggested that the Hopkins Statistic is more sensitive to departures from complete spatial randomness than most nearest-neighbor statistics because clustering simultaneously reduces the distances between observations (W_i) and increases the distances between observations and random points within the site (U_i) (Holgate 1965). The sampling distribution of W_F was approximated by the F distribution for the random population with numerator and denominator $df = 2 \times$ the number of mussels located in each sampling unit (Cressie 1993).

Morisita's Index (Morisita 1959, Cressie 1993) is

$$I_d = m\left\{\frac{\sum_{j=1}^{m}[y_j(y_j - 1)]}{[\hat{T}_{\mathrm{SRS}}\cdot(\hat{T}_{\mathrm{SRS}} - 1)]}\right\} \qquad [10]$$

where m is the number of sampling units in the sample, y_j is the number of individuals found in each sampling unit, and $\hat{T}_{\mathrm{SRS}}$ is estimated abundance. Morisita's Index has been proposed as being unaffected by changes in density (Morisita 1971). This characteristic suggests that surveys of varying levels of effort are likely to yield consistent results, provided a minimum level of effort is attained.

The Clumping Index (Cressie 1993) is

$$\mathrm{IC} = [\mathrm{var}(\hat{T}_{\mathrm{SRS}})/\hat{T}_{\mathrm{SRS}}] - 1. \qquad [11]$$

The clumping index, the most straightforward of the 3 statistics compared here, is a slight variation of the variance/mean ratio. Morisita's Index and the Clumping Index were calculated assuming a simple random sample so that these measures could be calculated for all designs in a consistent manner. As a consequence, differences in bias of these measures for different designs are attributable solely to the locations of the individual mussels found by the respective designs.

Measures that quantify clustering produce values that range from those representing a uniform population to those representing an extremely clumped distribution (Ludwig and Reynolds 1988). Values representing random distributions tend to fall towards the uniform end of the spectrum. Relative bias of measures of clumping was calculated by dividing the difference between the true value (calculated from the real population) and the estimated value for each measure by the true value of the measure. For example, the relative bias of the Hopkins statistic for a given design was calculated as $(H_F - \hat{H}_F)/H_F$. Relative bias for each measure was calculated for each design to understand how different design options affected accuracy in estimating degree of clustering.

Predicting spatial distribution

The effect of survey design on predictions of spatial distribution was examined using a small subset of all survey designs at a fixed level of sampling effort (~3% of the site sampled). These designs included 1 SRS design and 5 SYS designs (each with 3 random starts) that represented the full spectrum of SYS options in terms of D_{AC} and D_{AL}. For each selected design, a sample was drawn and kriging was used to predict the mussel count for each of the 5280 0.25-m^2 sampling units using GS+ (version 5, Gamma Design Software, Plainwell, Michigan). The sampling and kriging process was repeated 50× for each of the 6 designs for a total of 300 kriged predictions. GS+ allows the user to plot semivariogram data and to choose the model that best fits the data by comparing different options using model residual sum of squares (RSS). Inspection of the spatial correlation in each the 300 sample realizations confirmed that modeling the semivariograms as isotropic (i.e., having approximately equal decrease in spatial correlation in all directions) was more appropriate than the alternative, anisotropic semivariogram models. Other semivariogram options, such as type of model (spherical, exponential, Gaussian, or linear) and the values of nugget variance, sill, and range were selected based on the model RSS.

Two summaries of prediction error were calculated for kriged predictions. Spatial prediction error was compared using statistics based on the mean squared error (MSE). The Summed Mean Squared Error (SMSE), which is analogous to the Integrated Mean Squared Error (Cox et al. 1997) but for discrete data, is

$$\mathrm{SMSE} = \sum_{j=1}^{m}[(\hat{y}_j - y_j)^2] \qquad [12]$$

where y_j is the observed count for the jth sampling unit in the study site and $\hat{y}_j$ is the predicted count for the jth sampling unit. The Maximum Mean Squared Error (MMSE) (Cox et al. 1997) is

TABLE 2. Summary of relative efficiency of designs using 0.25-m² sampling units vs designs using 1-m² sampling units. See equation 1 in text for calculation of relative efficiency. SRS = simple random sampling. SYS = systemic sampling.

Designs	Population	No. of designs	Minimum	25th percentile	Median	Mean	75th percentile	Maximum
SYS and SRS	Real	83	0.07	1.29	1.65	1.78	2.14	3.72
SYS and SRS	Simulated	83	0.43	1.02	1.20	1.28	1.52	2.84
SYS only	Real	73	0.07	1.26	1.59	1.78	2.22	3.72
SYS only	Simulated	73	0.43	1.05	1.30	1.33	1.54	2.84

$$\text{MMSE} = \max_{y_j \in M} E[(\hat{y}_j - y_j)^2], \qquad [13]$$

which compares designs based on the maximum squared difference between the predicted and observed counts.

Results

Estimating abundance

In most cases, designs that used 0.25-m² sampling units had smaller variances than designs that used 1-m² sampling units (Table 2). Eighty-one percent (67 out of 83) of the designs using 0.25-m² sampling units were more efficient (i.e., had smaller variance) than the same design using 1-m² sampling units when sampling the real population. In contrast, 55% (46 out of 83) of the designs using 0.25-m² sampling units were more efficient than designs using 1-m² sampling units when sampling the simulated population. The superior efficiency of 0.25-m² sampling units was even more apparent when the comparison was limited to SYS designs. Sampling units that were 0.25 m² were equally or more efficient than 1-m² sampling units for 86% (63 out of 73) of the SYS designs when sampling the real population and 84% (61 out of 73) of the SYS designs when sampling the simulated population (Table 2).

SYS designs were consistently more efficient than SRS designs when sampling the real population and when distances between sampling units in SYS designs were chosen appropriately and >1 random starting point was used (Table 3). Variability in density across the stream was greater than variability along the stream (Fig. 1). Therefore, appropriate systematic samples were those for which $D_{AC} \leq D_{AL}$. When $D_{AC}/D_{AL} \leq 1$ and ≥2 random starts were used, SYS designs were more efficient than SRS for 100% of the 30 designs using 1-m² sampling units and 98% (53 out of 54) of the designs using 0.25-m² sampling units. When $D_{AC}/D_{AL} > 1$ and ≥2 random starts were used, SYS designs were more efficient than SRS designs in only 35% (7 out of 20) of the designs using 1-m² sampling units and 67% (26 out of 39) of the designs using 0.25-m² sampling units (Table 3).

Relative bias (RB) of abundance estimates was ≤0.01 except when equation 4 (abundance estimate based on cluster sampling) was used with SYS designs with 1 random start and 1-m² sampling units. In those instances, mean RB was 0.204 when $D_{AC}/D_{AL} \leq 1$ and 0.41 when $D_{AC}/D_{AL} > 1$. These designs also showed some bias in abundance estimates when equation 2 (abundance estimate based on SRS) was used with SYS designs with 1 random start and 1-m² sam-

TABLE 3. Summary of relative efficiency of systematic sampling designs vs simple random sampling designs for the freshwater mussel population. Abbreviations and design descriptors as in Tables 1 and 2. See equation 1 in text for calculation of relative efficiency. Random starts = ≥2 for all situations.

D_{AC}/D_{AL}	Sampling-unit size (m²)	No. of designs	Minimum	25th percentile	Median	Mean	75th percentile	Maximum
≤1	1	30	1.14	1.44	1.58	1.61	1.77	2.15
≤1	0.25	58	0.98	1.07	1.24	1.24	1.31	2.05
>1	1	20	0.57	0.89	1.08	1.02	1.16	1.37
>1	0.25	39	0.45	0.62	0.72	0.92	1.27	1.68

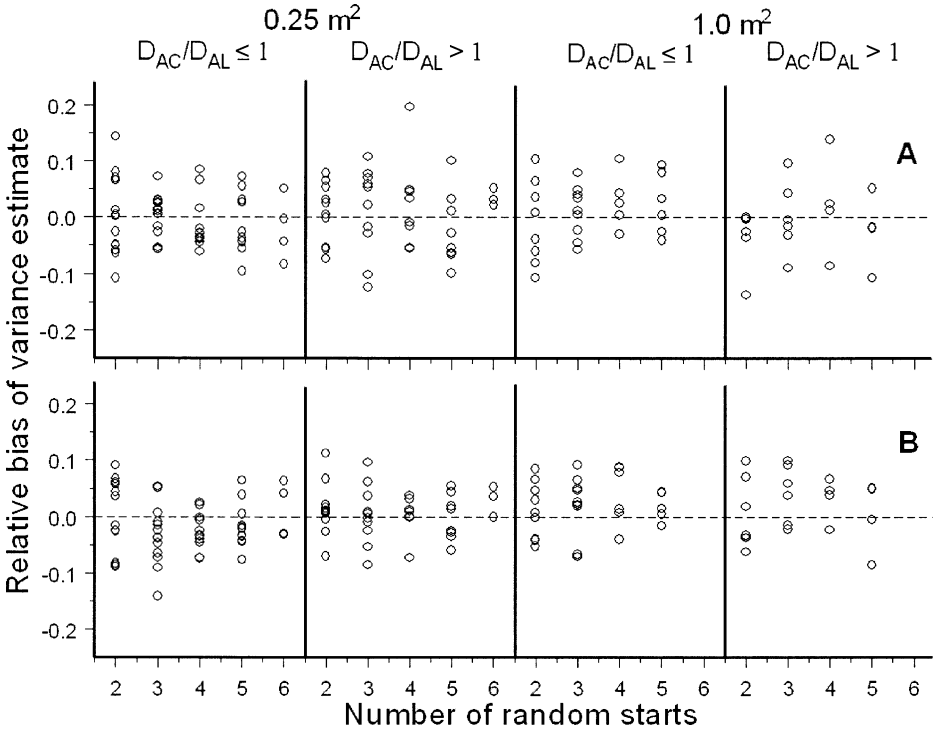

FIG. 2. Relative bias of variance (RBV) for systematic sampling (SYS) designs with ≥2 random starts. Mean RBV is shown for each design for the clustered real population (A) and the random simulated population (B). Abbreviations and design descriptors are as in Table 1. Bias was calculated using estimates of variance appropriate for SYS (equation 5). The dashed horizontal line at 0 represents no bias.

pling units. In those instances, mean RB was 0.048 when $D_{AC}/D_{AL} \leq 1$ and 0.122 when $D_{AC}/D_{AL} > 1$.

When $D_{AC}/D_{AL} \leq 1$, the relative bias of variance (RBV) estimates for the real population decreased substantially when SYS designs had 3 random starts instead of 2, regardless of which sampling unit size was used, but increasing the number of random starts beyond 3 did not result in any additional decrease in bias (Fig. 2A). When $D_{AC}/D_{AL} > 1$, using >2 random starts did not decrease RBV estimates for the real population (Fig. 2A). RBV estimates for the simulated population did decrease as number of random starts increased (Fig. 2B), but the decrease was less substantial for designs that used 1-m² sampling units and a $D_{AC}/D_{AL} > 1$ than for designs that used 0.25-m² sampling units and a $D_{AC}/D_{AL} \leq 1$.

CIs calculated using SRS formulae deviated from the expected coverage (0.95) when SYS designs (Table 4) were used to sample the real and simulated populations, but deviations from expected (both above and below 0.95) were more pronounced when the real population was sampled. CI coverage based on variance formulae appropriate for SYS performed better and gave CI coverage much closer to the expected 95% than CI coverage based on variance formulae appropriate for SRS. The systematic formulae, however, did result in substantially lower coverage than the expected 95%, when sampling from the real clustered population using 1-m² sampling units.

Quantifying degree of spatial clustering

All 3 measures of spatial clustering showed some bias, but severity and direction of bias (i.e., positive or negative) varied among the 3 statistics (Fig. 3). Positive bias indicates that the spatial distribution of a particular sample is likely to be classified as spatially clustered when it really is random. Negative bias indicates that a

TABLE 4. Mean confidence interval (CI) coverage for each group of sampling designs used to simulate sampling a spatially random simulated population or a clustered real population (see text for details). Abbreviations and design descriptors are as in Tables 1 and 2. Number of designs indicates how many designs were examined for each group specified. CIs were calculated using formulae appropriate for SRS and for SYS designs. Equations used for finding values in each column are noted in column headings. Bold font indicates that the mean CI coverage differed from the expected 95% by >1.5%. Blank spaces indicate equations 4, 5, and 6 could not be used because the designs had only 1 RS (see text for details).

					Confidence interval coverage			
					Simulated population		Real population	
Design	Sampling-unit size (m^2)	D_{AC}/D_{AL}	RS	No. of designs	SRS (Eqn 2, 3, 6)	SYS (Eqns 4, 5, 6)	SRS (Eqns 2, 3, 6)	SYS (Eqns 4, 5, 6)
SRS	0.25			10	0.947		0.936	
SRS	1			10	0.950		**0.922**	
SYS	0.25	≤1	1	19	0.938		0.939	
SYS	0.25	>1	1	13	**0.934**		**0.892**	
SYS	0.25	≤1	>1	54	0.946	0.945	0.948	0.942
SYS	0.25	>1	>1	39	0.954	0.939	**0.913**	0.941
SYS	1	≤1	1	14	**0.931**		**0.967**	
SYS	1	>1	1	9	0.940		0.954	
SYS	1	≤1	>1	30	0.947	0.944	**0.972**	**0.920**
SYS	1	>1	>1	20	0.947	0.943	**0.882**	**0.935**

spatial distribution is likely to be classified as random when it really is clustered. When sampling the real population, the Hopkins Statistic was negatively biased when a 0.25-m^2 sampling unit was used and positively biased when a 1-m^2 sampling unit was used, and relative bias of the Hopkins Statistic decreased slightly as sampling effort increased (Fig. 3). When sampling the simulated population, the Hopkins Statistic was positively biased at low levels of sampling effort but had little or no bias as effort increased. The Clumping Index was less biased when sampling the real population than when sampling the simulated population (Fig. 3). SYS designs using 1-m^2 sampling units caused the largest bias in the Clumping Index. Morisita's Index was the most reliable method for quantifying spatial clustering and resulted in the least bias overall. By reliable, we mean that bias decreased as effort increased in a similar manner for both the simulated and real populations and for both sample-unit sizes. When sampling the real population, Morisita's Index showed some negative bias with SYS designs using 1-m^2 sampling units, but it performed very well with SYS designs using 0.25-m^2 sampling units (Fig. 3).

The distributional assumptions of Hopkins Statistic also were examined. Our empirical results indicated that the Hopkins Statistic was overly sensitive to departures from spatial randomness. The probability of Type I error (probability of classifying a random population as clustered) was $\gg$5% (*F*-test). This over-sensitivity did not appear to decrease when sampling effort was increased, the number of random starts in SYS designs was increased, or an optimal D_{AC}/D_{AL} was chosen (Table 5). The Type II error (probability of classifying the clustered real population as spatially random) showed substantial variability when α = 5%. Overall, SYS designs with 1 random start using 0.25-m^2 sampling units had the highest Type II error rates.

Predicting spatial distribution

The SMSE was lowest for SRS designs, SYS designs with $D_{AC}/D_{AL} \leq 1$, and SYS designs with $D_{AC}/D_{AL} = 1$ (Table 6). MMSE was lowest for SYS designs with $D_{AC}/D_{AL} = 1$ (Table 6), but its performance was more variable than SMSE, possibly because of the sensitivity of the sample maximum to outliers. Moreover, the subjectivity inherent to the process of selecting a semivariogram model for kriging was a primary source of error in these results, but one that we could not quantify.

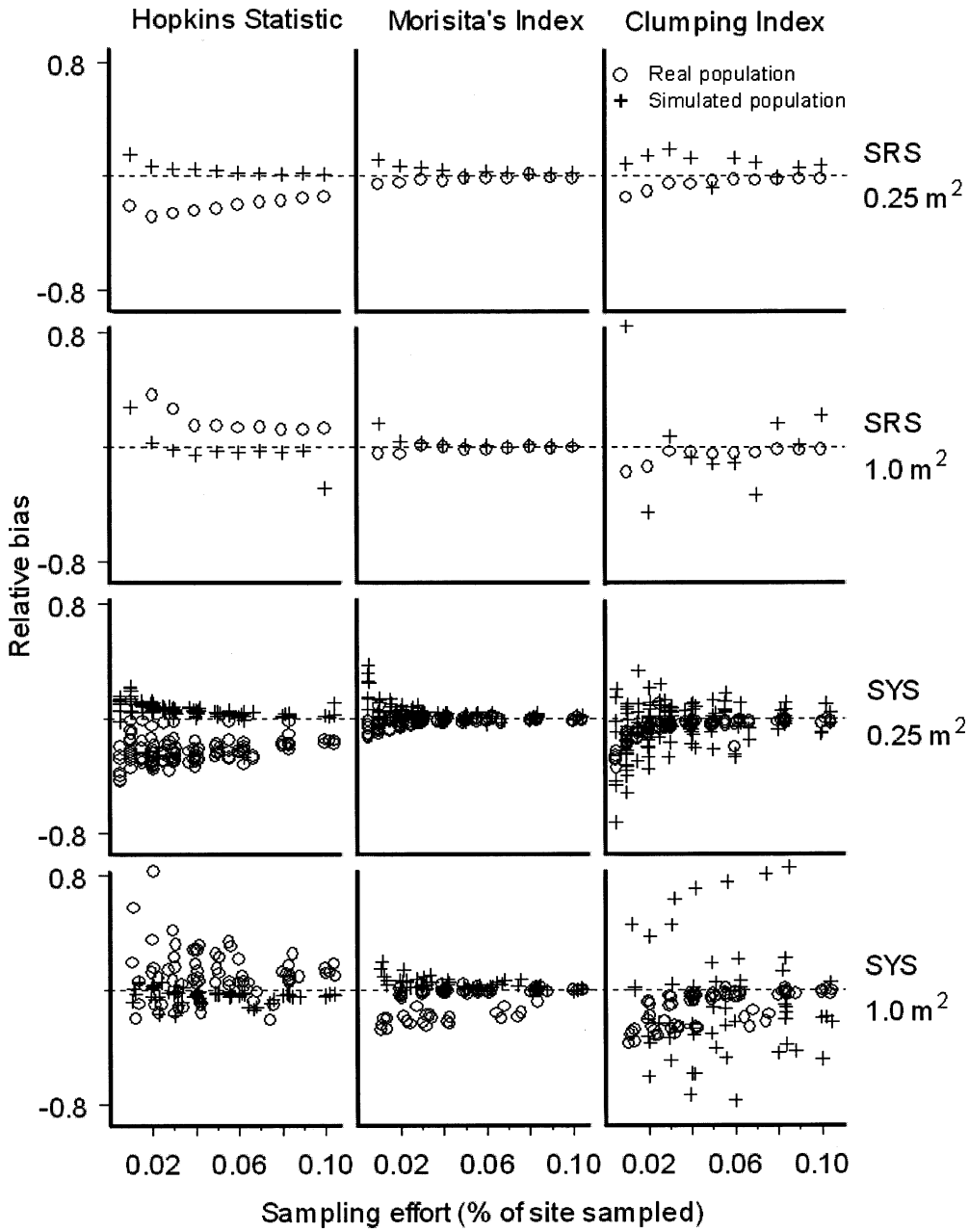

FIG. 3. Bias in 3 measures of spatial clustering as a function of sampling effort using simple random sampling (SRS) and systematic sampling (SYS) designs with different sampling-unit sizes (0.25 m^2 and 1 m^2) for real and simulated populations. Dashed horizontal lines at 0 represent no bias.

Discussion

Optimal design

SYS designs with 2 or 3 random starts using 0.25-m^2 sampling units provided reliable information for estimating abundance, quantifying spatial clustering, and predicting spatial distribution. Comparisons of RBV estimates indicated that using 3 random starts provides more accuracy than using 2 random starts when esti-

TABLE 5. Type I and Type II error rates for Hopkins Statistic test of complete spatial randomness. Abbreviations and design descriptors are as in Tables 1 and 2. Type I error rate is the probability that a random distribution is incorrectly classified as clustered and was calculated using samples from the random simulated population. Type II error rate is the probability that a clustered population is incorrectly classified as random if $\alpha = 0.05$ (Type I error rate). Type II error rate was calculated using samples from the clustered real population. Groups of designs are identical to those in Table 4.

Design	Sampling-unit size (m^2)	D_{AC}/D_{AL}	RS	Hopkins Statistic	
				Type I error rate	Type II error rate
SRS	0.25			0.0891	0.3400
SRS	1			0.1021	0.0338
SYS	0.25	≤1	1	0.0959	0.2775
SYS	0.25	>1	1	0.1026	0.2373
SYS	0.25	≤1	>1	0.0870	0.0848
SYS	0.25	>1	>1	0.0897	0.0844
SYS	1	≤1	1	0.0698	0.0708
SYS	1	>1	1	0.0523	0.0919
SYS	1	≤1	>1	0.0886	0.0137
SYS	1	>1	>1	0.0872	0.0255

mating variance of abundance in a clustered population. However, using >3 random starts does not provide substantially more accuracy in estimating RBV in clustered distributions. These same design recommendations also provide the most accurate CI coverage. The disparities in CI coverage shown in Table 4 may seem relatively slight, but even small inaccuracies can lead to incorrect conclusions in tests of hypotheses.

Size of sampling unit

We recommend 0.25-m^2 sampling units over 1-m^2 sampling units because the smaller unit requires less effort. Thus, more sampling units can be observed for the same amount of effort, assuming that travel cost between units is not overwhelming. A survey with more sampling units will provide better spatial coverage of the study area, a larger sample size, and more precise estimates than a survey with fewer sampling units. Our comparisons among designs differing only in sampling-unit size confirmed previous studies indicating that surveys using small sampling units are efficient for clustered populations (Elliott 1977). Strayer and Smith (2003) simulated SRS of the Cacapon River freshwater mussel populations to estimate species-specific density and compared efficiencies of sampling-unit sizes of 0.25, 0.5, 1, 2, and 4 m^2 and of 1-m-wide transects oriented across or with the current. They found that the coefficient of variation was lowest for 0.25-m^2 sampling units and that the advantage of the smallest sampling unit was greatest for the most clustered population. However, they also found that high per unit setup costs would reduce the comparative advantage of smaller units.

Systematic vs random sampling

Our results corroborate the utility of SYS for spatial prediction and to identify optimal de-

TABLE 6. Summed Mean Square Error (SMSE) and Maximum Mean Square Error (MMSE) for 6 sampling designs. Designs consisted of simple random sampling (SRS) or systematic sampling (SYS), and ~3% of the site was sampled with each design. All SYS designs had 3 random starts and were classified by the distance between sampling units across the stream (D_{AC}) and the distance between units along the stream (D_{AL}). SMSE and MMSE are means based on kriged predictions taken from each of 50 sampling realizations for each design. Values within a column with the same letter are not significantly different ($p > 0.05$).

Design	$D_{AC} \times D_{AL}$	D_{AC}/D_{AL}	SMSE	MMSE
SRS			2295.10^{a}	85.62^{d}
SYS	5 × 20	0.25	$2298.05^{a,b}$	83.19^{c}
SYS	20 × 5	4.0	2386.97^{c}	86.50^{e}
SYS	10 × 10	1.0	2302.87^{b}	79.83^{a}
SYS	6 × 17	0.35	2336.48^{c}	82.08^{b}
SYS	17 × 6	2.83	2336.48^{c}	81.17^{b}

signs in terms of multiple starts and distance between units when applied to freshwater mussel surveys. Smith et al. (2001) used kriging based on SYS for spatial interpolation to identify locations of mussel concentration in a large site (18,600 m^2) and overlaid the predicted spatial distribution with the locations of potential impacts from bridge replacement. Such analyses of spatial distribution can be used to assess the likelihood of negative impacts and to protect critical populations. Our findings verify that a survey design that provides uniform effort throughout the site is preferred for studying spatial distribution. SYS lends itself well to interpolation methods such as kriging (Thompson 1992) because it satisfies the uniformity condition needed. As a practical consideration, SYS with multiple random starts is easier to implement than SRS, and it retains a random component that allows valid statistical inference. Distance between systematically sampled units across the current should be equal to or less than the distance between units along the current if no information about directional variation is available. If information on directional variation within the study area is known, sampling units can be set closer together in the direction of more variation. Variation was greater across the current than along the current in the freshwater mussel population we examined, and we expect that spatial pattern to hold generally for freshwater mussels in lotic environments.

Spatial clustering

The erratic bias of the Hopkins Statistic raises some interesting questions about using nearest-neighbor methods to quantify degree of spatial clustering. The distribution of the individuals in relation to the sampling units might possibly make the Hopkins Statistic inappropriate for quadrat sampling. Cressie (1993) emphasized the extreme sensitivity of the Hopkins statistic when sampling a clustered population, but he did not discuss its accuracy when sampling a spatially random population. Cressie (1993) discussed the effectiveness of using the Hopkins statistic with respect to a sampling a population of trees in which each sampling unit was an individual observation (one tree). Cressie's (1993) findings are more likely to be applicable to other sample surveys where the goal is to sample from a large population of individuals (e.g., a large animal survey) in which animal locations are the sampling units In contrast, we counted individuals within sampling units. The calculations, which involve distances to nearest neighbors, may be affected by some distances between individual mussels that are within the same sampling unit. These relatively small distances would not be present if the sampling units were the individuals themselves rather than the sampling units within which individuals were found.

In conclusion, designing a sample survey is an important and sometimes complex task. Factors such as prevalence, spatial distribution of the sample population, and idiosyncrasies of the area to be sampled must be taken into consideration. In addition, almost all sampling studies are limited by cost, labor, and, in many cases, weather constraints. In light of these issues, SYS provides a flexible and efficient method for field sampling. The recommendations we have made for sampling a clustered freshwater mussel population (e.g., using 0.25-m^2 sampling units) are ideal for the research goals presented here, but may not fit the goals of studies involving other species. However, regardless of species, it makes sense to place sampling units closer together in the direction of more variability. If no information about directional variation is known, placing sampling units equal distances apart in all directions is appropriate for a wide variety of research goals. We think that almost all sampling studies could benefit greatly from doing a relatively small pilot study ahead of time to obtain estimates of variability. Information gained from a pilot study is often invaluable in designing an appropriate sampling scheme.

Acknowledgements

Rita Villella, David Lemarié, Craig Snyder, David Weller, and Priscilla Young assisted in data collection and data management. We thank Eric Smith for a helpful review prior to submission. Two anonymous referees and David Strayer provided detailed critiques of earlier versions of our manuscript.

Literature Cited

BEALL, G. 1939. Methods of estimating the population of insects in a field. Biometrika 26:422–439.

BYTH, K., AND B. D. RIPLEY. 1980. On sampling spatial

patterns by distance methods. Biometrics 36:279–284.

CAWLEY, E. T. 1993. Sampling adequacy in population studies of freshwater mussels. Pages 169–172 *in* K. S. Cummings, A. C. Buchanan, and L. M. Koch (editors). Conservation and management of freshwater mussels. Upper Mississippi River Conservation Committee, Rock Island, Illinois. (Available from: Illinois Natural History Survey, 607 East Peabody Drive, Champaign, Illinois 61820 USA.)

COX, D. D., L. H. COX, AND K. B. ENSOR. 1997. Spatial sampling and the environment: some issues and directions. Environmental and Ecological Statistics 4:219–233.

CRESSIE, N. 1993. Statistics for spatial data. 2nd edition. Wiley, New York.

CHRISTMAN, M. C. 2000. A review of quadrat-based sampling of rare, geographically clustered populations. Journal of Agricultural, Biological, and Environmental Statistics 5:168–201.

DIGGLE, P. J., J. E. BESAG, AND J. T. GLEAVES. 1976. Statistical analysis of spatial point patterns by means of distance methods. Biometrics 32:659–667.

DOWNING, J. A., AND W. L. DOWNING. 1992. Spatial aggregation, precision, and power in surveys of freshwater mussel populations. Canadian Journal of Fisheries and Aquatic Sciences 49:985–991.

ELLIOTT, J. M. 1977. Some methods for the statistical analysis of samples of benthic invertebrates. Scientific Publication No. 25. Freshwater Biological Association, Ambleside, UK.

FINNEY, D. J. 1946. Field sampling for the estimation of wireworm populations. Biometrics 2:1–7.

HEDAYAT, A. S., AND B. K. SINHA. 1991. Design and inference in finite population sampling. Wiley, New York.

HOLGATE, P. 1965. Tests of randomness based on distance methods. Biometrika 52:345–353.

LUDWIG, J. A., AND J. F. REYNOLDS. 1988. Statistical ecology, a primer on methods and computing. Wiley, New York.

MCARTHUR, R. D. 1987. An evaluation of sampling designs for estimating locally concentrated pollutants. Communications in Statistics: Simulation and Computation 16:735–759.

MORISITA, M. 1959. Measuring of the dispersion and analysis of distribution patterns. Memoirs of the Faculty of Science, Biology, Kyushu University, Series E 2:215–235.

MORISITA, M. 1971. Composition of the I_d index. Researches on Population Ecology 13:1–27.

RICCIARDI, A., AND J. B. RASMUSSEN. 1999. Extinction rates of North American freshwater fauna. Conservation Biology 13:1220–1222.

SALEHI, M. M., AND D. R. SMITH. 2005. Two-stage sequential sampling: a neighborhood-free adaptive sampling procedure. Journal of Agricultural, Biological, and Environmental Statistics 10:84–103.

SMITH, D. R., M. J. CONROY, AND D. H. BRAKHAGE. 1995. Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. Biometrics 51:777–788.

SMITH, D. R., R. F. VILLELLA, AND D. P. LEMARIÉ. 2001. Survey protocol for assessment of endangered freshwater mussels in the Allegheny River, Pennsylvania. Journal of the North American Benthological Society 20:118–132.

SMITH, D. R., R. F. VILLELLA, AND D. P. LEMARIÉ. 2003. Application of adaptive cluster sampling to low-density populations of freshwater mussels. Environmental and Ecological Statistics 10:7–15.

STRAYER, D. L., S. CLAYPOOL, AND S. SPRAGUE. 1997. Assessing unionid populations with quadrats and timed searches. Pages 163–169 *in* K. S. Cummings, A. C. Buchanan, C. A. Mayer, and T. J. Naimo (editors). Conservation and management of freshwater mussels II. Initiatives for the future. Upper Mississippi River Conservation Committee, Rock Island, Illinois. (Available from: Illinois Natural History Survey, 607 East Peabody Drive, Champaign, Illinois 61820 USA.)

STRAYER, D. L., AND D. R. SMITH. 2003. A guide to sampling freshwater mussel populations. American Fisheries Society Monograph 8. American Fisheries Society, Bethesda, Maryland.

TAYLOR, C. C. 1953. Nature of variablility in trawl catches. Fishery Bulletin 54:145–166.

UPTON, G. J. G., AND B. FINGLETON. 1985. Spatial data analysis by example, volume 1: point pattern and quantitative data. Wiley, New York.

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

THOMPSON, S. K., AND G. A. F. SEBER. 1996. Adaptive sampling. Wiley, New York.

Received: 10 November 2004
Accepted: 3 June 2005