J. N. Am. Benthol. Soc., 2006, 25(3):701–711

Survey design for detecting rare freshwater mussels

David R. Smith[1]

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

Abstract. A common objective when surveying freshwater mussels is to detect the presence of rare populations. In certain situations, such as when endangered or threatened species are potentially in the area of a proposed impact, the survey should be designed to ensure a high probability of detecting species presence. Linking survey design to probability of detecting species presence has been done for quantitative surveys, but commonly applied designs that are based on timed searches have not made that connection. I propose a semiquantitative survey design that links search area and search efficiency to probability of detecting species presence. The survey can be designed to protect against failing to detect populations above a threshold abundance (or density). I illustrate the design for surveys to detect clubshell (*Pluerobema clava*) and northern riffleshell (*Epioblasma torulosa rangiana*) in the Allegheny River. Monte Carlo simulation indicated that the proposed survey design performs well under a range of spatial distributions and low densities (<0.05 m^2) where search area is sufficient to ensure that the probability of detecting species presence is predicted to be ≥ 0.85.

Key words: unionid, probability of species detection, detectability, qualitative sampling, rare populations, species presence, timed search, occupancy.

A common objective of surveys of freshwater mussels is to detect the presence of rare populations, e.g., when assessing site-specific impacts on endangered or threatened species (Wilcox et al. 1993, Smith et al. 2001a) or when delineating the range of a rare species (Strayer et al. 1996). An important application of this objective is determining the presence of an endangered or threatened species in an area of a proposed impact. In that case, confirmation of species presence would halt or influence the activity that would cause the impact, whereas failure to detect a species when it was in fact present (analogous to a Type II error) could permit an adverse impact to occur. Thus, a survey designed to achieve this objective should ensure a high probability of detecting species presence.

Intuition tells us that the probability of detecting species presence is related to species abundance and spatial distribution, sampling effort, search efficiency within the area sampled (i.e., detectability), and the distribution of sampling effort within a study site. McArdle (1990) and Green and Young (1993) related detection of rare species to the number of sampling units taken in a quantitative, quadrat-based survey assuming perfect search efficiency. Near-perfect search efficiency would be achieved in a freshwater mussel survey by sediment excavation (Hornbach and Deneka 1996, Smith et al. 2001b). Green and Young (1993) provided guidelines for designing a quantitative survey that would ensure a high probability of detecting rare species. However, their guidelines have not been widely adopted for freshwater mussel surveys, in part because quantitative sampling is perceived as time-consuming and expensive (Obermeyer 1998), and timed-search surveys result in more species detections per unit time than quadrat-based surveys (Strayer et al. 1997, Vaughn et al. 1997, Obermeyer 1998).

Timed searches are qualitatively more efficient than quadrat-based surveys, but an explicit method to relate search time to the probability of detecting species presence does not appear to exist. Strayer et al. (1997) calculated probability of detection for timed searches for *Elliptio complanata*, but cited high variance of catch per unit effort statistics as a limitation on the generality of a timed-search-based detection curve. Metcalfe-Smith et al. (2000) found that >50% of species present are missed when typical search times are used and that increased search time resulted in more species detections. However, the essential question of how much search time is enough to ensure a high

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

probability of detecting a rare species remains unanswered.

I propose an alternative survey design that is intermediate between a timed search and Green and Young's (1993) quadrat-based sampling. The design relates probability of detecting species presence to search area and search efficiency. The semiquantitative approach does not require sediment excavation, but does require a priori information on search efficiency. Search effort is constrained to defined areas (i.e., sampling units), so the survey design can be linked to probability of detecting species presence. I describe an example survey designed to detect clubshell (*Pluerobema clava*) and northern riffleshell (*Epioblasma torulosa rangiana*) in the Allegheny River. Last, I evaluate the design using a Monte Carlo simulation that includes spatially clustered populations because the survey design relies on assumptions about spatial distribution of rare populations.

Survey Design

I developed the survey design by specifying the survey objective and applying a model to link the objective to elements of the design. In particular, I considered factors that affect search efficiency (e.g., detectability) because it is an important element in mussel survey design. I also considered relevant statistical principles that could guide how best to distribute the area to be searched within a site.

Survey objective

Clear, specific, and quantifiable objectives are central to successful survey design (Strayer and Smith 2003, McDonald 2004). The primary objective of our survey was to detect the presence of a rare population, but a survey objective should be defined further and stated quantitatively to allow for evaluating whether a proposed design will meet the objective. For example, the objective might be stated quantitatively: "To detect the presence of any of the endangered or candidate species in a site with probability $\geq$0.85 given that species abundance is $\geq$100 individuals." This statement has 2 important elements: 1) the minimum threshold for the probability of detecting presence of a species, and 2) a species abundance or density that is deemed biologically meaningful. I used an abundance of 100 individuals only as an example. The determination of a biologically meaningful threshold should involve multiple considerations including legal mandates, life history, population viability, and comparisons of densities throughout a local watershed, region, or range.

Modeling the sampling process

A model of the sampling process is needed to relate the proposed objective to the survey design. The model represents the expected survey results (counts of mussels) as a function of the controlling factors—mussel abundance, search area, and search efficiency. Search efficiency, which is also termed detectability, is the probability of detecting an individual mussel given that it is within the search area.

The expected number of individuals counted in a survey of a site can be represented as

$$E(C) = \alpha\beta T \qquad [1]$$

where C is the count of individuals, E(C) is the expected count based on a repeatable sampling process, α is the fraction of the site that is searched, β is the probability of detecting an individual given that it is in the search area, and T is the total number of individuals in the site (Williams et al. 2002:244). The expected number of individuals in the search area is $\alpha\, T = a\mu$ where a is the search area and μ is species density. Note that the fraction of the site that is searched is $\alpha = a/A$ where A denotes the area of the site. The search area is the sum of the areas of each unit in the sample, i.e., $a = \sum_{i=1}^{n} a_i$ where n is the sample size and a_i is the area of the i^{th} sampling unit (typically a_i is the same or nearly the same for all sampling units).

Search efficiency, which refers to the probability of detecting an individual given that it is in the search area, is a function of search rate (time per unit area) and search area (Fig. 1). In eq. 1, search efficiency is denoted by β. In theory, if one spends enough time and effort searching an area, all individuals that are present within the search area will be detected, in which case $\beta = 1$. However, in actual sampling situations, search time and effort are restricted so that not all individuals in the sample area are detected and $\beta < 1$.

Mussel sampling techniques have been classified as quantitative, semiquantitative, or qualitative (Strayer and Smith 2003). This classification can be related to the parameters in eq. 1 (Table 1). Quantitative and semiquantitative sampling are distinguished from qualitative sampling by α. α is known when sampling is quantitative or semiquantitative, but α is not known when sampling is qualitative. Quantitative and semiquantitative sampling are distinguished by β. Quantitative sampling is the case where $\beta = 1$ or $\beta < 1$ and is estimated. In either case, β can be accounted for in eq. 1. Semiquantitative sampling is the case where β is unknown. Unbiased estimation of abundance or density is possible only when α and β are known or

TABLE 1. Contrast of sampling techniques (classified as qualitative, semiquantitative, or quantitative) based on fraction of site searched (α), search efficiency or detectability (β), and which parameter(s) are known or estimated. C = count of mussels in a sample at a site, T = total number of individuals in the population at a site, $\hat{T}$ = estimated total number of individuals at a site, $\hat{\beta}$ = estimated probability of detecting an individual given that it is within the search area.

Sampling technique	α	β	Survey result
Qualitative	Unknown	Unknown	Incomplete count
Semiquantitative	Known	Unknown	Incomplete count within searched area
Quantitative	Known	Known or estimated	Abundance estimate: $\hat{T} = C/\left(\alpha\hat{\beta}\right)$

estimated, i.e., $\hat{T} = C/(\alpha\hat{\beta})$ where $\hat{T}$ is the estimated total number of individuals in the study site (i.e., the abundance estimate) and $\hat{\beta}$ is an unbiased estimate of the probability of detecting an individual given that it is in the search area.

Detecting the presence of a rare species within a site is equivalent to detecting at least one individual of that species, and it follows from eq. 1 that this event is a function of α, β, and T. That is:

$$\text{Prob(detecting at least one individual)} = \text{Prob}(C > 0) = f(\alpha\beta T). \quad [2]$$

Green and Young (1993) considered sampling rare populations of freshwater mussels in quadrats and derived a formula for the probability of detecting the presence of a low-density population (i.e., $\mu < 0.10/\text{m}^2$) using a Poisson probability distribution:

$$\text{Prob(detecting at least one individual)} = 1 - e^{-mn} \quad [3]$$

where m is the number of individuals within a sampling unit and n is the number of random sampling units searched. The Poisson assumption implies that mussels at very low density have a spatially random distribution. This assumption does not imply an absence of underlying ecological relationships, such as habitat associations and dispersal mechanisms, which affect distribution (Downing and Downing 1991). Rather, it indicates that when mussels are geographically rare at a site (i.e., $\mu < 0.1/\text{m}^2$), their low density masks underlying ecological relationships and their spatial distribution is random from a statistical perspective. Green and Young (1993) presented empirical data to support this contention. In addition, Smith et al. (2003) found that low-density mussels on the Cacapon River, West Virginia, had random distributions as evidenced by variance-to-mean ratios. A variance-to-mean ratio of 1 indicates a Poisson distribution (Elliott 1977). Downing and Downing (1991) presented a formula for variance as a function of the mean number of individuals collected that was developed empirically from surveys in lentic and lotic habitats. The Downing and Downing (1991) formula indicates that the variance-to-mean ratio approaches 1 (spatial randomness) as the mean approaches 0.10, the threshold for rarity used by Green and Young (1993). I used data from Smith et al. (2001b) and found variance-to-mean ratios for 60 species/site combinations (31 species at 14 sites) that indicated mussel distributions were statistically spatially random for $\mu \leq 0.10/\text{m}^2$. The same relationship between density and spatial distribution has been found in other populations (McArdle 1990, Welsh et al. 1996). Therefore, I propose eq. 3 as a useful approximation for guiding survey design, and I evaluate its use in a simulation that includes spatially clustered populations and sampling units other than quadrats (see below).

Equation 3 can be revised to account for search efficiency by including the parameter β, thereby making a connection to the sampling-process model in eq. 1. The expected number of individuals detected is $\beta nm = \beta\alpha T$. Thus:

$$\text{Prob(detecting at least one individual)} = 1 - e^{-\beta\alpha T} = 1 - e^{-\beta a T/A} = 1 - e^{-\beta a \mu}. \quad [4]$$

Equation 4 can be used to examine the effect of search efficiency (β), search area (a), and density (μ) on the probability of detecting at least one individual or, analogously, the probability of detecting species presence. Figure 2 shows the probability of detecting species presence for $\mu = 0.01$, 0.05, and $0.10/\text{m}^2$, $\beta =$ 0.2, 0.4, 0.6, and 0.8, and $a = 100$ to $1000\ \text{m}^2$. Equation 4 also could be used to examine the effect of abundance (T) for a given study site area (A) instead of μ. Table 2 shows probability of detecting species presence for $T =$ 100 to 500 and $A = 16{,}000$ and $32{,}000\ \text{m}^2$.

Factors that affect search efficiency

Search efficiency is a function of search area and search time (Fig. 1). The exact form of that relationship is not known and will vary over time and area. For a given search area, the more time spent searching, the higher the search efficiency. It is likely that search efficiency will increase quickly as search time is

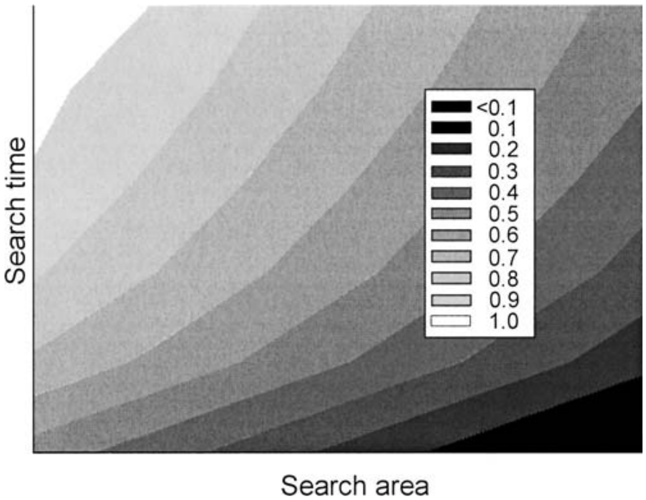

FIG. 1. Search efficiency (β; legend) as a function of search time and search area (*a*). The axes are not labeled because the exact form of the relationship is determined by a variety of factors involving mussel biology, physical environment, and observer capabilities.

increased from low to moderate levels and the rate of increase in search efficiency will slow as it approaches complete detection, exhibiting a point-of-diminishing-returns-type phenomenon. These relationships between search time and search efficiency also have been shown empirically (Metcalfe-Smith et al. 2000).

The exact form of the relationship between search efficiency and search time will depend on a number of factors (Strayer et al. 1997), some of which are inherent to the biology and natural history of the mussel species. For example, some species are more cryptic than others by virtue of their size, coloration, or reproductive behavior (Miller and Payne 1993, Obermeyer 1998, Haag and Warren 2000). Mussels exhibit seasonal patterns in vertical migration associated with day length and water temperatures (Amyot and Downing 1991, Watters et al. 2001, Perles et al. 2003). Other biological factors include gender and demographics. For example, female northern riffleshell (*Epioblasma torulosa rangiana*) are more visible than males (Smith et al. 2001a), and small mussels are difficult to detect (Miller and Payne 1988, Hornbach and Deneka 1996, Richardson and Yokley 1996, Smith et al. 2001b). Other factors, such as turbidity, hydrologic variability, substrate, and vegetative cover, are associated with the physical environment (Di Maio and Corkum 1997, Smith et al. 2001b). Last, some factors, such as observer experience, visual acuity, and fatigue, are associated with the observer (Strayer et al. 1997).

Only those mussels that are epibenthic or not buried can be found in a search restricted to the substrate surface (Amyot and Downing 1991). If an area is searched thoroughly so that all mussels on the substrate surface have been found, then search efficiency will be capped at the proportion of mussels that are on the surface. Beyond that level of effort, excavation would be required to increase search efficiency to the point that all or nearly all mussels within the searched area are found (Smith et al. 2001b).

Impact of search efficiency on survey design

Because search efficiency directly affects the probability of detecting species presence, it should be considered when designing a survey. Two approaches could be used to incorporate search efficiency in survey design. First, one could be conservative and assume that search efficiency (β) was low. Then the relationship from eq. 4 (Table 2, Fig. 2A–D) could be used as a guide to find the search area (*a*) that would ensure that the probability of detecting species presence is sufficiently high (Fig. 2A–D). For example, if β were assumed to be ≤ 0.2, then *a* would have to be >1000 m^2 to have a probability of detecting at least one individual = 0.85 for $\mu = 0.01/\text{m}^2$ (Fig. 2A). This *a* would be equivalent to ten 1-m-wide × 100-m-long transects (distribution of search effort throughout the site is discussed below). The assumed β could be based on life-history traits, such as likelihood that an individual would be endobenthic (Amyot and Downing 1991). This approach would be precautionary.

Second, β could be estimated at another time and place where the rare species was numerous or by a pilot survey based on a related, but more common, species. For example, β could be estimated by searching the surface of quadrats before excavating sediment (cf. Haukioja and Hakala 1974, Smith et al. 2001b). In this case, the estimate of β and eq. 4 could be used to predict the *a* that would result in the desired probability of detecting species presence. For example, if β for a search rate of 2 min/m^2 were estimated as 0.4, then $a = 500$ m^2 would ensure a probability of detecting species presence = 0.85 for $\mu = 0.01/\text{m}^2$ (Fig. 2B), and 1000 min (16.67 h) of search time would be required. The shortcoming of using an estimate from another time and place is that β would be estimated under one set of conditions and applied under a similar, but not identical, set of conditions. If an overestimate of β were used in survey design, then the probability of detecting species presence also would be overestimated, and the design would not be precautionary. The number of quadrats needed to estimate β would depend on μ at the site and the

TABLE 2. Probability of detecting species presence given the study site area (*A*), search efficiency (β), abundance (*T*), and search area (*a*). Bold font indicates probability of species detection ≥0.85.

			a (m²)								
A (m²)	β	*T*	100	200	300	400	500	600	700	800	900
16,000	0.2	100	0.12	0.22	0.31	0.39	0.46	0.53	0.58	0.63	0.68
		200	0.22	0.39	0.53	0.63	0.71	0.78	0.83	**0.86**	**0.89**
		300	0.31	0.53	0.68	0.78	**0.85**	**0.89**	**0.93**	**0.95**	**0.97**
		400	0.39	0.63	0.78	**0.86**	**0.92**	**0.95**	**0.97**	**0.98**	**0.99**
		500	0.46	0.71	**0.85**	**0.92**	**0.96**	**0.98**	**0.99**	**0.99**	**1.00**
	0.4	100	0.22	0.39	0.53	0.63	0.71	0.78	0.83	**0.86**	**0.89**
		200	0.39	0.63	0.78	**0.86**	**0.92**	**0.95**	**0.97**	**0.98**	**0.99**
		300	0.53	0.78	**0.89**	**0.95**	**0.98**	**0.99**	**0.99**	**1.00**	**1.00**
		400	0.63	**0.86**	**0.95**	**0.98**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**
		500	0.71	**0.92**	**0.98**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
	0.6	100	0.31	0.53	0.68	0.78	**0.85**	**0.89**	**0.93**	**0.95**	**0.97**
		200	0.53	0.78	**0.89**	**0.95**	**0.98**	**0.99**	**0.99**	**1.00**	**1.00**
		300	0.68	**0.89**	**0.97**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		400	0.78	**0.95**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		500	**0.85**	**0.98**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
32,000	0.2	100	0.06	0.12	0.17	0.22	0.27	0.31	0.35	0.39	0.43
		200	0.12	0.22	0.31	0.39	0.46	0.53	0.58	0.63	0.68
		300	0.17	0.31	0.43	0.53	0.61	0.68	0.73	0.78	0.82
		400	0.22	0.39	0.53	0.63	0.71	0.78	0.83	**0.86**	**0.89**
		500	0.27	0.46	0.61	0.71	0.79	**0.85**	**0.89**	**0.92**	**0.94**
	0.4	100	0.12	0.22	0.31	0.39	0.46	0.53	0.58	0.63	0.68
		200	0.22	0.39	0.53	0.63	0.71	0.78	0.83	**0.86**	**0.89**
		300	0.31	0.53	0.68	0.78	**0.85**	**0.89**	**0.93**	**0.95**	**0.97**
		400	0.39	0.63	0.78	**0.86**	**0.92**	**0.95**	**0.97**	**0.98**	**0.99**
		500	0.46	0.71	**0.85**	**0.92**	**0.96**	**0.98**	**0.99**	**0.99**	**1.00**
	0.6	100	0.17	0.31	0.43	0.53	0.61	0.68	0.73	0.78	0.82
		200	0.31	0.53	0.68	0.78	**0.85**	**0.89**	**0.93**	**0.95**	**0.97**
		300	0.43	0.68	0.82	**0.89**	**0.94**	**0.97**	**0.98**	**0.99**	**0.99**
		400	0.53	0.78	**0.89**	**0.95**	**0.98**	**0.99**	**0.99**	**1.00**	**1.00**
		500	0.61	**0.85**	**0.94**	**0.98**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**

proportion of individuals on the substrate surface (Smith et al. 2001b, Strayer and Smith 2003). Therefore, the environmental conditions in the pilot survey should be as close as possible to the conditions likely to be encountered at the site where species presence will be determined. Information on species-specific densities and search efficiencies are available in the literature in some cases (e.g., Smith et al. 2001a), and unpublished agency surveys are likely to provide relevant data.

Statistical principles guiding the distribution of search effort within the site

Two statistical principles, in particular, are useful for guiding distribution of search effort. First, spatially balanced sampling has been recognized as efficient for sampling natural resources (Christman 2000, Stevens and Olsen 2004). A spatially balanced sample is one that is distributed throughout a site or population. Various systematic or grid sampling methods qualify as spatially balanced. Second, it is generally more efficient (reduces sampling error) to distribute effort among many small units than a few large units. This principle is particularly relevant when the population is spatially clustered (Elliott 1977). The mitigating factor is the effort required to move among units. Many small units require more between-unit travel than few large units. Thus, the challenge is to find a sampling-unit size that represents a compromise between cost and sampling error. These principles can be combined with stratification to allocate effort efficiently and to ensure that sampling is done in all habitats. For example, a site can be stratified by macrohabitat (e.g., riffle, run, pool) and search area can be allocated proportionately or according to anticipated habitat value (i.e., more effort in better habitat). On the other hand, the survey could be conducted in phases as suggested by Kovalak et al. (1986) and implemented recently by Villella and Smith (2005). During the 1st phase, an informal search or surveillance can be conducted to delineate mussel beds or

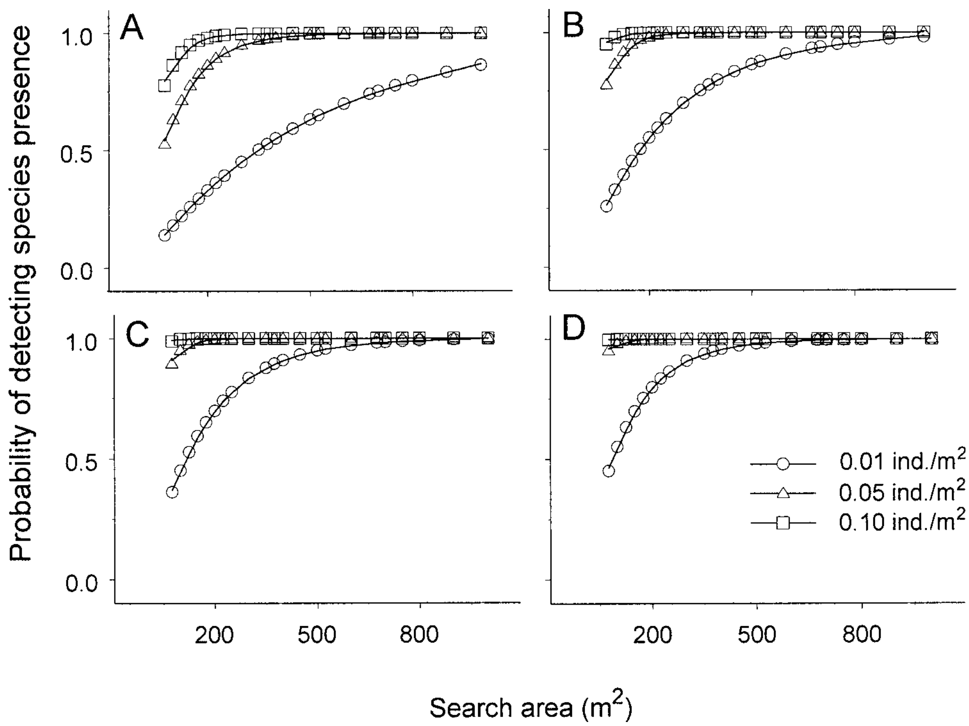

FIG. 2. Probability of detecting species presence as a function of search area (*a*) and density (0.01, 0.05, 0.10 individual/m^2) of mussels when search efficiency (β) was 0.2 (A), 0.4 (B), 0.6 (C), and 0.8 (D).

habitat. During the 2nd phase, the semiquantitative approach can be applied after the search area (a) has been determined to ensure a sufficiently high probability of detecting species presence. The predetermined *a* should be allocated so that most, but not all, of the area occurs within the bed or habitat identified during the 1st phase.

A cautionary note is warranted regarding the distribution of sampling effort according to an explicit or implied habitat model. If the habitat model is a good approximation, then it can be helpful in distributing search area. Depth and hydrological variability are useful predictors of mussel density (Haukioja and Hakala 1974, Strayer and Ralley 1993, Di Maio and Corkum 1995). However, if the model is a poor approximation, as Strayer and Ralley (1993) found for microhabitat variables, then model-based distribution can be inefficient at best and misleading at worst. A poor habitat model could lead to omission of the actual habitat from the area searched.

Detection of Clubshell (*Pleurobema clava*) and Northern Riffleshell (*Epioblasma torulosa rangiana*) in the Allegheny River

I used data from the clubshell and northern riffleshell in the Allegheny River to illustrate the design of a survey to detect their presence. In previous surveys on the Allegheny River, Smith et al. (2001a) reported that a thorough search of the substrate surface required 2 min/m^2 of search time. At the West Hickory bridge site, ~30% and 50% of clubshell and northern riffleshell were found at the substrate surface, respectively.

Suppose the goal was to protect a site against adverse impact if either species was present at $\mu \geq 0.01/m^2$ with a probability of detecting species presence $\geq$ 0.85. (Tolerance for risk is a subjective decision that often would be set during the regulatory process.) To protect either species, the β corresponding to the least detectable species, the clubshell, would be used. In this case, we assume that the substrate surface within *a* will be searched thoroughly so that β is the proportion of mussels on the substrate surface. Given this information, we can design a survey using eq. 4:

$$0.85 = 1 - e^{-0.30a0.01}$$

and solve for *a:*

$$a = \frac{\ln(1 - 0.85)}{-0.003}$$
$$= 632\text{m}^2.$$

Based on the principle of spatially balanced sampling, at least 632 m^2 of search area should be distributed throughout the site. A reasonable design would be to search within transects oriented perpendicular to shoreline or the thalweg. Following the rule that more small units are better, use of 0.5-m-wide transects would allow greater spatial dispersion of sampling effort; however, logistics and tradition might favor 1-m-wide transects, especially at sites where SCUBA is required. Transect length would depend on site dimensions. For example, if the site was 100 m across the river, then seven 1-m-wide transects would be required. Good spatial balance and coverage would be achieved by selecting a random start and placing transects at equal intervals. An improvement on that plan would include 2 random starts. To increase probability of detecting species presence to 0.95, ten 1×100 m transects would be required.

After a has been determined based on $\hat{\beta}$ and a μ that is to be protected, the time required to conduct the survey can be calculated. Based on 2 min/m^2 to search the surface substrate thoroughly, searching seven $1 \times$ 100-m transects would require ~23 h, which could be divided among multiple observers. The survey could be accomplished in ~1 d with a crew of 4. This time and effort does not seem to be an unreasonable survey cost when the objective is to detect a rare or endangered species before an adverse impact occurs. Budgets for construction projects, for example, can amount to hundreds of thousands to millions of dollars. The cost to conduct a rigorous mussel survey is trivial by comparison.

TABLE 3. Abundance (T), search efficiency (β), and cluster size for the populations used to simulate the proposed survey design. The study site was 16,000 m^2 (160 m $\times$ 100 m). Variance-to-mean ratios were calculated for individuals within 1 m $\times$ 100 m transects.

			Variance-to-mean ratio	
T	β	Cluster size	Entire population	Detectable portion of the population
100	0.2	1	1.19	0.91
		3	2.03	1.11
		5	2.33	1.04
	0.4	1	1.21	1.17
		3	1.93	1.40
		5	2.17	1.29
	0.6	1	1.05	1.04
		3	1.91	1.58
		5	2.84	1.77
300	0.2	1	0.87	0.85
		3	2.44	1.08
		5	2.53	1.20
	0.4	1	0.83	1.03
		3	1.85	1.39
		5	2.11	1.36
	0.6	1	0.94	1.08
		3	1.77	1.31
		5	2.62	1.92
500	0.2	1	1.13	1.15
		3	1.84	1.19
		5	2.91	1.44
	0.4	1	1.03	1.04
		3	1.61	1.11
		5	2.19	1.39
	0.6	1	1.03	0.80
		3	2.01	1.45
		5	2.81	2.04

Monte Carlo Simulation

To evaluate the proposed survey design, a computer program was used to generate locations for individual mussels within a site of 16,000 m^2 (100 m $\times$ 160 m), apply search efficiencies so that different proportions of the mussels were detectable, and count detectable mussels within systematically placed 1-m transects. Abundance at the site was a Poisson random variable with means of 100, 300, and 500 mussels representing population densities of 0.006, 0.02, and 0.03 (individuals/m^2). Individual mussels were in clusters with mean sizes of 1, 3, or 5 individuals (a cluster size of 1 represented complete spatial randomness). The location of the cluster center was random within the site, and individuals were distributed from the cluster center at a uniform random angle and exponential random distance, with mean distance of 1 m. Search efficiencies of 0.2, 0.4, or 0.6 were applied to determine whether each individual in the population was detectable. Detectable individuals were counted within 1-m transects oriented across the short axis of the site (100 m). Areas searched were 400, 600, 800, and 1000 m^2. The probability of detecting species presence was calculated as the proportion of 1000 replications where at least one individual was counted. Computations were done in SAS (version 9.1 SAS Institute, Cary, North Carolina).

The populations showed differing degrees of spatial clustering (Table 3, Fig. 3). Variance-to-mean ratios increased with cluster size were lower when calculated using detectable individuals only. Thus, the detectable portion of the population appears less spatially clustered than the actual population.

Simulated probabilities of detecting species presence generally tracked the probabilities predicted from eq. 4 (Table 4). Variability in the simulated probabilities was caused by variability in abundance, search efficiency, cluster size, and sample selection. This result is relevant because abundance, search efficiency, and spatial distribution would not be known exactly when using eq. 4 for survey design. The simulations indicated that eq. 4 is a useful guide under a range of conditions. Most important, the survey design

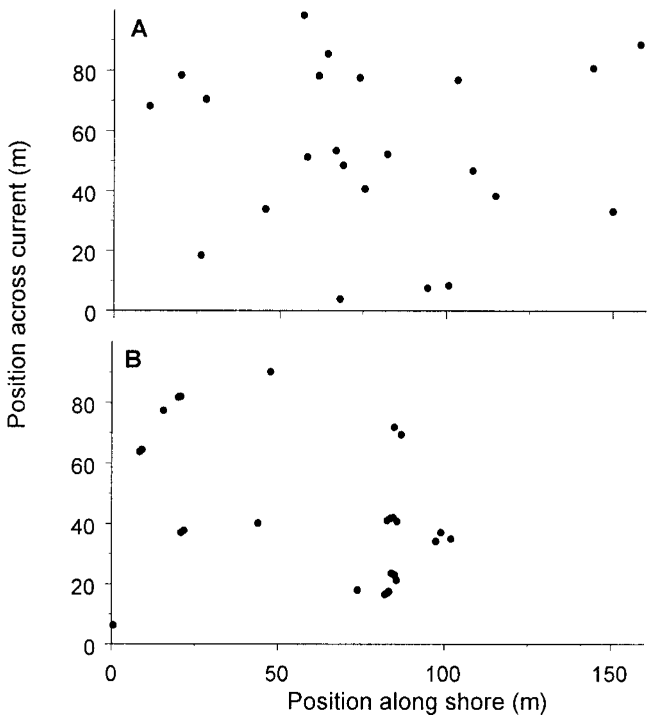

Fig. 3. Example spatial distributions of detectable mussels used to evaluate the survey design when simulated abundance was 100, search efficiency was 0.2, and cluster sizes were 1 (A) and 5 (B). Detectable mussels were a random subset of the abundance determined by the search efficiency. There were 23 detectable mussels in A and 29 in B.

performed well when a was predicted to result in a high probability of detecting species presence. Simulated probabilities of detecting species presence were $\geq$0.85 in 92% (77 of 84) of cases where eq. 4 predicted the probabilities would be $\geq$0.85 (Table 4).

Discussion

Clear, specific, and quantitative objectives are prerequisites to a successful survey design (McDonald 2004). For example, the objective for a pre-dredging survey could be to detect the presence of any endangered or candidate species with probability $\geq$0.85 given that species density is $\geq 0.01/m^2$. An important question to ask when designing a survey is whether the proposed design will meet the stated objective (Strayer and Smith 2003). The survey design described here provides a method for answering that question by linking survey elements, i.e., search area and search efficiency, to the probability of detecting species presence.

The proposed survey design, which is intermediate between timed search and quadrat methods, requires that the search area be constrained within sampling units, but excavation is not required because search efficiency is assumed to be less than perfect. Distribution of the search area within the site is flexible within guidelines. Based on well-established principles of sampling natural resources, it is best to distribute sampling effort throughout a study site in relatively small sampling units. The size of the sampling units is mitigated by logistic considerations with transects recommended in some cases because of ease of field application. A Monte Carlo simulation confirmed that use of systematically placed transects is a good approach for the objective of species detection. However, use of transects would not be a good approach when the objective is to estimate abundance or density because some amount of excavation would be required and, therefore, quadrats would be required (Smith et al. 2001b, Strayer and Smith 2003). Information on habitat or mussel beds can be used to stratify the site and to allocate the search area within strata either proportionately or with more of the search effort allocated to better habitats. More complex sample-selection procedures, such as unequal probability sampling, could be applied. However, ease of application should be an overarching concern, and simple selection procedures, such as systematic sampling, would be preferable.

Some population abundances or densities are unlikely to be detected without substantial sampling effort by increasing search efficiency or search area (Table 2). This constraint is unavoidable in any protocol. The proposed survey design incorporates sampling techniques (i.e., transect-based, semiquantitative sampling) that are part of many existing protocols. However, the user of the proposed design can be fully aware of population sizes that are likely to be detected by explicitly stating the probability of detecting species presence for given population size and sampling effort. As one reviewer noted, a main advantage of the proposed design is that the user has an answer to the question: "How much sampling effort is enough?"

A reasonable concern with the proposed design is the cost to survey a site. The recommended sampling effort is likely to exceed the costs associated with currently applied protocols. Few protocols for rare species detection have been published; however, Young et al. (2001) recommended at least 2 person-hours of search time in optimal habitat before concluding that a species was absent if no individuals were detected. At a search rate of 2 min/m^2, a 2-h search would be equivalent to <100 m^2 of search area, which appears to be an insufficient effort for detecting rare species. A search area of 100 m^2 resulted in a probability of detecting species presence as low as 0.12

TABLE 4. Probabilities of detecting species presence observed from a computer simulation and predicted by eq. 4. Abundance (T), search efficiency (β), and cluster size are mean values used in the simulation, but were random variables in the simulation. Cluster locations were random within a 16,000-m^2 study site. Searches were conducted within 1 m $\times$ 100 m transects. The search area (a) was the sum of the transect areas. Bold font indicates combinations with predicted probabilities $\geq$0.85.

			a (m^2)							
			400		600		800		1000	
T	β	Cluster size	Simulated	Predicted	Simulated	Predicted	Simulated	Predicted	Simulated	Predicted
100	0.2	1	0.45	0.39	0.78	0.53	0.55	0.63	0.51	0.71
		3	0.36	0.39	0.73	0.53	0.61	0.63	0.89	0.71
		5	0.43	0.39	0.57	0.53	0.70	0.63	0.75	0.71
	0.4	1	0.76	0.63	0.95	0.78	**0.82**	**0.86**	**0.91**	**0.92**
		3	0.63	0.63	0.89	0.78	**0.79**	**0.86**	**1.00**	**0.92**
		5	0.49	0.63	0.65	0.78	**0.82**	**0.86**	**0.92**	**0.92**
	0.6	1	0.89	0.78	**0.97**	**0.89**	**1.00**	**0.95**	**0.95**	**0.98**
		3	0.67	0.78	**0.87**	**0.89**	**0.96**	**0.95**	**0.94**	**0.98**
		5	0.72	0.78	**0.86**	**0.89**	**0.95**	**0.95**	**0.73**	**0.98**
300	0.2	1	0.75	0.78	**0.97**	**0.89**	**0.95**	**0.95**	**1.00**	**0.98**
		3	0.73	0.78	**0.78**	**0.89**	**1.00**	**0.95**	**1.00**	**0.98**
		5	0.79	0.78	**0.69**	**0.89**	**0.97**	**0.95**	**0.96**	**0.98**
	0.4	1	**0.92**	**0.95**	**1.00**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**
		3	**0.83**	**0.95**	**1.00**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**
		5	**0.94**	**0.95**	**1.00**	**0.99**	**0.93**	**1.00**	**1.00**	**1.00**
	0.6	1	**1.00**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		3	**0.88**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		5	**0.95**	**0.99**	**0.97**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
500	0.2	1	**0.92**	**0.92**	**1.00**	**0.98**	**1.00**	**0.99**	**1.00**	**1.00**
		3	**0.90**	**0.92**	**0.92**	**0.98**	**1.00**	**0.99**	**1.00**	**1.00**
		5	**0.95**	**0.92**	**0.96**	**0.98**	**1.00**	**0.99**	**1.00**	**1.00**
	0.4	1	**1.00**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		3	**0.96**	**0.99**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		5	**1.00**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
	0.6	1	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		3	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		5	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**

and <0.85 for all but one combination of abundance and search efficiency in Table 2. If this result is any indication, using the proposed survey design would lead to increased sampling effort and higher survey costs than currently practiced. A legitimate and reasonable question is whether the added cost is worthwhile and affordable. Ultimately, that question will have to be answered on a case-by-case basis by the organizations that are funding the survey. One counterbalancing consideration is the cost of failing to detect the presence of a rare population within the area of a pending adverse impact. Cost would be reduced if searching stopped as soon as one individual of the rare species was detected; however, that practice would limit the utility of the survey. There certainly are circumstances when designing a survey to achieve a high probability of detecting species presence will be worthwhile. Surveys of federally endangered species in areas of proposed adverse impacts would probably be one of those circumstances.

Acknowledgements

This paper was motivated by my participation in a review of survey protocols for detecting rare populations in the Allegheny River prior to sand and gravel dredging. The review was organized by Bob Carline and sponsored by the Pennsylvania Fish and Boat Commission. I appreciate the informative and thoughtful discussions on mussel sampling with the review panel members: Bob Carline, Drew Miller, Dick Neves, and Gerald Dinkins. I thank Paul Geissler, Jim Nichols, Penelope Pooler, Jennifer Brown, Rita Villella, Caryn Vaughn, and 3 anonymous referees for helpful comments on the paper.

Literature Cited

AMYOT, J.-P., AND J. A. DOWNING. 1991. Endo- and epibenthic distribution of the unionid mollusk *Elliptio complanata*. Journal of the North American Benthological Society 10: 280–285.

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.

Di Maio, J., and L. D. Corkum. 1995. Relationship between the spatial distribution of freshwater mussels (Bivalvia: Unionidae) and the hydrological variability of rivers. Canadian Journal of Zoology 73:663–671.

Di Maio, J., and L. D. Corkum. 1997. Patterns of orientation in unionids as a function of rivers with differing hydrological variability. Journal of Molluscan Studies 63:531–539.

Downing, J. A., and W. L. Downing. 1991. 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. Freshwater Biological Association Scientific Publication 25. Freshwater Biological Association, Ambleside, UK.

Green, R. H., and R. C. Young. 1993. Sampling to detect rare species. Ecological Applications 3:351–356.

Haag, W. R., and M. L. Warren. 2000. Effects of light and presence of fish on lure display and larval release behaviours in two species of freshwater mussels. Animal Behaviour 60:879–886.

Haukioja, E., and T. Hakala. 1974. Vertical distribution of freshwater mussels (Pelecypoda, Unionidae) in southwestern Finland. Annales Zoologici Fennici 11:127–130.

Hornbach, D. J., and T. Deneka. 1996. A comparison of a qualitative and a quantitative collection method for examining freshwater mussel assemblages. Journal of the North American Benthological Society 15:587–596.

Kovalak, W. P., S. D. Dennis, and J. M. Bates. 1986. Sampling effort required to find rare species of freshwater mussels. Pages 34–45 *in* B. G. Isom (editor). Rationale for sampling and interpretation of ecological data in the assessment of freshwater ecosystems. ATSM STP 894. American Society for Testing and Materials, Philadelphia, Pennsylvania.

McArdle, B. H. 1990. When are rare species not there? Oikos 57:276–277.

McDonald, L. L. 2004. Sampling rare populations. Pages 11–42 *in* W. L. Thompson (editor). Sampling rare or elusive species: concepts, designs, and techniques for estimating population parameters. Island Press, Covelo, California.

Metcalfe-Smith, J. L., J. Di Maio, S. K. Staton, and G. L. Mackie. 2000. Effect of sampling effort on the efficiency of the timed search method for sampling freshwater mussel communities. Journal of the North American Benthological Society 19:725–732.

Miller, A. C., and B. S. Payne. 1988. The need for quantitative sampling to characterize size demography and density of freshwater mussel communities. American Malacological Bulletin 6:49–54.

Miller, A. C., and B. S. Payne. 1993. Qualitative versus quantitative sampling to evaluate population and community characteristics at a large-river mussel bed. American Midland Naturalist 130:133–145.

Obermeyer, B. K. 1998. A comparison of quadrats versus timed snorkel searches for assessing freshwater mussels. American Midland Naturalist 139:331–339.

Perles, S. J., A. D. Christian, and D. J. Berg. 2003. Vertical migration, orientation, aggregation, and fecundity of the freshwater mussel *Lampsilis siliquoidia*. Ohio Journal of Science 103:73–78.

Richardson, T. D., and P. Yokley. 1996. A note on sampling technique and evidence of recruitment in freshwater mussels (Unionidae). Archiv für Hydrobiologie 137:135–140.

Smith, D. R., R. F. Villella, and D. P. Lemarié. 2001a. Survey protocol for assessment of endangered freshwater mussels in the Allegheny River. 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.

Smith, D. R., R. F. Villella, D. P. Lemarié, and S. von Oettingen. 2001b. How much excavation is needed to monitor freshwater mussels? Pages 203–218 *in* R. A. Tankersley, D. I. Warmolts, G. T. Watters, B. J. Armitage, P. D. Johnson, and R. S. Butler (editors). Freshwater Mollusk Symposium Proceedings. Ohio Biological Survey, Columbus, Ohio.

Stevens, D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262–278.

Strayer, D. L., S. Claypool, and S. J. 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 J. Ralley. 1993. Microhabitat use by an assemblage of stream-dwelling unionaceans (Bivalvia), including two rare species of *Alasmidonta*. Journal of the North American Benthological Society 12:247–258.

Strayer, D. L., and D. R. Smith. 2003. A guide to sampling freshwater mussel populations. American Fisheries Society Monograph No. 8. American Fisheries Society, Bethesda, Maryland.

Strayer, D. L., S. Sprague, and S. Claypool. 1996. A range-wide assessment of populations of *Alasmidonta heterodon*, an endangered freshwater mussel (Bivalvia: Unionidae). Journal of the North American Benthological Society 15: 308–317.

Vaughn, C. C., C. M. Taylor, and K. J. Eberhard. 1997. A comparison of the effectiveness of timed searches vs. quadrat sampling in mussel surveys. Pages 157–162 *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.)

VILLELLA, R. F., AND D. R. SMITH. 2005. Two-phase sampling to estimate river-wide populations of freshwater mussels. Journal of the North American Benthological Society 24: 357–368.

WATTERS, G. T., S. H. O'DEE, AND S. CHORDAS. 2001. Patterns of vertical migration in freshwater mussels (Bivalvia: Unionida). Journal of Freshwater Ecology 16:541–550.

WELSH, A. H., R. B. CUNNINGHAM, C. F. DONNELLY, AND D. B. LINDENMAYER. 1996. Modelling the abundance of rare species: statistical models for counts with extra zeros. Ecological Modelling 88:297–308.

WILCOX, D. B., D. D. ANDERSON, AND A. C. MILLER. 1993. Survey protocol and decision criteria for estimating the likelihood that *Lampsilis higginsi* is present in areas within the Upper Mississippi River System. Pages 163–167 *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.)

WILLIAMS, B. K., J. D. NICHOLS, AND M. J. CONROY. 2002. Analysis and management of animal populations. Academic Press, New York.

YOUNG, M. R., P. J. COSGROVE, L. C. HASTIE, AND B. HENNIGER. 2001. A standardised method for assessing the status of freshwater mussels in clear, shallow rivers. Journal of Molluscan Studies 67:395–405.

Received: 25 August 2005
Accepted: 17 March 2006