

Designing occupancy studies: general advice and allocating survey effort

DARRYL I. MACKENZIE* and J. ANDREW ROYLE†

*Proteus Wildlife Research Consultants, PO Box 5193, Dunedin, New Zealand; †US Geological Survey, Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, MD 20708–4017, USA

Summary

1. The fraction of sampling units in a landscape where a target species is present (occupancy) is an extensively used concept in ecology. Yet in many applications the species will not always be detected in a sampling unit even when present, resulting in biased estimates of occupancy. Given that sampling units are surveyed repeatedly within a relatively short timeframe, a number of similar methods have now been developed to provide unbiased occupancy estimates. However, practical guidance on the efficient design of occupancy studies has been lacking.

2. In this paper we comment on a number of general issues related to designing occupancy studies, including the need for clear objectives that are explicitly linked to science or management, selection of sampling units, timing of repeat surveys and allocation of survey effort. Advice on the number of repeat surveys per sampling unit is considered in terms of the variance of the occupancy estimator, for three possible study designs.

3. We recommend that sampling units should be surveyed a minimum of three times when detection probability is high ($> 0.5 \text{ survey}^{-1}$), unless a removal design is used.

4. We found that an optimal removal design will generally be the most efficient, but we suggest it may be less robust to assumption violations than a standard design.

5. Our results suggest that for a rare species it is more efficient to survey more sampling units less intensively, while for a common species fewer sampling units should be surveyed more intensively.

6. *Synthesis and applications.* Reliable inferences can only result from quality data. To make the best use of logistical resources, study objectives must be clearly defined; sampling units must be selected, and repeated surveys timed appropriately; and a sufficient number of repeated surveys must be conducted. Failure to do so may compromise the integrity of the study. The guidance given here on study design issues is particularly applicable to studies of species occurrence and distribution, habitat selection and modelling, metapopulation studies and monitoring programmes.

Key-words: detection probability, habitat modelling, occupancy models, patch occupancy, proportion of area occupied, species distribution, species occurrence, study design

Journal of Applied Ecology (2005) **42**, 1105–1114

doi: 10.1111/j.1365-2664.2005.01098.x

Introduction

As a general concept, the fraction of sampling units in a landscape where a target species is present (occupancy) is of considerable interest in ecology. Applications that focus on occupancy-related metrics include measures

of species occurrence, range and distribution (Brown 1995; Wikle 2003; Engler, Guisan & Rechsteiner 2004), habitat selection and modelling (Reunanen *et al.* 2002; Scott *et al.* 2002; Bradford *et al.* 2003; Gibson *et al.* 2004), metapopulation studies (Hanski 1994, 1999) and wildlife monitoring programmes (Zielinski & Stauffer 1996; Trenham *et al.* 2003; Weber, Hinterman & Zangger 2004). Occupancy has also been used in relation to a wide range of taxa, including owls in north-western USA (Azuma, Baldwin & Noon 1990), salamanders in eastern USA

Correspondence: Darryl I. MacKenzie, Proteus Wildlife Research Consultants, PO Box 5193, Dunedin, New Zealand (fax + 64 3486 1168; e-mail darryl@proteus.co.nz).

(Bailey, Simons & Pollock 2004), butterflies in Finland (Hanski 1994) and Wales (Cabeza *et al.* 2004) and tigers *Panthera tigris* in India (Nichols & Karanth 2002) and Malaysia (Kawanishi & Sunquist 2004). Therefore, the use of appropriate field and analytic methods in occupancy studies is applicable to scientists working in a wide range of ecological disciplines, on a large number of different taxa.

However, it has long been acknowledged that a species may go undetected in a survey of a sampling unit even when the species is actually present within that unit. Unaccounted for 'false absences' will lead to underestimates of the true level of occupancy, and estimates of the relative change in occupancy will only be valid if the surveyors' ability to detect the species is identical in the two time periods (MacKenzie 2005a). The imperfect detection of a species also has serious consequences for habitat models. Tyre *et al.* (2003) and Gu & Swihart (2004) found that false absences caused estimates of habitat effects to be biased even at modest levels, particularly if detection probability varied between habitats. Importantly, resulting inferences about the 'value' of different habitats could be severely misleading if detection probabilities are correlated with occupancy probabilities (Gu & Swihart 2004; MacKenzie 2006). In a meta-population context, Moilanen (2002) assessed the effect of a range of assumption violations on the estimation of incidence function parameters via computer simulation. Moilanen (2002) concluded that false absences were a greater source of bias than inaccurately recorded patch sizes or unknown habitat patches present within the study area. Furthermore, Moilanen (2002) recommended that field studies should be designed to minimize false absences and suggested that, given the level of bias encountered, the additional effort would be worthwhile.

To counter the effect of imperfect detection, one solution has been to conduct multiple surveys of the sampling units within a relatively short timeframe to minimize the possibility of a false absence. After a number of surveys, if the species has not been detected then one may assume that the species is absent from the sampling unit. However, there have been a number of independently developed techniques that advocate a more efficient use of such data, estimating detection probabilities explicitly that lead to unbiased estimates of occupancy (Giessler & Fuller 1987; Azuma, Baldwin & Noon 1990; MacKenzie *et al.* 2002; Royle & Nichols 2003; Tyre *et al.* 2003; Stauffer, Ralph & Miller 2004). In our opinion, the conceptual modelling framework as described by MacKenzie *et al.* (2002; its extension in Royle & Nichols 2003), is the most flexible and all the other approaches could be considered special cases (MacKenzie 2005a). In its most general form, the modelling approach of MacKenzie *et al.* (2002) could be considered as performing simultaneous logistic regression analyses on both occupancy and detection probabilities.

While the advent of such estimation techniques provides robust methods for analysing such data, little has been published on practical steps that should be taken

to ensure the study has been designed appropriately and efficiently. A key aspect of designing occupancy studies is the number of repeated surveys that should be conducted. Stauffer, Ralph & Miller (2002) took the approach of determining the number of surveys required to have 0.95 probability of detecting the species at a site if it was present, i.e. a 0.05 probability of declaring the species as falsely absent. MacKenzie *et al.* (2002) attempted to give guidance on the number of surveys required to provide a 'reasonable' estimate of occupancy based upon a simulation study. They suggested that a minimum of two repeated surveys could be used if occupancy was > 0.7 and detection probabilities (in a single survey) > 0.3 , but the precision of an estimate of occupancy may be poor. Tyre *et al.* (2003) also used simulation results in an attempt to give some guidance on the number of surveys required at each sampling unit. They recommended that when the probability of a false absence is low (i.e. detection probabilities are high) it is better to survey more units rather than increasing the number of surveys per sampling unit, but as detection probabilities decrease then more surveys per unit should be conducted. Most recently, Field, Tyre & Possingham (2005) have used simulation methods to investigate issues related to study design with respect to the power of a study to detect a decline in the level of occupancy over a 3-year period, subject to realistic budgetary constraints. They concluded that generally two to three repeat surveys per site would generally be sufficient unless occupancy was high or detection probability was low.

In this paper we assume that the ultimate goal of an occupancy-type study is to obtain as precise an estimate of occupancy as possible for a given level of total survey effort, or to achieve a desired level of precision with minimal effort. Our guidance on the most efficient allocation of survey effort was developed by considering the asymptotic variance of the occupancy estimator under three different sampling schemes. Analytic results were used to investigate how the variance of the estimator changes with respect to the number of sampling units surveyed and number of surveys per unit. However, we begin this paper by providing advice on general issues that need to be considered when designing an occupancy study, based upon our own experiences and those of our colleagues. These general issues are very important if one wishes to make reliable inferences about the population of interest as they can greatly aid in the design process or subtly alter how one should interpret parameter estimates.

Points for general consideration

THE WHY, WHAT AND HOW OF DESIGNING A STUDY

When designing an occupancy study, the general problem is how to select the sampling units (generically referred to as sites henceforth) from the population or area of interest, and the number of surveys per site, in order to

achieve the objective of the study. This simple sentence highlights one of the main issues that in our experience is often poorly addressed in many wildlife studies and monitoring programmes: the need for a clear objective.

Yoccoz, Nichols & Boulinier (2001) note that the three key questions that must be considered when designing a monitoring programme (which are also relevant for any scientific study) are why, what and how: why collect the data; what type of data to collect; and how should the data be collected in the field and then analysed. Formation of a clear objective is a critical part of addressing the why question, and is aided by the realization that data collection as a stand-alone activity is not of great inherent utility, but should be viewed as merely a component of advancing science or good management. A good study objective should be explicitly linked to how the data will be used to discriminate between scientific hypotheses about a system or how the data will be used to make management decisions (Pollock *et al.* 2002). Note that the need for a clear objective is exemplified when using a decision-theoretic approach to management (Williams, Nichols & Conroy 2002; Field, Tyre & Possingham 2005), but it is equally important in other contexts as different study designs may be more appropriate for some objectives than for others. For example, two potential objectives for a study might be to: (i) determine the overall level of occupancy for a species in a region; (ii) compare the level of occupancy in two different habitat types within that region. Two potential designs that could be used are: (a) randomly select sites to survey from throughout the entire region; (b) stratify the region according to habitat type, then randomly select sites only from the two strata of interest. Design (a) could be used for both objectives, although it may be an inefficient design for objective (ii) as some effort will be used to survey sites in habitat types not of interest, effectively reducing the sample size for the comparison. Design (b) would therefore be a much more efficient design for objective (ii) but would not be useful for objective (i) as some areas of the region will be excluded from the sampling.

In consideration of the what question, there are three levels at which data could generally be collected in demographic studies of wildlife populations: (i) individual organisms; (ii) individual species; (iii) the ecological community or multiple species. The respective state variables that can be used to characterize the current overall status of a population are abundance, occupancy (or extent of occurrence) and species richness. With a well-specified objective, the question of what data to collect should be a relatively simple one, although logistical considerations should also be taken into account with abundance-type state variables requiring the highest level of field effort. Here we primarily focus on design issues with respect to an occupancy state variable, but some comments are relevant more generally.

Typically the how question is the one that receives the greatest attention when designing a study, and the remainder of this paper is largely devoted to providing

guidance on how to design an occupancy study. However, we stress that attention should only be devoted to how once the why and what questions have been suitably addressed. In our experience, the lack of clear objectives will often lead to endless debate about design issues as there has been no specification for how the collected data will be used in relation to science and/or management; hence judgements about whether the 'right' data will be collected can not be made.

SITE SELECTION

Generally, the sites from which data are collected will only represent a fraction of the greater collection of sites of which the occupancy state is of interest. For example, interest might focus on all ponds in a national park, or all quadrats within a contiguous habitat, but only a relatively small fraction of ponds or quadrats will be surveyed. It is therefore necessary that the manner in which sites are selected allows the results of the data analysis to be generalized to the entire population. The estimation method of MacKenzie *et al.* (2002) (and similar methods mentioned above) implicitly assumes sites are randomly sampled from the greater population, although they could be easily generalized should stratified random sampling be used (by analysing the data for each stratum separately then combining them using standard stratified random sampling results; Cochran 1977). If other probabilistic sampling schemes are used to select sites (e.g. unequal probability sampling or adaptive sampling), then conceivably the above methods could be adapted to provide reliable inferences about occupancy in the greater population. However, if a non-probabilistic sampling scheme is used (e.g. sites are selected haphazardly or arbitrarily) then generalization of the results to the greater population involves a 'leap of faith' and is no longer based on any statistical theory.

In order to obtain unbiased estimates of occupancy for the entire population, it is preferable that sites are not selected based upon pre-existing knowledge about their potential occupancy state. A common example of this is studies where only sites of historic occupancy within the population are surveyed (i.e. sites that were once known to be occupied by the species). For example, the level of occupancy would probably be higher in a sample of 100 historic sites than in a sample of 100 sites randomly selected from the entire population. This issue also has important ramifications for studying change in occupancy over time. If the entire population is governed by the same processes of change in occupancy (colonizations of previously unoccupied sites and local extinctions of occupied sites; MacKenzie *et al.* 2003), then at a sample of sites that has an initially biased estimate of occupancy for the population (i.e. by surveying only historic sites) an apparent trend will be observed even if occupancy is stable for the population in general (MacKenzie *et al.* 2005). However, surveying only historic sites may be appropriate in situations where

the collection of historic sites actually represents the population of prime interest (e.g. to study the effects of human-induced disturbances at sites with historic records of a species). Hence, note the importance of the objective for determining whether sites have been selected in an appropriate manner for a given study.

TIMING OF REPEAT SURVEYS

Repeated surveys of the sites are often conducted as multiple discrete visits (e.g. on different days); however, discrete visits may not always be necessary. Other options include conducting multiple surveys within a single visit; using multiple observers to conduct independent surveys, either on the same or a different visit; surveying multiple plots within a larger site on a single visit (e.g. short randomly located transects within a 5-ha area of forest). The decision of which approach is most practical depends upon the study objective, whether the model assumptions are likely to be satisfied given the biology of the species, and the logistical considerations of sampling the species.

The first issue to consider is the timeframe over which the repeated surveys are to be conducted across the entire study region. Generally the intent of such studies is to provide a snapshot of the system at a given point in time; therefore it would seem reasonable to survey all sites as quickly as possible. The potential for change in the system increases the longer it takes to collect survey data from all sites, blurring understanding of the system. A basic assumption of the estimation methods noted above is that sites are closed to changes in occupancy for the duration of the repeat surveys. MacKenzie *et al.* (2004) suggest that this assumption can be relaxed and that, provided changes in the occupancy status of sites occur at random (i.e. the probability of occupancy in one time interval does not depend upon the occupancy status of a site in the previous time interval), the above estimation methods are valid, except that 'occupancy' should now be interpreted as 'use'. From a design perspective, the relaxation of this assumption may be important, as it changes the meaning and interpretation of the occupancy parameter. The proportion of area 'used' by a species will often be larger than the proportion of area where the species physically occurs. If occupancy is being employed as a surrogate for abundance estimation, the level of 'use' may be irrelevant and even misleading. For example, for a low-density highly mobile species such as a large carnivore, or a species with relatively large home ranges compared with the size of the sampling units, the proportion of area 'used' over a longer timeframe may be close to 100% even though population size is very small. Hence, if 'occupancy' at a single point in time is truly desired, then repeat surveys need to be conducted as quickly as possible (possibly within the same visit) to reduce the chance of the species moving among sites.

One of the main issues to be aware of when considering whether repeat surveys should be conducted during a

single visit or whether multiple visits are required, is the potential for introducing heterogeneity into the data. If the probability of detecting the species (given presence) varies among sites, then occupancy will be underestimated (MacKenzie *et al.* 2002; Royle & Nichols 2003). While heterogeneity can be accommodated either by covariates (MacKenzie *et al.* 2002) or by assuming a distributional form for the detection probabilities (MacKenzie *et al.* 2005), a better approach is to design a study to minimize the potential effects of heterogeneity from the outset. In most practical situations detection probability could be expected to vary at some time scale (e.g. days). If the time scale at which sites are visited corresponds with the time scale at which detection probabilities vary (e.g. one site is visited per day) and the repeat surveys are conducted within a single visit, then the design induces a form of heterogeneity. That is, if site A is only surveyed on day 1 and site B is only surveyed on day 2, and detection probabilities are different on days 1 and 2; then, as a result of the design, the detection probabilities for sites A and B will be different. Even with multiple visits the study design can induce heterogeneity in a similar manner, for example if the same observers always survey the same sites, or if the specific sites are always surveyed at the same times of day. In general, when determining how the repeat surveys should be conducted, it is necessary to consider the potential sources of variation in detection probabilities and how these may be correlated under potentially different designs. The intent should then be to use a design that breaks any such correlation structure. For instance, MacKenzie *et al.* (2004) suggest that, to reduce the effect of heterogeneity as a result of observer and 'time of day' effects, observers should be rotated amongst the sites that are to be surveyed on any given day and that the order in which sites are surveyed be changed each day.

Allocation of survey effort

Throughout the following section we make the simplifying assumption that occupancy (ψ) and detection probabilities (p) are constant across both space and time. While these assumptions may not always be reasonable in practice, it is usually necessary to make some simplifications of reality when designing a study. It should also be noted that, when designing studies, initial values need to be assumed for the population parameters of interest (here ψ and p).

We consider three general sampling schemes that have been used or proposed in the literature: (i) a standard design where s sites are each surveyed K times; (ii) a double sampling design where s_K sites are surveyed K times and s_1 sites surveyed once (note it is 'double' sampling in the sense that a second round of sampling is performed to select the sites at which the repeated surveys will be conducted); and (iii) a removal design where s sites are surveyed up to a maximum of K times, but surveying halts at a site once the species is detected.

For each design we also consider the impact of a cost function, where the cost of conducting subsequent surveys may be different from that of an initial survey (which may arise, for example, if multiple surveys are conducted during the same visit or if there is a set-up cost for establishing a new site).

We assume a general situation where the study is to be designed with an objective based on the variance of the occupancy estimator $\hat{\psi}$ ($\text{var}(\hat{\psi})$). Specifically, the study is to be designed either to (i) achieve a desired level of precision for minimal total survey effort; or (ii) minimize the variance for a given total number of surveys. The intent is therefore to determine what values of s and K will most efficiently achieve the study's objective, given the assumed values of ψ and p .

STANDARD DESIGN

For a standard design (where all sites are surveyed K times), using the MacKenzie *et al.* (2002) maximum likelihood approach to estimating ψ , it can be shown that the asymptotic variance for $\hat{\psi}$ (derived from the Fisher's information matrix; Williams, Nichols & Conroy 2002) is:

$$\text{var}(\hat{\psi}) = \frac{\psi}{s} \left[(1 - \psi) + \frac{(1 - p^*)}{p^* - Kp(1 - p)^{K-1}} \right] \quad \text{eqn 1}$$

where $p^* = 1 - (1 - p)^K$ is the probability of detecting the species at least once during K surveys of an occupied site. Note that as p^* approaches 1.0, $\text{var}(\hat{\psi}) \rightarrow \psi(1 - \psi)/s$, the variance for a simple binomial proportion. Further, the total number of surveys (TS) in a standard design will be:

$$TS = s \times K \quad \text{eqn 2}$$

If the study is to be designed such that $\hat{\psi}$ should achieve a desired level of precision, then to find the optimum combination of s and K the basic procedure would be to rearrange equation 1 to make s the subject and substitute into equation 2, to give:

$$TS = \frac{\psi K}{\text{var}(\hat{\psi})} \left[(1 - \psi) + \frac{(1 - p^*)}{p^* - Kp(1 - p)^{K-1}} \right] \quad \text{eqn 3}$$

As values for ψ and p will be assumed, K is the only unknown in equation 3. Therefore the minimum number of surveys required to obtain a given level of precision can be found by differentiating equation 3 with respect to K , setting to zero and solving for K . This may be done analytically or numerically. Once an optimum value for K has been found, this can be substituted into the rearranged equation 1 to give the optimum number of sites to survey.

Alternatively, if the study is to be designed in terms of minimizing the variance for a fixed total number of surveys, then equation 2 should be rearranged to make s the subject, and substituted in equation 1, giving:

$$\text{var}(\hat{\psi}) = \frac{\psi K}{TS} \left[(1 - \psi) + \frac{(1 - p^*)}{p^* - Kp(1 - p)^{K-1}} \right] \quad \text{eqn 4}$$

Equation 4 would then be minimized with respect to K , set to zero and solved for K . The optimum value of K could then be substituted into the rearranged equation 2 to give the optimum number of sites to survey.

However, consider the similar forms of equations 3 and 4. In both cases they could be expressed as:

$$f(K) = CK \left[(1 - \psi) + \frac{(1 - p^*)}{p^* - Kp(1 - p)^{K-1}} \right] \quad \text{eqn 5}$$

where C is a constant with respect to K . Therefore, the value of K that minimizes $f(K)$ will not depend upon C . This means that, regardless of whether the study is designed to minimize total survey effort to achieve a specified value of $\text{var}(\hat{\psi})$ or to minimize $\text{var}(\hat{\psi})$ for a fixed level of total survey effort, the optimum value of K will be the same. How the study is designed only determines the optimum number of sites to survey. This result is useful as it means standard tables can be used to give the optimum number of surveys that should be conducted at each site for given values of ψ and p .

The above results assume that the cost of conducting surveys is immaterial, or that the cost is equal across all sites and all surveys. However, in reality, costs will often be one of the major limiting factors when designing a study. When the cost associated with conducting a survey may vary either between sites or between survey occasions, then the above approach can be generalized such that an optimal design can be found given a specific cost function. Here we only consider cost functions of the form:

$$\text{Cost} = c_0 + s[c_1 + c_2(K - 1)]$$

where c_0 is a fixed overhead cost, c_1 is the cost of conducting the first survey of a site, and c_2 is the cost of conducting subsequent surveys, although other cost functions could be considered (Field, Tyre & Possingham 2005; MacKenzie *et al.* 2005).

Once the cost function has been defined, then it is possible to design a study either in terms of (i) minimizing cost while obtaining a desired level of precision, or (ii) minimizing the variance given a fixed total budget. For situations where the cost of conducting surveys does not vary among sites (as is the case here), then a similar result to above holds where the optimal number of surveys to conduct per site is independent of whether a study is designed in terms of minimizing total cost or minimizing the variance of the occupancy estimate. This means tables can again be constructed of the optimal number of surveys to conduct at each site for given relative costs of an initial to a subsequent survey.

In Table 1 we present the optimal number of surveys per site (K) where the cost of an initial survey is equal to, five times greater and 10 times greater than the cost

p	x	ψ								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	1	14	15	16	17	18	20	23	26	34
	5	16	17	18	19	20	22	24	28	35
	10	18	19	20	21	22	24	26	30	37
0.2	1	7	7	8	8	9	10	11	13	16
	5	9	9	9	10	11	11	12	14	17
	10	10	10	11	11	12	13	14	15	18
0.3	1	5	5	5	5	6	6	7	8	10
	5	6	6	6	7	7	8	8	9	11
	10	7	7	7	8	8	9	9	10	12
0.4	1	3	4	4	4	4	5	5	6	7
	5	4	5	5	5	5	6	6	7	8
	10	5	5	6	6	6	6	7	8	9
0.5	1	3	3	3	3	3	3	4	4	5
	5	4	4	4	4	4	4	5	5	6
	10	4	4	4	5	5	5	5	6	7
0.6	1	2	2	2	2	3	3	3	3	4
	5	3	3	3	3	3	4	4	4	5
	10	3	3	4	4	4	4	4	5	5
0.7	1	2	2	2	2	2	2	2	3	3
	5	2	2	3	3	3	3	3	3	4
	10	3	3	3	3	3	3	3	4	4
0.8	1	2	2	2	2	2	2	2	2	2
	5	2	2	2	2	2	2	2	3	3
	10	2	2	2	2	2	3	3	3	3
0.9	1	2	2	2	2	2	2	2	2	2
	5	2	2	2	2	2	2	2	2	2
	10	2	2	2	2	2	2	2	2	2

of a subsequent survey for selected values of ψ and p . The first thing to note with Table 1 is that the optimal value for K is never 1. That is, whenever the probability of detecting a species is < 1 , the most efficient use of resources is never to survey all sites only once (in fact, occupancy and detection probabilities are not identifiable without auxiliary information if such a design was used). Further, Table 1 suggests that the optimal number of surveys required for each site decreases as detection probability increases. However, an interesting aspect of Table 1 is that the optimal value for K increases as the probability of occupancy also increases. This implies that an optimal strategy for rare species is to conduct fewer surveys at more sites, while for a common species the optimal strategy is to conduct more surveys at fewer sites. A further interesting point related to Table 1 is that, given the optimal values for K it is possible to calculate the optimal probability of detecting the species at least once at an occupied site (p^* ; i.e. the probability of confirming the target species is present at a site). While not presented here, generally the optimal surveying strategy requires a reasonable degree of confirmation that the target species occupies a site ($0.85 < p^* < 0.95$). The optimal value for K generally changes little for the type of cost function considered here, although note that when subsequent surveys can be conducted relatively cheaply, an optimal strategy is to increase the number of repeat surveys.

DOUBLE SAMPLING

A double sampling design (where repeat surveys are conducted at a subset of sites only) is completely compatible with the modelling approach of MacKenzie *et al.* (2002). Initially, double sampling appears attractive as it seems reasonable that at some point the collection of additional information about detectability (by repeated surveys) may be inefficient, and there is greater benefit (in terms of precision) in increasing the total number of sites that are surveyed. Such a design has been proposed by MacKenzie *et al.* (2002, 2003), MacKenzie, Bailey & Nichols (2004) and MacKenzie (2005b). The above approaches can be generalized to determine the optimal allocation of sampling effort between sites and repeated surveys, when a double sampling scheme is being used.

When a double sampling scheme is used with s_K sites surveyed K times and s_1 sites surveyed once, assuming detection probability is constant, it can be shown that the asymptotic variance for $\hat{\psi}$ is:

$$\text{var}(\hat{\psi}) = \frac{\psi}{s_K + s_1} [(1 - \psi) + D_1 + D_2] \tag{eqn 6}$$

where

$$D_1 = \frac{(s_K + s_1)[(1 - p^*)s_K K(1 - \psi p) + (1 - p)s_1(1 - \psi p^*)]}{s_K[p^* - Kp(1 - p)^{K-1}][s_K K(1 - \psi p) + s_1[(1 - p) + (1 - \psi)Kp]]}$$

and

Table 2. Optimal fraction of total survey effort (expressed as a percentage) that should be used to survey s_1 sites only once using a double sampling design, where cost of the first and subsequent surveys are equal

p	ψ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0	0	0	0	0	0	0	0	0
0.2	0	0	0	0	0	0	0	0	0
0.3	0	0	0	0	0	0	0	0	0
0.4	0	3	0	0	0	0	0	0	0
0.5	6	1	0	0	0	0	0	0	0
0.6	0	0	0	12	4	0	0	0	0
0.7	9	5	0	0	0	0	0	0	0
0.8	33	30	26	21	14	5	0	0	0
0.9	56	54	51	48	44	39	31	17	0

$$D_2 = \frac{s_1(K-1)(1-\psi)(1-p)}{s_K K(1-\psi p) + s_1[(1-p) + (1-\psi)Kp]}$$

While appearing unwieldy, note that the general form of equation 6 is similar to equation 1: a simple binomial proportion variance with additional penalty terms as a result of the imperfect detection of the species. As for $\text{var}(\hat{\psi})$ from a standard design, note that $\text{var}(\hat{\psi}) \rightarrow \psi(1-\psi)/(s_K + s_1)$ as $p^* \equiv 1$ and furthermore that, if $s_1 = 0$, then equation 6 equates to equation 1 (with $s_K = s$), as would be expected.

When the cost of initial and subsequent surveys are equal (or not an issue), it was found that generally there is little advantage in using an optimal double sampling design compared with an optimal standard design with the same number of total surveys. Table 2 presents the optimal fraction of total survey effort that should be used to survey s_1 sites only once. However, even when it is suggested that a reasonable fraction of the total survey effort should be used, the percentage improvement in the standard error compared with an optimal standard survey is small unless ψ is small and p is large. Hence, speculation by MacKenzie *et al.* (2002, 2003), MacKenzie, Bailey & Nichols (2004) and MacKenzie (2005b) that a double sampling scheme may generally be more efficient is unsubstantiated.

Given these results, it would be expected that a double sampling design would only become more efficient than a standard design in situations where the cost of surveying a new site for the first time is lower than resurveying a site that had been surveyed previously (i.e. where $c_1 < c_2$), which was confirmed using numerical approaches. As it is difficult to imagine a situation where this may occur in practice, a double sampling scheme may not be a good design in most circumstances.

REMOVAL SAMPLING

The logic behind a removal sampling scheme (where surveying of a site halts once the species is detected or K surveys have been conducted) is that the main piece of information with respect to occupancy has been collected once the species has been confirmed at a site. We refer to this type of design as a removal sampling scheme

as sites are removed from the pool of sites being surveyed once the species has been detected, and also because of the analogy with removal studies conducted on animal populations (where individual animals are physically removed from the population upon first capture; Otis *et al.* 1978; Williams, Nichols & Conroy 2002).

Using a removal sampling design, it can be shown that the asymptotic variance for $\hat{\psi}$ is:

$$\text{var}(\hat{\psi}) = \frac{\psi}{s} \left[(1-\psi) + \frac{p^*(1-p^*)}{(p^*)^2 - K^2 p^2 (1-p)^{K-1}} \right] \quad \text{eqn 7}$$

Again, note the general form of the equation and the fact that $\text{var}(\hat{\psi}) = \psi(1-\psi)/s$ as $p^* \equiv 1$.

As with a standard design, regardless of whether the study is to be designed in terms of achieving a specified level of precision for minimal effort or to minimize the variance for a fixed level of effort, equation 7 could be re-arranged into a form similar to equation 5, i.e.:

$$f(K) = C \left[(1-\psi) + \frac{p^*(1-p^*)}{(p^*)^2 - K^2 p^2 (1-p)^{K-1}} \right]$$

Again, this implies that there is an optimal value for K (now the maximum number of repeat surveys) that is consistent for either design approach. Note that these optimal values (Table 3) are generally larger than the values for the standard design (Table 1). To compare the relative efficiency of an optimal removal design with the optimal standard design, Table 4 presents the ratio of the expected standard errors for $\hat{\psi}$ for these two designs with the same (expected) total number of surveys. Values < 1 indicate situations where the optimal standard design is more efficient in terms of obtaining a smaller standard error, which only occurs when the level of occupancy is < 0.3 . This suggests that, generally, an optimal removal design is more efficient than an optimal standard design; however, the implication is that one must be prepared to conduct a greater maximum number of surveys in order to realize fully the gain in efficiency. For example, if $\psi = 0.8$ and $p = 0.3$ the standard error of an optimal standard design with eight repeat surveys per site will be 42% greater than that of an optimal removal design, but sites would have to be surveyed up to a maximum of 12 times.

Table 3. Optimal maximum number of surveys to conduct at each site for a removal design where all sites are surveyed until the species is first detected, where cost of the first and subsequent surveys are equal

<i>p</i>	ψ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	23	24	25	26	28	31	34	39	49
0.2	11	11	12	13	13	15	16	19	23
0.3	7	7	7	8	8	9	10	12	14
0.4	5	5	5	6	6	6	7	8	10
0.5	4	4	4	4	4	5	5	6	8
0.6	3	3	3	3	3	4	4	5	6
0.7	2	2	2	3	3	3	3	4	5
0.8	2	2	2	2	2	2	3	3	4
0.9	2	2	2	2	2	2	2	2	3

Table 4. Ratio of standard errors for optimal standard and removal designs, where cost of the first and subsequent surveys are equal. Values < 1 indicate situations where an optimal standard design has a smaller standard error

<i>p</i>	ψ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.90	0.94	0.98	1.04	1.10	1.18	1.30	1.46	1.74
0.2	0.91	0.94	0.99	1.04	1.10	1.18	1.28	1.44	1.71
0.3	0.92	0.95	0.99	1.04	1.10	1.17	1.27	1.42	1.68
0.4	0.93	0.96	0.99	1.03	1.09	1.17	1.26	1.40	1.64
0.5	0.93	0.96	1.00	1.04	1.08	1.16	1.24	1.37	1.60
0.6	0.94	0.97	1.01	1.06	1.09	1.15	1.22	1.35	1.55
0.7	0.95	0.96	0.97	1.01	1.07	1.13	1.22	1.31	1.48
0.8	1.00	1.02	1.04	1.07	1.09	1.11	1.15	1.25	1.45
0.9	1.02	1.05	1.07	1.10	1.13	1.17	1.20	1.24	1.31

The effect of incorporating differential costs for initial and subsequent surveys here is of a similar magnitude to that for the standard design. Unless detection probability is high (> 0.8), the optimal maximum number of surveys increases if the cost of an initial survey is substantially higher than the cost of a subsequent survey (e.g. $c_1 \geq 5c_2$) and decreases slightly if the subsequent survey cost is higher than the cost of an initial survey.

EXAMPLE: DESIGNING A STUDY TO ACHIEVE A SPECIFIED PRECISION FOR $\hat{\psi}$

Consider a situation where we wish to conduct a study where it is thought that $\psi \approx 0.7$ and $p \approx 0.4$, and it is assumed all surveys will have the same cost. The study is to be designed such that the estimated level of occupancy has a standard error of 0.04, using as few surveys as possible. If a standard design is used, then from Table 1 the optimal number of repeat surveys per site is 5, and the probability of detecting the species at least once is $p^* = 1 - (1 - 0.4)^5 = 0.92$. To determine the number of sites to survey, the respective values can be inserted into equation 1, and solved for s , i.e.:

$$0.04^2 = \frac{0.7}{s} \left[(1 - 0.7) + \frac{(1 - 0.92)}{0.92 - 5 \times 0.4(1 - 0.4)^{5-1}} \right]$$

$$s = \frac{0.7}{0.04^2} \left[0.3 + \frac{0.08}{0.92 - 0.26} \right]$$

$$= 437.5[0.3 + 0.12]$$

$$\approx 183$$

(note there may be some small discrepancy as a result of rounding errors). Based on the above results, this suggests that surveying 183 sites, each five times (915 total surveys), should be the most efficient allocation of resources for a standard design, provided the assumed values for occupancy and detectability are reasonable.

If a removal design was to be used, then from Table 3 the maximum number of surveys per site is 7. Given the assumed values, now $p^* = 0.97$, which gives the number of sites to survey as (from equation 7):

$$0.04^2 = \frac{0.7}{s} \left[(1 - 0.7) + \frac{0.97(1 - 0.97)}{0.97^2 - 7^2 \times 0.4^2(1 - 0.4)^{7-1}} \right]$$

$$s = \frac{0.7}{0.04^2} \left[0.3 + \frac{0.03}{0.97^2 - 0.37} \right]$$

$$= 437.5[0.3 + 0.05]$$

$$\approx 152$$

Therefore surveying 152 sites until first detection of the species will give a design with an expected standard error for $\hat{\psi}$ of 0.04. As the decision of when to stop surveying a site relies upon an element of chance, the total number of surveys required for a removal design is actually a random variable but, for the above situation, the expected number of surveys required is 578.

In this instance, using a standard design may require 58% more surveys than a removal design to obtain a standard error of 0.04, although surveyors must be prepared to survey sites up to seven times rather than consistently surveying all sites five times. Note that if it was

decided that a maximum of five surveys could be conducted per site, to obtain the desired standard error 226 sites would need to be used for a removal design, requiring an expected 703 total number of surveys.

Discussion and general recommendations

We stress that attempting to survey as many sites as possible may not be the most efficient use of resources, and that surveying fewer sites more often may result in a more precise estimate of occupancy. For example, if $\psi \approx 0.4$ and $p \approx 0.3$ the asymptotic standard error for occupancy from a design where 200 sites are each surveyed twice is 0.11, but surveying 80 sites five times would provide an occupancy estimate with a standard error of 0.07. That is, by allocating resources more efficiently the standard error has been reduced by 36%. To achieve the same gain in precision with only two surveys of each site, then 500 sites would need to be surveyed, or the total survey effort would have to be increased by 250%!

Designing a study is as much an art as a science. Theoretical and simulation results provide useful guidance about the expected outcome of a study given certain assumptions, analytic techniques and designs. But these results must be tempered with common sense, expert knowledge of the system under study and, occasionally, lateral thinking. The results presented in Tables 1 and 3 indicate that, when detection probability is low, the optimal choice of K is very large. In most practical situations we doubt it would be likely that surveyors could conduct that many surveys within a timeframe short enough to ensure the closure assumption is satisfied, at enough sites to have meaningful results. However, we suggest these values can be used as a gold-standard to measure how much less efficient a design might be if an alternative value for K was to be used. We also point out that a relatively simple model for detection probability was assumed, and that if p varies in time (which we would suspect to often be the norm) then a greater number of repeat surveys is likely to be required. We therefore recommend that in general researchers should consider $K = 3$ as a minimum value when $p > 0.5$, and a greater number when p is smaller.

If detection probability is constant, then there is strong evidence that a removal design will be much more efficient than a standard design for estimating occupancy. The data yielded by a removal study, however, provide less flexibility for modelling, particularly for exploring potential sources of variation in detection probability. Hence we suspect that a removal design is likely to be less robust than a standard design in general. As a compromise between efficiency and robustness, it may be feasible to consider a hybrid design, with half of the total survey effort (say) used to conduct a standard design and the remaining effort used with a removal design. Such a hybrid design would provide greater flexibility for modelling the collected data, yet be more efficient than a full standard design.

While the above results are based upon the consideration of a relatively simple occupancy model, they are a useful starting point to give some indication of the likely number of repeated surveys per site that should be used for various values of ψ and p . As these optimal values do not depend upon the number of sites, they are valid even for a single site, which means if it is suspected that ψ and p vary across the population of interest (e.g. because of habitat or distance from the core of the species distribution), the population could be stratified and different values of K be used within each stratum. Furthermore, as a general strategy the results suggest that, when occupancy is low, more effort should be devoted to surveying more sites, while when occupancy is high more effort should be devoted to repeated surveys. We suggest this arises as p must be estimated in order to estimate occupancy accurately and information regarding p can only be gathered from occupied sites. Hence, when occupancy is low expending a lot of survey effort at relatively few sites may yield little information about p as most of the sites will be unoccupied. When occupancy is high, a lot of information about p can be garnered by surveying fewer sites more often. Field, Tyre & Possingham (2005) noted a similar result with a similar explanation.

Finally, we are firmly of the opinion that the best and most useful study designs arise through the close collaboration of biologists, statisticians and other relevant parties. This collaboration should begin at the embryonic stage of study design, when the study objective is being developed. The biologists have the expert knowledge of the species and system of interest, with an appreciation of the field techniques that could be employed, while the statisticians have the knowledge of appropriate analytic techniques and awareness of the data requirements for such methods. The quality of the inference about the scientific and/or management questions at the heart of the study lies heavily on the quality of the collected data. Therefore, only by careful consideration of all aspects of a proposed study design, and how they relate to the study's objective, can one hope to make reliable inferences about the species of interest.

Acknowledgements

We would like to thank Jim Nichols, Nigel Yoccoz and an anonymous referee for their comments on an earlier draft of this manuscript. This research was supported by a grant from the Royal Society of New Zealand Marsden Fund.

References

- Azuma, D.L., Baldwin, J.A. & Noon, B.R. (1990) *Estimating the Occupancy of Spotted Owl Habitat Areas by Sampling and Adjusting for Bias*. General Technical Report PSW-124. US Department of Agriculture Forest Service, Pacific Southwest Research Station, Berkeley, CA, USA.
- Bailey, L.L., Simons, T.R. & Pollock, K.H. (2004) Estimating site occupancy and species detection probability parameters for terrestrial salamanders. *Ecological Applications*, **14**, 692–702.

- Bradford, D.F., Neale, A.C., Nash, M.S., Sada, D.W. & Jaeger, J.R. (2003) Habitat patch occupancy by toads (*Bufo punctatus*) in a naturally fragmented desert landscape. *Ecology*, **84**, 1012–1023.
- Brown, J.H. (1995) *Macroecology*. University of Chicago Press, Chicago, IL.
- Cabeza, M., Araújo, M.B., Wilson, R.J., Thomas, C.D., Cowley, M.J.R. & Moilanen, A. (2004) Combining probabilities of occurrence with spatial reserve design. *Journal of Applied Ecology*, **41**, 252–262.
- Cochran, W.G. (1977) *Sampling Techniques*. Wiley, New York, NY.
- Engler, R., Guisan, A. & Rechsteiner, L. (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, **41**, 263–274.
- Field, S.A., Tyre, A.J. & Possingham, H.P. (2005) Optimizing allocation of monitoring effort under economic and observational constraints. *Journal of Wildlife Management*, **69**, 473–482.
- Geissler, P.H. & Fuller, M.R. (1987) Estimation of the proportion of area occupied by an animal species. *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 1986, pp. 533–538.
- Gibson, L.A., Wilson, B.A., Cahill, D.M. & Hill, J. (2004) Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *Journal of Applied Ecology*, **41**, 213–223.
- Gu, W. & Swihart, R.K. (2004) Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. *Biological Conservation*, **116**, 195–203.
- Hanski, I. (1994) A practical model of metapopulation dynamics. *Journal of Animal Ecology*, **63**, 151–162.
- Hanski, I. (1999) *Metapopulation Ecology*. Oxford University Press, Oxford, UK.
- Kawanishi, K. & Sunquist, M.E. (2004) Conservation status of tigers in a primary rainforest of peninsular Malaysia. *Biological Conservation*, **120**, 329–344.
- MacKenzie, D.I. (2005a) What are the issues with ‘presence/absence’ data for wildlife managers? *Journal of Wildlife Management*, in press.
- MacKenzie, D.I. (2005b) Was it there? Dealing with imperfect detection for species presence/absence data. *Australia and New Zealand Journal of Statistics*, **47**, 65–74.
- MacKenzie, D.I. (2006) Modeling the probability of resource use: the effect of, and dealing with, detecting a species imperfectly. *Journal of Wildlife Management*, in press.
- MacKenzie, D.I., Bailey, L.L. & Nichols, J.D. (2004) Investigating species co-occurrence patterns when species are detected imperfectly. *Journal of Animal Ecology*, **73**, 546–555.
- MacKenzie, D.I., Nichols, J.D., Hines, J.E., Knutson, M.G. & Franklin, A.D. (2003) Estimating site occupancy, colonization and local extinction when a species is detected imperfectly. *Ecology*, **84**, 2200–2207.
- MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A. & Langtimm, C.A. (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, **83**, 2248–2255.
- MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L. & Hines, J.E. (2005) *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier, San Diego, CA.
- MacKenzie, D.I., Royle, J.A., Brown, J.A. & Nichols, J.D. (2004) Occupancy estimation and modeling for rare and elusive populations. *Sampling Rare or Elusive Species* (ed. W.L. Thompson), pp. 149–172. Island Press, Washington, DC.
- Moilanen, A. (2002) Implications of empirical data quality for metapopulation model parameter estimation and application. *Oikos*, **96**, 516–530.
- Nichols, J.D. & Karanth, K.U. (2002) Statistical concepts: assessing spatial distributions. *Monitoring Tigers and Their Prey: A Manual for Researchers, Managers, and Conservationists in Tropical Asia* (eds K.U. Karanth & J.D. Nichols), pp. 29–38. Center for Wildlife Studies, Bangalore, India.
- Otis, D.L., Burnham, K.P., White, G.C. & Anderson, D.R. (1978) Statistical Inference from Capture Data on Closed Animal Populations. *Wildlife Monographs* **62**.
- Pollock, K.H., Nichols, J.D., Simons, T.R., Farnsworth, G.L., Bailey, L.L. & Sauer, J.R. (2002) Large scale wildlife monitoring studies: statistical methods for design and analysis. *Environmetrics*, **13**, 105–119.
- Reunanen, P., Nikula, A., Monkkonen, M., Hurme, E. & Nivala, V. (2002) Predicting occupancy for the Siberian flying squirrel in old-growth forest patches. *Ecological Applications*, **12**, 1188–1198.
- Royle, J.A. & Nichols, J.D. (2003) Estimating abundance from repeated presence absence data or point counts. *Ecology*, **84**, 777–790.
- Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Rafael, M.G., Wall, W.A. & Samson, F.B. (2002) *Predicting Species Occurrences: Issues of Accuracy and Scale*. Island Press, Washington, DC.
- Stauffer, H.B., Ralph, C.J. & Miller, S.L. (2002) Incorporating detection uncertainty into presence–absence surveys for marbled murrelet. *Predicting Species Occurrences: Issues of Accuracy and Scale* (eds J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Rafael, W.A. Wall & F.B. Samson), pp. 357–365. Island Press, Washington, DC.
- Stauffer, H.B., Ralph, C.J. & Miller, S.L. (2004) Ranking habitat for marbled murrelets: a new conservation approach for species with uncertain detection. *Ecological Applications*, **14**, 1374–1383.
- Trenham, P.C., Koenig, W.D., Mossman, M.J., Stark, S.L. & Jagger, L.A. (2003) Regional dynamics of wetland-breeding frogs and toads: turnover and synchrony. *Ecological Applications*, **13**, 1522–1532.
- Tyre, A.J., Tenhumberg, B., Field, S.A., Niejalke, D., Parris, K. & Possingham, H.P. (2003) Improving precision and reducing bias in biological surveys by estimating false negative error rates in presence–absence data. *Ecological Applications*, **13**, 1790–1801.
- Weber, D., Hinterman, U. & Zangger, A. (2004) Scale and trends in species richness: considerations for monitoring biological diversity for political purposes. *Global Ecological Biogeography*, **13**, 97–104.
- Wikle, C.K. (2003) Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology*, **84**, 1382–1394.
- Williams, B.K., Nichols, J.D. & Conroy, M.J. (2002) *Analysis and Management of Animal Populations*. Academic Press, San Diego, CA.
- Yoccoz, N.G., Nichols, J.D. & Boulinier, T. (2001) Monitoring of biological diversity in space and time. *Trends in Ecology and Evolution*, **16**, 446–453.
- Zielinski, W.J. & Stauffer, H.B. (1996) Monitoring *Martes* populations in California: survey design and power analysis. *Ecological Applications*, **6**, 1254–1267.

Received 26 September 2004; final copy received 24 July 2005
Editor: Phil Stephens