Ecology, 83(8), 2002, pp. 2248–2255
© 2002 by the Ecological Society of America

ESTIMATING SITE OCCUPANCY RATES WHEN DETECTION PROBABILITIES ARE LESS THAN ONE

DARRYL I. MACKENZIE,[1,5] JAMES D. NICHOLS,[2] GIDEON B. LACHMAN,[2,6] SAM DROEGE,[2] J. ANDREW ROYLE,[3] AND CATHERINE A. LANGTIMM[4]

[1]*Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695-8203 USA*
[2]*U.S. Geological Survey, Patuxent Wildlife Research Center, 11510 American Holly Drive, Laurel, Maryland 20708-4017 USA*
[3]*U.S. Fish and Wildlife Service, Patuxent Wildlife Research Center, 11510 American Holly Drive, Laurel, Maryland 20708-4017 USA*
[4]*U.S. Geological Survey, Florida Caribbean Science Center, Southeastern Amphibian Research and Monitoring Initiative, 7920 NW 71st Street, Gainesville, Florida 32653 USA*

Abstract. Nondetection of a species at a site does not imply that the species is absent unless the probability of detection is 1. We propose a model and likelihood-based method for estimating site occupancy rates when detection probabilities are <1. The model provides a flexible framework enabling covariate information to be included and allowing for missing observations. Via computer simulation, we found that the model provides good estimates of the occupancy rates, generally unbiased for moderate detection probabilities (>0.3). We estimated site occupancy rates for two anuran species at 32 wetland sites in Maryland, USA, from data collected during 2000 as part of an amphibian monitoring program, Frogwatch USA. Site occupancy rates were estimated as 0.49 for American toads (*Bufo americanus*), a 44% increase over the proportion of sites at which they were actually observed, and as 0.85 for spring peepers (*Pseudacris crucifer*), slightly above the observed proportion of 0.83.

Key words: *anurans; bootstrap;* Bufo americanus; *detection probability; maximum likelihood; metapopulation; monitoring; patch occupancy;* Pseudacris crucifer; *site occupancy.*

INTRODUCTION

We describe an approach to estimating the proportion of sites occupied by a species of interest. We envision a sampling method that involves multiple visits to sites during an appropriate season during which a species may be detectable. However, a species may go undetected at these sites even when present. Sites may represent discrete habitat patches in a metapopulation dynamics context or sampling units (e.g., quadrats) regularly visited as part of a large-scale monitoring program. The patterns of detection and nondetection over the multiple visits for each site permit estimation of detection probabilities and the parameter of interest, proportion of sites occupied.

Our motivation for considering this problem involves potential applications in (1) large-scale monitoring programs and (2) investigations of metapopulation dynamics. Monitoring programs for animal populations and communities have been established throughout the world in order to meet a variety of objectives. Most programs face two important sources of variation that must be incorporated into the design (e.g., see Thompson 1992, Lancia et al. 1994, Thompson et al. 1998, Yoccoz et al. 2001, Pollock et al. 2002).

The first source of variation is space. Many programs seek to provide inferences about areas that are too large to be completely surveyed. Thus, small areas must be selected for surveying, with the selection being carried out in a manner that permits inference to the entire area of interest (Thompson 1992, Yoccoz et al. 2001, Pollock et al. 2002).

The second source of variation important to monitoring program design is detectability. Few animals are so conspicuous that they are always detected at each survey. Instead, some sort of count statistic is obtained (e.g., number of animals seen, heard, trapped, or otherwise detected), and a method is devised to estimate the detection probability associated with the count statistic. Virtually all of the abundance estimators described in volumes such as Seber (1982) and Williams et al. (*in press*) can be viewed as count statistics divided by estimated detection probabilities. Not allowing for detectability and solely using the count statistic as an index to abundance is unwise. Changes in the count may be a product of random variations or changes in detectability, so it is impossible to make useful inference about the system under investigation.

The methods used to estimate detection probabilities of individual animals (and hence abundance) at each site are frequently expensive of time and effort. For

Manuscript received 29 May 2001; revised 11 October 2001; accepted 22 October 2001.

[5] Present address: Proteus Research and Consulting Ltd., P.O. Box 5193, Dunedin, New Zealand.
E-mail: darryl@proteus.co.nz

[6] Present address: International Association of Fish and Wildlife Agencies, 444 N. Capitol Street NW, Suite 544, Washington, D.C., 20001 USA.

this reason, these estimation methods are often used in detailed experiments or small-scale investigations, but are not as widely used in large-scale monitoring programs. The methods proposed here to estimate the proportion of sites (or more generally, the proportion of sampled area) occupied by a species can be implemented more easily and less expensively than the methods used for abundance estimation. For this reason, our proposed method should be attractive as a basis for large-scale monitoring programs, assuming that the proportion of sites or area occupied is an adequate state variable with respect to program objectives.

The second motivation for considering this estimation problem involves the importance of patch occupancy data to the study of metapopulation dynamics. The proportion of patches occupied is viewed as a state variable in various metapopulation models (e.g., Levins 1969, 1970, Lande 1987, 1988, Hanski 1992, 1994, 1997). So-called ''incidence functions'' (e.g., see Diamond 1975, Hanski 1992) depict the probability of occurrence of a species in a patch, expressed as a function of patch characteristics such as area. Under the assumption of a stationary Markov process, incidence function data are sometimes used to estimate patch extinction and colonization probabilities (e.g., Hanski 1992, 1994, 1997, Moilanen 1999). Given the relevance of patch occupancy data to metapopulation investigations and models, it seems important to estimate patch occupancy probabilities properly. For most animal sampling situations, detection of a species is indeed indicative of ''presence,'' but nondetection of the species is not equivalent to absence. Thus, we expect most incidence function estimates of the proportion of patches occupied to be negatively biased to some unknown degree because species can go undetected when present.

In this paper, we first present general sampling methods that permit estimation of the probability of site occupancy when detection probabilities are <1 and may vary as functions of site characteristics, time, or environmental variables. We then present a statistical model for site occupancy data and describe maximum likelihood estimation under this model. We illustrate use of the estimation approach with empirical data on site occupancy by two anuran species at 32 wetland sites in Maryland collected during 2000. Finally we discuss extending this statistical framework to address other issues such as colony extinction/colonization, species co-occurrence, and allowing for heterogeneous detection and occupancy probabilities

Methods

Notation

We use the following notation throughout this article: ψ_i, probability that a species is present at site i; p_{it}, probability that a species will be detected at site i at time t, given presence; N, total number of surveyed sites; T, number of distinct sampling occasions; n_t number of sites where the species was detected at time t; $n.$, total number of sites at which the species was detected at least once.

Our use of p, to signify detection probabilities, differs from its customary use in the metapopulation literature, where it is used to denote the probability of species presence (our ψ). However, our notation is consistent with the mark–recapture literature which provides the foundation of our approach.

Basic sampling situation

Here we consider situations in which surveys of species at N specific sites are performed at T distinct occasions in time. Sites are occupied by the species of interest for the duration of the survey period, with no new sites becoming occupied after surveying has begun, and no sites abandoned before the cessation of surveying (i.e., the sites are ''closed'' to changes in occupancy). At each sampling occasion, investigators use sampling methods designed to detect the species of interest. Species are never falsely detected at a site when absent, and a species may or may not be detected at a site when present. Detection of the species at a site is also assumed to be independent of detecting the species at all other sites. The resulting data for each site can be recorded as a vector of 1's and 0's denoting detection and nondetection, respectively, for the occasions on which the site was sampled. The set of such detection histories is used to estimate the quantity of interest, the proportion of sites occupied by the species.

General likelihood

We propose a method that parallels a closed-population, mark–recapture model, with an additional parameter (ψ) that represents the probability of species presence. In closed-population models, the focus is to estimate the number of individuals never encountered by using information garnered from those individuals encountered at least once (e.g., see Otis et al. 1978, Williams et al., *in press*). In our application, sites are analogous to individuals except that we observe the number of sites with the history comprising T 0's (sites at which the species is never detected over the T sampling occasions); hence, the total population size of sites is known, but the focus is to estimate the fraction of those sites that the species actually occupies. One could recast this problem into a more conventional closed mark–recapture framework by only considering those sites where the species was detected at least once. Use of such data with closed-population, capture–recapture models (e.g., Otis et al. 1978) would yield estimates of population size that correspond to the number of sites where the species is present. However, the following method enables additional modeling of ψ to be investigated (such as including covariate information).

A likelihood can be constructed using a series of

probabilistic arguments similar to those used in mark–recapture modeling (Lebreton et al. 1992). For sites where the species was detected on at least one sampling occasion, the species must be present and was either detected or not detected at each sampling occasion. For example, the likelihood for site i with history 01010 would be

$$\psi_i(1 - p_{i1})p_{i2}(1 - p_{i3})p_{i4}(1 - p_{i5}).$$

However, nondetection of the species does not imply absence. Either the species was present and was not detected after T samples, or the species was not present. For site k with history 00000, the likelihood is

$$\psi_k \prod_{t=1}^{5} (1 - p_{kt}) + (1 - \psi_k).$$

Assuming independence of the sites, the product of all terms (one for each site) constructed in this manner creates the model likelihood for the observed set of data, which can be maximized to obtain maximum likelihood estimates of the parameters.

Note that, at this stage, presence and detection probabilities have been defined as site specific. In practice, such a model could not be fit to the data because the likelihood contains too many parameters: the model likelihood is over-parameterized. However, the model is presented in these general terms because, in some cases, the probabilities may be modeled as a function of site-specific covariates, to which we shall return.

When presence and detection probabilities are constant across monitoring sites, the combined model likelihood can be written as

$$L(\psi, \mathbf{p}) = \left[\psi^{n.} \prod_{t=1}^{T} p_t^{n_t} (1 - p_t)^{n. - n_t}\right] \times \left[\psi \prod_{t=1}^{T} (1 - p_t) + (1 - \psi)\right]^{N - n.}. \quad (1)$$

Using the likelihood in this form, our model could be implemented with relative ease via spreadsheet software with built-in function maximization routines, because only the summary statistics ($n_1, \ldots, n_T, n.$) and N are required. Detection probabilities could be time specific, or reduced forms of the model could be investigated by constraining p to be constant across time or a function of environmental covariates.

We suggest that the standard error of ψ be estimated using a nonparametric bootstrap method (Buckland and Garthwaite 1991), rather than the asymptotic (large-sample) estimate involving the second partial derivatives of the model likelihood (Lebreton et al. 1992). The asymptotic estimate represents a lower bound on the value of the standard error, and may be too small when sample sizes are small. A random bootstrap sample of N sites is taken (with replacement) from the N monitored sites. The histories of the sites in the bootstrap sample are used to obtain a bootstrap estimate of ψ. The bootstrap procedure is repeated a large number of times, and the estimated standard error is the sample standard deviation of the bootstrap estimates (Manly 1997).

Extensions to the model

Covariates.—It would be reasonable to expect that ψ may be some function of site characteristics such as habitat type or patch size. Similarly, p may also vary with certain measurable variables such as weather conditions. This covariate information ($\mathbf{X}$) can be easily introduced to the model using a logistic model (Eq. 2) for ψ and/or p (denote the parameter of interest as θ and the vector of model parameters as $\mathbf{B}$:

$$\boldsymbol{\theta} = \frac{\exp(\mathbf{XB})}{1 + \exp(\mathbf{XB})}. \quad (2)$$

Because ψ does not change over time during the sampling (the population is closed), appropriate covariates would be time constant and site specific, whereas covariates for detection probabilities could be time varying and site specific (such as air or water temperature).

This is in contrast to mark–recapture models in which time-varying individual covariates cannot be used. In mark–recapture, a time-varying individual covariate can only be measured on those occasions when the individual is captured; the covariate value is unknown otherwise. Here, time-varying, site-specific covariates can be collected and used regardless of whether the species is detected. It would not be possible, however, to use covariates that change over time and cannot be measured independent of the detection process.

If ψ is modeled as a function of covariates, the average species presence probability is

$$\bar{\hat{\psi}} = \frac{\sum_{i=1}^{N} \hat{\psi}_i}{N}. \quad (3)$$

Missing observations.—In some circumstances, it may not be possible to survey all sites at all sampling occasions. Sites may not be surveyed for a number of reasons, from logistic difficulties in getting field personnel to all sites, to the technician's vehicle breaking down en route. These sampling inconsistencies can be easily accommodated using the proposed model likelihood.

If sampling does not take place at site i at time t, then that occasion contributes no information to the model likelihood for that site. For example, consider the history 10_11, where no sampling occurred at time 3. The likelihood for this site would be:

$$\psi p_1(1 - p_2)p_4 p_5.$$

Missing observations can only be accounted for in this manner when the model likelihood is evaluated separately for each site, rather than using the combined form of Eq. 1.

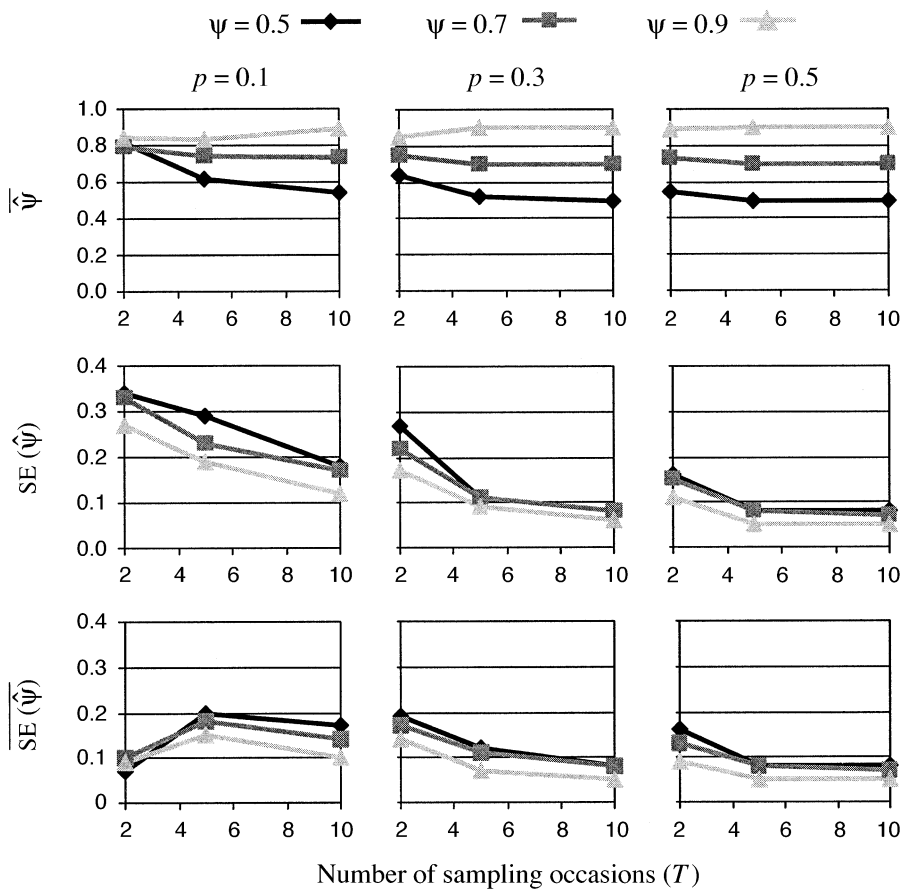


FIG. 1. Results of the 500 simulated sets of data for $N = 40$, with no missing values. Indicated are the average value of $\hat{\psi}$, $\bar{\hat{\psi}}$; the replication-based estimate of the true standard error of $\hat{\psi}$, SE($\hat{\psi}$); and the average estimate of the standard error obtained from 200 nonparametric bootstrap samples, $\overline{\text{SE}}(\hat{\psi})$, for various levels of T, p, and ψ.

SIMULATION STUDY

Simulation methods

A simulation study was undertaken to evaluate the proposed method for estimating ψ. Data were generated for situations in which all sites had the same probability of species presence, and the detection probability was constant across time and sites, $\psi(\cdot)p(\cdot)$. The effects of five factors were investigated: (1) $N = 20$, 40, or 60; (2) $\psi = 0.5$, 0.7, or 0.9; (3) $p = 0.1$, 0.3, or 0.5; (4) $T = 2$, 5, or 10; (5) probability of a missing observation = 0.0, 0.1, or 0.2.

For each of the 243 scenarios, 500 sets of data were simulated. For each site, a uniformly distributed, pseudo-random number between 0 and 1 was generated (y), and if $y \leq \psi$ then the site was occupied. Further pseudo-random numbers were generated and similarly compared to p to determine whether the species was detected at each time period, with additional random numbers being used to establish missing observations. The $\psi(\cdot)p(\cdot)$ model was applied to each set of simulated data. The resulting estimate of ψ was recorded and the nonparametric bootstrap estimate of the standard error was also obtained using 200 bootstrap samples.

Simulation results

Fig. 1 presents the simulation results for scenarios where $N = 40$ with no missing values only, but these are representative of the results in general. The full simulation results are included in the Appendix.

Generally, this method provides reasonable estimates of the proportion of sites occupied. When detection probability is 0.3 or greater, the estimates of ψ are reasonably unbiased in all scenarios considered for $T \geq 5$. When $T = 2$, only when detection probability is at least 0.5 do the estimates of ψ appear to be reasonable. For low detection probabilities, however, ψ tends to be overestimated when the true value is 0.5 or 0.7, but underestimated when ψ equals 0.9. A closer examination of the results reveals that, in some situations in which detection probability is low, $\hat{\psi}$ tends to 1.

In most cases, the nonparametric bootstrap provides a good estimate of the standard error for $\hat{\psi}$, the exception being for situations with low detection probabil-

ities. Again, this is caused by ψ estimates close to 1; in such situations, the bootstrap estimate of the standard error is very small, which overstates the precision of $\hat{\psi}$.

In general, increasing the number of sampling occasions improves both the accuracy and precision of $\hat{\psi}$, although in some instances there is little gain in using 10 occasions rather than five. If only two occasions are used, however, accuracy tends to be poor unless detection probabilities are high, and even then the standard error of $\hat{\psi}$ is approximately double that of using five sampling occasions.

Similarly, increasing the number of sites sampled, N, also improves both the accuracy and precision of $\hat{\psi}$.

Not presented here are the simulation results for scenarios with missing observations. The proposed method appears to be robust to missing data, with the only noticeable effect being (unsurprisingly) a loss of precision. In this study, on average, the standard error of $\hat{\psi}$ increased by 5% with 10% missing observations, and by 11% with 20% missing observations. The bootstrap standard error estimates also increased by a similar amount, accounting well for the loss of information.

Field Study of Anurans at Maryland Wetlands

Field methods and data collection

We illustrate our method by considering monitoring data collected on American toads (*Bufo americanus*) and spring peepers (*Pseudacris crucifer*) at 32 wetland sites located in the Piedmont and Upper Coastal Plain physiographic provinces surrounding Washington, D.C., and Baltimore, Maryland, USA. Volunteers enrolled in the National Wildlife Federation/U.S. Geological Survey's amphibian monitoring program, FrogwatchUSA, visited monitoring sites between 19 February 2000 and 12 October 2000. Sites were chosen nonrandomly by volunteers and were monitored at their convenience. Observers collected information on the species of frogs and toads heard calling during a 3-min counting period taken sometime after sundown. Each species of calling frog and toad was assigned a three-level calling index, which, for this study, was truncated to reflect either detection (1) or nondetection (0).

The data set was reduced by considering only the portion of data for each species between the dates of first and last detection exclusive. Truncating the data in this manner ensures that species were available to be detected throughout that portion of the monitoring period, thus satisfying our closure assumption. Including the dates of first and last detection in the analysis would bias parameter estimates because the data set was defined using these points; hence, they were excluded.

Three sites were removed after the truncation because they were never monitored during the redefined period. Fewer than eight of the 29 sites were monitored on any given day and the number of visits per site varied tremendously, with a very large number of missing observations (~90%). Note that in the context of this sampling, the entire sampling period included the interval between the date at which the first wetland was sampled and the date at which all sampling ended. A missing observation was thus any date during this interval on which a wetland was not sampled. Each time a site was visited, air temperature was recorded. Sites were defined as being either a distinct body of water (pond, lake) or other habitat (swamp, marsh, wet meadow). These variables were considered as potential covariates for detection and presence probabilities, respectively. The data used in this analysis have been included in the Supplement.

TABLE 1. Relative difference in AIC (ΔAIC), AIC model weights (w_i), overall estimate of the fraction of sites occupied by each species ($\bar{\hat{\psi}}$), and associated standard error (SE($\bar{\hat{\psi}}$)).

Model, by species	ΔAIC	w_i	$\bar{\hat{\psi}}$	SE($\bar{\hat{\psi}}$)
American toad				
ψ(Habitat) p(Temperature)	0.00	0.36	0.50	0.13
$\psi(\cdot)$ p(Temperature)	0.42	0.24	0.49	0.14
ψ(Habitat) $p(\cdot)$	0.49	0.22	0.49	0.12
$\psi(\cdot)$ $p(\cdot)$	0.70	0.18	0.49	0.13
Spring peeper				
ψ(Habitat) p(Temperature)	0.00	0.85	0.84	0.07
$\psi(\cdot)$ p(Temperature)	1.72	0.15	0.85	0.07
ψ(Habitat) $p(\cdot)$	40.49	0.00	0.84	0.07
$\psi(\cdot)$ $p(\cdot)$	42.18	0.00	0.85	0.07

Results of field study

American toad.—Daily records for the 29 sites, monitored between 9 March 2000 and 30 May 2000, were included for analysis. Sites were visited 8.9 times on average (minimum = 2, maximum = 58 times), with American toads being detected at least once at 10 locations (0.34). Three models with covariates and one without were fit to the data (Table 1) and ranked according to AIC (Burnham and Anderson 1998). The four models considered have virtually identical weight, suggesting that all models provide a similar description of the data, despite the different structural forms. Therefore we cannot make any conclusive statement regarding the importance of the covariates, but there is some suggestion that detection probabilities may increase with increasing temperature and occupancy rates may be lower for habitats consisting of a distinct body of water. However, all models provide very similar estimates of the overall occupancy rate (~0.49), which is 44% larger than the proportion of sites where toads were detected at least once. The standard error for the estimate is reasonably large and corresponds to a coefficient of variation of 27%.

Spring peeper.—Daily records for the 29 sites, monitored between 27 February 2000 and 30 May 2000, were included for analysis. Sites were visited, on av-

erage, 9.6 times (minimum = 2, maximum = 66 visits), with spring peepers being detected at least once at 24 locations (0.83). The same models as those for the American toad were fit to the spring peeper data and the results are also displayed in Table 1. Here the two $p(\cdot)$ models have virtually zero weight, indicating that the p(Temperature) models provide a much better description of the data. We suspect that this effect is due, partially, to a tapering off of the calling season as spring progresses into summer. The ψ(Habitat) p(Temperature) model clearly has greatest weight and suggests that estimated occupancy rates are lower for distinct bodies of water (0.77) than for other habitat types (1.00). This is not unexpected, given spring peepers were actually detected at all sites of the latter type. Regardless of how the models ranked, however, all models provide a similar estimate of the overall occupancy rate that is only marginally greater than the number of sites where spring peepers were detected at least once. This suggests that detection probabilities were large enough that spring peepers probably would be detected during the monitoring if present.

Discussion

The method proposed here to estimate site occupancy rate uses a simple probabilistic argument to allow for species detection probabilities of <1. As shown, it provides a flexible modeling framework for incorporating both covariate information and missing observations. It also lays the groundwork for some potentially exciting extensions that would enable important ecological questions to be addressed.

From the full simulation results for scenarios with low detection probabilities, it is very easy to identify circumstances in which one should doubt the estimates of ψ. We advise caution if an estimate of ψ very close to 1 is obtained when detection probabilities are low (<0.15), particularly when the number of sampling occasions is also small (<7). In such circumstances, the level of information collected on species presence/absence is small, so it is difficult for the model to distinguish between a site where the species is genuinely absent and a site where the species has merely not been detected.

Our simulation results may also provide some guidance on the number of visits to each site required in order to obtain reasonable estimates of occupancy rate. If one wishes to visit a site only twice, then it appears that the true occupancy rate needs to be >0.7 and detection probability (at each visit) should be >0.3. Even then, however, precision of the estimate may be low. Increasing the number of visits per site improves the precision of the estimated occupancy rate, and the resulting increase in information improves the accuracy of the estimate when detection probabilities are low. We stress that whenever a survey (of any type) is being designed, some thought should be given to the likely results and method of analysis, because these considerations can provide valuable insight on the level of sampling effort required to achieve ''good'' results.

Logistical considerations of multiple visits will probably result in some hesitancy to use this approach, but we suggest that the expenditure of extra effort to obtain unbiased estimates of parameters of interest generally will be preferable to the expenditure of less effort to obtain biased estimates. If travel time to sites is substantial, then multiple searches or samples may be conducted by multiple observers, or even by a single observer, at a single trip to a site, e.g., conduct two or more 3-min amphibian calling surveys in a single night at the same pond. If large numbers of patches must be surveyed, then it may be reasonable to conduct multiple visits at a subset of sites for the purpose of estimating detection probability, and perhaps associated covariate relationships. Then this information on detection probability, perhaps modeled as a function of site-specific covariates, could be applied to sites visited only once. Issues about optimal design require additional work, but it is clear that a great deal of flexibility is possible in approaches to sampling.

Site occupancy may well change over years or between seasons as populations change; new colonies could be formed or colonies could become locally extinct. When sites are surveyed on more than one occasion between these periods of change, for multiple periods, the approach described here could be combined with the robust design mark–recapture approach (Pollock et al. 1990). For example, suppose that the anuran sampling described in our examples is continued in the future, such that the same wetland sites are surveyed multiple times each summer, for multiple years. During the periods when sites are closed to changes in occupancy, our approach could be used to estimate the occupancy rate as in our example. The change in occupancy rates over years could then be modeled as functions of site colonization and extinction rates, analogous with the birth and death rates in an open-population mark–recapture study. Such Markov models of patch occupancy dynamics will permit time-specific estimation and modeling of patch extinction and colonization rates that do not require the assumptions of $p = 1$ or process stationarity invoked in previous modeling efforts (e.g., Erwin et al. [1998] required $p = 1$; Hanski [1992, 1994] and Clark and Rosenzweig [1994] required both assumptions).

Often monitoring programs collect information on the presence/absence of multiple species at the same sites. An important biological question is whether species co-occur independently. Does the presence/absence of species A depend upon the occupancy state of species B? Our method of modeling species presence could be extended in this direction, enabling such important ecological questions to be addressed. The model could be parameterized in terms of ψ_{AB} (in addition to ψ_A and ψ_B): the probability that both species A and species B are present at a site. However, the number

of parameters in the model would increase exponentially with the number of species, so reasonably good data sets might be required. For example, four additional parameters would be required to model co-occurrences between species A, B, and C (ψ_{AB}, ψ_{AC}, ψ_{BC}, ψ_{ABC}), but if six species were being modeled, 57 extra parameters would need to be estimated.

Not addressed are situations in which presence and detection probabilities are heterogeneous, varying across sites. Some forms of heterogeneity may be accounted for with covariate information such as site characteristics or environmental conditions at the time of sampling. On other occasions, however, the source of heterogeneity may be unknown. We foresee that combining our method with the mixture model approach to closed-population, mark–recapture models of Pledger (2000) would be one solution, which enables the problem to be contained within a likelihood framework. It may also be possible to combine our method with other closed-population, mark–recapture methods such as the jackknife (Burnham and Overton 1978) or coverage estimators (Chao et al. 1992). For different sampling frameworks, where monitoring is performed on a continuous or incidental basis rather than at discrete sampling occasions, combining our methods with the Poisson family of models (Boyce et al. 2001, MacKenzie and Boyce 2001) may also be feasible, particularly for multiple years of data.

The three extensions to the proposed methods are currently the focus of ongoing research on this general topic of estimating site occupancy rates.

Software to perform the above modeling has been included in the Supplement.

Acknowledgments

We would like to thank Christophe Barbraud, Mike Conroy, Ullas Karanth, Bill Kendall, Ken Pollock, John Sauer, and Rob Swihart for useful discussions of this general estimation problem and members of the USGS Southeastern Amphibian Research and Monitoring Initiative for discussions on amphibian monitoring. Atte Moilanen provided a constructive review of the manuscript, as did a second anonymous reviewer. Our thanks also go to the Frogwatch USA volunteers directed by Sue Muller at Howard County Department of Parks and Recreation for collection of the data.

Literature Cited

Boyce, M. S., D. I. MacKenzie, B. F. J. Manly, M. A. Haroldson, and D. Moody. 2001. Negative binomial models for abundance estimation of multiple closed populations. Journal of Wildlife Management **65**:498–509.

Buckland, S. T., and P. H. Garthwaite. 1991. Quantifying precision of mark–recapture estimates using the bootstrap and related methods. Biometrics **47**:255–268.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference—a practical information-theoretic approach. Springer-Verlag, New York, New York, USA.

Burnham, K. P., and W. S. Overton. 1978. Estimation of the size of a closed population when capture probabilities vary among animals. Biometrika **65**:625–633.

Chao, A., S.-M. Lee, and S.-L. Jeng. 1992. Estimating population size for capture–recapture data when capture probabilities vary by time and individual animal. Biometrics **48**:201–216.

Clark, C. W., and M. L. Rosenzweig. 1994. Extinction and colonization processes: parameter estimates from sporadic surveys. American Naturalist **143**:583–596.

Diamond, J. M. 1975. Assembly of species communities. Pages 342–444 *in* M. L. Cody and J. M. Diamond, editors. Ecology and evolution of communities. Harvard University Press, Cambridge, Massachusetts, USA.

Erwin, R. M., J. D. Nichols, T. B. Eyler, D. B. Stotts, and B. R. Truitt. 1998. Modeling colony site dynamics: a case study of Gull-billed Terns (*Sterna nilotica*) in coastal Virginia. Auk **115**:970–978.

Hanski, I. 1992. Inferences from ecological incidence functions. American Naturalist **139**:657–662.

Hanski, I. 1994. A practical model of metapopulation dynamics. Journal of Animal Ecology **63**:151–162.

Hanski, I. 1997. Metapopulation dynamics: from concepts and observations to predictive models. Pages 69–91 *in* I. A. Hanski and M. E. Gilpin, editors. Metapopulation biology: ecology, genetics, and evolution. Academic Press, New York, New York, USA.

Lancia, R. A., J. D. Nichols, and K. H. Pollock. 1994. Estimating the number of animals in wildlife populations. Pages 215–253 *in* T. Bookhout, editor. Research and management techniques for wildlife and habitats. The Wildlife Society, Bethesda, Maryland, USA.

Lande, R. 1987. Extinction thresholds in demographic models of territorial populations. American Naturalist **130**:624–635.

Lande, R. 1988. Demographic models of the northern spotted owl (*Strix occidentalis caurina*). Oecologia **75**:601–607.

Lebreton, J. D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals. A unified approach with case studies. Ecological Monographs **62**:67–118.

Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America **15**:237–240.

Levins, R. 1970. Extinction. Pages 77–107 *in* M. Gerstenhaber, editor. Some mathematical questions in biology. Volume II. American Mathematical Society, Providence, Rhode Island, USA.

MacKenzie, D. I., and M. S. Boyce. 2001. Estimating closed population size using negative binomial models. Western Black Bear Workshop **7**:21–23.

Manly, B. F. J. 1997. Randomization, bootstrap and Monte Carlo methods in biology. Second edition. Chapman and Hall, London, UK.

Moilanen, A. 1999. Patch occupancy models of metapopulation dynamics: efficient parameter estimation using implicit statistical inference. Ecology **80**:1031–1043.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs **62**.

Pledger, S. 2000. Unified maximum likelihood estimates for closed capture–recapture models using mixtures. Biometrics **56**:434–442.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture–recapture experiments. Wildlife Monographs **107**.

Pollock, K. H., J. D. Nichols, T. R. Simons, G. L. Farnsworth, L. L. Bailey, and J. R. Sauer. 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis. Environmetrics **13**:1–15.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. MacMillan Press, New York, New York, USA.

Thompson, S. K. 1992. Sampling. John Wiley, New York, New York, USA.
Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic Press, San Diego, California, USA.
Williams, B. K., J. D. Nichols, and M. J. Conroy. *In press.* Analysis and management of animal populations. Academic Press, San Diego, California, USA.
Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time; concepts, methods and designs. Trends in Ecology and Evolution **16**: 446–453.

APPENDIX

Full results of the simulation study are available in ESA's Electronic Data Archive: *Ecological Archives* E083-041-A1.

SUPPLEMENT

Software, source code, and the sample data sets are available in ESA's Electronic Data Archive: *Ecological Archives* E083-041-S1.