

OCCUPANCY ESTIMATION AND MODELING WITH MULTIPLE STATES AND STATE UNCERTAINTY

JAMES D. NICHOLS,^{1,4} JAMES E. HINES,¹ DARRYL I. MACKENZIE,² MARK E. SEAMANS,³ AND R. J. GUTIÉRREZ³

¹United States Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA

²Proteus Wildlife Research Consultants, P.O. Box 5193, Dunedin, New Zealand

³Department of Fisheries, Wildlife, and Conservation Biology, University of Minnesota, St. Paul, Minnesota 55108 USA

Abstract. The distribution of a species over space is of central interest in ecology, but species occurrence does not provide all of the information needed to characterize either the well-being of a population or the suitability of occupied habitat. Recent methodological development has focused on drawing inferences about species occurrence in the face of imperfect detection. Here we extend those methods by characterizing occupied locations by some additional state variable (e.g., as producing young or not). Our modeling approach deals with both detection probabilities <1 and uncertainty in state classification. We then use the approach with occupancy and reproductive rate data from California Spotted Owls (*Strix occidentalis occidentalis*) collected in the central Sierra Nevada during the breeding season of 2004 to illustrate the utility of the modeling approach. Estimates of owl reproductive rate were larger than naïve estimates, indicating the importance of appropriately accounting for uncertainty in detection and state classification.

Key words: California Spotted Owl; misclassification; multiple states; occupancy models; reproductive state; *Strix occidentalis occidentalis*; uncertainty.

INTRODUCTION

The proportion of sites or area occupied by a species is an important topic relevant to a variety of ecological and evolutionary questions involving such topics as species range dynamics, metapopulation dynamics, species–habitat relationships, and interspecific relationships. Because of its importance, recent efforts have focused on development of methods for estimating and modeling occupancy and associated dynamics (see MacKenzie et al. 2006 for summary). These methods are based on multiple visits to sites, where species detection and nondetection are noted, and the proportion of sites that is occupied by the species is estimated in the face of imperfect detection. Here, we extend these methods to a large class of questions that seek to classify sites by different categories of occupancy. For example, one of the most common uses of this method will be to classify occupied sites by whether young animals are produced at the sites. This distinction is relevant to ideas about source and sink sites (Pulliam 1988), and, more generally, about contributions of local sites to metapopulation dynamics (Runge et al. 2006). In addition, this method has relevance to evolutionary questions, because it can help elucidate site quality through its focus on fitness components such as reproduction (Fretwell 1972, Franklin et al. 2000). The concept of categories of occupancy is very general and can include animal behaviors at sites

(e.g., sites used for resting vs. foraging), sites with and without diseased or parasitized animals, and so on.

The concepts underlying our modeling are general and readily extend to variable numbers and types of states. However, we illustrate the approach with a specific problem involving estimation of reproductive rate of California Spotted Owls (*Strix occidentalis occidentalis*) in the face of state classification uncertainty and variable sampling protocols. Because potential owl territories are inhabited by a single breeding pair, the problem of estimating the proportion of sites at which young are produced can be translated into the problem of estimating the proportion of pairs that produce young. Franklin et al. (2004) concluded that this problem was sufficiently substantial to preclude detailed meta-analysis of Spotted Owl reproductive rates in a recent synthetic treatment of demographic data from five large study sites. Subsequent analyses of these data by Seamans (2005) highlighted the dependence of reproductive rate estimates for this species on area-specific sampling protocols and led to the recommendation that investigators (1) better standardize protocols and (2) develop an estimation approach that can estimate reproductive rate in the face of variable numbers and kinds of visits to owl territories. Here, we first develop a general three-state model for use with detection–nondetection data that include state classification of sites with detections. We then apply the model to reproductive rate data for the California Spotted Owl.

GENERAL SAMPLING SITUATION

We will deal with situations in which multiple visits are made to each of a set of sites or locations over a

Manuscript received 31 August 2006; revised 18 December 2006; accepted 8 January 2007; final version received 8 February 2007. Corresponding Editor: S. T. Buckland.

⁴ E-mail: jnichols@usgs.gov

fairly short time frame within a season. The true states of the sites are unoccupied (state = 0), occupied with no production of young (state = 1), and occupied with successful reproduction (state = 2). At each visit to a site, the result of the site visit is classified as either no detection of the species (denote as 0), detection of the species in state 2, with no uncertainty associated with state assignment (denote as 2), or detection of the species with uncertain reproductive state assignment (denote as 1). It is assumed that certain kinds of observations (e.g., detections of young in nest or fledged young in trees within the territory) lead to unambiguous classification of the site as belonging to state 2. In contrast, failure to make such observations (1) does not necessarily indicate the absence of successful reproduction, but admits the possibility that the site is in state 1 (species present but no production of young) or state 2. To summarize, observations at site visits can be scored as 0 (nondetection; possible true states are 0, 1, 2), 1 (detection with no evidence of production of young; possible true states are 1, 2), or 2 (detection with evidence of successful reproduction; true state is 2).

Some number, S , of sites is visited K times during a season yielding detection history data, a vector of observation data (0, 1, or 2) for each visited site. For example, detection history $h_i = 01021$ for site i indicates nondetection at sampling periods (visits) 1 and 3, detection with no evidence of young at periods 2 and 5, and detection with evidence of successful reproduction at period 4. We assume that the sites are closed to changes in occupancy during the season (each site is either occupied or not). Evidence of production of young is not necessarily assumed to be potentially available at each sample period (i.e., it may not be possible to observe young early in the breeding season), although such evidence must be available for at least two periods.

BASIC MODEL

We model the process that gives rise to the detection history data using the following parameters:

ψ_i^1 = probability that site i is occupied regardless of reproductive state, $\Pr(\text{true state} = 1 \text{ or } 2)$;

ψ_i^2 = probability that young occurred, given that the site is occupied, $\Pr(\text{true state} = 2 | \text{true state} = 1 \text{ or } 2)$;

p_{it}^1 = probability that occupancy is detected for site i , period t , given that true state = 1, $\Pr(\text{detection} | \text{true state} = 1)$;

p_{it}^2 = probability that occupancy is detected for site i , period t , given that true state = 2, $\Pr(\text{detection} | \text{true state} = 2)$;

δ_{it} = probability that evidence of successful reproduction is found, given detection of occupancy at site i , period t , with successful reproduction, $\Pr(\text{classified state } 2 | \text{detection, true state} = 2)$.

Using these parameters, the unconditional probability that a site is occupied by successful breeders is given by the product: $\psi_i^{1*2} = \psi_i^1 \psi_i^2$. Note that if state assignment

is not characterized by uncertainty, then we simply set $\delta_{it} = 1$ and proceed with the subsequent modeling.

Consider the modeling of detection history $h_i = 1021$

$$\begin{aligned} \Pr(h_i = 1021) \\ = \psi_i^1 \psi_i^2 p_{i1}^2 (1 - \delta_{i1}) (1 - p_{i2}^2) p_{i3}^2 \delta_{i3} p_{i4}^2 (1 - \delta_{i4}). \end{aligned}$$

The site was known to be occupied and it was known that successful reproduction occurred (observed at period 3). The probability associated with this event is $\psi_i^1 \psi_i^2$. The species was detected at the site at period 1 (probability p_{i1}^2), but no evidence of successful reproduction was detected at that time (probability $1 - \delta_{i1}$). The species was not detected at period 2 ($1 - p_{i2}^2$). At period 3, the species was detected (p_{i3}^2) and evidence of successful reproduction found (δ_{i3}). Finally, at period 4, the species was detected (p_{i4}^2), but no evidence of successful reproduction was found ($1 - \delta_{i4}$). This is the simplest form of detection history, because it includes a 2, so there is no uncertainty about true occupancy status or reproductive state.

Next, consider modeling detection history $h_i = 0101$

$$\begin{aligned} \Pr(h_i = 0101) \\ = \psi_i^1 [(1 - \psi_i^2) (1 - p_{i1}^1) p_{i2}^1 (1 - p_{i3}^1) p_{i4}^1 \\ + \psi_i^2 (1 - p_{i1}^2) p_{i2}^2 (1 - \delta_{i2}) (1 - p_{i3}^2) p_{i4}^2 (1 - \delta_{i4})]. \end{aligned}$$

The site was known to be occupied (ψ_i^1), but evidence of successful reproduction was never observed, so uncertainty exists about the true reproductive state. Because of this uncertainty, the detection history model must account for the fact that the true reproductive state could have been either 1 (first additive term within brackets) or 2 (second additive term). If the pair at the site was in true reproductive state 1 ($1 - \psi_i^2$), then occupancy was detected at periods 2 (p_{i2}^1) and 4 (p_{i4}^1), but not at periods 1 ($1 - p_{i1}^1$) or 3 ($1 - p_{i3}^1$). There are no δ_{it} parameters in this first term, because the true reproductive state in this part of the expression is 1 and, by assumption, is not prone to misclassification. However, if the true reproductive state was 2 (ψ_i^2), then detection at periods 2 and 4 is accompanied by misclassification [$p_{i2}^2 (1 - \delta_{i2})$, $p_{i4}^2 (1 - \delta_{i4})$]. Nondetection at periods 1 and 3 is modeled using the appropriate detection parameters [$(1 - p_{i1}^2)$, $(1 - p_{i3}^2)$].

Finally, consider the detection history with the greatest degree of uncertainty, $h_i = 0000$:

$$\begin{aligned} \Pr(h_i = 0000) \\ = (1 - \psi_i^1) + \psi_i^1 (1 - \psi_i^2) \prod_{t=1}^4 (1 - p_{it}^1) \\ + \psi_i^1 \psi_i^2 \prod_{t=1}^4 (1 - p_{it}^2). \end{aligned}$$

The first of the additive terms corresponds to the possibility that the site was unoccupied ($1 - \psi_i^1$). The second of the additive terms corresponds to the possibility that the site was occupied but that successful

reproduction did not occur at the site [$\psi_i^1(1 - \psi_i^2)$]. The true reproductive state was thus 1 for the second term, and the species went undetected in all four sampling periods [$\prod_{t=1}^4 (1 - p_{it}^1)$]. The third additive term corresponds to the final possibility that the site was occupied and that successful reproduction occurred ($\psi_i^1\psi_i^2$), and that the species went undetected during all four periods [$\prod_{t=1}^4 (1 - p_{it}^2)$].

Missing observations (e.g., all sites are not visited the same number of times) can be dealt with by omitting any modeling of that portion of the detection history. For example, consider the history $h_i = 102.$, indicating that site i was not visited at the final period of a four-period study. The model corresponding to this detection history is

$$\text{Prob}(h_i = 102.) = \psi_i^1\psi_i^2p_{i1}^2(1 - \delta_{i1})(1 - p_{i2}^2)p_{i3}^2\delta_{i3}.$$

This probability can be contrasted with that provided above for similar history $h_i = 1021$, in which the observation at the final sampling period is modeled.

The likelihood function for all of the detection histories in a data set (the detection histories at all S sites) will be proportional to the product of the site-specific probabilities:

$$L(\boldsymbol{\psi}^1, \boldsymbol{\psi}^2, \mathbf{p}^1, \mathbf{p}^2, \boldsymbol{\delta}|\mathbf{h}) \approx \prod_{i=1}^S \text{Pr}(h_i)$$

where bold type denotes vectors or matrices. A simulation study indicated that resulting estimators show little bias (Appendix A). Simulation results also demonstrated that precision of $\hat{\psi}^2$, but not $\hat{\psi}^1$, is strongly influenced by the misclassification parameter, with larger δ associated with more precise estimates (Appendix A). As is the case for standard occupancy models, likelihoods with parameters corresponding to individual sites contain too many parameters and do not permit estimation without extra information. If individual site covariates associated with model parameters are available, then estimation of site-specific parameters is typically possible. Site-specific covariates may influence either occupancy or detection or classification parameters. Indeed, interesting biological questions are often addressed by modeling the different occupancy states as functions of site-specific covariates, for example asking what site characteristics are associated with occupancy and successful reproduction. Covariates specific to sampling periods may be used to model detection and classification parameters as well. Covariate modeling can use the linear-logistic relationship as in standard occupancy modeling (e.g., MacKenzie et al. 2002, 2006).

CALIFORNIA SPOTTED OWL REPRODUCTION: DATA

We apply the above model to data from daytime visits to California Spotted Owl territories during April through mid-August of 2004 at the Eldorado study area in the central Sierra Nevada, California, USA. Information about Spotted Owl population dynamics and sampling protocols at this study area can be found in

Seamans et al. (2001) and Franklin et al. (2004). Visits to territories include efforts to locate adult owls. Once located, an owl is offered a live mouse and then visually followed after the mouse is taken. Definitive evidence of reproductive activity is provided when owls take mice to the nest or to young, and young are observed. Non-reproducing owls usually eat or cache mice, but such behavior at a single visit may not be indicative of a failure to reproduce. Investigators thus develop classification criteria to conclude whether an owl reproduces successfully or not. Seamans (2005) evaluated criteria used by different research groups at different California Spotted Owl study areas and concluded that variation among groups precluded reasonable comparative analyses of reproductive data. Indeed, that result was a primary motivation of the current work, to provide a means of estimating the proportion of sites at which successful reproduction occurred in the face of uncertain state assignment.

We applied the above modeling approach to 54 owl territories believed to be occupied during the breeding season of 2004. Thus, ψ^1 was believed to be 1, or nearly so, for the 54 sites, and ψ^2 was the parameter of primary interest. The assessment involved a maximum of five visits to each site (one visit per month, April through mid-August), with an unequal number of visits (one to five) among the different sites. At each visit, the investigator recorded one of three possible values for owl detection: 0 = no detection of an owl, 1 = detection of an adult owl but no detection of young, and 2 = detection of at least 1 young owl. Thus, as in the development above, a 2 indicated occupancy and reproduction, a 1 indicated an occupied territory with uncertainty about reproductive status, and a 0 indicated uncertainty about occupancy and reproductive status.

Detection histories are shown in Appendix B. Owls were detected at 47 of the 54 sites yielding a naïve occupancy probability of $\tilde{\psi}^1 \approx 0.87$. Young (successful reproduction) were detected at least once at 19 of the 47 sites known to be occupied, yielding a naïve reproductive estimate of $\tilde{\psi}^2 \approx 0.40$. The naïve estimate for overall probability of successful reproduction for a site was $\tilde{\psi}^{1*2} = \tilde{\psi}^1\tilde{\psi}^2 = 0.35$.

CALIFORNIA SPOTTED OWL REPRODUCTION: MODELS AND ESTIMATION

We considered a set of 12 a priori models to describe the processes that gave rise to the detection history data. We did not explore the use of covariate data with this analysis, so we drop the site-specific subscript i and model parameters as common to the set of 54 sites. We modeled detection parameters as a single parameter for all sample periods and both true occupancy states ($p_{it}^s = p$), as time-independent but state-dependent ($p_{it}^s = p^s$, $s = 1, 2$), as time-dependent but independent of state ($p_{it}^s = p_t$, $t = 1, \dots, 5$), and as dependent on both time and state ($p_{it}^s = p_t^s$, $s = 1, 2$; $t = 1, \dots, 5$). Classification parameters were modeled as constant over all time periods ($\delta_{it} = \delta$),

TABLE 1. Model selection statistics for 12 multi-state occupancy models fit to California Spotted Owl reproductive success data from 54 owl territories, 2004 breeding season, in the central Sierra Nevada, California, USA.

Model	No. parameters	ΔQAIC_c	Weight, w_i
$\psi^1, \psi^2, p, \delta_2$	5	0.00	0.50
$\psi^1, \psi^2, p_i, \delta_2$	9	1.09	0.29
$\psi^1, \psi^2, p^s, \delta_2$	6	2.47	0.14
$\psi^1, \psi^2, p, \delta_i$	8	4.84	0.04
$\psi^1, \psi^2, p_i, \delta_i$	12	7.50	0.01
$\psi^1, \psi^2, p^s, \delta_i$	9	7.65	0.01
$\psi^1, \psi^2, p_i^s, \delta_2$	14	15.68	<0.01
$\psi^1, \psi^2, p_i^s, \delta_i$	17	25.35	<0.01
$\psi^1, \psi^2, p, \delta$	4	25.69	<0.01
$\psi^1, \psi^2, p_i, \delta$	8	26.33	<0.01
$\psi^1, \psi^2, p^s, \delta$	5	28.11	<0.01
$\psi^1, \psi^2, p_i^s, \delta$	13	39.86	<0.01

as one parameter for periods 1–2 and another parameter for periods 3–5 ($\delta_{it} = \delta_{1-2}, \delta_{3-5}$; denoted as δ_2 for model designation), and as five time-specific parameters ($\delta_{it} = \delta_i$). The rationale underlying the δ_2 model concerned the time-dependence of reproduction, and the likelihood that successful reproduction would be impossible or very difficult to detect in the early portion of the breeding season, but much more readily detected later. The two occupancy parameters (ψ^1, ψ^2) were always modeled as distinct. For high-ranking models we used a different parameterization to directly estimate the product $\psi^{1*2} = \psi^1\psi^2$. This product parameter is the probability that successful reproduction occurs at a randomly selected site (within the set of 54 sites) or, alternatively, the expected proportion of territorial owls that successfully reproduced. Models were fit and maximum likelihood estimates obtained using program SURVIV (White 1983).

Pearson goodness-of-fit statistics could not be computed for the most general models (the cell pooling algorithm left fewer cells than estimated parameters), so we used a deviance-based goodness-of-fit statistic to compute quasiliikelihood-adjusted small-sample Akaike’s Information Criterion, QAIC_c (see Burnham and Anderson 2002), where number of sites was used as the effective sample size. The model with the smallest QAIC_c value was selected as providing the best description of the data. Model weights, w_i , were also computed based on the QAIC_c values and reflect the relative weights of evidence in favor of the respective models being the most appropriate model among the members of the model set (see Burnham and Anderson 2002). For the parameter of primary interest, ψ^{1*2} , we computed a model-averaged estimate using models with $w_i > 0.01$. The associated variance estimate incorporated model uncertainty and was computed as recommended by Buckland et al. (1997).

CALIFORNIA SPOTTED OWL REPRODUCTION: RESULTS

The deviance-based estimate of the variance inflation factor used in quasiliikelihood adjustments was $\hat{c} = 1.74$,

providing evidence of lack of fit of the most general model ($\psi^1, \psi^2, p^s, \delta_i$). The lack of fit was not severe, yet it warranted use of \hat{c} in model selection and variance estimation. Model selection statistics indicated that model ($\psi^1, \psi^2, p, \delta_2$) provided the best description of the data, with models ($\psi^1, \psi^2, p_i, \delta_2$) and ($\psi^1, \psi^2, p^s, \delta_2$) also receiving substantial support (Table 1). Detection probability under the top model was estimated at about 0.74 (Table 2), very high but also substantially <1. Under the third model ($\psi^1, \psi^2, p^s, \delta_2$), detection probability was estimated to be slightly higher for sites at which successful reproduction had occurred than for non-reproducing sites. Under all three of the top models, classification probability was estimated as 0 for sample periods 1 and 2 and then 0.87 for the final three periods of the breeding season, a result generally consistent with expectation (Table 2).

The probability that a site was occupied, ψ^1 , was estimated to be 0.98 for the top models, whereas no owls were ever detected at seven of the 54 sites, yielding the naïve occupancy estimate $\hat{\psi}^1 \approx 0.87$. Examination of detection history data shows that five sites at which no detections were made were visited only once, one site was visited twice and the other three times. Indeed, the ability to incorporate relatively under-sampled sites into the analysis in a reasonable manner is one of the strengths of the approach presented here. The probability that successful reproduction occurred at an occupied site was estimated at about 0.45, again somewhat larger than the naïve estimate of proportion of occupied sites at which reproduction was detected $\hat{\psi}^2 \approx 0.40$. Parameter estimates for $\hat{\psi}^{1*2}$ under the three top models were nearly identical: $\hat{\psi}^{1*2} = 0.44$, $\widehat{\text{SE}}(\hat{\psi}^{1*2}) = 0.052$, $\hat{\psi}^{1*2} = 0.44$, $\widehat{\text{SE}}(\hat{\psi}^{1*2}) = 0.051$, and $\hat{\psi}^{1*2} = 0.43$, $\widehat{\text{SE}}(\hat{\psi}^{1*2}) = 0.052$ (Table 2). The model-averaged estimate of this product parameter is $\hat{\psi}^{1*2} = 0.44$, $\widehat{\text{SE}}(\hat{\psi}^{1*2}) = 0.069$, with the larger standard error reflecting model uncertainty. These estimates are larger than the naïve estimate of the proportion of sites at which successful reproduction was detected, $\hat{\psi}^{1*2} = 0.35$. Our approach thus yields a somewhat different assessment of reproductive success than provided by the raw count data.

DISCUSSION

Our approach to occupancy modeling and estimation with multiple occupancy states should be useful in cases where investigators not only can assess occupancy, but also can collect additional information about the status of an occupied site. Information about production of young at a site is certainly not limited to Spotted Owl sampling, and can be obtained for many species of birds (finding nests with eggs or young), amphibians (e.g., locating egg masses, dip netting tadpoles), mammals (e.g., observing tracks of young animals, photographing young animals in camera traps), and other taxa. Inference about not only occupancy, but also site productivity, should be useful for a variety of purposes.

TABLE 2. Parameter estimates under top three multi-state occupancy models fit to California Spotted Owl reproductive success data from 54 owl territories, 2004 breeding season in the central Sierra Nevada, California, USA.

Parameter	Model ($\psi^1, \psi^2, p, \delta_2$)		Model ($\psi^1, \psi^2, p_t, \delta_2$)		Model ($\psi^1, \psi^2, p^s, \delta_2$)	
	Estimate	SE	Estimate	SE	Estimate	SE
$\hat{\psi}^1$	0.98	0.038	0.98	0.032	0.98	0.038
$\hat{\psi}^2$	0.45	0.076	0.45	0.079	0.44	0.082
$\hat{\psi}^{1*2}$	0.44	0.052	0.44	0.051	0.43	0.052
\hat{p}_1^1	0.74	0.036	0.58	0.081	0.73	0.053
\hat{p}_2^1	0.74	0.036	0.82	0.067	0.73	0.053
\hat{p}_3^1	0.74	0.036	0.92	0.043	0.73	0.053
\hat{p}_4^1	0.74	0.036	0.70	0.077	0.73	0.053
\hat{p}_5^1	0.74	0.036	0.62	0.111	0.73	0.053
\hat{p}_1^2	0.74	0.036	0.58	0.081	0.76	0.058
\hat{p}_2^2	0.74	0.036	0.82	0.067	0.76	0.058
\hat{p}_3^2	0.74	0.036	0.92	0.043	0.76	0.058
\hat{p}_4^2	0.74	0.036	0.70	0.077	0.76	0.058
\hat{p}_5^2	0.74	0.036	0.62	0.111	0.76	0.058
$\hat{\delta}_{1-2}$	0.00	†	0.00	†	0.00	†
$\hat{\delta}_{3-5}$	0.87	0.072	0.87	0.072	0.87	0.073

† Parameter estimated at boundary of parameter space. Standard error cannot be estimated.

Estimation and subsequent mapping of occupancy state can be used for large-scale, albeit crude, assessments of viability of occupied sites and can be useful in land use and acquisition decisions (e.g., Stith and Kumar 2002). The modeling of occupancy as a function of habitat covariates can now be extended to investigate habitat variables associated with successful reproduction. Our parameterization of this problem expresses the overall probability that an area produces young as a product of two components, occupancy and reproduction given occupancy ($\hat{\psi}^{1*2} = \hat{\psi}_i^1 \hat{\psi}_i^2$), leading to considerations about the possibility of covariates and management being relevant to one or both components.

These models are somewhat related to those developed by Royle (2004) and Royle and Link (2005) for a very different purpose. Their focus was on categorical index data in surveys of calling anurans, with an aim to inference about different categories of abundance. Their models differ in several ways from the approach we present here, but the use of extra information to categorize occupied sites into different states is conceptually similar. The multiple species models of MacKenzie et al. (2004) can also be viewed as multi-state models, although again the nature of the states themselves and the approach to dealing with classification uncertainty differ from the approach presented here. Our inclusion of classification parameters to deal with the uncertainty of state assignment is related to the use of similar parameters in capture-recapture studies to deal with uncertainties associated with an animal's reproductive state (Kendall et al. 2003), sex (Fujiwara and Caswell 2002, Nichols et al. 2004), and even species (Runge et al. 2007).

Our example analysis merits discussion on both general and specific levels. The primary point of general interest is the relevance of dealing with detection probability and misclassification for proper inference. Detection probabilities for occupied owl territories were relatively high (averaging about 0.74), but still sufficiently small that nondetection at an occupied site was not a rare event. Correct classification probabilities at sites where young were produced were 0 for sampling occasions early in the season, as expected, and about 0.87 later in the season. These two sources of uncertainty, imperfect detection and classification, led to naïve estimates of the proportion of sites with productive pairs of 0.35, whereas our approach yielded estimates of about 0.44. If detection and/or correct classification probabilities had been smaller, as would be more typical of many ecological studies, the difference between naïve and model-based estimates would have been even larger. Incorporation of this uncertainty into estimation and modeling is not simply a statistical fine point, but is needed for reasonable inference.

With respect to Spotted Owls, meta-analyses of Spotted Owl survival rates have used estimation approaches that deal explicitly with detection probabilities <1 (Franklin et al. 1996, 2004, Anthony et al. 2006). Misclassification and nondetection have been recognized and discussed by analysts of Spotted Owl reproductive rate data. However, approaches that deal with these uncertainties have been based on standardization of field protocols rather than on direct modeling of the sampling processes involved. The result is that differences in protocol lead to difficulties in comparing reproductive rate data (Franklin et al. 2004, Seamans

2005). We recommend that approaches such as that developed here be strongly considered for use in Spotted Owl future analyses.

We are currently extending this sort of modeling to multiple seasons using a robust design approach (Pollock 1982) similar to that of MacKenzie et al. (2003). Such models will include not only probabilities of colonization and local extinction but also state transition probabilities for occupied sites. For example, we can envision situations in which occupancy itself might remain relatively constant over time, yet the proportion of sites producing young could decrease over time. In such a situation, inferences about occupancy state would be especially useful from conservation or ecological perspectives. Indeed, probabilities of transition from productive to nonproductive states (e.g., as resulting from the influence of environmental change on site quality) might prove to be more important than local extinction probabilities.

ACKNOWLEDGMENTS

We thank Ken Pollock, Wayne Thogmartin, and Gary White for constructive comments on a draft of this manuscript.

LITERATURE CITED

- Anthony, R. G., et al. 2006. Status and trends in demography of Northern Spotted Owls, 1985–2003. *Wildlife Monographs* 163.
- Buckland, S. T., K. P. Burnham, and N. H. Augustin. 1997. Model selection: an integral part of inference. *Biometrics* 38: 469–477.
- Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer-Verlag, New York, New York, USA.
- Franklin, A. B., D. R. Anderson, E. D. Forsman, K. P. Burnham, and F. W. Wagner. 1996. Methods for collecting and analyzing demographic data on the Northern Spotted Owl. *Studies in Avian Biology* 17:12–20.
- Franklin, A. B., D. R. Anderson, R. J. Gutierrez, and K. P. Burnham. 2000. Climate, habitat quality, and fitness in Northern Spotted Owl populations in northwestern California. *Ecological Monographs* 70:539–590.
- Franklin, A. B., et al. 2004. Population dynamics of the California Spotted Owl: A meta analysis. *American Ornithologists' Union Monograph*. No. 54.
- Fretwell, S. D. 1972. Populations in a seasonal environment. Princeton University Press, Princeton, New Jersey, USA.
- Fujiwara, M., and H. Caswell. 2002. Estimating population projection matrices from multi-stage mark–recapture data. *Ecology* 83:3257–3265.
- Kendall, W. L., J. E. Hines, and J. D. Nichols. 2003. Adjusting multi-state capture–recapture models for misclassification bias: manatee breeding proportions. *Ecology* 84:1058–1066.
- MacKenzie, D. I., L. L. Bailey, and J. D. Nichols. 2004. Investigating species co-occurrence patterns when species are detected imperfectly. *Journal of Animal Ecology* 73:546–555.
- MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology* 84:2200–2207.
- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83:2248–2255.
- MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. Academic Press, New York, New York, USA.
- Nichols, J. D., W. L. Kendall, J. E. Hines, and J. A. Spendlow. 2004. Estimation of sex-specific survival from capture–recapture data when sex is not always known. *Ecology* 85: 3192–3201.
- Pollock, K. H. 1982. A capture–recapture design robust to unequal probability of capture. *Journal of Wildlife Management* 46:752–757.
- Pulliam, H. R. 1988. Sources, sinks and population regulation. *American Naturalist* 132:652–661.
- Royle, J. A. 2004. Modeling abundance index data from anuran calling surveys. *Conservation Biology* 18:1378–1385.
- Royle, J. A., and W. A. Link. 2005. A general class of multinomial mixture models for anuran calling survey data. *Ecology* 86:2505–2512.
- Runge, J. P., J. E. Hines, and J. D. Nichols. 2007. Estimating species-specific survival and movement when species identification is uncertain. *Ecology* 88:xx–xx.
- Runge, J. P., M. C. Runge, and J. D. Nichols. 2006. The role of local populations within a landscape context: defining and classifying sources and sinks. *American Naturalist* 167:925–938.
- Seamans, M. E. 2005. Population biology of the California Spotted Owl in the central Sierra Nevada. Dissertation. University of Minnesota, Minneapolis, Minnesota, USA.
- Seamans, M. E., R. J. Gutiérrez, C. A. Moen, and M. Z. Perry. 2001. Spotted Owl demography in the central Sierra Nevada. *Journal of Wildlife Management* 65:425–431.
- Stith, B. M., and N. S. Kumar. 2002. Spatial distributions of tigers and prey: mapping and the use of GIS. Pages 51–59 in K. U. Karanth and J. D. Nichols, editors. *Monitoring tigers and their prey. A manual for wildlife managers, researchers, and conservationists*. Centre for Wildlife Studies, Bangalore, India.
- White, G. C. 1983. Numerical estimation of survival rates from band recovery and biotelemetry data. *Journal of Wildlife Management* 47:716–728.

APPENDIX A

A figure showing results of the simulation study (*Ecological Archives* XXXX-XXX-XX).

APPENDIX B

Site detection histories (*Ecological Archives* XXXX-XXX-XX).