

## A BAYESIAN STATE-SPACE FORMULATION OF DYNAMIC OCCUPANCY MODELS

J. ANDREW ROYLE<sup>1,3</sup> AND MARC KÉRY<sup>2</sup>

<sup>1</sup>*U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA*

<sup>2</sup>*Swiss Ornithological Institute, 6204 Sempach, Switzerland*

**Abstract.** Species occurrence and its dynamic components, extinction and colonization probabilities, are focal quantities in biogeography and metapopulation biology, and for species conservation assessments. It has been increasingly appreciated that these parameters must be estimated separately from detection probability to avoid the biases induced by non-detection error. Hence, there is now considerable theoretical and practical interest in dynamic occupancy models that contain explicit representations of metapopulation dynamics such as extinction, colonization, and turnover as well as growth rates. We describe a hierarchical parameterization of these models that is analogous to the state-space formulation of models in time series, where the model is represented by two components, one for the partially observable occupancy process and another for the observations conditional on that process. This parameterization naturally allows estimation of all parameters of the conventional approach to occupancy models, but in addition, yields great flexibility and extensibility, e.g., to modeling heterogeneity or latent structure in model parameters. We also highlight the important distinction between population and finite sample inference; the latter yields much more precise estimates for the particular sample at hand. Finite sample estimates can easily be obtained using the state-space representation of the model but are difficult to obtain under the conventional approach of likelihood-based estimation. We use R and WinBUGS to apply the model to two examples. In a standard analysis for the European Crossbill in a large Swiss monitoring program, we fit a model with year-specific parameters. Estimates of the dynamic parameters varied greatly among years, highlighting the irruptive population dynamics of that species. In the second example, we analyze route occupancy of Cerulean Warblers in the North American Breeding Bird Survey (BBS) using a model allowing for site-specific heterogeneity in model parameters. The results indicate relatively low turnover and a stable distribution of Cerulean Warblers which is in contrast to analyses of counts of individuals from the same survey that indicate important declines. This discrepancy illustrates the inertia in occupancy relative to actual abundance. Furthermore, the model reveals a declining patch survival probability, and increasing turnover, toward the edge of the range of the species, which is consistent with metapopulation perspectives on the genesis of range edges. Given detection/non-detection data, dynamic occupancy models as described here have considerable potential for the study of distributions and range dynamics.

*Key words:* Bayesian analysis; detection probability; heterogeneity; hierarchical models; presence/absence data; range dynamics; site occupancy; state-space models; turnover; WinBUGS.

### INTRODUCTION

Species distribution and patch occupancy are fundamental concepts in biogeography, landscape ecology and metapopulation biology and are also of considerable interest in the conservation and management of animal and plant species. Moreover, studies of factors that influence occurrence and distribution, such as habitat or landscape structure, are the focus of many ecological investigations. However, in surveys of most species it is not possible to observe occurrence without error. That is, a species may go undetected where it is present yielding a false negative (detection) error. For instance,

in breeding bird surveys in North America and Switzerland, only ~76% and 89%, respectively, of all avian species present were estimated to be detected in a year (Boulinier et al. 1998b, Kéry and Schmid 2006). Thus, explicit attention to imperfect detection in animal sampling is necessary in order to obtain summaries of occurrence or distribution that have a clear conceptual interpretation. The problem of estimating occurrence probability, proportion of area occupied, or “site occupancy,” of a species subject to imperfect detection is a problem that has received considerable attention recently (Bayley and Peterson 2001, Kéry 2002, MacKenzie et al. 2002, Nichols and Karanth 2002, Royle and Nichols 2003, Tyre et al. 2003, Wintle et al. 2003, Gu and Swihart 2004, Pellet and Schmidt 2005, Royle and Link 2005, Schmidt 2005, Weir et al. 2005). Site occupancy models are becoming widely adopted in

Manuscript received 21 April 2006; revised 17 November 2006; accepted 6 December 2006. Corresponding Editor: F. He.

<sup>3</sup> E-mail: aroyle@usgs.gov

many survey and monitoring activities due to, at least in part, the ease of establishing surveys based on presence/absence (or rather, detection/non-detection) data, that such models facilitate an explicit accounting for detectability of the species in question, and the extensibility of the site occupancy modeling framework. In addition, site occupancy models can be used as a framework for estimating or modeling abundance either in a model-based framework (He and Gaston 2000a, b, Royle and Nichols 2003) or if the area of sample units approximates the home range of individuals. This greatly extends the range of their application.

A basic class of models can be developed (e.g., MacKenzie et al. 2002) under the assumption that the occupancy status of sample sites does not change across replicate (temporal) samples, i.e., that sites are closed to local extinction and colonization events. While such models are sufficient for many applications, there exists considerable interest in dynamic attributes of metapopulations. For example, colonization and extinction are central to metapopulation theory (Hanski 1999) and overlap and turnover are important summaries of community similarity and stability (Boulinier et al. 1998a, 2001, Nichols et al. 1998a, b, Doherty et al. 2003a, b). Thus, one important extension of the site occupancy modeling framework is to the situation in which a site's occupancy status may change through time, i.e., to the situation in which the metapopulation system is "open" to local extinction and colonization events. MacKenzie et al. (2003) provided a general characterization of open models, and described a likelihood-based framework for inference about model parameters. Under these open models, the metapopulation system is assumed to be closed within, but not across primary periods. The primary periods have been colloquially referred to as "seasons," and the models are commonly referred to as "multi-season" occupancy models. Here we adopt the adjective *dynamic* as being more descriptive because it reflects the fact that the models explicitly incorporate the metapopulation dynamical processes of extinction and colonization. Conversely, the term "multi-season" can equally well apply to a model that does not explicitly acknowledge these dynamic processes, i.e., one that has a time-varying occupancy probability, in which occupancy status is assumed to be independent across time.

In this paper, we describe a hierarchical, or state-space, representation of dynamic models for site occupancy in which the model for the observed detection/non-detection data is expressed as the product of two component models, first a submodel for the data conditional on the latent (unobserved or partially observed) process, and, secondly, a submodel for the latent occupancy process. We refer to this representation of the model as the state-space representation, as it is conceptually equivalent to that found in time series and similar settings (Jones 1993, Berliner 1996).

Whereas classical likelihood-based solutions to the problem (e.g., MacKenzie et al. 2003) carry out inference by removing the latent state process from the model by marginalization, we retain the state process as this representation yields itself naturally to several important extensions, such as the incorporation of random effects to accommodate spatial or temporal variation and structure in model parameters. In addition, inference in many applications of site occupancy models does not focus on some larger statistical population of sites, but rather on the particular sample of sites at hand. This is an instance wherein estimation of finite sample quantities is of interest, and it has perhaps not been widely appreciated in the ecological community that this is an inference problem distinct from the estimation of the parameters in a larger population. Explicitly accounting for a finite sample results in greatly improved precision of estimators and is easily achieved with a state-space formulation of dynamic occupancy models. While we believe this hierarchical representation of dynamic site occupancy models yields a more general formulation of such models, it is also a simple and natural way to express the models, and a Bayesian analysis of the hierarchical parameterization can be achieved directly with popular software for Bayesian analysis (WinBUGS).

We provide examples using avian survey data from the Swiss Survey of Common Breeding Birds, and an example with North American Breeding Bird Survey (BBS) data on Cerulean Warblers in which the model allows for spatial heterogeneity in parameters among sites. While such models have great biological appeal, they have not been explicitly addressed in the literature because of the absence of a practical framework for the analysis of these complex models. Therefore, we also include scripts to fit several versions of our models using freely available software.

#### SAMPLING DESIGN AND THE DYNAMIC MODEL

We consider data obtained from repeated presence/absence surveys of  $i = 1, 2, \dots, R$  spatial units (patches or "sites"). We suppose that each site is surveyed  $j = 1, 2, \dots, J$  times within each of  $t = 1, 2, \dots, T$  primary periods and that each site is closed with respect to its occupancy status within but not across primary periods. A typical case would be surveys repeated several times both within the breeding season of a species and over several years. This situation is that for which the "robust design" (Pollock 1982, Kendall et al. 1995, Williams et al. 2002: chapter 19) has been developed in conventional capture-recapture applications, and has also been exploited by MacKenzie et al. (2003) for developing site occupancy models for open populations.

Denote the observed occupancy status of site  $i$  for survey  $j$  within primary period  $t$  as  $y_j(i, t)$ . We suppose that  $y_j(i, t)$ ,  $j = 1, 2, \dots, J$  are independent and identically distributed for each site ( $i$ ) and primary

period ( $t$ ). Let  $z(i, t)$  denote the true occupancy status of unit  $i$  during primary period  $t$ , having possible states “occupied” ( $z = 1$ ) or “not occupied” ( $z = 0$ ). One parameter of interest is the probability of site occupancy (or the probability of occurrence) for period  $t$ ,  $\psi_t = \Pr(z(i, t) = 1)$ . Changes in occupancy over time can be parameterized explicitly in terms of local extinction and colonization processes, analogous to population demographic processes of survival and recruitment. Let  $\phi_t$  be the probability that an occupied site “survives” (i.e., remains occupied) from period  $t$  to  $t + 1$ , i.e.,  $\phi_t = \Pr(z(i, t + 1) = 1 | z(i, t) = 1)$ . Local extinction probability ( $\varepsilon_t$ ) in the parameterization used by MacKenzie et al. (2003), is the complement of  $\phi_t$ , i.e.,  $\varepsilon_t = 1 - \phi_t$ . In metapopulation systems, local colonization is the analog of the recruitment process. Let  $\gamma_t$  be the local colonization probability from period  $t$  to  $t + 1$ , i.e.,  $\gamma_t = \Pr(z(i, t + 1) = 1 | z(i, t) = 0)$ .

*State-space representation*

This model is naturally formulated as a state-space model, in which we express the model by its two component processes: a submodel for the observations conditional on the unobserved state process, i.e.,  $y(i, t) | z(i, t)$  and, secondly, a submodel for the unobserved or partially observed state process  $z(i, t); i = 1, 2, \dots, t = 1, 2, \dots$ . The classical likelihood formulation of this problem exploits an identical model structure, but removes the latent indicators of occupancy from the likelihood by integration. In adopting a Bayesian framework for analysis and inference (described in *Bayesian Analysis*), retention of the latent variables in the model does not pose any difficulty. This is attractive since for many inference problems, retaining the latent  $z(i, t)$ 's in the model is necessary because they are fundamental to the object of inference. Discussion of specific examples when this is the case is given in *Metapopulation summaries*, below.

*State model.*—The state model has a simple formulation in terms of initial occupancy probability, i.e., at  $t = 1$ , which we will designate  $\psi_1$ , local survival probability,  $\{\phi_t\}_{t=1}^{T-1}$ , and the recruitment (colonization) parameters  $\{\gamma_t\}_{t=1}^{T-1}$ . The initial occupancy states are assumed to be *iid* Bernoulli random variables, denoted as

$$z(i, 1) \sim \text{Bernoulli}(\psi_1) \text{ for } i = 1, 2, \dots, R \quad (1)$$

whereas, in subsequent periods

$$z(i, t) | z(i, t - 1) \sim \text{Bernoulli}\{z(i, t - 1)\phi_{t-1} + [1 - z(i, t - 1)]\gamma_{t-1}\} \quad (2)$$

for  $t = 2, 3, \dots, T$ . Thus, for a site that is occupied at  $t - 1$  (i.e.,  $z(i, t - 1) = 1$ ), the survival component in Eq. 2 is operative (the first component of the Bernoulli success probability), and  $z(i, t)$  is a Bernoulli outcome with probability  $\phi_{t-1}$ . Conversely, if a site is not occupied at time  $t - 1$ , then the recruitment

component in Eq. 2 becomes operative and  $z(i, t)$  is a Bernoulli outcome with parameter  $\gamma_{t-1}$ . The expressions in Eqs. 1 and 2 define the state process model. Generalizations, where  $\phi$  and  $\gamma$  may be structured spatially or temporally, for instance, are described later.

*Observation model.*—The observation model, specified conditional on the latent process  $\{z(i, t)\}$ , is given by

$$y_j(i, t) | z(i, t) \sim \text{Bernoulli}[z(i, t)p_{jt}].$$

Thus, if a site is occupied at time  $t$ , the data are Bernoulli trials with parameter  $p_{jt}$ . If a site is unoccupied at time  $t$ , then the data are Bernoulli trials with  $\Pr(y(i, t) = 1) = 0$ . Obvious generalizations to accommodate structure in  $p$  can be obtained directly. For example, in many applications, sampling covariates are measured during each survey such as search effort, weather, or other environmental conditions. In this case, the observation model is Bernoulli with index  $z(i, t)p_{ijt}$  and covariate effects are modeled on a suitable transformation of  $p_{ijt}$ .

*Metapopulation summaries*

The primary parameters of the dynamic occupancy model are the initial occupancy probability,  $\psi_1$ , the survival probabilities  $\{\phi_t\}_{t=1}^{T-1}$ , the colonization probabilities  $\{\gamma_t\}_{t=1}^{T-1}$  and the detection probabilities  $\{p_{jt}\}$ . In addition, a number of derived parameters are of interest. First of all, the occupancy probability at  $t$  can be computed recursively according to

$$\psi_t = \psi_{t-1}\phi_{t-1} + (1 - \psi_{t-1})\gamma_{t-1} \quad (3)$$

for  $t = 2, \dots, T$ . MacKenzie et al. (2003) defined the growth rate as

$$\lambda_t = \frac{\psi_{t+1}}{\psi_t}.$$

Nichols et al. (1998a) define “turnover” as “the probability that an occupied quadrat picked at random is a newly occupied one.” That is, turnover is the probability  $\Pr(z(t - 1) = 0 | z(t) = 1)$ . Bayes’ rule yields an expression for this in terms of previously defined model parameters:

$$\tau_t = \frac{\gamma_{t-1}(1 - \psi_{t-1})}{\gamma_{t-1}(1 - \psi_{t-1}) + \phi_{t-1}\psi_{t-1}} \quad (4)$$

for  $t = 2, \dots, T$ . The denominator here is equal to  $\psi_t$ , by Eq. 3.

One useful summary of the dynamical system is the equilibrium occupancy probability (Hanski 1994, MacKenzie et al. 2005: chapter 7; Brander et al 2007), i.e., the stable-state (occupancy) distribution. This is related to local survival and colonization according to

$$\psi_t^{(eq)} = \frac{\gamma_t}{\gamma_t + (1 - \phi_t)}.$$

This is the leading element of the dominant eigenvector of the state transition matrix. When  $\gamma$  and  $\phi$  are year

specific, the associated collection of year-specific values of  $\psi_{\text{eq}}$  provide a convenient summary of occupancy dynamics, e.g.,  $\psi_{t+1}^{(\text{eq})}/\psi_t^{(\text{eq})}$  summarizes the change in equilibrium occupancy as a result of the change in survival and colonization from  $t$  to  $t + 1$ .

#### Finite sample estimation

It is perhaps not widely understood that there is a distinction between estimation of population parameters (e.g., the probability that a site is occupied, occupancy), and estimation of finite-sample manifestations of those quantities such as the number of occupied sites in the actual sample (Dorazio and Royle 2005, MacKenzie et al. 2005: section 4.5).

Site occupancy parameters,  $\psi_t$ , as defined previously, are population averages and associated estimates apply to a theoretically infinite population of sites from which the sample of size  $R$  was drawn. However, given a sample of  $R$  sites, the actual proportion of sample sites occupied is a function of the latent state variables:

$$\psi_t^{(\text{fs})} = \frac{1}{R} \sum_i z(i, t) \quad (5)$$

where “fs” is for finite sample. While the expected values of the two quantities may be equivalent, the uncertainty of the point estimates will typically be very different (we demonstrate this in an example, below).

Finite-sample equivalents of occupancy or of other quantities can easily be computed as functions of the state-variables,  $z$ , whereas estimation of such quantities is difficult in the conventional framework of MacKenzie et al. (2002, 2003). For example, the estimator of sample growth rate is

$$\lambda_t^{(\text{fs})} = \frac{\sum_{i=1}^R z(i, t+1)}{\sum_i z(i, t)} \quad (6)$$

and the sample turnover rate is

$$\tau_t^{(\text{fs})} = \frac{\sum_{i=1}^R [1 - z(i, t-1)]z(i, t)}{\sum_{i=1}^R z(i, t)} \quad (7)$$

One of the main benefits of the state-space parameterization of occupancy models is that it permits the construction of these finite-sample estimators, as functions of the latent  $z(i, t)$  variables. That some of the  $z(i, t)$  are unobserved does not pose any difficulty in estimating  $\psi_t^{(\text{fs})}$  and related quantities using common methods of Bayesian analysis based on Markov chain Monte Carlo (MCMC). Using MCMC, one obtains a Monte Carlo sample of each model parameter, and each latent variable, from the target posterior distribution. The Monte Carlo samples are then used to obtain the desired summaries of the posterior distribution, such as the mean, standard deviation, or quantiles used to construct confidence intervals. Details are provided in *Estimation by Markov chain Monte Carlo*. Conversely, estimation and inference of these finite-sample estimands is difficult using conventional likelihood-based

methods. Typically, one must resort to empirical Bayes estimators of such quantities (Dorazio and Royle 2005).

The distinction between  $\psi_t$  and  $\psi_t^{(\text{fs})}$  has to do primarily with the scope of inference. If one is interested in the particular sites for which data were collected, then the estimator should be the finite-sample estimator of the proportion of occupied sites. Conversely, if one is interested in a (much) larger collection of sites from which the sampled sites are representative (e.g., randomly selected), then the estimator  $\psi_t$  is appropriate. While the distinction may seem subtle or esoteric, there can be very important consequences. If the scope of inference is restricted to the sites for which data are available, we can expect the variance of the estimate of the proportion of occupied sites to be much smaller because some of the occupancy state variables were observed, and they do not contribute to the variance of the estimated total (number of occupied sites). In other words, in Eq. 5, some of the  $z(i, t)$  are known (i.e., at sites where the species was observed) and so the variation of this quantity is less than that of  $R$  independent Bernoulli trials. This problem of estimating sample quantities is analogous to classical small-area estimation and the estimation of random effects common throughout statistics (Laird and Ware 1982, Robinson 1991).

In *European Crossbill: Variation among years* we illustrate the distinction between population quantities and their finite-sample manifestations with a brief example using data on the crossbill from the Swiss Survey of Common Breeding Birds.

#### BAYESIAN ANALYSIS

The simple conditional specification of the dynamic site occupancy model yields easily to Bayesian analysis using conventional methods of Markov chain Monte Carlo (MCMC). MCMC implementation can be achieved using Gibbs sampling (Geman and Geman 1984), which is based on drawing samples of each unknown quantity (i.e., parameters and latent variables) from their “full-conditional” distributions, i.e., the distribution of a parameter conditional on all other unknown quantities and the data. For the dynamic occupancy models, the Gibbs sampler is remarkably simple when certain natural prior distributions are used. Details are provided in the Appendix for the case where survival and colonization probabilities are allowed to vary by primary period. For the applications described subsequently, we use the freely available software package WinBUGS 1.4. (Gilks et al. 1994) to conduct the analyses, and we provide the WinBUGS model specification for those examples.

#### Implementation in WinBUGS

While the simple structure of the MCMC algorithm outlined in the Appendix is potentially useful, in practice, the dynamic model can be implemented easily in the popular software package WinBUGS, with little

more effort required than providing a “pseudo-code” description of the model. For example, the WinBUGS model specification for a model in which all parameters vary by year, say  $\{\psi_1, p_t, \phi_t, \gamma_t\}$ , is given in the Supplement, along with code for calculation of the year-specific occupancy probabilities, growth rates, and also the (finite) sample occupancy rates.

For the analysis in the example of the next section, the data were organized in the free software package R version 2.2 (Ihaka and Gentleman 1996, R Development Core Team 2005) and WinBUGS 1.4 was executed in batch mode using the R add-on library R2WinBUGS.

APPLICATIONS

We provide two examples of dynamic site occupancy models. First is a basic example, part of which could as well be done by conventional likelihood methods in the widely used programs PRESENCE or MARK. Secondly, we provide a more complicated example for which likelihood analysis is not straightforward. We consider an occupancy analysis of Cerulean Warblers (*Dendroica cerulea*) from the North American Breeding Bird Survey (BBS) using models in which demographic parameters are regarded as spatially and temporally varying random effects.

*European Crossbill: Variation among years*

We consider data from the Swiss Survey of Common Breeding Birds (Monitoring Häufige Brutvögel; MHB) over four years (2001–2004) for the European Crossbill (*Loxia curvirostra*). This program has been conducted annually since 1999 by the Swiss Ornithological Institute (Schmid et al. 2004, Kéry et al. 2005, Kéry and Schmid 2006). A systematic random sample of 267 1-km<sup>2</sup> quadrats across Switzerland is surveyed three times during the breeding season (15 April–15 July) on a quadrat-specific, irregular, transect route averaging 5 km using the territory mapping method (Bibby et al. 1992). The Crossbill is a medium-sized (34 g) pine-seed-eating finch widespread in Switzerland. Its abundance and occurrence depend greatly on the cone set of conifers, and in most years, Crossbills appear irruptively in many regions where otherwise they do not occur or are scarce. This species may have the greatest dynamics in site occupancy among all regular Swiss breeding birds.

We fitted the model in which each of site survival rate, colonization and detection probability were year specific, and computed the year-specific occupancy, growth, and turnover rates. The WinBUGS model specification is given in the Supplement. Posterior summaries were based on 20 000 MCMC iterations after discarding 2000 samples as “burn-in.” The canonical parameter estimates and certain derived parameters, as well as both population and finite sample occupancy estimates are given in Table 1.

There were large interannual changes in Crossbill occupancy in Switzerland with strong growth between

TABLE 1. Posterior summaries of model and derived parameters (beginning with  $\psi_2$  and below) from the dynamic occupancy model fitted to the Swiss Breeding Bird Survey occupancy data on the European Crossbill (*Loxia curvirostra*) from 2001 to 2004.

Parameter	Mean	SD	$q_{0.025}$	$q_{0.500}$	$q_{0.975}$
$p_1$	0.584	0.044	0.493	0.584	0.666
$p_2$	0.493	0.037	0.422	0.493	0.564
$p_3$	0.566	0.033	0.504	0.566	0.629
$p_4$	0.574	0.037	0.499	0.574	0.643
$\phi_1$	0.806	0.069	0.656	0.812	0.931
$\phi_2$	0.855	0.046	0.758	0.858	0.938
$\phi_3$	0.682	0.053	0.576	0.682	0.791
$\gamma_1$	0.259	0.037	0.189	0.257	0.334
$\gamma_2$	0.190	0.041	0.114	0.189	0.273
$\gamma_3$	0.071	0.029	0.022	0.068	0.133
$\psi_1$	0.242	0.029	0.190	0.241	0.300
$\psi_2$	0.391	0.035	0.323	0.390	0.461
$\psi_3$	0.450	0.034	0.386	0.450	0.517
$\psi_4$	0.346	0.032	0.286	0.345	0.409
$\psi_1^{(fs)}$	0.240	0.0124	0.222	0.237	0.271
$\psi_2^{(fs)}$	0.389	0.0210	0.353	0.387	0.436
$\psi_3^{(fs)}$	0.449	0.0149	0.425	0.447	0.481
$\psi_4^{(fs)}$	0.345	0.0148	0.320	0.342	0.380
$\lambda_1$	1.635	0.201	1.284	1.619	2.085
$\lambda_2$	1.157	0.099	0.978	1.153	1.371
$\lambda_3$	0.770	0.065	0.651	0.768	0.908
$\tau_1$	0.499	0.058	0.384	0.500	0.610
$\tau_2$	0.259	0.056	0.151	0.258	0.371
$\tau_3$	0.113	0.046	0.032	0.110	0.213
$\tau_1^{(fs)}$	0.498	0.032	0.429	0.500	0.555
$\tau_2^{(fs)}$	0.254	0.041	0.170	0.256	0.328
$\tau_3^{(fs)}$	0.104	0.034	0.034	0.103	0.170

Note: For the column headings,  $q_x$  is the  $100 \times x$ th percentile of the posterior distribution.

2001–2002 and 2002–2003, followed by a decline in occupancy between 2003 and 2004. Correspondingly, the growth rate declined steadily between 2001–2002 and 2003–2004. Estimation of survival and colonization rates enabled one to decompose this rate of change: patch survival probability remained high during years when the population expanded (2001 through 2003) and declined in 2003–2004, while colonization probabilities declined steadily over the years considered.

The high turnover rates confirm the highly erratic occupancy pattern of the Crossbill. Detection probability,  $p$ , did not change much across years, and this insight might be useful in the interpretation of surveys that do not correct for variations of detection probability. These results indicate that in the Swiss program, the number of quadrats found occupied by the crossbill may be a somewhat valid index to true occupancy by that species.

TABLE 2. Estimates of parameters from the year- and site-effects model fit to the Cerulean Warbler data.

Parameter	Mean	SD	$q_{0.025}$	$q_{0.500}$	$q_{0.975}$
$\psi$	0.600	0.068	0.465	0.600	0.737
$\sigma_\phi$	1.909	0.535	1.066	1.833	3.204
$\sigma_\gamma$	1.620	0.322	1.016	1.613	2.280
$\sigma_p$	1.453	0.070	1.319	1.453	1.592
$\phi_{\min}^*, t = 3$	0.647	0.132	0.365	0.415	0.878
$\phi_{\text{med}}^*, t = 4$	0.874	0.107	0.607	0.664	0.996
$\phi_{\max}^*, t = 31$	0.976	0.025	0.908	0.929	0.999
$\gamma_{\min}^*, t = 34$	0.128	0.128	0.003	0.006	0.478
$\gamma_{\text{med}}^*, t = 29$	0.321	0.192	0.026	0.049	0.751
$\gamma_{\max}^*, t = 4$	0.528	0.138	0.261	0.300	0.795
$\lambda_{\min}^{(fs)}, t = 32$	0.850	0.063	0.731	0.851	0.973
$\lambda_{\text{med}}^{(fs)}, t = 8$	0.988	0.103	0.817	0.977	1.227
$\lambda_{\max}^{(fs)}, t = 16$	1.174	0.159	0.914	1.162	1.53

Notes: For the column headings,  $q_x$  is the  $100 \times x$ th percentile of the posterior distribution. For parameters that are year specific, the minimum, median, and maximum values are given. The realized change in occupancy from year  $t$  to  $t + 1$  is  $\lambda_t^{(fs)}$ .

In comparing the annual estimates of  $\psi_t$  for the four years, along with the estimated proportion of sample sites occupied (i.e.,  $\psi_t^{(fs)}$ ), we note that the point estimates are equivalent (to within Monte Carlo error), though this need not generally be the case when the posterior distribution of these quantities is asymmetric. The main difference is that the finite-sample estimates are typically about 50% more precise, e.g., the 95% posterior interval for  $\psi_1$  is (0.190, 0.300), whereas the interval for  $\psi_1^{(fs)}$  is (0.222, 0.271). In Table 2, we have also provided both estimates of turnover: the population estimate  $\tau_t$ , in Eq. 4, and also the finite population estimator,  $\tau_t^{(fs)}$  given by Eq. 7. As before, the finite-sample point estimate is not very different from the population analog, but it is much more precise.

*Cerulean Warblers: Heterogeneity among sites*

Here we consider data on Cerulean Warbler (*Dendroica cerulea*) from the North American Breeding Bird Survey (BBS) using 40 years of data (1966–2005), and 446 BBS routes covering substantially all of the Cerulean Warbler range. For more details on the species and the data see Link and Sauer (2002) and Thogmartin et al. (2004) and the references therein. Cerulean Warblers are experiencing population declines throughout most of their range and there has been considerable attention focused on assessing the nature and cause of such declines. Here, we develop a model of route level occupancy. Following Boulmier et al. (1998b) and Nichols et al. (1998a, b) (among others), we use detection/non-detection observations at the observation level which yields 50 replicate observations (subsamples) per route. The use of spatial subsamples as replicates

yields a definition of occupancy as a route-level attribute (Nichols et al. 1998a, b), i.e., relevant to the aggregation of spatial subsamples.

Models with so-called “individual heterogeneity” are of some interest in classical population size estimation problems (Dorazio and Royle 2003), and have been considered also in the context of site occupancy models (Royle 2006). In the present situation, we expect that heterogeneity in parameters among sites should exist due to geographic variation in abundance across the species’ range. For example, survival probability should be higher in regions of relatively higher abundance. The abundance-induced mechanism underlying heterogeneity (and also correlation among parameters) was also used in multi-species occupancy by Dorazio and Royle (2005). Various other ecological hypotheses could be developed that imply structure among parameters of the model. Indeed, a recent focus in metapopulation studies has been on the spatial structuring of metapopulations (González-Megías et al. 2005, Johnson 2005).

We consider a model for the Cerulean Warbler data which allows for heterogeneity among sites in all model parameters (detection, survival, and colonization probabilities), in addition to fixed year effects. Specifically, we suppose that

$$\text{logit}(\phi_{it}) = a_t + u_i$$

and

$$\text{logit}(\gamma_{it}) = b_t + v_i$$

where  $(u_i, v_i)$  are the site effects, assumed to be normal with standard deviations  $\sigma_\phi$  and  $\sigma_\gamma$ . The parameters  $a_t$  and  $b_t$  are the fixed year effects for each year interval  $t = 1, 2, \dots, T - 1$ . The observation model is

$$y_j(i, t) | z(i, t) \sim \text{Bernoulli}[z(i, t)p_{i,t}]$$

and we have assumed a similar site and year-effects structure on the logit-transform of  $p_{i,t}$ . Specifically, the logit-transformed  $p_{i,t}$  is additive in year and site effects:

$$\text{logit}(p_{i,t}) = c_t + w_i$$

where  $c_t$  are the fixed year effects and  $w_i \sim \text{Normal}(0, \sigma_p^2)$ . For the  $\{a_t, b_t, c_t\}$  parameters, we assume priors that are customarily regarded as being suitably vague. Specifically,  $\text{logit}^{-1}(a_t) \sim \text{Uniform}(0, 1)$ , and similarly for  $b_t$  and  $c_t$ . Standard inverse-gamma priors were assumed for the inverse of all variance components. The WinBUGS model specification is given in the Supplement.

The MCMC algorithm was used to obtain 100 000 posterior draws of model parameters after a 10 000 burn-in period. Every third draw was retained for summarization. Posterior summaries of some model parameters are displayed in Table 2. In this table, the year-specific parameters  $a_t$  and  $b_t$  are summarized by their inverse-logits, e.g.,  $\phi_t^* = \exp\{a_t\} / (1 + \exp\{a_t\})$ , for

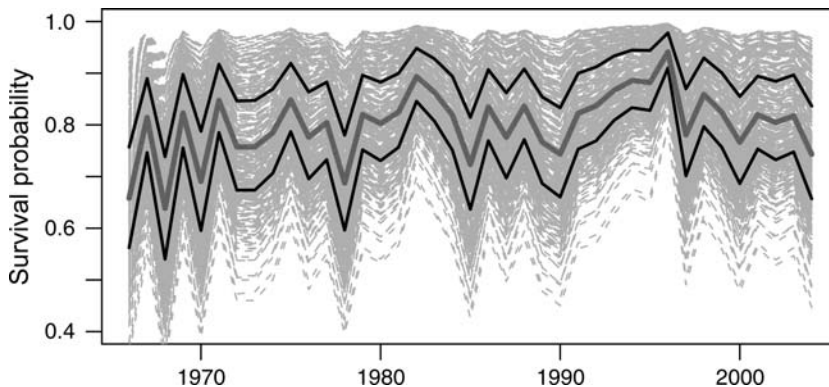


FIG. 1. Variation in patch- or site-survival probability among sites and across years for the Cerulean Warbler data. Solid lines are the mean (center, thick medium-gray line) and quartiles (black lines). Site-specific time-series are given in light gray.

the minimum, median, and maximum values of  $a_t$  across the 39 annual periods.

The network of sampling locations, when viewed as a metapopulation, is characterized by very high survival probabilities across years ( $\phi_{med}^* = 0.874$ ), relatively low colonization probabilities ( $\gamma_{med}^* = 0.321$ ), but considerable spatial heterogeneity (among routes) in both survival (Fig. 1) and colonization probability (Fig. 2). Annual changes in occupancy as measured by the sample growth rate (Eq. 6) were fairly stable. Disregarding the first six years which were highly variable and imprecise due to small sample sizes, occupancy growth rate has ranged from 0.851 (1997) to 1.174 (1981) with the median growth rate of 0.99 occurring in 1973. There appears to be little systematic change in occupancy on the sample of 446 BBS routes. This suggests that the distribution or range of Cerulean warblers is fairly static which is not necessarily inconsistent with previous findings that the “average count” is decreasing throughout most of the Cerulean range (Link and Sauer 2002, Thogmartin et al. 2004). This is apparently not being accompanied by a contraction in apparent range (when range is determined by the collection of BBS routes).

The posterior probability of detecting Cerulean Warblers on a hypothetical “new” route, given that the route is occupied, was computed for each year from  $1 - (1 - \bar{p}_t)^{50}$ , where  $\bar{p}_t$  is obtained by integrating the expression

$$p_t = \frac{\exp\{c_t + w_t\}}{1 + \exp\{c_t + w_t\}}$$

over the posterior distribution of  $c_t$  and also the posterior predictive distribution of  $w_t$ , which is Normal(0,  $\sigma_p^2$ ). The resulting posterior means are depicted in Fig. 3.

We note that the detection probability has declined over time, from about 0.85 in 1966 to 0.55 in 2005, consistent with decreasing average local population sizes under the model described by Royle and Nichols (2003).

It is the ecologists’ never-ending passion to explain such variation as depicted by Figs. 1 and 2 explicitly. Models allowing for spatial heterogeneity can help elucidate metapopulation effects. For example, the mean across years of the 446 site-specific survival probabilities ( $\phi_{it}$ ) are plotted as a function of distance from the center of the range in Fig. 4 (top panel). The corresponding plot of site-specific colonization (omitted) was less

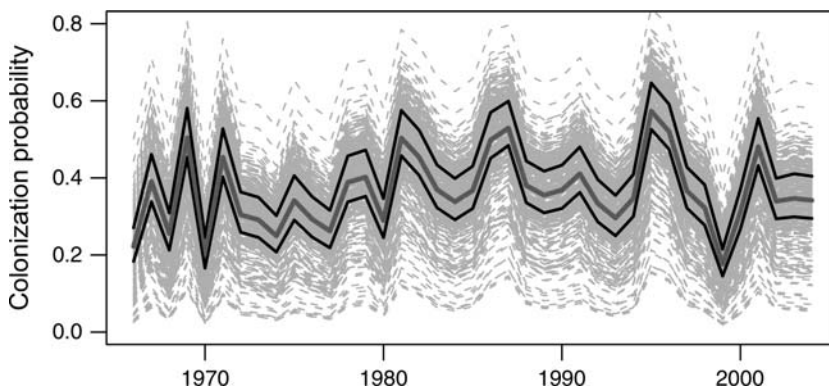


FIG. 2. Variation in patch- or site-colonization probability among sites and across years for the Cerulean Warbler data. Solid lines are the mean (center, thick medium-gray line) and quartiles (black lines). Site-specific time-series are given in light gray.

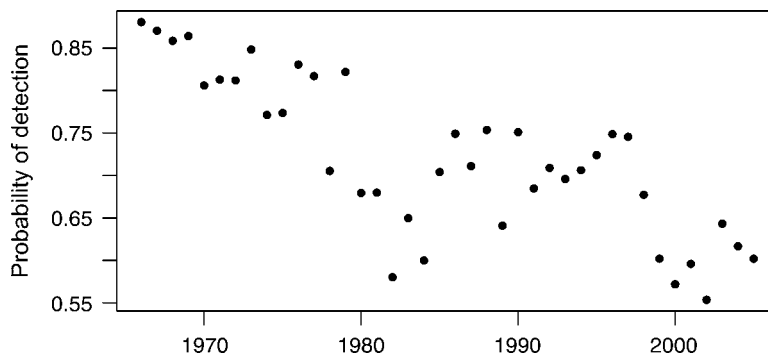


FIG. 3. Estimated probability of detecting a Cerulean Warbler on an occupied route, for each year 1966–2005. Plotted values are  $1 - (1 - \bar{p}_t)^{50}$ , where  $\bar{p}_t$  was obtained by integrating (using the Markov chain Monte Carlo output) the expression  $p_t = (\exp\{c_t + w_t\}) / (1 + \exp\{c_t + w_t\})$  over the posterior distribution of  $c_t$  and also the posterior predictive distribution of  $w_t$ , which is  $\text{Normal}(0, \sigma_p^2)$ .

pronounced. This structure is manifest also in the derived parameters, yielding a corresponding increase in turnover probability (Fig. 4, lower panel) toward the periphery of the range. These observations are consistent with niche centered notions about the determination of ranges. At the range edge conditions may not allow a positive population growth rate (Crozier 2004), and this may also be associated with lower survival and greater turnover of occupied areas (Mehlman 1997). The turnover result is also consistent with that reported by Doherty et al. (2003a) and Karanth et al. (2006) (see *Discussion*).

DISCUSSION

Occupancy and distribution are important state variables in biogeography, metapopulation biology, landscape ecology, and conservation biology. In addition, the components of change in occupancy, i.e., probabilities of colonization and survival, and derived quantities such as rates of growth or turnover are all a primary focus in many such studies. However, a pervasive problem in all such studies is our acute inability to observe occupancy status of patches or spatial sample units free of error. It has been shown

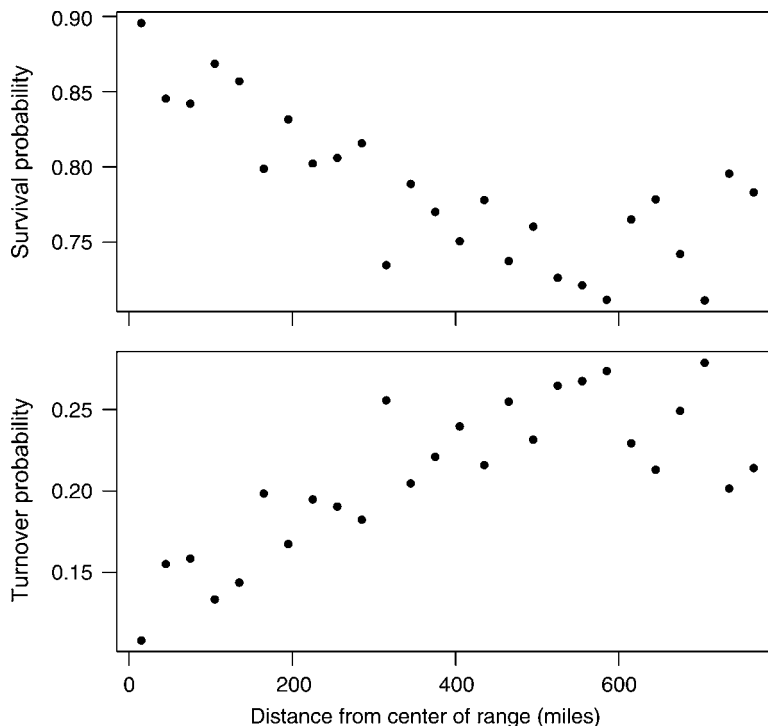


FIG. 4. Mean (across years) site-specific survival (top panel) and turnover (bottom panel) probabilities for Cerulean Warblers as a function of distance from center of range (average location of sample routes). Plotted values are batch means, computed by averaging the site-specific 40-year means in 30-mile intervals (1 mile = 1.6 km).



(Moilanen 2002) that zero detection error of occupancy is the most critical assumption for metapopulation models and is also essential for habitat suitability and other species distribution models (Gu and Swihart 2004). A likelihood-based framework for modeling occupancy dynamics (local extinction, colonization) in the presence of imperfect detection was described by MacKenzie et al. (2003).

In this paper, we propose a natural and more general state-space formulation of dynamic occupancy models. In this formulation, the model is described by two component models: one for the latent (unobserved, or only partially so) occupancy state variable (the “process model”) and another for the observations conditional on the state variable (the “observation model”). Both constituent models have remarkably simple forms, yielding a clear segregation of parameters governing ecological processes of interest from those which are responsible for dealing with nuisance sampling artifacts. The resulting conditional probability structure of the state-space formulation is ideally suited for Bayesian analysis, and implementation of these models is easily achieved in the freely available software package WINBUGS.

The state-space formulation of the model yields several important inferential advantages over the likelihood-based approach described by MacKenzie et al. (2003). First, it yields a generic, flexible and practical framework for modeling individual (“site”) effects, or other latent structure in parameters (e.g., random year effects), in dynamic occupancy models. Similar problems are of some interest in more conventional demographic models (Burnham and Rexstad 1993, Cam et al. 2001, 2004) of animal populations, but flexible and practical solutions to the analysis of such models have proven to be elusive.

Second, although the main object of inference may be the parameters of the process model (i.e., local extinction and colonization probabilities), sometimes summaries of the partial realization of the latent process itself will be of interest. One example is the number of occupied sites from among those that were sampled, a summary that is distinct from the site occupancy parameter, the distinction being that of random variables and realizations of random variables. There can be a large difference in precision between estimators of population parameters, and those of sample occupancy summaries, with sample quantities typically being estimated more precisely. The distinction between the population and finite-sample quantities has generally been disregarded in the literature, both in the development of methods, and also applications where the distinction might be important. The finite-sample estimator would generally be preferred when the number of sample sites is small, or the sample constitutes a large fraction of available sites, or when their selection was non-random, arbitrary, or opportunistic such that they are not representative of the population of potential

sample sites. In this case, generalizing the inference to a large, theoretically infinite, population might be unwarranted, whereas interpreting the sample occupancy as a detection-bias adjusted “index” to metapopulation structure might be adequate for some practical purposes.

Finally, the state-space approach generalizes recent advances in species distribution models. Hierarchical models have recently been described for analysis of species distributions (e.g., for an example with plants see Latimer et al. [2006]), where the true occupancy and detection in response to variations in sampling effort are dealt with at different levels of the hierarchy of a spatial model. These models outperformed more conventional, non-spatial and non-hierarchical models such as GLMs and GAMs in terms of the prediction of observed occurrences. The model framework we describe here generalizes in two dimensions these static occupancy models. First, variation in sampling intensity is modeled explicitly via incorporation of detection probability and hence, not relative but absolute patterns in occurrence are modeled. And second, the dynamic process of colonization and extinction that is at the root of any change in occupancy patterns can be studied directly. In addition, all advantages of hierarchical models (Latimer et al. 2006) such as spatially correlated random effects for the parameters can be incorporated as well.

Analysis of quadrat occupancy dynamics over four years for the Crossbill fully confirmed the erratic population dynamics of that species and enabled to decompose the overall changes in occupancy into contributions of patch survival and patch colonization probability. We note here that the crossbill does not display a classical metapopulation where patches are separated by unsuitable or inferior habitat. Instead, it could perhaps be called a sampling metapopulation, that is, spatially separate samples are drawn from a more or less continuous distribution. Therefore, the scope of inference would be the population of all possible quadrats and in the crossbill example, the population estimates for occupancy and dynamic parameters are more appropriate than the finite sample estimates of these quantities. However, for instance in the case of a classical metapopulation consisting of an amphibian occupying a series of ponds, the desired scope of inference might be exactly this set of ponds. Consequently, use of the finite sample estimators for these quantities would bring great benefits in terms of precision.

In the Cerulean Warbler, we found fairly low turnover in general. That the range extent is rather static contrasts with the earlier finding of a population decline (Link and Sauer 2002, Thogmartin et al. 2004). It may indicate that most of the decline has taken place at high density areas in the interior of the range. This pattern has been found for a majority of declining American bird species (Rodriguez 2002). From a survey design standpoint, it shows that abundance is more sensitive to population changes than occupancy. In addition, there was greater

instability in site occupancy, i.e., declining survival probabilities and increasing turnover, toward the range edge. This result is consistent with findings by Doherty et al. (2003a) and Karanth et al. (2006), who both used methods that account for imperfect species detection. For example, Doherty et al. (2003a) found that, for groups of species classified as being interior or edge-of-range, local extinction probabilities (among species) were higher at the edges than the interiors. They also noted that such differences are artificially magnified if imperfect detection is not taken into consideration in the estimation. The distinction of our study, however, is that we demonstrate this result “organically,” i.e., by fitting a model to the particular species in question. Both Doherty et al. (2003a) and Karanth et al. (2006) generate replication by defining a pool of species that are all assumed to have the same survival and colonization probabilities. More specifically, the community dynamics modeling problem is represented as a conventional *population* capture-recapture problem, wherein species identity takes the role of an individual animal. The model used by them for estimation allows for heterogeneity in detection probability among species (so-called “model  $M_h$ ”). Our result also concurs with a similarly greater variability in abundance at the edge of ranges (Williams et al. 2003). Altogether, these patterns provide support for a metapopulation view of range limits (Lennon et al. 1997, Holt and Keitt 2000).

The state-space formulation of dynamic occupancy models yields a flexible and practical framework for modeling spatial and temporal variation in occupancy dynamics. Analysis of these models is easily implemented in the freely available software WinBUGS. We believe that where multi-year detection/nondetection data are available, dynamic occupancy models as described here have an enormous potential for the study of distributions and range dynamics of animals, plants and diseases.

#### ACKNOWLEDGMENTS

We thank J. D. Nichols for helpful discussion relating to the concept of turnover in the context of models for occupancy. We also thank Hans Schmid for providing the crossbill data.

#### LITERATURE CITED

- Bayley, P. B., and J. T. Peterson. 2001. An approach to estimate probability of presence and richness of fish species. *Transactions of the American Fisheries Society* 130:620–633.
- Berliner, L. M. 1996. Hierarchical Bayesian time series models. Pages 15–22 in K. M. Hanson and R. N. Silver, editors. *Maximum entropy and Bayesian methods*. Kluwer, Dordrecht, The Netherlands.
- Bibby, C. J., N. D. Burgess, and D. Hill. 1992. *Bird census techniques*. Academic Press, London, UK.
- Boulinier, T., J. D. Nichols, J. E. Hines, J. R. Sauer, C. Flather, and K. H. Pollock. 1998a. Higher temporal variability of forest bird communities in fragmented landscapes. *Proceedings of the National Academy of Sciences (USA)* 95:7497–7501.
- Boulinier, T., J. D. Nichols, J. E. Hines, J. R. Sauer, C. Flather, and K. H. Pollock. 2001. Forest fragmentation and bird community dynamics: inference at regional scales. *Ecology* 82:1159–1169.
- Boulinier, T., J. D. Nichols, J. R. Sauer, J. E. Hines, and K. H. Pollock. 1998b. Estimating species richness: the importance of heterogeneity in species detectability. *Ecology* 79:1018–1028.
- Brander, S. M., J. A. Royle, and M. Eames. 2007. Evaluation of the status of anuran metapopulations on a refuge in suburban Maryland. *Journal of Herpetology* 41:51–59.
- Burnham, K. P., and E. A. Rexstad. 1993. Modeling heterogeneity in survival rates of banded waterfowl. *Biometrics* 49:1194–1208.
- Cam, E., W. A. Link, E. G. Cooch, J.-Y. Monnat, and E. Danchin. 2001. Individual covariation in life-history traits: seeing the trees despite the forest. *American Naturalist* 159: 96–105.
- Cam, E., J.-Y. Monnat, and J. A. Royle. 2004. Dispersal and individual quality in a long lived species. *Oikos* 106:386–398.
- Crozier, L. 2004. Warmer winters drive butterfly range expansion by increasing survivorship. *Ecology* 85:231–241.
- Doherty, P. F., T. Boulinier, and J. D. Nichols. 2003a. Local extinction and turnover rates at the edge and interior of species ranges. *Annales Zoologici Fennici* 40:145–153.
- Doherty, P. F., G. Sorci, J. A. Royle, J. E. Hines, J. D. Nichols, and T. Boulinier. 2003b. Sexual selection affects local extinction and turnover. *Proceedings of the National Academy of Science (USA)* 100:5858–5862.
- Dorazio, R. M., and J. A. Royle. 2003. Mixture models for estimating the size of a closed population when capture rates vary among individuals. *Biometrics* 59:351–364.
- Dorazio, R. M., and J. A. Royle. 2005. Estimating size and composition of biological communities by modeling the occurrence of species. *Journal of the American Statistical Association* 100:389–398.
- Geman, S., and D. Geman. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6: 721–741.
- Gilks, W. R., A. Thomas, and D. J. Spiegelhalter. 1994. A language and program for complex Bayesian modelling. *Statistician* 43:169–178.
- González-Megías, A., J. M. Gómez, and F. Sánchez-Piñero. 2005. Consequences of spatial autocorrelation for the analysis of metapopulation dynamics. *Ecology* 86:3264–3271.
- Gu, W., and R. K. Swihart. 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.
- He, F., and K. J. Gaston. 2000a. Occupancy–abundance relationships and sampling scales. *Ecography* 23:503–511.
- He, F., and K. J. Gaston. 2000b. Estimating species abundance from occurrence. *American Naturalist* 156:553–559.
- Holt, R. D., and T. H. Keitt. 2000. Alternative causes for range limits: a metapopulation perspective. *Ecology Letters* 3:41–47.
- Ihaka, R., and R. Gentleman. 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5:299–314.
- Johnson, D. M. 2005. Metapopulation models: an empirical test of model assumptions and evaluation methods. *Ecology* 86:3088–3098.
- Jones, R. H. 1993. *Longitudinal data with serial correlation: a state-space approach*. Chapman and Hall, London, UK.
- Karanth, K. K., J. D. Nichols, J. R. Sauer, and J. E. Hines. 2006. Comparative dynamics of avian communities across edges and interiors of North American ecoregions. *Journal of Biogeography* 33:674–682.

- Kendall, W. L., K. H. Pollock, and C. Brownie. 1995. A likelihood-based approach to capture-recapture estimation of demographic parameters under the robust design. *Biometrics* 51:293–308.
- Kéry, M. 2002. Inferring the absence of a species: a case study of snakes. *Journal of Wildlife Management* 66:330–338.
- Kéry, M., J. A. Royle, and H. Schmid. 2005. Modeling avian abundance from replicated counts using binomial mixture models. *Ecological Applications* 15:1450–1461.
- Kéry, M., and H. Schmid. 2006. Estimating species richness: calibrating a large avian monitoring program. *Journal of Applied Ecology* 43:101–110.
- Laird, N. M., and J. H. Ware. 1982. Random-effects models for longitudinal data. *Biometrics* 38:963–974.
- Latimer, A. M., S. Wu, A. E. Gelfand, and J. A. Silander, Jr. 2006. Building statistical models to analyze species distributions. *Ecological Applications* 16:33–50.
- Lennon, J. J., J. R. G. Turner, and D. Connell. 1997. A metapopulation model of species boundaries. *Oikos* 78:486–502.
- Link, W. A., and J. R. Sauer. 2002. A hierarchical analysis of population change with application to Cerulean Warblers. *Ecology* 83:2832–2840.
- MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. D. 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 1. *Ecology* 83:2248–2255.
- MacKenzie, D., J. D. Nichols, J. A. Royle, K. H. Pollock, L. A. Bailey, and J. E. Hines. 2005. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier, Burlington, Massachusetts, USA.
- Mehlman, D. W. 1997. Change in avian abundance across the geographic range in response to environmental change. *Ecological Applications* 7:614–624.
- Moilanen, A. 2002. Implications of empirical data quality to metapopulation model parameter estimation and application. *Oikos* 96:516–530.
- Nichols, J. D., T. Boulinier, J. E. Hines, K. H. Pollock, and J. R. Sauer. 1998a. Estimating rates of species extinction, colonization and turnover in animal communities. *Ecological Applications* 8:1213–1225.
- Nichols, J. D., T. Boulinier, J. E. Hines, K. H. Pollock, and J. R. Sauer. 1998b. Inference methods for spatial variation in species richness and community composition when not all species are detected. *Conservation Biology* 12:1390–1398.
- Nichols, J. D., and K. U. Karanth. 2002. Statistical concepts; assessing spatial distribution. Pages 29–28 in K. U. Karanth and J. D. Nichols, editors. *Monitoring tigers and their prey*. Centre for Wildlife Studies, Bangalore, India.
- Pellet, J., and B. R. Schmidt. 2005. Monitoring distributions using call surveys: estimating site occupancy, detection probabilities and inferring absence. *Biological Conservation* 123:27–35.
- Pollock, K. H. 1982. A capture-recapture design robust to unequal probability of capture. *Journal of Wildlife Management* 46:757–760.
- R Development Core Team. 2005. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Robinson, G. K. 1991. That BLUP is a good thing: the estimation of random effects (with discussion). *Statistical Science* 6:15–51.
- Rodriguez, J. P. 2002. Range contractions in declining North American bird populations. *Ecological Applications* 12:238–248.
- Royle, J. A. 2006. Site occupancy models with heterogeneous detection probabilities. *Biometrics* 62:97–102.
- Royle, J. A., and W. A. Link. 2005. A general class of multinomial mixture models for anuran calling survey data. *Ecology* 86:2505–2512.
- Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence-absence data or point counts. *Ecology* 84:777–790.
- Schmid, H., N. Zbinden, and V. Keller. 2004. Überwachung der Bestandsentwicklung häufiger Brutvögel in der Schweiz. Swiss Ornithological Institute, Sempach, Switzerland.
- Schmidt, B. R. 2005. Monitoring the distribution of pond-breeding amphibians when species are detected imperfectly. *Aquatic Conservation: Marine and Freshwater Ecosystems* 15:681–692.
- Thogmartin, W. E., J. R. Sauer, and M. G. Knutson. 2004. A hierarchical spatial model of avian abundance with application to Cerulean Warblers. *Ecological Applications* 14:1766–1779.
- Tyre, A. J., B. Tenhumberg, S. A. Field, D. Niejalke, K. Parris, and H. P. Possingham. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. *Ecological Applications* 13:1790–1801.
- Weir, L. A., J. A. Royle, P. Nanjappa, and R. E. Jung. 2005. Modeling anuran detection and site occupancy on North American Amphibian Monitoring Program (NAAMP) routes in Maryland. *Journal of Herpetology* 39:627–639.
- Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, London, UK.
- Williams, C. K., A. R. Ives, and R. D. Applegate. 2003. Population dynamics across geographical ranges: time-series analyses of three small game species. *Ecology* 84:2654–2667.
- Wintle, B. A., M. A. McCarthy, K. M. Parris, and M. A. Burgman. 2003. Precision and bias of methods for estimating point survey detection probabilities. *Ecological Applications* 14:703–712.

#### APPENDIX

Development of an MCMC algorithm by Gibbs sampling for a simple model (*Ecological Archives* E088-108-A1).

#### SUPPLEMENT

WinBUGS model specification for European Crossbill and Cerulean Warbler examples in the paper (*Ecological Archives* E088-108-S1).