

Hierarchical Models of Animal Abundance and Occurrence

J. Andrew ROYLE and Robert M. DORAZIO

Much of animal ecology is devoted to studies of abundance and occurrence of species, based on surveys of spatially referenced sample units. These surveys frequently yield sparse counts that are contaminated by imperfect detection, making direct inference about abundance or occurrence based on observational data infeasible. This article describes a flexible hierarchical modeling framework for estimation and inference about animal abundance and occurrence from survey data that are subject to imperfect detection. Within this framework, we specify models of abundance and detectability of animals at the level of the local populations defined by the sample units. Information at the level of the local population is aggregated by specifying models that describe variation in abundance and detection among sites. We describe likelihood-based and Bayesian methods for estimation and inference under the resulting hierarchical model. We provide two examples of the application of hierarchical models to animal survey data, the first based on removal counts of stream fish and the second based on avian quadrat counts. For both examples, we provide a Bayesian analysis of the models using the software WinBUGS.

Key Words: Bayesian analysis; Capture-recapture; Empirical Bayes; Multiple observer sampling; Occurrence; Point-count sampling; Removal sampling; Site occupancy.

1. INTRODUCTION

Investigations of animal populations often require estimates of animal abundance or occurrence. These estimates may be used to monitor spatial or temporal changes in the population or to determine whether relationships exist between animals and one or more environmental characteristics (i.e., to identify habitat). A common difficulty in these studies is that animals which are available to be sampled often go undetected; thus, the number or count of animals observed in a survey provides a negatively biased index of abundance. For some inference problems, the bias is ignorable if it is reasonable to assume that detection or capture rate is identical at all sample locations and all sample times, or if covariates

J. Andrew Royle is Research Statistician, U.S. Geological Survey Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, MD 20708 (E-mail: aroyle@usgs.gov). Robert M. Dorazio is Research Statistician, U.S. Geological Survey, Florida Integrated Science Center, Department of Statistics, University of Florida, P.O. Box 110339, Gainesville, FL 32611 (E-mail: bdorazio@usgs.gov).

In the Public Domain

Journal of Agricultural, Biological, and Environmental Statistics, Volume 11, Number 3, Pages 249–263

DOI: 10.1198/108571106X129153

responsible for variation in detectability can be identified and their effects modeled indirectly (e.g., Link and Sauer 2002); however, such is often not the case, even when “standardized” methods of data collection are employed. Moreover, explicit consideration of detectability is essential when occurrence-based summaries of population status are the focus of inference. In such cases, putative absences (i.e., observed counts of zero) are ambiguous and this ambiguity can only be resolved by explicit consideration of detectability. In actual surveys, rates of detection can vary among collection times, collection locations, and even among individual animals (say, owing to differences in animal behavior or coloration).

To account for the various sources of heterogeneity in detection, several sampling protocols and statistical models have been devised to estimate the abundance of animals *and* their probabilities of detection simultaneously. These protocols include distance sampling, sampling with multiple observers, capture-recapture, and removal sampling (Williams, Nichols, and Conroy 2002). The class of statistical models developed for each of these protocols represents, in some sense, a mechanism for “transforming” the observed animal counts into an index of abundance that is unbiased with respect to variation in detectability.

These sampling protocols and models have proven to be enormously useful in surveys of animal populations. However, in some situations the conventional methods of data collection and analysis can be difficult to apply. Consider, for example, surveys in which a particular sampling protocol is used at a large number of relatively small, but spatially distinct, sample units. For example, in many surveys of bird populations “point counts” are conducted at a set of randomly selected locations within a park, refuge, or forest. In these types of surveys the size of the population that is available to be detected within each sample unit may be small (even zero); consequently, the observed counts are generally sparse and yield little information about “local abundance” (i.e., at the scale of the sample unit) of the species under study. A common solution to this problem is to pool observations among sample units and to develop elaborate models of heterogeneity in detection that ignore the spatial indexing of the sample units. Thus, focus shifts from modeling variation in abundance of animals to modeling variation in their detectability, and the estimator of total abundance (i.e., among all sample units) essentially amounts to an adjustment of the observed aggregate count that conditions on estimated nuisance parameters of the detection model.

In this article we develop an alternative solution to the problem by augmenting the model of detectability with a model that specifies variation in local abundance of animals. We adopt a hierarchical approach wherein prior distributions for abundance or detection are used to specify stochastic dependence among sample units. Our approach also allows spatial covariates that are thought to be informative of detection or abundance to be included; thus, systematic sources of variability (e.g., habitat features) may be explored by specifying models with or without such covariates. A hierarchical modeling framework also allows abundance to be estimated or predicted for individual locations or groups of locations (regions). These estimates are often used to assess the consequences of location-specific management actions. Another conceptual benefit of the hierarchical framework that we propose is its unified treatment of inference problems for both animal abundance and occurrence. Inferences about these two quantities are often viewed independently, leading

some investigators to estimate occurrence using models of counts reduced to binary presence/absence data [e.g., logistic regression or site-occupancy, see MacKenzie et al. (2002, 2005); Manly et al. (2002); Sargeant, Sovada, Slivinski, and Johnson (2005)]. In contrast, in the modeling framework that we develop, estimators of occurrence are derived from the fundamental equivalence between occurrence and the event that at least one animal is present (i.e., abundance $N > 0$).

Section 2 describes various sampling protocols and observation models that are widely used in surveys of animal populations. Section 3 develops the hierarchical framework, and Section 4 describes methods of estimation (frequentist and Bayesian). Section 5 illustrates their application in two animal sampling problems. In the first example, removal sampling was used to sample a population of stream fishes at 28 locations. At each location fishes were removed in three successive passes. In this example, we require a model for estimating the local abundance of fishes while accounting for sources of spatial variation in mean abundance. One important feature of these data is the substantial heterogeneity in abundance (the number of fish removed from each sample location ranged from 1 to 177). We develop a hierarchical model in which mean abundance varies systematically among six distinct drainages and stochastically within each drainage using an exchangeable prior for site-specific random effects. In the second example, we consider quadrat counts of birds from an avian monitoring program in Switzerland. More than 200 1 km^2 quadrats were surveyed three times yielding counts of territories for each quadrat. The observation model is thus a product-binomial. We develop a hierarchical abundance model describing variation in abundance as a function of landscape structure (elevation and forest coverage). We also account for variation in sampling intensity (sample route length) among quadrats. Section 6 concludes with a discussion of advantages and shortcomings of the hierarchical approach and suggests various modeling extensions.

2. SAMPLING PROTOCOLS AND OBSERVATION MODELS

Let N_i denote the unknown number of animals that are available to be detected or captured within sample unit i , and let \mathbf{x}_i denote the vector of counts observed when a particular sampling protocol is used to detect the N_i animals. The sampling protocol determines the precise structure of the observation vector \mathbf{x}_i and corresponding model.

Many common sampling protocols induce a dependence among the elements of \mathbf{x}_i that may be specified with a multinomial distribution. For example, a capture-recapture protocol conducted on T sample occasions potentially yields $2^T - 1$ observable encounter histories, often summarized by a vector of ones and zeros to indicate the occasions where animals were captured (i.e., detected) or not captured, respectively. As an example, $T = 2$ implies three observable encounter histories: $(1, 1)$ for animals captured on both occasions, $(1, 0)$ for animals captured on the first occasion but not the second, and $(0, 1)$ for animals captured on the second occasion but not the first. In this example each element of \mathbf{x}_i corresponds to the number of animals associated with a distinct encounter history. By conditioning on N_i ,

a multinomial distribution may be used to model the observed vector of counts as follows:

$$f(\mathbf{x}_i | N_i, \boldsymbol{\pi}_i) = \frac{N_i!}{(N_i - x_{i\cdot})! \prod_j x_{ij}!} \left(\prod_j \pi_{ij}^{x_{ij}} \right) \pi_{i0}^{N_i - x_{i\cdot}}, \quad (2.1)$$

where $x_{i\cdot} = \sum_j x_{ij}$ is the total number of animals detected at sample unit i and π_{ij} denotes the probability that animals present within sample unit i experience encounter history j . Note that $j = 0$ corresponds to the unobservable encounter history of the $N_i - x_{i\cdot}$ animals not detected within sample unit i . Each multinomial cell probability π_{ij} is formulated as a function of parameters θ_i that often include site-specific detection probabilities. For example, in capture-recapture surveys with $T = 2$, animals with encounter history $(1, 0)$ may be modeled using $\pi_{ij} = \theta_i(1 - \theta_i)$.

Analogous data-model conjunctions arise under other sampling protocols. For example, if sampling is conducted by T observers such that the identity of each animal is recorded independently, then again there are $2^T - 1$ observable encounter histories. Another common sampling protocol, removal sampling, yields the number of animals removed from the population of N_i animals on T successive occasions. Here the form of the multinomial cell probability is $\pi_{ij} = \theta_i \prod_{k=1}^{j-1} (1 - \theta_i)$. Thus, while the exact form of the multinomial cell probabilities $\boldsymbol{\pi}_i$ generally depends on the particular sampling protocol employed, the multinomial distribution in (2.1) is suitable for modeling dependence among counts observed within a sample unit.

Perhaps the most common sampling protocol employed in avian surveys is based on simple point counts. This protocol yields a univariate count statistic that is usually modeled as a realization from a binomial distribution parameterized by index N_i and detection probability π_i . The utility of such counts is a matter of some debate owing to the inherent confounding between variation in detectability and variation in abundance (although the nature of this depends on consideration of a model). However, a minor extension of this protocol can provide information on both π_i and N_i . Consider a sampling protocol wherein each local population (i.e., each sample unit) is visited on T separate occasions, yielding a sequence of independent binomial counts x_{ij} ($j = 1, \dots, T$). In this case, the likelihood function is a product-binomial distribution

$$f(\mathbf{x}_i | N_i, \boldsymbol{\pi}_i) = \prod_j \frac{N_i!}{(N_i - x_{ij})! x_{ij}!} \pi_i^{x_{ij}} (1 - \pi_i)^{N_i - x_{ij}}. \quad (2.2)$$

The instability (acute sensitivity to minor perturbations of the data) of (2.2) is well-known (Olkin, Petkau, and Zidek 1981; Carroll and Lombard 1985; Raftery 1988), and this deficiency may explain why the sampling protocol is not widely used in practice. However, when this sampling protocol is applied at replicate locations and additional model structure is imposed on the location-specific abundance parameters (see Section 3.1), the estimation problem becomes more tractable (Royle 2004).

3. HIERARCHICAL MODELS

In surveys where the local abundance of animals within a sample unit is relatively small, the observed counts \mathbf{x}_i are almost certain to be low, regardless of sampling protocol. In this situation inferences about N_i based on using (2.1) or (2.2) as a likelihood function are therefore unlikely to be satisfactory. Our proposed solution to this problem is to formulate a model that sensibly links the abundance and detection parameters of different sample units, thereby producing an explicit model-based aggregation of the data and a description of their spatial heterogeneity.

We consider surveys wherein a particular sampling protocol is applied at each of n locations (sample units) that constitute a representative sample of the entire population of animals. We assume that these locations and the method and duration of sampling can be selected to ensure that the local abundance of animals within each sample unit is demographically closed during the sampling activity. This does not preclude the interaction among location populations through time. Our model of the population therefore corresponds to sampling a metapopulation of animals (Hanski and Simberloff 1997), a collection of spatially distinct subpopulations (local populations) that occasionally intermix. In the following subsections we develop a hierarchical framework that connects the local population parameters to those of the metapopulation.

3.1 MODELS OF ABUNDANCE

The Poisson distribution is a natural choice for modeling local abundance; therefore, we assume $[N_i | \lambda_i] \sim \text{Poisson}(\lambda_i)$, where λ_i denotes the mean abundance of animals within the i th sample unit. In addition, we use a Poisson-regression formulation, $\log \lambda_i = \mathbf{w}_i' \boldsymbol{\alpha}$, to model systematic variation in abundance among sample locations as a function of covariates \mathbf{w}_i and metapopulation-level parameters $\boldsymbol{\alpha}$. Here, \mathbf{w}_i is an $m \times 1$ vector of covariates for sample unit i and $\boldsymbol{\alpha}$ is the corresponding $m \times 1$ vector of coefficients. In the absence of such covariates, a model with constant mean abundance $\lambda_i = \exp(\alpha)$ may be specified by assigning $w_i = 1$ ($\forall i$).

Stochastic sources of variation in the mean abundance (which imply extra-Poisson variation in local abundance) also may be specified by way of additional distributional assumptions. In particular, a multivariate normal prior, $\boldsymbol{\phi} \sim \text{Normal}(\mathbf{W}\boldsymbol{\alpha} \boldsymbol{\Sigma}_\phi)$, where $\phi_i = \log \lambda_i$ and $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n)'$, allows great flexibility in specifying various forms of spatial variation. For example, an exchangeable prior may be considered by assuming $\boldsymbol{\Sigma}_\phi = \sigma_\phi^2 \mathbf{I}$ (where \mathbf{I} denotes an $n \times n$ identity matrix) and thus estimating only 1 additional parameter σ_ϕ^2 . Alternatively, we may assume spatial dependence among the mean abundance parameters by specifying inter-site correlations in abundance as a function of the distance between sites and a few model parameters, for example, $\boldsymbol{\Sigma}_\phi = \sigma_\phi^2 \mathbf{R}_\phi$. Such models are routinely applied in geostatistics (Diggle, Tawn, and Moyeed 1998). Alternative models of extra-Poisson variation, such as the negative binomial, may seem more natural. However, the Poisson prior can yield computational savings (see Section 4.1) because the N_i parameters may be removed from the likelihood analytically to express it directly in terms of the Poisson mean parameters λ_i .

3.2 MODELS OF DETECTION PROBABILITY

Spatial variation in detection probability may be specified similar to the way that spatial variation in abundance was modeled. In particular we use a logistic-regression formulation, $\text{logit } \theta_i = \mathbf{v}_i' \boldsymbol{\beta}$, to model systematic variation in detection among sample locations as a function of covariates \mathbf{v}_i and metapopulation-level parameters $\boldsymbol{\beta}$. Specification of stochastic variation in detectability requires additional distributional assumptions. In particular, we use a multivariate normal prior, $\boldsymbol{\eta} \sim \text{Normal}(\mathbf{V}\boldsymbol{\beta}, \boldsymbol{\Sigma}_\eta)$, where $\eta_i = \text{logit } \theta_i$ and $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)'$. Thus, whereas the multivariate mean specifies systematic variation in detection probability among sample locations, $\boldsymbol{\Sigma}_\eta$ parameterizes unobserved sources of heterogeneity in detection.

4. METHODS OF ESTIMATION

4.1 MAXIMUM LIKELIHOOD

The hierarchical framework developed in Section 3 includes a wide variety of models. Some of these models, particularly those with exchangeable priors, may be fitted to count data by maximizing the marginal likelihood obtained by integrating away the parameters associated with local abundance or detection. For example, let $[\mathbf{x}_i | N_i, \theta_i]$ represent either a multinomial or product-binomial distribution (described in Section 2) that conditions on parameters N_i and θ_i . Now suppose that N_i is assumed to be adequately modeled by $[N_i | \boldsymbol{\alpha}] \sim \text{Poisson}(\exp(\mathbf{w}_i' \boldsymbol{\alpha}))$. In other words, there is no stochastic variation in mean abundance among sample units ($\sigma_\phi^2 = 0$). Then the marginal distribution of the observed counts is obtained by eliminating (through summation) the local abundance parameter N_i as follows:

$$[\mathbf{x}_i | \boldsymbol{\alpha}, \theta_i] = \sum_{N_i=y_i}^{\infty} [\mathbf{x}_i | N_i, \theta_i] [N_i | \boldsymbol{\alpha}], \quad (4.1)$$

where $y_i = x_i$ if $[\mathbf{x}_i | N_i, \theta_i]$ is multinomial; otherwise, $y_i = \max(\mathbf{x}_i)$. In practice, a suitably large integer may be used as the upper limit of summation; however, if $[\mathbf{x}_i | N_i, \theta_i]$ is multinomial, the summation in (4.1) is equivalent to the product of a finite number of Poisson probabilities, which can provide substantial computational savings. This can be demonstrated analytically by carrying out the summation in Equation (4.1). In either case, maximum-likelihood estimates (MLEs) of the model parameters may be computed by determining those values $\hat{\boldsymbol{\alpha}}$ and $\{\hat{\theta}_i\}$ that maximize the product of marginal probability densities among all n sample units: $L(\boldsymbol{\alpha}, \{\theta_i\} | \{\mathbf{x}_i\}) = \prod_{i=1}^n [\mathbf{x}_i | \boldsymbol{\alpha}, \theta_i]$. This approach has been successfully used in analyses of simple point counts (Dodd and Dorazio 2004; Royle 2004) and counts observed in removal samples (Dorazio, Jelks, and Jordan 2005), distance samples (Royle, Dawson, and Bates 2004), spatially referenced capture-recapture samples (Royle, Kéry, Gautier, and Schmid 2006), and simple presence/absence data (Royle and Nichols 2003).

Once the MLEs have been computed, Bayes’s theorem may be used to estimate the conditional (on $\hat{\alpha}$ and $\hat{\theta}_i$) posterior probability of local abundance

$$[N_i | \mathbf{x}_i, \hat{\alpha}, \hat{\theta}_i] = \frac{[N_i | N_i, \hat{\theta}_i] [N_i | \hat{\alpha}]}{[\mathbf{x}_i | \hat{\alpha}, \hat{\theta}_i]} \tag{4.2}$$

which is useful in estimating the conditional mean abundance of animals within the i th sample unit or in computing inferences for particular values of N_i or functions of N_i . For example, the probability that at least one animal is present within the i th sample unit (i.e., the probability of occurrence), may be estimated by $1 - [N_i = 0 | \mathbf{x}_i, \hat{\alpha}, \hat{\theta}_i]$. Also, the metapopulation average of this quantity, referred to as “site occupancy,” is the probability of occurrence averaged over replicate sites. Note that the mean of Equation (4.2) is the (estimated) best unbiased predictor typically used as an estimator of random effects in classical statistics. For the case where $[N_i | \alpha]$ is Poisson and \mathbf{x}_i is multinomial, $[N_i | \mathbf{x}_i, \alpha, \theta_i] = x_{i.} + \text{Poisson}(\{1 - \pi_{i.}(\theta_i)\} \lambda_i(\alpha))$. Thus, a point estimate of N_i may be obtained as, for example, the mean of this distribution evaluated at the MLEs of α and θ_i . Measures of uncertainty may also be obtained by conventional means (e.g., as suitable functions of estimated variances of the MLEs).

A deficiency of these empirical Bayes estimators of animal abundance or animal occurrence is that by conditioning on the MLEs the estimators fail to account for uncertainty in the MLEs. Laird and Louis (1987) developed a parametric bootstrapping procedure to correct for this source of uncertainty, but the procedure requires additional computation. An alternative, as described in Section 4.2, is to conduct a fully Bayesian analysis wherein inferences are based on the marginal posterior distribution $[N_i | \mathbf{x}_i]$ that integrates over the uncertainty in model parameters.

Maximum-likelihood also may be used to fit models with exchangeable priors on detection probability. For example, let $[\theta_i | \beta, \sigma_\eta^2]$ denote a conditional distribution that specifies exchangeable spatial heterogeneity in θ_i . Adding this distributional assumption to the previous example, we may estimate the model parameters $(\alpha, \beta, \sigma_\eta^2)$ by computing a marginal likelihood that removes both N_i and θ_i by integration:

$$[\mathbf{x}_i | \alpha, \beta, \sigma_\eta^2] = \sum_{N_i=y_i}^{\infty} \left[\int_0^1 [\mathbf{x}_i | N_i, \theta_i] [\theta_i | \beta, \sigma_\eta^2] d\theta_i \right] [N_i | \alpha]. \tag{4.3}$$

Although particular choices of the distribution $[\theta_i | \beta, \sigma_\eta^2]$ allow analytical evaluation of the term in brackets, an accurate numerical approximation of the inner integral can be obtained at only minor computational expense (see Dorazio et al. 2005 for examples).

4.2 BAYESIAN ANALYSIS

In the previous section we showed that maximum-likelihood estimation can be used to fit many models. Here we develop more general estimators of animal abundance that are particularly useful in fitting models with spatial correlations in animal abundance or detection. Let $[\mathbf{x}_i | N_i, \eta_i]$ represent either a multinomial or product-binomial distribution

(described in Section 2) that conditions on the abundance N_i and the logit-scale detection probability η_i within the i th sample unit. As described in Section 3.1, the local abundance of animals is modeled by conditioning on a log-scale Poisson mean parameter ϕ_i : $[N_i | \phi_i] \sim \text{Poisson}(\exp(\phi_i))$. In addition, let $[\phi | \mathbf{W}\alpha, \Sigma_\phi]$, and $[\eta | \mathbf{V}\beta, \Sigma_\eta]$ denote the multivariate normal distributions used to model spatial heterogeneity in mean abundance and detection, respectively.

The joint posterior distribution of the model's parameters is proportional to the product of our conditional distributions and a set of mutually independent prior distributions for the metapopulation-level parameters $(\alpha, \beta, \Sigma_\phi, \Sigma_\eta)$:

$$[\mathbf{N}, \phi, \eta, \alpha, \beta, \Sigma_\phi, \Sigma_\eta | \mathbf{x}] \propto \left(\prod_{i=1}^n [\mathbf{x}_i | N_i, \eta_i] [N_i | \phi_i] \right) [\phi | \mathbf{W}\alpha, \Sigma_\phi] [\eta | \mathbf{V}\beta, \Sigma_\eta] [\alpha] [\beta] [\Sigma_\phi] [\Sigma_\eta],$$

where $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)'$ denotes the entire matrix of observed counts. Generally speaking, stochastic simulation methods, such as Markov chain Monte Carlo (MCMC) sampling (Gilks, Richardson, and Spiegelhalter 1996), are required to approximate the joint posterior distribution. It is then straightforward to compute inferences about animal abundance or animal occurrence from the simulated sample of the joint posterior. We illustrate these calculations in the following section, where models of various types of count data are fitted using the BUGS language (Gilks, Thomas, and Spiegelhalter 1994) and WinBUGS software (<http://www.mrc-bsu.cam.ac.uk/bugs/>).

5. EXAMPLES

5.1 REMOVAL SAMPLING OF STREAM FISHES

In 2004 a species of stream fishes, the Okaloosa darter (*Etheostoma okaloosae*), was surveyed by removal sampling. Samples were collected at 28 locations considered to be representative of the darter's native range. At each location fishes were removed in three successive passes of a 20-meter stream reach. Only a few hours were required to collect the fish from each location. Given the short duration of sampling and the strong affinity of this species for its habitat, we may reasonably assume that each local population of fishes was demographically closed during the survey.

Estimates of local abundance are highly relevant in this population of fishes because management actions are limited to particular regions of the darter's habitat (as opposed to its entire range). We therefore require a model for estimating the local abundance of fishes while accounting for probable sources of spatial variation in mean abundance. The number of fish removed from each sample location ranged from 1 to 177 darters, suggesting substantial spatial heterogeneity in abundance. Using the hierarchical framework developed in Section 3, the mean abundance of fishes within each sample unit was assumed to differ systematically among six distinct drainages (specified by a covariate matrix \mathbf{W}) and stochastically within each drainage using an exchangeable prior (i.e., $\Sigma_\phi = \sigma_\phi^2 \mathbf{I}$).

```

model {
  # Specify priors for metapopulation-level parameters
  for (k in 1:p) {
    alpha[k] ~ dnorm(0.0, 1.0E-4)
    lambda.mean[k] <- exp(alpha[k])
  }
  sigma.phi ~ dunif(0, 100)
  tau.phi <- 1/(sigma.phi*sigma.phi)
  theta ~ dunif(0,1)
  beta <- log(theta/(1-theta))

  # Specify observation model and priors for local parameters
  for (i in 1:n) {
    x[i,1] ~ dbin(theta, N[i])
    z[i,1] <- N[i] - x[i,1]
    for(j in 2:m) {
      x[i,j] ~ dbin(theta, z[i,j-1])
      z[i,j] <- N[i] - sum(x[i,1:j])
    }
    N[i] ~ dpois(lambda[i])
    log(lambda[i]) <- phi[i]
    mu.phi[i] <- inprod(w[i,], alpha[])
    phi[i] ~ dnorm(mu.phi[i], tau.phi)
  }
}

```

Figure 1. WinBUGS code for fitting the hierarchical model of removal counts of stream fishes

The WinBUGS specification of this model appears in Figure 1. Summaries of the posterior distribution were calculated from four independent Markov chains initialized with random starting values, run 50,000 iterations after a 10,000 burn-in and thinning every five draws. This yielded four sets of 10,000 posterior draws. We computed the Brooks-Gelman-Rubin convergence diagnostic (Gelman and Rubin 1992; Brooks and Gelman 1998) using the output from the four chains. Values of this statistic, referred to as the scale reduction factor, near 1.0 indicate convergence. For these data, the multivariate potential scale reduction factor was 1.000247, and for each parameter the potential scale reduction factors differed from 1.0 by less than 10^{-4} .

Inferences based on the MCMC sample of the joint posterior suggest that while the (posterior) log-mean abundance of fishes differed considerably among the six drainages ($\hat{\alpha} = (2.7, 3.0, 3.4, 4.0, 4.2, 4.3)'$), some within-drainage variation also existed ($\hat{\sigma}_\phi = 0.96$; 95% credible interval = (0.67, 1.40)). We estimated the local abundance (and 95% credible intervals) for each of the 28 sample locations (Figure 2) and compared these estimates with MLEs obtained by fitting a multinomial likelihood (in (2.1)) separately to each location's counts.

5.2 REPEATED POINT COUNTS OF BIRDS

Here we consider data from the Swiss monitoring program for common breeding birds. This survey is conducted annually by the Swiss Ornithological Institute (Schmid et al. 2001). More than 200 1 km^2 quadrats across Switzerland are sampled during the breeding

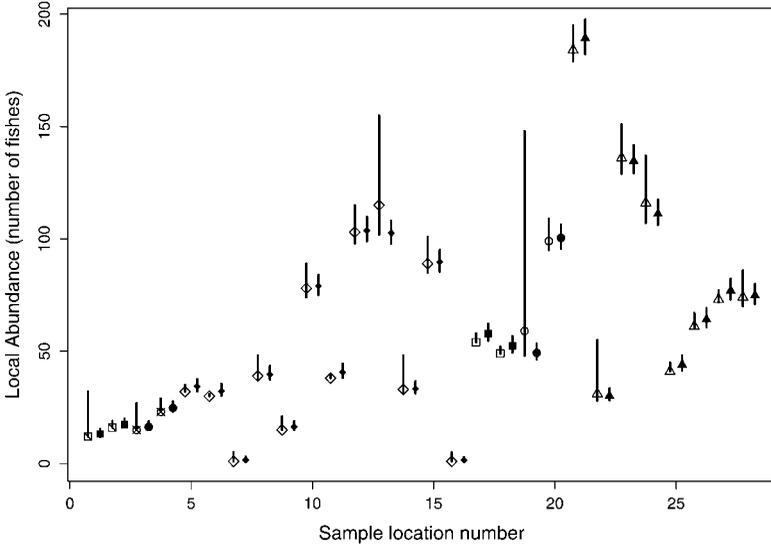


Figure 2. Estimated abundance of fishes at each of 28 sample locations. Different symbols represent different drainages. Maximum-likelihood estimates (open symbols) and Bayesian estimates (closed symbols) are plotted side-by-side with their respective 95% profile-likelihood intervals and 95% credible intervals.

season. Each quadrat is surveyed three times annually by an experienced observer along a quadrat-specific route, thus yielding the vector of counts $\mathbf{x}_i = (x_{i1}, x_{i2}, x_{i3})$ for quadrats $i = 1, 2, \dots, n$. A method known as “territory mapping” is used, yielding data that are interpreted as territory counts. The data used here are counts of the blackbird (*Turdus merula*) collected on 239 quadrats during 2002. Two important covariates are available to describe variation in abundance across Switzerland: elevation, the response to which we suppose is quadratic, and forest cover (both covariates were standardized to have mean zero and unit variance). For additional details on the data and survey, and alternative analyses, consult Kéry, Royle, and Schmid (2005).

The basic sampling strategy yields data consistent with the product-binomial likelihood given by (2.2). Thus, we might suppose that x_{ij} is binomial with index N_i and constant detection probability π . However, an important consideration affecting interpretation of these data is that the sample route length is not constant among quadrats. Each observer chooses a route length that remains constant across replicate samples of the same quadrat, but varies among quadrats. Kéry et al. (2005) argued that route length is a surrogate for sampled area and parameterized the effect of the logarithm of route length as being additive on the Poisson log-mean. We adopt a slightly different formulation of the effect of route length, and introduce another level into the hierarchical model that yields an “adjustment” for incomplete quadrat coverage, and a more desirable interpretation of estimates. First, we define N_i to be the abundance of the *quadrat*, and suppose that N_i is Poisson with mean λ_i and

$$\log(\lambda_i) = \alpha_0 + \alpha_1 \text{Forest}_i + \alpha_2 \text{Elev}_i + \alpha_3 \text{Elev}_i^2.$$

```

model {
  # Specify priors for metapopulation-level parameters
  alpha0 ~ dnorm(0, .0001)
  alpha1 ~ dnorm(0, .0001)
  alpha2 ~ dnorm(0, .0001)
  alpha3 ~ dnorm(0, .0001)
  alpha4 ~ dnorm(0, .0001)
  theta ~ dunif(0,1)

  # Specify observation model and priors for local parameters
  for(i in 1:n){
    log(lambda[i]) <- alpha0 + alpha1*forest[i] +alpha2*elev1[i]
+alpha3*elev2[i]
    log(g[i]) <- alpha4*length[i]
    N[i] ~ dpois(lambda[i])
    M[i] ~ dbin(g[i],N[i])
    for(j in 1:T){
      x[i,j] ~ dbin(theta,M[i])
    }
  }
  Ntot <- sum(N[])
  Mtot <- sum(M[])
}

```

Figure 3. WinBUGS code for fitting the hierarchical model of avian counts.

Then, we define $M_i \leq N_i$ to be the population of quadrat i that is *exposed* to sampling by a route of length L_i . We suppose then that M_i is binomial with index N_i and parameter $g(L_i)$. Sensible parameterizations of $g(L_i)$ have the property that $g(L_i) \rightarrow 1$ as $L_i \rightarrow \infty$ so that, as route length increases, the exposed population tends to the actual quadrat population. Thus, N_i is the population that one would be sampling if quadrat i was saturated with effort. To model the relationship between M_i and N_i we suppose that $g(L_i) = \exp\{\alpha_4/L_i\}$ which has the desirable property mentioned previously if $\alpha_4 < 0$ (note that one could formally impose the restriction that $\alpha_4 < 0$). Finally, the observed counts, x_{ij} , are independent binomial samples with index M_i and detection probability θ . This yields a hierarchical model of the form $[x_i | M_i, \theta][M_i | N_i, g(L_i)][N_i | \alpha][\alpha, \theta]$ from which we are primarily interested in making inferences about the N_i parameters (i.e., abundance at the scale of the sample quadrat) including factors that influence variation in N_i .

Our WinBUGS code for specifying the model containing the three abundance covariates and the additional hierarchical structure relating M_i to N_i appears in Figure 3. The two-dimensional (sites \times replicates) is represented by $x[i, j]$ in Figures 1 and 3. Thus, for example, the line of code $x[i, 1] \sim \text{dbin}(\text{theta}, N[i])$ specifies the conditional probability distribution of the count obtained from the first sample at site i , given the parameters θ and local abundance, N_i . Two additional estimands are defined in this model specification: the total abundance across all 239 sample quadrats unadjusted for incomplete quadrat coverage (“Mtot”) and adjusted for incomplete quadrat coverage (“Ntot”). Summaries of the posterior distribution were calculated from three independent Markov chains initialized with random starting values, run 100,000 iterations after a 10,000 itera-

Table 1. Posterior Summaries of Model Parameters for the Swiss Avian Survey Data

<i>Parameter</i>	<i>Mean</i>	<i>Std. dev.</i>	<i>2.5%</i>	<i>Median</i>	<i>97.5%</i>
α_0	2.87	0.100	2.67	2.88	3.06
α_1	0.11	0.021	0.07	0.11	0.16
α_2	-1.21	0.044	-1.30	-1.21	-1.13
α_3	-0.30	0.047	-0.39	-0.30	-0.21
α_4	-2.69	0.420	-3.53	-2.69	-1.88
θ	0.57	0.017	0.54	0.57	0.60
Ntot	5615.00	484.600	4732.00	5589.00	6638.0
Mtot	3331.00	97.800	3153.00	3326.00	3537.00

tion burn-in, and every fifth draw was retained for analysis. This yielded three sets of 20,000 posterior draws. We computed the Brooks-Gelman-Rubin convergence diagnostic (Gelman and Rubin 1992; Brooks and Gelman 1998) using the output from the three chains. The multivariate potential scale reduction factor was 1.000145, and the scale reduction factor for each parameter was within 10^{-4} of 1.0.

Estimates of model parameters and these two estimands are provided in Table 1. We note a slight positive effect of forest cover and a negative effect of elevation on abundance. Under this model, the estimated probability of detection (during a single sample) is $\hat{\theta} = 0.57$. Consequently, the net probability of detection (i.e., over three independent sample periods) is about 0.92. Thus, we can expect to detect nearly all of the birds exposed to sampling by a route. Note, however, that the effect of incomplete quadrat coverage is substantial, yielding an adjustment from 3,331 exposed territories to 5,586, aggregated over all 239 quadrats. We note that various other model embellishments are possible such as those considered in the Darter example, and also see Kéry et al. (2005).

6. DISCUSSION

Hierarchical models provide a general and flexible framework for addressing inference problems about animal abundance and occurrence from spatially referenced survey data that are subject to imperfect detection. By specification of models that describe variation in abundance and detection among sites, hierarchical models admit a formal decomposition of variation in the observed data into that due to abundance, and that due to detectability, and they permit direct inferences about ecological processes of interest.

Hierarchical models permit a unified treatment of modeling and inference for a vast array of animal sampling methods. For example, such models have been used to estimate animal abundance from binary “presence/absence” data (Royle and Nichols 2003), and simple (binomial) counts (Dodd and Dorazio 2004; Royle 2004). Hierarchical models have been used to resolve important inference problems from multinomial counts resulting from various protocols including small-area estimation in removal sampling (Dorazio et al. 2005), modeling spatial variation on local abundance in distance sampling (Royle et al. 2004), development of spatial models of abundance and occurrence from capture-recapture samples (Royle et al. 2006), and inference from abundance “index” data (Royle 2004b; Royle and

Link 2005). The hierarchical modeling framework can be directly extended to other common sampling protocols including multiple observer sampling, and double sampling and, we believe, provides a natural framework for developing models for spatio-temporal dynamics (Wikle 2003) and multiple species systems (Dorazio and Royle 2005).

Estimation and inference under such models can be achieved by conventional (integrated) likelihood methods (Royle and Nichols 2003; Dodd and Dorazio 2004; Royle 2004) or, as described in this article, by adopting a fully Bayesian formulation of the model. The Bayesian formulation provides an accurate assessment of uncertainty of estimates of model parameters, predictions, and functions of model parameters. In both examples considered in this article, prediction of the latent abundance parameters is a central focus of the analyses (e.g., Figure 2). Although a number of non-Bayesian methods of inference in such problems can be described (see, e.g., Kéry et al. 2005; Dorazio et al. 2005), the lack of generality of such methods contrasts with the flexibility and rigor of the hierarchical Bayesian formulation. Moreover, the shrinkage behavior of Bayesian estimators (that Bayesian estimators “borrow strength” from the ensemble) is evident in the narrower width of the Bayesian intervals of Figure 2. This is critical in applications such as those considered here, in which the sample size of each spatial sample unit is relatively small, yielding little information about each sample unit, yet the overall sample size is such that considerable information is provided about the metapopulation structure. Such is also the motivation behind conventional small-area estimation problems. Generalizations of the models that admit additional hierarchical structure can render likelihood-based estimation and inference inefficient, but are relatively straightforward in a Bayesian framework (Dorazio and Royle 2005; Royle et al. 2006). We have demonstrated that Bayesian analysis of hierarchical models is remarkably straightforward using the software package WinBUGS, requiring little more than a “pseudo-code” representation of the model.

The models described here are based on the assumptions that local populations are distinct (i.e., nonoverlapping) and closed to demographic processes such as recruitment, mortality, and movement. These are important considerations because, when satisfied, they yield desirable interpretations of model parameters (e.g., of local abundance) and, under suitably strict standardization of survey protocols, relatively simple models of abundance and detection at the level of the local population. Although these assumptions are often not strictly met in practice, their importance is largely related to timing and spatial extent of sampling. Moreover, certain departures from these assumptions can be accommodated by extensions of the hierarchical model. For example, overlapping local populations can be parameterized by conditioning on a finer partitioning of abundance such that, when aggregated to the level of a sample unit, dependence among local population (sizes) is induced. Also, temporary movements of animals into and out of the local population can be modeled by defining larger “super-populations” from which independent draws of N_i are sampled (similar to the construction used in the Swiss point count data described in Section 5.2). The development and application of generalizations such as these is an important area of ongoing research.

ACKNOWLEDGMENTS

The authors are grateful to Howard Jelks (Florida Integrated Science Center, U.S. Geological Survey, Gainesville, FL) and Frank Jordan (Department of Biological Sciences, Loyola University New Orleans, New Orleans, LA) for sharing their removal counts of Okaloosa darters. The authors also thank Hans Schmid and Marc Kéry (Swiss Ornithological Institute, Sempach, Switzerland) for the avian point count data, and the dedicated field-naturalists that annually conduct the Swiss breeding bird survey. The authors thank the editor, associate editor, and two anonymous referees for many helpful comments and suggestions.

[Received April 2005. Revised March 2006.]

REFERENCES

- Bayley, P. B., and Peterson, J. T. (2001), "An Approach to Estimate Probability of Presence and Richness of Fish Species," *Transactions of the American Fisheries Society*, 130, 620–633.
- Brooks, S. P., and Gelman, A. (1998), "General Methods for Monitoring Convergence of Iterative Simulations," *Journal of Computational and Graphical Statistics*, 7, 434–455.
- Carroll, R. J., and Lombard, F. (1985), "A Note on N Estimators for the Binomial Distribution," *Journal of the American Statistical Association*, 80, 423–426.
- Diggle, P. J., Tawn, J. A., and Moyeed, R. A. (1998), "Model-Based Geostatistics" (with discussion), *Journal of the Royal Statistical Society, Series C*, 47, 299–350.
- Dodd, C. K., and Dorazio, R. M. (2004), "Using Counts to Simultaneously Estimate Abundance and Detection Probabilities in a Salamander Community," *Herpetologica*, 60, 468–478.
- Dorazio, R. M., and Jelks, H. L., and Jordan, F. (2005), "Improving Removal-Based Estimates of Abundance by Sampling a Population of Spatially Distinct Subpopulations," *Biometrics*, 61, 1093–1101.
- Dorazio, R. M., and Royle, J. A. (2005), "Estimating Size and Composition of Biological Communities by Modeling the Occurrence of Species," *Journal of the American Statistical Association*, 100, 389–398.
- Gelman, A., and Rubin, D. B. (1992), "Inference From Iterative Simulation Using Multiple Sequences," *Statistical Science*, 7, 457–511.
- Gilks, W. R., Thomas, A., and Spiegelhalter, D. J. (1994), "A Language and Program for Complex Bayesian Modelling," *The Statistician*, 43, 169–178.
- Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996), *Markov Chain Monte Carlo in Practice*, London: Chapman and Hall.
- Hanski, I., and Simberloff, D. (1997), "The Metapopulation Approach, its History, Conceptual Domain, and Application to Conservation," in *Metapopulation Biology: Ecology, Genetics, and Evolution*, eds. I. Hanski and M. E. Gilpin, New York: Academic Press, pp. 1–11.
- Kéry, M., Royle, J. A., and Schmid, H. (2005), "Modeling Avian Abundance from Replicated Counts Using Binomial Mixture Models," *Ecological Applications*, 15, 1450–1461.
- Laird, N. M., and Louis, T. A. (1987), "Empirical Bayes Confidence Intervals Based on Bootstrap Samples" (with discussion), *Journal of the American Statistical Association*, 82, 739–757.
- Link, W. A., and Sauer, J. R. (2002), "A Hierarchical Model of Population Change with Application to Cerulean Warblers," *Ecology*, 83, 2832–2840.
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., and Langtimm, C. A. (2002), "Estimating Site Occupancy Rates when Detection Probabilities are Less than One," *Ecology*, 83, 2248–2255.
- MacKenzie, D., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. A., and Hines, J. E. (2005), *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*, Burlington, MA: Elsevier.

- Magnusson, W. E., Caughley, C. G., and Grigg, G. C. (1978), "A Double-Survey Estimate of Population Size From Incomplete Counts," *Journal of Wildlife Management*, 42, 174–175.
- Manly, B. F. J., McDonald, L. L., Thomas, D. L., McDonald, T. L., and Erickson, W. P. (2002), *Resource Selection by Animals: Statistical Design and Analysis for Field Studies* (2nd ed.), Dordrecht: Kluwer Academic Publishers.
- Olkin, I., Petkau, A. J., and Zidek, J. V. (1981), "A Comparison of n Estimators for the Binomial Distribution," *Journal of the American Statistical Association*, 76, 637–642.
- Raftery, A. E. (1988), "Inference for the Binomial N Parameter: A Hierarchical Bayes Approach," *Biometrika*, 75, 223–228.
- Royle, J. A. (2004), "N-mixture Models for Estimating Population Size from Spatially Replicated Counts," *Biometrics*, 60, 108–115.
- (2004b), "Modeling Abundance Index Data from Anuran Calling Surveys," *Conservation Biology*, 18, 1378–1385.
- Royle, J. A., Dawson, D. K., and Bates, S. (2004), "Modeling Abundance Effects in Distance Sampling," *Ecology*, 85, 1591–1597.
- Royle, J. A., Kéry, M., Gautier, R., and Schmid, H. (2006), "Hierarchical Spatial Models of Avian Abundance and Occurrence from Imperfect Survey Data," unpublished manuscript.
- Royle, J. A., and Link, W. A. (2005), "A General Class of Multinomial Mixture Models for Anuran Calling Survey Data," *Ecology*, 86, 2505–2512.
- Royle, J. A., and Nichols, J. D. (2003), "Estimating Abundance from Repeated Presence-Absence Data or Point Counts," *Ecology*, 84, 777–790.
- Sargeant, G. A., Sovada, M. A., Slivinski, C. C., and Johnson, D. H. (2005), "Markov chain Monte Carlo Estimation of Species Distributions: A Case Study of the Swift Fox in Western Kansas," *Journal of Wildlife Management*, 69, 483–497.
- Schmid, H., Luder, R., Naef-Daenzer, B., Graf, R., and Zbinden, N. (2001), "Schweizer Brutvogelatlas. Verbreitung der Brutvögel in der Schweiz und im Fürstentum Liechtenstein 1993–1996, Schweizerische Vogelwarte, Sempach, Switzerland.
- Wikle, C. K. (2003), "Hierarchical Bayesian Models for Predicting the Spread of Ecological Processes," *Ecology*, 84, 1382–1394.
- Williams, B. K., Nichols, J. D., and Conroy, M. J. (2002), *Analysis and Management of Animal Populations*, San Diego, CA: Academic Press.