STATISTICAL REPORTS

MODELING ABUNDANCE EFFECTS IN DISTANCE SAMPLING

J. Andrew Royle,[1,4] Deanna K. Dawson,[2] and Scott Bates[3]

[1]*U.S. Fish and Wildlife Service, Division of Migratory Bird Management, 11510 American Holly Drive,*
Laurel, Maryland 20708 USA
[2]*USGS Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, Maryland 20708 USA*
[3]*National Park Service, Center for Urban Ecology, 4598 MacArthur Boulevard, NW, Washington, D.C. 20007-4227 USA*

Abstract. Distance-sampling methods are commonly used in studies of animal populations to estimate population density. A common objective of such studies is to evaluate the relationship between abundance or density and covariates that describe animal habitat or other environmental influences. However, little attention has been focused on methods of modeling abundance covariate effects in conventional distance-sampling models. In this paper we propose a distance-sampling model that accommodates covariate effects on abundance. The model is based on specification of the distance-sampling likelihood at the level of the sample unit in terms of local abundance (for each sampling unit). This model is augmented with a Poisson regression model for local abundance that is parameterized in terms of available covariates. Maximum-likelihood estimation of detection and density parameters is based on the integrated likelihood, wherein local abundance is removed from the likelihood by integration. We provide an example using avian point-transect data of Ovenbirds (*Seiurus aurocapillus*) collected using a distance-sampling protocol and two measures of habitat structure (understory cover and basal area of overstory trees). The model yields a sensible description (positive effect of understory cover, negative effect on basal area) of the relationship between habitat and Ovenbird density that can be used to evaluate the effects of habitat management on Ovenbird populations.

Key words: abundance estimation; avian point counts; distance-sampling methodology; mixture models; Ovenbird; random effects.

INTRODUCTION

The use of distance-sampling methods to estimate density is widespread in studies of animal populations (Buckland et al. 2001, Williams et al. 2002). Such methods use information on observed distances of animals from transects or points of observation to characterize the detection probability of individuals. Under the hypothesis that detection probability is related to the distance between animals and the point of observation, one may obtain an estimate of density that is, in effect, adjusted for nondetection bias.

Distance-sampling methods are attractive in many animal-sampling problems because they do not require that individuals be uniquely marked and recaptured (or resighted) through time. In avian counting applications based on point counts (Buckland et al. 2001, Rosenstock et al. 2002), distances are recorded from a point of observation (instead of a transect), and this is usually referred to as a "point transect" (Buckland et al. 2001).

In many field situations, birds are detected only by their vocalizations and so it can be difficult to obtain precise distance information. Consequently, distances are frequently recorded by grouping into discrete distance intervals. We focus subsequent discussion on these point-transect situations, due to our interest in modeling avian point-count data described in *Application . . .* , below.

In distance sampling, as with other common sampling protocols, considerable modeling effort is focused on describing variation in detection probability (Buckland et al. 2001, Marques and Buckland 2003, Ramsey and Harrison 2004). To this end, very complex models of detection probability are often considered. However, many applications have explicit objectives that involve understanding mechanisms that affect abundance. In particular, estimation of the effect of spatial covariates on abundance is fundamental to many investigations of animal populations. For example, landscape or habitat characteristics associated with each spatial sample unit are often collected, and interest is in modeling the relationship between abundance (or density) and these measured covariates. Most frequently, applications that focus on modeling the effect of

abundance covariates do not explicitly address the issue of detectability, and use conventional techniques such as regression or generalized linear models for modeling the observed counts or detection/nondetection (e.g., Vincent and Haworth 1983, Robbins et al. 1989, Friesen et al. 1995, Hutto 1995, Young and Hutto 1999, 2002, Fewster et al. 2000, Brand and Georgea 2001, Vernier et al. 2002, Beadell et al. 2003). This can lead to biased estimators of habitat effects (Gu and Swihart 2004). There have been few methods developed for accommodating covariate effects on abundance into distance sampling or other procedures that directly account for detectability. Application of conventional distance sampling (and related) methods to point-transect data is based on pooling data collected from multiple point counts. This yields an estimate of average density, but results in a loss of information at the level of sample units (e.g., point locations) due to pooling.

We propose a model that allows for incorporation of abundance covariate effects within distance-sampling models. We pose the likelihood for data from each point as a function of "local abundance" in the vicinity of that point. Then local abundance may be related to covariates using a Poisson or other generalized linear regression model. Local abundance is regarded as a random effect, and analysis is based on the integrated likelihood, which is a function of parameters of the detection function, density, and relevant density covariates. We apply the proposed model to avian point-transect data that are typical of those to which distance-sampling methods are often applied.

DATA AND MODEL

We suppose that point counts of a fixed radius are conducted at $i = 1, 2, \ldots, R$ points in some region (e.g., a park or forest). Each point count represents a sample of fixed area, which we will refer to as a "sample unit" or "site." We consider grouped data here wherein distances are recorded in discrete intervals from the central point of observation for each site. Let $k = 1, 2, \ldots, K$ index the distance classes, with end points $(c_1, c_2), (c_2, c_3), \ldots, (c_K, c_{K+1})$. Here, c_{K+1} is the maximum distance at which birds were counted, or the radius of the point count. Let y_{ik} be the observed count of individuals in distance class k for site $i = 1, 2, \ldots, R$.

Let $g(x; \theta)$ denote some function that describes the relationship between detection probability and distance, x, from the point of observation. This detection probability function is parameterized by the (possibly vector-valued) parameter θ. See Buckland et al. (2001) for a discussion of possible detection functions and underlying assumptions.

The key departure from the traditional formulation of distance-sampling estimators of density that we entertain here is that we consider estimation based on the *unconditional* likelihood of the observed counts as opposed to the likelihood which is "conditioned on detection" (see Sanathanan [1972] for a discussion of the

distinction). This allows direct parameterization of the likelihood for the data from each site in terms of the abundance at that site, N_i. The main reason for consideration of the unconditional likelihood here is that modeling spatial variation in N_i (or density) is straightforward under the unconditional formulation, as demonstrated subsequently. Under the conventional conditional formulation, there is no site-specific quantity that can be regarded as abundance or density, and that can be related to site-specific covariates. However, Hedley and Buckland (2004) recently have developed an approach based on the conditional likelihood (see *Discussion*, below).

In terms of N_i, the site-specific likelihood for data $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{iK})$ is

$$f(\mathbf{y}_i \mid \theta) = \frac{N_i!}{\left(\prod_k y_{ik}!\right)(N_i - y_{i\cdot})!} \left[\prod_k \pi_k(\theta)^{y_{ik}}\right]$$
$$\times \left[1 - \sum_k \pi_k(\theta)\right]^{N_i - y_{i\cdot}} \quad (1)$$

where $y_{i\cdot} = \Sigma_k y_{ik}$.

Construction of cell probabilities, $\pi_k(\theta)$, is based on conventional considerations (e.g., Buckland et al. 2001: chapter 3). Specifically, in Eq. 1, $\pi_k(\theta)$ is the probability that an individual occurs and is detected in distance class k, which depends on the detection function under consideration. For point transects, this can be computed by integrating $g(x; \theta)$ over the area of the circle between c_k and c_{k+1} (e.g., see Buckland et al. 2001:54). To be precise,

$$\pi_k(\theta) = \int_{c_k}^{c_{k+1}} \frac{2\pi x g(x; \theta)}{\pi c_{K+1}^2} \, dx.$$

We emphasize that $\pi_k(\theta)$ here are unconditional cell probabilities, in contrast to those considered in most distance-sampling situations (e.g., the π's given in Buckland et al. [2001: chapter 3]).

Modeling variation in abundance among sites

Fundamental to the situation under consideration here is that abundance may vary among sites due to measurable environmental characteristics that are also site specific (e.g., habitat). Thus, we require an extension of Eq. 1 that makes this idea precise. For that, we augment the model with an additional model containing potential sources of variation in N_i. A natural choice is based on conventional generalized-linear-modeling (GLM) ideas. Suppose N_i were *observed,* then consider the Poisson regression model specified by

$$N_i \sim \text{Poisson}(\lambda_i) \quad (2)$$

where λ_i is the expected value of N_i, typically assumed to be linearly related to available covariates according to the following (for a single covariate):

TABLE 1. Results of fitting the four candidate models of mean density to the Ovenbird data, using the Poisson and negative-binomial abundance models.

Candidates	np	AIC	β	α_0	UFC	BA	k
Poisson							
Model 0	2	340.82	14.23	-0.551			
Model 1 (UFC)	3	340.27	14.23	-0.567	1.859		
Model 2 (BA)	3	339.30	14.23	-0.574		-0.829	
Model 3 (both)	4	340.67	14.23	-0.578	1.042	-0.643	
Negative binomial							
Model 0	3	342.73	14.23	-0.551			0.053
Model 1 (UFC)	4	342.26	14.23	-0.567	1.866		0.016
Model 2 (BA)	4	341.30	14.23	-0.574		-0.832	0.010
Model 3 (both)	5	342.67	14.23	-0.578	1.042	-0.643	0.000

Notes: Explanation of column heads: np, number of parameters; AIC, Akaike information criterion; β, scale parameter of distance function; α_0, intercept of the abundance model; UFC, understory foliage cover coefficient estimate; BA, basal area coefficient estimate; k, negative binomial over-dispersion parameter.

$$\log(\lambda_i) = \alpha_0 + \alpha_1 z_i \qquad (3)$$

where z_i is the value of the covariate measured at site i. Note that in the usual terminology of Poisson models, λ_i is *density* per ''sample unit'' (assuming the sample units are of the same area), and so this model now includes a parameter that can be interpreted in a manner consistent with traditional distance-sampling ideas, but where density is now spatially explicit. If covariates are centered to have mean zero, then $\exp(\alpha_0)$ is the ''mean density.'' One referee suggested incorporating the term $\log(\pi c_{K+1}^2)$ as an additive offset in Eq. 3 so that λ_i has the more precise interpretation of individuals per unit area (recall that c_{k+1} is the outer limit of observation beyond which birds are not counted). Also, if point-count areas are not constant, then this should be accounted for in a similar manner.

We note that alternative abundance models may also be considered. An appealing and common model for describing variation in counts (e.g., animal abundance) in the presence of over-dispersion (or ''excess-Poisson variation'') is the negative-binomial (Boyce et al. 2001) model. A common parameterization of the negative binomial (we use this parameterization in *Application . . .* , below) is that in which $E[N] = \lambda$ and $\text{Var}[N] = \lambda + k\lambda^2$ where $k > 0$ is the ''over-dispersion'' parameter. Thus, $\text{Var}[N] \geq E[N]$ and in limit (as k goes to 0) the negative binomial is equivalent to the Poisson distribution. As with the Poisson, a log-linear model relating covariates to mean abundance may be used. Abundance models other than Poisson or negative binomial that allow for more flexibility in modeling over-dispersion may also be considered (e.g., see Bhattacharya and Holla 1965, Agarwal et al. 2002, Puig 2003). In practice it may be desirable to formally choose among several plausible models of abundance given observational data. When inference is based on the integrated likelihood described subsequently, model selection may be carried out using AIC (Akaike information criterion; Burnham and Anderson 1998).

In contrast to many conventional applications of Poisson regression, the N_i are *not* observable in distance-sampling problems. A solution that yields a precise treatment of the likelihood (Eq. 1) in the context of the Poisson regression model (expression 2) is to regard N_i as unobserved *random effects* with distribution given by expression 2 and, following conventional notions of the treatment of random effects (e.g., Laird and Ware 1982, Robinson 1991), integrate them from the likelihood (Eq. 1). This yields a marginal likelihood (often referred to as the ''integrated likelihood'') that is only a function of detection parameters (contained in $g(x; \theta)$), and density parameters α_0 and α_1.

The integrated likelihood

The integrated likelihood for the data from site i is formulated by integrating Eq. 1 over the random-effects distribution (expression 2):

$$L(\boldsymbol{\alpha}, \theta \,|\, \mathbf{y}_i) = \sum_{N_i = y_{i.}}^{\infty} \left\{ \frac{N_i!}{\left(\prod_k y_{ik}!\right)(N_i - y_{i.})!} \left[\prod_k \pi_k(\theta)^{y_{ik}} \right] \right.$$
$$\left. \times \left[1 - \sum_k \pi_k(\theta) \right]^{N_i - y_{i.}} \right\} \frac{e^{-\lambda_i(\boldsymbol{\alpha})} \lambda_i(\boldsymbol{\alpha})^{N_i}}{N_i!}.$$
$$(4)$$

Note that the dependence of λ_i on the parameters $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)$ is emphasized in this expression.

One benefit of the Poisson abundance model is that the summation in Eq. 4 can be done analytically, yielding the product Poisson likelihood:

$$L(\boldsymbol{\alpha}, \theta \,|\, \mathbf{y}_i) = \prod_{k=1}^{K} \text{Poisson}[\, y_{ik}; \lambda_i(\boldsymbol{\alpha})\pi_k(\theta)]. \qquad (5)$$

The simple likelihood obtained under the Poisson abundance model is more computationally efficient than the

general mixture form given by Eq. 4. For other abundance models (such as the negative binomial) Eq. 4 does not simplify in any meaningful way.

Assuming the data are independent across sites, the joint likelihood for all data is then

$$L(\alpha, \theta \mid \mathbf{y}_1, \ldots, \mathbf{y}_R) = \prod_{i=1}^{R} L(\alpha, \theta \mid \mathbf{y}_i). \qquad (6)$$

Practically speaking, independence here implies that point counts are made sufficiently far apart so as to ensure that the same individuals are not being counted at multiple sites. Consequently, this may only be reasonable for species that can be sampled at a time when individuals are territorial or sedentary. Note that independence underlies all of the conventional distance-sampling methods and so it should not be viewed as an assumption that is specific to the proposed model.

The integrated likelihood (Eq. 6) can be maximized using conventional numerical methods. For the analyses reported in the following section, we developed our own routines using the free software **R** (Ihaka and Gentleman 1996). Note that for abundance models other than the Poisson, the upper bound appearing in the summation of Eq. 4 must be truncated at some finite integer. One can check that the chosen cut-off was sufficient by increasing its value and verifying that the MLEs (maximum-likelihood estimators) are unchanged.

As a final comment, note that it is possible to estimate individual N_i's using what is usually referred to as best unbiased prediction (BUP), although there is often no obvious need to do so since most inferential problems focus on density ("average abundance") effects. See Royle (2004) for details.

Evaluating model goodness of fit

Under the Poisson model, the conventional deviance statistic for Poisson data (Agresti 2002) may be used in sufficiently large samples. However, small counts such as those generated from our study (and perhaps most studies based on avian point-count surveys) render the asymptotic null distribution of this statistic invalid. Also, a convenient deviance statistic (that is, with a known asymptotic null distribution) under other abundance distributions has not been developed. Consequently, we considered assessment of goodness of fit using conventional parametric bootstrap procedures (Dixon 2002).

APPLICATION TO AVIAN POINT-COUNT DATA

Here we consider avian point-count data collected in late May and June of 2002 at 70 sites in the Catoctin Mountains, Frederick County, Maryland, USA. To assess the possible impacts of white-tailed deer (*Odocoileus virginianus*) on natural resources in the Catoctin Mountain Park, a unit of the National Park Service, breeding-bird populations and vegetation were sampled in the park and in the nearby Frederick City Watershed Cooperative Wildlife Management Area, where deer

numbers are controlled by hunting. Thirty-five sites in each area were randomly selected for sampling from a grid of points, spaced at 250-m intervals, generated in ArcInfo (ESRI, Redlands, California, USA).

At each site, a 12-min count was made of all birds seen or heard. Distances to detected birds were recorded in 25-m distance classes out to 100 m from the point. Observations of two covariates thought to influence local bird abundance were also collected: percent understory-foliage cover (UFC) and basal area of overstory trees (BA; in square meters). Understory vegetation was sampled by counting the number of height intervals (0–0.1 m, >0.1–0.3 m, >0.3–0.5 m, >0.5–1.0 m, >1.0–1.5 m, >1.5–2.0 m, >2.0–2.5 m, and >2.5–3.0 m aboveground) in which live vegetation intersected a pole planted at 3-m intervals along 25-m transects that radiated from the point in the cardinal directions (north, south, east, west). For each point, UFC is the percentage of the height-interval samples in which vegetation was present. A forestry prism (basal area factor: 10) was used to sample overstory basal area at each point (Hovind and Rieck 1970).

The goal is to evaluate the relationship between bird density and these habitat characteristics. In this illustration, we consider counts of male Ovenbirds (*Seiurus aurocapillus*), which are detected primarily by song. Ovenbirds are a ground-nesting species, and so it is suspected that understory vegetation serves to provide nesting cover. Conversely, excessive overstory (measured by BA) impedes the development of understory cover. Consequently, we expect a positive relationship between N_i (abundance at site i) and UFC and a negative relationship between N_i and BA. These covariates were standardized to have mean zero, and incorporated into the abundance model as in Eq. 3, with the additive offset $\log(\pi)$. Because the point-count radius was 100 m, λ_i is interpreted as the number of individuals per hectare (10 000 m^2). We used a "half-normal" detection function to model the relationship between detection probability and distance:

$$g(x; \theta) = \exp\left(-\frac{x^2}{25\beta}\right).$$

Note that the denominator is scaled so that the (equally spaced) distance intervals are bounded by the integers (0, 1), (1, 2), (2, 3), and (3, 4).

We considered four possible models of density to describe the Ovenbird data using both the Poisson and negative-binomial models for abundance (eight models in all). Model 0 contained neither covariate. Models 1 and 2 contained only UFC and BA, respectively, and Model 3 contained both covariates. Results of fitting these models are given in Table 1. To evaluate the relative merits of each model, the AIC score (Burnham and Anderson 1998) for each model is also given.

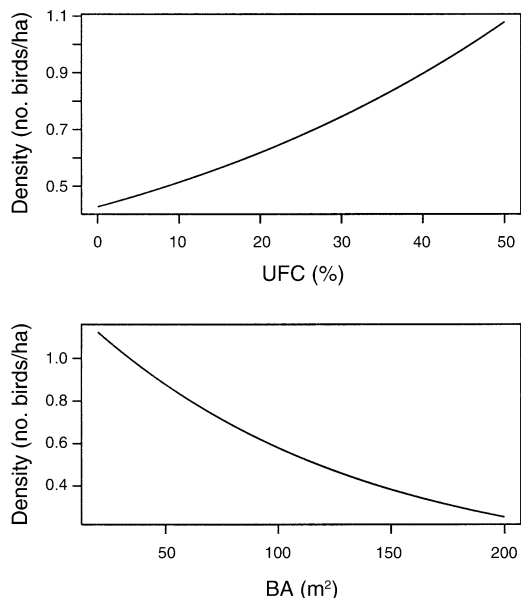We note that the estimated over-dispersion parameter of the negative-binomial model is near 0 in all cases

FIG. 1. Estimated Ovenbird density as a function of understory foliage cover (UFC; top panel) and basal area (BA; bottom panel).

(Table 1). Consequently, the estimates of the other parameters in the negative-binomial models are essentially equivalent to those of the corresponding Poisson models. Thus, AIC favors the simpler (more parsimonious) Poisson abundance model. Among the Poisson models, those with one or the other covariate (but not both) are preferred. Note that the two covariates are negatively correlated with one another (sample correlation, $\rho = -0.49$), and thus they contain redundant information. This explains why the model with both covariates has a higher AIC despite the apparent importance of each covariate individually.

The signs of the coefficient estimates are as anticipated (positive relationship between abundance and UFC, negative for BA). The estimated effect of each covariate (on the density scale) is depicted graphically in Fig. 1 where λ is plotted as a function of UFC and BA using estimates obtained from the corresponding models. Thus, for example, as UFC increases from 0 to 50%, density increases from approximately 0.4 to 1.1 Ovenbirds/ha.

Goodness of fit was evaluated for the best fitting Poisson model using the parametric bootstrap procedure based on model deviance. This indicated no significant lack of fit ($P = 0.726$).

Finally, note that the estimated scale parameter (β in Table 1) is the same to within two decimal places for all eight models considered. Apparently, it is unaffected by variation in N as might be expected (note that the N_i's can be removed by conditioning).

DISCUSSION

We have proposed a modeling approach that allows for incorporation of covariate effects on abundance in distance-sampling models. The key idea underlying this model is the formulation of the (multinomial) distance-sampling likelihood in terms of site-specific abundance parameters, N_i. These are regarded as random effects and assigned a distribution wherein the mean depends on the covariates under consideration. The marginal likelihood of the data is constructed by integrating the multinomial likelihood for the data over this random-effects distribution.

The model proposed here expands on that considered by Royle (2004) for estimating abundance from point-count data. That model is based on *simple* point-count data (i.e., with no distance information), and also uses an integrated likelihood wherein the binomial likelihood for the observed counts is integrated over a Poisson prior distribution on N_i. See also Royle and Nichols (2003) for a similar strategy for modeling heterogeneity in detection in occupancy surveys that is due to variation in abundance. That idea has been used here to facilitate modeling structure in site-specific abundance under a multinomial distance-sampling model.

Distance sampling is widely used to estimate abundance of animal populations, and the underlying theory is well developed and intricate. Despite this, until recently, relatively little attention has been focused on the development of distance-sampling methods that allow for modeling covariate effects. Note that Buckland et al. (2001) contains a single section (3.8) devoted to the topic. Much of what has been done has focused on modeling covariates that affect *detection probability* (e.g., Ramsey et al. 1987, Marques and Buckland 2003, Ramsey and Harrison 2004). To the best of our knowledge, only Hedley et al. (1999) and Hedley and Buckland (2004) consider models that allow for covariate effects on abundance. These recent efforts focus on formulating distance-sampling models in terms of the spatial-intensity function of a Poisson point process and adopt the more common conditional formulation of the likelihood for the observation locations of individuals. The resulting likelihood is not analytically tractable and appears unstable for complex intensity functions (Hedley and Buckland 2004). This motivates Hedley and Buckland (2004) to suggest a more informal approach that first estimates the detection function, and then uses a ''plug-in'' type of procedure based on the partial likelihood of the observed counts. Our approach is similar in the sense that the abundance model is derived from consideration of a spatially varying intensity function. However, formulation of the problem in terms of the unconditional multinomial likelihood (a function of local abundance) and then integrating over a prior distribution for N (containing covariate effects), yields a well-behaved likelihood. In the case where the prior is Poisson, the likelihood is of closed form (a product Poisson likelihood). Generality to other abundance distributions (determining mean–variance relationships) can also be considered, and it is not clear how this can be accommodated in

the likelihood considered by Hedley and Buckland (2004).

A natural result of a spatially explicit formulation of abundance models is that one may produce maps of the density surface over the sample domain, or predict abundance at unsampled sites. This is necessary if characterizing the total population size is required, and requires that the spatial covariates are measured everywhere. Hedley and Buckland (2004) consider spatial mapping of the intensity function in their application, which involved covariates that were functions of spatial location (e.g., spatial regression functions).

While our focus here has been on point-transect data collected according to a distance-sampling protocol, the general notion is clearly applicable to more common distance-sampling situations that make use of transects, when landscape or habitat covariates can be measured for each transect, or segments of each transect. Also note that the general modeling strategy proposed here can be applied regardless of the form of the multinomial likelihood. In particular, it can be applied to problems where the multinomial likelihood is based on other sampling protocols such as multiple observer (Nichols et al. 2000), removal (Farnsworth et al. 2002), and even conventional capture–recapture when there are spatially indexed samples. As with distance sampling, modeling covariate effects on abundance has been a deficiency in modeling data collected under these sampling approaches and, consequently, the proposed model can be used to address this objective in a formal manner.

Our application of the proposed model to Ovenbird counts provides a compelling illustration of the potential management utility of being able to assess covariate effects within a distance-sampling framework. In particular, the model provides a concise linkage between measurable habitat structure and Ovenbird density. Consequently, it allows for evaluation of the effect of habitat-management actions on bird density. For example, the density of Ovenbirds increases as understory foliage cover increases (Fig. 1). We note that the duration of the point count should be an important consideration prior to data collection and the 12-min count duration used in our study is probably excessive. Abundance estimation methods based on point counting must be regarded as instantaneous in time. As sampling duration increases, the effect of "temporary emigration" (Kendall et al. 1997) will lead to some bias in mean density but will not lead to biased covariate effect estimates.

One referee was concerned about potential bias in density estimates obtained in our study due to potential heterogeneity as a result of habitat effects on detectability. That is, while we have biological reasons to suggest that these habitat covariates affect abundance, it is plausible that they also affect detectability. To evaluate this, we considered the distance class frequency of detections for each of three levels (low, moderate, high) of the BA (basal-area) covariate (the model with the best AIC score). The frequencies of detections in each of the four distance classes were (3,6,16,5), (4,7,11,9), and (0,3,7,6) for low, moderate, and high values of the BA covariate. A simple chi-square test of homogeneity yields $\chi^2 = 5.097$ (df = 6, $P = 0.531$), suggesting no difference between these distance-frequency distributions. We note also that our bootstrap goodness-of-fit procedure indicates that there is no significant lack of fit, suggesting that there is no additional heterogeneity, such as due to habitat effects on detection probability. Consistent with this point, note that the fitted negative-binomial models are essentially equivalent to the Poisson models. Finally, we note that detection covariates can be modeled easily in the half-normal distance function (Marques and Buckland 2004) and so this idea could conceivably be used to extend the model considered here to include covariate effects on both detection and density. However, we believe that the identifiability of habitat effects in both detection *and* abundance (i.e., of the same habitat variables) will be highly sensitive to model structure. Thus, when a particular covariate affects both detection and abundance, we believe that biologists will have to make a priori judgments about the most sensible way to partition variance and attempt to design surveys that minimize this possibility.

LITERATURE CITED

Agarwal, D. K., A. E. Gelfand, and S. Citron-Pousty. 2002. Zero-inflated models with application to spatial count data. Environmental and Ecological Statistics **9**:341–355.

Agresti, A. 2002. Categorical data analysis. Second edition. Wiley Interscience, Hoboken, New Jersey, USA.

Beadell, J., R. Greenberg, S. Droege, and J. A. Royle. 2003. Distribution, abundance, and habitat affinities of the coastal plain swamp sparrow. Wilson Bulletin **115**:38–44.

Bhattacharya, S. K., and M. S. Holla. 1965. On a discrete distribution with special reference to the theory of accident proneness. Journal of the American Statistical Association **60**:1060–1066.

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

Brand, L. A., and T. L. Georgea. 2001. Response of passerine birds to forest edge in coast redwood forest fragments. Auk **118**:678–686.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to distance sampling. Oxford University Press, New York, New York, USA.

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

Dixon, P. M. 2002. Bootstrap resampling. Encyclopedia of Environmetrics **1**:212–220.

Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. Auk **119**:414–425.

Fewster, R. M., S. T. Buckland, G. M. Siriwardena, S. R. Baillie, and J. D. Wilson. 2000. Analysis of population trends for farmland birds using generalized additive models. Ecology **81**:1970–1984.

Friesen, L. E., P. F. J. Eagles, and R. J. Mackay. 1995. Effects of residential development on forest-dwelling Neotropical migrant songbirds. Conservation Biology **9**:1408–1414.

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.

Hedley, S. L., and S. T. Buckland. 2004. Spatial models for line transect sampling. Journal of Agricultural, Biological and Environmental Statistics, *in press*.

Hedley, S. L., S. T. Buckland, and D. L. Borchers. 1999. Spatial modelling from line transect data. Journal of Cetacean Research and Management **1**:255–264.

Hovind, H. J., and C. E. Rieck. 1970. Basal area and point-sampling: interpretation and application. Technical Bulletin Number 23. Wisconsin Department of Natural Resources, Madison, Wisconsin, USA.

Hutto, R. L. 1995. Composition of bird communities following stand-replacement fires in Northern Rocky Mountain (U.S.A.) conifer forests. Conservation Biology **9**:1041–1058.

Ihaka, R., and R. Gentleman. 1996. R: a language for data analysis and graphics. Journal of Computational and Graphical Statistics **5**:299–314.

Kendall, W. L., J. D. Nichols, and J. E. Hines. 1997. Estimating temporary emigration using capture–recapture data with Pollock's robust design. Ecology **78**:563–578.

Laird, N. M., and J. H. Ware. 1982. Random-effects models for longitudinal data. Biometrics **38**:963–974.

Marques, F. F. C., and S. T. Buckland. 2003. Incorporating covariates into standard line transect analyses. Biometrics **59**:924–935.

Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk **117**:393–408.

Puig, P. 2003. Characterizing additively closed discrete models by a property of their maximum likelihood estimators, with an application to generalized Hermite distributions. Journal of the American Statistical Association **98**:687–692.

Ramsey, F. L., and K. Harrison. 2004. A closer look at detectability. Environmental and Ecological Statistics **11**:73–84.

Ramsey, F. L., V. Wildman, and J. Engbring. 1987. Covariate adjustments to effective area in variable-area wildlife surveys. Biometrics **43**:1–11.

Robbins, C. S., D. K. Dawson, and B. A. Dowell. 1989. Habitat area requirements of breeding forest birds of the Middle Atlantic States. Wildlife Monographs **103**:1–34.

Robinson, G. K. 1991. That BLUP is a good thing: the estimation of random effects (with discussion). Statistical Science **6**:15–51.

Rosenstock, S. S., D. R. Anderson, K. M. Giesen, T. Leukering, and M. F. Carter. 2002. Land-bird counting techniques: current practices and an alternative. Auk **119**:46–53.

Royle, J. A. 2004. N-mixture models for estimating population size from spatially replicated counts. Biometrics **60**:108–115.

Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence–absence data or point counts. Ecology **84**:777–790.

Sanathanan, L. 1972. Estimating the size of a multinomial population. Annals of Mathematical Statistics **43**:142–152.

Vernier, P. R., F. K. A. Schmiegelow, and S. G. Cumming. 2002. Modeling bird abundance from forest inventory data in the boreal mixed-wood forests of Canada. Pages 559–571 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.

Vincent, P. J., and J. M. Haworth. 1983. Poisson regression models of species abundance. Journal of Biogeography **10**:153–160.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, San Diego, California, USA.

Young, J. S., and R. L. Hutto. 1999. Habitat and landscape factors affecting Cowbird distribution in the Northern Rockies. Studies in Avian Biology **18**:41–51.

Young, J. S., and R. L. Hutto. 2002. Use of a landbird monitoring database to explore effects of partial-cut timber harvesting. Forest Science **48**:373–378.

STATISTICAL REPORTS