Contributed Paper

The Importance of Being Spatial (and Reserved): Assessing Northern Spotted Owl Habitat Relationships with Hierarchical Bayesian Models

CARLOS CARROLL* AND DEVIN S. JOHNSON†

*Klamath Center for Conservation Research, Orleans, CA 95556, U.S.A., email carlos@klamathconservation.org
†National Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA Fisheries, 7600 Sand Point Way NE, Seattle, WA 98115, U.S.A.

Abstract: *Regional conservation planning increasingly draws on habitat suitability models to support decisions regarding land allocation and management. Nevertheless, statistical techniques commonly used for developing such models may give misleading results because they fail to account for 3 factors common in data sets of species distribution: spatial autocorrelation, the large number of sites where the species is absent (zero inflation), and uneven survey effort. We used spatial autoregressive models fit with Bayesian Markov Chain Monte Carlo techniques to assess the relationship between older coniferous forest and the abundance of Northern Spotted Owl nest and activity sites throughout the species' range. The spatial random-effect term incorporated in the autoregressive models successfully accounted for zero inflation and reduced the effect of survey bias on estimates of species–habitat associations. Our results support the hypothesis that the relationship between owl distribution and older forest varies with latitude. A quadratic relationship between owl abundance and older forest was evident in the southern portion of the range, and a pseudothreshold relationship was evident in the northern portion of the range. Our results suggest that proposed changes to the network of owl habitat reserves would reduce the proportion of the population protected by up to one-third, and that proposed guidelines for forest management within reserves underestimate the proportion of older forest associated with maximum owl abundance and inappropriately generalize threshold relationships among subregions. Bayesian spatial models can greatly enhance the utility of habitat analysis for conservation planning because they add the statistical flexibility necessary for analyzing regional survey data while retaining the interpretability of simpler models.*

Keywords: Bayesian inference, focal species, habitat relationships, Northwest Forest Plan, spatial autoregressive model, species distribution model, Spotted Owl, *Strix occidentalis*

La Importancia de Ser Espacial (y Reservado): Evaluación de las Relaciones del Hábitat del Búho *Strix occidentalis* con Modelos Bayesianos Jerárquicos

Resumen: *Cada vez más, la planificación regional de la conservación utiliza modelos de aptitud de hábitat para sostener las decisiones relacionadas con la adjudicación y manejo de tierras. Sin embargo, las técnicas estadísticas utilizadas comúnmente para desarrollar dichos modelos pueden producir resultados engañosos porque no toman en cuenta 3 factores comunes en los conjuntos de datos de la distribución de especies: autocorrelación espacial, la gran cantidad de sitios donde la especie está ausente (inflación cero) y un esfuerzo de muestreo desigual. Utilizamos modelos espaciales autoregresivos adaptados con técnicas Bayesianas Cadena Markov Monte Carlo para evaluar la relación entre el bosque de coníferas viejo y la abundancia de nidos del búho* Strix occidentalis *y sitios de actividad en el área de distribución de la especie. El término espacial de efecto aleatorio incorporado a los modelos autoregresivos explicó la inflación cero exitosamente y redujo el efecto del sesgo de muestreo sobre estimaciones de las asociaciones especie-hábitat. Nuestros resultados sustentan la hipótesis de que la relación entre la distribución del búho y el bosque viejo varía con la latitud.*

1026

Una relación cuadrática entre la abundancia de búhos y el bosque viejo fue evidente en la porción sur del área de distribución y una relación pseudo umbral fue evidente en la porción norte. Nuestros resultados sugieren que los cambios propuestos a la red de reservas de hábitat de S. occidentalis *reducirían, hasta en un tercio, la proporción de la población protegida, y que las directrices propuestas para el manejo del bosque dentro de las reservas subestiman la proporción de bosque viejo asociado con la máxima abundancia de búhos e inadecuadamente generalizan relaciones umbral entre las subregiones. Los modelos Bayesianos espaciales pueden resaltar la utilidad del análisis de hábitat para la planificación de la conservación porque agregan la flexibilidad estadística necesaria para analizar datos de muestreos regionales y mantienen la interpretabilidad de modelos más simples.*

Palabras Clave: búho, especies focales, inferencia Bayesiana, modelos de distribución de especies, Plan Forestal del Noroeste, relaciones del hábitat, *Strix occidentalis*

Introduction

Habitat models aid conservation planning because they increase understanding of factors limiting species distribution, facilitate the development of guidelines for habitat protection, and help evaluate suitability of unsurveyed or currently unoccupied areas. Nevertheless, statistical techniques commonly used for developing such models may give misleading results because they fail to account for factors common in data sets of species distribution: spatial autocorrelation, the large number of sites where the species is absent (zero inflation), and uneven survey effort. To address these factors and develop an example of new techniques with wide applicability in habitat modeling, we applied hierarchical Bayesian spatial models to evaluate habitat relationships and conservation options for the Northern Spotted Owl (*Strix occidentalis caurina*), which has served as a de facto "umbrella" species for conserving the community of species associated with older forests in the Pacific Northwest (USA) (Noon & Blakesley 2006).

The Northern Spotted Owl (henceforth referred to as "the owl") was listed in 1990 as a threatened species under the U.S. Endangered Species Act (ESA) due to declining population trends related to the loss of older coniferous forest habitat to timber harvest (USFWS 1992). The Northwest Forest Plan (NFP), initiated in 1994, sought to ensure viable populations of the owl and other old-growth–associated species by coordinating regional habitat management across multiple ownerships encompassing the range of the owl within the United States (Noon & Blakesley 2006). The NFP based its guidelines for the size and spacing of habitat reserves primarily on simulation models of owl viability that elucidated general reserve-design rules without reference to data on the current distribution of older forest habitat (Noon & Blakesley 2006).

Before the advent of the NFP, forested federal lands in the region were divided between congressional reserves (e.g., parks and wilderness areas) and nonreserved lands generally open to timber harvest. With the listing of the owl, previously unreserved lands that were thought to be necessary for owl persistence were designated "critical habitat" as required under the ESA. The NFP subsequently created an additional management category, late-successional reserves, of similar total area as congressional reserves, where only limited timber harvest was allowed. Subsequent recovery plans propose replacing the NFP's management designations with a modified reserve network (USFWS 2007). Two alternative reserve networks, managed owl conservation areas (option 1) and habitat blocks (option 2), are under consideration (USFWS 2007).

Although the owl's survival and fecundity has been examined within intensive demographic study areas (Franklin et al. 2000; Olson et al. 2004; Dugger et al. 2005; Anthony et al. 2006), the relationship between habitat and owl persistence at broader spatial scales remains a subject of debate, due in part to uncertainty regarding the effects of latitudinal variation in prey community composition. Most researchers have found that annual survival rate is positively correlated with the amount of old-growth forest within the vicinity of nest sites (Noon & Blakesley 2006). Nevertheless, results of studies in the southern portion of the range, where woodrat (*Neotoma fuscipes*) is the primary prey, suggest that edge habitat between early- and late-seral forest stands may increase woodrat abundance, prey availability to owls, and owl fecundity (Franklin et al. 2000). This relationship may not occur in the northern portion of the range, where old-forest–associated prey species such as northern flying squirrel (*Glaucomys sabrinus*) dominate the owl's diet (Forsman et al. 1984). Recent recovery planning for the owl proposes major changes in land allocation and management guidelines on the basis of the assumption that the role of a territory as source habitat is maximized throughout the owl's range at intermediate proportions of older forest (USFWS 2007). Because demographic data are too expensive to collect throughout the range of a species, rangewide habitat models are a key tool with which to evaluate this assumption and compare the efficacy of alternate reserve allocations in protecting high-quality habitat.

Based on a range-wide analysis, vegetation in the 2-ha neighborhoods surrounding owl nest sites have higher conifer tree size and canopy closure, greater complexity of stand structure, and lower canopy closure of deciduous trees than the landscape as whole (Davis & Lint 2005). Nevertheless, statistical techniques commonly used for developing species distribution models, such as logistic regression and the Biomapper software used by Davis and Lint (2005), may be poorly suited for the analysis of spatially autocorrelated survey data. Sites that are near one another often have more-similar abundance values than widely separated sites, due to biological processes such as dispersal that tend to result in aggregated distributions or to spatial autocorrelation of environmental factors (Clark 2007). Such spatially autocorrelated data violate the assumption of independence in standard statistical tests, which could lead to inclusion of nonsignificant variables, poor interpretability of model structure and coefficients, and a lack of general applicability to new areas or novel future conditions.

Spatial autoregressive models attempt to overcome these problems by fitting a model that may contain both environmental covariates and a spatial random effect (Gelman et al. 2004). The presence of a spatial random effect component (ρ) allows response values at a site to depend on the sites' characteristics and the influence of values at neighboring sites, thus incorporating local habitat factors and spatial (e.g., metapopulation) effects (Latimer et al. 2006). Such spatial models also help address problems of survey bias inherent in "found" data (i.e., collected without an overall randomized or systematic sampling design). Spatial autoregressive models are generally more robust to geographical survey bias because such bias can be incorporated within the ρ term, thus reducing its influence on estimation of the effects of environmental variables. Survey data also often show a higher proportion of sites with zero values (where the species is absent) than is expected by standard statistical distributions (such as the Poisson for count data) (Martin et al. 2005). This may occur when suitable habitat is rare or the species does not occupy all suitable habitats due to stochastic processes. Alternate models such as the negative binomial and zero-inflated Poisson have been used to avoid problems of poor fit and underestimation of model uncertainty with zero-inflated data (Martin et al. 2005).

Spatial autoregressive models, although challenging for standard statistical software, can be fit with Bayesian Markov Chain Monte Carlo (MCMC) techniques (e.g., as implemented by WinBUGS [Spiegelhalter et al. 2003]). Nevertheless, MCMC models carry substantial costs in computational time and complexity. By comparing spatial models with nonspatial analogues, we sought to provide guidance as to when such spatial modeling is necessary and what additional information is provided by such models.

Methods

Study Area and Survey Data

The study area encompassed the U.S. range of the Northern Spotted Owl, which stretches from the Cascade Range to the Pacific Ocean in Washington and Oregon and covers portions of northwestern California (Noon & Blakesley 2006). We excluded the extreme southern portion of the range for which we lacked vegetation data. We derived locations of owl nest sites or activity centers (henceforth owl sites) from digital databases provided by federal and state agencies (Davis & Lint 2005; California Department of Fish and Game [CDFG], unpublished data). Locations had been collected from the late 1980s through 2000 for Oregon and Washington and from 1974 through 2006 for California. Standardized survey protocols were used to collect data to determine owl occupancy and reproductive status, with 71.7, 22.0, and 6.3% of the owl sites located within U.S. Forest Service, Bureau of Land Management, and other public and private lands, respectively (USFS 1988). Although survey effort was among the largest for any ESA-listed species, surveys occurred primarily during the planning process for activities such as timber harvest that potentially would disturb owls, on private lands and those federal lands open to such activities (i.e., lands not within parks or wilderness areas), and within long-term demographic study areas established as part of the NFP's monitoring program (Anthony et al. 2006). Survey data were available from all jurisdictions within California. In Oregon and Washington, the data set was as used in Davis and Lint (2005), with the exception that sites from lands not within Forest Service or Bureau of Land Management land (approximately 10% of the locations) were not available due to confidentiality restrictions. This data gap was addressed by inclusion of a variable representing management jurisdiction as described later.

Vegetation Data

We used data on conifer forest age class developed through classification of Landsat Thematic Mapper satellite imagery (Strittholt et al. 2006). We structured our set of candidate models to provide comparability with earlier research that addressed the relationship between owl occurrence and old-growth habitat (Zabel et al. 2003). Our predictor variables were the proportion of old conifer forest (>150 years) and mature conifer forest (50–150 years) within an analysis unit. This proportion was measured for habitat-capable lands only, following the methodology of Davis and Lint (2005), which excluded from consideration lands that could not sustain closed-canopy conifer forest (e.g., barren rock or serpentine) or above the elevation zone inhabited by the owl. We did not explore candidate models containing a larger set of environmental variables (e.g., climate) because we wished to

maintain comparability with previous work (Zabel et al. 2003; Davis & Lint 2005) and assess relationships between owl distribution and older forest with a highly interpretable and generalizable model.

Modeling Methodology

Wildlife distribution models are commonly based on a form of generalized linear model (GLM; McCullagh & Nelder 1989). Counts such as the owl site data we used can be modeled with a GLM with a log link, the Poisson model. Spatial autoregressive models build on the GLM framework through the addition of a spatial random effect, ρ. One form of spatial autoregressive model, the intrinsic conditional autoregressive (ICAR) model, makes the simplifying assumption that the spatial random effect in cell i depends only on the neighboring cells of i (N_i), and that all such neighbors have equal influence (weight of 1) (Gelman et al. 2004). The model is defined as follows. If y_i is the measured distribution variable (here abundance) at site i, then y_i is distributed as a Poisson variable with mean $\lambda_i = \exp\{\eta_i\}$, where $\eta_i = \beta_0 + \beta_1 x_{1i} \times \ldots + \beta_p x_{pi} + \rho_i$. The variables $x_{1i}, \ldots, x_{pi}$ are site-specific environmental covariates and the ρ_i are jointly distributed as a multivariate normal ICAR spatial model. The ICAR is defined by the conditional normal distributions $\rho_i \sim N(\mu_i, \sigma^2/n_i)$, where

$$\mu_i = \frac{1}{n_i} \sum_{j \in N_i} \rho_j$$

and n_i is the number of adjacent cells. Because we used a first-order neighborhood and hexagonal cells, N_i indexed the 6 immediately adjacent cells. We tested the adequacy of the first-order neighborhood by testing residuals from spatial and nonspatial models for significant remaining spatial autocorrelation with the Moran I statistic.

We used the program WinBUGS (version 1.4.2, Spiegelhalter et al. 2003) to fit both spatial ICAR and nonspatial GLM analogues (see Supplementary Material for BUGS code). The WinBUGS simulations used 3 chains, each with a burn-in period of 10,000 iterations followed by 40,000 iterations for estimation. We evaluated convergence with the Brooks–Gelman–Rubin statistic (Spiegelhalter et al. 2003). We assumed no preexisting knowledge of model coefficients and thus specified "vague" or uninformative priors (Gelman et al. 2004).

We adapted the modeling framework of Latimer et al. (2006; models 2 and 4), which overlays a regular grid of cells on the study area and considers each cell a sample unit potentially containing a number of survey locations. Our sample unit grid was derived from the grid of 24 km^2 hexagons used by the Forest Inventory and Analysis program (FIA; http://fia.fs.fed.us/) to monitor forest structure and plant communities at broad scales. The FIA hexagons form a seamless regular lattice across the contiguous United States. This modeling framework reduces

problems arising from uneven sampling intensity or survey effort, reduces the number of sample units to a level (here <4000) that makes regional-scale ICAR modeling computationally feasible, facilitates parameterization of spatial neighborhood effects, and makes results comparable with other species distribution models that are based on this sampling grid.

Because we did not have access to survey data from nonfederal lands in Oregon and Washington, we added an additional hierarchical level to the model structure that allowed incorporation of a variable representing the proportion of such federal lands within a hexagon. Although an imperfect surrogate for survey effort, this variable allowed us to assess alternate models that incorporated known sources of survey bias. Survey effort (relative area or RA) modified the predicted abundance term such that the observed Poisson mean $\lambda^{\text{obs}}_i = \lambda_i \times \exp\{\alpha \times \log(RA_i)\}$. The addition of survey effort to a second hierarchical level, rather than at the same level as the habitat covariates, allows one to easily extract estimates of abundance under uniform survey effort (Latimer et al. 2006 [model 4]; Supplementary Material S3).

For those models that incorporated a survey effort variable, proportion of old-growth and mature forest was measured for the surveyed area alone (FS and BLM lands), rather than for the cell as a whole, to avoid biasing estimates of the coefficients of environmental variables. With one exception, the 2275 cells with <1.5% of their area surveyed showed no owl site records. Therefore, we reduced computational time by categorizing these cells as unsurveyed (i.e., as lacking response data in the simulations). In our data set, 51.3, 19.5, and 29.2% of the 5501 cells with survey effort above this level contained 0, 1, and 2 or more owl sites, respectively. Among surveyed cells, mean number of owl sites was 1.42. To assess potential geographic variation in owl habitat relationships, we divided the analysis area into 3 subregions on the basis of contrasting prey communities and a strong dispersal barrier (the Columbia River): (1) southern, northwestern California/southwestern Oregon, where woodrats dominate the owl diet (Forsman et al. 1984; USFWS 2007), (2) central, the remainder of Oregon, and (3) northern, Washington.

We first selected an a priori set of candidate models of 9 alternate model structures to be evaluated with both ICAR and nonspatial structures (Supplementary Material S1). Models representing linear, pseudothreshold, and quadratic relationships were evaluated with variables for proportion of old-growth forest only (OG), combined old-growth and mature forest (MAT), and separate old-growth and mature forest variables. The pseudothreshold model was structured as a logarithmic relationship, for example, $\beta_{1i} \times \log((x_i) + 1 \times 10^{-8})$ (Franklin et al. 2000). We fitted candidate models in WinBUGS and ranked competing models by their deviance information criterion (DIC) value (Spiegelhalter et al. 2003). In

addition to the use of DIC, we sought the model with lowest mean-squared predictive error (MSPE; Gelfand & Ghosh 1998) and with a posterior predictive *p* value (PPPV) between 0.05 and 0.95, which indicated that it was not unlikely that the model predictions were drawn from the same distribution as the observed data (Gelman et al. 2004). Because zero inflation is common in count data from surveys of rare species (Martin et al. 2005), we compared the relative fit of the best model selected from the set of 9 standard Poisson ICAR models with analogous models from 2 alternative distributions, negative binomial and zero-inflated Poisson (ZIP).

We interpreted the coefficients of best spatial models and the spatial patterns of model results (predicted abundance, spatial random effect (ρ), and predicted abundance with the ρ term removed) in the light of previous studies of the ecology of the species. Evaluating predicted abundance without the ρ term allowed us to draw conclusions about species–habitat relationships that are more robust to the influence of survey bias or unmeasured environmental variables. To remove the influence of survey effort on predictions of abundance without the ρ term, we also recalculated predictions with relative area set to 1 (100%) and with proportion of old-growth and mature forest measured for the cell as a whole, rather than for the surveyed area alone.

Lastly, we assessed the proportion of predicted owl abundance (with the ρ term removed) and total Biomapper habitat score (Davis & Lint 2005) captured by 5 alternate reserve network proposals: 1992 and 2007 critical habitat (USFWS 1992, 2007), the NFP's late-successional reserves, and options 1 and 2 of the 2007 draft Northern Spotted Owl recovery plan (USFWS 2007). Although

unlike managed owl conservation areas, habitat blocks would be delineated by regional managers on the basis of rules governing size and spacing after adoption of the recovery plan, we used reserve boundaries from an example of the habitat-block reserve network published in the draft recovery plan (USFWS 2007).

Results

Spatial Poisson Models

Poisson ICAR models with a variable representing survey effort (proportion of cell surveyed) performed better than those without this variable in all subregions (Supplementary Material S1). Representing survey effort in a second hierarchical level of the model, rather than on the same level as habitat covariates, also increased model fit (ΔDIC of 23.8–35.6 for best subregional models, C.C., unpublished data). Nevertheless, selection of the best model from among the 9 candidate models was consistent between model sets that did or did not contain the survey effort variable. The structure of the best model (e.g., linear, pseudothreshold, or quadratic) differed between the southern subregion and other subregions (Table 1). A quadratic model based on the combined proportion of old-growth and mature forest (model 8; Supplementary Material S1) showed the lowest DIC in the southern subregion (northwestern California and southwestern Oregon) (Fig. 1). In the central (northern Oregon) and northern (Washington) subregions, the best model contained a pseudothreshold relationship between owl site abundance and the proportion of old-growth and mature forest, represented separately (Fig. 2b; model 6,

Table 1. Model coefficients for best (1) and closest competing models (2) for the relationship between Northern Spotted Owl site occurrence and old-growth (OG) and mature (MAT) forest within 3 subregions in the Pacific Northwest.*

				Model component			
		intercept	*OG*	*MAT*	*OG + MAT*	*α*	*DIC*
Southern							
(1) quadratic: (og + mat)	mean	−1.739			5.056, −2.666	0.873	4931.0
	SD	0.158			0.674, 0.689	0.116	
(2) quadratic: (og, mat)	mean	−1.446	4.575, −3.486	2.873, −1.448		0.872	4934.5
	SD	0.121	0.769, 1.454	0.760, 1.204		0.112	
Central							
(1) threshold: (og, mat)	mean	1.547	0.511	0.295		0.962	5425.3
	SD	0.178	0.084	0.094		0.072	
(2) threshold: (og + mat)	mean	0.961			0.807	0.991	5430.4
	SD	0.118			0.108	0.071	
Northern							
(1) threshold: (og + mat)	mean	0.573			1.271	1.364	2455.7
	SD	0.181			0.143	0.123	
(2) threshold: (og, mat)	mean	1.128	0.914	0.206		1.316	2464.2
	SD	0.239	0.132	0.104		0.122	

Mean estimates from the Markov Chain Monte Carlo simulations are followed by standard deviations (SD). The variable α modifies the influence of survey effort on observed counts. DIC is deviance information criterion. Sample size (n) = 2151, 2827, and 3248 for the southern, central, and northern subregions, respectively.
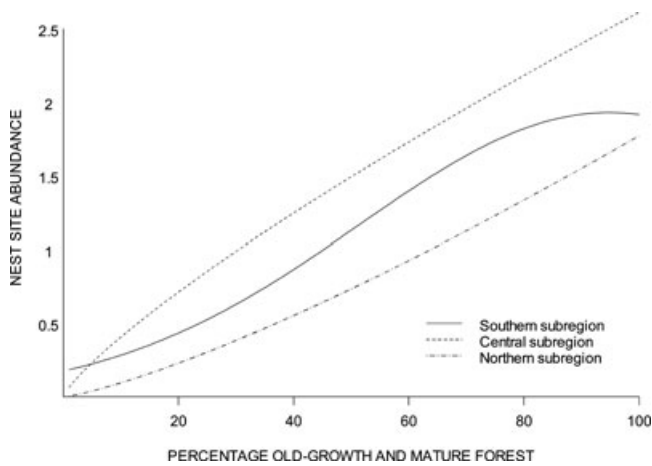
Figure 1. Predicted response curves of the relationship of Northern Spotted Owl site abundance to combined proportion of old-growth and mature forest within 24-km² cells. Response curves for the 3 geographic subregions are based on the following models: southern, model 8; central, model 5; northern, model 5 (which are [a], [b], and [c], respectively, in Fig. 2). We identified a model with a single, combined old-growth/mature forest variable as the best model in the southern and northern subregions and the closest competing model in the central subregion.

Supplementary Material S1) or as a combined variable (Fig. 1; model 5, Supplementary Material S1), in the central and northern subregions, respectively. We used the set of Poisson ICAR models that contained survey effort as the base set of models from which to draw comparisons. Within this base set, model 8 (southern subregion), model 6 (central subregion), and model 5 (northern subregion) are referred to as the selected set of models (Table 1).

Nonspatial Poisson models exhibited much poorer performance (higher DIC) than Poisson ICAR models in all subregions, but especially in the central and northern subregions (Table 2). The rank of models was relatively robust between spatial models and their non-spatial analogues. Although the magnitude of the coefficient for old-growth was similar between analogous spatial and nonspatial models (model 6), the coefficient for mature forest was smaller in the nonspatial versions for the northern and central subregions. Residuals from non-spatial Poisson models showed significant spatial auto-correlation (p value < 0.01 for Moran *I* statistics) at all lag distances. Residuals from Poisson ICAR models generally showed no significant spatial autocorrelation. The exception, between hex cells at the largest lag distances in the northern subregion, is likely due to edge effects arising from the disjunct distribution of surveyed lands in Washington.

Performance of Poisson, Negative Binomial, and ZIP Models

Nonspatial ZIP and negative binomial models performed better than nonspatial Poisson models, as judged by DIC and the ability to approximate the observed proportion of cells in the 3 abundance classes (cells with 0, 1, and 2 or more sites; Supplementary Material S2). Nevertheless, the Poisson ICAR models performed better than either ZIP or negative binomial ICARs. The ZIP ICARs performed better than negative binomial ICARs.

Nonspatial Poisson models greatly underestimated the number of cells with zero sites (Supplementary Material S2). In contrast, Poisson ICARs predicted a similar number of cells with 0 sites as was observed, but over-predicted the number of cells with 1 site and underpre-dicted the number of cells with 2 or more sites. Nevertheless, predictions from the negative binomial and ZIP ICARs also differed from the observed data in this respect. The overdispersion coefficients were 1.32, 2.51, and 3.79 for the southern, central, and northern subregions, respectively, in the nonspatial negative binomial
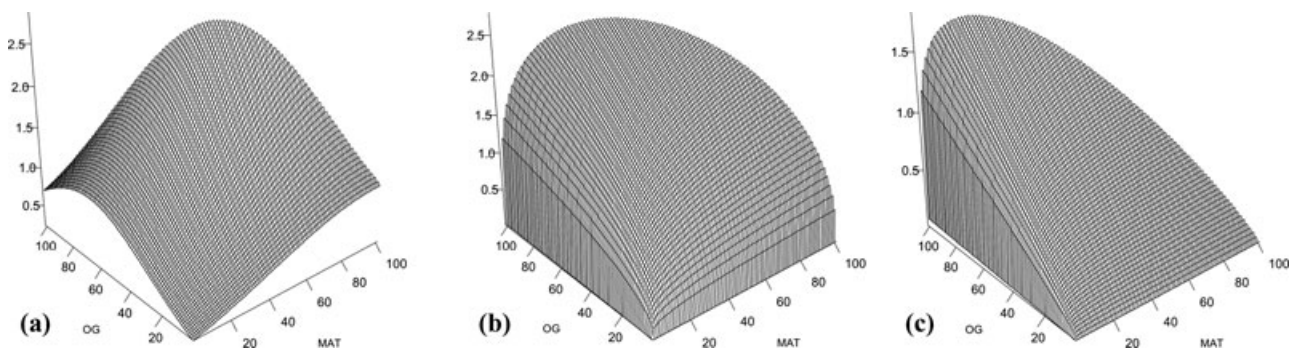
Figure 2. Predicted response curves of the relationship of Northern Spotted Owl site abundance (z-axis) to proportion of old-growth (OG) (y-axis) and mature (MAT) forest (x-axis) within 24-km² cells. Response curves for the 3 geographic subregions are derived from the following models: (a) southern, model 9; (b) central, model 6; (c) northern, model 6. We identified a model with distinct variables for old-growth and mature forest as the best model in the central subregion and the closest competing model in the southern and northern subregions.

	Subregion		
	southern	*central*	*northern*
Model			
Poisson ICAR	4931.0	5425.3	2455.7
Poisson nonspatial	5685.8	10079	4579.9
negative binomial ICAR	4940.1	5548.4	2559.3
negative binomial nonspatial	5575.7	6757.0	3160.2
ZIP ICAR	4933.9	5524.2	2529.9
ZIP nonspatial	5620.2	9098.0	3893.9

Abbreviations: ICAR, intrinsic conditional autoregressive model; ZIP, zero-inflated Poisson model.

models. The overdispersion coefficients in the negative binomial ICARs (1.01, 1.03, and 1.12, respectively), however, approximated those from a standard Poisson model (overdispersion = 1). Similar contrasts were evident between the nonspatial and ICAR ZIP models. Estimates of q (probability of the observation at a cell being generated from a Poisson distribution) were 0.852, 0.680, and 0.448 in the nonspatial ZIP models, but were 0.997, 0.944, and 0.896 (for the southern, central, and northern subregions, respectively) in the ZIP ICARs.

Interpretation of Model Diagnostics, Coefficients, and Spatial Predictions

The DIC and MSPE suggested a generally similar ranking among the set of Poisson ICAR models (Supplementary Material S1). The best model as judged by DIC also consistently showed the PPPV closest to 0.50 (thus showing an optimal balance between over- and underfitting). Nevertheless, PPPV values for the selected model in the southern subregion were >0.95, suggesting that all candidate models were overfit. The DIC and MSPE rankings were also generally consistent within the set of nonspatial models. Nevertheless, PPPV was uniformly near zero, indicating lack of fit.

Among the base set of Poisson ICAR models, for models that contained old-growth and mature forest as separate variables, the coefficient for old-growth increased in magnitude from the southern to northern subregion. In all subregions the coefficient for old-growth was greater than that for mature forest, and this contrast increased from the southern to northern subregion (Table 1). Graphical comparison of the predicted response surface from the selected models across subregions (Figs. 1 & 2) showed the effect of contrasting model structures (quadratic [southern] vs. pseudothreshold [central and northern]) and contrasting treatment of old-growth and mature forest as 2 separate (central) or a combined (southern and northern) variable.

On the basis of the model's environmental component (excluding ρ and with survey effort set uniformly to 100%), predicted owl abundance declined with increasing latitude because of the contrast in models between the 3 subregions (Fig. 3a). Areas of high predicted abundance occurred in low to mid elevation valleys of the southern subregion and in the southern Oregon Coast Range and portions of the Oregon and Washington Cascades (Fig. 3a). Areas with strongly positive values of ρ were generally associated with demographic study areas (Fig. 3b).

Abundance within Alternate Reserve Networks

The 5 alternate reserve networks (1992 and 2007 critical habitat, late-successional reserves, managed owl conservation areas, and habitat blocks) contained between 17.8% and 28.7% of the region's federal land and between 22.0% and 35.8% of the total owl abundance derived from predicted abundance values that did not include ρ (Table 3). The 2 options proposed in the 2007 recovery plan—managed owl conservation areas and habitat blocks—reduced the proportion of predicted owl abundance captured by 20.3% and 32.7%, respectively, from that captured within the late-successional reserve network (Table 3). The 2007 critical habitat protected 38.0% less of the predicted owl abundance than the 1992 designation. Figures based on the percentage of habitat value predicted by the Biomapper model (Davis & Lint 2005) were comparable to those from this study except for Biomapper's more positive assessment of 2007 proposed critical habitat (Table 3).

Discussion

Role of Spatial Models

Spatial autocorrelation is a pervasive characteristic of data sets of species distribution, because of spatially autocorrelated environmental factors and biological processes such as dispersal that tend to result in aggregated distributions of individuals (Clark 2007). It may be especially prevalent for species such as the Northern Spotted Owl that are area and dispersal limited (Noon & Blakesley 2006) and thus show strong effects of population processes on distribution. Our analysis attempts to bridge the gap between rangewide nonspatial modeling of species distributions derived from found data and abundance modeling derived from mark–recapture data from intensive studies of limited geographic extent. Because ICAR models facilitate partitioning of the spatial component from environmental effects, they increase the generality of the habitat relationships described in the model and improve estimation of their coefficients.

Presence-only data, such as that used here, are more likely to contain survey bias than systematic presence
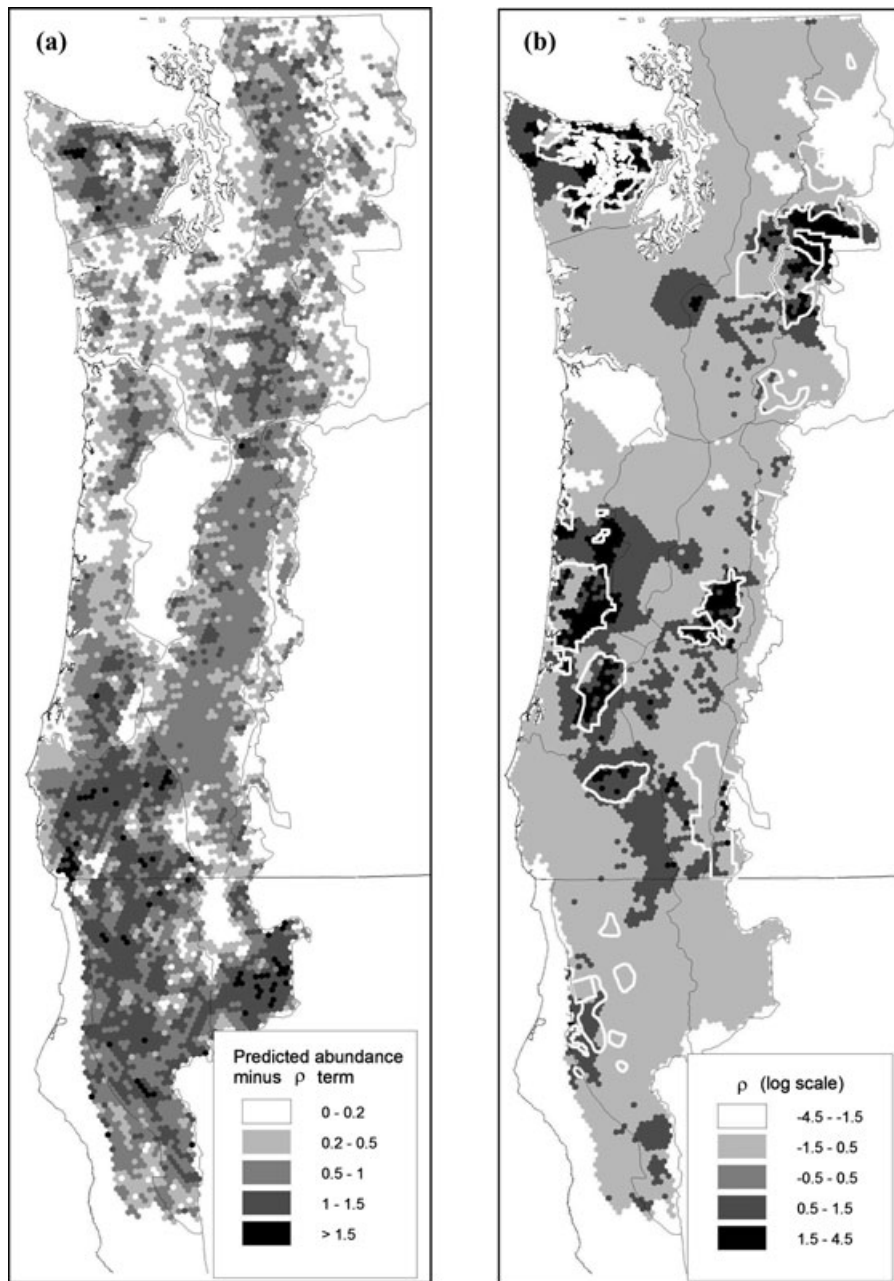
Figure 3. Map of predicted Northern Spotted Owl abundance based on a probability surface composed of the best model for each subregion: (a) predicted probability without spatial random effect (ρ) and (b) spatial random effect (ρ). In (b), demographic study areas are outlined in white.

and absence data sets. By allowing survey bias to be accounted for in the ρ term, our ICAR models provide a more robust surrogate for true abundance than would a nonspatial model developed from found data. A more complex hierarchical model incorporating a sampling model and a spatially varying "migrant pressure" model would be required to explicitly distinguish effects of unmeasured environmental variables, sampling bias, or population processes. Nevertheless, the strong pattern of higher ρ in demographic study areas (Fig. 3b) suggests that survey bias is a major component of the spatial random effect. Because owl site data from California had been preprocessed to remove some duplicate observations (CDFG, unpublished data) and thus reduce survey

bias, the contrast in fit between spatial and nonspatial models was less in that subregion. Although a model containing a binary variable representing demographic study areas could be evaluated post hoc, we sought here to demonstrate how the ICAR framework allows more flexible exploration of and increased robustness to spatial effects that are not included in the a priori model set. If the spatial random effect term is dominated by nuisance factors such as survey bias, probability values can be mapped without ρ to assess which elements of the distributional patterns are linked to known environmental variation.

The ICAR models we used are more difficult to construct and require more computational time than

Table 3. Percentage of the total Northern Spotted Owl population predicted to be contained within alternate proposed reserve networks, as predicted by a probability surface composed of the best model for each subregion and by habitat scores from a previous study in which the Biomapper model was used (Biomapper habitat score) (Davis & Lint 2005).[a]

	Percentage of area[b]	Predicted abundance[c]	Biomapper habitat score
1992 critical habitat	28.72/27.79*	35.48	35.77
2007 critical habitat	20.01/19.86	21.99	25.65
LSRs	27.06/26.65	32.68	32.53
MOCAs (option 1)	20.39/20.17	26.06	25.96
Habitat blocks (option 2)	17.78/17.61	21.99	22.27

[a]*Abbreviations: LSRs, late-successional reserves; MOCAs, managed owl conservation areas.*
[b]*Percentage of area differs slightly between predicted abundance (first value) and Biomapper habitat value calculations (second value) due to differing extents of the input vegetation data.*
[c]*Predicted abundance was calculated without the spatial random effect (ρ). Figures are for federal lands only, excluding congressionally reserved areas, which are reserved from timber harvest under all alternatives and represent 27% of federal land area, 18.83% of predicted abundance, and 19.53% of Biomapper habitat value.*

equivalent nonspatial models. The fact that model structure (e.g., linear, pseudothreshold, or quadratic) did not contrast strongly between the best spatial and nonspatial models might suggest that such investment is unnecessary. Nevertheless, the additional cost of spatial models is justified by the large contrast in the magnitude of coefficients of certain variables between spatial and nonspatial models and the greatly increased fit of spatial models. Authors of a recent review used simulated data and found that CAR models performed well when compared with alternative methods for modeling spatially autocorrelated species distribution data (Dormann et al. 2007). The large proportion of sites where a species is absent or not detected may also confound application of standard statistical methods (Martin et al. 2005). Although our data set had relatively little zero inflation, standard Poisson models were clearly inferior to the negative binomial and ZIP alternatives in a nonspatial context. Poisson ICAR models inherently incorporate overdispersion through their spatial random-effect term (Gelman et al. 2004). For our data this spatial structure was sufficient to obviate the need for additional remedies for zero inflation such as ZIP ICARs. Although Poisson ICAR models outperformed negative binomial and ZIP ICAR models in our results, this may not be true for all species, making it advisable to evaluate all 3 model structures when building spatial distribution models for rare species.

Although the DIC (Spiegelhalter et al. 2003) is the most commonly used model-selection criterion for Bayesian models, criteria derived from posterior predictive loss (MSPE and PPPV; Gelfand & Ghosh 1998; Gelman et al.

2004) may be a useful alternate diagnostic. In our results, both DIC and MSPE suggested a generally similar ranking of models, whereas PPPV provided an absolute rather than relative evaluation of overall model fit.

Owl Habitat Associations

As is typical of found data sets, we lacked comprehensive information on survey effort. We accounted for a portion of the geographic variation in survey effort with our "relative area of federal land" term. The Bayesian MCMC methodology allowed inclusion of survey effort as an additional hierarchical level in the model rather than simply as an additional covariate (Latimer et al. 2006; Clark 2007). The flexibility of the hierarchical Bayes structure thus facilitated a more straightforward interpretation of predicted abundance values and better model fit as compared with a single-level model. Encouragingly, model ranking was robust to inclusion of the survey-effort covariate. Although coefficients were of lower magnitude in models with versus without the survey-effort term, they showed similar contrasts between old-growth and mature forest and between regions. Thus, conclusions from models with a survey effort term may be conservative in their estimates of the strength of the owl's association with older forest.

Our results support the hypothesis of a more linearly increasing relationship between owl abundance and older forest in the northern than in the southern portion of the range. This may be due to contrast in prey communities between the southern subregion, where the woodrat is the primary prey species, and the central and northern subregions, where old-forest–associated species dominate the owl's diet.

In the southern subregion the best model predicted a quadratic relationship between owl sites and the combined proportion of old-growth and mature forest (Table 1). This could occur if edge habitat between early- and late-seral forest stands increases woodrat abundance and availability and owl fecundity (Franklin et al. 2000). The closest competing model estimated separate quadratic functions of old-growth and mature forest, with old-growth having a stronger effect than mature forest on owl abundance.

In the central subregion the best model predicted a pseudothreshold relationship between owl sites and old-growth and mature forest, estimated separately. This suggests that the habitat value of old-growth for owls is more distinguishable from that of mature forest than in the southern subregion and that there is no reduction in value in cells with the highest proportions of older forest. The closest competing model in this subregion was derived from a pseudothreshold function of the combined proportion of old-growth and mature forest (Table 1).

In the northern subregion the best model predicted a pseudothreshold relationship between owl sites and

the combined proportion of old-growth and mature forest. The closest competing model contained separate thresholds for old-growth and mature forest (Table 1). Although the best models were clearly superior to the closest competing models (ΔDIC of 3.5–8.5; Table 1), results more conclusively distinguished between model structures (quadratic or pseudothreshold) than between the habitat value of old-growth and mature forest (Supplementary Material S1). This may be because our mature-forest age class (50–150 years) includes some stands (e.g., those over 100 years in age) that are often included in the old forest age class in owl habitat studies (Dugger et al. 2005). In addition, interpretation of the predicted habitat associations for each subregion (Figs. 1 & 2) depends equally on the form of the best model (quadratic or pseudothreshold) and the coefficients of the model. For example, because the quadratic inflection in the model for the southern subregion occurred in landscapes with 95% old-growth and mature forest, it effectively portrays a threshold relationship at levels of greater than 80% old-growth and mature forest (Fig. 1).

Reserve Design Implications

Although our results support the hypothesis of latitudinal variation in owl habitat relationships, they do not support the draft recovery plan's generalization of a quadratic habitat relationship to the central and northern subregions (USFWS 2007). They also suggest that the recovery plan's goals for maintaining the proportion of "high-quality" habitat within reserves at 50% (in the south) to 70% (in the north) (USFWS 2007) may be too low to maximize owl abundance within reserves. In contrast, model results would suggest maintenance of over 80% of old-growth and mature age classes within owl conservation areas would be optimal for maximizing owl abundance in the southern subregion. Within the central and northern subregions, no such threshold is evident from our results because owl abundance was predicted to continuously increase with increasing proportion of old-growth and mature forest.

Our results also suggest that the 2 options proposed in the 2007 recovery plan would protect one-quarter to one-third fewer owls than would the current late-successional reserve network. At this scale, ICAR results were consistent with those from the Biomapper model (Table 3). Nevertheless, the ICAR model results provided a clearer interpretation of how habitat relationships vary latitudinally and of patterns of potential survey bias. Although results from our static models cannot directly address how habitat loss affects population viability, they do facilitate better parameterization of dynamic models that address this question. Given the negative mean annual population growth rate of 12 of the 13 demographic study areas (Anthony et al. 2006), maintaining or strengthening, rather

than reducing, the current reserve system appears advisable.

Spotted Owl abundance may be negatively affected by competition with its congeneric, the Barred Owl (*S. varia*), which has expanded its range from eastern North America in the last 40 years (Noon & Blakesley 2006). Although we did not have access to data on Barred Owl distribution of sufficient quality to use as a covariate in the models, model results can help generate hypotheses regarding their effects. The greater abundance of barred owls in the northern subregion (Noon & Blakesley 2006) may be linked to the Spotted Owl's lower abundance and greater association with old-growth forest there in our results.

Our results suggest that the significantly greater cost of spatial models in computational time and effort may be well justified for regional analyses of species distribution. Autoregressive models, by explicitly incorporating a spatial random effect term, provide a more flexible substitute for the trend surface variables and other surrogates for spatial trend commonly used in species distribution models and help account for the zero inflation (the large number of sample units where the species was absent) commonly found in such data without resorting to the use of more complex probability distributions. In habitat modeling, as more broadly in ecology (Clark 2007), the increased availability of flexible software for Bayesian analysis has the potential to transform our understanding of the relationship between species and their environment.

Supplementary Material

Model selection and diagnostic results for all candidate models (Appendix S1); the number of cells per abundance class for the selected set of Northern Spotted Owl habitat models with contrasting model structures (Appendix S2); and WinBUGS code for Poisson, negative binomial, and ZIP ICAR models (Appendix S3) are available as part of the on-line article from http://www.blackwell-synergy.com/. The author is responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

Literature Cited

Anthony, R. G., et al. 2006. Status and trends in demography of Northern Spotted Owls, 1985–2003. Wildlife Monographs **163:**1–48.

Clark, J. S. 2007. Models for ecological data: an introduction. Princeton University Press, Princeton, New Jersey.

Davis, R. J., and J. B. Lint. 2005. Habitat status and trend. Pages 21–82 in J. B. Lint, technical coordinator. Northwest forest plan—the first ten years (1994–2003): status and trend of Northern Spotted Owl populations and habitat. General technical report PNW-GRT-648. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, Oregon.

Dormann, C. F., et al. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography **30:**609–628.

Dugger, K. M, F. Wagner, R. G. Anthony, and G. S. Olson. 2005. The relationship between habitat characteristics and demographic performance of Northern Spotted Owls in southern Oregon. Condor **107:**863–878.

Forsman, E. D., E. C. Meslow, and H. M. Wight. 1984. Distribution and biology of the Spotted Owl in Oregon. Wildlife Monographs **87:**1–64.

Franklin, A. B., D. R. Anderson, R. J. Gutiérrez, and K. P. Burnham. 2000. Climate, habitat quality, and fitness in Northern Spotted Owl populations in northwestern California. Ecological Monographs **70:**539–590.

Gelfand, A. E., and S. K. Ghosh. 1998. Model choice: a minimum posterior predictive loss approach. Biometrika **85:**1–11.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2004. Bayesian data analysis. 2nd edition. Chapman & Hall/CRC Press, Boca Raton, Florida.

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.

Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Chey, A. J. Tyre, and H. Possingham. 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecology Letters **8:**1235–1246.

McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. 2nd edition. Chapman & Hall, New York.

Noon, B. R., and J. A. Blakesley. 2006. Conservation of the Northern Spotted Owl under the northwest forest plan. Conservation Biology **20:**288–296.

Olson, G. S., E. Glenn, R. G. Anthony, E. D. Forsman, J. A. Reid, P. J. Loschl, and W. J. Ripple. 2004. Modeling demographic performance of Northern Spotted Owls relative to forest habitat in Oregon. Journal of Wildlife Management **68:**1039–1053.

Spiegelhalter, D., A. Thomas, N. Best, and D. Lunn. 2003. WinBUGS user manual. Version 1.4. MRC Biostatistics Unit, Cambridge, United Kingdom. Available from http://www.mrc-bsu.cam.ac.uk/bugs (accessed November 2007).

Strittholt, J. R., D. A. Dellasala, and H. Jiang. 2006. Status of mature and old-growth forests in the Pacific Northwest. Conservation Biology **20:**363–374.

USFS (U.S. Forest Service). 1988. Spotted Owl inventory and monitoring handbook. USFS, Portland, Oregon, and San Francisco, California.

USFWS (U.S. Fish and Wildlife Service). 1992. Final draft recovery plan for the Northern Spotted Owl. USFWS, Portland, Oregon.

USFWS (U.S. Fish and Wildlife Service). 2007. Draft recovery plan for the Northern Spotted Owl (*Strix occidentalis caurina*). USFWS, Portland, Oregon.

Zabel, C. J., J. R. Dunk, H. B. Stauffer, L. M. Roberts, B. S. Mulder, and A. Wright. 2003. Northern Spotted Owl habitat models for research and management application in California (USA). Ecological Applications **13:**1027–1040.