



RESEARCH
PAPER

Spatial and temporal variations in species occurrence rate affect the accuracy of occurrence models

Lesley Bulluck^{1,*}, Erica Fleishman², Chris Betrus¹ and Rob Blair^{1,3}

¹Department of Zoology, Miami University, Oxford, Ohio 45056, USA, ²Center for Conservation Biology, Department of Biological Sciences, Stanford University, Stanford, CA 94305–5020, USA, ³Department of Fisheries, Wildlife, and Conservation Biology, 200 Hodson Hall, University of Minnesota, St. Paul, MN 55108, USA

ABSTRACT

Aim Predictive models of species occurrence have potential for prioritizing areas for competing land uses. Before widespread application, however, it is necessary to evaluate performance using independent data and effective accuracy measures. The objectives of this study were to (1) compare the effects of species occurrence rate on model accuracy, (2) assess the effects of spatial and temporal variation in occurrence rate on model accuracy, and (3) determine if the number of predictor variables affected model accuracy.

Location We predicted the distributions of breeding birds in three adjacent mountain ranges in the Great Basin (Nevada, USA).

Methods For each of 18 species, we developed separate models using five different data sets — one set for each of 2 years (to address the effects of temporal variation), and one set for each of three possible pairs of mountain ranges (to address the effects of spatial variation). We evaluated each model with an independent data set using four accuracy measures: discrimination ability [area under a receiver operating characteristic curve (AUC)], correct classification rate (CCR), proportion of presences correctly classified (sensitivity), and proportion of absences correctly classified (specificity).

Results Discrimination ability was not affected by occurrence rate, whereas the other three accuracy measures were significantly affected. CCR, sensitivity and specificity were affected by species occurrence rate in the evaluation data sets to a greater extent than in the model-building data sets. Discrimination ability was the only accuracy measure affected by the number of variables in a model.

Main conclusions Temporal variation in species occurrence appeared to have a greater impact than did spatial variation. When temporal variation in species distributions is great, the relative costs of omission and commission errors should be assessed and long-term census data should be examined before using predictive models of occurrence in a management setting.

Keywords

Accuracy measures, birds, Great Basin, Nevada, occurrence rate, receiver operating characteristic curves, species occurrence models.

*Correspondence: Lesley Bulluck, University of Tennessee, Department of Forestry, Wildlife and Fisheries, Knoxville, TN 37996, USA.
E-mail: lbulluck@utk.edu

INTRODUCTION

The accuracy of predictive models of species distributions has become a fundamental concern for both ecologists and land managers. Interest is due, in part, to the numerous direct applications of models that predict where a species is likely to be present or absent. Of particular interest are those applications with relevance to species conservation. The occurrence of rare and endangered species can be modelled with the aim of determining which land use practices are suitable for their

protection (Scott *et al.*, 1993; Caicco *et al.*, 1995) or of prioritizing areas for different land uses (Peterson *et al.*, 2000). Predictive models of occurrence can also be utilized in conjunction with potential restoration strategies to predict how certain species might respond to various revegetation scenarios (Fleishman & Mac Nally, 2004). Other applications of modelling species occurrences include predicting where invasive species are likely to colonize (Wadsworth *et al.*, 2000) and shifts in species distributions as a result of climate change (Aspinall & Matthews, 1994).

Predictive models have become popular because of the need for efficiency in the design and implementation of conservation management. Traditionally, conservation research has been based on the collection of detailed natural history data (Stauffer, 2002). More recently, methods in quantitative ecology have offered the potential to increase efficiency in the land use decision-making process. This change has been facilitated in part by the use of Geographic Information Systems (GIS) and computationally intensive statistical procedures. GIS allows for the management of spatially explicit data sets at state- or province-wide, regional, or continental scales (Boone & Krohn, 2000; Peterson, 2001), and thus affords the ability to develop and evaluate models of species distributions at multiple levels. In many parts of the world, landscape and regional scenarios will be the most useful to managers because land use decisions are often made at these spatial extents. To make land use decisions based on models of species distributions, one must assess the degree of spatial and temporal variations in a system and thereby understand the limits of a model's usefulness. A model built using data from one time or place may not be suitable for use in another time or place. However, if model predictions are focused on a location or time period similar to that used to develop the model, and the species being modelled are not exceptionally plastic in their habitat requirements, models of occurrence may be sufficiently reliable in many situations.

Predictive models are generally static, having no intrinsic ability to account for variation in the distribution of species or environmental variables in space and time (Guisan & Zimmerman, 2000). In contrast, ecosystems and the organisms inhabiting them are dynamic, such that prediction errors based on static models are inevitable (Fielding, 2002). The usefulness of static models in terms of relative reliability across space and time, however, should be assessed. Furthermore, the selection and presentation of accuracy measures for predictive models should incorporate the understanding that models are estimations of reality and are affected by natural variation.

Despite the widespread application of species distribution models, there has been little evaluation of model performance in a predictive capacity (Manel *et al.*, 2001). The few existing evaluations often simply state the proportion of sites in which the presence or absence of the species is predicted correctly. This measure of accuracy is a correct classification rate (CCR) and is calculated from a two-by-two classification table (Table 1). Measures of CCR can be misleading because they are affected by the relative occurrence rate (proportion of occupied sites) of the

species being modelled (Pearce & Ferrier, 2000a; Manel *et al.*, 2001). In addition, a species is predicted to be present or absent at a given site based on an arbitrarily selected probability threshold. For example, if the probability threshold is set at 0.5, then a species is predicted to be present whenever the predicted probability of occurrence is greater than or equal to 0.5 and to be absent whenever the predicted probability of occurrence is less than 0.5. This confounds assessment of model performance because every probability threshold is associated with a different level of accuracy (Pearce *et al.*, 2002; Fleishman *et al.*, 2003b).

Regardless of the probability threshold chosen, it can be helpful to separately assess a model's ability to predict presences and absences. This can be accomplished by calculating sensitivity and specificity from the values in a classification table for any given probability threshold. Sensitivity is the proportion of true positives, or sites correctly predicted as occupied (Table 1). Specificity is the proportion of true negatives, or sites correctly predicted as being unoccupied (Table 1).

Another measure of model performance is the area under a receiver operating characteristic (ROC) curve. In the past, ROC curves were used primarily in medical applications, but recently, they have been used in ecological studies (Manel *et al.*, 2001). ROC curves are excellent accuracy measures because they are not based on arbitrary thresholds and are less sensitive to occurrence rate than other commonly used evaluation methods (Pearce & Ferrier, 2000a; Manel *et al.*, 2001; Pearce *et al.*, 2002). The area under an ROC curve (AUC) provides a measure of discrimination ability (i.e. how well a model can distinguish between occupied and vacant sites) (Pearce & Ferrier, 2000b). Indeed, Hanley and McNeil (1982) demonstrated that an AUC value could be interpreted as a probability that a model will distinguish correctly between occupied and unoccupied sites. The ROC curve plots sensitivity against the proportion of false negatives (i.e. 1-specificity) for several probability thresholds (Pearce & Ferrier, 2000a). AUC values range from 0.5 (no discrimination ability) to 1.0 (perfect discrimination). Values greater than 0.9 indicate excellent discrimination ability; values between 0.7 and 0.9 indicate a usable, reasonable level of discrimination; and values between 0.5 and 0.7 indicate poor discrimination ability (Pearce & Ferrier, 2000a).

Study system

The Great Basin is a useful region within which to develop and evaluate models of species distribution based on vegetation and topography. The Great Basin includes more than 200 mountain ranges, most in a north-south orientation. These ranges were isolated from each other and the intervening valleys as the region became warmer and drier following the Pleistocene. Hence, for many species, these mountain ranges represent distinct habitat islands in the Great Basin landscape (Brown, 1978; Grayson, 1993). The numerous canyons that drain the east and west slopes of these mountain ranges may also act as habitat islands for some species due to the riparian vegetation and the resources they harbour relative to the intervening uplands (McDonald & Brown, 1992).

Table 1 Classification of the observed and predicted occurrences of a species. $A + B + C + D$ is equal to the total number of observations in the evaluation data set. Correct classification rate (CCR) = $(A + D) / (A + B + C + D)$. Sensitivity = $A / (A + C)$. Specificity = $D / (B + D)$

	Observed present	Observed absent
Predicted present	A (true positive)	B (false positive)
Predicted absent	C (false negative)	D (true negative)

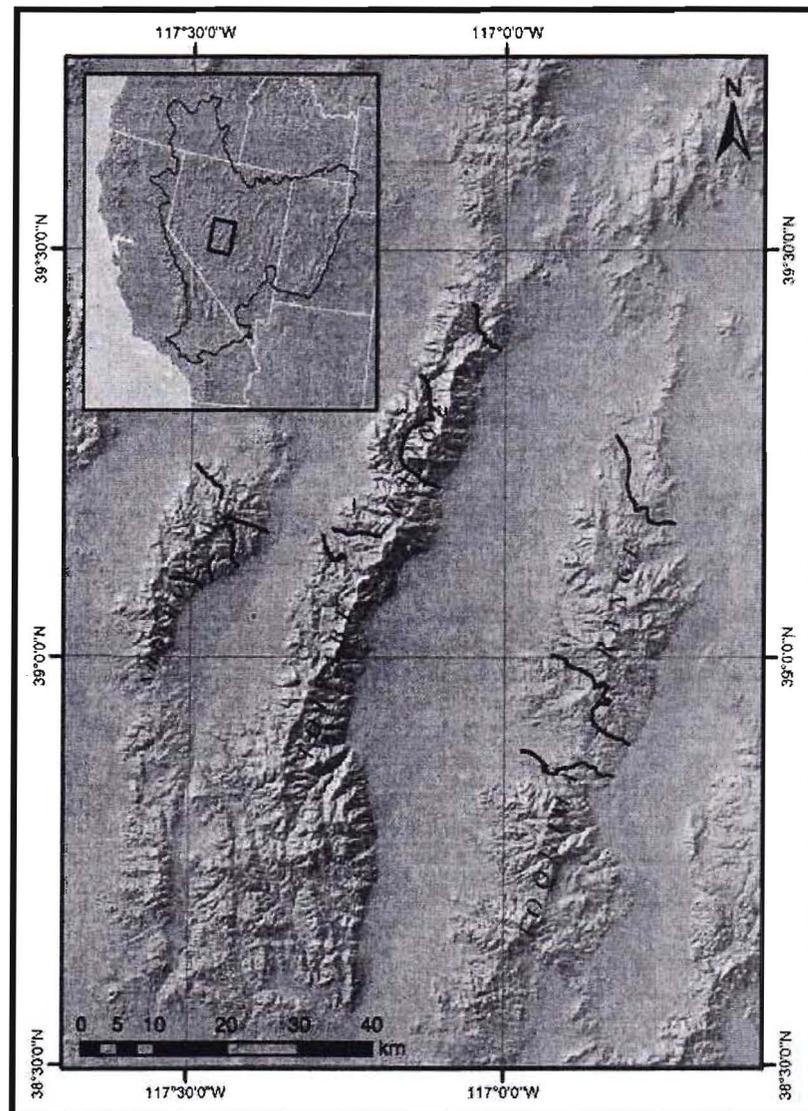


Figure 1 Location of the Shoshone Mountains, Toiyabe Range and Toiquima Range (black rectangle, see inset) in the Great Basin (black border, see inset). The inventoried canyons within the three mountain ranges are illustrated with black lines. Three pairs of canyons in the Toiquima Range and one pair of canyons in the Toiyabe Range connect at the crest of these ranges.

Three adjacent mountain ranges, the Toiquima Range, Toiyabe Range and the Shoshone Mountains (Lander and Nye counties, Nevada), were the focus of our study (Fig. 1). These ranges are in the same biogeographical subregion (Austin & Murphy, 1987; Grayson, 1993) and contain similar biotic communities. However, differences in elevation and precipitation among mountain ranges lead to differences in land cover (Grayson, 1993). For example, the Toiyabe Range has higher mean elevation and receives more precipitation annually than the other two ranges. Consequently, the canyons that drain its east and west slopes have more riparian vegetation than the relatively dry canyons of the Toiquima Range and the Shoshone Mountains. Thus, these three mountain ranges offer an opportunity to test hypotheses regarding the performance of predictive models of species occurrence when faced with spatial variation in the distribution of land cover.

Fleishman *et al.* (2001, 2003b) constructed butterfly occurrence models using data from the Toiquima Range and evaluated the models in the Shoshone Mountains. These butterfly models

predicted both presences and absences fairly well (Fleishman *et al.*, 2001, 2003b), suggesting that the spatial applicability of the models extended beyond the mountain range in which they were developed. Predictions of absence consistently had higher success rates than predictions of presence, which is often the case for species with occurrence rates of < 0.5 . We expected that models for breeding birds with occurrence rates similar to butterflies, constructed and evaluated within the same region, would also yield good results. We anticipated this success for two reasons. First, the species composition of Great Basin birds exhibits a distinct stratification across an elevational gradient (Medin *et al.*, 2000), making their distributions relatively easy to predict based on topographic variables in addition to vegetation. Second, other researchers have successfully modelled bird distributions in diverse landscapes around the world (Dettmers & Bart, 1999; Manel *et al.*, 2001; Pearce *et al.*, 2001; Pearce *et al.*, 2002).

In this study, we developed and evaluated predictive models of breeding bird distributions in the Toiquima Range, Toiyabe

Table 2 Data sets used to build and evaluate models of species occurrence. See Fig. 1 for geographical locations. Data for the temporal models were drawn from all three mountain ranges

Model type	Model-building data set	Number of sites	Evaluation data set	Number of sites
Temporal	2001	176	2002	176
	2002	176	2001	176
Spatial	Toiyabe Range and Toquima Range	125	Shoshone Mountains	51
	Toiyabe Range and Shoshone Mountains	120	Toquima Range	56
	Toquima Range and Shoshone Mountains	107	Toiyabe Range	69

Range and Shoshone Mountains. For each of 18 species, we constructed predictive models using five different subsets of the available data, two temporal (2001 and 2002) and three spatial (Toiyabe/Toquima, Toiyabe/Shoshone and Toquima/Shoshone), and then evaluated each model using three different measures of accuracy and independent data (i.e. the remaining data not used to develop the model) (Table 2).

We used the 90 models of species occurrence (18 species by five data sets) (Table 2) to address three main issues. First, we compared four different measures of model accuracy and determined whether those measures were correlated and whether they were affected by occurrence rate. The four accuracy measures were discrimination ability (AUC), CCR, sensitivity (proportion of true positives) and specificity (proportion of true negatives). Second, we assessed whether the accuracy of predictive models of species distributions was affected by either spatial or temporal variation in occurrence rates. Third, we determined whether the number of variables in a model affected each measure of model accuracy.

METHODS

Field methods

Our data collection incorporated established techniques that reliably detect species presence and permit assessment of distributional trends (Bibby *et al.*, 2000; Buckland *et al.*, 2001; Siegel *et al.*, 2001). These methods have previously been described in detail (Mac Nally *et al.*, 2004). In brief, we conducted inventories for birds in 176 locations throughout five canyons in both the Shoshone Mountains and the Toiyabe Range, and six canyons in the Toquima Range, during the breeding seasons of 2001 and 2002 using 75-m fixed radius point counts (Ralph *et al.*, 1995; Dobkin & Rich, 1998). Survey effort was equal between years. Within each canyon, we placed point counts at least 200 m apart along an elevational gradient in each of the three major vegetation types: (1) riparian areas with aspen groves or willow and rose thickets, (2) piñon-juniper woodland, and (3) sagebrush-dominated shrubland. Within each year, we conducted at least one of three point counts at each site very early, and at least one later in the morning. This enabled us to account for daily temporal variation in detection probability; most birds tend to sing less as the daily temperature increases and different birds tend to sing at different intensities throughout the morning. We allowed for at least 10 days between each visit to each site to account for

temporal variation in species detectability throughout the breeding season.

We collected vegetation data during the breeding season of 2002. All data were collected within three 11.3-m radius circles (0.04 ha) near the centre of each 75-m radius point count site (Martin *et al.*, 1997). We gathered vegetation density readings at each cardinal direction in each circle using a density board (Noon, 1981). We also measured height and diameter at breast height (d.b.h.) of all trees within each 11.3-m radius circle. We collected 22 ocular tube readings from each point count site and used these to calculate percent cover of canopy, shrub and ground vegetation (Noon, 1981). We also recorded the presence, length (or height) and d.b.h. of dead wood (logs and snags) and the presence, mean width and type of water (i.e. seep, flowing and standing) present at each point count site.

Model development

We used multiple logistic regression to construct predictive models of species occurrence (presence/absence). Subsets of potential explanatory variables were selected based on the ecology of the species being modelled. We included environmental variables derived from a digital elevation model (see below) and vegetation data collected at each point count site. Using a GIS, we derived environmental values from several grids within a 90-m buffer around the centre of each point (Table 3). All except one of our GIS layers were composed of 30-m grid cells; thus, a buffer of 90 m allowed for complete representation of the point count area. The grid for mean annual precipitation has a 4.5-km grid cell resolution and is derived from PRISM (Daly *et al.*, 1994); due to a lack of weather stations in this region, this is one of the best estimations of precipitation available. Grids representing distance to water, elevation, slope, aspect as described by mean eastness and mean northness, and topographic exposure were derived using the spatial analyst extension in ArcView 3.2 (ESRI, 1996). We considered quadratic terms for precipitation, elevation, topographic exposure, and slope during model construction to allow for the possibility of a non-linear relationship between occurrence and these explanatory terms. Finally, we used a measure of primary productivity, the Normalized Difference Vegetation Index (NDVI), which was obtained from a satellite image taken on 2 June 2000. Seto *et al.* (2004) demonstrated that NDVI had a significant relationship with avian species richness in the same mountain ranges used in this study.

Table 3 Independent variables derived from a digital elevation model using GIS. Asterisk (*) indicates that quadratic terms were also used in model building as potential predictor variables

Grid	Value derived within a 90-m buffer around each point count centre
*Digital elevation model (DEM)	Mean elevation
*Slope	Mean slope
Northness	Mean northness on a scale from -100 (south facing) to 100 (north facing)
Eastness	Mean eastness on a scale from -100 (west facing) to 100 (east facing)
Distance to water	Mean distance to water (running or standing) in metres
Normalized difference vegetation index (NDVI)	Mean NDVI
Topographic exposure	Mean topographic exposure where negative numbers represent valleys and positive values represent ridge tops
*PRISM (Daly <i>et al.</i> , 1994)	Mean annual precipitation

We developed models of species occurrence using five different data sets (two temporal and three spatial) and validated each of the five sets of models using an independent set of data not used to develop the model (see Table 2). Because of the large number of potential independent variables, we took measures to avoid multicollinearity. Prior to model construction for each species, we conducted Pearson correlation analyses and defined any combination of two independent variables as highly correlated if $r > 0.40$. We then conducted univariate logistic regression analyses with each independent variable and the occurrence of the modelled species. We retained all variables with log-likelihood $P < 0.25$ as potential variables for the final multiple regression model (Hosmer & Lemeshow, 1989). If two independent variables were correlated in the Pearson correlation analyses, we retained the variable with the lower P value. However, if log-likelihood P values for the two variables were the same, we ran a separate multiple logistic regression model with each variable.

We developed multiple logistic models for each of the 18 species of breeding birds for which, in all five subsets of data, the number of candidate predictor variables was less than the total number of sites where the species was present. Across the landscape, the number of presences of these species ranged from 3 to 130 (mean 38.2) across the 176 sites. During model construction, we used a stepwise selection procedure followed by a best subset selection. In order to obtain a potential model for each significance level, we used stepwise selection (SAS version 8.0, SAS Institute, 2002) with high alpha values (0.9) (Shtatland *et al.*, 2001). We then used the smallest Akaike's Information Criterion (AIC) value to select the optimal model (Burnham & Anderson, 1998). AIC is a model selection criterion that attempts to balance model complexity (overfitting) and bias (underfitting) using the principle of parsimony (Burnham & Anderson, 2001), and can be used with numerous and diverse modelling approaches, from genetic algorithms (Jefferson *et al.*, 1997; Vinterbo & Ohno-Machado, 1999) to semivariograms (Banerjee *et al.*, 2004).

Next, we conducted a best subset regression with all models containing one predictor variable more or less than the AIC-optimal model from the stepwise procedure (Shtatland *et al.*, 2001). For example, if the model with the smallest AIC value from the stepwise procedure had three predictor variables, we ran a best subset regression with all models having two, three and

four predictor variables. Model hierarchy was maintained such that models with a quadratic term were considered only if the corresponding linear term for that variable also entered the model. We again selected the model with the lowest AIC value; this model was identified as the final explanatory model for each species in each model-building data set. The use of AIC values precludes the selection of a final, best model such that practitioners should compare a subset of candidate models and evaluate reoccurring predictors therein. However, for the purposes of this paper regarding model accuracy, we selected a final explanatory model to facilitate model evaluation. The explanatory models were used to make predictions for locations or years not included in the model-building data set; these were the predictions that we subsequently evaluated.

Model evaluation and comparisons

Following model development and prior to model evaluation using independent data sets, we assessed the accuracy of the explanatory models using the same four measures: the area under a receiver operating characteristic curve (AUC) (SAS version 8.0, SAS Institute, 2002), correct classification rate (CCR) for a probability threshold of 0.7, sensitivity (the proportion of true positives Table 1) and specificity (proportion of true negatives, Table 1). We used a probability threshold of 0.7 rather than 0.5 because we preferred to have fewer overall predictions of higher quality (Fleishman *et al.*, 2003b).

In the same manner in which we assessed the accuracy of the explanatory models, we calculated the AUC, CCR, sensitivity and specificity for each of the models tested with independent, evaluation data sets. We performed a simple linear regression between each pair of accuracy measures (AUC, CCR, sensitivity and specificity). For example, we wanted to know whether there was a positive correlation between CCR and AUC. We also used simple linear regression to assess the relationship between each accuracy measure and species occurrence rates in both the model-building and evaluation data sets. In other words, we examined whether species occurrence rates affect the different measures of accuracy. To assess the effects of spatial and temporal variations in species occurrence rates on model accuracy, we compared mean values of the different accuracy measures across model type (spatial vs.

temporal) using two-tailed *t*-tests. In addition, we used simple linear regression to test whether the number of variables in a model affected the four measures of model accuracy.

RESULTS

Overall model performance and correlations among accuracy measures

We recorded 76 species of breeding birds during the 2001 and 2002 censuses. Species richness was similar between years with 65 species recorded in 2001 and 66 recorded in 2002, whereas species composition was more variable between years (Fleishman *et al.*, 2003a). We developed occurrence models for the 18 species of breeding birds that met our criteria with respect to model stability using each of the five data sets (two temporal and three spatial) for a total of 90 models. We present the results of 89 models evaluated with independent data sets; a model for sage thrasher (*Oreoscoptes montanus*) developed using data from the Toquima Range and Shoshone Mountains was omitted because of the small number of occurrences ($n = 3$) in the evaluation data set.

The mean discrimination ability of each of the five sets of models (i.e. the average discrimination ability of the 18 models constructed using each different set of data) was good (mean AUC ≥ 0.7) when evaluated with an independent data set (Table 4). Furthermore, > 55% of the individual occurrence models within each data set had good discrimination ability (Table 4). Only 3% of all models had excellent discrimination ability (AUC > 0.9), and nearly 40% of the models had poor discrimination ability (AUC < 0.7). Model discrimination in the evaluation data sets (when tested on independent data) was not significantly correlated with discrimination ability of the explanatory models ($n = 89$, $R^2 = 0.00$, $P = 0.70$). In other words, performance of the explanatory models does not necessarily indicate that the same models will perform well when predicting occurrence in another time or place.

The CCR (probability threshold = 0.7) of model predictions was > 70% for all five sets of models (Table 4). In other words, > 70% of the predictions were correct when the predicted probability of presence was ≥ 0.7 . Four of the five data sets had a

mean CCR of > 77%, and one of these (2002) had a mean CCR of 85%. In each of the data sets, specificity was much greater than sensitivity, which means that species absences were predicted more accurately than species presences (Table 4). Indeed, there was a negative relationship between sensitivity and specificity across all 89 models. There was also a negative relationship between sensitivity and CCR. In contrast, there was a positive relationship between a model's ability to predict absences (specificity) and CCR, and a positive relationship between a model's ability to predict presences (sensitivity) and a model's discrimination ability (AUC).

Effect of occurrence rate on model accuracy

Species occurrence rate (proportion of sites occupied) was positively correlated between the model-building and evaluation data sets for each of the five sets of models ($n = 89$, $R^2 = 0.37$, $P < 0.0001$). This suggests that a species present in a high proportion of sites in the model-building data set was likely to be present in a high proportion of sites in the evaluation data set. There was a significant relationship between model accuracy and occurrence rate in both the model-building and the evaluation data. For example, there was a significant positive quadratic relationship between CCR and occurrence rate across the model-building data sets and between CCR and occurrence rate across the evaluation data sets ($P < 0.0001$ in both cases) (Fig. 2). However, occurrence rate in the evaluation data set explained 80% of the variation in CCR, whereas occurrence rate in the model-building data set explained only 22%.

There was a positive linear relationship between sensitivity and occurrence rate in the model-building data sets ($n = 89$, $R^2 = 0.29$, $P < 0.0001$) and between sensitivity and occurrence rate in the evaluation data sets ($n = 89$, $R^2 = 0.81$, $P < 0.0001$). In general, a model's ability to predict presences increased as occurrence rate increased (Fig. 2). However, there was a negative quadratic relationship between model specificity (ability to predict absences) and occurrence rate in both the model-building and evaluation data sets ($P < 0.0001$ for both data sets), with more variance in specificity explained by occurrence rate in the evaluation data sets (84%) than in the model-building data sets (33%) (Fig. 3). The differential effects of occurrence rate in the model-

Table 4 Accuracy measures across all data sets. Thresholds for 'poor', 'good' and 'excellent' were < 0.7, ≥ 0.7 –0.9 and > 0.9, respectively (Pearce & Ferrier, 2000a). TQ, Toquima Range; SH, Shoshone Mountains; TY, Toiyabe Range

Model type	Model-building data set	Evaluation data set	Mean CCR	Mean Sensitivity	Mean Specificity	Discrimination ability			
						Mean AUC	Poor (AUC < 0.7)	Good (0.7 < AUC < 0.9)	Excellent (AUC > 0.9)
Temporal	2001	2002	78.6	0.148	0.963	0.721	8 (44.4%)	10 (55.6%)	0
	2002	2001	85.1	0.001	0.996	0.735	8 (44.4%)	10 (55.6%)	0
Spatial	TYSH	TQ	77.7	0.158	0.902	0.735	7 (38.9%)	10 (55.6%)	1 (5.6%)
	TYTQ	SH	70.6	0.239	0.899	0.731	6 (33.3%)	10 (55.6%)	2 (11.1%)
	TQSH	TY	77.2	0.183	0.867	0.718	7 (41.2%)	10 (58.8%)	0
All models	NA	NA	77.8	0.145	0.926	0.728	36 (40.4%)	50 (56.2%)	3 (3.3%)

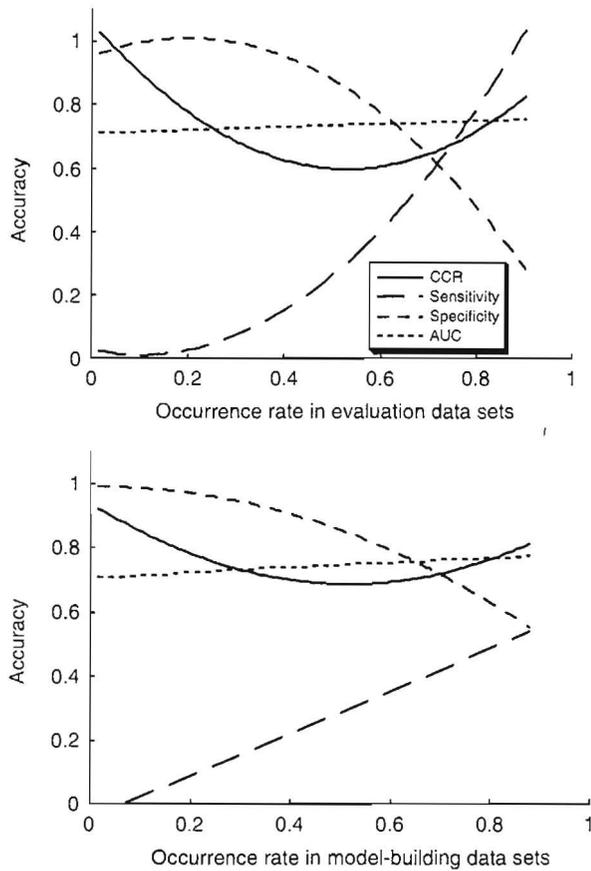


Figure 2 Effects of observed species occurrence rate on model accuracy in the model-building and evaluation data sets.

building and evaluation data sets on specificity and sensitivity are illustrated in Fig. 3. There was more variation in model specificity and sensitivity for species in the model-building data set than for species in the evaluation data set (Figs 2 & 3). Table 5 shows the values for mean occurrence rate and mean accuracy for each species modelled. The relationship between model discrimination (AUC) and species occurrence rate was not statistically significant in either the model-building data ($n = 89$, $R^2 = 0.03$, $P = 0.11$) or the evaluation data ($n = 89$, $R^2 = 0.01$, $P = 0.31$) (Fig. 2).

Although the occurrence rates in the model-building and evaluation data sets were correlated, some species had higher rates of occurrence in one of the two data sets. The difference in species occurrence rates between model-building and evaluation data sets helps explain variability in model accuracy across space or time (i.e. in spatial or temporal models, respectively). Because the majority of species were present in more sites in 2002 than in 2001, all differences in species occurrence rates between the 2001 model-building and 2002 evaluation data sets were negative (Fig. 4).

Across all 89 models, AUC did not change as the absolute value of the difference in occurrence rates between the model-building and evaluation data sets increased ($n = 89$, $R^2 = 0.03$, $P = 0.10$). However, when each of the five sets of models was considered

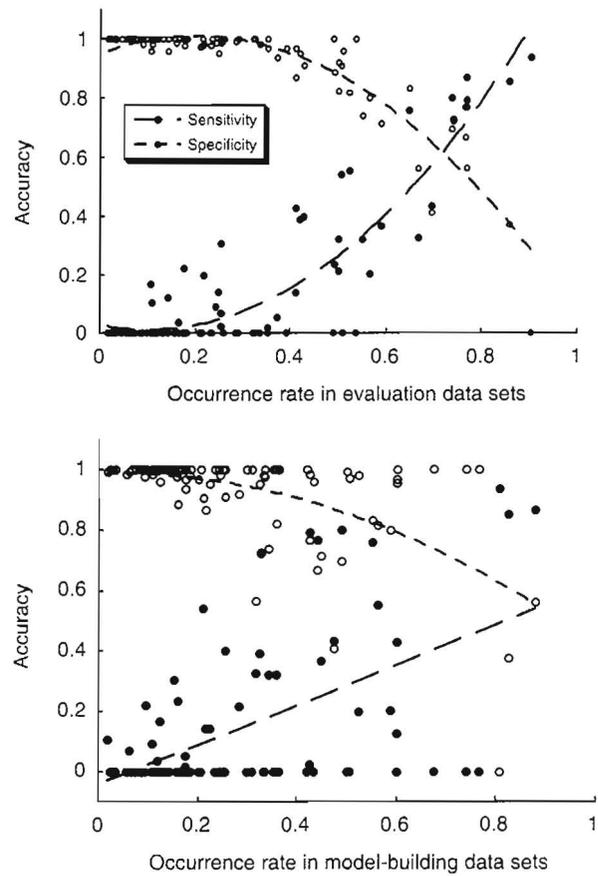


Figure 3 Effects of observed species occurrence rate on sensitivity and specificity in the model-building and evaluation data sets.

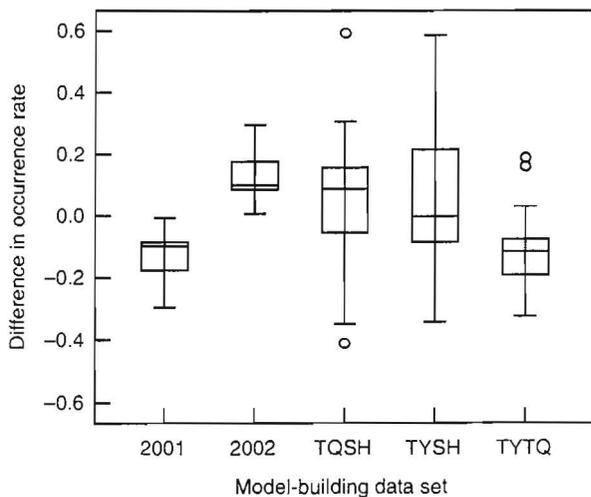


Figure 4 Difference in species occurrence rates between all model-building and evaluation data sets: TQSH, Toquima/Shoshone; TYSH, Toiyabe/Shoshone, TYTQ, Toiyabe/Toquima. Error bars indicate the maximum and minimum values except in the case of outliers. Outliers are defined as values more than 1.5 times the interquartile range.

Table 5 Mean and standard deviation of occurrence rate and model accuracy for 18 species of breeding birds. Values for each species are averaged across the five data sets (three spatial and two temporal). AUC, area under a receiver operating characteristic curve; CCR, correct classification rate

Species	Mean occurrence rate	Mean AUC	Mean CCR	Mean sensitivity	Mean specificity
American robin, <i>Turdus migratorius</i>	0.376 (0.16)	0.687 (0.10)	63.9 (19.8)	0.09 (0.14)	0.90 (0.19)
Blue-grey gnatcatcher, <i>Poliophtila caerulea</i>	0.154 (0.07)	0.793 (0.04)	85.2 (7.69)	0.06 (0.14)	0.99 (0.01)
Brewer's sparrow, <i>Spizella breweri</i>	0.476 (0.11)	0.791 (0.10)	64.9 (12.6)	0.36 (0.21)	0.83 (0.24)
Broad-tailed hummingbird, <i>Selasphorus platycercus</i>	0.241 (0.11)	0.762 (0.07)	77.4 (12.4)	0.06 (0.10)	0.97 (0.05)
Black-throated grey warbler, <i>Dendroica nigrescens</i>	0.470 (0.20)	0.820 (0.06)	76.9 (3.98)	0.41 (0.40)	0.86 (0.14)
Bushtit, <i>Psaltriparus inimus</i>	0.102 (0.05)	0.640 (0.10)	89.5 (7.10)	0 (0)	1 (0)
Cassin's finch, <i>Carpodacus cassinii</i>	0.194 (0.08)	0.680 (0.10)	79.6 (12.3)	0.04 (0.06)	0.98 (0.03)
Chipping sparrow, <i>Spizella passerina</i>	0.230 (0.11)	0.730 (0.08)	76.2 (14.7)	0.04 (0.06)	0.97 (0.06)
Dark-eyed junco, <i>Junco hyemalis</i>	0.097 (0.05)	0.785 (0.08)	90.6 (4.70)	0.10 (0.10)	0.99 (0.02)
Green-tailed towhee, <i>Pipilo chlorurus</i>	0.748 (0.15)	0.743 (0.09)	74.2 (13.2)	0.69 (0.39)	0.53 (0.37)
McGillivray's warbler, <i>Oporornis tolmiei</i>	0.426 (0.16)	0.803 (0.05)	71.6 (14.3)	0.30 (0.30)	0.87 (0.14)
Mountain bluebird, <i>Sialia currucoides</i>	0.108 (0.06)	0.726 (0.10)	88.5 (8.40)	0 (0)	0.99 (0.01)
Mountain chickadee, <i>Poecile gambeli</i>	0.349 (0.12)	0.724 (0.05)	67.3 (13.9)	0.17 (0.16)	0.90 (0.12)
Mourning dove <i>Zenaida macroura</i>	0.108 (0.04)	0.610 (0.07)	88.9 (7.09)	0 (0)	1 (0)
Red-shafted flicker, <i>Colaptes auratus</i>	0.210 (0.08)	0.685 (0.04)	78.3 (13.1)	0.01 (0.01)	0.99 (0.01)
Sage thrasher, <i>Oreoscoptes montanus</i>	0.085 (0.04)	0.840 (0.10)	90.4 (6.04)	0 (0)	1 (0)
Spotted towhee, <i>Pipilo maculatus</i>	0.505 (0.13)	0.699 (0.15)	59.8 (14.5)	0.24 (0.31)	0.91 (0.10)
Western tanager, <i>Piranga ludoviciana</i>	0.191 (0.07)	0.615 (0.05)	80.4 (12.2)	0 (0)	0.99 (0.01)

independently, we found that the difference in occurrence rates between the 2001 model-building and 2002 evaluation data sets had a positive relationship with discrimination ability ($n = 36$, $R^2 = 0.25$, $P = 0.04$). Across all models, the relationship between CCR and the difference in occurrence rates between the model building and evaluation data sets was not significant ($n = 89$, $R^2 = 0.04$, $P = 0.08$). However, when each data set was considered individually, the difference in species occurrence rates between the 2001 model-building data and the 2002 evaluation data showed a significant negative relationship with CCR ($n = 36$, $R^2 = 0.50$, $P = 0.001$), as did the difference in occurrence rates between the 2002 model-building and the 2001 evaluation data sets with CCR ($n = 36$, $R^2 = 0.41$, $P = 0.003$). As the difference in occurrence rates between the model-building and evaluation data sets increased, there was no change in a model's ability to predict presences (sensitivity) ($n = 89$, $R^2 = 0.03$, $P = 0.08$) or its ability to predict absences (specificity) ($n = 89$, $R^2 = 0.01$, $P = 0.27$), even when each data set was assessed individually.

Effect of model predictor variables

Of the 18 predictor variables used during model development, 11 were derived from GIS and seven were measured in the field. Four of the five predictors that entered occurrence models most frequently were GIS-derived (elevation, precipitation, NDVI and slope). The most frequent field-based predictor was canopy cover. Elevation and canopy cover were the most common predictors in species occurrence models: elevation entered 28 models (31%), whereas canopy cover entered 26 (29%). Quadratic terms for elevation, slope and precipitation entered 9%, 4% and 6% of the models, respectively. Six models (7%) included only

predictors that were measured in the field and 17 models (19%) included only GIS-derived predictors.

The number of predictors per model ranged from one to seven. Most models had three or four predictors (31% and 28%, respectively). Across all models, as the number of predictor variables in a model increased, the model discrimination (AUC) increased as well ($n = 89$, $R^2 = 0.21$, $P < 0.0001$). This relationship was driven by a positive association between AUC and the number of predictor variables in two sets of models: the models built with the 2001 data set and evaluated with 2002 data, and the models built with the Toiyabe/Toquima data set and tested with Shoshone data. Similarly, across all models, there was a positive relationship between sensitivity and the number of predictor variables ($n = 89$, $R^2 = 0.06$, $P = 0.02$), but the positive relationship was attributable solely to the models developed with 2001 data and evaluated with the 2002 data. The relationships between the number of predictor variables and either CCR or specificity were not statistically significant.

Spatial and temporal variations

Model discrimination ability was not affected by whether a model predicted species occurrence in space (Toiyabe/Toquima, Toiyabe/Shoshone and Toquima/Shoshone) or species occurrence in time (2001 and 2002). The mean AUC for spatial models and temporal models was the same (0.73) (Fig. 5). However, temporal models had higher average CCR and a higher average specificity than spatial models (i.e. temporal models classified absences better than spatial models). In contrast, spatial models exhibited higher average sensitivity than temporal models (i.e. spatial models classified species presences better than temporal models)

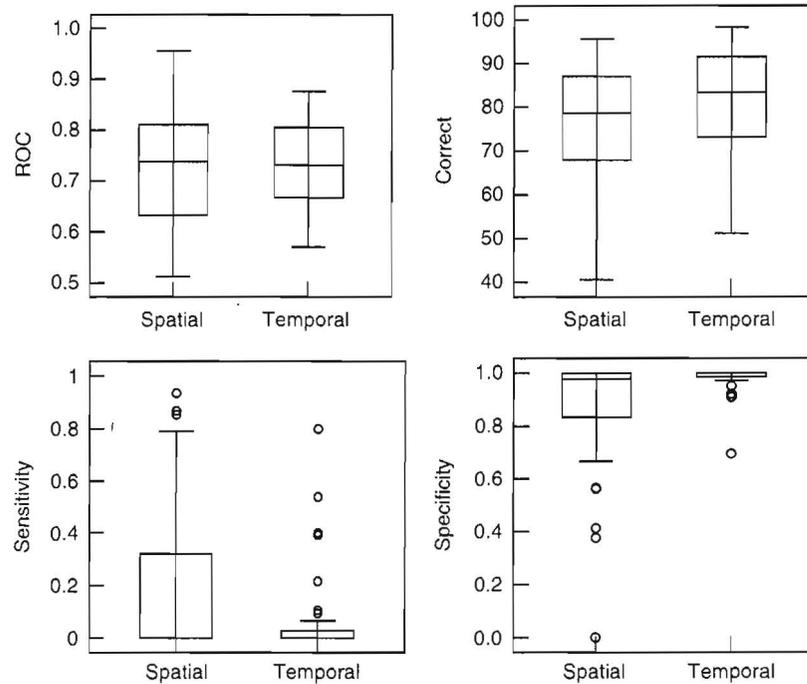


Figure 5 Comparison of the distributions of accuracy measures for models assessing the effects of both spatial and temporal variation in occurrence rate. Note that the scale for sensitivity and specificity is different from the scale for AUC and CCR. Error bars indicate the maximum and minimum values except in the case of outliers. Outliers are defined as values more than 1.5 times the interquartile range.

(Fig. 5). Overall, the variation in model accuracy was greater in spatial models than in temporal models (Fig. 5).

DISCUSSION

Overall model performance

Several of the avian occurrence models presented here appeared to perform fairly well when evaluated with independent data from another time period or location. The spatial and temporal extent across which these models apparently can be applied successfully makes their use practical given the scales of most management decisions, at least in western North America. The Great Basin, for example, covers more than 400,000 km². If effective models of species occurrence can be built on the basis of data collected from sections of three mountain ranges, those models should be cost-effective even within the 30,000-km² biogeographical sub-region where our work was centred (Austin & Murphy, 1987). Within a region, it is reasonable to assume that land use decisions might be based on occurrence models derived from census data collected in a previous year (or years) within a subset of the same landscape. Therefore, the framework we present for the development and evaluation of species occurrence models can potentially be applied in any region and to a variety of taxonomic groups.

Virtually all distributional patterns reflect an ecological response to abiotic or biotic phenomena. Nonetheless, those responses can be exaggerated by sources of non-independence such as spatial autocorrelation (Legendre, 1993; Diniz-Filho *et al.*, 2003; Tobin, 2004). We recently examined patterns of spatial autocorrelation in avian species composition and three measures of birds' habitats (vegetation composition, vegetation structure, and estimated productivity) in the same study system

(E. Fleishman and R. Mac Nally, unpublished manuscript). We found that avifaunal composition was spatially autocorrelated at lag distances (distances between sites) of approximately 0 to 30 km. The three measures of habitat were also spatially autocorrelated, but to different distances than each other or avifaunal composition. Although our results suggested that the spatial pattern of avifaunal composition was influenced more strongly by vegetation composition (variables not emphasized heavily in this study) than by vegetation structure or estimated productivity, patterns in the bird data were not closely concordant with patterns in any of our three measures of bird habitats. Thus, we believe that spatial autocorrelation was unlikely to have a strong influence on model selection and performance.

Our results illustrate the importance of calculating more than one measure of model accuracy and of assessing separately the ability of a model to predict presences and absences (Fielding & Bell, 1997; Manel *et al.*, 1999). However, there are two major drawbacks associated with measures of accuracy derived from classification tables (Forbes, 1995; Fielding & Bell, 1997; Manel *et al.*, 1999). First, they are often based on an arbitrarily chosen probability threshold value, and second, as our work demonstrates, their accuracy can be greatly affected by occurrence rate. Nonetheless, it is useful to conduct some type of comparison between the ability of a model to predict presences vs. absences.

Effect of occurrence rate

Species occurrence rates in the evaluation data sets had a much stronger relationship with CCR, sensitivity and specificity than did occurrence rates in the model-building data sets (Figs 2 & 3). The fact that so much of the variation in these accuracy measures is explained by species occurrence alone makes them inferior

measures compared with AUC (area under the ROC curve). In general, accuracy measures describe how well models perform when evaluated with independent data. Therefore, the observed occurrence rates in the evaluation data affected these accuracy measures more so than the occurrence rates in the model-building stage. Unfortunately, one cannot anticipate the magnitude of temporal variation in occurrence rate of a species unless multiple years of survey data exist. We suggest that whenever possible, surveys of breeding birds in temperate ecosystems be conducted approximately every 3 to 4 years to ensure that occurrence rates have not changed dramatically, which would render predictive models less reliable.

Like other researchers (Manel *et al.*, 1999; Fleishman & Mac Nally, 2004), we found that species absences were predicted more accurately than species presences (i.e. specificity was greater than sensitivity), even when the overall predictive success (CCR) was considered to be good. One possible explanation is that occurrence rate has considerable influence on accuracy measures. Species occurrence rate in this study ranged from 0.01 to 0.85 (mean = 0.28). The trend presented here and commonly seen in the literature — that low occurrence rate often corresponds to low model sensitivity — highlights the difficulty in predicting the presence of rare species.

Our choice of decision threshold could also explain why the ability of our models to predict absences was consistently much greater than their ability to predict presences. We used a threshold of 0.7, whereby a species was considered to be present when the probability of occurrence was ≥ 0.7 , as opposed to a more traditional threshold of 0.5. Stricter classification rules (i.e. decision thresholds) may produce fewer overall predictions of higher quality compared with more lenient classification rules (Fleishman *et al.*, 2003b). Zweig and Campbell (1993) suggested that researchers determine the optimal threshold value for each individual circumstance based on 'costs' assigned to commission errors [predicting that a species is present when it is in fact absent (false positive)] and omission errors [predicting that a species is absent when it is in fact present (false negative)]. Depending on the purpose of any one occurrence model, the relative 'costs' associated with commission and omission errors can be determined (Zweig & Campbell, 1993; Fielding, 2002). A probability threshold of 0.5 assumes that the costs of commission and omission errors are equal, whereas a threshold of 0.7 assumes that omission errors are more serious. Increasing this threshold tends to decrease sensitivity and increase specificity. An optimal threshold that considers both occurrence rate and the relative costs of commission and omission for the situation at hand can be estimated by calculating the slope of a line tangent to an ROC curve that incorporates occurrence rates and costs associated with the errors (Zweig & Campbell, 1993; Fielding, 2002).

Effect of number of predictor variables

Although discrimination ability (AUC) is a useful measure of model accuracy that is not affected by occurrence rate, we found that it was significantly correlated with the number of predictor variables in a model. Including a large number of variables in a

model limits the model's temporal or spatial generality. In this study, Akaike's Information Criterion (AIC) was used in model selection. This criterion attempts to balance model fit (ability to explain observed variation in the dependent variable) and model complexity (number of independent variables). Therefore, if we had used a method of model selection that did not constrain the number of independent variables, the relationship between discrimination ability and k (the number of predictors in the model) might have been different.

Effects of spatial and temporal variations in occurrence rate

The differences in accuracy between temporal and spatial models may also be attributed to species occurrence rates. Temporal models had significantly higher mean CCR and specificity and significantly lower sensitivity than spatial models. In our study, there was less variation in species occurrence rates in space (among mountain ranges) than in time (between years). The mean occurrence rate in 2002 for the 18 species modelled (0.27) was nearly twice the mean occurrence rate in 2001 (0.15).

Temporal models consistently predicted absences well while predicting presences poorly. Spatial models also classified absences better than presences, but their accuracy values were much more variable; some presences were classified well and some absences were classified poorly (Fig. 5). In our study system, variation in occurrence rates of breeding birds from year to year may be greater than variation among mountain ranges, and consequently may constrain the accuracy of predicted presences. Other researchers have documented considerable temporal variation in avian occurrence in desert systems (Johnson, 1995; Gutzwiller & Barrow, 2001). In systems where large annual fluctuations in occurrence are the norm, it is advantageous to collect multiple years of occurrence data to develop and evaluate models.

If management plans are to be based in part on predictive models of occurrence, then follow-up censuses across the landscape are essential to determine whether management is indeed sufficient to protect species of concern, and to trigger changes in land use if current management is not achieving ecological objectives. Such an adaptive management framework helps to ensure that land uses and available funding are being appropriately assigned. Ecosystems and the species that inhabit them are inherently dynamic and no model is universally applicable across space and time. Nonetheless, there is a need to understand the distributions of species in order to protect them and their habitats. Predictive models of occurrence, and more importantly their subsequent evaluation using independent data, provide substantial insight into spatial and temporal variations in species occurrences. Such insight better equips ecologists and land managers to make responsible decisions.

ACKNOWLEDGEMENTS

We are thankful to John Fay for providing the GIS grids used to derive predictor variables and to Dennis Murphy, Eric Porter and

Ralph Mac Nally for comments and support. We are also thankful to Jason Bulluck and Melissa Betrus for their tireless help collecting field data. Financial assistance for this project was provided by Miami University, the Nevada Biodiversity Research and Conservation Initiative, and the Joint Fire Sciences Program via the Rocky Mountain Research Station, Forest Service, US Department of Agriculture.

REFERENCES

- Aspinall, R. & Matthews, K. (1994) Climate-change impact on distribution and abundance of wildlife species — an analytical approach using GIS. *Environmental Pollution*, **86**, 217–223.
- Austin, G.T. & Murphy, D.D. (1987) Zoogeography of Great Basin butterflies: patterns of distribution and differentiation. *Great Basin Naturalist*, **47**, 186–201.
- Banerjee, S., Carlin, B.P. & Gelfand, A.E. (2004) *Hierarchical modeling and analysis for spatial data*. Chapman & Hall / CRC, Boca Raton, Florida.
- Bibby, C.J., Burgess, N.D., Hill, D.A. & Mustoe, S. (2000) *Bird census techniques*. Academic Press, London.
- Boone, R.B. & Krohn, W.B. (2000) Predicting broad scale occurrences of vertebrates in patchy landscapes. *Landscape Ecology*, **15**, 63–74.
- Brown, J. (1978) The theory of insular biogeography and the distribution of boreal mammals and birds. *Great Basin Naturalist Memoirs*, **2**, 209–228.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001) *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press, New York.
- Burnham, K.P. & Anderson, D.R. (1998) *Model selection and inference: a practical, information-theoretic approach*. Springer-Verlag, New York.
- Burnham, K.P. & Anderson, D.R. (2001) Kullback-Leiber information as a basis for strong inference in ecological studies. *Wildlife Research*, **28**, 111–119.
- Caicco, S.L., Scott, J.M., Butterfield, B. & Csuti, B. (1995) A gap analysis of the management status of the vegetation of Idaho (USA). *Conservation Biology*, **9**, 498–511.
- Daly, C., Neilson, R.P. & Phillips, D.L. (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, **33**, 140–158.
- Dettmers, R. & Bart, J. (1999) A GIS modeling method applied to predicting forest songbird habitat. *Ecological Applications*, **9**, 152–163.
- Diniz-Filho, J.A.F., Bini, L. & Hawkins, B.A. (2003) Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, **12**, 53–64.
- Dobkin, D.S. & Rich, A.C. (1998) Comparison of line-transect, spot-map, and point-count surveys for birds in riparian areas of the Great Basin. *Journal of Field Ornithology*, **69**, 430–443.
- ESRI—Environmental Systems Research Institute Inc. (1996) *Using ArcView GIS 3.2*. Redlands, CA.
- Fielding, A.H. (2002) What are the appropriate characteristics of an accuracy measure? *Predicting species occurrence: issues of accuracy and scale* (ed. by J.M. Scott, P.J. Heglund and M.L. Morrison), pp. 271–280. Island Press, Washington D.C.
- Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.
- Fleishman, E. & Mac Nally, R. (2004) Linking models of species occurrence and landscape reconstruction 2002–03. *Transactions of the Western Section of the Wildlife Society*, **38/39**, 1–4.
- Fleishman, E., Mac Nally, R., Fay, J.P. & Murphy, D.D. (2001) Modeling and predicting species occurrences using broad-scale environmental variables: an example with butterflies of the Great Basin. *Conservation Biology*, **15**, 1674–1685.
- Fleishman, E., Betrus, C. & Penfield, L. (2003a) Annual variability of species richness and composition of breeding birds in the central Great Basin. *Great Basin Birds*, **6**, 36–44.
- Fleishman, E., Mac Nally, R. & Fay, J.P. (2003b) Validation tests of predictive models of butterfly occurrence based on environmental variables. *Conservation Biology*, **17**, 806–817.
- Forbes, D.A. (1995) Classification-algorithm evaluation: five performance measures based on confusion matrices. *Journal of Clinical Monitoring*, **11**, 189–206.
- Grayson, D.K. (1993) *The desert's past: a natural prehistory of the Great Basin*. Smithsonian Institution Press, Washington, D.C.
- Guisan, A. & Zimmerman, N. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Gutzwiller, K.J. & Barrow, W.C. (2001) Bird-landscape relations in the Chihuahuan desert: coping with uncertainties about predictive models. *Ecological Applications*, **11**, 1517–1532.
- Hanley, J.A. & McNeil, B.J. (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **143**, 29–36.
- Hosmer, D.W. & Lemeshow, S. (1989) *Applied logistic regression*. Wiley-Interscience Publication, New York.
- Jefferson, M.F., Pendleton, N., Lucas, S.B. & Horan, M.A. (1997) Comparison of a genetic algorithm neural network with logistic regression for predicting outcome after surgery for patients with nonsmall cell lung carcinoma. *Cancer*, **79**, 1338–1342.
- Johnson, N.K. (1995) Seven avifaunal censuses spanning one-half century on an island of white firs (*Abies concolor*) in the Mojave Desert. *Southwestern Naturalist*, **40**, 76–85.
- Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74**, 1659–1673.
- MacNally, R., Fleishman, E., Bulluck, L. & Betrus, C. (2004) Comparative influence of spatial scale on beta diversity within regional assemblages of birds and butterflies. *Journal of Biogeography*, **31**, 917–929.
- Manel, S., Dias, J.M., Buckton, S.T. & Ormerod, S.J. (1999) Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology*, **36**, 734–747.
- Manel, S., Williams, H.C. & Ormerod, S.J. (2001) Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, **38**, 921–931.
- Martin, T.E., Paine, C.R., Conway, C.J., Hochachka, W.M., Allen, P.

- & Jenkins, W. (1997) *BBIRD field protocol*. Montana Cooperative Wildlife Research Unit, University of Montana, Missoula, Montana, USA.
- McDonald, K.A. & Brown, J.H. (1992) Using montane mammals to model extinctions due to climate change. *Conservation Biology*, **6**, 409–415.
- Medin, D.E., Welch, B.L. & Clary, W.P. (2000) Bird habitat relationships along a Great Basin elevational gradient. Research Paper **RMRS-RP-23**. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO. p. 22.
- Noon, B. (1981) Techniques for sampling avian habitats. *The use of multivariate statistics in studies of wildlife habitat* (ed. by D.E. Capen), pp. 42–52. USDA Forest Service General Technical Report. **RM-89**. Rocky Mountain Forest and Range Experimental Station, Fort Collins, CO.
- Pearce, J.L. & Ferrier, S. (2000a) An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling*, **128**, 127–147.
- Pearce, J.L. & Ferrier, S. (2000b) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, **133**, 225–245.
- Pearce, J.L., Ferrier, S. & Scotts, D. (2001) An evaluation of the predictive performance of distributional models for flora and fauna in north-east New South Wales. *Journal of Environmental Management*, **62**, 171–184.
- Pearce, J.L., Venier, L.A., Ferrier, S. & McKenney, D. (2002) Measuring prediction uncertainty in models of species distribution *Predicting species occurrence: issues of accuracy and scale* (ed. by J.M. Scott, P.J. Heglund and M.L. Morrison), pp. 383–390. Island Press, Washington D.C.
- Peterson, A.T. (2001) Predicting species' geographic distributions based on ecological niche modeling. *The Condor*, **103**, 599–605.
- Peterson, A.T., Egbert, S.L., Sanchez-Cordero, V. & Price, K.P. (2000) Geographic analysis of conservation priority: endemic birds and mammals in Veracruz, Mexico. *Biological Conservation*, **93**, 85–94.
- Ralph, J.C., Sauer, J.R. & Droege, S. (1995) *Monitoring bird populations by point counts*. General Technical Report. PSW-GTR-149. Pacific Southwest Research Station, Forest Service, US Department of Agriculture, Albany, CA.
- SAS Institute (2000) *SAS version 8.0*. SAS Institute Inc, Cary, NC.
- Scott, J.M., Davis, F., Csuti, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S., D'Erchia, F., Edwards, T.C., Jr, Ulliman, J. & Wright, R.G. (1993) Gap analysis: a geographic approach to protection of biological diversity. *Wildlife Monographs*, **123**, 1–41.
- Seto, K.C., Fleishman, E., Fay, J.P. & Betrus, C.J. (2004) Predicting spatial patterns of butterfly and bird species richness with Landsat TM derived NDVI. *International Journal of Remote Sensing*, **25**, 4309–4324.
- Shtatland, E.S., Barton, M.B. & Cain, E. (2001) The perils of stepwise logistic regression and how to escape them using information criteria and the output delivery system. SUGI 2001. Proceedings, Paper 222, SAS Institute Inc., Cary, NC.
- Siegel, R.B., Desante, D.F. & Nott, M.P. (2001) Using point counts to establish conservation priorities: how many visits are optimal? *Journal of Field Ornithology*, **72**, 228–235.
- Stauffer, D.F. (2002) Linking populations and habitats: Where have we been? Where are we going? *Predicting species occurrence: issues of accuracy and scale* (ed. by J.M. Scott, P.J. Heglund and M.L. Morrison), pp. 53–61. Island Press, Washington D.C.
- Tobin, P.C. (2004) Estimation of the spatial autocorrelation function: consequences of sampling dynamic populations in space and time. *Ecography*, **27**, 767–775.
- Vinterbo, S. & Ohno-Machado, L. (1999) A genetic algorithm to select variables in logistic regression: example in the domain of myocardial infarction. *Journal of the American Medical Informatics Association*, 1999 Symposium Supplement, 984–988.
- Wadsworth, R.A., Collingham, Y.C., Willis, S.G., Huntley, B. & Hulme, P.E. (2000) Simulating the spread and management of alien riparian weeds: are they out of control? *Journal of Applied Ecology*, **37** (Suppl. 1), 28–32.
- Zweig, M.H. & Campbell, G. (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, **39**, 561–577.

BIOSKETCHES

Lesley Bulluck is currently in the Department of Forestry, Wildlife and Fisheries at the University of Tennessee where she is developing population and habitat-use models for golden-winged warblers (*Vermivora chrysoptera*) in the Cumberland Mountains.

Erica Fleishman is at the Center for Conservation Biology at Stanford University. She aims to develop scientifically reliable, cost-effective approaches to understand the spatial and temporal distributions of assemblages and species, underlying mechanisms and the potential influence of human actions on those patterns and processes.

Christopher Betrus is in the Zoology Department at Miami University of Ohio where he is developing an umbrella index model for species conservation.

Rob Blair is in the Fisheries, Wildlife and Conservation Biology Department at the University of Minnesota where he studies the effects of land-use on birds and butterflies.

Editor: Tim Blackburn

