Ecological Applications, 16(3), 2006, pp. 1090–1102
© 2006 by the Ecological Society of America

PREDICTING ABUNDANCE OF DESERT RIPARIAN BIRDS: VALIDATION AND CALIBRATION OF THE EFFECTIVE AREA MODEL

L. Arriana Brand,[1,3] Barry R. Noon,[1] and Thomas D. Sisk[2]

[1]*Department of Fishery and Wildlife Biology, Colorado State University, Fort Collins, Colorado 80523-1474 USA*
[2]*Center for Environmental Sciences and Education, Northern Arizona University, P.O. Box 5694, Flagstaff, Arizona 86011-5694 USA*

Abstract. Reliable prediction of the effects of landscape change on species abundance is critical to land managers who must make frequent, rapid decisions with long-term consequences. However, due to inherent temporal and spatial variability in ecological systems, previous attempts to predict species abundance in novel locations and/or time frames have been largely unsuccessful. The Effective Area Model (EAM) uses change in habitat composition and geometry coupled with response of animals to habitat edges to predict change in species abundance at a landscape scale. Our research goals were to validate EAM abundance predictions in new locations and to develop a calibration framework that enables absolute abundance predictions in novel regions or time frames. For model validation, we compared the EAM to a null model excluding edge effects in terms of accurate prediction of species abundance. The EAM outperformed the null model for 83.3% of species ($N = 12$) for which it was possible to discern a difference when considering 50 validation sites. Likewise, the EAM outperformed the null model when considering subsets of validation sites categorized on the basis of four variables (isolation, presence of water, region, and focal habitat). Additionally, we explored a framework for producing calibrated models to decrease prediction error given inherent temporal and spatial variability in abundance. We calibrated the EAM to new locations using linear regression between observed and predicted abundance with and without additional habitat covariates. We found that model adjustments for unexplained variability in time and space, as well as variability that can be explained by incorporating additional covariates, improved EAM predictions. Calibrated EAM abundance estimates with additional site-level variables explained a significant amount of variability ($P < 0.05$) in observed abundance for 17 of 20 species, with R^2 values >25% for 12 species, >48% for six species, and >60% for four species when considering all predictive models. The calibration framework described in this paper can be used to predict absolute abundance in sites different from those in which data were collected if the target population of sites to which one would like to statistically infer is sampled in a probabilistic way.

Key words: abundance; birds; calibration; Effective Area Model (EAM); model testing; null model; prediction; San Pedro River, Arizona, USA; validation.

Introduction

Given the pace and extent of human-induced landscape change, conservation decisions often need to be made more rapidly than relevant data can be collected (Côté and Reynolds 2002). While there have been many studies that attempt to explain patterns of species occurrence or abundance in relation to habitat variables, few studies have attempted to predict abundance in novel locations (Stauffer 2002). An ability to predict species occurrence or abundance in previously unstudied locations is critical to the ability of land managers to make reliable, rapid decisions in the face of potential landscape alterations (Noon et al. 1980).

Manuscript received 21 January 2005; revised 30 August 2005; accepted 14 September 2005. Corresponding Editor: T. R. Simons.

[3] E-mail: arriana@cnr.colostate.edu

Predictions of species occurrence patterns have experienced modest success (Block et al. 1994, Stauffer 2002), but abundance has been much more difficult to predict (Rotenberry 1986, Morrison et al. 1987, Gutzwiller and Barrow 2001). However, accurate predictions of species abundance are much more useful than mere occurrence. Occurrence predictions enable managers to identify areas in which species are present, but cautions notwithstanding (Van Horne 1983), abundance predictions enable managers to prioritize conservation areas, and to develop specific management objectives pertaining to population size and trends (see Bock and Jones 2004). Absolute (vs. relative) abundance predictions also enable land managers to prioritize areas based on results obtained from different studies and regions. Despite its importance to conservation planning, accurate prediction of absolute abundance remains a largely unachieved goal (Rotenberry 1986, Stauffer 2002, Wiens 2002).

One of the principal difficulties in predicting species abundance under novel conditions is the inherent variability in ecological systems (Rotenberry 1986, Stauffer 2002, Wiens 2002). Predictions are often hampered by random or unpredictable fluctuations in time and/or space due to weather, food, habitat, or disturbance processes (Gutzwiller and Barrow 2001). Given the inherent variability or "noise" in ecological systems, the ability to predict abundance in novel locations, or at different times at the same location, is commonly believed to be difficult or impossible (Rotenberry 1986, Stauffer 2002).

One of the principal ways to begin to understand this otherwise unexplained variability, and to evaluate the ability of models to accurately predict abundance, is to investigate prediction error (Starfield 1997, Wiens 2002). Prediction error may be caused by the inability of models to predict under novel conditions due to environmental variability as well as interannual fluctuations in abundance (Gutzwiller and Barrow 2001). If the goal is to accurately predict abundance for conservation planning, prediction error is an important means for assessing the utility of a given model (model validation). In addition, empirical estimation of prediction error, and adjustment for prediction error, may provide a means to effectively adjust predictions of absolute abundance to novel locations or to different time frames (model calibration).

The Effective Area Model (EAM; Sisk et al. 1997, 2002) uses both habitat composition and habitat geometry in a model structure that can be parameterized with data collected in traditional bird population surveys. In order to predict species abundance in novel landscapes or regions, the EAM uses edge response functions (e.g., change in density of a given bird species as a function of distance from edge) along with broad-scale vegetation (habitat) maps. Maps may be based on remotely sensed imagery or Global Positioning System coordinates. As used here, edge refers to an abrupt transition in vegetation structure and/or composition that has the potential to affect the distribution and abundance of bird species. The EAM operates in a Geographic Information System (GIS) framework (ArcView 3.2, ESRI 2001) and was developed to be a "desktop" management tool to predict population level responses to potential changes in landscape configuration (Sisk et al. 2002).

In order to assess whether the EAM is an improvement to less complex nonspatial models, we compared the ability of the EAM to predict abundance in novel locations relative to a null model (Gotelli and Graves 1996). Models were assessed in terms of the difference between prediction and observation, referred to as prediction error. The null model provides a prediction of abundance based solely on habitat types and areas without considering edge effects, and thus ignores the effects of within-habitat heterogeneity resulting from proximity to adjacent habitat types (Sisk et al. 1997, 2002). The null model provides a benchmark for comparison to assess whether the incorporation of edge response functions into the model structure decreases prediction error.

In addition to model validation, we also wanted to assess whether the EAM predictions of abundance in novel locations could be improved via model calibration. Calibration is often used in the general sciences as a means to improve measurement of a biased or imprecise parameter by adjusting its estimate with a highly accurate measurement process that is often more difficult or expensive (Denham and Brown 1993). For ecological applications, calibration has been used to convert relatively easy to estimate population indices into estimates of abundance for a variety of taxa (Eberhardt and Simmons 1987, Graham 2002, Hamm et al. 2002, Yoo et al. 2003). Calibration may also be used to adjust for inherent spatial and temporal variability in ecological systems. It is in this sense that the term "calibration" is used in this paper. Currently, parameterization of the EAM is limited to edge response functions that remain constant in time and space. Here we propose a framework to improve the accuracy of abundance predictions and to extend usage of the EAM by adjusting the model parameterization to fit the unique attributes of new locations and/or time frames. Calibration of EAM predictions to new regions and inclusion of covariates describing site-level habitat characteristics may be an efficient way to further incorporate habitat models along with the EAM to improve prediction of absolute abundance.

The overall goals of this paper were to validate and calibrate the Effective Area model for the purpose of predicting absolute abundance in novel locations. For model validation, our primary objective was to compare the relative predictive ability of the EAM to a null model, by species, in terms of prediction error. Comparisons were made for all validation sites, as well as subsets of validation sites from different regions, and varying degrees of isolation, presence of water, and different habitat types. Our second objective was to assess whether the consideration of some site-level factors consistently reduced EAM prediction error and would be candidate variables to include in a calibration. Our final objective was to explore a framework for producing field-calibrated models to improve the prediction of absolute abundance in new locations.

Methods

Predicted abundance

We predicted abundance for the 25 most common bird species in 50 validation sites for both the Effective Area Model (EAM) and the null model. In order to parameterize the EAM, we estimated edge response functions from data obtained on the San Pedro River, Arizona, USA, along with habitat maps developed for each validation site. For the null model, parameter estimates were obtained as the average interior density

of a given species in a given habitat type, along with the validation site maps, to predict abundance for each site.

Field protocol.—We collected data to parameterize both the EAM and null model on the upper and middle San Pedro River in Cochise County, Arizona, USA. Study sites extended from 2 km north of the U.S.–Mexico border to 145 km north of the international border at Cascabel, Arizona. Elevation varied from 1300 m at the international boundary to 900 m in the most northern study site. Study sites included 16 areas within the San Pedro Riparian National Conservation Area (NCA) managed by the Bureau of Land Management (BLM) as well as seven sites on privately owned land north of the NCA. The study sites provided an excellent opportunity to study mostly natural edges between adjacent habitats. By surveying birds within distinct habitat types and across edge types, it was possible to develop edge response functions and to estimate average density within the interiors of specific habitat types for 25 species.

The San Pedro River watershed has two primary zones of riparian vegetation extending perpendicular from the river thalweg, or channel low point, upslope to expansive desert scrub communities. Adjacent to the river is a primary riparian zone that consists of gallery forests dominated by Freemont cottonwood (*Populus fremontii*) and Goodding's willow (*Salix gooddingii*). A secondary riparian zone consisted of mesquite (*Prosopis* spp.) interspersed with patches of Sacaton grass (*Sporobolus wrightii*). Distinct natural edges occurred between the primary and secondary riparian and desert scrub communities.

We investigated eight edge types representing both sides of four distinct habitat edges. Four edge types occurred at the primary to secondary riparian interface, including both sides of the edge where cottonwood was adjacent to mesquite, and both sides of the edge where cottonwood was adjacent to grassland. Two edge types occurred between the secondary riparian habitats at both sides of the mesquite–grassland edge. Two edge types occurred at the secondary riparian to nonriparian interface, including both sides of the mesquite–desert scrub edge. Maximum widths of sampled cover types were 170 and 210 m for cottonwood adjacent to mesquite and grassland; 352, 298, and 271 m for mesquite adjacent to cottonwood, grassland, and desert scrub; 532 and 508 m for grassland adjacent to cottonwood and mesquite; and 660 m for desert scrub adjacent to mesquite.

To estimate bird density across the different types of edges, we established 210 point transect locations on 23 sites in the upper and middle reaches of the San Pedro River. Each site consisted of 11–14 point transect locations on two or three transects. Points were located 100 m apart along the habitat gradient perpendicular from the river extending through primary riparian, to secondary riparian, to desert scrub vegetation communities. We used points located at varying distances from habitat edges to construct edge response functions. While the majority of points contained only one habitat type, points close to edges contained the focal as well as varying proportions of the adjacent habitat type, which appropriately accounted for the influence of the adjacent habitat type close to edges on bird density.

We used distance sampling to estimate detection functions (Buckland et al. 2001). At each point, we estimated distances from the observer (i.e., point center) to individual birds by sight or sound with the aid of a Yardage Pro 400 Laser Rangefinder (Bushnell Performance Optics, Lenexa, Kansas, USA). In order to accurately assign birds to the appropriate distance from edge category, we recorded sighting $\leq$60 m from the survey point. Point transect locations were surveyed for five minutes after an initial one-minute wait period. Additionally, in order to meet the distance sampling assumption of perfect detection at the point, and to reduce potential bias resulting from evasive movement in response to an observer, birds were recorded as an observer approached a point. Movement of birds was recorded to avoid double counting. We conducted surveys from 10 minutes before sunrise until three hours after sunrise. The order in which transects were walked was varied systematically to avoid bias related with time of day.

Each point was visited between six and 15 times during the 1998–2001 field seasons for a total of 2082 point transect surveys. Ten experienced observers conducted surveys during the 4-yr study. At the beginning of each field season, and prior to data collection, observers had a minimum of 2.5 weeks of training on survey techniques including identification of birds by sight and sound. Observers were rotated between sites so that each point transect location was surveyed approximately the same number of times by each observer, each year.

Distance sampling analysis.—The distribution of horizontal sighting distances from the observer to individual birds was entered into program Distance 3.5 to estimate detection functions separately for each species (Thomas et al. 1998, Buckland et al. 2001). The detection function provides an estimate of detection probability for each species out to a given radius from the survey point. Detection probability, combined with the mean count by point of a species, provides a density estimate for each point location. Encounter rate data were combined for a given species at each point location for all visits within and between years, though sighting distances used to estimate detection functions were pooled across many points. Differing numbers of visits to each point transect were accounted for by incorporating a survey effort multiplier in the analysis (Buckland et al. 2001). To improve estimation of the detection function, 5–15% of the distance data were truncated at distances near 60 m (Buckland et al. 2001).

Many factors affect the detection probability of a species (Yoccoz et al. 2001). We attempted to account

for as much of this heterogeneity as possible in the study design by balancing observers and time of day across survey points. Additionally, we evaluated candidate detection functions that included habitat type as a potentially significant source of heterogeneity in detectability. In general, the detection function for each species was allowed to vary by each of the four major habitat types: cottonwood, mesquite, grassland, and desert scrub. At the other extreme, points within all four habitats were pooled to obtain a single detection function for a given species. Intermediate approaches included pooling two or more habitats. Whether habitat types were pooled in relation to the detection function depended on the audio and visual characteristics of a given bird species in each habitat type as well as sample size considerations (>50 detections). It was also necessary to pool the detection function across habitat types when a particular bird species was rare in one or more habitat, which tends to underestimate differences in density between habitat types since the same detection function is used. The best detection function selected via Akaike information criterion (AIC; Burnham and Anderson 2002) was subsequently used to estimate density at all points where a species was detected.

Distances of point transect locations to the closest edge were based on differentially corrected Universal Transverse Mercator (UTM) coordinates of points and edges obtained with a Trimble GPS unit (Trimble Navigation, Sunnyvale, California, USA) accurate to 2–5 m, a classified image of the upper San Pedro (U.S. Army Topographic Engineering Center 2001), and field reconnaissance. The number of point transect locations varied from 18 to 42 locations per edge type. Estimates of density (dependent variable) as a function of distance from edge (independent variable) were used to estimate continuous edge response functions.

EAM and null model parameterization.—Using piecewise linear regression or simple linear regression we estimated 200 edge response functions for 25 species at eight edge types. Piecewise linear regression (Toms and Lesperance 2003) was evaluated when density appeared to show a threshold response to distance from edge. Of the total species–edge type combinations, the piecewise linear regression model converged on parameter estimates in 135 cases. In the remaining 65 species–edge type combinations where a clear breakpoint was not discernable, simple linear regression analysis was used to estimate the edge response function (Fig. 1). All regression models were estimated with Stata 8.0 (StataCorp 2003).

The coefficients estimated via regression modeling were used as input parameters to the EAM and null model. Input parameters included: (1) density at the edge, (2) density in the interior, and (3) $D_{\max}$, representing the maximum distance of edge influence (Fig. 2). Piecewise linear regression provided direct estimates of these parameters (Fig. 2). With simple linear regression, $D_{\max}$ was set equal to the maximum distance from edge sampled. Edge density was estimated by the intercept (density when distance = 0), and the interior density was calculated by using the linear regression equation: interior density = (edge density) + (slope parameter) × (maximum distance sampled). Even though density estimates from point transect locations may show spatial dependence, we concluded that use of estimates pooled across sites did not greatly affect the parameterization of the EAM and the null model since least squares methods give unbiased estimates of parameter coefficients even with dependency in the data (Zar 1999, Greene 2000). Similarly, the assumption of equal variance could be relaxed when the primary goal was to estimate regression parameters rather than conduct a statistical test on the regression equation (Zar 1999). To parameterize the null model, we estimated the interior density for each of the 200 species–edge type combinations. Interior density was considered to be constant across an entire habitat patch regardless of distance from adjacent habitats. EAM and null model density inputs, adjusted by the vegetation areas with and without edge effects, respectively, provided estimates of predicted abundance.

Validation study sites.—We used 50 sites located from <1 km to >200 km from the San Pedro River to evaluate the predictive performance of the two models. Validation sites were established in four distinct regions: (1) San Pedro River sites that were not used to estimate the edge response functions, (2) within the east and west ranges of Fort Huachuca, (3) Sonoita Creek Preserve managed by The Nature Conservancy, and (4) Empire Cienega National Riparian Conservation Area managed by the BLM. Study sites were located in Cochise, Pima, and Santa Cruz Counties and ranged in elevation from 900 to 1200 m and in latitude from 31°30′ N to 32°20′ N. Seven sites were set up in 2000 for pilot validation work, and 47 additional sites were set up in 2001. Four sites that had been used in 2000 were no longer accessible, and were dropped in 2001.

Validation sites varied by four site-level factors: region, focal habitat, degree of isolation, and presence/absence of surface water. First, sites were classified as within (18 sites) or beyond (32 sites) the San Pedro River riparian corridor. Thirty four of the validation sites were dominated by cottonwood and 16 by mesquite. Cottonwood habitat was adjacent to mesquite, grassland, or both. Mesquite habitat was adjacent to desert scrub, or both cottonwood and desert scrub. Cottonwood sites were classified as isolated if they were >100 m from the primary riparian corridor of the San Pedro River, Sonoita River, or the Cienega River. All cottonwood sites on Fort Huachuca were considered isolated, since they were all narrow drainages eventually discharging into the San Pedro River. Mesquite sites were classified as isolated if they were narrow strips of vegetation occurring in washes (<50 m wide), and nonisolated if they were part of the much wider secondary riparian corridor immediately adjacent to the primary riparian

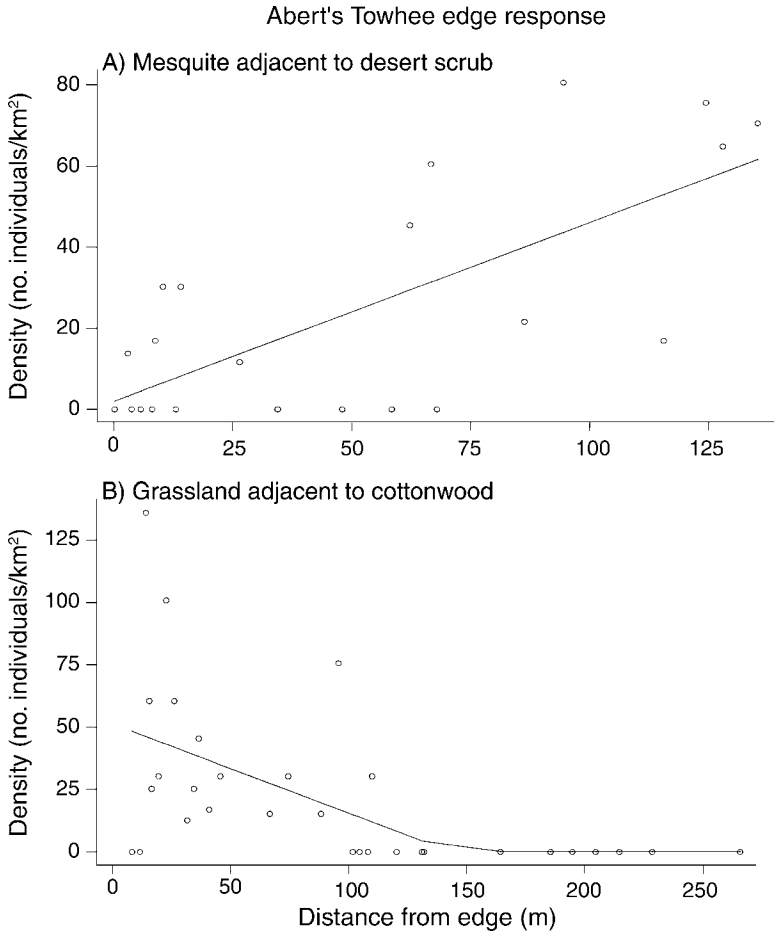

FIG. 1. Edge response functions (density vs. distance from edge) for Abert's Towhee (A) in mesquite adjacent to desert scrub using simple linear regression, and (B) in grassland adjacent to cottonwood using piecewise linear regression.

corridor. If surface water was not present during any of the visits to the site, sites were classified as surface water absent; otherwise they were classified as surface water present.

We developed habitat maps for the EAM and null model using a Trimble GPS unit to record UTM coordinates along edges between habitat types. The perimeter of habitat types within validation sites was grouped and differentially corrected in Pathfinder Office (Trimble software) to obtain accuracy of 2–5 m. Location data were brought into ArcView 3.2 for the creation of habitat patch polygons.

At each validation site, point transect locations were established and each point location was buffered to 60 m. To standardize survey area between predicted and observed abundance estimates, the buffered polygon was used as the outer limit of predicted abundance for both models.

Observed abundance

The number of bird survey points at each site varied positively with habitat patch size. Of the 50 sites, five sites had two point transect locations and one site had 1, while the remaining 44 sites contained three point transect locations. There were 95 plots in 34 cottonwood sites, and 47 plots in 16 mesquite sites, for a total of 142 plots. Each site was visited between three and six times for ~450 surveys across the 2000–2001 field seasons. The order in which sites were visited was varied systematically in order to avoid bias related to time of day.

Distance sampling (see *Methods: Predicted abundance: Distance sampling analysis*) was used to obtain an abundance estimate for each species in each of the 50 sites. The area of each validation site, calculated within ArcView 3.2, was used to obtain abundance estimates per site. To obtain an observed per-site abundance estimate to compare with the predicted abundance for each site, data were combined for a given species at a given site for all visits within and between years. Two different observers conducted surveys during the two years, with a single observer (L. A. Brand) doing 92% of the validation surveys in both years. Sufficient detections were obtained to enable estimation of abundance

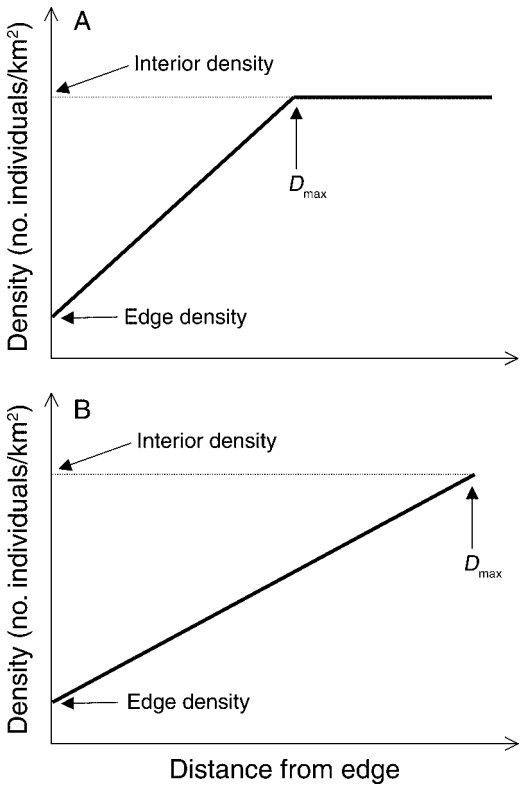

FIG. 2. The three EAM input parameters: edge density, interior density, and D_{max} (representing the maximum distance of edge influence), as derived from (A) piecewise linear regression and (B) simple linear regression.

of 20 species by validation site. All data were analyzed in the program Distance 3.5 (Thomas et al. 1998).

Akaike information criterion (AIC) model selection (Burnham and Anderson 2002) was used to select the best detection function to estimate per-site abundance for each of the 20 species. Eight candidate detection functions were assessed, including detection functions allowed to vary by habitat type, riparian corridor width, and adjacent habitat type. In addition, a global detection function was estimated based on pooling across sites. We selected the best model via AIC with a minimum of 50 detections. All candidate detection functions were assessed for each species (Buckland et al. 2001). Due to relatively small sample sizes for the species-specific analyses, only two detection functions were selected using AIC: focal habitat (separate detections for cottonwood vs. mesquite) or global (pooling across all sites).

Comparison of observed and predicted

Model validation.—We evaluated the relative predictive performance of the EAM and null model from the validation study sites for 20 species with adequate data. In order to assess the relative performance of the models, we compared prediction error between observed and predicted abundance using both models for each of the 20 species. Mean absolute prediction error (henceforth, prediction error, PE) was calculated separately for each species for each of the two models (EAM and null) as

$$\mathrm{PE} = \frac{\sum_{i=1}^{50} |E(\mathrm{obs}_i) - E(\mathrm{pred}_i)|}{50} \qquad (1)$$

where E equals the expected value and i indexes the validation site.

To obtain an estimate of the expected value of the observed abundance for each validation site, we regressed observed abundance (dependent variable) onto predicted abundance (independent variable) for each species in each validation site separately for each model. The fitted line from the linear regression model represents the expected value of observed abundance for different levels of the predicted abundance (i.e., different validation sites). We call these lines the regression of observed vs. predicted (illustrated without x marked lines, Fig. 3). The rationale for this approach is that the regression estimates the expected value of observed abundance, at a given predicted abundance, and smoothes the inherent variability in the observed data. Much of this variability is believed to be related to sampling variation and should not be attributed to prediction error. Finally, we compare the observed vs. predicted line to the line of perfect fit (y intercept = 0 and slope = 1 illustrated with x marked lines; Fig. 3). Prediction error is now estimated, separately for each model, in terms of the distance between the observed vs. predicted regression line and the line of perfect fit. The difference between the observed vs. predicted line and the line of perfect fit provides a clear comparison between the prediction error of the EAM relative to the prediction error of the null model.

Model comparisons were based on 10 000 bootstrap replicates of the 50 validation sites with replacement to estimate the distribution of prediction error for the EAM and null model, and the difference in prediction error between the models (Efron and Tibshirani 1993). It was necessary to use the bootstrap since the independence assumption of a paired t test was not met. The bootstrap process computed the mean absolute value of the difference between the observed vs. predicted line and the line of perfect fit (analogous to the residuals) across the 50 validation sites. We then calculated bootstrap 95% confidence intervals and P values to test the hypothesis that the prediction error of the EAM is less than the prediction error of the null model (Efron and Tibshirani 1993).

Model calibration.—When attempting to predict abundance in locations (validation sites) different from those used to parameterize the model (San Pedro River), substantial prediction error may result from many sources of spatial or temporal variability. To assess whether calibration would improve prediction of absolute abundance in novel locations, we calculated EAM prediction error across all validation sites for each

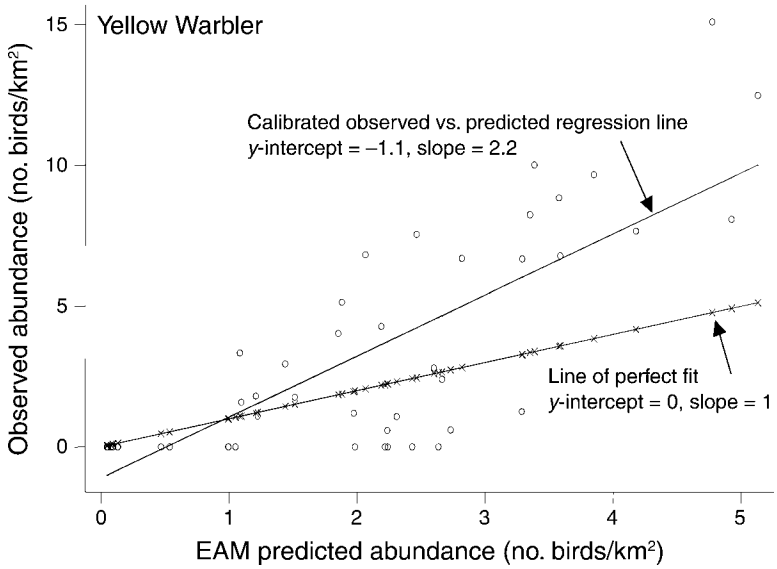

FIG. 3. Yellow Warbler observed abundance vs. EAM predicted abundance for each validation site showing the calibrated observed vs. predicted regression line (without x marks), and the line of perfect fit (with x marks).

species. Additionally, we used subsets of the 50 validation sites to investigate whether inclusion of four site-level variables (i.e., isolation, presence of water, region, and focal habitat) in a calibration improves the ability of the EAM to predict abundance in validation sites that are categorically different from those in which the models were parameterized.

Calibration can be achieved by regressing the predicted abundance on the observed abundance using a basic regression model:

$$\hat{y} = \beta_0 + \beta_1 \times \mathrm{EAM}_{\mathrm{abun}} \quad (2)$$

where y = observed abundance, $\mathrm{EAM}_{\mathrm{abun}}$ = predicted EAM abundance, and β_1 = slope coefficient. Regression of observed vs. predicted abundance enables empirical estimation of the slope and intercept parameters to adjust or correct abundance predictions for novel locations. Additional covariates (e.g., continuous covariates such as canopy height or other variables easily estimated from remotely sensed data) could be used, but we restricted our analyses to binary site-level factors in the following ANCOVA model:

$$\hat{y} = \beta_0 + \beta_1 \times \mathrm{EAM}_{\mathrm{abun}} + \beta_2 \times \mathrm{habitat} \quad (3)$$

where y = observed abundance, $\mathrm{EAM}_{\mathrm{abun}}$ = predicted EAM abundance, β_2 = site factor coefficient, and habitat = focal habitat. The additional site-level factors in the ANCOVA model shift the intercept but do not change the slope. For example, incorporation of focal habitat as an additional covariate in a model calibration enables adjustment for different average abundance associated with the new location depending on vegetation type. To evaluate improvement we calculated the coefficient of determination (R^2 values) and model P values by species as a function of EAM predictions alone, EAM predictions separately by site-level variable, and EAM predictions with all four site-level variables.

RESULTS

Model validation

Based on all validation sites, the Effective Area Model (EAM) and null model had significant ($P < 0.05$) differences in mean absolute prediction error for 12 of 20 species (Fig. 4, Table 1). The EAM was not always the best model. For example, for the Yellow-breasted Chat the EAM outperformed the null model, but for the Black-throated Sparrow the null model outperformed the EAM (Fig. 4). Of the 12 species for which it was possible to discern a significant difference between models, the EAM performed better than the null model for 10 species (Table 1).

When comparing models separately for subsets of the site-level variables, the EAM generally outperformed the null model across species and validation sites. However, three of the four site-level variables affected the relative performance of the models. The EAM showed better prediction relative to the null model in cottonwood focal habitat, isolated patches, and in sites where water was absent (Table 2). There was no difference based on region in terms of the percentage of species for which the EAM outperformed the null model (Table 2). The EAM most outperformed the null model in isolated patches and in patches without water; 91% of the species were better predicted by the EAM when it was possible to discern a difference between the two approaches (Table 2). No difference in prediction error occurred between the two models for eight of 20 species when considering all validation sites, and for 9–15 of 20 species when

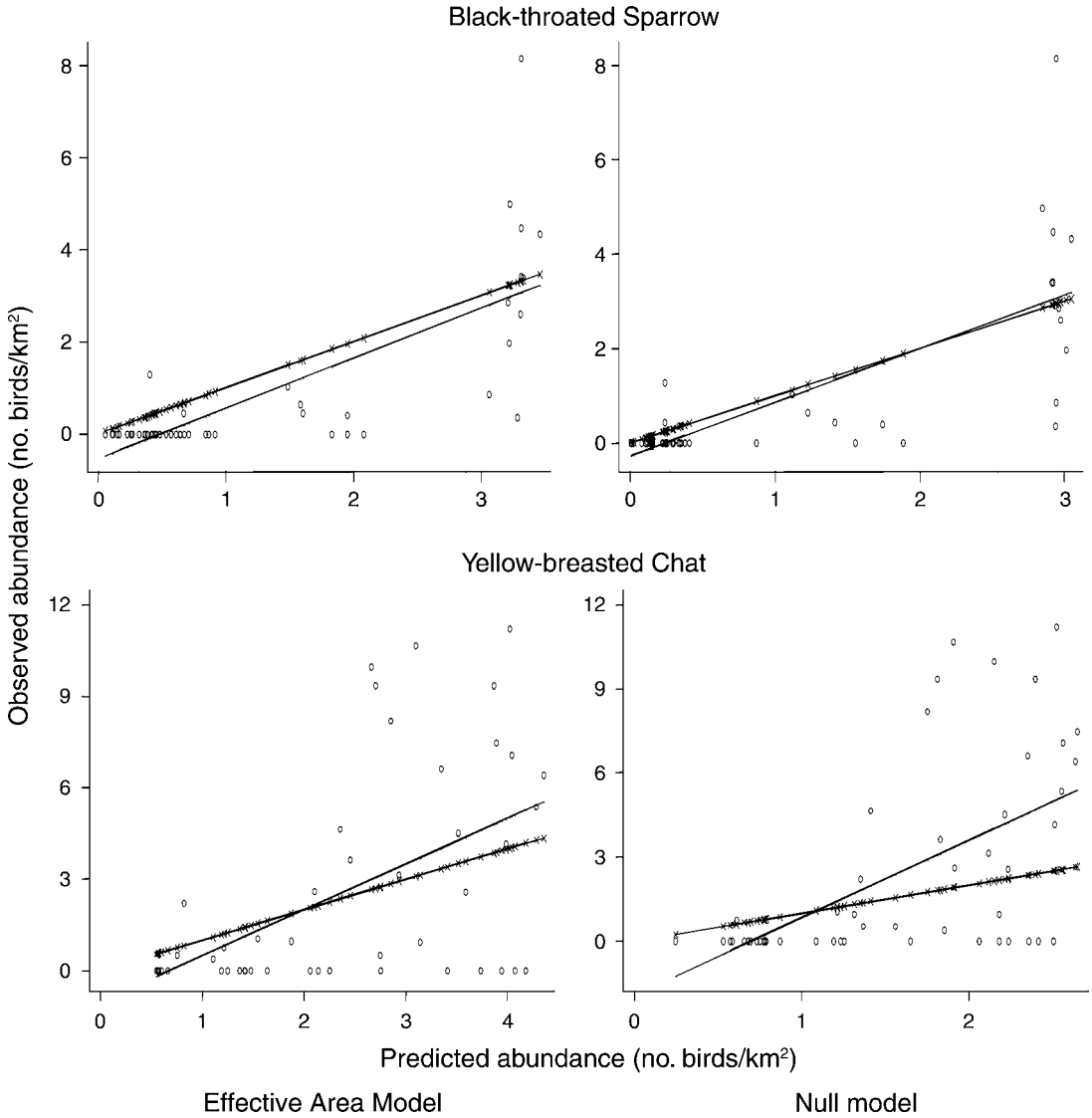

FIG. 4. Black-throated Sparrow and Yellow-breasted Chat observed vs. predicted abundance regression line without x marks and line of perfect fit with x marks, for both the EAM and the null model.

considering subsets of validation sites, though samples sizes were quite low for some subsets (Tables 1, 2).

Model calibration

When considering all 50 validation sites, the 95% confidence interval of EAM prediction error did not overlap zero for any of the species (Fig. 5). In particular, five species (Cassin's Kingbird, Lucy's Warbler, Brown-crested Flycatcher, Song Sparrow, and Yellow Warbler) showed substantial prediction error (Fig. 5). These results indicate that calibration was needed to improve the ability to predict abundance for common desert riparian bird species in novel locations.

When subsets of validation sites were selected using site-level variables, nine of 20 species showed a significant ($P < 0.05$) change in EAM prediction error as a function of one or more of the four site-level covariates (focal habitat, isolation, presence of water, and region). Prediction error was reduced in both isolated (vs. contiguous) sites and sites where water was absent (vs. present) for five of the 20 species (Common Yellowthroat, Song Sparrow, Summer Tanager, Yellow-breasted Chat, and Yellow Warbler; Fig. 6). Prediction error was lower in mesquite (vs. cottonwood) focal habitat for six species (Brown-crested Flycatcher, Cassin's Kingbird, Lesser Goldfinch, Song Sparrow, Summer Tanager, and Yellow Warbler; Fig. 6) and lower in cottonwood for only Abert's Towhee. Only Song Sparrow had increased prediction error on sites away from the San Pedro River. These results indicate that absolute abundance predictions in novel locations may be improved by incorporating additional site-level

TABLE 1. Prediction error (PE) by species for null model, Effective Area Model (EAM), and difference between null and EAM models (ΔPE).

	PE		ΔPE			
Species	Null	EAM	Null − EAM	95% CL	P	Best model†
ABTO	0.457	0.463	−0.005	−0.466, 0.266	0.456	either
ATFL	0.404	0.462	−0.058	−0.214, 0.086	0.187	either
BCFL	1.776	1.646	0.130	0.105, 0.151	0.001	EAM
BEWR	0.738	0.591	0.147	−0.103, 0.224	0.103	either
BHCO	1.021	0.852	0.169	0.079, 0.257	0.005	EAM
BLGR	0.586	0.459	0.126	0.085, 0.163	0.000	EAM
BTSP	0.197	0.425	−0.228	−0.321, −0.002	0.022	null
BUOR	0.757	0.529	0.228	0.072, 0.272	0.007	EAM
CAKI	4.134	3.979	0.155	0.100, 0.221	0.000	EAM
COYE	1.188	1.123	0.064	−0.029, 0.164	0.089	either
GIWO	0.759	0.743	0.016	−0.127, 0.159	0.385	either
HOFI	0.777	0.568	0.210	0.158, 0.252	0.000	EAM
LEGO	1.127	1.146	−0.019	−0.098, 0.059	0.309	either
LUWA	3.490	2.152	1.339	1.013, 1.693	0.000	EAM
MODO	0.622	0.614	0.008	−0.013, 0.030	0.234	either
SOSP	2.397	2.480	−0.083	−0.409, 0.291	0.299	either
SUTA	0.177	0.650	−0.473	−0.530, −0.134	0.005	null
WWDO	0.566	0.467	0.099	0.069, 0.131	0.000	EAM
YBCH	1.312	0.568	0.743	−0.068, 0.989	0.030	EAM
YWAR	1.752	1.601	0.151	−0.031, 0.379	0.053	EAM

Note: Species codes are as in Table 3.
† Best model associated with ΔPE.

variables, and that the EAM predictions were robust across regions for the majority of species.

Of the six models examined for improvement in prediction by the inclusion of covariates, we found at least one significant model for 17 of the 20 species examined, with R^2 values >25% for 12 species, >48% for six species, and >60% for four species (Table 3). The EAM by itself explained a significant amount of variability in observed abundance for 11 of 20 species, with R^2 values >23% for eight species, and >45% for three species (Table 3). At least one additional site-level variable significantly improved predictions for the majority of species examined (Table 3).

TABLE 2. Number of species for which the EAM or null model performed better in terms of the difference in prediction error (as in Table 1).

Site factors	N	EAM better	Null model better	%†
All sites	50	10	2	83.3
Isolation				
Contiguous	36	6	4	60.0
Isolated	14	10	1	90.9
Water				
Present	13	6	4	60.0
Absent	37	10	1	90.9
Focal habitat				
Cottonwood/deciduous	34	11	3	78.6
Mesquite	16	4	2	66.7
Region				
San Pedro River	18	4	1	80.0
Not San Pedro	32	8	2	80.0

Notes: Results are presented for all validation sites, as well as for a subset of sites based on four site-level factors (isolation, presence of water, focal habitat, and region).
† Percentage of species for which EAM gave the better prediction.

DISCUSSION

Abundance is strongly related with distance from habitat edge for many bird species due to abiotic and biotic factors operating at edges (Brand 2004, Ries et al. 2004). Given the strong relationships between abundance, habitat type, and distance from edge, it is essential to incorporate these factors into predictive models (Sisk et al. 1997, 2002, Ries et al. 2004). Despite the large amount of effort given to the study of bird response to habitat edges (e.g., Brand 2004, Battin 2004, Fletcher 2005), the Effective Area Model (EAM) is the only tool currently available that projects information from edge response studies to population level responses at the landscape scale (Ries et al. 2004). Quantitative predictions that incorporate the spatial arrangement of vegetation types in the landscape coupled with the responses of a species to edges represent an important tool for conservation planning.

Comparison of the performance of the EAM with the null model is a means to empirically assess the importance of incorporating edge response and patch context for predicting avian abundance. Comparing the relative predictive performance of the EAM to a null model that ignores edge effects, we found the EAM to have lower mean absolute prediction error than the null model. We found this to be true for all validation sites as well as subsets of validation sites. Overall, it appears that including information on the response of bird to

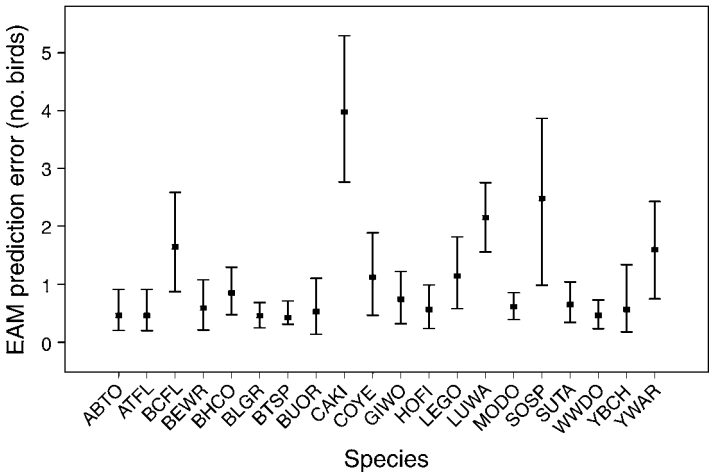

FIG. 5. EAM absolute prediction error (no. birds; defined in Eq. 1), means with 95% CI, for 20 species from 50 validation sites. Species codes are as in Table 3.

habitat edges, as well as to habitat composition, provides more accurate predictions of abundance in novel locations.

The EAM outperformed the null model but not consistently. For example, the null model predicted abundance of Black-throated Sparrow and Summer Tanager more accurately than the EAM when considering all validation sites. The reasons are unclear since both species showed a significant edge response in three of the eight edge types sampled in the validation sites (Brand 2004). One explanation is that some species may not consistently demonstrate edge responses across regions. Currently it is unknown why, and to what extent, edge responses change for a given species from one region to another with the same habitat types (see Ries et al. 2004). Additionally, situations in which the models don't predict well may be due to limited power to detect subtle responses, lack of model generality, and failure to include key predictor variables.

No difference between the predictive performance of the two models for certain species could be due to lack of importance of edge effect for those species. All species are expected to show neutral responses at some edge types (Ries et al. 2004). For the eight species for which both models were equally valid, two (Bewick's Wren, Song Sparrow) had neutral edge responses in seven of the eight edge types, three (Gila Woodpecker, Lesser Goldfinch, and Mourning Dove) had neutral edge responses in six of the eight edge types, two (Common Yellowthroat, Ash-throated Flycatcher) had neutral edge responses in five of eight edge types, and one species (Abert's Towhee) had four of eight neutral edge responses (Brand 2004). Given the relatively high number of species for which it was not possible to discern a difference between the models for subsets of validation sites, it is also likely that lack of statistical power made it impossible to predict all but pronounced edge effects.

Even though the EAM reduced predicted error in novel locations when compared with the null model,

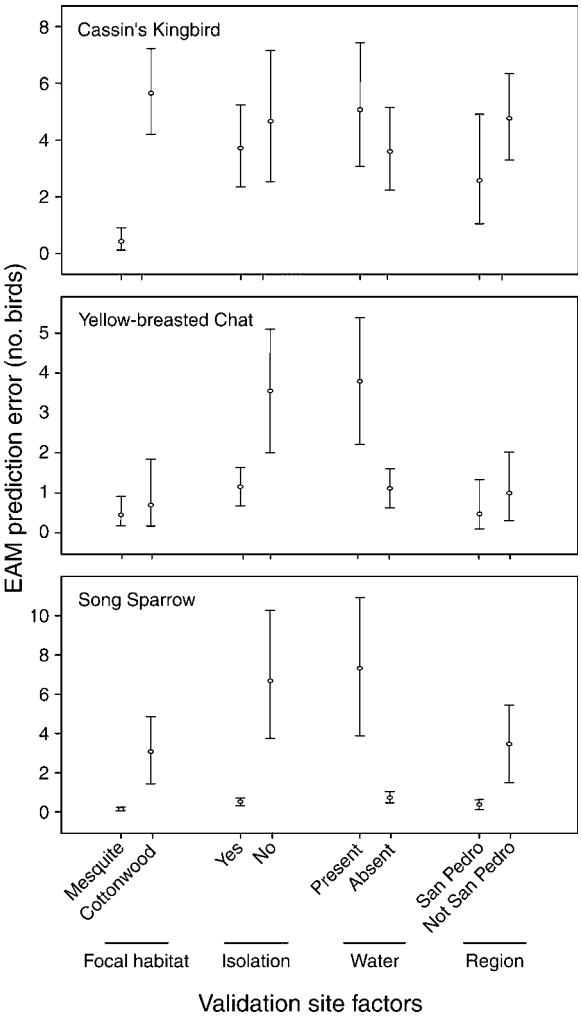

FIG. 6. EAM absolute prediction error (no. birds; defined in Eq. 1), means with 95% CI, for Cassin's Kingbird, Yellow-breasted Chat, and Song Sparrow as a function of four site-level variables: focal habitat, isolation, presence of water, and region.

TABLE 3. Coefficient of determination ($r^2 \times 100$) representing percentage of variability explained in observed abundance, by species, as a function of EAM prediction alone and EAM prediction plus site-level variables for calibrated models.

Species code	Species name	EAM	EAM + focal habitat	EAM + isolation	EAM + region	EAM + water	EAM + all variables
ABTO	Abert's Towhee *Pipilo aberti*	0.7	6.5	2.9	20.2**	1.0	25.6*
ATFL	Ash-throated Flycatcher *Myiarchus cinerascens*	3.2	3.9	15.1*	4.7	10.9	16.4
BCFL	Brown-crested Flycatcher *Myiarchus tyrannulus*	1.0	14.1*	4.5	4.2	4.1	16.8
BEWR	Bewick's Wren *Thryomanes bewickii*	23.3***	27.2***	27.1***	24.4**	25.4**	34.7**
BHCO	Brown-headed Cowbird *Molothrus ater*	1.0	3.0	1.4	1.4	2.4	10.3
BLGR	Blue Grosbeak *Guiraca caerulea*	9.8*	9.8	10.3	9.8	10.7	13.8
BTSP	Black-throated Sparrow *Amphispiza bilineata*	58.3***	61.4***	58.5***	59.3***	58.3***	61.7***
BUOR	Bullock's Oriole *Icterus bullockii*	0.0	4.3	8.0	0.1	1.0	13.7
CAKI	Cassin's Kingbird *Tyrannus vociferans*	26.1***	35.4***	26.1***	26.3***	26.1***	37.0***
COYE	Common Yellowthroat *Geothlypis trichas*	5.7	9.9	32.2***	8.5	40.8***	48.5***
GIWO	Gila Woodpecker *Melanerpes uropygialis*	28.1***	28.1***	28.2***	28.2***	28.5***	29.1**
HOFI	House Finch *Carpodacus mexicanus*	0.4	3.7	0.6	8.4	4.5	10.9
LEGO	Lesser Goldfinch *Carduelis psaltria*	27.5***	31.2***	28.1***	28.2***	27.8***	32.4**
LUWA	Lucy's Warbler *Vermivora luciae*	7.2*	25.4**	7.3	13.9*	9.4	28.8**
MODO	Mourning Dove *Zenaida macroura*	9.8*	9.8	11.3	11.9*	13.8*	16.2
SOSP	Song Sparrow *Melospiza melodia*	45.9***	52.0***	59.6***	46.2***	61.2***	67.5***
SUTA	Summer Tanager *Piranga rubra*	0.0	36.8***	33.4***	11.1	34.3***	49.5***
WWDO	White-winged Dove *Zenaida asiatica*	5.0	5.1	5.4	7.6	12.5*	19.4
YBCH	Yellow-breasted Chat *Icteria virens*	29.9***	30.0***	64.7***	31.5***	67.8***	77.6***
YWAR	Yellow Warbler *Dendroica petechia*	63.6***	66.4***	72.0***	63.6***	68.2***	73.1***
Mean across species		17.3	23.2	24.8	20.5	25.4	34.2

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

prediction error was observed for all 20 species, an outcome commonly found in other studies (Rotenberry 1986, Morrison et al. 1987, Wiens 2002, Whittingham et al. 2003). Many sources of environmental variability may contribute to prediction error. The most obvious of these arises from the fundamental differences between San Pedro River sites used to parameterize the model and validation sites used to test the model. Regional differences between model parameterization and validation sites included differences in the landscape surrounding the riparian corridor, with the San Pedro riparian corridor surrounded by Chihuahuan or Sonoran desert and validation sites on the Sonoita Creek Preserve, Fort Huachuca, and Empire Cienega surrounded by high desert grassland. In addition, interannual variability may have been a source of prediction error in that parameterization and validation sampling occurred during different years (1998–2001 and 2000–2001 field seasons, respectively). Year-to-year variation in regional abundance has been found to occur in many other breeding bird studies (Noon et al. 1985, Villard et al. 1995, Wesolowski and Tomialojc 1997, Gutzwiller and Barrow 2001) and is likely to have affected prediction errors. Interannual or regional differences may have affected the species pool or species relative abundances, which in turn could affect interspecific interactions such as competition or predation. In summary, there were many sources of potential variability that likely affected the magnitude of bias between observed and predicted abundance.

In most cases, the accuracy of EAM predictions improved with the inclusion of site-level variables. For example, the relative performance of the EAM vs. null model was greatly improved by the inclusion of information distinguishing isolated dry sites from contiguous wet sites (91% vs. 60%, EAM prediction

better than the null model). The significance of local covariates suggests that edge response may change due to subtle differences in habitat characteristics and resources, even if gross habitat composition remains constant.

Other attempts to predict absolute abundance in novel locations or at different time periods have been largely unsuccessful (Rotenberry 1986, Morrison et al. 1987, Wiens 2002, Whittingham et al. 2003), perhaps due to inherent variability between time frames or locations. For example, Morrison et al. (1987) found that predicted abundances at the same locations at a different time underestimated observed abundances by 25–50%. Rotenberry (1986) and Wiens (2002) attempted to use habitat models that had explained up to 70% of the variation in observed abundance to predict abundance in new locations and time frames and found these models did not predict well. Whittingham et al. (2003) applied predictive habitat models to new locations and found that models developed in one region performed poorly when applied in other regions. However, Whittingham et al. (2003) also noted that observed Skylark territories were significantly correlated with predicted Skylark territories, even if they did not predict a 1:1 relationship. From this result, they cautioned that models developed in one place cannot be used to make absolute abundance predictions in other regions. However, they apparently did not recognize that estimating slope and intercept from a regression of observed vs. predicted abundance, as done in this paper, enables empirical estimation and adjustment for prediction error that can be used to calibrate models to new locations.

Given the strong causal relationships between animal abundance and a large number of environmental factors (Rotenberry 1986, Lowe and Bolger 2002), it is unlikely that a model based on habitat composition and edge response alone can make highly accurate predictions. However, it is encouraging that easily acquired site-level covariates helped explain components of variability in bird abundance and reduced EAM prediction error. Environmental factors that differently affect bird abundance between parameterization and validation sites likely contributed to the observed prediction error for many species. In addition to helping explain EAM prediction error, the four site-level covariates improved abundance predictions after model calibration. Observed abundance explained by calibrated models with and without site-level factors (R^2 values >25% for 12 species and >48% for six species) was quite good compared with other studies (Morrison et al. 1987). Thus, use of the site-level habitat covariates, along with calibrated EAM predictions, significantly improved prediction success.

Increased accuracy of abundance predictions using the calibration framework presented here requires close attention to model parameterization and calibration. To predict abundance in novel locations requires that a model such as the EAM be parameterized in terms of absolute (not relative) abundance. Given the large amount of theoretical development of abundance estimation methods, edge response functions are best parameterized using distance sampling (Buckland et al. 2001) or mark–recapture methods (Seber 1982, White et al. 1982) that enable absolute abundance estimation by adjusting for detection probabilities that are heterogeneous and <1.0. Additionally, in order to calibrate predictive models to local site conditions, the target population must be clearly identified. If the EAM or other models are used to predict abundance in the calibration sample, then empirical estimation of the slope and intercept of the regression between observed vs. predicted abundance (calibration Eqs. 2 and 3) can be used to predict absolute abundance, adjusted for known site-level factors as well as for interannual, regional, or other unmeasured sources of variation.

Acknowledgments

We thank Kenneth P. Burnham, Kenneth R. Wilson, John A. Wiens, Dean F. Stauffer, and one anonymous reviewer for providing helpful comments on this manuscript. Leslie Ries, James Battin, and Haydee Hampton were very helpful during the early stages of project development. We appreciate the hard work of the field crew that assisted with data collection on the San Pedro River: Seth Factor, Devin Biggs, Carey Hill, Mark Faherty, Brian Acord, Dolly Crawford, Christopher Putnum, Janine McCabe, and Brynne Langan. The Bureau of Land Management, San Pedro Project Office, was particularly helpful during all stages of the data collection effort and provided access to the San Pedro and Empire Cienega Riparian National Conservation Areas, as well as housing for the four summers of data collection effort. In addition to the BLM, permission to work on particular study areas was granted by numerous private landowners as well as Fort Huachuca and The Nature Conservancy Sonoita Creek Preserve. Rocky Coleman and Brian Cade assisted by loaning field equipment. This project was funded by the Strategic Environment Research and Development Project (SERDP) CS-1100.

Literature Cited

Battin, J. 2004. When good animals love bad habitats: ecological traps and the conservation of animal populations. Conservation Biology **18**:1482–1491.

Block, W. M., M. L. Morrison, J. Verner, and P. N. Manley. 1994. Assessing wildlife–habitat–relationships models: a case study with California oak woodlands. Wildlife Society Bulletin **22**:549–561.

Bock, C. E., and Z. F. Jones. 2004. Avian habitat evaluation: should counting birds count? Frontiers in Ecology and the Environment **8**:403–410.

Brand, L. A. 2004. Prediction and assessment of edge response and abundance for desert riparian birds in southeastern Arizona. Dissertation. Colorado State University, Fort Collins, Colorado, USA.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. J. Thomas. 2001. An introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford, UK.

Burnham, K. P., and D. A. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer-Verlag, New York, New York, USA.

Côté, I. M., and J. D. Reynolds. 2002. Predictive ecology to the rescue? Science **298**:1181–1182.

Denham, M. C., and P. J. Brown. 1993. Calibration with many variables. Applied Statistics **42**:515–528.

Eberhardt, L. L., and M. A. Simmons. 1987. Calibrating population indices by double sampling. Journal of Wildlife Management **51**:665–675.

Efron, B., and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, New York. New York, USA.

ESRI (Environmental Systems Research Institute). 2001. ArcView. Version 3.2. ESRI, Redlands, California, USA.

Fletcher, J., Jr. 2005. Multiple edge effects and their implications in fragmented landscapes. Journal of Animal Ecology **74**:342–352.

Gotelli, N. J., and G. R. Graves. 1996. Null models in ecology. Smithsonian Institution Press, Washington, D.C., USA.

Graham, I. M. 2002. Estimating weasel *Mustela nivalis* abundance from tunnel tracking indices at fluctuating field vole *Microtus agrestis* density. Wildlife Biology **8**:279–287.

Greene, W. H. 2000. Econometric analysis. Fourth edition. Prentice Hall, Englewood Cliffs, New Jersey, USA.

Gutzwiller, K. J., and W. C. Barrow, Jr. 2001. Bird–landscape relations in the Chihuahuan desert: coping with uncertainties about predictive models. Ecological Applications **11**:1517–1532.

Hamm, K. A., L. V. Diller, and D. W. Kitchen. 2002. Comparison of indices to estimate abundance of dusky-footed woodrats. Wildlife Society Bulletin **30**:64–70.

Lowe, W. H., and D. T. Bolger. 2002. Local and landscape-scale predictors of salamander abundance in New Hampshire headwater streams. Conservation Biology **16**:183–193.

Morrison, M. L., I. C. Timossi, and K. A. With. 1987. Development and testing of linear regression predictive bird–habitat relationships. Journal of Wildlife Management **51**: 247–253.

Noon, B. R., D. K. Dawson, D. B. Inkley, C. S. Robbins, and S. H. Anderson. 1980. Consistency in habitat preference of forest bird species. Transactions of the North American Wildlife and Natural Resources Conference **45**:226–244.

Noon, B. R., D. K. Dawson, and J. P. Kelly. 1985. A search for stability gradients in North American breeding bird communities. Auk **102**:64–81.

Ries, L., R. J. Fletcher, Jr., J. Battin, and T. D. Sisk. 2004. Ecological responses to habitat edges: mechanisms, models, and variability explained. Annual Review of Ecology, Evolution, and Systematics **35**:491–522

Rotenberry, J. T. 1986. Habitat relationships of shrub steppe birds: even "good" models cannot predict the future. Pages 217–222 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. University of Wisconsin Press, Madison, Wisconsin, USA.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Second edition. Macmillan, New York, New York, USA.

Sisk, T. D., N. M. Haddad, and P. R. Ehrlich. 1997. Bird assemblages in patchy woodlands: modeling the effects of edge and matrix habitats. Ecological Applications **7**:1170–1180.

Sisk, T. D., B. R. Noon, and H. M. Hampton. 2002. Estimating the effective area of habitat patches in heterogeneous landscapes. Pages 713–725 *in* J. M. Scott, P. J. Heglund, and M. L. Morrison, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.

Starfield, A. M. 1997. A pragmatic approach to modeling for wildlife management. Journal of Wildlife Management **61**: 261–270.

StataCorp. 2003. Stata Statistical Software. Release 8.0. Stata, College Station, Texas, USA.

Stauffer, D. F. 2002. Linking populations and habitats: Where have we been? Where are we going? Pages 53–61 *in* J. M. Scott, P. J. Heglund, and M. L. Morrison, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.

Thomas, L., et al. 1998. Distance 3.5. Research unit for wildlife population assessment, University of St. Andrews, UK. ⟨http://www.ruwpa.st-and.ac.uk/distance⟩

Toms, J. D., and M. L. Lesperance. 2003. Piecewise regression: a tool for identifying ecological thresholds. Ecology **84**:2034–2041.

U.S. Army Topographic Engineering Center. 2001. Vegetation map of the San Pedro Riparian National Conservation Area and Babocomari River. Final report submitted to U.S. Army Garrison at Fort Huachuca, Arizona, USA.

Van Horne, B. 1983. Density as a misleading indicator of habitat quality. Journal of Wildlife Management **47**:893–901.

Villard, M.-A., G. Merriam, and B. A. Maurer. 1995. Dynamics in subdivided populations of neotropical migratory birds in a fragmented temperate forest. Ecology **76**:27–40.

Wesolowski, T., and L. Tomialojc. 1997. Breeding bird dynamics in a primaeval temperate forest: long-term trends in Bialowieza National Park (Poland). Ecography **20**:432–453.

White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture–recapture and removal methods for sampling closed populations. Los Alamos National Laboratory Report LA-8787-NERP, Los Alamos, New Mexico, USA.

Whittingham, M. J., J. D. Wilson, and P. F. Donald. 2003. Do habitat association models have any generality? Predicting Skylark *Alauda arvensis* abundance in different regions of southern England. Ecography **26**:521–531.

Wiens, J. A. 2002. Predicting species occurrences: progress, problems, and prospects. Pages 739–749 *in* J. M. Scott, P. J. Heglund, and M. L. Morrison, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution **16**:446–453.

Yoo, H. J., A. Stewart-Oaten, and W. W. Murdoch. 2003. Converting visual census data into absolute abundance estimates: a method for calibrating timed counts of a sedentary insect population. Ecological Entomology **28**: 490–499.

Zar, J. H. 1999. Biostatistical analysis. Fourth edition. Prentice Hall, Englewood Cliffs, New Jersey, USA.