Published in final edited form as:
. 2012 February 1; 18(2): 158–167. doi:10.1111/j.1472-4642.2011.00831.x.

Predicting patch occupancy in fragmented landscapes at the rangewide scale for an endangered species: an example of an American warbler

Bret A. Collier[1,*], **Julie E. Groce**[2], **Michael L. Morrison**[3], **John C. Newnam**[4], **Andrew J. Campomizzi**[3], **Shannon L. Farrell**[3], **Heather A. Mathewson**[1], **Robert T. Snelgrove**[1], **Raymond J. Carroll**[5], and **Robert N. Wilkins**[1]

[1]Institute of Renewable Natural Resources, Texas A and M University, College Station, TX 77843, USA

[2]Institute of Renewable Natural Resources, Texas A and M University, San Antonio, TX 77843, USA

[3]Department of Wildlife and Fisheries Sciences, Texas A and M University, College Station, TX 77843, USA

[4]Texas Department of Transportation, PO Box 15426, Austin, TX 78761, USA

[5]Department of Statistics, Texas A and M University, College Station, TX 77843, USA

Abstract

Aim—Our objective was to identify the distribution of the endangered golden-cheeked warbler (*Setophaga chrysoparia*) in fragmented oak–juniper woodlands by applying a geoadditive semiparametric occupancy model to better assist decision-makers in identifying suitable habitat across the species breeding range on which conservation or mitigation activities can be focused and thus prioritize management and conservation planning.

Location—Texas, USA.

Methods—We used repeated double-observer detection/non-detection surveys of randomly selected ($n = 287$) patches of potential habitat to evaluate warbler patch-scale presence across the species breeding range. We used a geoadditive semiparametric occupancy model with remotely sensed habitat metrics (patch size and landscape composition) to predict patch-scale occupancy of golden-cheeked warblers in the fragmented oak–juniper woodlands of central Texas, USA.

Results—Our spatially explicit model indicated that golden-cheeked warbler patch occupancy declined from south to north within the breeding range concomitant with reductions in the availability of large habitat patches. We found that 59% of woodland patches, primarily in the

*Correspondence: Bret A. Collier, Institute of Renewable Natural Resources, Texas A&M University, College Station, TX 77843, USA. bret@tamu.edu.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Author Contributions: All authors contributed to the development of the study ideas and design presented in this work. J.E.G., S.L.F., A.J.C. and H.A.M. directed field operations; R.T.S. directed spatial data analysis, B.A.C. and R.J.C. analysed the data; B.A.C., M.L.M. and R.N.W. led the writing.

northern and central portions of the warbler's range, were predicted to have occupancy probabilities ≤0.10 with only 3% of patches predicted to have occupancy probabilities >0.90. Our model exhibited high prediction accuracy (area under curve = 0.91) when validated using independently collected warbler occurrence data.

Main conclusions—We have identified a distinct spatial occurrence gradient for golden-cheeked warblers as well as a relationship between two measurable landscape characteristics. Because habitat-occupancy relationships were key drivers of our model, our results can be used to identify potential areas where conservation actions supporting habitat mitigation can occur and identify areas where conservation of future potential habitat is possible. Additionally, our results can be used to focus resources on maintenance and creation of patches that are more likely to harbour viable local warbler populations.

Keywords

Bayesian inference; golden-cheeked warbler; habitat conservation; occupancy; semiparametric regression; *Setophaga chrysoparia*

INTRODUCTION

Species inhabiting human-dominated environments often exist in locations where habitat loss and fragmentation have reduced patch contiguity, patch size, and increased edge and isolation effects (Marzluff, 2001; Bolger, 2002). Such changes in structural features at the local scale also influence dynamics in surrounding areas (Forman, 1995; Saab, 1999). Moreover, fragmentation can create ecological thresholds for demographic values such as population size (Homan *et al.*, 2004; Betts *et al.*, 2007), survival (Ruiz-Gutiérrez *et al.*, 2008), dispersal (Bayne & Hobson, 2002) and reproduction (Robinson *et al.*, 1995; Lloyd *et al.*, 2005). Such thresholds can potentially increase local extinction risk for species in areas under intensive land use pressures (Lande, 1998). Whether or not a species is structured as a metapopulation, knowledge of how habitat is occupied through space allows predictions to be made on colonization and extinction probabilities (Mills, 2007:265).

A core issue of conservation biology is the distribution of individuals through space, typically with a focus on the relationship between availability, size and proximity of potential habitats to one another. Information on how environmental metrics predict habitat use across space and time provides the foundation for population management and conservation strategies (Sagarin *et al.*, 2006; Brotons *et al.*, 2007). As such, environmental metrics from locations where presence–absence surveys have been conducted represent the basis for predictive modelling of species distributions and for monitoring changes in distributions as environmental conditions change (Moore & Swihart, 2005; Elith *et al.*, 2006; MacKenzie, 2006; Syphard & Franklin, 2009). Species inhabiting human-dominated environments present challenges for habitat modelling because they often exist in locations where habitat loss and fragmentation has reduced patch contiguity and patch size and increased edge and isolation effects (Bolger, 2002). Thus, it is likely that the underlying distribution process varies nonlinearly in space wherein we would expect, for instance, that conditions will become less similar as spatial proximity declines (e.g. spatial autocorrelation; Augustin *et al.*, 1996; Royle *et al.*, 2007). In general, the impact of incorporating spatial relationships into predictive models is an attempt to create a proxy for addressing unmeasurable or unidentifiable environmental metrics, which otherwise would not be incorporated into model structure (Tognelli & Kelt, 2004).

Our goal was to predict patch occupancy in a fragmented landscape for the golden-cheeked warbler (*Setophaga chrysoparia*) across the entirety of its breeding range. The goldencheeked warbler was listed as federally endangered in the United States in 1990 owing to concerns about habitat loss within the warbler's restricted breeding range (88,878 km^2 in central Texas, USA; Fig. 1a; U.S. Fish and Wildlife Service 1992). The warblers' endemism to central Texas during breeding is driven by its relationship to the oak (*Quercus* spp.)–Ashe juniper (*Juniperus ashei*) woodland communities that provide foraging habitat, nesting cover, and shredded bark from Ashe juniper for nest construction (Pulich, 1976; Ladd & Gass, 1999). The breeding range of the warbler in central Texas has seen an increase in human populations of approximately 50% since species listing (Groce *et al.*, 2010). Previous studies of the warbler in Texas have focused on public lands (Anders & Dearborn, 2004; Peak, 2007; Reidy *et al.*, 2008), which represent <5% of land within the warblers range; hence, few data exist for the accurate assessment of rangewide warbler distribution and factors affecting distribution (DeBoer & Diamond, 2006; Collier *et al.*, 2010). Several recent attempts to map the distribution of potential warbler habitat within the breeding range (Diamond, 2007; SCWA 2007, Loomis Austin 2008) used remotely sensed vegetative conditions that were deemed appropriate for warblers and classified patches into qualitative categories representing habitat quality (e.g. high or low quality) based on expert opinion and limited field data. However, these qualitative assessments of potential habitat were unable to quantitatively estimate likelihood of warbler presence across the species range, thus limiting their usefulness for conservation planning.

We developed a geoadditive semiparametric occupancy model (Ruppert *et al.*, 2003; Crainiceanu *et al.*, 2005; Gimenez *et al.*, 2006) for repeated detection/non-detection survey data to predict patch-specific occupancy of the golden-cheeked warbler across the breeding range in Texas. We corrected for imperfect detection (Royle & Kéry, 2007; Royle *et al.*, 2007) and allowed the spatial relationship between patch occupancy probabilities to be determined as a function of a nonparametric interaction (Ruppert *et al.*, 2003; Gimenez *et al.*, 2006; Grosbois *et al.*, 2009). The geoadditive component of our model represents the merging of the underlying semiparametric model with the spatial aspect provided by the spline basis functions as detailed in Kammann & Wand (2003) and Ruppert *et al.* (2003). Our objective was to predict those areas of high and low occurrence probability to better assist decision-makers in prioritizing conservation, management or mitigation activities across the golden-cheeked warbler's breeding range. We also detail how our sampling and analytical approach can be broadly applied to many other species in habiting fragmented landscapes.

METHODS

Study area and biological covariates

Golden-cheeked warblers are endemic to oak–juniper woodlands (Pulich, 1976; Ladd & Gass, 1999); thus, we defined woodland patches as the sampling unit on which we measured occurrence (Collier *et al.*, 2010). We delineated woodland patches based on spring 2007 and 2008 LANDSAT five imagery (30 m pixel) that maximized vegetative spectral differences after leaf emergence in spring and prior to the stresses associated with heat and drought in the summer (Collier *et al.*, 2010). We conducted an unsupervised classification of woodlands across central Texas, aggregating land cover types into two classes (oak–juniper woodland and other) and used the 2001 National Land Cover Dataset (NLCD) to mask any notable areas (e.g. croplands, wetlands) that were misclassified as woodlands by our unsupervised classification. We created breaks between patches traversed by paved roads using the Texas strategic mapping program (STRATMAP) by intersecting road data with our woodland classification and deleting any woodland classified pixels that intersected the road layer. We assigned each patch of woodland habitat to an administrative unit based on US Fish and Wildlife Service Recovery Regions (USFWS 1992). For each woodland patch, we calculated patch size (Collier *et al.*, 2010) and landscape composition (Magness *et al.*,

2006) using ESRI ArcGIS 10 as both are known to influence warbler presence (see Appendices S1 and S2 in Supporting Information). We estimated landscape composition for each patch as the mean percentage of woodlands within a 400-m radius circle surrounding a given pixel as this radius was determined to capture landscape variation relevant to warbler presence at the patch scale (Magness *et al.*, 2006). The mean value for all pixels within a patch was used as the landscape composition estimate of the patch for analysis.

Sample size and survey methodology

We determined minimum sample size following MacKenzie and Royle (2005). We used a probability proportional to size sampling design (PPS; Thompson 2002) and selected patches randomly in proportion to size for surveying so not to over- or under-weight our sample frame because of the non-normal distribution of patches sizes across the range. We focused our survey efforts on patches ≤200 ha owing to our knowledge of size-based threshold effects in species presence (He & Gaston, 2000; Butcher *et al.*, 2010; Collier *et al.*, 2010). Because we anticipated access restrictions to patches on private lands (Hilty & Merenlender, 2003), we created a randomly selected sampling frame five times greater than the minimum sample size. If we were unable to obtain access to the selected patch, we contacted landowners of the next randomly selected patch akin to assuming that missing or inaccessible properties were missing completely at random (Stevens & Jensen, 2007). We supplemented our random sampling with patches on public and private properties that fell within the bounds of our sampling design on which we currently had access. We assumed throughout our study that access to property was not influenced by known or perceived warbler presence or absence and that local management and access were unrelated, and we made no assumptions that any one patch surveyed would have a greater or lesser likelihood of warbler presence. We used standardized protocols for surveying regardless of patch ownership. We attempted to survey the entirety of every accessed patch, and a patch was contained on multiple properties where some access was restricted, and we assumed that the probability of detecting warblers did not differ between accessible and inaccessible areas of the patch.

We conducted auditory and visual surveys for warblers between mid-March and late May 2009 to determine patch occupancy (Collier *et al.*, 2010). Two simultaneous independent observers surveyed each patch systematically for warbler presence–absence (MacKenzie, 2006). If at least one observer detected a warbler within a patch during a survey period, we did not revisit the patch (e.g. a removal approach, MacKenzie, 2006). Based on previous research (Collier *et al.*, 2010), if no warblers were detected during a survey, it was necessary to resurvey patches up to a maximum of six times (three double-observer surveys) in an attempt to detect warblers. While most patches (>95%) were surveyed ≤6 times, several were surveyed one additional time by double observers, so we used a maximum number of survey occasions of 8 for our analysis. Thus, our data represented repeated detection/non-detection surveys of $i = 1, 2,\ldots, R$ patches of warbler habitat where each site was surveyed $j = 1, 2,\ldots, J$ times to determine warbler detection/non-detection during the 2009 survey season. Data resulting from our repeated presence–absence surveys represent a capture history (e.g. 0011….) wherein a positive detection is given a 1, no detection is given a 0, and no survey conducted is designated using ‘.’ notation (White and Burnham 1999).

Analysis

Although there are many statistical methods used to predict species distributions (Guisan & Zimmermann, 2000; Elith *et al.*, 2006; Guisan *et al.*, 2006; MacKenzie, 2006), generalized linear (GLM) and additive (GAMs) models represent two popular approaches for presence-absence data (Guisan & Zimmermann, 2000; Guisan *et al.*, 2006; Syphard & Franklin, 2009). Models such as these are useful for modelling linear and nonlinear effects of biological covariates (Kammann & Wand, 2003; Wood, 2003) and can be used to address spatial variation in species distribution data (Augustin *et al.*, 1996, Kammann & Wand, 2003; Knapp *et al.*, 2003; Gimenez *et al.*, 2009). However, GAMs and GLMs often do not provide predictions of the probability of occurrence within a specific location, often only allowing for statements on relative suitability (Elith *et al.*, 2006; Royle *et al.*, 2007). Thus, we used geoadditive semiparametric regression (Ruppert *et al.*, 2003; Gimenez *et al.*, 2006, 2009) to model golden-cheeked warbler patch occupancy and associated detection probabilities across the species breeding range in Texas.

The dependent variable in our analysis was the presence or absence of golden-cheeked warblers in patches of oak–juniper woodland. We modelled patch-specific occupancy probability with patch size, landscape composition and their interaction entering the model linearly with spatial predictors (latitude and longitude) incorporated into the model as a nonparametric interaction (Ruppert *et al.*, 2003; Crainiceanu *et al.*, 2005). We predicted occupancy probability (Ψ_i) as a function of those covariate data and spatial location within the range expressed as

$$\log_{\text{it}}(\Psi_i) = \beta_0 + \beta_1 X_1 + \sum_{k=1}^{20} u_k \left(Location_i - \kappa_k\right) + \varepsilon_l$$

where the $\beta_l X_1$ represents a vector of *l* predictor variables (patch size, landscape composition, patch size–landscape composition interaction, and patch-specific UTMs) entering the model linearly and ($Location_i - \kappa_k$) represents the spatial effect for each surveyed habitat patch (Gimenez *et al.*, 2009). Spatial relationships in our model used radial basis penalized splines (Ruppert *et al.*, 2003), which can be efficiently modelled in a generalized linear mixed model framework and have the added benefit of being rotationally invariant when used for geographical smoothing (Ruppert *et al.*, 2003). We fit the above model using $k = 20$ knots (Ruppert *et al.*, 2003; Crainiceanu *et al.*, 2005), which ensured adequate flexibility and used the space-filling algorithm of Nychka & Saltzman (1998) and the *R* package fields (Fields Development Team 2006) to select knot locations within our landscape. Our model was adapted for use as a single-season occupancy model (MacKenzie, 2006) and relied on the penalized spline structure detailed by Ruppert *et al.* (2003), Crainiceanu *et al.* (2005) and Gimenez *et al.* (2006, 2009). We have provided annotated WinBUGS code adapted to our specific study as an Appendix S3.

Detection model

Many species distribution modelling approaches (Knapp *et al.*, 2003; and reviews by Elith *et al.*, 2006; Guisan *et al.*, 2006) do not mention detection rates when discussing modelling methods even though variable detection rates can have significant impacts on distribution model predictions (MacKenzie, 2006; Royle & Kéry, 2007; Kéry *et al.*, 2010). Because our model was hierarchical in nature, we accounted for the impact of imperfect detection on our occupancy predictions by modelling the detection process (Royle *et al.*, 2007). We used a temporal covariate representing survey date for detection modelling as date of survey has been shown to adequately predict detection rates of warblers at the patch scale (Collier *et al.*, 2010). Thus, we addressed issues associated with imperfect detection using the linear logistic relationship (Kéry, 2008; Royle & Dorazio, 2008)

$$\log_{\text{it}}(p_i) = \alpha_0 + \alpha_1 \text{Day},$$

where Day represents the numeric day since 15 March 2009 and continuing through the end of the breeding season.

Bayesian inference

We adopted a Bayesian approach that has been shown to be computationally efficient for hierarchical generalized linear mixed models with radial basis splines (Wood *et al.*, 2002; Ruppert *et al.*, 2003; Crainiceanu *et al.*, 2005; Gimenez *et al.*, 2009; King *et al.*, 2010). We provided a set of prior distributions for all model parameters to fully specify our model (Royle & Dorazio, 2008; King *et al.*, 2010). We used normal prior's N (0,100) on the β's that enter linearly into our model and specified independent, normal priors on random effect parameters $u_k \sim (0, \tau_b)$ where $\tau_b = \tau(0, 1, 0.1)$ (Ruppert *et al.*, 2003; Crainiceanu *et al.*, 2005). We used normal priors $N(0,100)$ for the intercept and slope of the detection sub-model. We standardized covariates prior to analysis to assist with model convergence (Crainiceanu *et al.*, 2005).

We performed all analysis using WinBUGS v. 1.4 (Spiegelhalter *et al.* 2003) and *R* (R Core Development Team, 2009) using R package R2WinBUGS (Sturtz *et al.*, 2005) for the Markov chain Monte Carlo (MCMC) simulations. Annotated R and WinBUGS code for running the semiparametric occupancy model and for predicting patch-specific occupancy is available as Appendix S3. We ran our MCMC algorithm for 1×10^6 iterations after a 50,000 iteration burn-in. We thinned every 100th iteration for model diagnostics and inference. We assessed model convergence based on the Gelman and Rubin statistic (Gelman & Rubin, 1992) and through residual evaluation (Ruppert *et al.*, 2003) using *R* packages boa (Smith, 2007) and coda (Plummer *et al.*, 2006).

Model evaluation

To evaluate our semiparametric model, we used independent survey data on golden-cheeked warblers in Texas. During 2003–2007 and 2010, we conducted observational studies in woodland patches across the warbler's breeding range where presence-absence data at the patch scale was collected using those methods detailed in *Sample size and survey methodology*, which represented an optimal, data-driven approach for model validation (Guisan & Zimmermann, 2000). We followed advice of Guisan & Zimmermann (2000) and Elith *et al.* (2006) and developed receiver operating curves (ROC) and sensitivity/specificity comparisons using the independent survey data to evaluate predictive accuracy of our modelling approach (Sing *et al.*, 2005).

RESULTS

We identified and assigned biological metrics to 63,616 patches of woodlands across our study area in Texas (Fig. 1a). Our remotely sensed habitat layer included approximately 1.678 million ha of woodlands. Approximately 70% of the patches were ≤10 ha in area and encompassed about 11% of total available habitat. In 2009, we surveyed 287 patches for warbler presence and had positive detections in 150 of the 287 patches, providing a naïve estimate of patch occupancy of 52%. Surveyed patches ranged from 2 to >11,000 ha in area, and we surveyed patches in 34 of the 35 counties in the warbler's range.

Semiparametric model results

We provide summaries of the posterior distribution of model parameters in Table 1. Brooks–Gelman–Rubin diagnostics indicated model convergence with scale reduction factors for all parameters between 1.00 and 1.10 and a multivariate scale reduction factor of 1.08 (Gelman & Rubin, 1992). We developed a predictive surface of the spatial effect by predicting occupancy probability for locations on a 100-m grid distributed across our study area while holding all model parameters at their mean covariate value. There was substantial evidence of spatial variability affecting warbler distribution, as we found areas of higher relative occurrence probability on both the eastern border of the region roughly between 30°50′N and 31°55′N latitude, as well as in the south-western third of the study region south of 30°10′N and west of 99°00′W (Fig. 1b). Habitat patches towards the northern and western edge of the species range showed a much lower predicted occurrence probability of warblers.

Detection and occupancy predictions

As expected, the effect of sample survey date on detection probability was negative, indicating that detection probability declines as the breeding season progresses. Using the mean date (43 days since 15 March) for all detections, the posterior mean detection probability was 0.701. Based on this value, we estimated that the probability of not detecting a warbler when one was present within a patch of warbler habitat when surveyed by the mean date would be approximately $(1–0.701)^8 = 0.00006$, or effectively 0.

Our model predicted that 86% of the patches (see *Methods: Study area and biological covariates*) had occupancy probabilities ≤50%, and 59% had a predicted occupancy ≤10% (Table 2). Only 2.9% of patches were predicted to have occupancy probabilities >90% (Table 2), most of which were large patches in the south-western portion of the species breeding range (Fig. 1a). Patches with occupancy probabilities between 0.50 and 0.90 were more widely distributed across the range although occupancy estimates declined in the northern regions. Overall, the expected occurrence distribution at the rangewide scale indicated significant variation in occurrence probabilities dependent upon both spatial location and patch-specific metrics (Fig. 2).

Model validation

Using survey data collected during 2003–2007 and 2010 in woodland patches ($n = 143$), we classified patch detection/non-detection for comparative purposes with our occupancy predictions. We used a scoring classifier (Sing *et al.*, 2005) to visualize a ROC graph showing model prediction accuracy versus false-positive rate and sensitivity versus specificity (Fig. 3a,b). Our area under curve (AUC) estimate was high (0.91), indicating our model predicted reality based on our field survey data.

DISCUSSION

Population distribution is often limited by the amount of appropriate habitat available (Hanski & Gilpin, 1991) and the proximity or isolation of habitat patches (Shanahan & Possingham, 2006). Identification of potential habitat distribution and the ability to distinguish among patches of varying occupancy probabilities are important for driving conservation actions for species (Guisan & Thuiller, 2005). Our general approach to study design and analysis is most directly applicable to species where individuals are difficult to follow, such as smaller-bodied animals that cannot be readily sampled by telemetry and other marking methods (Fahrig, 2007:74–75).

We found that the majority (59%) of woodland patches within the breeding range of the golden-cheeked warbler were predicted to have ≤0.10 of being occupied. Thus, as it is infeasible to maintain all current habitat for warblers in Texas in perpetuity, our results can be used to (1) focus resources on maintenance of those patches with higher occupancy estimates that may be likely to harbour viable local warbler populations (He & Gaston, 2000; Collier *et al.*, 2010) and (2) identify locations where habitat management actions can assist in creating, maintaining or linking available habitat.

Our results indicate that warbler occurrence declined from south to north across the breeding range, which corresponded with a decrease in the proportion of large patches from south to north. The decrease in patch size was correlated with an overall decrease in environmental conditions supporting large patches with high canopy cover in the northernmost portion of the species range, and additionally, greater residential and commercial development in the south-east portion of the range (Groce *et al.*, 2010). Our data indicate that with an overall shift to smaller and more fragmented patches within the northern portions of the range, the probability of warbler occurrence declines significantly, even for large patches of woodland habitats.

Species distribution models incorporating imperfect detection and spatial relatedness are known to outperform standard GLMs in predictive accuracy (Royle & Kéry, 2007). Using our data as an example, consider two patches where one was in the northern edge of the warblers range and the other was in the south-western region of range. Both patches had similar patch size (25.4 and 25.9 ha) and landscape composition (57.1 and 57.2) but differed significantly in predicted occupancy from our model (northern patch occupancy, 0.10; south-western patch occupancy, 0.72). Based on our values, we can infer obvious differences in patch occupancy and hence use that information for focusing conservation at the landscape scale. However, if one were to use the same biological metrics in a standard GLM, occurrence probabilities for both patches would be 0.63, with a potential consequence of conservation effort focused naïvely, and possibly ineffectively, on the northern patch.

We adopted a set of biological covariates (patch size and landscape composition) that previous research had indicated was useful for predicting warbler occurrence at the patch scale (Magness *et al.*, 2006; Collier *et al.*, 2010). Although we are confident that the covariates we used were appropriate for determining presence-absence across the range, it is likely that additional environmental metrics, such as within-patch tree communities, juniper density, relative age of oak–juniper woodland, or other metrics may help refine occupancy in small patches as well as provide information for better prediction of warbler abundance and fecundity. However, model evaluation indicated that our model's predictive accuracy was more than adequate overall; thus, we are comfortable with the environmental metrics used. We assumed a fairly simple process for species detection rates, assuming that detection was solely a function of observation date (Collier *et al.*, 2010). However, it is plausible that more detailed determination of the detection function could be accomplished via incorporation of additional variables for factors such as observers (e.g. differing abilities), patch metrics or potential interactions between spatial location and survey date. However, given the high detection rates of warblers during our surveys, we thought that additional metrics would add unnecessary complexity to our modelling approach.

Our approach used variables relevant to a landscape scale, which translated into a model that was insensitive to fine-scale variability in habitat composition. While we acknowledge the fact that site-specific (e.g. within patch) variation in habitat characteristics can influence how warblers distribute themselves within patches of habitat (Ladd & Gass, 1999) and could potentially affect between-patch distribution, we stress that our work was not focused on evaluating mechanisms driving within-patch differences in local selection (Manly *et al.*, 2002). Rather, our approach shows how robust models of species distributions using coarse-scale metrics can be developed for supporting conservation decisions at a scale not easily attainable with local-scale models.

Our application of a geoadditive semiparametric regression occupancy model to golden-cheeked warbler breeding range survey data provides a flexible framework for predicting warbler distributions while addressing latent spatial variation and issues associated with imperfect species detection. Thus, our work builds upon others who have incorporated spatial relationships into models (Elith *et al.*, 2006), while accounting for imperfect detection (Royle *et al.*, 2007), and provides an additional framework for model-based prediction of species distributions. We applied our geoadditive model as a single-season occupancy model, where occurrence does not change; thus, our example provides a snapshot prediction of golden-cheeked warbler patch occupancy distribution across the species range. Our model can be useful for other questions linking space to demography (Grosbois *et al.*, 2009) as the general structure allows for direct incorporation of spatial relationships into future distribution models. Thus, our approach could easily be applied to dynamic models to address temporal variation in golden-cheeked warbler habitat patch occupancy state (MacKenzie *et al.*, 2002, Royle & Kéry, 2007), or given our model's structure, it could easily be modified to fit a variety of additional hierarchical models focusing on golden-cheeked warbler abundance estimation across the species breeding range (Royle, 2004; Kéry *et al.*, 2005; Royle *et al.*, 2005; Thogmartin *et al.*, 2006; Conroy *et al.*, 2008).

As reviewed by Mills (2007:265), one limitation of patch-occupancy models is they ignore local population dynamics largely because they are data-intensive. Although we acknowledge this limitation, a core issue of conservation biology is the distribution of individuals through space, with a focus on the size of potential habitat patches and their proximity to one another. Whether or not a species is structured as a metapopulation, knowledge of how habitat is occupied through space allows predictions to be made on colonization and extinction probabilities. Thus, our broad-scale approach serves as a template for addressing modelling of colonization and extinction based on the occupancy of habitat, hence a more thorough understanding of population viability. For example, we are using our resulting model to develop more regionally focused studies of patch abundance, productivity and dispersal. If properly designed, regionally focused studies can then be expanded more broadly to the population of interest.

In summary, our study provides an approach for developing a broad-scale assessment of the potential distribution of a species. By using detection/non-detection surveys and remotely sensed habitat metrics within a spatially explicit modelling context, we identified a spatial gradient of occurrence for golden-cheeked warblers as well as relationships between two measurable landscape characteristics that can be used for further conservation planning. As opposed to methods typically used to predict species distributions where habitats are predicted to be either usable or unusable for a species (e.g. Elith *et al.*, 2006), our approach allows for a probabilistic prediction of the likelihood that a patch would harbour the target species. Our model was accurate when evaluated using an independent dataset, suggesting that our predictions were both robust and applicable rangewide.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank J. Andrew Royle, Marc Kéry, Frank Thompson and two anonymous reviewers for constructive comments on earlier versions of this manuscript. We appreciate discussions with J. Andrew Royle, Ciprian Crainiceanu, Marc Kéry and Marc Genton during model development. Support for our research was provided by Texas Department of Transportation. Additionally, we appreciate support from Texas Parks and Wildlife Department. B.A. Collier and R.J. Carroll acknowledge partial support for their work from Award No. KUS-C1-016-04, made by the King Abdullah University of Science and Technology (KAUST). We gratefully acknowledge the private landowners who allowed us access to their lands during the course of this project as well as managers of public lands surveyed during this study.

BIOSKETCH

Bret A. Collier is a research ecologist with the Institute of Renewable Natural Resources at Texas A&M University where his research focuses on quantitative ecology and wildlife population dynamics. Work on this project was conducted by the Research and Management System for Endangered Species (RAMSES) research group at the Institute of Renewable Natural Resources at Texas A&M University (irnr.tamu.edu/RAMSES/).

REFERENCES

Anders AD, Dearborn DC. Population trends of the endangered golden-cheeked warbler at Fort Hood, Texas, from 1992-2001. Southwestern Naturalist. 2004; 49:39–47.

Augustin NH, Mugglestone MA, Buckland ST. An autologistic model for the spatial distribution of wildlife. Journal of Applied Ecology. 1996; 33:339–347.

Bayne EM, Hobson KA. Apparent survival of male Ovenbirds in fragmented and forested boreal landscapes. Ecology. 2002; 83:1307–1316.

Betts MG, Forbes GJ, Diamond AW. Thresholds in songbird occurrence in relation to landscape structure. Conservation Biology. 2007; 21:1046–1058. [PubMed: 17650254]

Bolger DT. Habitat fragmentation effects on birds in southern California: contrast to the "top-down" paradigm. Studies in Avian Biology. 2002; 25:141–157.

Brotons L, Herrando S, Pla M. Updating bird species distribution at large spatial scales: applications of habitat modelling to data from long-term monitoring programs. Diversity and Distributions. 2007; 13:276–288.

Butcher JA, Morrison ML, Ransom D Jr, Slack RD, Wilkins RN. Evidence of a minimum patch threshold of reproductive success in an endangered songbird. Journal of Wildlife Management. 2010; 74:133–139.

Collier BA, Morrison ML, Farrell SL, Campomizzi AJ, Butcher JA, Hays KB, MacKenzie DI, Wilkins RN. Monitoring endangered species occupying private lands: case study using the golden-cheeked warbler. Journal of Wildlife Management. 2010; 74:1–12.

Conroy MJ, Runge JP, Barker RJ, Schofield MR, Fonnesbeck CJ. Efficient estimation of abundance for patchily distributed populations via two-phase, adaptive sampling. Ecology. 2008; 89:3362–3370. [PubMed: 19137943]

Crainiceanu CM, Ruppert D, Wand MP. Bayesian analysis for penalized spline regression using WinBUGS. Journal of Statistical Software. 2005; 14:1–24. Available at http://www.jstatsoft.org/v14/i14.

DeBoer TS, Diamond DD. Predicting presence-absence of the endangered golden-cheeked warbler (Dendroica chrysoparia). Southwestern Naturalist. 2006; 51:181–190.

Diamond, DD. Range-wide modeling of golden-cheeked warbler habitat. University of Missouri; Columbia, Missouri: 2007.

Dormann CF. Effects of incorporating spatial auto-correlation into the analysis of species distribution data. Global Ecology and Biogeography. 2007; 16:129–138.

Elith J, Graham CH, Anderson RP, et al. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 2006; 29:12–151.

Fahrig, L. Estimating minimum habitat for population persistence. In: Lindenmayer, BD.; Hobbs, RJ., editors. Managing and designing landscapes for conservation: moving from perspectives to principles. Blackwell Publishing; Oxford: 2007. p. 64-80.

Fields Development Team. Tools for spatial data. National Center for Atmospheric Research; Boulder, Colorado, USA: 2006. http://www.cdg.ucar.edu/Software/Fields

Forman, RTT. Land mosaics: the ecology of landscapes and regions. Cambridge University Press; New York: 1995.

Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Statistical Science. 1992; 7:457–511.

Gimenez O, Crainiceanu C, Barbraud C, Jenouvrier S, Morgan BTJ. Semiparametric regression in capture-recapture modeling. Biometrics. 2006; 62:691–698. [PubMed: 16984309]

Gimenez O, Grégoire A, Lenormand T. Estimating and visualizing fitness surfaces using mark-recapture data. Evolution. 2009; 63:3097–3105. [PubMed: 19656185]

Groce, JE.; Mathewson, HA.; Morrison, ML.; Wilkins, RN. Scientific evaluation for the 5-year state review of the Golden-cheeked warbler. College Station; Texas A&M Institute of Renewable Natural Resources: 2010.

Grosbois V, Harris MP, Anker-Nilsenn T, McCleery RH, Shaw DN, Morgan BJT, Gimenez O. Modeling survival at multi-population scales using mark-recapture data. Ecology. 2009; 90:2922–2932. [PubMed: 19886500]

Guisan A, Thuiller W. Predicting species distributions: offering more than simple habitat models. Ecology Letters. 2005; 8:993–1009.

Guisan A, Zimmermann NE. Predictive habitat distribution models in ecology. Ecological Modeling. 2000; 135:147–186.

Guisan A, Lehmann A, Ferrier S, Austin M, Overton JMC, Aspinall R, Hastie T. Making betterbiogeographical predictions of species' distributions. Journal of Applied Ecology. 2006; 43:386–392.

Hanski I, Gilpin M. Metapopulation dynamics: a brief history and conceptual domain. Biological Journal of the Linnean Society. 1991; 42:3–16.

He F, Gaston KJ. Occupancy-abundance relationships and sampling scales. Ecography. 2000; 23:503–511.

Hilty J, Merenlender AM. Studying biodiversity on private lands. Conservation Biology. 2003; 17:132–137.

Homan RN, Windmiller BS, Reed JM. Critical thresholds associated with habitat loss for two vernal pool-breeding amphibians. Ecological Applications. 2004; 14:1547–1553.

Kammann EE, Wand MP. Geoadditive models. Applied Statistics. 2003; 52:1–18.

Kéry M. Estimating abundance form bird counts: binomial mixture models uncover complex covariate relationships. Auk. 2008; 125:336–345.

Kéry M, Royle JA, Schmid H. Modeling avian abundance from replicated counts using binomial mixture models. Ecological Applications. 2005; 15:1450–1461.

Kéry M, Gardner B, Monnerat C. Predicting species distributions from checklist data using site-occupancy models. Journal of Biogeography. 2010; 37:1851–1862.

King, R.; Morgan, BJT.; Gimenez, O.; Brooks, SP. Bayesian analysis for population ecology. CRC Press; Boca Raton: 2010.

Knapp RA, Matthews KR, Preisler HK, Jellison R. Developing probabilistic models to predict amphibian site occupancy in a patchy landscape. Ecological Applications. 2003; 13:1069–1082.

Ladd, C.; Gass, L. Golden-cheeked warbler (Dendroica chrysoparia). Account 420. The Birds of North America. In: Poole, A.; Gill, F., editors. The Academy of Natural Sciences. The American Ornithologists' Union, Washington, D.C., USA; Philadelphia, Pennsylvania: 1999.

Lande R. Anthropogenic, ecological, and genetic factors in extinction and conservation. Research in Population Ecology. 1998; 40:259–269.

Lloyd P, Martin TE, Roland LR, Langner U, Hart MM. Linking demographic effects of habitat fragmentation across landscapes to continental source-sink dynamics. Ecological Applications. 2005; 15:1504–1514.

Austin, Loomis. Mapping potential golden-cheeked warbler breeding habitat using remotely-sensed forest canopy cover data. Austin: 2008. Report LAI Project Number 051001

MacKenzie DI. Modeling the probability of resource use: the effect of, and dealing with, detecting a species imperfectly. Journal of Wildlife Management. 2006; 70:367–374.

MacKenzie DI, Royle JA. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology. 2005; 42:1105–1114.

MacKenzie DI, Nichols JD, Lachmann GB, Droege S, Royle JA, Langtimm CA. Estimating site occupancy rates when detection probabilities are less than one. Ecology. 2002; 83:2248–2255.

Magness DR, Wilkins RN, Hejl SJ. Quantitative relationships among golden-cheeked warbler occurrence and landscape size, composition, and structure. Wildlife Society Bulletin. 2006; 34:473–479.

Manly, BFJ.; McDonald, LL.; Thomas, DL.; McDonald, TL.; Erikson, WP. Resource selection by animals. Kluwer Academic Publishers; Dordrecht: 2002.

Marzluff, JM. Worldwide urbanization and its effects on birds. In: Marzluff, JM.; Bowman, R.; Donnelly, R., editors. Avian ecology and conservation in an urbanizing world. Kluwer Academic Publishers; Norwell, Massachusetts, USA: 2001. p. 19-48.

Mills, LS. Conservation of wildlife populations: demography, genetics, and management. Blackwell Publishing; Oxford, England: 2007.

Moore JE, Swihart RK. Modeling patch occupancy by forest rodents: incorporating detectability and spatial autocorrelation with hierarchically structured data. Journal of Wildlife Management. 2005; 69:933–949.

Nychka, D.; Saltzman, N. Design of air-quality monitoring networks. Case studies in environmental statistics. In: Nychka, D.; Piegorsch, WW.; Cox, LH., editors. Lecture notes in statistics. Vol. 132. Springer-Verlag; New York: 1998. p. 51-76.

Peak RG. Forest edges negatively affect Golden-cheeked warbler nest survival. Condor. 2007; 109:628–637.

Plummer M, Best N, Cowles K, Vines K. CODA: convergence diagnosis and output analysis for MCMC. R News. 2006; 6:7–11.

Pulich, WM. The golden-cheeked warbler: a bioecological study. Texas Parks and Wildlife; Austin: 1976.

R Core Development Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2009. ISBN 3-900051-07-0, Available at http://www.R-project.org

Reidy JL, Stake MM, Thompson FR III. Golden-cheeked warbler nest mortality and predators in urban and rural landscapes. Condor. 2008; 110:458–466.

Robinson SK, Thompson FR III, Donovan TM, Whitehead DR, Faaborg J. Forest fragmentation and the regional population dynamics of songbirds. Science. 1995; 267:1987–1990. [PubMed: 17770113]

Royle JA. N-mixture models for estimating population size from spatially replicated counts. Biometrics. 2004; 60:108–115. [PubMed: 15032780]

Royle, JA.; Dorazio, RM. Hierarchical modeling and inference in ecology. Academic Press; London: 2008.

Royle JA, Kéry M. A Bayesian state-space formulation of dynamic occupancy models. Ecology. 2007; 88:1813–1823. [PubMed: 17645027]

Royle JA, Nichols JD, Kéry M. Modelling occurrence and abundance of species when detection is imperfect. Oikos. 2005; 110:353–359.

Royle JA, Kéry M, Gautier R, Schmid H. Hierarchical spatial models of abundance and occurrencefrom imperfect survey data. Ecological Monographs. 2007; 77:465–481.

Ruiz-Gutiérrez V, Gavin TA, Dhondt AA. Habitat fragmentation lowers survival of a tropical forest bird. Ecological Applications. 2008; 18:838–846. [PubMed: 18536246]

Ruppert, D.; Wand, MP.; Carroll, RJ. Cambridge University Press; New York: 2003. Semiparametric regression.

Saab V. Importance of spatial scale to habitat use by breeding birds in riparian forests: a hierarchical analysis. Ecological Applications. 1999; 9:135–151.

Sagarin RD, Gaines SD, Gaylord B. Moving beyond assumptions to understand abundance distributions across the range of species. Trends in Ecology and Evolution. 2006; 21:524–530. [PubMed: 16815588]

SCWA Environmental Consultants. Preliminary deliverable golden-cheeked warbler status review. San Antonio: 2007.

Shanahan DF, Possingham HP. Predicting avian patch occupancy in a fragmented landscape: do we know more than we think? Journal of Applied Ecology. 2006; 46:1026–1035.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Sing T, Snader O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics. 2005; 21:3940–3941. [PubMed: 16096348]

Smith BJ. R Package for MCMC Output Convergence Assessment and Posterior Inference. Journal of Statistical Software. 2007; 21:1–37.

Spiegelhalter, DJ.; Thomas, A.; Best, NG.; Lunn, D. WinBUGS user manual (version 1.4). MRC Biostatistics Unit; Cambridge: 2003. Available at: http://www.mrc-bsu.cam.ac.uk/bugs/

Stevens DL Jr, Jensen SF. Sample design, execution, and analysis for wetland assessment. Wetlands. 2007; 27:515–523.

Sturtz S, Ligges U, Gelman A. R2WinBUGS: a package for running WinBUGS from R. Journal of Statistical Software. 2005; 12:1–16.

Syphard AD, Franklin J. Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. Ecography. 2009; 32:907–918.

Thogmartin WE, Knutson MG, Sauer JR. Predicting regional abundance of rare grassland birds with a hierarchical spatial count model. Condor. 2006; 108:25–46.

Thompson, SK. Sampling. 2nd edn. John Wiley & Sons, Inc.; New York: 2002.

Tognelli MF, Kelt DA. Analysis of determinants of mammalian species richness in South America using spatial autoregressive models. Ecography. 2004; 27:427–436.

United States Fish and Wildlife Service. Golden-cheeked warbler (Dendroica chrysoparia) recovery plan. U.S. Fish and Wildlife Service; Albuquerque: 1992.

White GC, Burnham KP. Program MARK: survival estimation from populations of marked animals. Bird Study. 1999; 46:120–138.

Wood SN. Thin plate regression splines. Journal of the Royal Statistical Society, Series B. 2003; 65:95–114.

Wood SA, Jiang W, Tanner M. Bayesian mixture of splines for spatially adaptive nonparametric regression. Biometrika. 2002; 89:513–528.

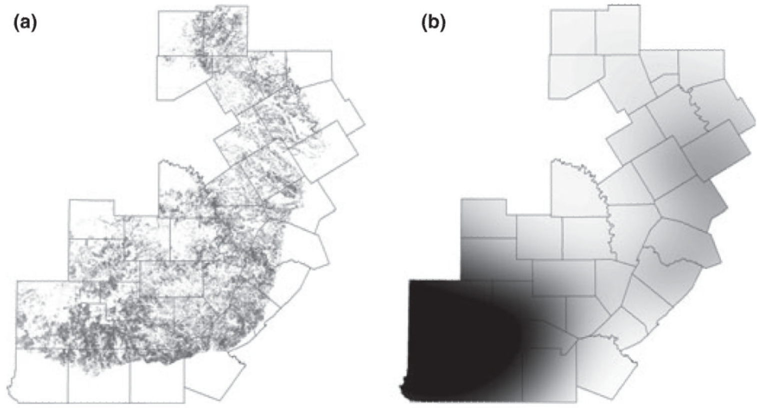

Figure 1.
Distribution (a) of woodland patches (n = 63,616) and posterior predicted spatial process (b) centred at the regression means for the geoadditive semiparametric model containing patch size, landscape composition, X and Y location, and the interaction between patch size and landscape composition within the 35-county breeding range of the golden-cheeked warbler in Texas, USA.

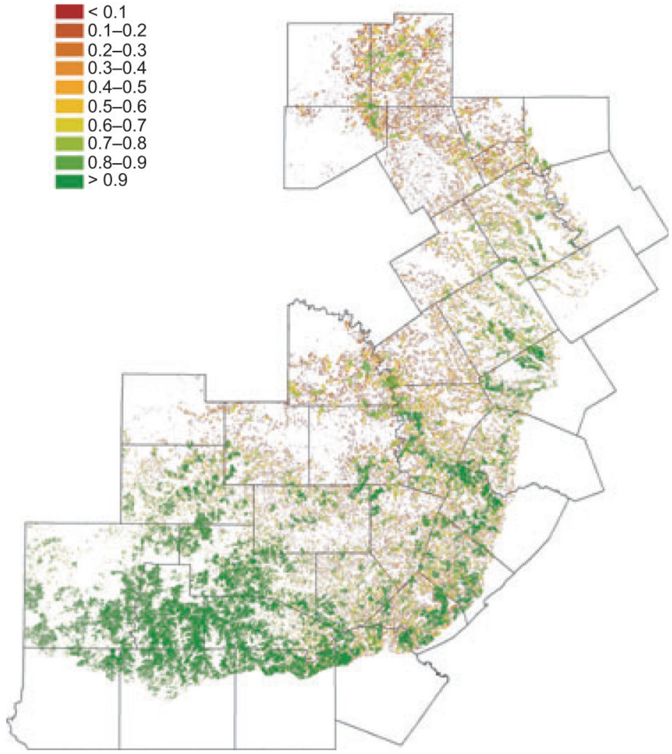

Figure 2.
Estimated patch-specific occupancy probability for the golden-cheeked warbler in the 35-county breeding range in Texas, USA.

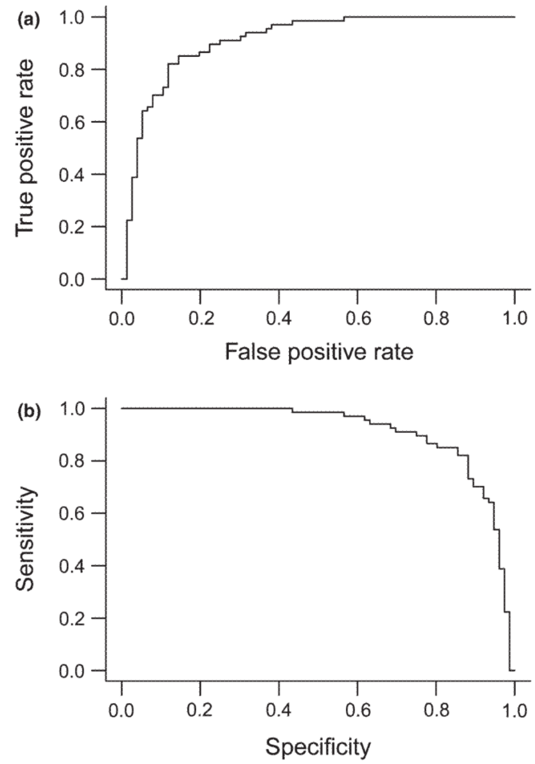

Figure 3.
Visualization of classification accuracy (area under curve =0.91) for patch-scale surveys (n = 143) relative to predictions from the semiparametric model showing (a) the ROC curve and (b) model sensitivity-specificity trade-off.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Posterior parameter estimates for the geoadditive semiparametric occupancy model applied to the golden-cheeked warbler survey data collected in Texas during 2009. The β_i terms represent the model intercept (β_1), latitudinal spatial location (β_2), longitudinal location (β_3), patch size (β_4), landscape composition (β_5) and patch size-landscape composition interaction (β_6). The li terms represent the random effect terms for the 20 knot locations relative to each patch, and α_0 and α_1 represent the intercept and slope for the linear-logistic detection model.

Parameter	Mean	SD
β_1	−11.56	3.530
β_2	0.044	3.440
β_3	−1.107	3.085
β_4	−0.338	5.961
β_5	0.676	6.004
β_6	1.152	5.957
μ_1	0.077	0.517
μ_2	−0.268	0.799
μ_3	0.086	0.965
μ_4	1.469	1.205
μ_5	0.007	1.151
μ_6	−0.891	1.154
μ_7	−0.321	1.346
μ_8	−0.725	1.352
μ_8	1.430	1.439
μ_{10}	−0.253	0.965
μ_{11}	0.099	1.231
μ_{12}	0.877	1.390
μ_{13}	−0.341	1.258
μ_{14}	1.546	1.344
μ_{15}	0.116	1.219
μ_{16}	−1.683	1.446
μ_{17}	0.026	0.494
μ_{18}	0.919	1.232
μ_{19}	0.225	1.219
μ_{20}	−0.816	1.236
α_0	1.101	0.141
α_1	−0.145	0.137

Table 2

Number of classified habitat patches, patch area (ha), and percentage of total potential patches and total area by predicted occupancy category based on a geoadditive semiparametric occupancy model for golden-cheeked warblers in Texas during 2009.

Predicted occupancy	No. of patches	Total ha	Percentage of total	
			Patches	Area
<0.10	37,717	200,654	59.3	12.0
0.10–0.20	7341	103,439	11.5	6.2
0.20–0.30	4129	85,883	6.5	5.1
0.30–0.40	3022	77,761	4.8	4.6
0.40–0.50	2600	86,655	4.1	5.2
0.50–0.60	2107	82,346	3.3	4.9
0.60–0.70	1841	101,792	2.9	6.1
0.70–0.80	1710	134,357	2.7	8.0
0.80–0.90	1447	175,870	2.3	10.5
>0.90	1702	629,941	2.7	37.5
Total	63,616	1,678,698		