

Impacts of forest fragmentation on species richness: a hierarchical approach to community modelling

Elise F. Zipkin^{*1}, Amielle DeWan² and J. Andrew Royle¹

¹USGS Patuxent Wildlife Research Center, 12100 Beech Rd, Laurel, MD 20708, USA; and ²Hudson River Estuary Program/Cornell University, 1 South Putt Corners Rd, New Paltz, NY 12561, USA

Summary

1. Species richness is often used as a tool for prioritizing conservation action. One method for predicting richness and other summaries of community structure is to develop species-specific models of occurrence probability based on habitat or landscape characteristics. However, this approach can be challenging for rare or elusive species for which survey data are often sparse.

2. Recent developments have allowed for improved inference about community structure based on species-specific models of occurrence probability, integrated within a hierarchical modelling framework. This framework offers advantages to inference about species richness over typical approaches by accounting for both species-level effects and the aggregated effects of landscape composition on a community as a whole, thus leading to increased precision in estimates of species richness by improving occupancy estimates for all species, including those that were observed infrequently.

3. We developed a hierarchical model to assess the community response of breeding birds in the Hudson River Valley, New York, to habitat fragmentation and analysed the model using a Bayesian approach.

4. The model was designed to estimate species-specific occurrence and the effects of fragment area and edge (as measured through the perimeter and the perimeter/area ratio, P/A), while accounting for imperfect detection of species.

5. We used the fitted model to make predictions of species richness within forest fragments of variable morphology. The model revealed that species richness of the observed bird community was maximized in small forest fragments with a high P/A . However, the number of forest interior species, a subset of the community with high conservation value, was maximized in large fragments with low P/A .

6. *Synthesis and applications.* Our results demonstrate the importance of understanding the responses of both individual, and groups of species, to environmental heterogeneity while illustrating the utility of hierarchical models for inference about species richness for conservation. This framework can be used to investigate the impacts of land-use change and fragmentation on species or assemblage richness, and to further understand trade-offs in species-specific occupancy probabilities associated with landscape variability.

Key-words: biodiversity, fragmentation, hierarchical modelling, imperfect detection, occupancy modelling, species richness

Introduction

With increasing development pressure on natural landscapes, managers and conservation scientists must determine the most effective ways to preserve the integrity of ecosystems and maintain biodiversity. Inherent in the goals of any conservation effort is the desire to protect as many ecosystem components and processes as possible (Williams *et al.* 1996; Margules &

Pressey 2000). Many conservation and monitoring programmes focus on species or population-level approaches because this can be a concrete and clear way to aid in the preservation of biodiversity (e.g. Myers *et al.* 2000; Pearce *et al.* 2008; Wilhere, Goering & Wang 2008). Species richness (the total number of species in a region) is another more comprehensive, and frequently used, state variable on which to base conservation and management decisions (Yoccoz, Nichols & Boulinier 2001). However, it can be difficult to gauge richness in variable environments (O'Dea, Whittaker & Ugland 2006;

*Correspondence author. E-mail: ezipkin@usgs.gov

Vellend, Lilley & Starzomski 2008). To prioritize conservation action it is critical to have reliable estimates of species richness and to understand how environmental factors affect species-specific patterns of occurrence across a landscape (Boulinier *et al.* 2001; Cabeza *et al.* 2004; Fischer, Lindenmayer & Cowing 2004; Soares & Brito 2007; Lepczyk *et al.* 2008).

There are two challenges in using community-level summaries such as species richness in conservation and management applications. First, species identity is not preserved in many standard analyses used for inference about richness, which are based on simple aggregate species numbers (species–accumulation curves, Gotelli & Colwell 2001; Cam *et al.* 2002) or encounter frequencies (capture–recapture methods, Boulinier *et al.* 1998). However, species-specific patterns of occurrence should be accounted for in modelling approaches (Fischer *et al.* 2004) because the response of species richness to features that can be manipulated (landscape, habitat) is necessarily species specific. A second issue is that in most practical situations species are detected imperfectly. The importance of addressing the biasing effects of imperfect detection on community assessments is widely acknowledged (Boulinier *et al.* 1998; Nichols *et al.* 1998; O’Dea *et al.* 2006; Kéry, Royle & Schmid 2008). Moreover, because detectability naturally varies by species (Boulinier *et al.* 1998), we expect that observed summaries of community structure (e.g. based on species lists) are biased towards abundant and widespread species, which are likely to show diminished response to ecological gradients.

One method for examining species richness in heterogeneous landscapes is to model species occurrence probabilities, or occupancy, based on localized habitat characteristics (Dorazio & Royle *et al.* 2005). Occupancy can be an effective assessment method (Manley *et al.* 2005), generally requires less effort and expense than estimating total abundance of all species (MacKenzie *et al.* 2006), and allows for imperfect detection of species (MacKenzie *et al.* 2002). Multi-species occupancy models have been used for inference in community studies in a number of situations, including estimation of richness and community overlap (Dorazio & Royle 2005), construction of individual-based species accumulation curves (Dorazio *et al.* 2006), and in determining the influence of habitat and landscape variation on richness (Kéry & Royle 2008, 2009; Russell *et al.* 2009).

In addition to understanding total species richness, inferences on the number of rare, endangered or functionally important species are frequently of interest in conservation planning and monitoring programmes (Samu, Csontos & Szinetar 2008). Occupancy estimates for rare species and guild or assemblage richness (number of species in a subset of the population) can be more informative about areas of high conservation priority than assessments on only species that are common. Unfortunately, it can be difficult to get reliable estimates of occupancy for rare and/or elusive species because traditional sampling efforts often do not generate enough data for standard analyses (Queheillalt *et al.* 2002; Stockwell & Peterson 2002). Some approaches to mitigating this problem combine data on rare, but functionally similar, species (e.g. by genus) or use indicator species to deduce occupancy of those species with limited data (O’Connell, Jackson & Brooks 2000;

Fleishman, Blair & Murphy 2001; Sergio *et al.* 2006). Such approaches discard valuable information about species-specific responses, and could be misleading or erroneous if rarely observed species respond differently than indicator species (Andelman & Fagan 2000; Kéry, Royle & Schmid 2008; Lawler & White 2008). The question remains regarding the most efficient and cost-effective method for estimating the occurrence and distribution of uncommon and elusive species (Thompson 2004; MacKenzie *et al.* 2005).

Our research is motivated by a desire to develop a community-level quantitative framework for predicting areas of conservation value, and to provide high-quality baseline data for vertebrate monitoring programmes in urbanizing landscapes. To this end, we present a recently developed approach for assessing community composition based on species-specific occupancy and detection (Dorazio & Royle 2005) in which individual species occurrence models are linked together within a hierarchical (or multi-level) model (Gelman & Hill 2007; Royle & Dorazio 2008). Many multi-species field studies and monitoring programmes have limited data on a large portion of observed species; as such, typical species-by-species analyses are simply unable to provide occurrence estimates or information about the effects of environmental factors on occurrence probabilities. An advantage of the hierarchical modelling framework over typical species richness analyses is that it accounts for both species-level effects as well as aggregated effects of landscape/habitat on the community as a whole (Kéry & Royle 2008, 2009), leading to a more efficient use of available data and increased precision in occupancy estimates, especially for infrequently observed species. We demonstrate the strengths of this approach by applying the hierarchical modelling framework to a bird community in forest fragments across the Hudson River Valley (HRV), New York (DeWan *et al.* 2009), a biologically diverse and ecologically significant region that is under intense development pressure, in the north-eastern United States (Finton *et al.* 2000; Smith *et al.* 2001). Efforts are underway to prioritize the landscape for conservation actions, yet little is known about many of the species in the region (DeWan *et al.* 2009). We focused our analyses on the community response to habitat fragmentation by modelling species-level changes in occupancy to two factors with well-established effects on the success of breeding birds: forest fragment area and edge-effects as measured by responses to perimeter, and perimeter/area ratio (P/A) (Rafe, Usher & Jefferson 1985; Helzer & Jelinski 1999).

Materials and Methods

BACKGROUND

We used a hierarchical model that links species-specific detection and occupancy, which are then related (across species, at the community level) through an additional component of the hierarchical model (Dorazio & Royle 2005; Dorazio *et al.* 2006). A hierarchical (sometimes referred to as multi-level or state-space) model is one in which various biological and sampling components are formally specified and related to one another (Gelman & Hill 2007; Royle & Dorazio 2008). For example, in the context of estimating occupancy,

hierarchical models can help distinguish absence from non-detection by explicitly incorporating models that specify presence vs. absence as one process and detection vs. non-detection as another process that is dependent upon whether or not the species is in fact present. Hierarchical models posit weak, stochastic relations rather than deterministic relations among parameters and processes (Link 1999; Link *et al.* 2002), resulting in improved estimation of individual parameters by considering them in the context of a group of related variables (Bayesian shrinkage: 'borrowing strength from the ensemble' Link & Sauer 1996). In the context of our community model, this allows for increased precision of occurrence estimates for rare or elusive species through utilization of collective community data (Russell *et al.* 2009) and improved 'composite' analyses of species groups (Sauer & Link 2002). With limited resources and budgets, many multi-species data collection efforts have very small sample sizes – to such an extent that it is not possible to carry out formal inference on a species-by-species basis. The hierarchical modelling approach allows for the most effective use of available data while not requiring *a priori* assumptions on group structure or relatedness among species.

STUDY SITE

The data come from a breeding bird survey collected over a 2-year period (15 May to 1 July 2006 and 15 May to 1 July 2007) at 72 randomly selected independent points in deciduous and mixed-deciduous forest fragments across the HRV, New York. The sampling locations ranged over the entire 9546 km² region which includes all or part of nine counties that border the Hudson River, north of New York City. Points were located at least 500 m apart using Hawth's stratified random sampling tool (Beyer 2004), and then mapped and field-checked, eliminating those that: (i) had recent disturbance that altered the cover classification ($n = 1$); (ii) were too dangerous to access (e.g. steep ravine) ($n = 4$); or (iii) did not receive private landowner permission to access the site ($n = 21$). Forest fragments ranged in size from 0.14 to 8677.4 ha ($\mu = 533.7$ ha), while P/A ranged from 0.08 to 1.5 km ha⁻¹ ($\mu = 0.2$). Two trained observers recorded the presence of all species seen or heard during the 10-min, 250 m fixed-radius point counts at each sampling station (Hutto, Pletschet & Hendricks 1986). Sites were visited on three separate occasions during the breeding season (once each per 2-week period), although not all sites were surveyed both years. The perimeter and area of the fragment in which the point occurred was recorded. A total of 78 species were observed in this study. Of these, the data for 32 species were particularly sparse with less than 20 detections each over the entirety of the sampling season (see Appendix S1 in Supporting Information for a complete species list). Because of the small size of the data set, typical single species approaches for estimating occupancy were inadequate for the majority of observed species. For more details on the sampling design and region, see DeWan *et al.* (2009).

THE MODEL

The repeated sampling protocol allows for non-detection to be discerned from point-level absence at each location (MacKenzie *et al.* 2002). We developed a hierarchical model which assumes that site-specific occupancy (i.e. 'true' presence/absence) for species $i = 1, 2, \dots, N$ at site $j = 1, 2, \dots, J$, denoted $z(i, j)$, where $z(i, j) = 1$ if species i occurs in site j and is zero otherwise. The model for occurrence is specified as $z(i, j) \sim \text{Bern}(\psi_{ij})$ where ψ_{ij} is the probability that species i occurs at site j . The state variable $z(i, j)$ is usually not known with certainty. Instead, we observe data $x(i, j, k)$ for species i at site j during sampling period k , which are also assumed to be

Bernoulli random variables if species i is present (i.e. if $z(i, j) = 1$); otherwise, if $z(i, j) = 0$, then $x(i, j, k) = 0$ with probability 1. The observation model is represented by $x(i, j, k) \sim \text{Bern}(\theta_{i,j,k} \cdot z(i, j))$ where $\theta_{i,j,k}$ is the detection probability of species i for the k th sampling period at site j , if species i is present at site j . Note that the model satisfies the condition that detection is a fixed zero when a species does not occur (because $z(i, j) = 0$).

In the simplest specification of the model, the occurrence and detection probabilities, ψ and θ , are determined by unspecified species and site-level effects (Dorazio *et al.* 2006). These effects are incorporated into the model linearly on the logit-probability scale: $\text{logit}(\psi_{ij}) = u_i + \alpha_j$ and $\text{logit}(\theta_{i,j,k}) = v_i + \beta_j$, where u_i and v_i are species-level effects and α_j and β_j are site-level effects on occurrence and detection respectively. Because high abundance species are likely to be both easier to detect and more prevalent across the landscape, we modelled a correlation (ρ) between occurrence and detection in the model by allowing u_i and v_i to be jointly distributed such that $[u_i, v_i | \Sigma] \sim N(0, \Sigma)$ where (σ_u^2, σ_v^2) are the variance components among species for occurrence and detection, respectively, and σ_{uv} is the covariance of the 2×2 matrix Σ (Dorazio & Royle 2005; Kéry & Royle 2008).

Extensions of this basic model have explicitly incorporated landscape and survey characteristics into the probabilities of occupancy and detection (Kéry & Royle 2009; Russell *et al.* 2009). We followed this approach, and modelled the occurrence probability for species i at j by incorporating site-specific habitat characteristics. In this case we used the size and relative shape of the forest fragment in which the point count occurred. As counts were conducted in a 250 m radius, occupancy and detection estimates for individual species are provided at the point (not fragment) level. Thus we are considering how occupancy at a random point is affected by the area and shape of the forest fragment in which it occurs. We incorporated fragment area, perimeter and P/A in the occupancy estimates by assuming that the logit transform of the occurrence probability was a linear combination of a species effect and the site-specific habitat characteristics as follows:

$$\text{logit}(\psi_{ij}) = u_i + \alpha_1 \text{perimeter}_j + \alpha_2 \text{area}_j + \alpha_3 \text{P/A}_j.$$

We standardized the covariates so that the means of the perimeter, area and P/A data were zero. Thus, the inverse-logit of u_i is the occurrence probability for species i in sites with 'average' habitat characteristics. The coefficients α_1 , α_2 , and α_3 are the effects of perimeter, area and P/A, for species i respectively. The detection probability for species i was assumed to vary based on the date of the survey (linear and squared effects) and the year of the survey. We assumed that the community was closed (i.e. the species pool remained constant) over the two years during which the survey was conducted, but added a year effect (constant across species) to account for shifting detection between the two years as a result of annual fluctuations in seasonality:

$$\text{logit}(\theta_{i,j,k}) = v_i + \beta_1 \text{date}_{j,k} + \beta_2 \text{date}_{j,k}^2 + \beta_3 \text{year}_{j,k}.$$

Our model contains seven parameters for each species in the community, and one (year effect) that is estimated across species. As observations were sparse for many species in the sample (Appendix S1), estimating all of these parameters would not be possible if the data were analysed on a species-by-species basis. As such, we added an additional hierarchical component of the model by assuming that the species-level parameters were random effects, each governed by community-level 'hyper-parameters'. For example, we assumed that $\alpha_1 \sim N(\mu_{\alpha_1}, \sigma_{\alpha_1})$ where μ_{α_1} is the community response (mean across

species) to perimeter and σ_{z1} is the standard deviation (among species), thus the hyper-parameters are simply the mean and variance for each covariate as measured across species (Kéry & Royle 2009).

We estimated model parameters and community summaries using a Bayesian analysis of the model with vague priors for the hyper-parameters (e.g. uniform distribution from 0 to 1 for community-level occupancy and detection covariates; normal distributions with mean zero and variance 1000 for community-level habitat and sampling covariates). Hierarchical models are naturally analysed using Bayesian methods (Gelman & Hill 2007). We carried out our analysis with WinBUGS (Spiegelhalter *et al.* 2003), general purpose software for Bayesian analysis that uses Markov chain Monte Carlo (MCMC). The advantage of WinBUGS is that it only requires specification of the model, and not a technical development of the MCMC algorithm. The complete model specification and additional details, including an assessment of model fit using a Bayesian *P*-value approach, are presented in Appendix S2.

Results

SPECIES RICHNESS AND COMMUNITY-LEVEL SUMMARIES

The mean estimates for the community response to fragment perimeter and area were negative, while the response to P/A was positive (Table 1). This suggests that, in general, the mean probability of occupancy across species in this community was higher at points in smaller, more irregularly shaped fragments than in larger fragments with less edge. The posterior intervals for each of the community hyper-parameters contain both positive and negative values (Table 1), which is a manifestation of the variability in the community (Appendix S1). In our study, which encompasses a diverse bird community, we would naturally expect the response of individual species to vary with landscape fragmentation. Thus, diffuse posterior distributions for the community-level habitat covariates are as expected and simply reflect the diversity within the community.

Table 1. Community-level summaries of the hyper-parameters for the detection and occupancy covariates

Community-level hyper-parameter		Mean	Standard deviation	95% Posterior intervals
μ_{z1}	Perimeter	-0.06	0.75	-1.65, 1.32
σ_{z1}	Perimeter	0.64	0.34	0.20, 1.51
μ_{z2}	Area	-0.25	0.83	-1.79, 1.48
σ_{z2}	Area	1.83	0.47	0.82, 2.75
μ_{z3}	P/A	0.07	0.12	-0.15, 0.32
σ_{z3}	P/A	0.54	0.18	0.26, 0.96
$\mu_{\beta1}$	Date effect (linear term)	-0.04	0.04	-0.12, 0.05
$\sigma_{\beta1}$	Date effect (linear term)	0.25	0.04	0.17, 0.33
$\mu_{\beta2}$	Date effect (squared term)	-0.03	0.05	-0.13, 0.08
$\sigma_{\beta2}$	Date effect (squared term)	0.25	0.05	0.16, 0.36
$\beta3$	Year effect	0.11	0.05	0.01, 0.21

P/A, perimeter/area ratio.

We used the model to make predictions of species richness at localized points across a landscape with heterogeneous forest fragments that varied by area and P/A (Fig. 1; for details on how richness was calculated and estimates on the precision of the predictions, see Appendix S3). Species richness was maximized in small areas with high perimeter to area ratios (large amounts of edge habitat) (Fig. 1 left panel). However, assemblage richness of forest interior breeding birds (17 species), a subset of the population with high conservation value, was maximized in large fragments with less edge (Fig. 1 right panel).

SPECIES-SPECIFIC RESPONSES

Mean probabilities of occurrence varied widely among species, ranging from 6.5% to 98.5%. Detection was low for many species and also varied widely (7.1–75.9%). There was a strong correlation between occupancy and detection (posterior mean for ρ was 0.73, 95% posterior interval: 0.52–0.88; Fig. 2), a phenomenon that is likely due to heterogeneity in abundance among species (Dorazio & Royle 2005). Posterior summaries of occupancy and detection for each species, as well as species-specific responses to the habitat covariates, are presented in Appendix S1.

Fragment area, compared with perimeter or P/A, had a large impact on mean estimates of occupancy for many species within the community (Appendix S1). Over the range of surveyed fragments, 24 species showed (on average) an increase in occurrence probability as area increased (greater than 10% change in mean estimates of occupancy from minimum to maximum fragment size in the survey), 31 species showed a decrease in occurrence probability (> 10%) with increasing area and 23 species showed no change in occurrence probability with area (< 10% change).

Many species whose mean occurrence probabilities increased in response to increased area were forest-dependent species of high conservation concern. On average, nine forest interior breeders (Acadian flycatcher *Empidonax virens*, black-and-white warbler *Mniotilta varia*, blackburnian warbler *Dendroica fusca*, black-throated blue warbler *Dendroica caerulescens*, black-throated green warbler *Dendroica virens*, cerulean warbler *Dendroica cerulea*, hooded warbler *Wilsonia citrina*, worm-eating warbler *Helmitheros vermivorum* and winter wren *Troglodytes troglodytes*) showed substantial increases in occupancy probabilities as fragment area increased, but less response to changes in perimeter or P/A ratio (Fig. 3; Appendix S1). Although the number of observations for these species was fairly low (6–36 for each), the community approach allowed us to obtain estimates of the response of each species to fragment area and regularity of shape. The precision on species-level estimates of occupancy and effects of fragmentation increased for most species in the community model compared with standard species-specific models (for selected results comparing the community model to a single species modelling approach, see Appendix S4). When modelling each species separately, occupancy estimates for species with sparse data could not

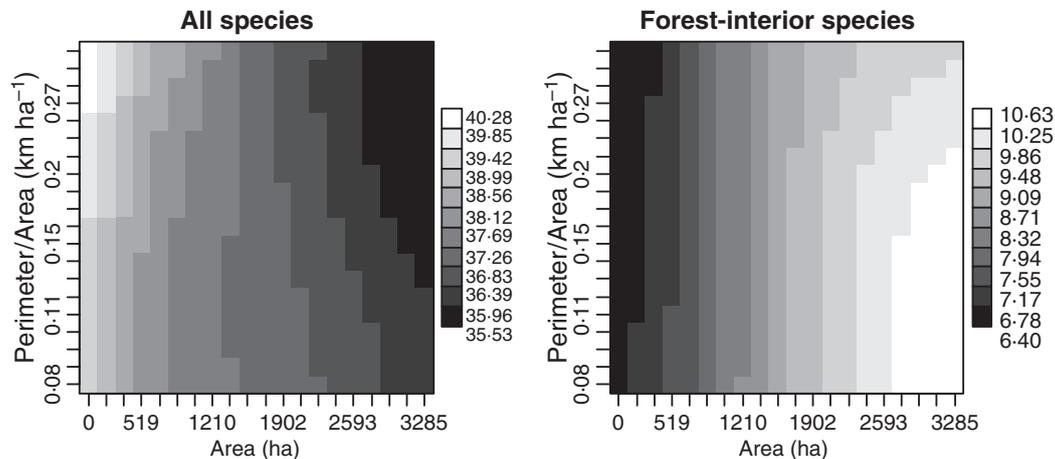


Fig. 1. Total estimated species richness (left) based on area and perimeter/area ratio (P/A) and estimated assemblage richness of forest interior species (right – 17 observed species: Acadian flycatcher, black-and-white warbler, blackburnian warbler, black-throated blue warbler, black-throated green warbler, brown creeper, Canada warbler, cerulean warbler, hooded warbler, northern parula, ovenbird, red-breasted nuthatch, scarlet tanager, veery, worm-eating warbler, winter wren and wood thrush).

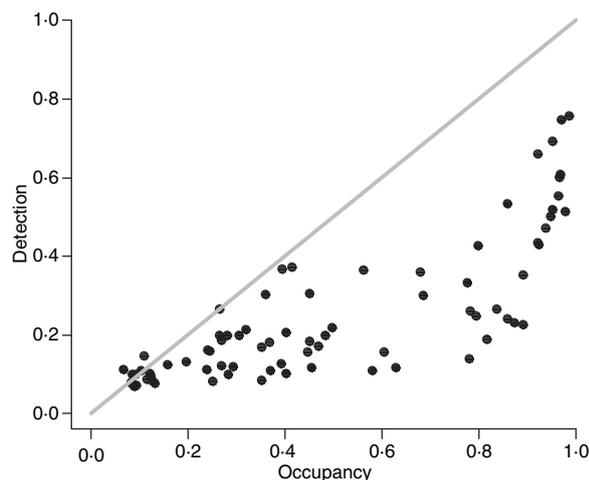


Fig. 2. Mean estimated values of occupancy and detection for the 78 observed species. The shaded line shows the one to one relationship.

be obtained without exhibiting extreme sensitivity to the prior. For the above nine forest interior species, the standard deviations on the estimated species-specific effects of area were generally lower using the hierarchical community model (range 1.24–1.83) than a standard species-level model (range 1.40–2.03; Appendix S4). Three species (ovenbird *Seiurus aurocapilla*, scarlet tanager *Piranga olivacea* and veery *Catharus fuscescens*) also had a positive response to area, but the effects were less discernible on estimates of occupancy because they were widely observed (e.g. occupancy was universally high). A few forest-dependent species (brown creeper *Certhia americana*, Canada warbler *Wilsonia canadensis*, northern parula *Parula americana*, red-breasted nuthatch *Sitta canadensis* and wood thrush *Hylocichla mustelina*) responded more closely to the community-level response with decreasing occupancy probabilities as fragment area increased (Fig. 4).

Discussion

Although reliable summaries of species occurrences and distributions are required for effective conservation, analysis of multi-species data can be challenging because sampling techniques often identify numerous species with few detections. One way to address this issue is to utilize models that integrate data across species, allowing for composite analyses of communities or groups of species. Hierarchical models are particularly valuable in this context, in part because they do not require *a priori* assumptions about community structure; any composite analysis will improve estimates on metrics of interest, regardless of relationships among species (Sauer & Link 2002). For conservation purposes, it is generally useful to consider species from one community or related communities; otherwise community-level summaries may not be meaningful. In some situations it may be possible to incorporate additional group structure into the model when relationships among species have been well established. Estimates for rarely observed species will naturally be drawn to group averages ('Bayesian shrinkage' toward the mean; Link 1999), but the precision of estimates can be improved with even a minimal number of observations (Appendix S4). Accuracy of species-specific estimates will always be limited by the amount of available data, which is reflected in the diffuse posterior distributions for many habitat covariates (Appendix S1). Such estimates can only be objectively improved through additional data collection efforts. However, as with meta-analysis in classical statistics (Osenberg *et al.* 1999), many 'weak' inferences can be combined to make a stronger collective response. Thus, by accounting for both species-level effects as well as the aggregated effects of landscape covariates on the community as a whole, hierarchical models provide a valuable alternative to single species analyses of community data.

Our model produced a number of key findings relevant to prioritizing conservation actions and was capable of making predictions of bird species richness based on fragment area

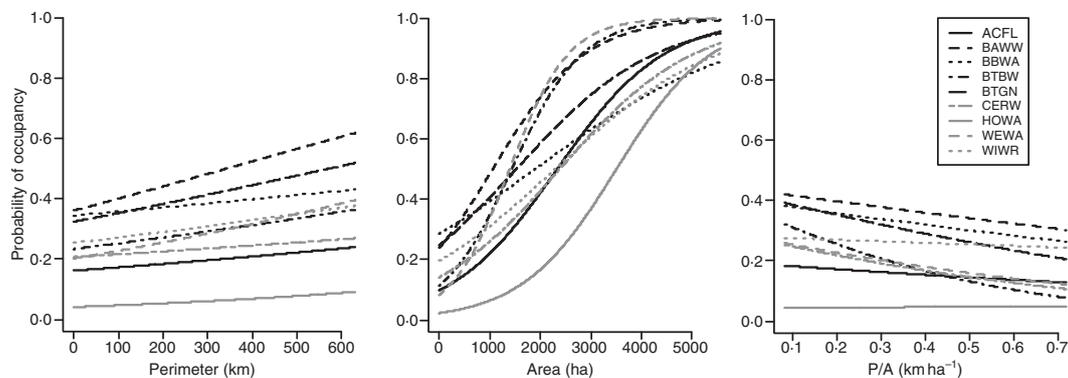


Fig. 3. Mean marginal probabilities of occupancy for nine forest interior breeding bird species (Acadian flycatcher, black-and-white warbler, blackburnian warbler, black-throated blue warbler, black-throated green warbler, cerulean warbler, hooded warbler, worm-eating warbler and winter wren) in relation to fragment perimeter, area and perimeter/area ratio (P/A).

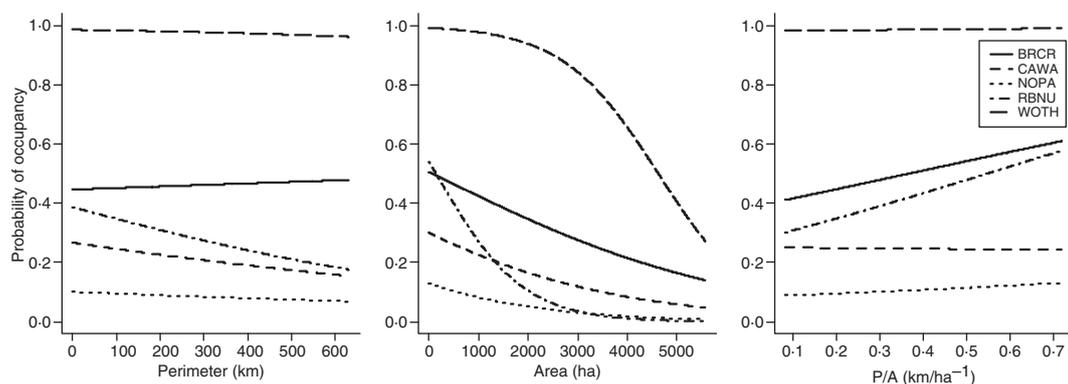


Fig. 4. Mean marginal probabilities of occupancy for five forest interior breeding bird species (brown creeper, Canada warbler, northern parula, red-breasted nuthatch and wood thrush) in relation to fragment perimeter, area and perimeter/area ratio (P/A).

and edge effects (Fig. 1), which should be verified through additional sampling. Understanding the relationship between environmental factors and species richness will improve the efficacy of conservation efforts in the protection of biodiversity in urbanizing landscapes. For example, our estimates of the community and species-level relationships between occupancy probabilities and habitat characteristics allows a direct valuation of forest fragments in terms of either total species richness (Fig. 1 left panel) or assemblage richness (Fig. 1 right panel), and illustrates an explicit trade-off between these two competing objectives. Overall, the community-level response to area and P/A suggests that many species increased in occupancy in response to fragmentation inducing a concomitant increase in species richness. These results are consistent with the intermediate disturbance hypothesis (Grime 1973; Horn 1975; Connell 1978) which suggests that diversity is maximized in areas of moderate disturbance. Similar to Lepczyk *et al.* (2008), we found that extremely large fragments with extensive forest interior may be less common (DeWan *et al.* 2009) and estimates of species richness would be expected to decline if sites were dominated by edge-tolerant or generalist species. In a conservation context, our overall estimates of species richness may not be particularly valuable; however, the hierarchical framework offered a means to acquire improved precision in estimates of occupancy for rarer species, which we used to determine

assemblage richness for a subset of the community with high conservation value.

Many of the forest-breeding species responded to increased fragmentation with decreased probabilities of occupancy (Fig. 3; Appendix S4). However, occupancy for some forest-breeding species responded negatively to fragment area (Fig. 4). Although this may not be surprising for more urban-tolerant species (e.g. red-breasted nuthatch), these results were not typical for others that are sometimes considered sensitive to fragmentation (e.g. Canada warbler, wood thrush). In addition, some area-sensitive species were so common that their relationship to area would not have been discernible through typical occupancy approaches. Scarlet tanager, ovenbird and veery were observed frequently during sampling and had high occupancy estimates. If we had *a priori* grouped these species together as an indicator of sensitivity to fragmentation, without testing the assumptions, we would have been unable to discern differences among species in their response to fragment area and P/A. We therefore suggest that researchers use caution in *a priori* grouping if the purpose is to understand how a species or assemblage may be responding to landscape variability.

Our approach allows for estimation of occupancy and detection probabilities of all observed species, even if they are poorly represented in the sample data. Detection probabilities were

very low for many species (Fig. 2; Appendix S1), further supporting a number of studies that have demonstrated the importance of accounting for detection in occupancy and abundance modelling (Bailey, Simons & Pollock 2004; MacKenzie *et al.* 2006; Kéry, Royle & Schmid 2008). Detection probability can also be significantly affected by abundance (Royle & Nichols 2003), which is evidenced in our analysis by the high correlation between detection and occupancy. Variance around species-specific estimates of occupancy, detection and the covariates will inevitably be high for species with limited data. However, the community-level approach typically provides more precise estimates for rare species than traditional species-level analyses (Appendix S4) and was especially valuable for the nine forest interior species that were sensitive to habitat fragmentation, yet would not have yielded reliable estimates of occupancy due to low sample size. Our analysis framework should be particularly effective in reducing cost and increasing efficiency for organizations where funding for field-based data collection is limited.

Many conservation and management decisions rely on estimates of species richness to prioritize areas for protection and monitoring. For example, DeWan *et al.* (2009) developed a map of high priority conservation areas in the HRV region based on indices of richness for a subset of forest interior bird species. Their analyses were limited to species that were neither too common nor too rare. The results from our community-level approach can be used to improve such maps and more accurately determine areas of high conservation value to protect from urban development. We demonstrated, using a diverse bird community, the applicability and relevance of our hierarchical modelling approach to: (i) assess species richness while accounting for individual species; (ii) improve the precision on estimates of occupancy and detection for many species, even species with relatively sparse data; and (iii) investigate the impacts of fragmentation on breeding birds at the community and species levels. Our hierarchical framework offers an exciting tool for wildlife agencies and conservation organizations who struggle to effectively monitor and protect biological diversity. Monitoring the status and distribution of biodiversity and rare species is a priority at local, national and international scales (Oberbillig 2008). Because of challenges in sampling and cost, lack of quality data has been identified as a serious challenge for biodiversity conservation, particularly for rarer species (The Heinz Center 2002). Many sampling designs already include data collection on multiple species (Heyer 1994; Wilson *et al.* 1996) and multi-species inventory techniques can reduce sampling costs and effort (Manley *et al.* 2005; Vesely *et al.* 2006). The community approach allows researchers to use data from all sampled species to improve estimates of species richness and generate previously unavailable estimates of occupancy for rare or elusive species. The flexibility of hierarchical modelling can provide greater insight into how a particular taxonomic community responds to environmental changes, while also accounting for species-specific differences. If incorporated into monitoring and assessment programmes, this framework could improve estimation of species richness and inferences for rare species,

and provide scientifically sound information to support conservation planning and action.

Acknowledgements

The authors thank Sarah Converse, Beth Gardner, Gonçalo Ferraz, Marc Kéry and Julien Martin for discussions on the topic and providing comments on earlier versions of the manuscript. We also thank three anonymous reviewers and Chris Elphick for many useful suggestions. This research would not be possible without the funding and support of the following: The Hudson River Estuary Program, Biodiversity Research Institute, and New York Cooperative Fish and Wildlife Research Unit at Cornell University.

References

- Andelman, S.J. & Fagan, W.F. (2000) Umbrellas and flagships: efficient conservation surrogates or expensive mistakes? *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 5954–5959.
- Bailey, L.L., Simons, T.R. & Pollock, K.H. (2004) Estimating site occupancy and species detection probability parameters for terrestrial salamanders. *Ecological Applications*, **14**, 692–702.
- Beyer, H.L. (2004) Hawth's Analysis Tools for ArcGIS. <http://www.spatia/ecology.com/htools>, accessed May 1 2008.
- Boulinier, T., Nichols, J.D., Sauer, J.R., Hines, J.E. & Pollock, K.H. (1998) Estimating species richness: the importance of heterogeneity in species detectability. *Ecology*, **79**, 1018–1028.
- Boulinier, T., Nichols, J.D., Hines, J.E., Sauer, J.R., Flather, C.H. & Pollock, K.H. (2001) Forest fragmentation and bird community dynamics: inference at regional scales. *Ecology*, **82**, 1159–1169.
- Cabeza, M., Araujo, M.B., Wilson, R.J., Thomas, C.D., Cowley, M.J.R. & Moilanen, A. (2004) Combining probabilities of occurrence with spatial reserve design. *Journal of Applied Ecology*, **41**, 252–262.
- Cam, E., Nichols, J.D., Sauer, J.R. & Hines, J.E. (2002) On the estimation of species richness based on the accumulation of previously unrecorded species. *Ecography*, **25**, 102–108.
- Connell, J.H. (1978) Diversity in tropical rain forests and coral reefs – high diversity of trees and corals is maintained only in a non-equilibrium state. *Science*, **199**, 1302–1310.
- DeWan, A.A., Sullivan, P.J., Lembo, A.J., Smith, C.R., Maerz, J.C., Lassoie, J.P. & Richmond, M.E. (2009) Using occupancy models of forest breeding birds to prioritize conservation planning. *Biological Conservation*, **142**, 982–991.
- Dorazio, R.M. & Royle, J.A. (2005) Estimating size and composition of biological communities by modeling the occurrence of species. *Journal of American Statistical Association*, **100**, 389–398.
- Dorazio, R.M., Royle, J.A., Söderström, B. & Glimskär, A. (2006) Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology*, **87**, 842–854.
- Finton, A.D., Novak, P.G., Schneider, K.J. & Weldy, T.W. (2000) *Rare Species and Significant Ecological Communities of the Counties Bordering the Hudson River Estuary North of New York City*. New York State Department of Environmental Conservation, New York.
- Fischer, J., Lindenmayer, D.B. & Cowing, A. (2004) The challenge of managing multiple species at multiple scales: reptiles in an Australian grazing landscape. *Journal of Applied Ecology*, **41**, 32–44.
- Fleishman, E., Blair, R.B. & Murphy, D.D. (2001) Empirical validation of a method for umbrella species selection. *Ecological Applications*, **11**, 1489–1501.
- Gelman, A. & Hill, J. (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York.
- Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, **4**, 379–391.
- Grime, J.P. (1973) Competition and diversity in herbaceous vegetation – reply. *Nature*, **244**, 311–312.
- Helzer, C.J. & Jelinski, D.E. (1999) The relative importance of patch area and perimeter-area ratio to grassland breeding birds. *Ecological Applications*, **9**, 1448–1458.
- Heyer, W.R. (1994) *Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians*. Smithsonian, Washington, DC.
- Horn, H.S. (1975) Markovian properties of forest succession. *Ecology and Evolution of Communities* (eds M.L. Cody & J.M. Diamond), pp. 196–211. Belknap Press, Boston.

- Hutto, R.L., Pletschet, S.M. & Hendricks, P. (1986) A fixed-radius point count method for nonbreeding and breeding season use. *Auk*, **103**, 593–602.
- Kéry, M. & Royle, J.A. (2008) Hierarchical Bayes estimation of species richness and occupancy in spatially replicated surveys. *Journal of Applied Ecology*, **45**, 589–598.
- Kéry, M. & Royle, J.A. (2009) Inference about species richness and community structure using species-specific occupancy models in the National Swiss Breeding Bird Survey MHB. *Modeling Demographic Processes in Marked Populations* (eds D.L. Thomson, E.G. Cooch & M.J. Conroy), pp. 639–656. Springer, New York.
- Kéry, M., Royle, J.A. & Schmid, H. (2008) Importance of sampling design and analysis in animal population studies: a comment on Sergio *et al.* *Journal of Applied Ecology*, **45**, 981–986.
- Lawler, J.J. & White, D. (2008) Assessing the mechanisms behind successful surrogates for biodiversity in conservation planning. *Animal Conservation*, **11**, 270–280.
- Lepczyk, C.A., Flather, C.H., Radeloff, V.C., Pidgeon, A.M., Hammer, R.B. & Liu, J.G. (2008) Human impacts on regional avian diversity and abundance. *Conservation Biology*, **22**, 405–416.
- Link, W.A. (1999) Modeling pattern in collections of parameters. *Journal of Wildlife Management*, **63**, 1017–1027.
- Link, W.A. & Sauer, J.R. (1996) Extremes in ecology: avoiding the misleading effects of sampling variation in summary analyses. *Ecology*, **77**, 1633–1640.
- Link, W.A., Cam, E., Nichols, J.D. & Cooch, E.G. (2002) Of BUGS and Birds: Markov chain Monte Carlo for hierarchical modeling in wildlife research. *Journal of Wildlife Management*, **66**, 277–291.
- MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A. & Langtimm, C.A. (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, **83**, 2248–2255.
- MacKenzie, D.I., Nichols, J.D., Sutton, N., Kawanishi, K. & Bailey, L.L. (2005) Improving inferences in population studies of rare species that are detected imperfectly. *Ecology*, **86**, 1101–1113.
- MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L. & Hines, J.E. (2006) *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier, Oxford.
- Manley, P.N., Schlesinger, M.D., Roth, J.K. & Van Horne, B. (2005) A field-based evaluation of a presence-absence protocol for monitoring ecoregional-scale biodiversity. *Journal of Wildlife Management*, **69**, 950–966.
- Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.
- Nichols, J.D., Boulenger, T., Hines, J.E., Pollock, K.H. & Sauer, J.R. (1998) Estimating rates of local species extinction, colonization, and turnover in animal communities. *Ecological Applications*, **8**, 1213–1225.
- O'Connell, T.J., Jackson, L.E. & Brooks, R.P. (2000) Bird guilds as indicators of ecological condition in the central Appalachians. *Ecological Applications*, **10**, 1706–1721.
- O'Dea, N., Whittaker, R.J. & Ugland, K.I. (2006) Using spatial heterogeneity to extrapolate species richness: a new method tested on Ecuadorian cloud forest birds. *Journal of Applied Ecology*, **43**, 189–198.
- Oberbillig, D. (2008) State wildlife action plans: working together to prevent wildlife from becoming endangered [WWW document]. http://www.wildlife-actionplans.org/pdfs/wildlife_action_plans_summary_report.pdf [accessed on October 1, 2008].
- Osenberg, C.W., Sarnelle, O., Cooper, S.D. & Holt, R.D. (1999) Resolving ecological questions through meta-analysis: goals, metrics, and models. *Ecology*, **80**, 1105–1117.
- Pearce, J.L., Kirk, D.A., Lane, C.P., Mahr, M.H., Walmsley, J., Casey, D., Muir, J.E., Hannon, S., Hansen, A. & Jones, K. (2008) Prioritizing avian conservation areas for the yellowstone to Yukon Region of North America. *Biological Conservation*, **141**, 908–924.
- Queheillalt, D.M., Cain, J.W., Taylor, D.E., Morrison, M.L., Hoover, S.L., Tuatoo-Bartley, N., Ruge, L., Christopherson, K., Hulst, M.D., Harris, M.R. & Keough, H.L. (2002) The exclusion of rare species from community-level analyses. *Wildlife Society Bulletin*, **30**, 756–759.
- Rafe, R.W., Usher, M.B. & Jefferson, R.G. (1985) Birds on reserves—the influence of area and habitat on species richness. *Journal of Applied Ecology*, **22**, 327–335.
- Royle, J.A. & Dorazio, R.M. (2008) *Hierarchical Modeling and Inference in Ecology*. Academic Press, Boston.
- Royle, J.A. & Nichols, J.D. (2003) Estimating abundance from repeated presence-absence data or point counts. *Ecology*, **84**, 777–790.
- Russell, R.E., Royle, J.A., Saab, V.A., Block, W.M., Lehmkühl, J.F. & Sauer, J.R. (2009) Modeling the response of wildlife communities to environmental disturbance: quantifying the effect of prescribed fire treatments on avian communities in a coniferous forest in Washington. *Ecological Applications*, in press.
- Samu, F., Csontos, P. & Szinetar, C. (2008) From multi-criteria approach to simple protocol: assessing habitat patches for conservation value using species rarity. *Biological Conservation*, **141**, 1310–1320.
- Sauer, J.R. & Link, W.A. (2002) Hierarchical modeling of population stability and species group attributes from survey data. *Ecology*, **86**, 1743–1751.
- Sergio, F., Newton, I., Marchesi, L. & Pedrini, P. (2006) Ecologically justified charisma: preservation of top predators delivers biodiversity conservation. *Journal of Applied Ecology*, **43**, 1049–1055.
- Smith, C.R., DeGloria, S.D., Richmond, M.E., Gregory, S.K., Laba, M.K., Smith, S.D., Braden, J.L., Feagraus, E.H., Fiore, J.J., Hill, E.A., Ogurcak, D.E. & Weber, J.T. (2001) *The New York Gap Analysis Project*. New York Cooperative Fish and Wildlife Research Unit, Cornell University, Ithaca, NY.
- Soares, C. & Brito, J.C. (2007) Environmental correlates for species richness among amphibians and reptiles in a climate transition area. *Biodiversity Conservation*, **16**, 1087–1102.
- Spiegelhalter, D.J., Thomas, A., Best, N.G. & Lunn, D. (2003) *WinBUGS Version 1.4 User Manual*. MRC Biostatistics Unit, Cambridge, UK.
- Stockwell, D.R.B. & Peterson, A.T. (2002) Effects of sample size on accuracy of species distribution models. *Ecological Modeling*, **148**, 1–13.
- The Heinz Center (2002) *The State of our Nation's Ecosystems*. Island Press, Washington, DC.
- Thompson, W.L. (ed.) (2004) *Sampling Rare or Elusive Species: Concepts, Designs, and Techniques for Estimating Population Parameters*. Island Press, Washington, DC.
- Vellend, M., Lillley, P.L. & Starzomski, B.M. (2008) Using subsets of species in biodiversity surveys. *Journal of Applied Ecology*, **45**, 161–169.
- Vesely, D., McComb, B.C., Vojta, C.D., Suring, L.H., Halaj, J., Holt-hausen, R.S., Zuckerberg, B. & Manley, P.M. (2006) *Development of Protocols to Inventory or Monitor Wildlife, Fish, or Rare Plants*. Gen. Tech. Rep. WO-72. U.S. Department of Agriculture, Forest Service, Washington, DC.
- Wilhere, G.F., Goering, M. & Wang, H. (2008) Average optimacy: An index to guide site prioritization for biodiversity conservation. *Biological Conservation*, **141**, 770–781.
- Williams, P., Gibbons, D., Margules, C., Rebelo, A., Humphries, C. & Pressey, R. (1996) A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. *Conservation Biology*, **10**, 155–174.
- Wilson, D.E., Cole, F.R., Nichols, J.D., Rudran, R. & Foster, S. (1996) *Measuring and Monitoring Biological Diversity: Standard Methods for Mammals*. Smithsonian, Washington, DC.
- Yoccoz, N.G., Nichols, J.D. & Boulenger, T. (2001) Monitoring of biological diversity in space and time. *Trends in Ecology and Evolution*, **16**, 446–453.

Received 28 January 2009; accepted 1 May 2009

Handling Editor: Chris Elphick

Supporting Information

Additional supporting information may be found in the online version of this article.

Appendix S1. Species-specific estimates of occupancy, detection and effects of fragmentation

Appendix S2. Hierarchical community model WinBUGS code

Appendix S3. Species richness calculation with precision estimates

Appendix S4. Comparison of the community model to species-level models

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.