ECOGRAPHY

Research

Integrating over uncertainty in spatial scale of response within multispecies occupancy models yields more accurate assessments of community composition

Luke Owen Frishkoff, D. Luke Mahler and Marie-Josée Fortin

L. O. Frishkoff (https://orcid.org/0000-0001-5738-2140) ✉ (luke.frishkoff@uta.edu), D. L. Mahler and M.-J. Fortin, Dept of Ecology and Evolutionary Biology, Univ. of Toronto, Toronto, ON, Canada. LOF also at: Dept of Biology, Univ. of Texas at Arlington, Arlington, TX, USA.

Ecography
42: 2132–2143, 2019
doi: 10.1111/ecog.04365

Subject Editor: Matthew Fitzpatrick
Editor-in-Chief: Miguel Araújo
Accepted 26 August 2019

Species abundance and community composition are affected not only by the local environment, but also by broader landscape and regional context. Yet, determining the spatial scales at which landscapes affect species remains a persistent challenge, hindering our ability to understand how environmental gradients shape communities. This problem is amplified by rare species and imperfect species detection. Here, we present a Bayesian framework that allows uncertainty surrounding the 'true' spatial scale of species' responses (i.e. changes in presence/absence) to be integrated directly into a community hierarchical model. This scale-selecting multispecies occupancy model (ssMSOM) estimates the scale of response, and shows high accuracy and correct levels of uncertainty in parameter estimates across a broad range of simulation conditions. An ssMSOM can be run in a matter of minutes, as opposed to the many hours required to run normal multispecies occupancy models at all queried spatial scales, and then conduct model selection – a problem that up to now has prohibited scale of response from being rigorously evaluated in an occupancy framework. Alternatives to the ssMSOM, such as GLM-based approaches frequently fail to detect the correct spatial scale and magnitude of response, and are often falsely confident by favoring the incorrect parameter estimates, especially as species' detection probabilities deviate from perfect. We further show how trait information can be leveraged to understand how individual species' scales of response vary within communities. Integrating spatial scale selection directly into hierarchical community models provides a means of formally testing hypotheses regarding spatial scales of response, and more accurately determining the environmental drivers that shape communities.

Keywords: habitat selection, imperfect detection, landscape, traits

Introduction

Features of the landscape beyond the local scale often affect the processes that give rise to patterns of community composition (Wiens 1989, Levin 1992, Kneitel and Chase 2004, Dray et al. 2012, Fortin et al. 2012, McGarigal et al. 2016). Ecologists have

NORDIC SOCIETY OIKOS

www.ecography.org

thus sought to quantify what landscape features explain the presence and abundance of species, and at what spatial scales. Yet, determining how species respond to the landscape has been challenging, in part because the relevant spatial scales at which environmental conditions affect species and communities are rarely known a priori. This difficulty has led to uncertainty regarding the conclusions of many landscape-level studies (Jackson and Fahrig 2015). The development of methods that more robustly incorporate scales of responses within the statistical analysis of communities (Borcard and Legendre 2002, Jombart et al. 2009, Matthiopoulos et al. 2011, Dray et al. 2012, Warton et al. 2016), and more accurately convey uncertainty regarding these scales (Chandler and Hepinstall-Cymerman 2016), has the potential to accelerate basic and applied ecological research. In particular, if spatial scale of response can be accurately assessed, and tied to particular species traits, it may grant greater predictability regarding how species in general will respond to landscape-level change and fragmentation.

When considering landscape-level effects on species presence, abundance, or biomass, two properties of species are generally of interest. First, at what spatial scales do species respond to the environment (Desrochers et al. 2010), and second, how do they respond (positively, negatively, or in more complex non-linear fashions)? The most commonly used approach for determining spatial scale of response to environmental gradients (i.e. the spatial context, or spatial contingency; Fortin et al. 2012) quantifies the average environmental value within buffers of various radii (Holland et al. 2004, Weaver et al. 2012, Zuckerberg et al. 2012, McGarigal et al. 2016), and then repeats a statistical analysis using the environmental covariate at each spatial scale (Fig. 1). For each species in turn (or for some community-level index like species richness), the most likely spatial scale (as quantified by information criteria, correlation coefficient, or slope parameter value) is selected.

This multi-scale analysis approach has been successful in elucidating species and community responses (McGarigal et al. 2016). By considering whole landscapes, it has helped quantify the benefits of small forest fragments to biological communities, and related ecosystem services (Karp et al. 2013, Mendenhall et al. 2016). More generally it has revealed that

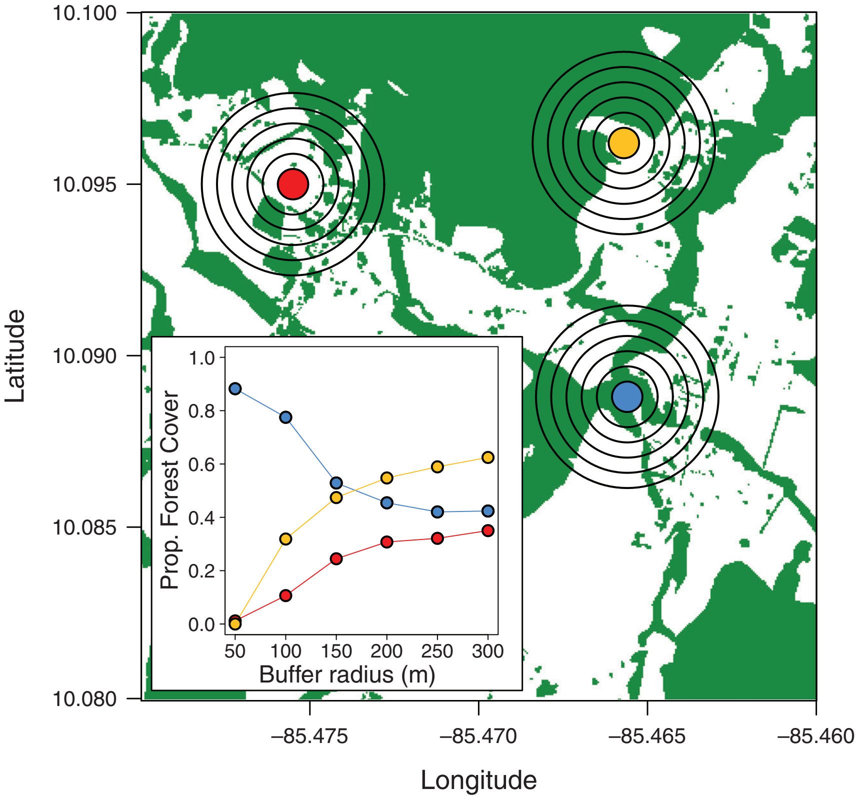

Figure 1. Map of empirical 5 m-resolution subsection of forest cover landscape in northwestern Costa Rica with three fictitious sampling points that break the correlation structure between local and broader landscape forest cover. Forest cover is indicated in green, non-forest in white. Colored points represent 50 m point count radii, and each successive buffer represents an increase in radius of 50 m. Only 50–300 m radii are shown, though for simulation analyses radii extended to 1500 m. Inset depicts the proportion of forest cover surrounding each point at the spatial scales investigated, with lines indicating linear interpolation between buffer sizes. Note that for analysis, the environmental variable should be standardized (centered and scaled to unit variance) within each spatial scale. This eliminates trends of decreasing variation at broader spatial scales, which could compromise model mixing and result in downstream consequences for other parameters (e.g. species' responses to the environment).

species respond to different environmental conditions at different spatial scales, and that species distribution models possess greater predictive power when these multiple scales are directly incorporated (Desrochers et al. 2010, Weaver et al. 2012). However, the current multi-scale approach presents a number of statistical problems, exacerbating uncertainty by treating species individually rather than the community as an integrated whole, and ignoring issues with species detectability. All of these may inflate error in estimating the true spatial scale of response, and in quantifying how species respond to environmental conditions.

The first of these problems is that commonly used single-species model comparison approaches select a single best model, and typically neither quantify nor integrate over uncertainty regarding scale selection. This means that other parameters may be biased if the 'most likely' scale is not the true scale. Relatedly, the set of scales analyzed is often quite small (Desrochers et al. 2010, Jackson and Fahrig 2015), and as a result is unlikely to even include the true spatial scale. Meta-analysis has shown that the most likely spatial scale is often at one of the extremes of those analyzed – suggesting that the true spatial scale is even more extreme (Jackson and Fahrig 2015).

The second major problem is that most analyses examining spatial scale of response rely on a 'stacked' approach – independently analyzing community members in a species-by-species fashion (Dubuis et al. 2011). However, this approach is often inadequate because it necessarily ignores rare species due to lack of power, even though rare species are typically of greatest conservation concern, and may be trophically influential. Ignoring rare species is especially problematic in tropical communities where species rarity is common (MacArthur 1969), and which are particularly threatened by environmental change. Further, a stacked approach is prone to both estimation bias and loss of power (Ovaskainen and Soininen 2011, Banks-Leite et al. 2014). Hierarchical community models have been proposed to move beyond stacked single species assessments (Ovaskainen and Soininen 2011, Warton et al. 2015, Ovaskainen et al. 2017). By assuming that species parameters come from common distributions, overall statistical error is minimized and rare species can be included in analyses.

Finally, imperfect detection of species is a problem for biological communities generally, especially when species traits, site characteristics, or the time or conditions of observation influence detectability. Multispecies occupancy models (MSOMs) are commonly employed to account for imperfect and variable detection within communities (Iknayan et al. 2013). MSOMs are typically implemented in a Bayesian framework (relying on MCMC to overcome challenges in maximizing likelihoods when numerous random effects exist). Yet, model comparison and selection is still difficult to implement with Bayesian models (but see Lele et al. 2007, Hooten and Hobbs 2015). Together this means that comparing models across a large range of spatial scales would be both extremely time consuming (because MCMCs are relatively slow), and non-trivial to implement. Because MSOMs are not easily amenable to adequate testing at multiple spatial scales, the scale of response has generally not been included into community analyses that incorporate imperfect detection. Consequently, their power has not been sufficiently utilized to understand how landscape features structure community-level processes.

Recently Chandler and Hepinstall-Cymerman (2016) suggested including a scale parameter directly within a generalized linear model (GLM) to efficiently estimate a single species' scale of response in a maximum likelihood framework. Here, we take a similar approach to integrate scale selection into a series of models of greater complexity, using a parameter (or group of parameters) that explicitly estimates the spatial scales of response to environmental covariates. When extended beyond single species cases to whole communities, incorporating scale selection results in 1) the appropriate spatial scale(s) being estimated directly from the data within a single model run, while 2) ensuring that other parameter estimates are not biased because they were analyzed at the wrong spatial scale, even for rare species. We first present a scale-selecting multispecies occupancy model (hereafter; ssMSOM), which brings parameter-based scale selection to the community level, while simultaneously taking into account the potential for imperfect detection. We then test the ssMSOM's performance in estimating both community-wide spatial scale of response, and individual species' strengths of response to the environment (i.e. a landscape-level covariate). To determine when the more complex models present meaningful reductions in error we compare the ssMSOM (scale selection, information sharing between species, detection probability between species and environments accounted for) to a series of simpler models. These models include 1) a series of species-by-species GLMs (no scale selection within model, stacked approach, detection probability not accounted for), 2) species-by-species scale selecting GLMs (scale selection, stacked, no detection; roughly equivalent to models from Chandler and Hepinstall-Cymerman (2016)), and 3) scale selecting generalized linear mixed effect models or GLMMs (scale selection, information sharing between species, no detection). Finally, we generalize the ssMSOM to ask a question that cannot be robustly addressed with existing methods: whether species' traits govern the spatial scale at which organisms interact with the environment. While the approach here is demonstrated with a simple single-season occupancy model for multiple species, the internalized scale selection is broadly applicable. For example, it could be directly integrated into abundance models, or combined with flexible hierarchical modeling approaches to query population dynamics through time, or the effects of phylogenetic relatedness or interspecific competition on community structure (Yackulic et al. 2014, Frishkoff et al. 2017).

Material and methods

Model overview

Scale selection is predicated on quantifying an environmental covariate at progressively greater distances from a point at which community composition was measured (e.g. the average proportion of forest cover within 50, 100, 150 m, etc., of survey sites; Fig. 1). These measures can be summarized in a site-by-spatial scale environment matrix, here denoted as **E**. The task then becomes selecting the column(s) of **E** that best describe the data, and biologically interpreting the column(s) in terms of the original spatial scale.

The scale-selecting multispecies occupancy model (ssMSOM) estimates parameters related to occupancy and detection probability in communities containing multiple species (indexed by *i*), across multiple sites (*j*), with multiple site visits (*k*). The observed detection histories ($X_{i,j,k}$) are assumed to derive from unobserved (latent) occupancy states ($Z_{i,j}$, where $Z_{i,j} = 1$ for presence, and 0 for absence) and detection probabilities determined by species, and site ($P_{i,j}$; visit-based variability in detection is ignored for simulations and models, but could be incorporated if desired). Specifically, they can be described as:

$$X_{i,j,k} \sim \text{Bernoulli}\left(Z_{i,j} \times P_{i,j,k}\right) \tag{1}$$

The occupancy state ($Z_{i,j}$) in turn is assumed to come from some underlying occupancy probability according to:

$$Z_{i,j} \sim \text{Bernoulli}\left(\psi_{i,j}\right) \tag{2}$$

Species' occupancy probabilities ($\psi_{i,j}$) can be described according to:

$$\text{logit}\left(\psi_{i,j}\right) = \beta 0_i + \beta 1_i \times \mathbf{w}_j + \gamma 0_j \tag{3}$$

The β terms are the species-specific occupancy intercept and slope with respect to the environment. Here, **w** is a vector of sites' environment values at the spatial scale deemed appropriate by the model, based on some function of the site-by-scale matrix **E**. In this model formulation all species are assumed to interact with the environment at the same spatial scale. A parameter ς is fit in the model to select the appropriate spatial scale (see below).

The detection process is modeled according to:

$$\text{logit}\left(\text{P}_{i,j,k}\right) = \alpha 0_i + \alpha 1_i \times \text{E}_{j,1} \tag{4}$$

Here α0 and α1 are the detection intercept and slope with regards to the environment, respectively. A species' detectabilities are influenced by local environmental conditions (i.e. the first column of the site-by-scale environment matrix **E**, which corresponds to the finest spatial scale at which environmental data were gathered).

All parameters in the α and β groups are estimated for each species, with species terms each drawn from normal distributions of means (μ) and variances (σ^2) estimated from the data ($\alpha 0_i \sim \text{Normal}\left(\mu_{\alpha 0}, \sigma^2_{\alpha 0}\right)$). The γ terms were random intercepts for each site (variances estimated from data around means of 0) designed to incorporate consistent differences across sites in occupancy probabilities for all species that are not accounted for by **E**.

The core purpose of the scale-selecting model is to translate the environment matrix ($\mathbf{E}_{j,s}$) into a vector of site values at the appropriate scale ($\mathbf{w}_j$). The indexing value *s* (representing columns of **E**) determines spatial scale (50 m increments for the example in Fig. 1). A parameter ς is estimated in the model to determine which scale (i.e. column of matrix **E**) best fits the data. Thus if ς takes integer values it acts as a stand in for *s*. For example, when $\varsigma = 1$, then $\mathbf{w}_j = \mathbf{E}_{j,1}$, when $\varsigma = 2$ $\mathbf{w}_j = \mathbf{E}_{j,2}$, etc. Multiple smoothing functions could be chosen when ς takes non-integer values. Here, we used linear interpolation between columns of **E** to generate continuous environment scores (**w**) based on values of ς. Note that when converting values of ς back to biologically interpretable scale values, one must multiply by the resolution at which the environment was measured to generate **E**. For example, if every column of **E** represents a 50 m increment around a focal point, a ς of 4.5 corresponds to 4.5 × 50 m = 225 m.

Because ssMSOMs are implemented using MCMC, model fitting results in a posterior distribution for values of ς, which fully integrates over the uncertainly regarding the proper spatial scale, and which further can be used to select the most appropriate spatial scale (e.g. posterior mean and/or highest posterior density) or an interval of spatial scales that well describe the data. Because **E** is centered and scaled, β1 represents the change in occupancy probability by increasing the value of the environmental variable by 1 SD, at the relevant spatial scale.

In addition to drawing species' parameters from common distributions (as done here), joint community models can also estimate covariances between species' parameters. We do not do this here because the estimation procedure significantly adds to computational time, because we do not simulate these covariances, and because estimation of covariances is unnecessary to demonstrate the functionality of integrated scale selection within occupancy models. For empirical analyses – particularly if there are specific hypotheses regarding how species parameters covary – adjusting the presented ssMSOM to accommodate these interspecific covariances is recommended (Karp et al. 2018).

For demonstration and testing purposes, we assume that a single environmental gradient affects community composition. However, the ssMSOM could be generalized to include multiple environmental conditions (multiple site-by-scale environment matrices, e.g. $\mathbf{E1}_{j,s}$, $\mathbf{E2}_{j,t}$, $\mathbf{E3}_{j,u}$, etc.), each affecting communities at different spatial scales (with scale parameters ς1, ς2, ς3, etc., each independently estimated from the data to provide indexing for *s*, *t*, *u*, etc.). Further, species could respond in non-linear fashions to a single environmental gradient. In the supplement, we report performance of a case in which species respond to the environment in a quadratic fashion (Supplementary material Appendix 1 Fig. A1). Finally, not all relevant aspects of the environment will always be measured. If these unmeasured variables affect species at a different spatial scale than the measured covariate of interest it could introduce error into scale selection. We evaluate the consequences of this possibility in the supplement (Supplementary material Appendix 1 Fig. A2).

Simulation conditions

To test the performance of the ssMSOM, we simulated communities using an environmental matrix (**E**) based on empirical landscape forest cover. Spatial forest cover data for simulation and analysis came from a recent study of how local and landscape-level habitat conversion affects community composition in northwestern Costa Rica (Karp et al. 2018). In that study, sites were selected to ensure that local forest cover varied independently from landscape-level forest cover. To measure surrounding forest cover, all tree cover within 1.5 km of sites was classified using high-resolution Google Earth images obtained from 2013 to 2016. The resulting 5 m-resolution tree cover map was verified based on ground-truthed data collected in the field. For analysis, site level forest cover proportion was calculated in radii from 50 m to 1500 m, in 50 m increments, resulting in **E** having 30 columns.

To test the performance of ssMSOM under a variety of conditions, we simulated 240 datasets (eight at each of the 30 spatial scales), each with 16 species (N_{sp}), across 50 sites (N_{site}), with three visits per site (N_{visit}). All species parameters were drawn from normal distributions, generating diversity in species' commonness, and species' responses to tree cover. This diversity of overall commonness and responses to the environment mimics patterns observed in many empirical systems. We repeated these simulations under five alternative detection regimes.

1) Perfect detection, where the probability of detecting a species at a site if the site is occupied is 1 (denoted 'perfect' below).

2) High detection probability: average detectability approximately 0.5 (denoted 'high' below).

3) Low detection probability: average detectability approximately 0.25 (denoted 'low' below).

4) Low detection with detection affected by local environment: average detectability approximately 0.25, but rising to approximately 0.5 under high local values of **E** and dropping to approximately 0.1 under low local values of **E** (denoted 'env' below).

5) Low detection with species-specific variation in detectability by local environment: Same as 4, but some species increase in detectability with increasing local values of **E**, while others decrease in detectability (denoted 'env. var' below, to indicate environment with variation between species).

These simulation conditions were chosen to represent a spectrum of biologically relevant detection probabilities. For most species detectability will be far below perfect (e.g. mean detectability of 0.4 for a study of bats (Frank et al. 2017) and 0.26 for a study of birds (Carrillo-Rubio et al. 2014)). However, detection may approach perfect for some sessile organisms. The environment effect on detectability is meant to capture situations such as when dense vegetation obfuscates visual detection or muffles sound. Similarly, the final case is representative of when species may differ in how detectable they are in specific environments based on camouflage mismatch (e.g. detection of green lizards is difficult against leaves, but easy in a rocky environment, whereas detection of brown lizards could show the opposite trend). For an overview of all simulation parameters see Supplementary material Appendix 1 Table A1.

Model comparison

To examine ssMSOM performance, we compared it to a series of simpler models: A GLMM approach with scale selection (referred to here as ssGLMM), single species GLMs implemented in a Bayesian framework with scale selection (ssGLM), and simple single species GLMs implemented in a maximum likelihood framework, but without integrated scale selection. These model comparisons assess the effects of accounting for detection probability (ssMSOM), information sharing between species (ssMSOM and ssGLMM), and internalized scale selection (ssMSOM, ssGLMM and ssGLM), in terms of accuracy and error for parameter estimation. These by no means represent the full set of potential comparisons, but are used to evaluate when particular model assumptions are important for accuracy in specific hypothesis testing scenarios.

We describe these models from simplest, to most complex. The simple single-species GLMs were fit using maximum likelihood, and rerun for each of the 30 spatial scales in **E**. These models (referred to as stacked GLMs throughout) do not account for detection probability and instead focus on naïve occupancy. More specifically multiple site visits are not assessed and a species is deemed to be present if it was ever detected at a site, and absent otherwise. The stacked GLMs are described by standard logistic function of the form:

$$\begin{aligned} &\text{logit}\left(\psi_j\right) = \beta 0 + \beta 1 \times \mathbf{e}_j \\ &Y_j \sim \text{Bernoulli}\left(\psi_j\right) \end{aligned} \tag{5}$$

here, **e** is a vector of environmental values at a single focal spatial scale (e.g. for a model examining the most local scale $\mathbf{e}_j = \mathbf{E}_{j,1}$), ψ_j is the naïve occupancy probability of the focal species and Y_j is the naïve occupancy state. We ran independent models at each of the 30 spatial scales, and then used AIC_c to choose the optimal spatial scale for each species in turn.

AIC_c is a variation of the Akaike information criterion that corrects for potential overfitting when sample sizes are small, but converges to AIC when sample size is large. Note that for single-species GLMs no random effect of site is needed, because only one observation comes from each site (species observed versus species not observed).

The ssGLMs take this framework, and integrates scale selection. They therefore require only a single model run for each species in turn. The ssGLM is described by:

$$\begin{aligned} &\mathrm{logit}\left(\psi_j\right) = \beta 0 + \beta 1 \times \mathbf{w}_j \\ &\mathrm{Y}_j \sim \mathrm{Bernoulli}\left(\psi_j\right) \end{aligned} \tag{6}$$

where **w** is fit as in the ssMSOM with the parameter ς to determine how **E** is translated into **w**.

Finally, we ran a scale-selecting GLMM. This model is described by:

$$\begin{aligned} &\mathrm{logit}\left(\psi_{i,j}\right) = \beta 0_i + \beta 1_i \times \mathbf{w}_j + \gamma 0_j \\ &\mathrm{Y}_{i,j} \sim \mathrm{Bernoulli}\left(\psi_{i,j}\right) \end{aligned} \tag{7}$$

and is identical to the ssMSOM except that it examines naïve occupancy, rather than quantifying detection probability.

We focus on two core questions when evaluating the ssMSOM versus other model approaches. First, how accurate and precise is parameter estimation in determining the scale of response? Second, even when estimating spatial scale is not the primary goal and is therefore considered a nuisance variable, when does the ssMSOM result in more accurate estimates of how species respond to the environment (β1) than simpler approaches?

To quantify the accuracy of the models, we calculated the root mean square error (RMSE) across the entire community of the family of parameter estimates from the true simulated value. For ssMSOMs, ssGLMMs, and ssGLMs, we considered the highest posterior density (posterior mode) of the scale parameter ς. For the stacked GLMs, we consider the spatial scale that minimizes information criteria (AIC_c). We use the highest posterior density to summarize ς because spatial scales are bounded by the broadest and finest scale queried. The posterior mean would have returned centrally biased summaries of the posterior for species that respond at very fine (or very broad) scales. To ensure the results are comparable across approaches, we calculate RMSE for each species in turn, even though in the case of the ssMSOM and ssGLMM the parameter ς is estimated for all species simultaneously, and is therefore identical for all species in the simple ssMSOM model (see below for a relaxation of this assumption).

We also examine coverage probability of the posterior estimates. If models are behaving as expected, the 95% CIs (or Bayesian credible intervals; BCIs) of the parameter estimates should contain the true value 95% of the time. Coverage probabilities around ς for the ssMSOMs, ssGLMMs and ssGLMs were calculated using the 95% highest posterior density interval of ς. For stacked GLMs, multiple models are compared to determine the most likely spatial scale. In order to calculate a value similar to coverage probabilities of the spatial scale in these cases, we first calculated the AIC_c weight for all spatial scales, and then asked whether the true spatial scale was within the top 95% of the cumulative model weights (summed from most likely to least likely). Such AIC_c weights represent the relative likelihood of each model, and are calculated for each model as $e^{-0.5\times\Delta\mathrm{AICc}}$ of the focal model divided by the sum of these values across all models.

We additionally compared the computational efficiency of the ssMSOM versus running a single MSOM on each of the 30 spatial scales, and using cross-validation to choose which spatial scale best described the data.

Testing the role of traits

In nature, species may not respond to the environment at the same spatial scale, and researchers may therefore wish to test which species traits predict the scale of response. For example, if body size determines scale of response then estimating the effect of this trait will provide predictability for the scale at which habitats must be preserved for rare species with known body sizes. To allow for trait effects, we relax the assumption that all species respond at the same scale by adding a hierarchy to the scale-selection procedure – specifically species scales of response are drawn from a shared distribution. To accomplish this we first replace **w** by a matrix of environmental conditions at the relevant spatial scale calculated separately for each species, denoted here by **W**:

$$\mathrm{logit}\left(\psi_{i,j}\right) = \beta 0_i + \beta 1_i \times \mathbf{W}_{i,j} + \gamma 0_j \tag{8}$$

W is a function of a species-specific scale parameter ς_i which is calculated:

$$\begin{aligned} &\varsigma_i = \eta \times \mathrm{trait}_i + \zeta_i \\ &\zeta_i \sim \mathrm{Normal}\left(\mu_{\mathrm{scale}}, \sigma^2_{\mathrm{scale}}\right) \end{aligned} \tag{9}$$

where η is the effect of the trait and ζ contains both the intercept, and variation around the trait effect. In the case of a model without trait effects but with variable species scales of response, then ζ_i is equal to ς_i. To accommodate scale selection we bound values of ς_i in the MCMC to be between 1 (equivalent to the finest spatial scale; 50 m) and the maximum number of spatial scales (in our case 30, corresponding to 1500 m).

We simulated and ran the trait model on 100 datasets (20 per detection regime) calculating coverage probabilities, statistical power, and accuracy of μ_{scale}, σ_{scale}, η and ς_i. These parameters all relate to ς and therefore reflect which columns of the **E** matrix are selected to evaluate the scale of response, and must be multiplied by the resolution of the **E** matrix to arrive at biologically meaningful scales. As such μ_{scale} and σ_{scale} representing the mean column and the standard deviation, expressed in number of columns, at which species respond after taking in account the trait effect. We ran these simulations for a spatial scale where μ_{scale} was 10 (i.e. the true average column of **E** was the 10th, corresponding to 500 m), a σ^2_{scale} of 25 (i.e. a variance of 62 500 m, or a standard deviation of 250 m) and a η of 3 (i.e. an expected scale-of-response increase of 150 m for every standard deviation of the trait value increased). Species trait values were simulated from a normal distribution of mean 0 and standard deviation 1. We evaluate model performance from a calibrated Bayes perspective (Little 2006), quantifying frequentist performance of the MCMC-based models – as such power for the trait effect was determined by the proportion of simulations for which the 95% BCI excluded 0. We further tested the role of sample size (number of species) on accuracy and power by repeating the simulation and fitting procedure on a community of 64 species. We compared accuracy and power of the ssMSOM to stacked GLMs. The role of traits in determining spatial scale of response can not be directly estimated from stacked GLMs, because doing so requires a multispecies model. Therefore in the case of η, we tested the ssMSOM against the performance of a secondary linear regression in which the predictor variable was species trait values, and the response variable was the best spatial scales derived from each species' GLM (as determined through AIC_c). Simulation parameters for the trait model test are summarized in Supplementary material Appendix 1 Table A2.

Model fitting

Simple GLM models were fit using maximum likelihood. More complex models, especially those with many random effects, have difficulty converging using ML optimizers. As such, ssMSOMs, ssGLMMs and ssGLMs were fit in a Bayesian framework using MCMC – specifically using JAGS through the R environment. AIC_c was used when selection between multiple GLMs fit by maximum likelihood.

We selected MCMC conditions to generate a sufficient number of converged models to analyze according to the following process. We first examined five test runs for the ssMSOM when individual species varied in how detection responds to the environment (the 'env.var' condition), which was the model we expected to converge with the most difficulty. These test runs indicated that models converged (Gelman–Rubin convergence diagnostic (Gelman and Rubin 1992) for all parameters < 1.1, and appropriate traceplots after 10 000 iterations). To better ensure convergence, we then quintupled this number of iterations for all subsequent model runs. Final MCMC run conditions included 50 000 iterations, with a burnin of 10 000 and a thin rate of 160, with four chains, yielding 1000 posterior samples. We then checked all parameters for convergence of the fitted models, and discarded datasets if convergence failed for any parameter in any of the models (ssMSOM, ssGLMM, ssGLM) for that dataset. This procedure resulted in the following number of datasets analysed for each model condition: Perfect detection: 211; High detection: 211, Low detection: 140, Environmentally variable detection ('env'): 151, Environmentally and species-specific variation in detection ('env.var'): 198. For MCMC analyses diffuse priors were used throughout.

Data deposition

Data and code available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.bk3j9kd6f> (Frishkoff et al. 2019).

Results

Inferring spatial scale

The posterior mode of spatial scales from the ssMSOM tended to accurately estimate the true spatial scale of response, and was typically off by less than 100 m under the conditions simulated (Fig. 2, 3). Performance was roughly equivalent as long as models shared information between species when assessing spatial scale (i.e. ssGLMM), even though they do not account for detection probability. In contrast, ssGLMs and stacked GLMs failed to consistently recover the true spatial scale for the majority of species in the community, even when detection was perfect. When detection itself varied along the environmental gradient in a species-specific manner, using GLM-based approaches resulted in error in estimating species response scales that was no better (and sometime worse) than guessing a scale at random (Fig. 3a).

Despite elevated RMSE in estimating spatial scale by GLM approaches, most models demonstrated roughly correct coverage probabilities when estimating scale of response, regardless of detection regime. Stacked GLMs had a slight tendency to exclude the true scale from the top 95% of AIC_c-weighted models, and this behavior was exacerbated as detection probability deviated from perfect (Fig. 3b).

Estimating species responses to the environment

Estimates of species responses to the environment ($\beta1$) tended to be more accurate in the ssMSOM and ssGLMM than in GLM-based approaches (Fig. 4a). In general, estimates are more accurate when detection probabilities are high. Detection regime does, however, strongly affect coverage probabilities of confidence intervals. If detection is perfect, most models types show roughly correct coverage probabilities (Fig. 4b). When detection deviates from perfect non-occupancy models generate falsely confident results, with the true values of species responses excluded from the 95% confidence intervals nearly 30% of the time under some simulated conditions (i.e. 6× the nominal error rate). In contrast the ssMSOMs possess 95% CIs that behave as expected, regardless of detection regime, even though in some individual simulation runs RMSE for the ssMSOM was quite high. These behaviors hold even when species respond to the

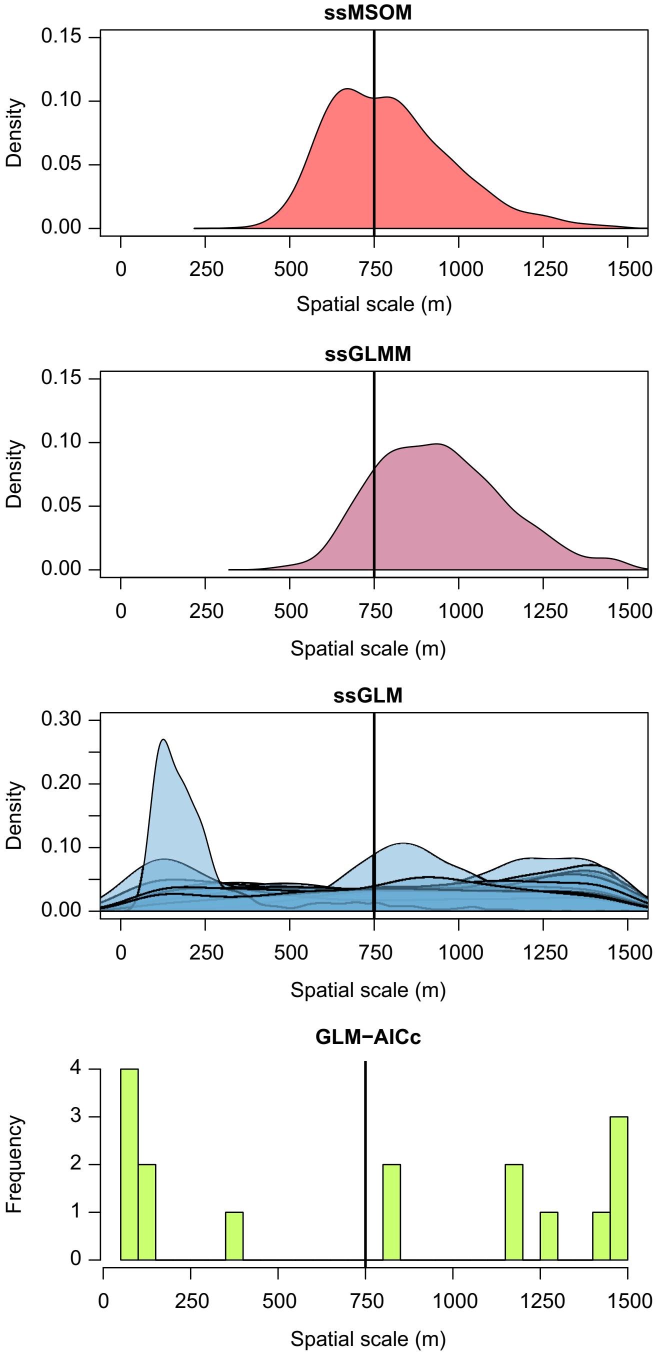

Figure 2. Scale selection from scale-selecting models. Example comes from the 'env.var' detection case (mean detection = 0.25, and species' detectabilities vary as a function of the environment). The simulated spatial scale of response was 750 m (marked by bold vertical lines). Panels for ssMSOM, ssGLMM, and ssGLM depict posterior distribution of the spatial scales that describes species' responses to the environment. In the case of ssGLM spatial scale of response is allowed to vary by species, and all species, posteriors are plotted. The GLM-AIC$_c$ panel depicts a histogram showing the spatial scale that minimizes AIC$_c$ for each single-species set of GLMs.

environment in non-linear fashions (Supplementary material Appendix 1 Fig. A1).

Efficiency

The ssMSOM, by incorporating a scale-selecting parameter, allows spatial scale to be directly estimated in a single model run. An alternative means of getting most of the same benefits would be to run a MSOM on every spatial scale of interest, and then use model selection techniques (e.g. *k*-fold cross-validation) to select the spatial scale that best explains the data (albeit without inferring the degree of uncertainty in this measure, or incorporating the consequences of this uncertainty for other parameter estimates). Under the conditions simulated, these alternatives appear to yield roughly the same conclusions regarding the spatial scale of response (Supplementary material Appendix 1 Fig. A5). However, the ssMSOM can be run in 5 min of cpu time on a 2.8 Ghz processor, while completing five-fold cross-validation on all 30 queried spatial scales takes ~5.5 h.

Detecting trait effects on species' scale of response

In a more complex version of an ssMSOM, each species could be allowed to respond to the environment at a unique scale (Fig. 5a), and these scales of response could be controlled by some species' trait. To present the approach, and keep comparisons simple, we evaluated this approach only against the stacked GLMs fit by maximum likelihood. In this case, the quality of inferences made using an ssMSOM depended heavily on the number of species in the dataset. With small numbers of species, the ssMSOM estimated species-specific spatial scale with more error than the stacked GLM approach. This increase in error likely comes about because hyper-parameters (mean and variance of community level community responses) cannot be estimated accurately with so few species. However, if numbers of species are sufficiently large (64 in case of simulations here), the ssMSOM has less overall error (Fig. 5b).

Despite problems detecting individual species' scales of response when species numbers are low, the ssMSOM always outperforms stacked GLMs when testing for the role of traits in mediating these responses. Using stacked GLMs and secondary regressions to test for the role of traits tends to lead to estimates of trait effects that are biased towards 0 (Fig. 5c). Bias was particularly pronounced when detection probability was low, and varied with the environment.

The power to detect significant trait effects was greater for ssMSOMs so long as the number of species analyzed was sufficiently large (64 species here). However, power is very low when only 16 species are analyzed, and degrades to approximately 0 as detection regimes depart from perfect. With 64 species, power increases substantially over the 16-species case, though more complex detection regimes lose power compared to perfect detection regimes (Fig. 5d).

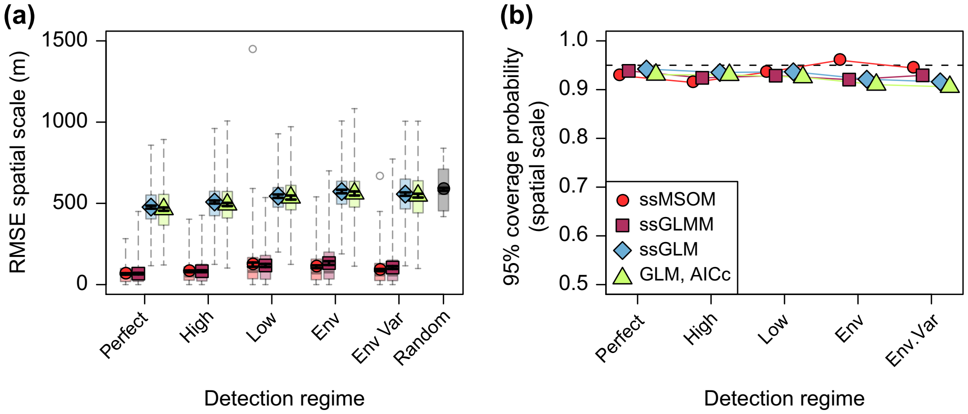

Figure 3. (a) Root mean square error for spatial scale (ς) across detection regimes, comparing posterior mode from ssMSOMs, ssGLMMs, ssGLMs, and stacked GLMs with spatial scale selected using AIC_c. 'Random' indicates the distribution of RMSEs that one would obtain if selecting spatial scales randomly along the uniform range from 50 m to 1500 m. Only when all species inform scale selection (as with ssMSOM and ssGLMM) in accuracy high. Means and standard errors (points and dark error bars) are plotted over light boxplots depicting full range of the data (outliers indicated by light grey points). (b) Probability that the true spatial scale is within the set selected by the models. For ssMSOMs, ssGLMMs and ssGLMs, coverage probability indicates the proportion of simulations for which the true spatial scale was within the 95% HPD interval of the spatial scale posterior. For stacked GLMs coverage probability indicates the proportion of species across simulations for which the true spatial scale was within the top 95% of AIC_c weighted models. The expected coverage probability of 0.95 is depicted as a black dotted line.

Discussion

Here, we described and tested the statistical properties of scale-selecting multispecies occupancy models (ssMSOMs) against simpler methods of ascertaining species' and communities' scales of response to the environment. We find that internalizing scale selection into multispecies models provides an accurate means to quantify the scale of response, and the uncertainty around that scale of response. Additionally, it allows new questions to be addressed, such as: how do species traits affect the scale of response? The key strength of the internalized scale selection is that it does not rely on setting the spatial scale a priori. Scale-selecting models could help researchers avoid the pitfall of selecting too few scales to investigate, and thus missing the true scale of response (Jackson and Fahrig 2015). This of course requires that researchers first extract landscape data from a sufficiently broad a range of scales, ideally in fine spatial increments. Such sampling allows spatial scale to be treated as continuous, such that 95% CIs can be created, and inferences made as with any other continuous parameter in the model.

Whether to consider species scale of response to be a species-specific property (as modeled in the trait-based ssMSOM approach), or an average community property (as modeled in the simple ssMSOM) will depend on the nature

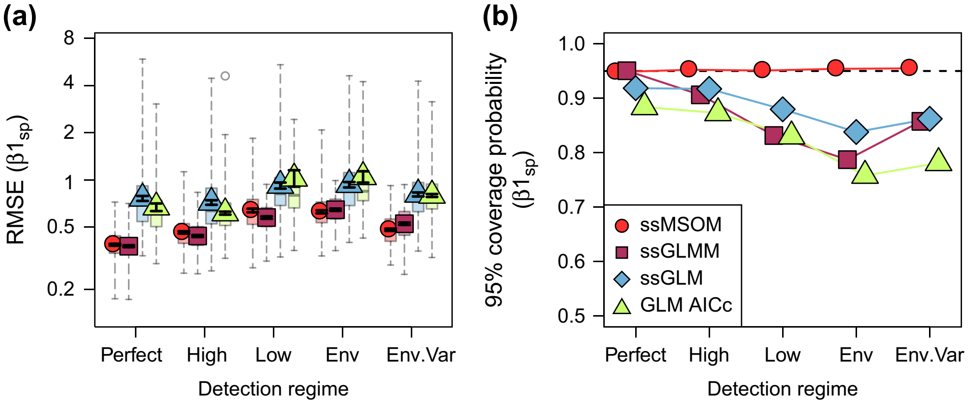

Figure 4. Accuracy of estimates of species responses to the environment ($\beta 1_{sp}$). (a) Community-level RMSE for each simulated dataset across detection regime. Means and standard errors across all datasets (points and dark error bars) are plotted over light boxplots depicting full range of the data. (b) Coverage probabilities around true value of response to the environment ($\beta 1_{sp}$). Each point represents the proportion of species across all simulations for which the true value fell within the 95% (B)CIs of the estimate. Values below 0.95, depicted as a black dotted line, indicate inflated error.

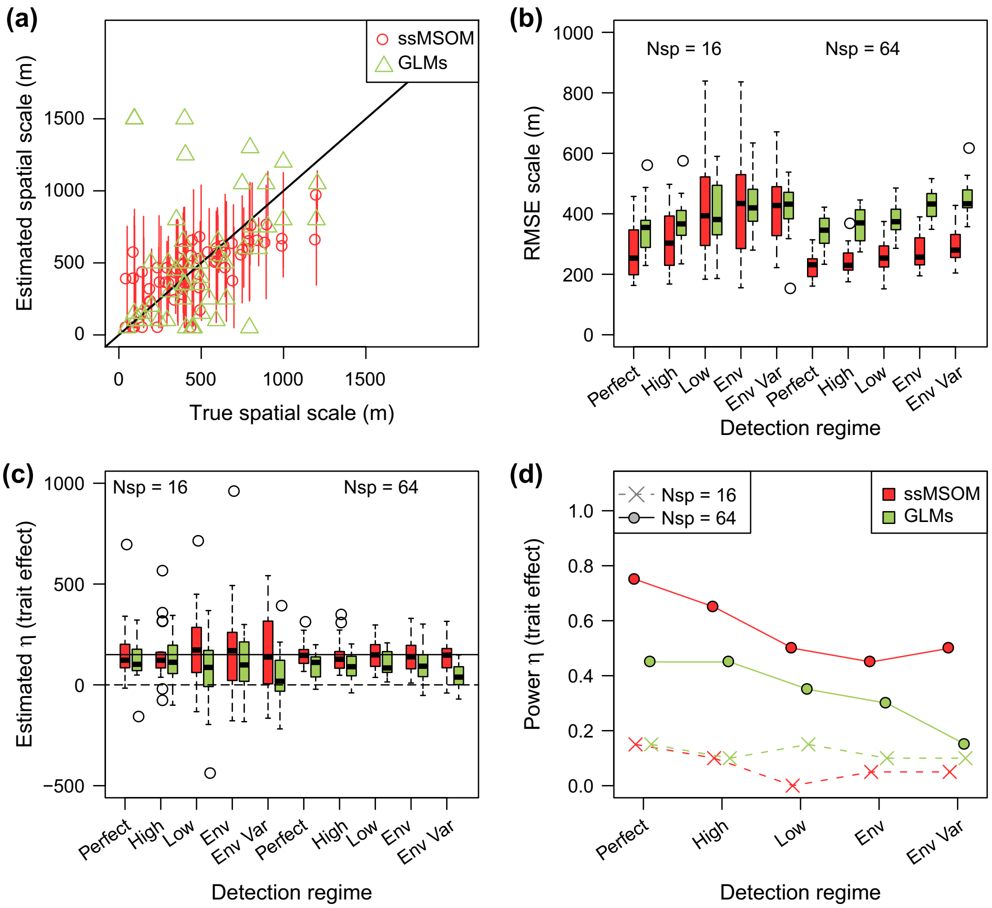

Figure 5. Accuracy of species-specific scale estimates and power to detect the effect of a trait on species' scales of response. (a) An example 64-species dataset and model-fitting (simulated under the 'low detection' regime) evaluating match between true species scale of response versus estimated scale (ς_i). Red points depict species' posterior modes from ssMSOMs, with lines depicting 95% BCIs. Triangles depict the minimum AIC$_c$ scale from stacked GLMs. Perfect estimation would fall on the 1–1 line. (b) Root mean square error of estimating ς_i for ssMSOMs, and stacked GLMs. Boxplots represent medians, interquartile ranges and full ranges of total RMSE (across all species) within each of 20 simulated datasets and model runs for each detection-by-species number combination. (c) Estimation of correlation of trait values with spatial scale of response. Boxplots represent posterior means for ssMSOMs, and MLEs in secondary linear regression for stacked GLMs. Solid horizontal line indicates the simulation value of the trait effect (150 m increase per 1 standard deviation of trait). (d) The proportion of dataset simulations and model fittings for which the 95% BCI around the trait effect excludes 0 (for ssMSOMs) or likelihood ratio test significance of the trait effect in secondary linear regression (for GLMs).

of the study and the system at hand. Empirical studies have shown that different species respond to the environment at different spatial scales (Chambers et al. 2016), and theoretical approaches suggest that some species traits may modulate the scale of response (Jackson and Fahrig 2012). However, we show in our simulations that empirical analyses conducted on a species-by-species basis (as many past studies have been) are often unable to recover the true spatial scale at which species respond, and show high heterogeneity in the scale of response even if all species are simulated to respond at the same spatial scale. While many species undoubtedly do respond at different scales, this finding casts some doubt on the specific estimates of scales of response presented in past empirical studies. The high degree of inaccuracy and low power inherent in the GLM approach may be partially responsible for the lack of correlation between empirically estimated scales of response, and species traits thought to modulate these scales (Jackson and Fahrig 2015).

Using the ssMSOM models presented here, future empirical work could address these hypotheses (i.e. by estimating the value of σ_{scale}, and determining the probability that this term is greater than 0). However, when taking a flexible scale-estimation approach it is essential to use as fine-scale environmental data as possible. If environmental data are coarse with respect to the resolution at which species interact with the environment, then the estimated spatial scale of response will be strongly upwardly biased and overall model performance will suffer (Mendenhall et al. 2011).

In real communities, species' responses to the broader landscape might be predicated on conditions at the local scale (i.e. an interaction between local and landscape scales). For example, a farmland bird species might benefit from landscape-level tree cover when the local habitat is agriculture, but might only exist in forest when there are low amounts of landscape-level tree cover because it uses forest edge habitat. Allowing interaction terms between landscape and local effects (e.g. forest cover as estimated within a point count radius) will allow these types of species interactions with the environment to be tested (Matthiopoulos et al. 2011, Paton and Matthiopoulos 2016).

When assessing real-world communities, accounting for residual spatial correlation between sampling points is often important. Adding spatial covariance structures to the ssMSOM is an obvious extension to address these issues, while also determining the scale at which species respond to measured environmental variables. Accounting for spatial autocorrelation is well established, with robust methods for both single, and multispecies hierarchical modeling frameworks (Latimer et al. 2009, Schliep et al. 2018).

Chandler and Hepinstall-Cymerman (2016) pointed out that the step function used to calculate proportion of focal habitat within a given radius has no theoretical basis, and instead favor a Gaussian weighting function. This approach could be easily implemented with the Bayesian framework presented here, by calculating **w** based on the output of the weighting function over incremental changes in its key parameter. While alternatives to the commonly used step functions are certainly appealing on theoretical grounds, at least one study that examined Gaussian weighting versus a step function radius method found that models performed roughly equivalently (Timm et al. 2016).

Humans are altering landscapes across the globe, such that the remaining extent of natural habitats are often much diminished and severely fragmented (Haddad et al. 2015, 2017). Such complex, heterogeneous landscapes challenge ecologists' abilities to discern the underlying environmental drivers of community composition. Achieving successful conservation strategies in these landscapes requires simultaneously describing and predicting how these spatially heterogeneous environments affect not just individual species, but entire communities. Internalizing spatial scale selection within community models offers a computationally efficient approach to uncover the environmental drivers behind such community change while accommodating the unavoidable uncertainty in the 'true' scale of species' responses. The ssMSOM possesses generally high accuracy and correct credible intervals when identifying both the spatial scale of response, and the direction and magnitude with which individual species respond to environmental gradients. Further, it correctly identifies how traits modulate species' scales of response, so long as a sufficient number of species are included in the analysis. This approach represents a promising path forward for understanding the ecological drivers of community composition, and the consequences of ongoing environmental change.

Acknowledgements – We would like to thank D. S. Karp and A. Echeverri for use of fine-scale forest cover data from northwestern Costa Rica.

Funding – This study was supported by a Univ. of Toronto Ecology and Evolutionary Biology Postdoctoral Fellowship to LOF and NSERC Discovery Grants to DLM and M-JF.

References

Banks-Leite, C. et al. 2014. Assessing the utility of statistical adjustments for imperfect detection in tropical conservation science. – J. Appl. Ecol. 51: 849–859.

Borcard, D. and Legendre, P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. – Ecol. Model. 153: 51–68.

Carrillo-Rubio, E. et al. 2014. Use of multispecies occupancy models to evaluate the response of bird communities to forest degradation associated with logging. – Conserv. Biol. 28: 1034–1044.

Chambers, C. L. et al. 2016. Influences of scale on bat habitat relationships in a forested landscape in Nicaragua. – Landscape Ecol. 31: 1299–1318.

Chandler, R. and Hepinstall-Cymerman, J. 2016. Estimating the spatial scales of landscape effects on abundance. – Landscape Ecol. 31: 1383–1394.

Desrochers, A. et al. 2010. Area-sensitivity by forest songbirds: Ttheoretical and practical implications of scale-dependency. – Ecography 33: 921–931.

Dray, S. et al. 2012. Community ecology in the age of multivariate spatial analysis. – Ecol. Monogr. 82: 257–275.

Dubuis, A. et al. 2011. Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. – Divers. Distrib. 17: 1122–1131.

Fortin, M. J. et al. 2012. Spatial statistics, spatial regression and graph theory in ecology. – Spat. Stat. 1: 100–109.

Frank, H. K. et al. 2017. Phylogeny, traits and biodiversity of a neotropical bat assemblage: close relatives show similar responses to local deforestation. – Am. Nat. 190: 200–212.

Frishkoff, L. O. et al. 2017. Phylogenetic occupancy models integrate imperfect detection and phylogenetic signal to analyze community structure. – Ecology 98: 198–210.

Frishkoff, L. O. et al. 2019. Data from: Integrating over uncertainty in spatial scale of response within multispecies occupancy models yields more accurate assessments of community composition. – Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.bk3j9kd6f>.

Gelman, A. and Rubin, D. B. 1992. Inference from iterative simulation using multiple sequences. – Stat. Sci. 7: 457–472.

Haddad, N. M. et al. 2015. Habitat fragmentation and its lasting impact on Earth's ecosystems. – Sci. Adv. 1: e1500052.

Haddad, N. M. et al. 2017. Connecting models, data and concepts to understand fragmentation's ecosystem-wide effects. – Ecography 40: 1–8.

Holland, J. D. et al. 2004. Determining the spatial scale of species' response to habitat. – Bioscience 54: 227.

Hooten, M. B. and Hobbs, N. T. 2015. A guide to Bayesian model selection. – Ecol. Monogr. 85: 3–28.

Iknayan, K. J. et al. 2013. Detecting diversity: emerging methods to estimate species diversity. – Trends Ecol. Evol. 29: 97–106.

Jackson, H. B. and Fahrig, L. 2012. What size is a biologically relevant landscape? – Landscape Ecol. 27: 929–941.

Jackson, H. B. and Fahrig, L. 2015. Are ecologists conducting research at the optimal scale? – Global Ecol. Biogeogr. 24: 52–63.

Jombart, T. et al. 2009. Finding essential scales of spatial variation in ecological data: a multivariate approach. – Ecography 32: 161–168.

Karp, D. S. et al. 2013. Forest bolsters bird abundance, pest control and coffee yield. – Ecol. Lett. 16: 1339–1347.

Karp, D. S. et al. 2018. Agriculture erases climate-driven β-diversity in Neotropical bird communities. – Global Change Biol. 24: 338–349.

Kneitel, J. M. and Chase, J. M. 2004. Trade-offs in community ecology: linking spatial scales and species coexistence. – Ecol. Lett. 7: 69–80.

Latimer, A. M. et al. 2009. Hierarchical models facilitate spatial analysis of large data sets: a case study on invasive plant species in the northeastern United States. – Ecol. Lett. 12: 144–154.

Lele, S. R. et al. 2007. Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. – Ecol. Lett. 10: 551–563.

Levin, S. A. 1992. The problem of pattern and scale in ecology. – Ecology 73: 1943–1967.

Little, R. J. 2006. Calibrated Bayes: a Bayes/frequentist roadmap. – Am. Stat. 60: 213–223.

MacArthur, R. H. 1969. Patterns of communities in the tropics. – Biol. J. Linn. Soc. 1: 19–30.

Matthiopoulos, J. et al. 2011. Generalized functional responses for species distributions. – Ecology 92: 583–589.

McGarigal, K. et al. 2016. Multi-scale habitat selection modeling: introduction to the special issue. – Landscape Ecol. 31: 1157–1160.

Mendenhall, C. D. et al. 2011. Predictive model for sustaining biodiversity in tropical countryside. – Proc. Natl Acad. Sci. USA 108: 16313–16316.

Mendenhall, C. D. et al. 2016. Quantifying and sustaining biodiversity in tropical agricultural landscapes. – Proc. Natl Acad. Sci. USA 113: 14544–14551.

Ovaskainen, O. and Soininen, J. 2011. Making more out of sparse data: hierarchical modeling of species communities. – Ecology 92: 289–295.

Ovaskainen, O. et al. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. – Ecol. Lett. 20: 561–576.

Paton, R. S. and Matthiopoulos, J. 2016. Defining the scale of habitat availability for models of habitat selection. – Ecology 97: 1113–1122.

Schliep, E. M. et al. 2018. Joint species distribution modelling for spatio-temporal occurrence and ordinal abundance data. – Global Ecol. Biogeogr. 27: 142–155.

Timm, B. C. et al. 2016. Multi-scale Mexican spotted owl (*Strix occidentalis lucida*) nest/roost habitat selection in Arizona and a comparison with single-scale modeling results. – Landsc. Ecol. 31: 1209–1225.

Warton, D. I. et al. 2015. So many variables: joint modeling in community ecology. – Trends Ecol. Evol. 30: 766–779.

Warton, D. I. et al. 2016. Extending joint models in community ecology: a response to Beissinger et al. – Trends Ecol. Evol. 31: 737–738.

Weaver, J. E. et al. 2012. An invasive species' relationship with environmental variables changes across multiple spatial scales. – Landscape Ecol. 27: 1351–1362.

Wiens, J. A. 1989. Spatial scaling in ecology. – Funct. Ecol. 3: 385–397.

Yackulic, C. B. et al. 2014. The roles of competition and habitat in the dynamics of populations and species distributions. – Ecology 95: 265–279.

Zuckerberg, B. et al. 2012. Overlapping landscapes: a persistent, but misdirected concern when collecting and analyzing ecological data. – J. Wildl. Manage. 76: 1072–1080.

Supplementary material (available online as Appendix ecog-04365 at <www.ecography.org/appendix/ecog-04365>). Appendix 1.