2017

A Bayesian method for assessing multi-scale species-habitat relationships

Erica F. Stuber
University of Nebraska, Lincoln, efstuber@gmail.com

Lutz F. Gruber
Nebraska Cooperative Fish and Wildlife Research Unit

Joseph J. Fontaine
U.S. Geological Survey, Nebraska Cooperative Fish and Wildlife Research Unit, jfontaine2@unl.edu

Stuber, Erica F.; Gruber, Lutz F.; and Fontaine, Joseph J., "A Bayesian method for assessing multi-scale species-habitat relationships" (2017). *Nebraska Cooperative Fish & Wildlife Research Unit -- Staff Publications*. 240.
https://digitalcommons.unl.edu/ncfwrustaff/240

Landscape Ecol (2017) 32:2365–2381
https://doi.org/10.1007/s10980-017-0575-y

RESEARCH ARTICLE

A Bayesian method for assessing multi-scale species-habitat relationships

Erica F. Stuber · Lutz F. Gruber · Joseph J. Fontaine

Received: 9 December 2016 / Accepted: 15 September 2017 / Published online: 3 October 2017

Abstract

Context Scientists face several theoretical and methodological challenges in appropriately describing fundamental wildlife-habitat relationships in models. The spatial scales of habitat relationships are often unknown, and are expected to follow a multi-scale hierarchy. Typical frequentist or information theoretic approaches often suffer under collinearity in multi-scale studies, fail to converge when models are complex or represent an intractable computational burden when candidate model sets are large.

Objectives Our objective was to implement an automated, Bayesian method for inference on the spatial scales of habitat variables that best predict animal abundance.

Methods We introduce Bayesian latent indicator scale selection (BLISS), a Bayesian method to select spatial scales of predictors using latent scale indicator variables that are estimated with reversible-jump Markov chain Monte Carlo sampling. BLISS does not suffer from collinearity, and substantially reduces computation time of studies. We present a simulation study to validate our method and apply our method to a case-study of land cover predictors for ring-necked pheasant *(Phasianus colchicus)* abundance in Nebraska, USA.

Results Our method returns accurate descriptions of the explanatory power of multiple spatial scales, and unbiased and precise parameter estimates under commonly encountered data limitations including spatial scale autocorrelation, effect size, and sample size. BLISS outperforms commonly used model selection methods including stepwise and AIC, and reduces runtime by 90%.

Conclusions Given the pervasiveness of scale-dependency in ecology, and the implications of mismatches between the scales of analyses and ecological processes, identifying the spatial scales over which species are integrating habitat information is an important step in understanding species-habitat relationships. BLISS is a widely applicable method for identifying important spatial scales, propagating scale uncertainty, and testing hypotheses of scaling relationships.

Keywords Abundance · Bayesian model selection · Habitat selection · Model uncertainty · Spatial scale

---

Electronic supplementary material The online version of this article (doi:10.1007/s10980-017-0575-y) contains supplementary material, which is available to authorized users.

E. F. Stuber (✉) · L. F. Gruber
Nebraska Cooperative Fish and Wildlife Research Unit,
School of Natural Resources, University of Nebraska-Lincoln, Lincoln, Nebraska 68583, USA
e-mail: efstuber@gmail.com

J. J. Fontaine
U.S. Geological Survey Nebraska Cooperative Fish and Wildlife Research Unit, School of Natural Resources,
University of Nebraska-Lincoln, Lincoln,
Nebraska 68583, USA

Introduction

Species distribution models (SDMs) are increasingly identified as a powerful tool for guiding conservation planning (Guisan and Thuiller 2005). By associating habitat variables, generally land-cover imagery, with species occurrence or abundance, SDMs allow managers and policy makers to explore how large scale conservation challenges such as climate change may shape future biological communities (Schwartz et al. 2006). One of the fundamental challenges in the development of SDMs is that the ecological processes that shape the niche of a species act at multiple spatial scales (Levin and Simon 1992; Boyce and Mark 2006; Sandel and Smith 2009; Chase and Jonathan 2011). Thus, habitat attributes that may predict where a species is found at one spatial scale may perform poorly at predicting species occurrence when considered at another spatial scale (Robinson 1950). Mismatches between the spatial scale of analysis and the spatial scale of ecological processes underlying the distribution of a species can contribute to incorrect inference and overall poor model performance (Henebry 1995; Keitt et al. 2002). Indeed, issues of spatial autocorrelation, biases in regression coefficients (Lennon 2000; Kuhn 2007; Bini et al. 2009), limited explanatory power (De Knegt et al. 2010), and Type 1 error (Anselin and Griffith 1988; Clifford et al. 1989; Legendre et al. 2002) are all exacerbated when mismatches of scale occur. Researchers must also be cautious when altering scale-related dimensions of a study, such as grain or extent during analysis, as it may impact statistical results (Turner et al. 1989; Dungan et al. 2002).

Despite an awareness of the issues associated with spatial scale mismatches in ecology (Horne and Schneider 1995; Chalfoun and Martin 2007; Hurlbert and Jetz 2007), identifying, incorporating, and quantifying influential processes at various scales in ecological modeling has proven difficult and is often ignored or over-simplified (Wheatley and Johnson 2009; Jackson and Fahrig 2015). Although there is presumably not a single 'best' spatial scale at which to associate species and their habitats for all species or ecological conditions, it is possible to reveal relevant ecological associations when the scale of analysis approaches the operational scale of the process of interest. Studies often use expert knowledge to set the scale of analysis, or in some cases use expert-informed designs to collect data at 'fine-' versus 'large-' spatial scales to compare their relative explanatory power (e.g., Saab 1999; Coppeto et al. 2006; Cunningham et al. 2014). Unfortunately, information regarding the ecological processes shaping the distribution or abundance of a species is often lacking, let alone the spatial scale at which these processes may act. To overcome such limitations of knowledge researchers can collect data on habitat conditions at multiple spatial scales and perform model selection to choose a single 'characteristic' scale of their species' response (e.g., Būhning-Gaese and Katrin 1997; Pope et al. 2000; van Langevelde 2000; Steffan-Dewenter et al. 2002; Holland et al. 2004; Gray et al. 2010; Thornton and Fletcher 2014). While potentially informative, the notion of a single 'characteristic' scale to which a species may associate fails to consider that, in general, the processes that define a species' niche often interact, with the possibility of multiple explanatory variables acting at different spatial scales (e.g., Cushman and McGarigal 2002). However, few studies evaluate the explanatory performance of different variables collected at multiple scales simultaneously (e.g., Grand and Cushman 2003; Pearson et al. 2004; Jorgensen et al. 2014), but such studies have the ability to provide insights into the hierarchical nature of spatial processes (Urban et al. 1987).

We conducted a literature review of manuscripts investigating multi-scale habitat-relationships from the time period 2013–2017 (see Supplementary Material for details). Our search identified 128 empirical articles aiming to identify the spatial scale at which to study ecological patterns (e.g., abundance, occupancy, resource-selection). Of these, 92 and 8% were conducted in frequentist and Bayesian frameworks, respectively, and were dominated by an information-theoretic approach to model selection (71% information theoretic, 16% fully frequentist (e.g., selection based on p-value or strength of coefficient), 7% model fit (e.g., variance explained), 5.5% combination of methods, 0.8% fully Bayesian (e.g., posterior probability)). Particularly, Bayesian techniques are common in studies that account for imperfect detection of subjects because Bayesian hierarchical models have the flexibility to accommodate highly complex, custom, multilevel model specifications, and multiple levels of error that are not possible using currently available frequentist tools. Indeed, this is reflected in our literature search where only 13% of multi-scale studies conducted in a frequentist framework accounted for imperfect detection, compared to 50% of Bayesian multi-scale studies, and compared to an overall average 35% of articles accounting for detection based on a separate literature review (including uni- and multi-scale studies; Kellner and Swihart 2014). Although there are numerous Bayesian methods for model selection (Hooten and Hobbs 2015) that have been making inroads into ecological methodology relatively recently, we hypothesize that the disproportionate underrepresentation of Bayesian techniques in multi-scale studies is due to the lack of an accessible method of model selection for spatial scales in a Bayesian framework.

We develop a method of Bayesian latent indicator scale selection (BLISS), estimated with reversible-jump Markov chain Monte Carlo (MCMC) to provide a fully-integrated model-based fitting and selection procedure for selecting important spatial scales and estimating the effects of predictors in a single step. Although the use of indicator variables for model selection is well-established in Bayesian literature (see Godsill 2001; O'Hara and Sillanpää 2009; Tenan et al. 2014), extensions for use in spatial scale determination in ecology have not been implemented, regardless of the efficiency of automated exploration of the candidate model space. There are practical reasons for employing Bayesian methods for scale-selection beyond any philosophical reasons for preferring a fully Bayesian approach. Although information criteria including AIC, BIC, and DIC are heavily used, they are not appropriate for hierarchical (AIC, BIC) or mixture (DIC) models which are increasingly used in ecological literature, particularly to combat issues associated with imperfect detection (Hooten and Hobbs 2015). Bayesian methods allow a priori weighting of covariates or models (e.g., when prior information is available), which is not possible with many other approaches, including AIC. Furthermore, given equal priors, it is simple to calculate Bayes factors, to use in model averaging, and model probabilities for model selection (Johnson and Hoeting 2011). In our literature review, 77% of studies created models separately by scale, including predictors measured at each particular scale. Assuming that the ecological processes governing predictors are limited to acting at a single ecological scale is likely an oversimplification. BLISS is able to reveal multimodal scale patterns across predictors, and interactions between predictors at multiple scales. Both cases are either impossible or intractable to efficiently explore with frequentist methods, but trivial to implement using our method once a candidate set of scales is defined. For example, in a study investigating four predictors at four spatial scales, BLISS would require coding and estimating only a single model whereas frequentist or information theoretic approaches would require the researcher to code and estimate all 256 possible models individually to achieve the same flexibility in selecting spatial scales of predictor variables. The number of models required to represent all possible spatial scale combinations increases exponentially with the addition of predictors or scales. Furthermore, common selection rules such as 'choose the scale with the lowest AIC', would ignore the possibility of multiple important scales, and would not provide an interpretable quantification of estimation uncertainty and the likelihoods of alternative model specifications, both of which are efficiently revealed with BLISS. We first describe our Bayesian latent indicator variable approach. Next, we demonstrate its efficacy through simulations, and illustrate its efficiency with comparisons to other approaches. Finally, we apply our approach to a case study of ring-necked pheasants in Nebraska, USA.

Methods

Hierarchical abundance-detection model

Single scale model

The single scale model is based on an N-mixture model for estimating abundance under repeated measurements and probabilistic detection (Royle 2004). True abundance at site s and season t, $N_{s,t}$, is treated as an unobserved (latent) variable, which follows a Poisson distribution with mean parameter $\lambda_{s,t}$. Conditional on abundance $N_{s,t}$, the observed detections at site s, during season t, and replication r, $d_{s,t,r}$, follow a binomial distribution with $N_{s,t}$ trials and detection probability $p_{s,t,r}$:

$$N_{s,t} \sim \mathrm{Poi}(\lambda_{s,t}) \quad (1)$$

$$d_{s,t,r}|N_{s,t} \sim \mathrm{Bin}(N_{s,t}, p_{s,t,r}). \tag{2}$$

The abundance parameter $\lambda_{s,t}$ and detection parameter $p_{s,t,r}$ are functionally dependent on a set of environmental or other predictors and their corresponding regression coefficients;

$$\lambda_{s,t} = \exp(\alpha_1 x_{s,t,1} + \alpha_2 x_{s,t,2} + \cdots + \alpha_k x_{s,t,k}) \tag{3}$$

$$p_{s,t,r} = \mathrm{logit}^{-1}(\beta_1 z_{s,t,r,1} + \beta_2 z_{s,t,r,2} + \cdots + \beta_l z_{s,t,r,l}), \tag{4}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)'$ represent coefficients of abundance predictors and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_l)'$ represent coefficients of detection predictors; potential random effects are omitted for notational simplicity.

Given the observed detections $d_{s,t,r}$ for all sites, seasons, and replicates, as well as the corresponding abundance predictors (e.g., landcover variables) $\mathbf{x}_{s,t} = (x_{s,t,1}, \ldots, x_{s,t,k})'$ and detection predictors (e.g., time of day) $\mathbf{z}_{s,t,r} = (z_{s,t,r,1}, \ldots, z_{s,t,r,l})'$, the unknown (latent) abundance $N_{s,t}$ and regression coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are estimated.

Incorporating multiple scales

We developed an extension from the single scale model to handle multiple candidate scales for each abundance predictor where each abundance predictor $x_{s,t,i}$, $i \in \{1, \ldots, k\}$, is measured on one or more spatial scales $\mathbf{sc}_i \in \{sc_{i,1}, \ldots, sc_{i,n_i}\}$. We model the scale of each predictor as a latent categorical variable, estimated along with the other unknown quantities of the single scale model.

Priors for regression coefficients and scales Throughout, we use weak, normally distributed priors for the coefficients α_i, $i \in \{1, \ldots, k\}$, and β_j, $j \in \{1, \ldots, l\}$ with large prior variances $\sigma^2_{\alpha_i}$ and $\sigma^2_{\beta_j}$, respectively.

Each predictor $x_{s,t,i}$, $i \in \{1, \ldots, k\}$ can be included with a different number of scales $n_i < \infty$ as defined by the study. We use a categorical distribution over all candidate scales $\mathbf{sc}_i \in \{sc_{i,1}, \ldots, sc_{i,n_i}\}$ with prior probabilities $w_{i,1}, \ldots, w_{i,n_i}$ such that $\sum_{m=1}^{n_i} w_{i,m} = 1$ and guarantee the existence of a posterior distribution. The discrete uniform distribution is contained as a special case where all weights are equal $(w_{i,m} = 1/n_i)$, and represents a proper non-informative distribution for spatial scales.

Our priors are summarized by:

$$\alpha_i \sim N(0, \sigma^2_{\alpha_i}) \text{ for all } i = 1, \ldots, k \tag{5}$$

$$\beta_j \sim N(0, \sigma^2_{\beta_j}) \text{ for all } j = 1, \ldots, l \tag{6}$$

$$\mathbf{sc}_i \sim \mathrm{cat}(w_{i,1}, \ldots, w_{i,n_i}) \text{ for all } i = 1, \ldots, k. \tag{7}$$

Ecological model The ecological model incorporating multiple spatial scales differs from the traditional single scale model in that the mean abundance parameters $\lambda_{s,t}$ depend on the (latent) scales $\mathbf{sc}_i$, $i \in \{1, \ldots, k\}$, of the predictors. Below, $x_{s,t,i}[\mathbf{sc}_i]$ denotes the value of ith abundance predictor at site s, time t, and measured on scale $\mathbf{sc}_i$. Because we use a separate categorical scale variable for each predictor, a posterior distribution for candidate spatial scales is produced for each predictor, which allows predictor-specific selection of the most informative spatial scales. The choice of scales $\mathbf{sc}_i$, $i \in \{1, \ldots, k\}$, propagates to the latent abundance variable $N_{s,t}$ via its mean parameter $\lambda_{s,t}$, resulting in

$$\lambda_{s,t}|\alpha_1, \ldots, \alpha_k, \mathbf{sc}_1, \ldots, \mathbf{sc}_k = \exp(\alpha_1 x_{s,t,1}[\mathbf{sc}_1] + \cdots + \alpha_k x_{s,t,k}[\mathbf{sc}_k]) \tag{8}$$

$$N_{s,t}|\lambda_{s,t} \sim \mathrm{Poi}(\lambda_{s,t}) \tag{9}$$

By only including one scale of each predictor at a time, collinearity among the different scales of the predictors is excluded. As a result, this model design avoids the need for regularization approaches used in the presence of correlated predictor variables to eliminate predictors with little marginal explanatory power. Regularization approaches are often computationally intensive and require additional input from the investigator in specifying regulator parameters (Hooten and Hobbs 2015).

Detection model The number of detected individuals $d_{s,t,r}$ are conditional on the latent abundance variables $N_{s,t}$, which, when incorporating multiple scales, depend on the scale indicators $\mathbf{sc}_i$, $i \in \{1, \ldots, k\}$, and thus are inherently linked to the scales of the abundance process.

$$p_{s,t,r}|\beta_1,\ldots,\beta_l = \text{logit}^{-1}(\beta_1 z_{s,t,r,1} + \cdots + \beta_l z_{s,t,r,l}) \tag{10}$$

$$d_{s,t,r}|N_{s,t}, p_{s,t,r} \sim \text{Bin}(N_{s,t}, p_{s,t,r}) \tag{11}$$

Posterior estimation using reversible jump MCMC

All models were estimated via posterior simulation with JAGS ("just another Gibbs sampler;" Plummer 2003) (package rjags; Plummer 2013); for example code, see Supplementary Material.

Simulation studies

To assess the performance of BLISS, we simulated multi-season abundance data, where abundance and detection data were generated as a function of three site-specific abundance covariates measured at four spatial scales with season-specific intercept terms. Season-specific intercepts imply that the population is assumed closed within seasons, but open between seasons. Abundance covariates were drawn from a uniform distribution on the interval (0, 1), and we investigated three scenarios with different strengths of correlations between different scales of the same predictor: low (correlation 0.3 between the four scales), medium (0.60), and high (0.90) (Table 1 Studies 1.1, 1.2, and 1.3). The true spatial scales for the three abundance covariates were chosen as 2, 3, and 4, respectively. The abundance coefficients were normally distributed about their means (SD) 1.25 (0.20), 0.50 (0.05), and − 0.80 (0.10), respectively. To assess the estimation uncertainty of BLISS under varying sample sizes we simulated data with low (2), medium (3), and high (6) within-season temporal replication (Table 1 Studies 2.1, 2.2, and 2.3), and low (50), medium (100), and high (200) within-season spatial replication (Table 1 Studies 3.1, 3.2, and 3.3). Within these simulation studies, we also evaluated the impact of the strength of the abundance covariates on scale selection performance.

Site-specific detection probabilities were simulated as a function of four covariates drawn independently from a standard normal distribution combined with normally distributed coefficients (means (SD) − 0.50 (0.10), 1.25 (0.20), 1.00 (1.00), − 0.90 (0.30)), one intercept term and an observer-level random effect. We modeled detection for a hypothetical 3-season study with multiple replicate abundance counts within each season (e.g., the 'robust design' following Williams et al. 2002), assuming that the population is closed during each of 3 simulated seasons.

For each of the simulation studies, we generated 100 data sets, and analyzed each using the model incorporating multiple scales and RJMCMC in JAGS. We present results based on 20,000 MCMC samples after burn-in (burn-in = 5000 iterations) of the posterior distribution of model parameters (coefficients) and latent variables (scales, abundance, random effects). For each predictor, we analyze scale selection accuracy and 'select' the scale with the highest posterior probability as the 'best' scale. This intuitive decision rule minimizes the probability of mis-specification under the margins of the posterior distribution. We summarized coefficient and abundance estimation for datasets by reporting the posterior mean or mode (N) estimates along with their 95%

Table 1 Conditions for simulation studies of scale autocorrelation (1.1, 1.2, 1.3), within-season temporal replication (2.1, 2.2, 2.3), and spatial replication (3.1, 3.2, 3.3)

Study	Scale Autocorrelation	Temporal Replicates	Spatial Replicates	Seasons	Total Observations
1.1	0.3	3	200	3	1800
1.2	0.6	3	200	3	1800
1.3	0.9	3	200	3	1800
2.1	0.6	2	200	3	1200
2.2	0.6	3	200	3	1800
2.3	0.6	6	200	3	3600
3.1	0.6	3	50	3	450
3.2	0.6	3	100	3	900
3.3	0.6	3	200	3	1800

credible intervals (CI) and evaluated model performance by comparing the estimated site and season-specific abundance estimates with the true simulated values by calculating the root mean squared error (RMSE) of the posterior mode estimates of $N_{s,t}$.

Comparison with existing methods

We compared the performance of BLISS to other methods commonly found in the multi-scale literature using simulated datasets. Additionally, because the Watanabe-Akaike or widely applicable information criterion (WAIC), which represents the posterior predictive likelihood minus an adjustment for the effective number of parameters, is a suitable information criterion for hierarchical models, but is not yet widely implemented, we include an evaluation of its performance here (Watanabe 2013 and Gelman et al. 2014)). We simulated 25 replicate datasets each under two scenarios of between-predictor correlation (e.g., collinear predictors; 0 correlation, and 0.5 between-predictor correlation) and the characteristics of simulation study 1.1 (Table 1). We analyzed each replicate dataset using five model selection techniques to select the spatial scale of each predictor: BLISS, backward stepwise selection, effect size selection, AIC model selection, and WAIC model selection. BLISS proceeded as previously described. Backward stepwise selection was performed by fitting models first containing all scales of a single predictor at once (e.g., three full models, one for each predictor) and dropping the least important scale based on the largest 'p-value' calculated using a normal approximation of the posterior distribution until only a single scale was left for each predictor. Coefficient effect size-based model selection was performed by estimating models containing all predictors separately by scale (i.e., each model contained all predictors measured at the same single scale) and selecting the scale with the strongest estimated coefficient for each predictor individually (e.g., Jorgensen et al. 2014). Information criterion-based model selection was performed by calculating AIC and WAIC for independent predictor by scale models (i.e., each model contained only 1 predictor measured at 1 scale for 12 total models) and selecting the scale model with the lowest AIC (Akaike and Hirotogu 1998) or WAIC (Watanabe 2013; Gelman et al. 2014) for each predictor (e.g., Kirol et al. 2015).

Study of the ring-necked pheasant *(Phasianus colchicus)* population in Nebraska

We demonstrate the applicability of our approach to species count data collected following a typical field design where data collection was initially conducted to estimate abundance and habitat relationships for creating predictive abundance maps. During April–July of 2016, we conducted multi-species 500 m fixed-radius aural point count surveys (Robbins et al. 1986) at roadside sites (Mccarthy et al. 2012) located across the State of Nebraska. Spatially balanced survey sites were selected to reflect the background landcover composition such that predictors of interest were sampled along a gradient from low to high proportions in amounts found within the State. We conducted surveys from 15 min before sunrise until approximately 10 a.m., when detection rates are highest and most consistent across species (Hutto et al. 1986), and recorded every individual seen or heard during a three-minute period. Additionally, during each survey we recorded the start time, temperature, percentage cloud cover, and windspeed, as these may affect our ability to detect individuals that are present at each location. We did not conduct surveys during inclement weather including high fog, prolonged rain, or in winds exceeding 20 km/h. Observers conducted between 10 and 15 surveys per day, and each site was sampled four times (replicates) during the breeding season unless inclement weather precluded surveying.

Landcover variables were derived from the 2010 Rainwater Basin Joint Venture Nebraska Landcover layer generalized into six classes expected to be important to pheasants (Conservation Reserve Program grasses, grass, row crops, small grains, trees, and wetlands, 95% overall accuracy; Bishop et al. 2011) Based on these landcover data, we derived the proportion of each habitat type surrounding survey points at seven different buffer sizes: 500 m, 1, 2, 3, 4, 5, and 10 km radii.

Landcover variables (each measured at seven spatial scales) were entered into our model as covariates (centered on the mean) for pheasant abundance, including their quadratic effects, while time of day, date, temperature, cloud cover, and windspeed were included as covariates in the detection process of the model along with a random effect for observer. BLISS and coefficient estimation was based on 75,000 iterations after 25,000 iterations of burn-in.

Results

Autocorrelation of predictors across spatial scales

Scale selection

Our simulation results demonstrated that when predictor scale autocorrelation was low ($\rho = 0.30$), BLISS had the highest success in identifying the correct spatial scales, compared with medium and high autocorrelation (average posterior probability of true scale = 0.99; range 0.56–1.00). Of 100 simulated datasets investigating the best-fitting scales of three predictor variables BLISS did not mis-select any incorrect scales. As scale autocorrelation increased to 0.6 (medium correlation) BLISS continued to perform well, correctly identifying 100% of scales from 100 simulated datasets. Similarly, the average posterior probability of the true scales remained high 0.99 (range 0.89–1.00) in datasets with medium scale autocorrelation. The percentage correctly selected scales began to decrease under high scale autocorrelation (12%) and average posterior probability for the true scales declined to 0.87 (range 0.01–1.00).

Covariate and abundance estimation

Posterior distributions of parameter estimates produced by BLISS always included the true coefficient value. Regardless of the level of scale autocorrelation in the predictors, our mean coefficient estimates were within 0.05 units of the true simulated values (Tables 2, 3, 4). RMSE was low, and consistent over the simulated levels of autocorrelation. BLISS could predict site-specific abundance within 1 simulated individual of the truth in approximately 79% of cases, and in 93% of cases predicted abundance within two individuals irrespective of scale autocorrelation in the predictors.

Temporal and spatial sample size

Scale selection

Scale selection was successful for low, medium, and high within-season replicate simulations. At medium- and low-replication, BLISS only misidentified the true scale in 0, and three instances (1%), respectively. With 6 within-season replicates, BLISS only misidentified the true scale in two instances (1%). Posterior probability of the true scale under high within-season replication was, on average, 0.99 (range 0.15–1.00) and remained high when within-season replicates were decreased to 3 (average: 0.99; range 0.89–1.00), and 2 (average: 0.98; range 0.13–1.00).

At high spatial replication, BLISS did not mis-select any scales, and the posterior probability of the true scale was high (average: 0.99; range 0.89–1.00). Scale selection success slightly declined as spatial replicates were lowered to 2 and 12% mis-selected under medium, and low replicates, respectively. Posterior probability in the true scale was high (average: 0.97; range 0.11–1.00) with medium replicates, but declined to 0.83, on average, (range 0.01–1.00) with only 50 sites sampled.

Covariate and abundance estimation

BLISS returned accurate estimates of predictor coefficients regardless of the number of within-season replicates. Furthermore, the precision of our estimates improved with the addition of within-season replicates, as increasing replicates from 3 to 6 narrowed the credible intervals of coefficient estimates by 0.03 (95% CI 0.01, 0.05). Average RMSE of abundance decreased as we increased our simulations to 6 replicates per season decreasing by 0.43 compared to 3 replicates (95% CI 0.14, 0.67) and 0.70 compared with 2 replicates (95% CI 0.44, 0.93). Similarly, increasing within-season replication improved BLISS's accuracy in predicting site-specific abundance. At low replication, predicted abundance was within 1 individual of the truth in 73% of cases, which increased to 90% at high replication, while predicted abundance within 2 individuals of the truth increased from 89 to 98% of cases.

Bias remained low in estimates of abundance coefficients across levels of spatial replication (Table 4); the averaged credible intervals across all 100 replications included the true coefficient values for all coefficients and scenarios. However, as sample size decreased, the precision of our estimates decreased significantly for all predictors. Decreasing the number of study sites from 200 to 100 widened the 95% credible intervals of coefficient estimates by 0.11, on average (95% CI 0.08, 0.16). Further decreasing sample size to 50 sites surveyed widened

Table 2 Results of simulation studies for scale autocorrelation (AC)

	0.3 AC	0.6 AC	0.9 AC	TRUE
Coefficient 1 (α_1)	1.29	1.27	1.30	1.25
	(1.14, 1.44)	(1.12, 1.42)	(1.15, 1.44)	(0.86, 1.64)
Coefficient 2 (α_2)	0.49	0.50	0.50	0.50
	(0.35, 0.64)	(0.36, 0.65)	(0.35, 0.64)	(0.40, 0.60)
Coefficient 3 (α_3)	− 0.79	− 0.80	− 0.82	− 0.80
	(− 0.94, − 0.64)	(− 0.95, − 0.65)	(− 0.97, − 0.67)	(− 1.00, − 0.60)
Scale 1 ($\mathbf{sc}_1$)	100/100	100/100	100/100	–
Scale 2 ($\mathbf{sc}_2$)	100/100	100/100	63/100	–
Scale 3 ($\mathbf{sc}_3$)	100/100	100/100	100/100	–
Total	300/300	300/300	263/300	–
RMSE	1.38	1.38	1.45	–
%Abundance ± 1	77	78	77	–
%Abundance ± 2	92	92	91	–

Estimated coefficients (mean of posterior distribution) and 95% credible intervals (2.5 and 97.5% quantiles of the posterior distribution) of the abundance predictors $\alpha_1, \alpha_2, \alpha_3$; root mean squared error; number of replications (out of 100) that correctly selected the scales $\mathbf{sc}_1, \mathbf{sc}_2, \mathbf{sc}_3$; and percentage of abundance posterior mode estimates within 1 or 2 of the true (simulated) abundance

Table 3 Results of simulation studies for within-season temporal replication

	2 replications	3 replications	6 replications	TRUE
Coefficient 1 (α_1)	1.28	1.27	1.31	1.25
	(1.10, 1.45)	(1.12, 1.42)	(1.18, 1.45)	(0.86, 1.64)
Coefficient 2 (α_2)	0.51	0.50	0.51	0.50
	(0.35, 0.66)	(0.36, 0.65)	(0.38, 0.64)	(0.40, 0.60)
Coefficient 3 (α_3)	− 0.79	− 0.80	− 0.80	− 0.80
	(− 0.94, − 0.63)	(− 0.95, − 0.65)	(− 0.94, − 0.67)	(− 1.00, − 0.60)
Scale 1 ($\mathbf{sc}_1$)	100/100	100/100	100/100	–
Scale 2 ($\mathbf{sc}_2$)	97/100	100/100	98/100	–
Scale 3 ($\mathbf{sc}_3$)	100/100	100/100	100/100	–
Total	297/300	300/300	298/300	–
RMSE	1.65	1.38	0.94	–
%Abundance ± 1	70	78	89	–
%Abundance ± 2	88	92	97	–

Estimated coefficients (mean of posterior distribution) and 95% credible intervals (2.5 and 97.5% quantiles of the posterior distribution) of the abundance predictors $\alpha_1, \alpha_2, \alpha_3$; root mean squared error; number of replications (out of 100) that correctly selected the scales $\mathbf{sc}_1, \mathbf{sc}_2, \mathbf{sc}_3$; and percentage of abundance posterior mode estimates within 1 or 2 of the true (simulated) abundance

the 95% credible intervals by 0.25, on average (95% CI 0.17, 0.35). Nevertheless, RMSE of these models were consistently low irrespective of spatial replicate sample size. Regardless of spatial sample size, BLISS could predict site-specific abundance within one simulated individual of the truth in approximately 80% of cases, and in 92% of cases predicted abundance within two individuals (Fig. 1).

Effect size of abundance predictors and scale selection

Scale selection

The posterior probability of the true scale was not different at low or medium spatial scale autocorrelation for the three predictors, which all had different

Table 4 Results of simulation studies for spatial replication

	50 sites	100 sites	200 sites	TRUE
Coefficient 1 (α_1)	1.33	1.30	1.27	1.25
	(1.01, 1.66)	(1.09, 1.51)	(1.12, 1.42)	(0.86, 1.64)
Coefficient 2 (α_2)	0.49	0.53	0.50	0.50
	(0.15, 0.82)	(0.32, 0.73)	(0.36, 0.65)	(0.40, 0.60)
Coefficient 3 (α_3)	− 0.79	− 0.79	− 0.80	− 0.80
	(− 1.13, − 0.45)	(− 0.99, − 0.59)	(− 0.95, − 0.65)	(− 1.00, − 0.60)
Scale 1 ($\mathbf{sc}_1$)	99/100	100/100	100/100	–
Scale 2 ($\mathbf{sc}_2$)	72/100	96/100	100/100	–
Scale 3 ($\mathbf{sc}_3$)	93/100	97/100	100/100	–
Total	264/300	293/300	300/300	–
RMSE	1.34	1.37	1.38	–
%Abundance ± 1	79	78	78	–
%Abundance ± 2	93	92	92	–

Estimated coefficients (mean of posterior distribution) and 95% credible intervals (2.5 and 97.5% quantiles of the posterior distribution) of the abundance predictors $\alpha_1, \alpha_2, \alpha_3$; root mean squared error; number of replications (out of 100) that correctly selected the scales $\mathbf{sc}_1, \mathbf{sc}_2, \mathbf{sc}_3$; and percentage of abundance posterior mode estimates within 1 or 2 of the true (simulated) abundance

size coefficients (95% CI of the difference in probability was not different from zero). Under high simulated scale autocorrelation between spatial scales, however, the predictor with the weakest coefficient had significantly lower posterior probability in the correct scale than the predictors with bigger coefficients (mean difference: 0.39; 95% CI 0.00, 0.99). Posterior probabilities in the true spatial scales were not impacted by coefficient effect sizes when varying the amount of temporal replication in datasets (mean differences always overlapped zero). In our spatial replication study, the predictor with the weakest coefficient had significantly lower posterior probability in the true scale compared with the strongest coefficient when the amount of spatial replication was 50 and 100 sites (mean difference 50 sites: 0.40; 95% CI − 0.02, 0.96; mean difference 100 sites: 0.07; 95% CI 0.00, 0.66).

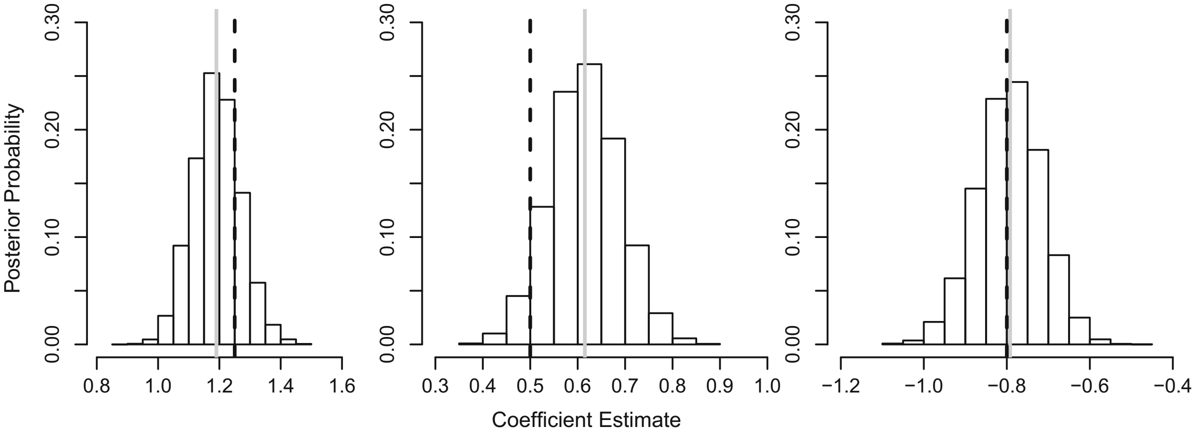

Fig. 1 Estimated posterior distribution of abundance coefficients for three abundance covariates (left, middle, and right columns) of a representative simulated dataset from study 1.1. The solid line indicates the mean of the posterior distribution and the dashed line indicates the mean of the true (simulated) value. Posterior distributions of scale selection were relatively unambiguous, and trace plots for abundance coefficient estimates were well-mixed

Table 5 Results of simulation studies comparing the spatial scale selection accuracy of four alternative model selection techniques and BLISS under no between-predictor correlation, and with moderate between-predictor correlation ($\rho = 0.50$). Values indicate the percentage of simulated datasets where the correct spatial scale was identified for each predictor variable. Predictor coefficients were generated following the parameters described in the simulation study with the mean coefficients $\alpha_1 = 1.25$, $\alpha_2 = 0.50$, and $\alpha_3 = -0.80$

	Between-predictor correlation $\rho = 0.0$			Between-predictor correlation $\rho = 0.50$		
	Predictor 1	Predictor 2	Predictor 3	Predictor 1	Predictor 2	Predictor 3
BLISS	100	100	100	100	100	100
Coefficient strength	100	100	100	100	92	100
WAIC	100	100	100	100	60	88
Backward stepwise	100	100	100	100	36	92
AIC	80	76	76	52	48	56

Comparison with existing methods

In our methods comparison simulation study, the BLISS method was 100% accurate in selecting the best spatial scales of predictors, out-performing all other methods compared here (Table 5). Although selection by backward stepwise elimination, coefficient strength, and WAIC accurately selected the spatial scales of predictors in all cases of the uncorrelated predictor scenario, their performance degraded under conditions of predictor collinearity (Table 4; $\rho = 0.50$). Model selection by AIC, the most commonly used method in current multi-scale studies, displayed the worst performance under both the uncorrelated and correlated predictor scenarios. Excluding the BLISS method, which had 100% accuracy, methods for model selection of spatial scales performed, on average, worst at selecting the spatial scale of the predictor with the smallest coefficient (predictor 2) and performed the best at selecting the spatial scale for the predictor with the strongest coefficient size (predictor 1).

Application: scale selection for ring-necked pheasant abundance in Nebraska

During one breeding season we completed 1464 point counts at 456 survey sites. Compared with conditions explored in our simulation studies, autocorrelation between the multiple spatial scales of landcover variables was high (> 0.85), temporal replication was medium (3–4 replicates per site), and spatial replication was high (> 200 sites).

Scale selection

Using BLISS, we identified 4 predictors that were clearly supported at well-discriminated spatial scales in predicting pheasant abundance. Proportion of grass and trees were estimated to predict pheasant abundance at very local scales (1 km, and 500 m radius scales, respectively) with high posterior probability (0.66, 0.89, respectively), while CRP best explained pheasant abundance at the 2 km radius scale (posterior probability: 0.99), and the proportion of small grains acted on pheasant abundance at the largest tested scale (posterior probability: 0.88; 10 km radius). Although the posterior probability for the best single spatial scale was lower than for other well-distinguished predictors, our model indicated that the proportion of row crops at 1km best explained pheasant abundance (posterior probability: 0.41). A single scale could not be unambiguously selected for the proportion of wetlands; however, posterior probability was approximately normally distributed around the 4 km radius scale (posterior probability = 0.34), while the second most supported scale (3km) had only a marginally lower posterior probability (0.29) (Fig. 2).

Covariate and abundance estimation

The number of detected pheasants per site ranged from 0 to 9 individuals, and the average modeled abundance was 17 individuals (estimated intercept: Table 6). We included linear and quadratic terms in the linear predictor to allow the relationship between abundance and the predictor to change signs, which was supported for row crops, CRP, small grain, and grasses. Wetlands had a strictly positive effect on abundance; while trees had a strictly negative effect on abundance (Fig. 3).

Discussion

Organisms use information arising from multiple spatial scales which shape distribution and space-use decisions, but statistical methods for determining such scales are still not widely applied in ecological studies. Our studies have demonstrated that Bayesian hierarchical models can be highly effective for identifying the spatial scales that best explain animal abundance while also returning accurate predictor and abundance estimates. BLISS was tested under various ecologically relevant data limitations encountered in field-based studies (sample size, # replicates,

Table 6 Estimated coefficients (posterior mean) and their associated 95% credible intervals (CI) based on BLISS applied to point counts of ring-necked pheasants in Nebraska, USA

Coefficient	Posterior mean and 95% CI
Intercept	2.84 (2.21, 3.86)
Crops	0.65 (− 0.28, 1.53)
Crops^2	− 1.43 (− 3.51, 0.61)
CRP	4.96 (3.31, 6.69)
CRP^2	− 9.43 (− 16.97, − 3.74)
Grasses	− 1.04 (− 1.75, − 0.26)
Grasses^2	− 2.59 (− 4.44, − 0.37)
Small grains	3.20 (1.87, 4.43)
Small grains^2	− 4.01 (− 7.22, − 0.97)
Wetland	5.34 (− 0.17, 10.94)
Wetland^2	− 0.36 (− 19.00, 18.83)
Trees	− 3.79 (− 6.48, − 1.43)
Trees^2	− 1.54 (− 12.03, 8.52)

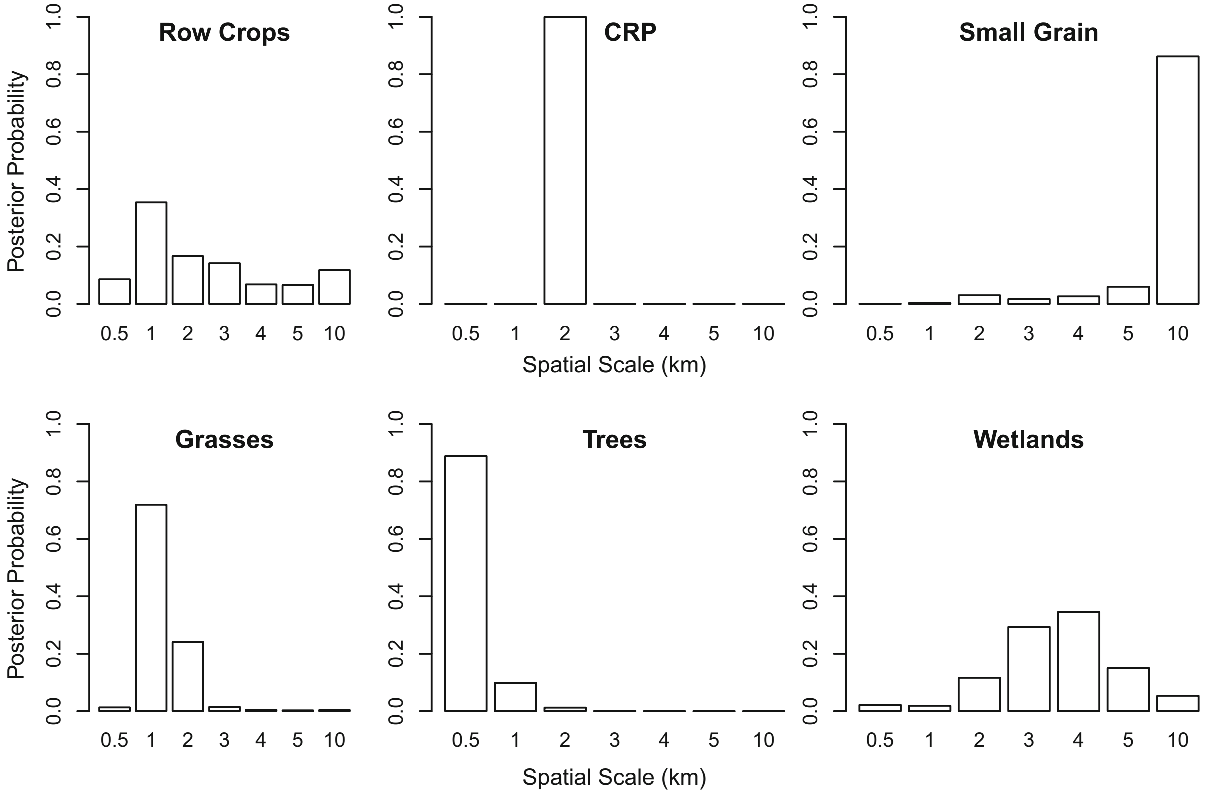

Fig. 2 Posterior distributions of the spatial scales (in km) of the land cover abundance predictors row crops, CRP, small grain, grasses, trees and wetlands for ring-necked pheasants breeding in Nebraska as estimated by BLISS

autocorrelation, confounding detection covariates). Unsurprisingly, scale selection accuracy approaches 100% with increasing sample size, and decreasing spatial autocorrelation between scales within predictors, as these represent ideal conditions. However, our simulations reveal the utility of BLISS even when study design and ecological conditions do not conform to this ideal. Our model-based method for describing spatial scale relationships is useful when information on the spatial scale of predictor relationships is uncertain, the aim of the study is model selection for scales, or for propagating scale uncertainty to subsequent analyses. However, insight gained from purely observational studies should be evaluated within the limits of correlational research.

Varying the degree of scale autocorrelation in our predictors introduced uncertainty in the identification of true scale relationships as correlations increased. This result is unsurprising given that neighboring scales will appear very similar at such high correlation and will explain the data reasonably well when substituted for the true scale. When neighboring scales are highly correlated, it would be necessary to consider whether a single ecological process governs the abundance-relationship at those scales (e.g., multiple scales are predicting the same ecological relationship), whether the species of interest can perceive a difference between the scales being considered (Nams 2005), and differences in efficiency of collecting data at those scales, to guide model-selection decisions or target further studies.

Strong correlation between different scales of the same predictor had no effect on BLISS's ability to retrieve accurate coefficient or abundance estimates. Unlike model selection procedures that evaluate models containing multiple scales of predictors simultaneously, BLISS does not suffer from collinearity within scales of predictor variables (Lennon 2000; Overmars et al. 2003). This is especially relevant when the research interest lies in estimating and interpreting sign and magnitude of coefficients, as the simultaneous inclusion of highly collinear predictors can lead to identification problems and severely biased estimates which can easily be avoided in our method.

BLISS had decreasing success in identifying scales when there was a relatively weak relationship (effect size) between the predictor and abundance; however, this was only detectable when scale autocorrelation

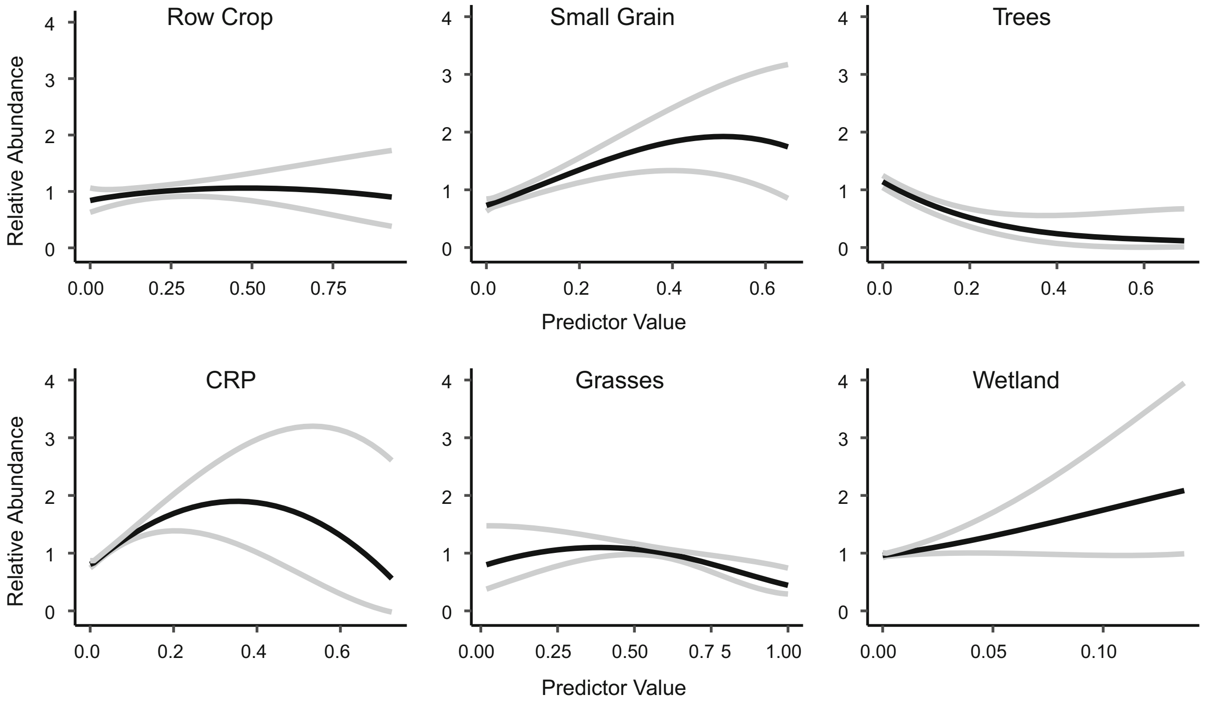

Fig. 3 Marginal effects (mean: black line; 95% CI: gray lines) of landcover predictors on relative pheasant abundance. The range of the x-axis represents values of landcover predictors measured as proportions sampled in Nebraska

was high, or spatial replication was medium or low, as variables that contribute little to the variation of the population mean $\lambda_{s,t}$ will only change the population mean and likelihood by small amounts. If the posterior mass is widely dispersed over several candidate scales, only limited confidence can be placed on any one single-scale model. Rather, uncertainty propagation, or model averaging should be employed, the former of which is straightforward in a Bayesian framework.

We did not see an effect of varying the number of within-season replicates on our ability to successfully identify scale relationships. This is fully expected in a closed population model, because there was only one set of predictors and abundance for each site and season, and additional detection data do not provide additional habitat information. We expect that extending the duration of study to multiple seasons would improve model performance, because this would provide additional unique habitat information. The effect of temporal replication manifested itself only in the accuracy of abundance estimates. While the accuracy and precision of coefficient estimates were consistent, there was a positive relationship between accuracy of site-specific abundance predictions and replication, resulting in lower RMSE with increased replication, similar to results presented in simulation studies of occupancy probabilities (Guillera-Arroita et al. 2014).

The number of field sites sampled seasonally had a small, but noticeable, impact on the success of spatial scale identification, as increasing the number of sites sampled adds additional data to the estimation of the ecological model. Furthermore, the results of this simulation demonstrate a significant increase in precision in predictor coefficient estimates with sample size, while bias remained low across simulations; additional sites provide additional data for the ecological model, which then yields a narrower high likelihood density area, and by extension, a narrower high posterior density (HPD) area.

In all simulation studies, some of the generated datasets performed poorly in the scale selection procedure (e.g., there was less than 50% of the posterior probability in the single simulated 'true' scale). This was noticeable primarily in simulation studies with high autocorrelation (0.90) between different scales of the predictors, or where sample size was low ($N = 50$ replicate sites). In most cases, particularly within the simulation study with low spatial replication, there was no detectable lack of convergence. However, we noticed that the posterior means of the abundance coefficients of each individual simulation replicate (out of the 100) were typically slightly biased, while the aggregated distributions of the 100 simulation posterior mean estimates were centered around the true coefficient means. This observation is indicative of the randomness contained in the simulated datasets and not a problem with the proposed estimation approach. When drawing a sample from a data generation process (DGP), each sample's moments and other characteristics are typically different from those of the DGP, but a single, noisy sample is the only data available on which to apply the estimation procedure. In small samples, noise is more pronounced, which leads to greater error in selection results. In a few cases, a lack of convergence in the MCMC samples was the likely driver of mis-selection. This situation occurred most frequently in scenarios with highly correlated variables (as in study 1.3). We re-ran a subset of these datasets, and in most of the re-runs, the MCMC output converged well and scale selection results were improved. This illustrates the need for researchers to perform algorithm-checking steps.

In two additional simulation studies, we demonstrated that BLISS outperforms other methods of model selection used to identify important spatial scales in ecological modeling. Not only did BLISS correctly identify the true scales of predictors in 100% of cases, BLISS required only estimating a single model (runtime (hh:mm): 00:13, on average 90% faster than other methods), compared with 9 models in backward stepwise (runtime: 02:21), four models in the coefficient strength method (runtime: 01:11), and 12 models in both AIC and WAIC selection (runtime: 02:52). To achieve the same flexibility in scale combinations, AIC and WAIC procedures would require estimating 64 models, representing all possible predictor by scale combinations (extrapolated runtime: 15:16). We expect that if all possible model combinations are estimated and selection is performed using AIC or WAIC, that the correct model would be identified, although at a substantial runtime cost. Although the use of AIC is known to be problematic for hierarchical or random effects models, it is disconcerting that AIC model selection performed worst, of the methods evaluated here, given that it is the most heavily used method in multi-scale studies. Indeed, while all methods excluding AIC could identify the true spatial scales when predictor variables were uncorrelated, AIC failed to select true scales both under moderate between-predictor correlation, and no correlation. Some degree of predictor collinearity is generally the rule in ecological studies, and particularly in investigations of species-habitat relationships where correlated environmental variables are hypothesized to predict abundance or occurrence at a location. Researchers often rely on rules of thumb (e.g., remove variables with $\rho > 0.5, 0.6, \text{or}\, 0.7$, or variable inflation factor greater than 10) when considering the impact of collinearity on model inference. Between-predictor correlations such as the one used in our simulation are not uncommon in field investigations. This highlights the necessity of simulation studies to identify under what conditions particular methods perform well, and further investigation into the performance properties of AIC model selection applied to scale-selection scenarios seems warranted. Selection based on estimated coefficient strength ranked second in performance after BLISS although it is not widely used. Also not commonly used, likely due to its recent introduction to ecological literature, WAIC was the third best model selection method considered here. WAIC is an appealing option for model selection as it can be interpreted much the same as AIC, which is commonly applied in ecological studies, and is a fully Bayesian method. However, because it is a relatively new information criterion, its properties are likely the least well studied and there is little guidance in calculating the metric in commonly used software. Because of their poor performance, we cannot recommend backward stepwise or AIC model selection in multi-scale studies without further investigation and validation of their use in similar studies.

We applied our method to a case study of ring-necked pheasants, investigating 6 predictors measured at seven spatial scales. Ring-necked pheasants are a culturally and economically important species in Nebraska that is heavily managed to maintain viable populations and provide adequate hunting opportunities. To provide the most effective habitat management to maintain or increase populations, wildlife managers must either acquire parcels of land with local and neighborhood characteristics that are conducive to supporting high pheasant abundance, or alter the environment to produce habitat that can sustain high pheasant abundance. Because it is rarely feasible to conduct habitat management at large spatial scales, our results suggest that managers might most efficiently focus their energy on altering row crops, CRP grasses, non-CRP grasses, or woodland, as pheasant abundance is associated with these landcover types at relatively small scales (e.g., 2 km radius or smaller). Contrariwise, managers are less likely to affect pheasant abundance through management of either small grains, or wetlands in Nebraska, as the relationship with these landcover types manifests at prohibitively large spatial scales. Using Bayesian model selection, we could determine important spatial scales by running a single model (runtime: 13 min). If we were to implement typical frequentist or information-theory model selection with the same flexibility in identifying scales, we would have had to estimate 117,649 separate models (representing all possible scale combinations; runtime 1062 days), an inefficient method of model selection given the available Bayesian alternative.

In a similar study of ring-necked pheasants in Nebraska, Jorgensen et al. (2014) created a mixed-scale model based on effect size selection from Bayesian hierarchical N-mixture models of predictors estimated separately at 2 scales (local management: 1 km radius, and landscape: 5 km radius). The authors chose to model predictors separately by scale because of autocorrelation, and avoided information criteria (i.e., AIC, BIC, DIC) because of their inappropriateness for hierarchical mixture models. Because of these issues, and because reversible-jump MCMC methods of model selection had not yet been introduced in the multi-scale ecology literature, Jorgensen et al. were unable to estimate optimal models. For example, the authors simplified model structure in the effect size selection procedure to ignore quadratic effects, which were expected and included in their final, mixed-scale model, and clearly supported in our analysis. Additionally, during model selection, the authors were limited to evaluating the effect sizes of predictors only when other predictors were included at the same scale. We demonstrate that nearly all landcover variables optimally explain pheasant abundance at different spatial scales, which was not possible to reflect with the previous analysis workflow. It is important to investigate a larger candidate model space when spatial scales are unknown a priori. Indeed, when allowed to select from a larger space, the optimal spatial scales of landcover predictors ranged from 0.5 to 10 km radii with only three landcover predictors with similar selected scales between the two methods (effect size selection/BLISS: CRP 1/2 km, Grasses 1/1 km, Wetlands 5/4 km).

Conclusions

We have clearly demonstrated the performance of BLISS for the selection of spatial scales and coefficient estimation of abundance predictors. The greatest utility of BLISS is providing investigators with an extremely efficient tool to explore candidate spatial scale space, coupled with its superior estimation accuracy. Current approaches to selecting the spatial scales of predictors are typically implemented under acknowledged suboptimal caveats because most familiar model selection methods were not developed to be used under conditions of collinearity, or with hierarchical, or mixture models, with serious implications for drawing appropriate inferences. However, with the development of analytical approaches, availability of software, and accessibility of data, hierarchical and mixture models are increasingly used to understand spatial ecological processes in uni-scale but not multi-scale investigations. Therefore, models such as BLISS, which overcome the drawbacks of typical model selection methods, are an important progression in ecology, allowing scientists to adequately conduct multi-scale studies. We think it is critical that studies relying on model selection techniques not developed for the implemented models of choice (e.g., AIC for selection in hierarchical models) first validate their performance under particular study designs and data structure. Only then can we be confident about results and inferences made from such analyses. BLISS has promise for substantial generalization and future research will include estimating species occurrence, distribution, and resource use, and in non-hierarchical settings or with different model error distributions.

Acknowledgements Funding for this project was received from Federal Aid in Wildlife Restoration projects W-98-R, administered by the Nebraska Game and Parks Commission. We would like to thank Chelsea Forehead, Caitlyn Gillespi, Anthony Jenniges, Amanda Lipinski, and Lindsey Messinger for their assistance in collecting the data presented here, Annie Madsen and Matthew Strassburg for assistance in conducting the literature review, and two anonymous reviewers for their valuable comments on earlier versions of this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The Nebraska Cooperative Fish and Wildlife Research Unit is supported by a cooperative agreement among the U.S. Geological Survey, the Nebraska Game and Parks Commission, the University of Nebraska, the U.S. Fish and Wildlife Service, and the Wildlife Management Institute. The authors declare no conflicts of interest.

References

Akaike H (1998) Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G (eds) Selected papers of Hirotugu Akaike. Springer, New York, pp 199–213

Anselin L, Griffith DA (1988) Do spatial effects really matter in regression analysis? Pap Reg Sci 65(1):11–34

Būhning-Gaese K (1997) Determinants of avian species richness at different spatial scales. J Biogeogr 24(1):49–60

Bini LM, Diniz JAF, Rangel TFLVB, Akre TSB, Albaladejo RG, Albuquerque FS, Aparicio A, Araújo MD, Baselga A, Beck J, Isabel Bellocq M, Böhning-Gaese K, Borges PAV, Castro-Parga I, Chey VK, Chown SL, De Marco Paulo Jr, Dobkin DS, Ferrer-Castán D, Field R, Filloy J, Fleishman E, Gómez JF, Hortal J, Iverson JB, Kerr JT, Daniel Kissling W, Kitching IJ, León-Cortés JL, Lobo JM, Montoya D, Morales-Castilla I, Moreno JC, Oberdorff T, Olalla-Tárraga MÁ, Pausas JG, Qian H, Rahbek C, RodrÍguez MÁ, Rueda M, Ruggiero A, Sackmann P, Sanders NJ, Terribile LC, Vetaas OR, Hawkins BA (2009) Coefficient shifts in geographical 565 ecology: an empirical evaluation of spatial and non-spatial regression. Ecography 32(2):193–204

Bishop A, Barenberg A, Volpe N, Riens J, Grosse R (2011) Nebraska land cover development. Rainwater Basin Joint Venture Report, Landcover Accuracy Assessment Report

Boyce MS (2006) Scale for resource selection functions. Divers Distrib 12(3):269–276

Chalfoun AD, Martin TE (2007) Assessments of habitat preferences and quality depend on spatial scale and metrics of fitness. J Appl Ecol 44(5):983–992

Chase JM, Myers JA (2011) Disentangling the importance of ecological niches from stochastic processes across scales. Philos Trans R Soc Lond 366(1576):2351–2363

Clifford P, Richardson S, Hémon D (1989) Assessing the significance of the correlation between two spatial processes. Biometrics 12:123–134

Coppeto SA, Kelt DA, Van Vuren DH, Wilson JA, Bigelow S (2006) Habitat associations of small mammals at two spatial scales in the northern Sierra Nevada. J Mammal 87(2):402–413

Cunningham RB, Lindenmayer DB, Crane M, Michael DR, Barton PS, Gibbons P, Okada S, Ikin K, Stein JAR (2014) The law of diminishing returns: woodland birds respond to native vegetation cover at multiple spatial scales and over time. Divers Distrib 20(1):59–71

Cushman SA, McGarigal K (2002) Hierarchical, multi-scale decomposition of species-environment relationships. Landscape Ecol 17(7):637–646

De Knegt HJ, van Langevelde FV, Coughenour MB, Skidmore AK, de Boer WF, Heitkönig IMA, Knox NM, van der Waal C, Prins HHT (2010) Spatial autocorrelation and the scaling of species-environment relationships. Ecology 91(8):2455–2465

Dungan JL, Perry JN, Dale MRT, Legendre P, Citron-Pousty S, Fortin M-J, Jakomulska A, Miriti M, Rosenberg MS (2002) A balanced view of scale in spatial statistical analysis. Ecography 25(5):626–640

Gelman A, Hwang J, Vehtari A (2014) Understanding predictive information criteria for Bayesian models. Stat Comput 24(6):997–1016

Godsill SJ (2001) On the relationship between Markov Chain Monte Carlo methods for model uncertainty. J Comput Graph Stat 10(2):1–19

Grand J, Cushman SA (2003) A multi-scale analysis of species-environment relationships: breeding birds in a pitch pine-scrub oak (*Pinus rigida–Quercus ilicifolia*) community. Biol Conserv 112(3):307–317

Gray TNE, Phan C, Long B (2010) Modelling species distribution at multiple spatial scales: gibbon habitat preferences in a fragmented landscape. Anim Conserv 13(3):324–332

Guillera-Arroita G, Lahoz-Monfort José JJ, MacKenzie DI, Wintle BA, McCarthy MA (2014) Ignoring imperfect detection in biological surveys is dangerous: a response to 'fitting and interpreting occupancy models'. PLoS ONE 9(7):71

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8(9):993–1009

Henebry GM (1995) Spatial model error analysis using autocorrelation indexes. Ecol Model 82(1):75–91

Holland JD, Bert DG, Fahrig L (2004) Determining the spatial scale of species' response to habitat. BioScience 54(3):227–233

Hooten MB, Hobbs NT (2015) A guide to Bayesian model selection for ecologists. Ecol Monogr 85(1):3–28

Horne JK, Schneider DC (1995) Spatial variance in ecology. Oikos 12:18–26

Hurlbert AH, Jetz W (2007) Species richness, hotspots, and the scale dependence of range maps in ecology and conservation. Proc Natl Acad Sci 104(33):13384–13389

Hutto RL, Pletschet SM, Hendricks P (1986) A fixed-radius point count method for nonbreeding and breeding season use. Auk 12:593–602

Jackson HB, Fahrig L (2015) Are ecologists conducting research at the optimal scale? Glob Ecol Biogeogr 24(1):52–63

Johnson DS, Hoeting JA (2011) Bayesian multimodel inference for geostatistical regression models. PLoS ONE 6(11):e25677

Jorgensen CF, Powell LA, Lusk JJ, Bishop AA, Fontaine JJ (2014) Assessing landscape constraints on species abundance: does the neighborhood limit species response to local habitat conservation programs? PLoS ONE 9(6):e99339

Keitt TH, Bjornstad ON, Dixon PM, Citron-Pousty S (2002) Accounting for spatial pattern when modeling organism-environment interactions. Ecography 25(5):616–625

Kellner KF, Swihart RK (2014) Accounting for imperfect detection in ecology: a quantitative review. PLoS ONE 9(10):e111436

Kirol CP, Beck JL, Huzurbazar SV, Holloran MJ, Miller SN (2015) Identifying greater Sage-Grouse source and sink habitats for conservation planning in an energy development landscape. Ecol Appl 25(4):968–990

Kuhn I (2007) Incorporating spatial autocorrelation may invert observed patterns. Divers Distrib 13(1):66–69

Legendre P, Dale MRT, Fortin M-J, Gurevitch J, Hohn M, Myers D (2002) The consequences of spatial structure for the design and analysis of ecological field surveys. Ecography 25(5):601–615

Lennon JJ (2000) Red-shifts and red herrings in geographical ecology. Ecography 23(1):101–113

Levin SA (1992) The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73(6):1943–1967

Mccarthy KP, Fletcher RJ Jr, Rota CT, Hutto RL (2012) Predicting species distributions from samples collected along roadsides. Conserv Biol 26(1):68–77

Nams VO (2005) Using animal movement paths to measure response to spatial scale. Oecologia 143(2):179–188

O'Hara RB, Sillanpää MJ (2009) A review of Bayesian variable selection methods: what, how and which. Bayesian Anal 4(1):85–117

Overmars KP, De Koning GHJ, Veldkamp A (2003) Spatial autocorrelation in multi-scale land use models. Ecol Model 164(2):257–270

Pearson RG, Dawson TP, Liu C (2004) Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. Ecography 27(3):285–298

Plummer M (2003) JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd international workshop on distributed statistical computing, vol 124, p 125

Plummer M (2013) rjags: Bayesian graphical models using MCMC. R package version 3

Pope SE, Fahrig L, Gray Merriam H (2000) Landscape complementation and metapopulation effects on leopard frog populations. Ecology 81(9):2498–2508

Robbins CS, Bystrak D, Geissler PH (1986) The breeding bird survey: its first fifteen years, 1965–1979. Report, DTIC Document

Robinson WS (1950) Ecological correlations and the behavior of individuals. Am Sociol Rev 15(3):351–357

Royle JA (2004) N-mixture models for estimating population size from spatially replicated counts. Biometrics 60(1):108–115

Saab V (1999) Importance of spatial scale to habitat use by breeding birds in riparian forests: a hierarchical analysis. Ecol Appl 9(1):135–151

Sandel B, Smith AB (2009) Scale as a lurking factor: incorporating scale-dependence in experimental ecology. Oikos 118(9):1284–1291

Schwartz MW, Iverson LR, Prasad AM, Matthews SN, O'Connor RJ (2006) Predicting extinctions as a result of climate change. Ecology 87(7):1611–1615

Steffan-Dewenter I, Münzenberg U, Bürger C, Thies C, Tscharntke T (2002) Scale-dependent effects of landscape context on three pollinator guilds. Ecology 83(5):1421–1432

Tenan S, O'Hara RB, Hendriks I, Tavecchia G (2014) Bayesian model selection: the steepest mountain to climb. Ecol Model 283:62–69

Thornton DH, Fletcher RJ (2014) Body size and spatial scales in avian response to landscapes: a meta-analysis. Ecography 37(5):454–463

Turner MG, O'Neill RV, Gardner RH, Mine BT (1989) Effects of changing spatial scale on the analysis of landscape pattern. Landscape Ecol 3(3–4):153–162

Urban DL, Robert VO, Shugart HH Jr (1987) A hierarchical perspective can help scientists understand spatial patterns. BioScience 37(2):119–127

van Langevelde F (2000) Scale of habitat connectivity and colonization in fragmented nuthatch populations. Ecography 23:614–622

Watanabe S (2013) A widely applicable Bayesian information criterion. J Mach Learn Res 14:867–897

Wheatley M, Johnson C (2009) Factors limiting our understanding of ecological scale. Ecol Complex 6(2):150–159

Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations. Academic Press, New York