Accounting for the space-varying nature of the relationships between temporal community turnover and the environment

Marta A. Jarzyna, Andrew O. Finley, William F. Porter, Brian A. Maurer, Colin M. Beier and Benjamin Zuckerberg

Ecological relationships are often explored using non-spatial regression models, even though this method is rarely adequate in settings where the processes operate at large spatial scales and when model covariates do not explain all variation present in the outcome variable. Given the complexity of ecological processes, it is often unrealistic to assume a set of stationary regression coefficients can capture space-varying and scale-dependent relationships between covariates and an ecological response. Spatially-varying coefficients (SVC) models fit within a Bayesian inferential framework provide a statistically robust method to explore potential space-varying and scale-dependent impacts of covariates. Our study objective was to assess the utility of SVC models for capturing non-stationary relationships between temporal community dynamics in avian assemblages and variation in environmental factors. We also wanted to compare the inference drawn from SVC models to that obtained from space-varying intercept models and also models that do not acknowledge any spatial structure beyond what is introduced by the covariates. Our analysis examines the temporal turnover, expressed as a proportion, of avian communities across New York State, USA. Given the expected outcome is non-Gaussian, we detail a generalized linear model specification of the proposed model structures. Our results show the SVC model outperformed the spatially-varying intercept and non-spatial models in terms of model fit and model predictive inference. Further, by fitting these models within a Bayesian inferential framework, we were able to make inferences about the spatial impact of covariates and other process parameters, as well as obtain full posterior predictive inference about the rate of turnover at new, unobserved locations. We conclude that SVC models provide a flexible framework for exploring and accounting for non-stationary mechanisms driving ecological patterns.
outcome) can improve inference, increase accuracy and precision of model predictions, and reveal missing covariates.

While adding a space-varying intercept to the model accounts for spatial dependency and often improves inference, it does not explicitly deal with spatial non-stationarity and its associated influence on model covariates. Spatial non-stationarity is present when the strength and nature of the relationship between a response variable and the predictor covariates vary across the spatial domain (Fortin and Dale 2009, Miller 2012). Interactions or feedbacks among unobserved and observed covariates or presence of inherently different mechanisms that impact the outcome in different parts of the study region can produce non-stationarity (Miller 2012). Non-stationary relationships are common in ecology (Bini et al. 2009). For example, Foody (2004) showed that the relationships between avian species richness and total annual precipitation, mean annual temperature and terrestrial land cover in sub-Saharan Africa were strongly non-stationary. Other large-scale studies support these findings (Grøtan et al. 2009, Martin-Queller et al. 2011, McNew et al. 2013). Thus, given this complexity of ecological phenomena across large spatial scales, it is often unrealistic to assume a set of stationary regression coefficients (i.e. coefficients that are constant across space) can capture space-varying and scale-dependent relationships between covariates and outcome variables (Finley 2011). In other words, a global ecological relationship is often affected by local processes yielding a heterogeneous pattern, and would be described more accurately by local model parameter values that differ from the global values (Miller 2012). Ignoring spatial non-stationarity can have similar consequences to those resulting from ignoring spatial autocorrelation. In a strongly non-stationary system, estimates of model parameters and inferences and predictions resulting from models that do not adequately account for spatial heterogeneity are likely to be flawed. Despite the importance of the problem and apparent ubiquity of the space-varying nature of ecological phenomena, few methods provide an opportunity to evaluate and account for potential spatially varying relationships between ecological responses and environmental covariates. Bayesian spatially-varying coefficients (SVC) models are one of the more flexible and robust approaches for accommodating non-stationarity (Gelfand et al. 2003). SVC models use a valid probability model that affords full posterior inferences for model parameters and subsequent prediction of the outcome and covariate processes at any new location. While spatially-varying coefficient models have been developed and tested for data that follow a Gaussian distribution (Finley 2011), models for non-Gaussian outcomes, e.g. binomial or Poisson distributions, have not yet been widely applied to ecological questions, with a few exceptions that focus on statistical methodology development (Finley et al. 2009, Finley et al. 2011). Here, we consider a SVC model for evaluating the relationships between the temporal turnover (i.e. changes in community composition across time) in avian biodiversity and a set of environmental covariates.

We chose to focus on temporal turnover in community composition for two reasons. First, patterns and changes in biological diversity are an indication of the underlying mechanisms that control biodiversity and are one of the fundamental fields of investigation in ecology. Consequently, understanding patterns of biodiversity and their environmental determinants have been at the center of recent ecological research (Jost 2007, Buckley and Jetz 2008, Basela 2010, Kraft et al. 2011). While traditionally community turnover has been used to describe biodiversity patterns across space rather than time, it can also be viewed in terms of a temporal change in species composition occurring over a specified time period. Indeed, describing temporal changes in communities is of increasing importance because of the global environmental threats such as climate change and land-cover change, both of which have a strong temporal component. Understanding temporal relationships between biodiversity patterns and environmental variability is necessary to accurately forecast the consequences of environmental change and design sound conservation strategies.

Second, the often observed space-varying nature of temporal community turnover and the kinds of inference we seek encourage the use of new, more flexible, tools such as generalized linear SVC models. Further, because biodiversity dynamics are often summarized using indices such as Jaccard, Sorensen, or Diamond–May, there is a need for development of the methods appropriate for such non-Gaussian outcomes.

Our objective was to test the utility of the SVC model to quantify the influence of non-stationarity on relationships between temporal community dynamics in avian assemblages and changes in environmental factors such as climate variability and patterns of landscape fragmentation. Specifically, we sought to assess the merits of a Bayesian spatially-varying coefficients approach in comparison with non-spatial and spatially-varying intercept models.

**Data and methods**

**Site description**

The study area is the State of New York, USA. New York covers 128,401 km², including 42,440 km² of inland water (excluding of the boundary-water areas of Long Island Sound, New York Harbor and lakes Ontario and Erie). Most of the state lies between latitudes 42 and 45° N and between longitudes 73.5 and 79.75° W. Climate of New York is affected by the state’s broad elevation gradient as well as by the proximity to lakes Erie and Ontario, and the Atlantic Ocean. Adirondack and Catskill Mountains in the east are among the highest regions of the state, with elevation varying between 600 and 1500 m. South-western New York is located in the northern portions of the Allegheny Plateau and its elevation ranges from approximately 300 to 900 m. North-western New York as well as Hudson River valley and New York City are low-elevation regions with elevation ranging from approximately 0 to 200 m.

The state’s land-cover is approximately 51.2% forest (deciduous, evergreen, and mixed forests), 13.4% pasture and hay, 8.2% cultivated crops, 2.9% scrub and shrub, 1.0% grassland and herbaceous, 8.0% wetland cover, 8.7% developed land, and 0.2% barren land (Homer et al. 2004). Forested areas dominate the Adirondack, Catskill and Alleghany regions, while agriculture is prevalent on historic glacial lake plains south of Lake Ontario.
New York offers a broad gradient of landscape fragmentation. The regions of the Adirondack and Catskill Mountains are the least fragmented. The Hudson River valley runs through the eastern part of the state and is more fragmented in terms of land-cover. The landscapes of western New York are more fragmented and characterized mostly by agriculture–forest mosaic. Because of these heterogeneous patterns in landscape fragmentation, elevation gradients, and diversity of ecosystems, we might expect space-varying impacts of covariates to explain variation in ecological outcomes of interest.

Breeding Bird Atlas

We used the New York State Breeding Bird Atlas (BBA) as our model dataset to characterize changes in avian communities through time. BBA is a statewide survey that documented the distribution of breeding birds in New York. The BBA has been conducted in two time periods, 1980–1985 (hereafter, BBA1980; Andrle and Carroll 1988) and 2000–2005 (hereafter, BBA2000; McGowan and Corwin 2008). For both BBAs, a grid system was used to define the basic unit for reporting data. The entire State of New York was divided into approximately 1300 squares, each measuring 10 × 10 km. The BBA reporting unit (a block) was one quarter of this square and measure 5 × 5 km and a total of 5,335 blocks covered the entirety of New York State. This data set represents one of the largest and finest resolution atlases in the world (Gibbons et al. 2007).

A total of 242 species were recorded for BBA1980 and 248 species were recorded for BBA2000 (see Supplementary material Appendix 1, Table A1 for the list of species used in this analysis). Avian breeding was recorded at three levels of certainty of breeding occurrence based on the behavior of birds observed: possible, probable, and confirmed (Andrle and Carroll 1988, McGowan and Corwin 2008). Observations were made by skilled birders who spent at least 8 h in each block, visited all cover types in each block, and included at least one nighttime visit to document nocturnal species. Observer effort was recorded for each BBA and reported as the number of person hours (McGowan and Zuckerberg 2008). A block was considered sufficiently surveyed when at least 76 species were documented (with exceptions of blocks that might be expected to have fewer species). The BBA represents a presence/absence dataset, although absence indicates that species could not be found given search criteria (McGowan and Corwin 2008).

Temporal turnover

While many different indices for quantifying community turnover have been developed over the years (Gaston et al. 2004, Magurran 2004, Tuomisto 2010a, b), we used the Diamond–May (DM) index (Gaston et al. 2004). DM index is calculated as a proportion and designates turnover as high when the proportion of species shared between two sites (or two time steps) is low (Gaston et al. 2004). From a statistical perspective, DM is a binomial outcome, where the number of species lost and gained across space or time is the number of successes, while the total number of species is the number of trials. DM index has traditionally been used to quantify spatial turnover, but it can be easily adapted to reflect temporal turnover as follows:

$$DM = \frac{E + C}{E + C + P}$$

where $E$ is the number of species that went extinct in the block between BBA1980 and BBA2000 (i.e. extinction), $C$ is the number of species that colonized the block between BBA1980 and BBA2000 (i.e. colonization), and $P$ is number of species in the same block common to both BBAs (i.e. persistence). The values of DM are bounded by 0 and 1; values approaching 0 indicate low temporal turnover in a block between BBA1980 and BBA2000, while values approaching 1 indicate high temporal turnover in a block between BBA1980 and BBA2000.

Model covariates

Environmental factors that we deemed especially important in shaping changes in community composition were temporal trends in climatic variables and landscape fragmentation. We also considered survey effort (number of survey hours) as a variable potentially affecting the observed temporal community turnover.

Climatic trends

The climate data was derived from the PRISM (Parameter-elevation Regressions on Independent Slopes Model) climate mapping system (Daly and Gibson 2002). PRISM consists of interpolated monthly maximum and minimum temperatures and precipitation at a 2.5-arcmin resolution from 1891–2010 for the entire contiguous United States.

We calculated the magnitude of the 25-yr (1980–2005) trend in average monthly maximum and minimum temperatures and in total monthly precipitation using ordinary least squares regression. The slope of the OLS regression indicates the magnitude of the trend and reflects the amount of change in climatic variables that occurred between 1980 and 2005. We then interpolated the trend magnitudes for each BBA block and averaged the monthly values to reflect breeding season (May through September) trend magnitude in maximum and minimum temperatures (TMAX and TMIN, respectively) and in total precipitation (PRECIP). The trend magnitudes of TMAX and TMIN were expressed in °C/25 yr, while the units of the trend magnitudes of PRECIP were in mm/25 yr.

Habitat fragmentation

Even though amount of suitable habitat is often cited as one of the most important factors driving patterns of species occurrence, it is difficult to quantify suitable habitat while dealing with a group of species with vastly different habitat requirements. Habitat fragmentation has been shown to also be a significant driver of songbird community dynamics (Kennedy et al. 2011). Hence, we chose landscape fragmentation as the primary land-cover factor associated with community turnover.
Habitat fragmentation variables were derived using the 30 × 30 m National Land Cover Data (NLCD, Homer et al. 2004) product. Because there is no land cover data readily available for the time period of BBA1980, we used a space-for-time substitution to assess whether temporal turnover is related to landscape fragmentation (Pickett 1989, Zuckerberg and Porter 2010). Space-for-time substitution assumes that spatial variation is equivalent to a temporal variation, and relationships between variables (e.g. species occurrence and land cover) derived for one time step across a large spatial extent will be equivalent to that derived for two time steps. Thus, we quantified landscape fragmentation for one time step only and looked for the relationship between the temporal turnover and the observed spatial variation in landscape fragmentation.

We used the 2001 NLCD, because it coincided well with the time of BBA2000. The 2001 NLCD consists of 16 land cover classes. Prior to landscape analysis, we consolidated open space developed, low intensity developed, medium intensity developed, high intensity developed into one class of developed land; cultivated crops and pasture/hay into one class of agriculture; and deciduous forest, evergreen forest, mixed forest, and forested wetland into one class of forest. We decided to consolidate these types of land cover to improve accuracy of classification and to simplify the environment for purposes of our evaluation. We ended up with nine land cover classes: water/ice, developed, barren land, forest, scrub/shrub, grassland/herbaceous, agriculture, and wetlands.

Landscape fragmentation was quantified using FRAGSTATS 4.1 (McGarigal et al. 2002) and the Geospatial Modelling Environment (GME, Beyer 2013). We recognize that there are multiple ways to measure fragmentation and that careful characterization of landscape fragmentation for a diverse suite of species often requires multiple metrics. However, our purpose was not comprehensive ecological analyses, but rather evaluation of the space-varying qualities of fragmentation. Therefore, we chose a landscape-scale variable to capture broad-scale variation in habitat fragmentation and best represent requirements of a diverse suite of species with varying habitat requirements. Specifically, we chose edge density (ED) as a measure of landscape fragmentation because an increase in habitat edge is a primary outcome of habitat fragmentation (Hargis et al. 1998). ED was also reported as an effective tool for evaluating landscape fragmentation and performed better than other popular landscape fragmentation indices (Hargis et al. 1998). However, in situations when landscape consists entirely of one cover type, the ED would be 0 regardless of the type of land cover present. Therefore, in order to differentiate between landscapes that consisted entirely of natural land cover (e.g., forest) and those mostly developed (e.g., urban areas), we also included the proportion of developed land (DEVEL) metric in our models.

### Effort

Increasing survey effort often results in a higher number of recorded species (Tobler et al. 2008). Different survey effort between BBA1980 and BBA2000 could decrease the value of the DM indicating high temporal turnover, which could be a result of different number of recorded species rather than actual changes in species identities throughout time. Despite the fact that concerted attention was invested to minimize the effects of survey effort in the BBA surveys, we wanted to recognize the need to control for survey effort bias. To account for potential survey effort bias, we calculated the absolute difference in the number of person hours between BBA1980 and BBA2000 (EFF = EFF1980 − EFF2000) for each BBA block and included it as a covariate in our models.

### Space-varying coefficient models

We evaluated three statistical models, each of which included main effects of all the covariates (TMAX, Tmin, PRECIP, ED, DEVEL, and EFF). We standardized all the covariates to ease the comparison and interpretation of the coefficient estimates. Since the temporal turnover (DM) is a proportion, we used a binomial regression model. We started with a non-spatial regression specification (hereafter, non-spatial model) and then increased the level of model complexity by adding spatially structured random effects first to the intercept to form a spatially-varying intercept (SVI) model, then to the regression parameters associated with the covariates to form the SVC model.

More formally, for a given BBA grid cell the regression outcome variable $y(s)$ is the numerator in Eq. 1, where $s$ represents the grid cell’s spatial coordinates. Given the denominator in Eq. 1, denoted $N(s)$, we assume $y(s)$ follows a binomial distribution. For the given grid cell, $y(s) \mid \eta(s) \sim \text{Binomial}(N(s), \eta(s))$, where $\eta(s)$ is the probability of success at location $s$ and $\eta(s)$ is model's regression equation, which for the SVC model is equal to $x(s)\beta + x(s) \cdot w(s)$ with $x(s)$ representing the vector comprising an intercept and location specific covariate values. The parameters to be estimated are the global regression coefficients, $\beta = (\beta_0, \beta_{\text{TMAX}}, \beta_{\text{DEVEL}}, \beta_{\text{PRECIP}}, \beta_{\text{ED}}, \beta_{\text{DEVEL}}, \beta_{\text{ED}})^T$, and associated spatially-varying adjustments, $w(s) = (w_0(s), w_{\text{TMAX}}(s), w_{\text{DEVEL}}(s), w_{\text{PRECIP}}(s), w_{\text{ED}}(s), w_{\text{DEVEL}}(s), w_{\text{ED}}(s))^T$. A logit link function $\eta(s) = \exp(\eta(s))/(1 + \exp(\eta(s)))$ was used for this model.

We assume each element in the spatially-varying coefficients vector, $w(s)$, arises from a spatial Gaussian process (GP) (see Banerjee et al. 2004 or Cressie and Wikle 2011 for more details). Specifically the $j$-th spatially-varying coefficient at location $s$ is $w_j(s) \sim \text{GP}(0, C_j(s, s'))$, where $s$ and $s'$ are any two locations within the study area, the spatial covariance $C_j(s, s') = \sigma_j^2 \rho_j(s, s'; \phi_j)$ with variance parameter $\sigma_j^2$, correlation function $\rho_j(\cdot; \phi_j)$, and spatial decay parameter $\phi_j$. The spatial decay parameter $\phi_j$ controls the decay in spatial correlation; lower values of the parameter indicate shorter range in spatial correlation. The exponential spatial correlation was assumed for $\rho_j(\cdot)$.

Prior distributions on the remaining parameters complete the hierarchical model, see Gelman et al. 2004 for detail on Bayesian model specification. Here, we specified uninformative prior distributions. The regression coefficients, $\beta_j$’s followed a Normal distribution $N(0, 100)$, while the variance components $\sigma_j^2$’s were assigned inverse-Gamma IG(2, 1) priors, where 2 and 1 are shape and scale hyper-parameters, respectively. This specification of the IG provides a prior distribution with infinite variance that is centered on 1.
The spatial decay parameters, \( \phi \)'s, followed an informative Uniform prior, with support that ranged from 1 km to the maximum distance between any two grid cells. Exploratory analysis using different hyper-parameters suggested prior specification had little influence on parameter and predictive inference. This is not surprising given the large size of the data set and choice of prior distributions. Our sensitivity analysis did suggest that given an IG shape hyper-parameter of 2, very large scale parameter values, e.g. greater than 10, resulted in poor MCMC chain convergence by forcing an unreasonable amount variability in the GP. For the SVC model, we ran three MCMC chains for 60 000 iterations each. We ran 10 000 iterations for the non-spatial and SVI models, as those required less time to converge. Convergence was diagnosed using the Gelman–Rubin diagnostic (Gelman and Rubin 1992).

Given we are working within a Bayesian framework, candidate models fit to the observed data were assessed using the deviance information criterion (DIC; Spiegelhalter et al. 2002). We computed the expected posterior deviance as \( D(\Omega) = E_{\text{Bayes}}[-2\log L(\text{Data}|\Omega)], \) where \( \Omega \) is the set of parameters estimated for each model and \( L(\text{Data}|\Omega) \) is the first stage Gaussian likelihood from the respective models. We computed the effective number of parameters as \( p_D = D(\Omega) - D(\bar{\Omega}) \), where \( \bar{\Omega} \) is the posterior mean of the model parameters. The DIC was then computed as \( D(\bar{\Omega}) + p_D \). Lower values of DIC indicate improved fit. Model accuracy was compared using root mean square error (RMSE) that was calculated as the square root of the mean squared deviations between fitted and observed outcomes. Lower values of RMSE indicate improved accuracy.

Candidate model performance for out-of-sample prediction was assessed using a randomly selected 10% holdout dataset. While there is no convention on how much of the original data should be held out for model validation, we believe that 10% of the data is sufficient as it amounts to approximately 500 observations. Also, while several other approaches to verify model's predictive ability are available, holding-out a set of randomly selected observations is a common model validation technique (Pearson et al. 2002, Finley 2011). An alternative approach to verifying model's predictive ability would be bootstrapping, where the original dataset is sampled randomly with replacement (Austin and Tu 2004). Bootstrapping requires building and running multiple models and in each case predictive performance is assessed against the corresponding test data. While bootstrapping is a robust approach, it is computationally costly, particularly in situations when models are complex. Given that fitting a single SVC model is very computationally intensive in itself, repeating this procedure multiple times for bootstrapping would be prohibitive and thus was not feasible in our case. Hold out set predictions were compared to the observed values using root mean square error (RMSE) where lower values indicate improved performance.

Prior to statistical analysis, we removed all blocks that did not have continuous land cover coverage; those were mainly blocks at the periphery of the study area, which extended farther than the NLCD layer. Also, blocks with more than 50% open water coverage and those that did not have survey effort data were removed. Ultimately, we used a total of \( n = 4271 \) blocks to fit the models and \( n_{\text{ho}} = 473 \) holdout blocks to test the models' predictive ability. The non-spatial and SVI models were run in R 2.15.1 statistical package (R Development Core Team) using package spBayes (Finley and Banerjee 2013). We wrote and compiled the SVC model in C++ programming language and R 2.15.1 (R Development Core Team). Summaries of parameter estimates were generated using the R coda package (Plummer et al. 2012).

**Results**

**Temporal turnover**

The mean grid cell value of the DM index was 0.39, which indicates that on average approximately 39% of species had either colonized or become extinct and 61% of the species were common to both BBAs (Fig. 1A). Northern regions of the state (i.e. Adirondack Mountains) experienced the highest temporal community turnover, with values reaching 1.0 (i.e. complete turnover in species assemblage) in several locations (Fig. 1A). Western (agricultural) and eastern (urbanized) parts of the state also showed slightly higher than average temporal turnover, while the central regions of the state generally underwent the lowest temporal turnover (Fig. 1A). The lowest values of temporal turnover was approximately 0.11.

**Covariates**

The north-eastern and south-western parts of the state experienced the largest warming trend, while the southern and central New York generally underwent decreases in average maximum temperatures (see Supplementary material Appendix 1, Fig. A1 for maps of covariates). Minimum temperatures increased across most of the state, though some small regions experienced a cooling trend (Supplementary material Appendix 1, Fig. A1).

Northern (i.e. region of the Adirondack Mountains), western and southern New York experienced drier conditions, while the rest of the state experienced wetter conditions (Supplementary material Appendix 1, Fig. A1).

Regions of Adirondack and Catskill Mountains as well as some locations in the south and south-western parts of the state had the lowest edge density (i.e. lowest habitat fragmentation; Supplementary material Appendix 1, Fig. A1). The percentage of developed land was low across most of New York, with exception of large urban centers such as New York City, Albany, Syracuse, Rochester, and Buffalo. Survey effort was relatively constant across the study region (Supplementary material Appendix 1, Fig. A1).

**Statistical analysis**

**Model fit and predictive ability**

As suggested by the DIC, RMSE, and RMSE\(_{\text{pred}}\) values provided in Table 1, adding spatial random effects to the posited model improves fit to the observed data, model accuracy, and predictive performance for new, unobserved, locations. The improvements to fit and model accuracy suggest there was
All candidate models provide a close visual approximation to the observed data (see Fig. 1 for the candidate model fitted value surfaces). For the SVI model (Fig. 1C), the addition of the spatial random effect to the model intercept provides a smooth process that captures spatial discrepancies between observed values and those estimated with the additive linear trend involving the covariates. The SVC model fitted surface (Fig. 1D) provides the closest approximation to the observed data because the regression coefficients processes are able to estimate the covariates' local and regional impact on temporal turnover. For example, the variation in the temporal turnover in the north-east New York is better represented on the SVC model fitted surface (Fig. 1D) than on surfaces resulting from the other two models (Fig. 1B, C).

In terms of model predictive performance, all models’ predicted surfaces approximated the holdout data relatively well (Fig. 2). However, by not relying only on the smooth process on the intercept, the SVC model is able to accommodate more local-scale variation than the SVI model (Fig. 2D). For example, several locations of high temporal turnover in the western part of the study area were picked up by the SVC predictive model, but not by either the non-spatial or the SVI predictive models (Fig. 2D).

Table 1. Comparison of the non-spatial, spatially-varying intercept (SVI), and spatially-varying coefficients (SVC) models using deviance information criterion (DIC), estimated number of parameters pD, and relative mean square error (RMSE). Predictive performance of the models was evaluated by calculating RMSEpred of the observed values from the hold-out set and the predicted values resulting from all three models.

<table>
<thead>
<tr>
<th>Model</th>
<th>ΔDIC</th>
<th>pD</th>
<th>RMSE</th>
<th>RMSEpred</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVC</td>
<td>0.00</td>
<td>458.90</td>
<td>0.076</td>
<td>0.072</td>
</tr>
<tr>
<td>SVI</td>
<td>1063.84</td>
<td>106.30</td>
<td>0.082</td>
<td>0.079</td>
</tr>
<tr>
<td>Non-spatial</td>
<td>2871.31</td>
<td>7.20</td>
<td>0.090</td>
<td>0.088</td>
</tr>
</tbody>
</table>

substantial spatial dependence among the non-spatial models residuals – a generalized linear model assumption violation – and clear indication that a more complex spatial model was warranted. The pD in Table 1 is the effective number of parameters and is used in the DIC calculation to penalize more complex models. For example, the non-spatial model has only seven regression coefficients, hence, pD is approximately seven. The increasing number of spatial random effects added to SVI and SVC, respectively, is reflected by larger penalties. Despite the larger number of effective parameters, the SVC model has lower DIC than the non-spatial and SVI models.
Coefficient estimates

All regression coefficients in the non-spatial model were statistically significant (Table 2), i.e. the 95% credible intervals did not include zero. For spatially-varying intercept models, all of the regression coefficients remained significant (Table 2), but their magnitude and credible intervals changed in comparison with the non-spatial model (the magnitude of the coefficient estimates was assessed by simply comparing the direction and absolute value of the mean estimates resulting from both models). The magnitude of $\beta_0$, $\beta_{TMAX}$, $\beta_{TMIN}$, $\beta_{PRECIP}$, and $\beta_{ED}$ decreased after the residual spatial dependence was accommodated, while $\beta_{DEVEL}$ and $\beta_{EFF}$ increased in magnitude.

For a SVC model, it is typically not instructive to focus on a covariate’s regression coefficient mean, i.e. $\beta$, but rather assess the variability of the coefficient over the study area. That is, we draw inference by looking to maps of $\beta + w(s)$ (see Fig. 3 for $\beta + w(s)$ associated with each covariate). All SVC model coefficient estimates were spatially heterogeneous both in terms of the magnitude of the coefficient estimate as well as the direction of the relationship between the covariate and the outcome variable (Fig. 3). Here too, grid cells where the 90% credible intervals of the given $\beta + w(s)$ include zero are identified with a black point. Or conversely, those regions with no black points differ significantly from

Table 2. Coefficient estimates resulting from the non-spatial and spatially-varying intercept (SVI) models for all model covariates: intercept ($\beta_0$), magnitude of the 25-yr (1980–2005) trend in average maximum temperature of the breeding season ($\beta_{TMAX}$), magnitude of the 25-yr (1980–2005) trend in average minimum temperature of the breeding season ($\beta_{TMIN}$), magnitude of the 25-yr (1980–2005) trend in average total precipitation of the breeding season ($\beta_{PRECIP}$), percent developed land ($\beta_{DEVEL}$), edge density ($\beta_{ED}$), and effort ($\beta_{EFF}$). 50% indicates the mean of the posterior distribution, while 2.5 and 97.5% are the lower and upper quantiles of the posterior distribution.

<table>
<thead>
<tr>
<th>Model</th>
<th>Coefficient</th>
<th>50%</th>
<th>2.5%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-spatial</td>
<td>$\beta_0$</td>
<td>-0.492</td>
<td>-0.499</td>
<td>-0.486</td>
</tr>
<tr>
<td></td>
<td>$\beta_{TMAX}$</td>
<td>0.063</td>
<td>0.055</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>$\beta_{TMIN}$</td>
<td>0.041</td>
<td>0.034</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>$\beta_{PRECIP}$</td>
<td>-0.052</td>
<td>-0.059</td>
<td>-0.045</td>
</tr>
<tr>
<td></td>
<td>$\beta_{DEVEL}$</td>
<td>0.017</td>
<td>0.010</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>$\beta_{ED}$</td>
<td>-0.092</td>
<td>-0.099</td>
<td>-0.086</td>
</tr>
<tr>
<td></td>
<td>$\beta_{EFF}$</td>
<td>-0.010</td>
<td>-0.016</td>
<td>-0.004</td>
</tr>
<tr>
<td>SVI</td>
<td>$\beta_0$</td>
<td>-0.375</td>
<td>-0.435</td>
<td>-0.275</td>
</tr>
<tr>
<td></td>
<td>$\beta_{TMAX}$</td>
<td>0.018</td>
<td>0.007</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>$\beta_{TMIN}$</td>
<td>0.026</td>
<td>0.014</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>$\beta_{PRECIP}$</td>
<td>-0.015</td>
<td>-0.031</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>$\beta_{DEVEL}$</td>
<td>0.032</td>
<td>0.022</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>$\beta_{ED}$</td>
<td>-0.047</td>
<td>-0.057</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td>$\beta_{EFF}$</td>
<td>-0.014</td>
<td>-0.020</td>
<td>-0.008</td>
</tr>
</tbody>
</table>
IBS special issue

zero in the direction indicated by the grid cell color. Some caution needs to be taken when considering the hypothesis tests implied in Fig. 3. In a testing situation, where we encounter a very large number of hypotheses, controlling type-1 errors would be required, hence raising the question of multiple comparisons. However, we do not pursue such a formulation here and simply look to Fig. 3 to identify regions of substantial deviations of the spatial coefficients from the global mean (see also Supplementary material Appendix 1, Fig. A2 for maps of deviations of the spatially-varying coefficient estimates from the global mean).

**Discussion and conclusions**

We found that the SVC model outperformed both the non-spatial and SVI models in terms of model fit, despite the large number of parameters it had to estimate, and predictive performance. We feel confident that the increase in model fit is not a result of over parameterization, although, like any highly flexible random effect model, the SVC models can occasionally over fit the observed data. However, compared to models with unstructured random effects, the flexible process imposed by the spatial Gaussian process reduces the possibility of over fitting. Additionally, the penalizing qualities of the DIC metric account for the potential favoring the parameter rich SVC model. Furthermore, over fitting generally results in reduced predictive performance compared to simpler models; however, we did not find lower predictive performance of the SVC model. Anecdotally, while working with other data sets, we have noticed that the propensity to over fit the observed data is greater when the number of observed locations is small (e.g. n < 100), sparsely sampled, and when the range of spatial dependence is short. In these settings, the interpolating qualities of the Gaussian processes reduce the predictive power of the SVC model. When the data set is large, however, observations densely cover the domain, and the parameters of the underlying spatial processes associated with the regression coefficients are well estimated, then the risk of over fitting is reduced and prediction does not suffer.

The improvement in model fit suggests the relationship between temporal community turnover and environmental covariates is more spatially-varying in nature than it is stationary. Indeed, by allowing the regression coefficients to vary spatially over the domain and accommodate local impact of the covariates, the SVC model was able to explain more local-scale variation in temporal community turnover than the other two models. The assumption that regression coefficients are stationary results in poor fit and misleading inference about the impact of the covariates on temporal turnover. The assumption of scalar regression coefficients is pervasive in the literature. In the SVC model, any component of model spatial residual pattern due to non-stationarity of the impact of covariates is apportioned to the coefficients. Thus, the SVC model, despite being an intrinsically correlational approach, provides the richer opportunity for ecological interpretation.

We commonly see that once residual spatial dependence is accommodated, e.g. via a spatial random effect, regression coefficients differ from those resulting from a non-spatial model. In the non-spatial model, all covariates were significant predictors of temporal community turnover. Once spatial dependence was taken into account in the SVI model, the covariates remained significant but the magnitude of their coefficient estimates changed. For example, the association of the TMAX, TMIN, PRECIP, and DEVEL with the temporal community turnover was weaker for the
SVI model than for the non-spatial model, while that of ED and EFF became stronger. Such results suggest the non-spatial model violated the assumption of independent and identically distributed residuals and imply that ecological inferences drawn from these models will be different. Other research corroborates our findings. For example, Record et al. (2013) found that climatic variables were significant predictors of presence–absence of two tree species in non-spatial models, but in the spatial models the magnitude and sign of some parameter estimates changed. Foody (2004) found the magnitude of relationships between species richness and the explanatory environmental variables changed when spatial dependency and non-stationarity were accounted for, even though statistically significant relationships were established using a conventional global regression analysis.

When the space-varying nature of the regression coefficients is accommodated, different regions within the study area might exhibit conflicting relationships between the outcome and covariate, causing the study-area-wide mean of the associated regression coefficient to be similar to those seen in the SVI models. In our case, spatial heterogeneity of the relationships between covariates and temporal community turnover became apparent after accounting for non-stationarity. Some regions were characterized by negative relationships, other regions displayed positive relationships, yet another showed no significant relationships at all. For example, a positive relationship between temporal community turnover and the magnitude of the trend in minimum temperatures was found in the eastern New York, while this relationship was negative in the central part of the state. Such spatial pattern might indicate that communities in eastern New York are responding stronger or faster to increasing temperatures resulting from climate change than those found in central New York. Similarly, we found that the strongest negative influences of landscape fragmentation were in eastern part of the state, i.e. in regions dominated by Adirondack Mountains and the Hudson River valley. Perhaps in those historically forested and contiguous habitats, small levels of landscape fragmentation have a much more detrimental effect on community composition than elsewhere within the state. Furthermore, given that the strongest influences of these two covariates were detected in similar geographic regions, it is also plausible that these two ecological factors interact in driving temporal community turnover. The remaining covariates showed similar heterogeneous patterns. These heterogeneous surfaces illustrate the complexity of impact and interplay among the covariates in explaining variability in temporal community turnover.

Other research provides evidence for non-stationarity of ecological relationships (Bini et al. 2009, Martin-Queller et al. 2011, McNew et al. 2013) and thus corroborates findings from our study. For example, Grøtan et al. (2009) found the proportion of variance in local recruitment of great tit explained by spring temperature differed nearly 10-fold among four Dutch populations. Sæther et al. (2008) showed that population dynamics of eight species of prairie ducks showed pronounced latitudinal gradient; depending on the location, influence of increased spring temperature or winter precipitation on population dynamics was either positive or negative. Anders and Post (2006) found that NAO affected populations of the migrant yellow-billed cuckoo Coccyzus americanus only in southern and eastern parts of the US, while El Nino Southern Oscillation (ENSO) affected populations in northern and central US. Martin-Queller et al. (2011) found relationships between species richness of woody plants and environmental and biotic factors to also be spatially heterogeneous. Given this apparent ubiquity of space-varying relationships in nature, we suggest that researchers studying large-scale ecological phenomena should be especially cognizant of the need to evaluate and account for potential non-stationarity.

The three models we tested yielded different results despite the fact that the same covariates were included in all of them. By ignoring spatial dependency in model residuals, the non-spatial model produced relationships between the covariates and temporal community turnover. In the SVI model, the relationships between covariate and outcome variable generally weakened as a result of accommodating the spatial auto-correlation in model covariates via spatial random effect. By allowing the coefficient estimates to vary spatially, the SVC model yielded heterogeneous patterns and often conflicting relationships between the outcome and covariate. The heterogeneity of the model covariates themselves (e.g. the magnitude of the climatic trends) might be partly responsible for driving such space-varying relationships. Applying the SVC model enabled us to detect and account for these spatial differences.

It must be noted that other approaches exist to evaluate and account for potential non-stationarity of ecological phenomena. One of the more common options is a frequentist method called geographically-weighted regression (GWR, Fotheringham et al. 2002), which uses spatial weights to estimate spatially adaptive coefficients. GWR has been the leading statistical method to account for spatial non-stationarity (Ma et al. 2012a, b, Miller 2012), in no small part due to its availability in several popular Geographic Information Systems, e.g. ESRI products, and statistical computing environments, e.g. spgwr package in R statistical program (Bivand and Yu 2013). However, recently it has been shown that GWR is not robust to collinearity among the covariates and the presence of complex spatial correlation structures (Wheeler and Waller 2009, Finley 2011) and it tends to provide less accurate and more correlated regression coefficient estimates than those resulting from SVC models (Wheeler and Cadler 2007). Furthermore, Ma et al. (2012a) used GWR to model species richness derived from the same dataset as ours and concluded that GWR performed poorly the larger the spatial scale of investigation. Another central shortcoming of GWR, from an inferential standpoint, is the lack of a legitimate probability model in the sense that the joint distribution linking the parameters and the data is not a valid probability distribution. This is problematic for inference because the standard errors computed from such models may not be justifiable. Asymptotic arguments may be supplied but are complicated in spatial contexts because of the divergent paradigms of infill and expanding domain asymptotics. Thus, GWR can be a useful tool for exploratory data analysis, but generally should not be used in settings where one seeks inference about the importance of model parameters or in prediction.
Given the growing wealth of space and time indexed ecological data (Kelling et al. 2009, Reichman et al. 2011, Schimel 2011) deeper insight into ecological processes is contingent upon our ability to specify valid and tractable models able to accommodate often complex spatio-temporal relationships between outcomes and posited environmental drivers. Because the propensity for non-stationary relationships increases over larger spatial domains and time periods, need for more sophisticated models also increase as we begin looking at macro-system questions. This study represents an important case assessing the utility of SVC models for capturing non-stationary relationships between non-Gaussian biodiversity dynamics and environmental changes. We suggest that SVC models are especially appropriate for studying macro-scale (i.e. regional or global) systems where we can expect the ecological processes and environmental drivers to be strongly heterogeneous across the spatial domain. As such, SVC models provide a unique opportunity to exploring environmental questions about impact of global environmental change on biodiversity.

Acknowledgements – We would like to thank the thousands of volunteers who participated in both New York State Breeding Bird Atlases. We also thank Kimberley Corwin and Kevin McGowan for supplying atlas databases and other information. The manuscript benefited from discussions with the members of the Quantitative Wildlife Laboratory at Michigan State Univ. We thank James H. Stagg for technical assistance with hyper-performance computing. This study received financial support from NASA Grant NNX09AK16G and Boone and Crockett Club. AOF’s work was supported by National Science Foundation Grants DMS-1106609, EF-1137309, EF-1241874 and EF-1253225, as well as NASA Carbon Monitoring System grants.

References


McGarigal, K. et al. 2002. FRAGSTATS: spatial pattern analysis program for categorical maps. – Computer software program produced by the authors at the Univ. of Massachusetts, Amherst, <www.umass.edu/landeco/research/fragstats/fragstats.html> accessed October 2012.
McGowan, K. J. and Corwin, K. 2008. The second atlas of breeding birds in New York State. – Cornell Univ. Press.