Application of Spatially-Varying Coefficient Models

Jeffrey Doser[1] & Andrew Finley[2]

May 15, 2023

[1]Department of Integrative Biology, Michigan State University.

[2]Department of Forestry, Michigan State University.

Spatially-Varying Coefficient Models

- Extension of spatial regression approaches that allow regression coefficients to vary across space, and not just the intercept
- SVC models are random slopes models, with spatial structure given to the random slopes

SVC GLMMs

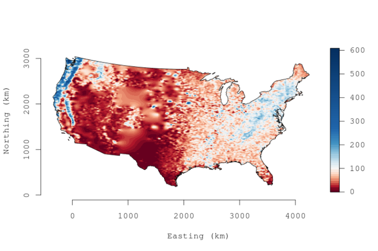
$$
\begin{aligned}
y(\mathbf{s}_i) &\sim f(\mu(\mathbf{s}_i), \tau) \\
\mu(\mathbf{s}_i) &= g^{-1}(\eta(\mathbf{s}_i)) = \mathbf{x}(\mathbf{s}_i)^\top \tilde{\beta}(\mathbf{s}_i) \\
\tilde{\beta}_r(\mathbf{s}_i) &= \beta_r + \mathrm{w}_r(\mathbf{s}_i) \text{ for each } r = 1, \dots, p
\end{aligned}
$$

- We can model $\mathbf{w}(s_i)$ using a GP, predictive process, or NNGP
- We can envision modeling $\mathbf{w}(s_i)$ in two ways:
    1. Multivariate NNGP (see previous forest biomass example)
    2. Independent NNGPs
- Here we focus on the latter
- Pros and cons to both approaches, similar to correlations between random slopes and intercepts in mixed models

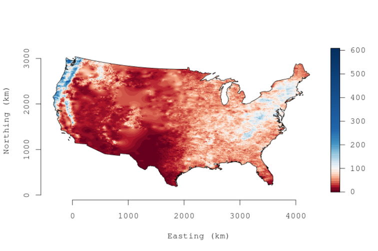
Potential benefits of SVC models

- Improved predictive performance
- Tremendous flexibility to accommodate spatial variability in effects
- Hypothesis testing and generation
- Accommodate highly non-linear relationships
- Model spatial variability in trends over time

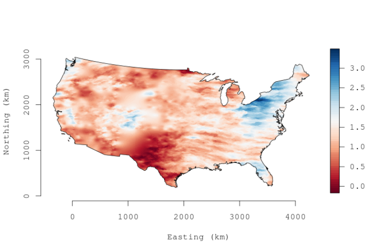
Improved predictive performance



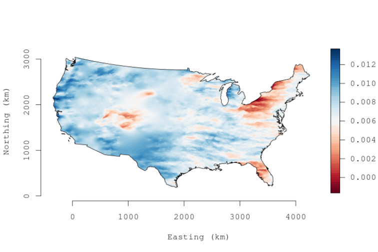
Observed biomass



Fitted biomass



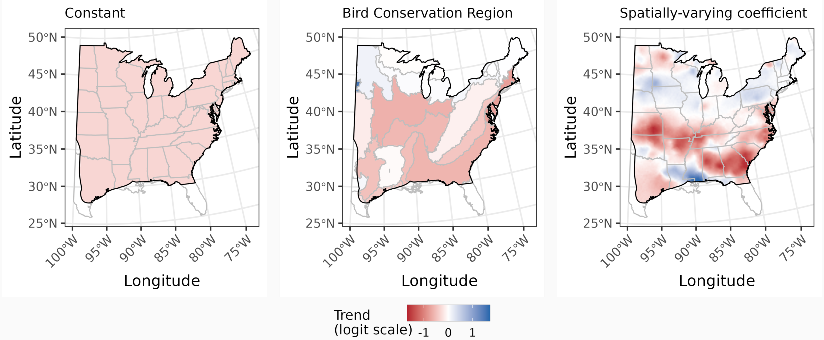
$\beta_0(s)$



$\beta_{NDVI}(s)$

More flexibility to accommodate spatial variability in effects

Gray Catbird occurrence trend across the eastern US from 2000-2019



Some cautionary notes: (1) Standardization of covariates

- Common recommendation to center and scale covariates is likely not appropriate. Why?

Some cautionary notes: (1) Standardization of covariates

- Common recommendation to center and scale covariates is likely not appropriate. Why?
- Consider a single SVC $\mathrm{w}(\mathbf{s}_i)$:

Some cautionary notes: (1) Standardization of covariates

- Common recommendation to center and scale covariates is likely not appropriate. Why?
- Consider a single SVC $\mathrm{w}(\mathbf{s}_i)$:
  - $\mathrm{var}(\mathrm{w}(\mathbf{s}_i)) = \mathrm{x}^2(\mathbf{s}_i)\sigma^2 + \tau^2$

Some cautionary notes: (1) Standardization of covariates

- Common recommendation to center and scale covariates is likely not appropriate. Why?
- Consider a single SVC $\mathrm{w}(\mathbf{s}_i)$:
  - $\mathrm{var}(\mathrm{w}(\mathbf{s}_i)) = \mathrm{x}^2(\mathbf{s}_i)\sigma^2 + \tau^2$
  - $\mathrm{cov}(\mathrm{w}(\mathbf{s}_i), \mathrm{w}(\mathbf{s}_j)) = \sigma^2\mathrm{x}(\mathbf{s}_i)\mathrm{x}(\mathbf{s}_j)\rho(\mid \mathbf{s}_i - \mathbf{s}_j \mid; \phi)$

Some cautionary notes: (1) Standardization of covariates

- Common recommendation to center and scale covariates is likely not appropriate. Why?
- Consider a single SVC $\mathrm{w}(\mathbf{s}_i)$:
  - $\mathrm{var}(\mathrm{w}(\mathbf{s}_i)) = \mathrm{x}^2(\mathbf{s}_i)\sigma^2 + \tau^2$
  - $\mathrm{cov}(\mathrm{w}(\mathbf{s}_i), \mathrm{w}(\mathbf{s}_j)) = \sigma^2 \mathrm{x}(\mathbf{s}_i)\mathrm{x}(\mathbf{s}_j)\rho(\mid \mathbf{s}_i - \mathbf{s}_j \mid; \phi)$
  - If x takes both positive and negative values, variance will be high at large and small x values and small around zero.

Some cautionary notes: (1) Standardization of covariates

- Common recommendation to center and scale covariates is likely not appropriate. Why?
- Consider a single SVC $\text{w}(\mathbf{s}_i)$:
  - $\text{var}(\text{w}(\mathbf{s}_i)) = \text{x}^2(\mathbf{s}_i)\sigma^2 + \tau^2$
  - $\text{cov}(\text{w}(\mathbf{s}_i), \text{w}(\mathbf{s}_j)) = \sigma^2 \text{x}(\mathbf{s}_i)\text{x}(\mathbf{s}_j)\rho(\mid \mathbf{s}_i - \mathbf{s}_j \mid; \phi)$
  - If x takes both positive and negative values, variance will be high at large and small x values and small around zero.
  - Further, negative covariances can arise, which may drive σ^2 to be very close to zero.

Some cautionary notes: (1) Standardization of covariates

- Common recommendation to center and scale covariates is likely not appropriate. Why?
- Consider a single SVC $\mathrm{w}(\mathbf{s}_i)$:
  - $\mathrm{var}(\mathrm{w}(\mathbf{s}_i)) = \mathrm{x}^2(\mathbf{s}_i)\sigma^2 + \tau^2$
  - $\mathrm{cov}(\mathrm{w}(\mathbf{s}_i), \mathrm{w}(\mathbf{s}_j)) = \sigma^2 \mathrm{x}(\mathbf{s}_i)\mathrm{x}(\mathbf{s}_j)\rho(\mid \mathbf{s}_i - \mathbf{s}_j \mid; \phi)$
  - If x takes both positive and negative values, variance will be high at large and small x values and small around zero.
  - Further, negative covariances can arise, which may drive σ^2 to be very close to zero.
- Recommendation following Gelfand et al. (2003) *JASA*: only use positive covariate values. Can lead to identifiability problems.
- For modeling spatially-varying trends, standardiziation is recommended.

Some cautionary notes: (2) Inference

- Estimating multiple GPs simultaneously is difficult.
- Use caution when making inference from SVC models, particularly with observational data.

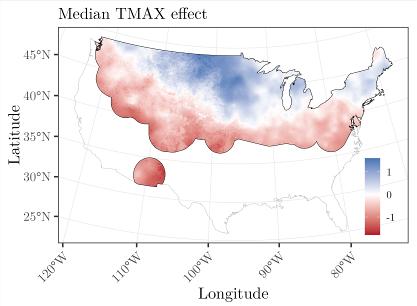
Some cautionary notes: (2) Inference

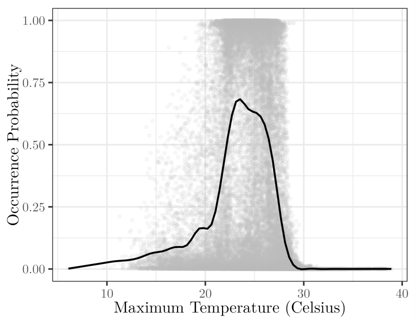
- Estimating multiple GPs simultaneously is difficult.
- Use caution when making inference from SVC models, particularly with observational data.
- The model only identifies the product of $\tilde{\beta}(\mathbf{s}_i)$ and $\mathrm{x}(\mathbf{s}_i)$. Prediction at $\mathbf{s}_i$ with a new covariate value (e.g., $\mathrm{x}^*(s_i)$) may not be appropriate.

Some cautionary notes: (2) Inference

- Estimating multiple GPs simultaneously is difficult.
- Use caution when making inference from SVC models, particularly with observational data.
- The model only identifies the product of $\tilde{\beta}(\mathbf{s}_i)$ and $\mathrm{x}(\mathbf{s}_i)$. Prediction at $\mathbf{s}_i$ with a new covariate value (e.g., $\mathrm{x}^*(s_i)$) may not be appropriate.
- Recommend plotting spatial maps of SVC effects together with a plot of the covariate vs. the predicted mean for all observed locations
- Interpretation can be difficult when both the effect and the covariate vary over space. More straightforward with something like a trend.

Example: Effect of max temperature on Bobolink occurrence



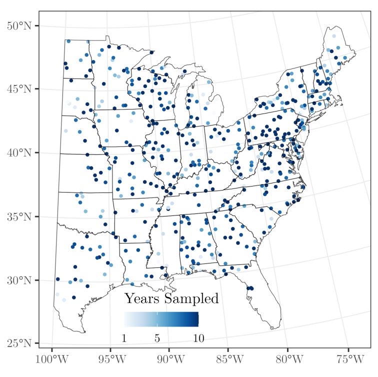


Software

- `spBayes`: univariate Gaussian SVC with full GPs
- `spOccupancy`: univariate Binomial SVC with NNGPs (multivariate on its way)
- `varycoef`: maximum likelihood Gaussian SVCs (Dambon et al. 2021 *Spatial Stats.*)
- `sdmTMB`: penalized likelihood and Bayesian SVC GLMMs (Anderson et al. 2022 *bioRxiv*)

Exercise: 10-year occurrence trend of Wood Thrush

- Data come from USGS North American Breeding Bird Survey
- We desire to account for observational biases in detection of the birds (i.e., false negatives).
- Add on an additional observational layer to our hierarchical model



Exercise: Process model

- Let $z_t(\mathbf{s}_i)$ denote the true presence (1) or absence (0) of the species at site $i = 1, \ldots, 500$ during year $t = 1, \ldots, 10$.

Exercise: Process model

- Let $z_t(\mathbf{s}_i)$ denote the true presence (1) or absence (0) of the species at site $i = 1, \ldots, 500$ during year $t = 1, \ldots, 10$.
- If the bird was detected at a site and year, we know $z_t(\mathbf{s}_i) = 1$. If not, it might be there and we just missed it during the surveys.

Exercise: Process model

- Let $z_t(\mathbf{s}_i)$ denote the true presence (1) or absence (0) of the species at site $i = 1, \ldots, 500$ during year $t = 1, \ldots, 10$.
- If the bird was detected at a site and year, we know $z_t(\mathbf{s}_i) = 1$. If not, it might be there and we just missed it during the surveys.
- We model $z_t(\mathbf{s}_i)$ just as before with a Bernoulli GLM, with a SVC for trend

$$z_t(\mathbf{s}_i) \sim \text{Bernoulli}(\psi_t(\mathbf{s}_i))$$
$$\text{logit}(\psi_t(\mathbf{s}_i)) = \tilde{\beta}_0(\mathbf{s}_i) + \tilde{\beta}_1(\mathbf{s}_i) \cdot \text{YEAR}_t$$

- $\tilde{\beta}_0(s_i)$ and $\tilde{\beta}_1(\mathbf{s}_i)$ are modeled as independent SVCs with NNGPs

Exercise: Observation model

- Let $y_{t,k}(\mathbf{s}_i)$ denote the observed detection (1) or nondetection (0) of the bird at site i during year t and survey $k = 1, \ldots, 5$.
- We model $y_{t,k}(\mathbf{s}_i)$ conditional on the true presence/absence of the species $z_t(\mathbf{s}_i)$

$$y_{t,k}(\mathbf{s}_i) \mid z_t(\mathbf{s}_i) \sim \text{Bernoulli}(p_{i,t,k} \cdot z_t(\mathbf{s}_i))$$
$$\text{logit}(p_{i,t,k}) = \alpha_{0,t} + \alpha_1 \cdot \text{DAY}_{i,t,k} + \alpha_2 \cdot \text{DAY}^2_{i,t,k}$$

- A key assumption for identifiability is that $z_t(\mathbf{s}_i)$ does not change across the 5 replicate surveys at site i during year t.