Hierarchical spatial modelling for applied population and community ecology

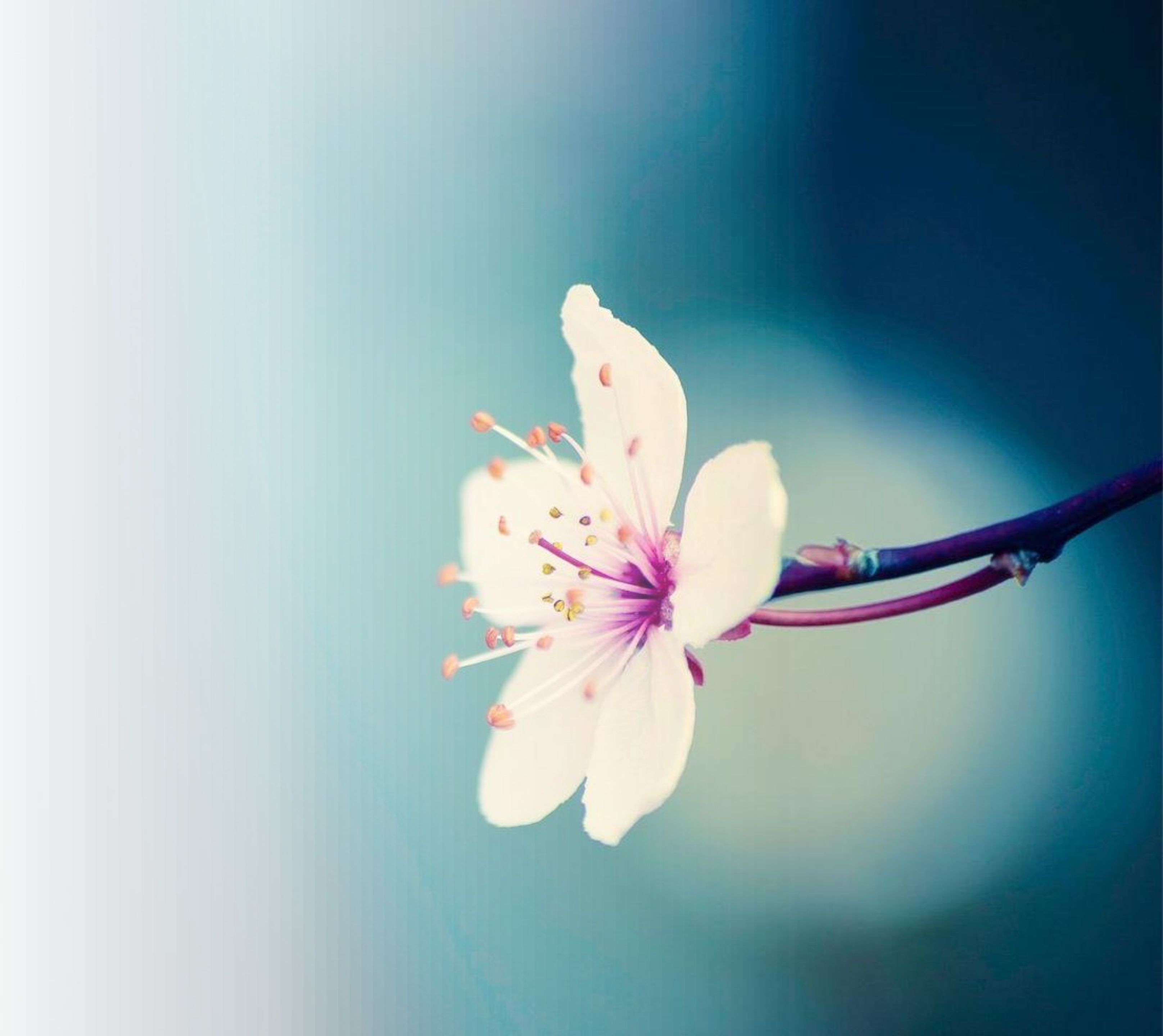
Jeffrey W. Doser, Marc Kéry, Gesa von Hirschheydt

24-27 June 2024

Spatial multi-species occupancy models

Jeffrey W. Doser

24-27 June 2024



Spatial autocorrelation in MSOMs

- Spatial autocorrelation may be more prevalent in MSOMs since different species are driven by different habitat requirements.
- Likely won't be able to include all covariates relevant to all species in the model.
- Can extend the MSOM just like we did with the SSOM to account for spatial autocorrelation in occupancy probability for each species.
- Two "flavors" of spatial MSOMs included in `spOccupancy`

Type 1: Spatial MSOMs

- Estimate a separate spatial random effect **w** for each species
- Each **w** is estimated with a NNGP as before
- Nothing new here, now just have N (the number of species) NNGP spatial random effects
- Would very rarely want to use a full GP instead of an NNGP, since we now need to estimate N spatial random intercepts
- `spMsPGOcc()` function

Type 1 Spatial MSOMs

$$
\begin{aligned}
z_{i,j} &\sim \text{Bernoulli}(\psi_{i,j}) \\
\text{logit}(\psi_{i,j}) &= \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j} + \text{w}_{i,j} \\
\beta_{r,i} &\sim \text{Normal}(\mu_{\beta_r}, \tau^2_{\beta_r}) \\
\mathbf{w}_i &\sim \text{Normal}(\mathbf{0}, \tilde{\boldsymbol{C}}(d, \phi_i, \sigma^2_i) \\
\mu_{\beta_r} &\sim \text{Normal}(\mu 0_r, \sigma^2_{\beta_r}) \\
\tau^2_{\beta_r} &\sim \text{Inverse gamma}(a_{r,\beta}, b_{r,\beta}) \\
\phi_i &\sim \text{Uniform}(a_{\phi,i}, b_{\phi,i}) \\
\sigma^2_i &\sim \text{Inverse gamma}(a_{\sigma,i}, b_{\sigma,i})
\end{aligned}
$$

Type 1 Spatial MSOMs

$$
\begin{aligned}
z_{i,j} &\sim \text{Bernoulli}(\psi_{i,j}) \\
\text{logit}(\psi_{i,j}) &= \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j} + \boxed{\text{w}_{i,j}} \\
\beta_{r,i} &\sim \text{Normal}(\mu_{\beta_r}, \tau^2_{\beta_r}) \\
\boxed{\mathbf{w}_i \sim \text{Normal}(\mathbf{0}, \tilde{\boldsymbol{C}}(d, \phi_i, \sigma^2_i)} \\
\mu_{\beta_r} &\sim \text{Normal}(\mu 0_r, \sigma^2_{\beta_r}) \\
\tau^2_{\beta_r} &\sim \text{Inverse gamma}(a_{r,\beta}, b_{r,\beta}) \\
\phi_i &\sim \text{Uniform}(a_{\phi,i}, b_{\phi,i}) \\
\sigma^2_i &\sim \text{Inverse gamma}(a_{\sigma,i}, b_{\sigma,i})
\end{aligned}
$$

Type 1 Spatial MSOMs

$$
\begin{aligned}
z_{i,j} &\sim \text{Bernoulli}(\psi_{i,j}) \\
\text{logit}(\psi_{i,j}) &= \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j} + \boxed{\mathrm{w}_{i,j}} \\
\beta_{r,i} &\sim \text{Normal}(\mu_{\beta_r}, \tau^2_{\beta_r}) \\
\boxed{\mathbf{w}_i \sim \text{Normal}(\mathbf{0}, \tilde{\boldsymbol{C}}(d, \phi_i, \sigma^2_i))} \\
\mu_{\beta_r} &\sim \text{Normal}(\mu 0_r, \sigma^2_{\beta_r}) \\
\tau^2_{\beta_r} &\sim \text{Inverse gamma}(a_{r,\beta}, b_{r,\beta}) \\
\phi_i &\sim \text{Uniform}(a_{\phi,i}, b_{\phi,i}) \\
\sigma^2_i &\sim \text{Inverse gamma}(a_{\sigma,i}, b_{\sigma,i})
\end{aligned}
$$

The detection sub-model is exactly the same as the nonspatial MSOM

Problems with Type 1 spatial MSOMs

- Can become quite slow when the number of species in the community is greater than 10 or so
- Species-specific spatial random effects can be very difficult to estimate for rare species
- Most useful when working with a modest number of spatial locations (e.g. 100-500) and a modest number of species (e.g., <10) that are not extremely rare.

Type 2: Spatial factor MSOMs

- More powerful and efficient than the previously discussed spatial MSOMs
- A multi-species modeling approach that simultaneously accounts for:
  - Imperfect detection
  - Spatial autocorrelation
  - Species correlations

Received: 6 April 2022 | Revised: 26 May 2023 | Accepted: 13 June 2023
DOI: 10.1002/ecy.4137

STATISTICAL REPORT

ECOLOGY
ECOLOGICAL SOCIETY OF AMERICA

Joint species distribution models with imperfect detection for high-dimensional spatial data

Jeffrey W. Doser[1,2] | Andrew O. Finley[2,3] | Sudipto Banerjee[4]

Spatial factor MSOMs

- An extension of the latent factor MSOM we discussed previously
- Identical to the latent factor MSOM, but now the factors are estimated as spatial NNGPs!
- **Foundational concept**: spatial autocorrelation in the distributions of individual species may arise from a set of shared, common factors that have a spatial pattern
- Only difference from the previous spatial MSOM is in how the spatial random effects are modelled

Spatial factor MSOMs

$$z_{i,j} \sim \text{Bernoulli}(\psi_{i,j})$$
$$\text{logit}(\psi_{i,j}) = \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j} + \text{w}^*_{i,j}$$

Spatial factor MSOMs

Species-specific spatial random effect

$$z_{i,j} \sim \text{Bernoulli}(\psi_{i,j})$$

$$\text{logit}(\psi_{i,j}) = \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j} + \text{w}^*_{i,j}$$

Spatial factor MSOMs

Species-specific spatial random effect

$$z_{i,j} \sim \text{Bernoulli}(\psi_{i,j})$$
$$\text{logit}(\psi_{i,j}) = \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j} + \text{w}^*_{i,j}$$
$$\text{w}^*_{i,j} = \lambda_{i,1} \cdot \text{w}_{1,j} + \lambda_{i,2} \cdot \text{w}_{2,j} + \cdots + \lambda_{i,q} \cdot \text{w}_{q,j}$$

Species specific spatial effect is the sum of a set of q spatial factors and their species-specific coefficients (loadings)

Spatial factor MSOMs

Species-specific spatial random effect

$$z_{i,j} \sim \text{Bernoulli}(\psi_{i,j})$$

$$\text{logit}(\psi_{i,j}) = \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j} + \text{w}^*_{i,j}$$

$$\text{w}^*_{i,j} = \lambda_{i,1} \cdot \text{w}_{1,j} + \lambda_{i,2} \cdot \text{w}_{2,j} + \cdots + \lambda_{i,q} \cdot \text{w}_{q,j}$$

Species specific spatial effect is the sum of a set of q spatial factors and their species-specific coefficients (loadings)

"Missing covariates" that account for residual spatial autocorrelation

Spatial factor MSOMs

Species-specific spatial random effect

$$z_{i,j} \sim \text{Bernoulli}(\psi_{i,j})$$

$$\text{logit}(\psi_{i,j}) = \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j} + \text{w}^*_{i,j}$$

$$\text{w}^*_{i,j} = \lambda_{i,1} \cdot \text{w}_{1,j} + \lambda_{i,2} \cdot \text{w}_{2,j} + \cdots + \lambda_{i,q} \cdot \text{w}_{q,j}$$

Species specific spatial effect is the sum of a set of q spatial factors and their species-specific coefficients (loadings)

Effects of the missing covariates

"Missing covariates" that account for residual spatial autocorrelation

Spatial factor MSOMs

$$z_{i,j} \sim \text{Bernoulli}(\psi_{i,j})$$
$$\text{logit}(\psi_{i,j}) = \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j} + \text{w}^*_{i,j}$$
$$\text{w}^*_{i,j} = \lambda_{i,1} \cdot \text{w}_{1,j} + \lambda_{i,2} \cdot \text{w}_{2,j} + \cdots + \lambda_{i,q} \cdot \text{w}_{q,j}$$

Identical to latent factor MSOMs, except we now model each w with a NNGP instead of from a Normal(0, 1).

Priors and identifiability considerations

- Same priors as we've seen before
- Normal priors on community-level means
- Inverse-gamma priors on community-level variances
- Uniform priors on spatial decay parameters
- Same restrictions on the factor loadings matrix as with latent factor models

Why use the spatial factor MSOM?

Performs well in a variety of scenarios while being more computationally efficient than alternative spatial MSOMs.

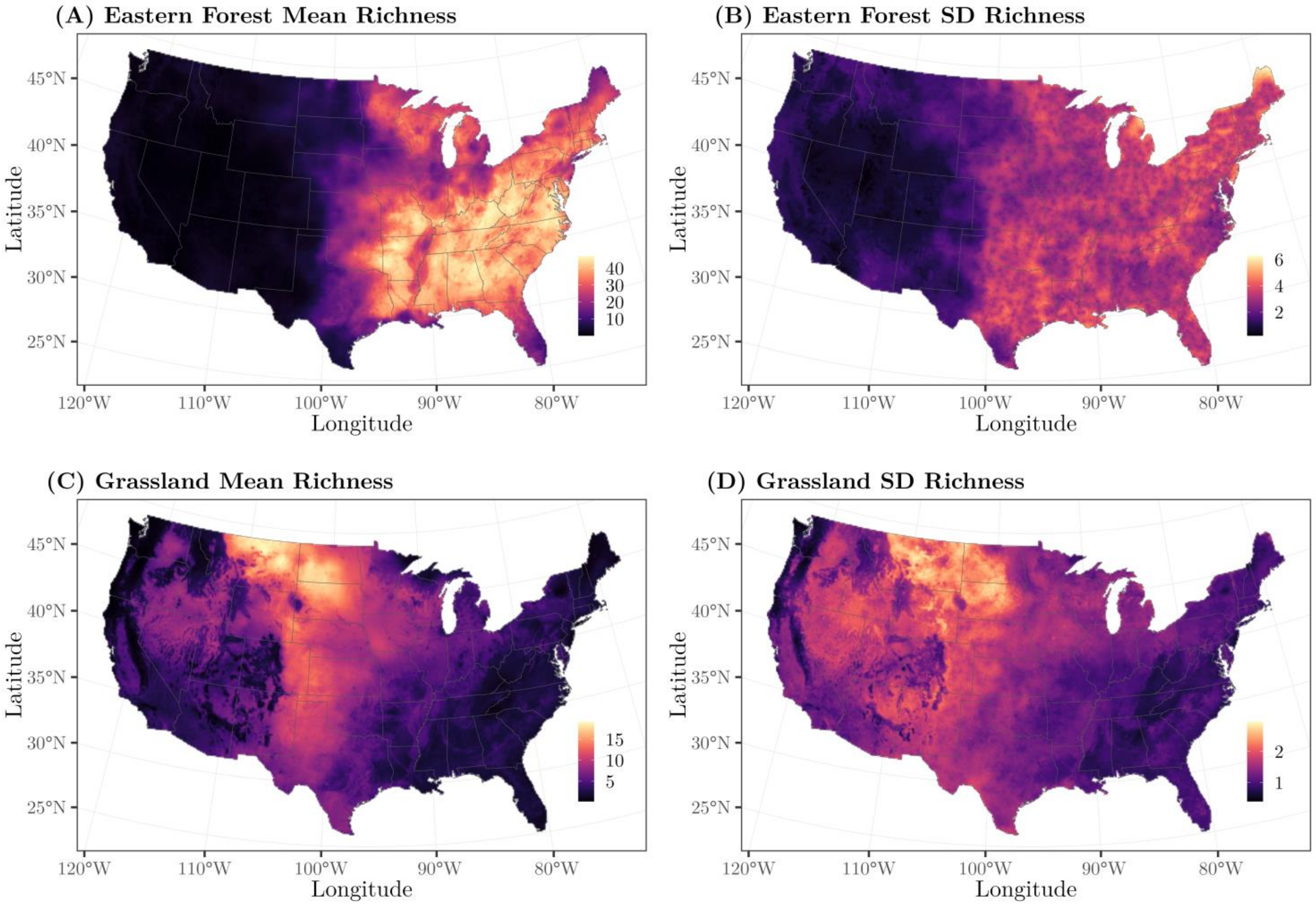
TABLE 2 Estimated coverage rates of simulated species-specific occurrence probabilities and covariate effects for six different simulation scenarios and six models of varying complexity, as well as average run time.

		Model					
Parameter	**Scenario**	**lfJSDM**	**sfJSDM**	**msPGOcc**	**spMsPGOcc**	**lfMsPGOcc**	**sfMsPGOcc**
$\psi_i(\boldsymbol{s}_j)$	1	91.5	90.8	68.9	88.1	95.6	95.3
	2	91.6	91.0	69.1	89.1	95.5	95.4
	3	85.6	84.8	95.0	96.4	95.5	95.5
	4	77.5	76.4	80.2	93.1	95.7	95.5
	5	75.3	74.2	71.3	88.5	95.5	95.3
	6	76.0	75.0	72.2	89.6	95.3	95.2
β_i	1	88.7	88.2	82.0	91.1	95.2	95.1
	2	88.8	88.2	82.2	91.7	94.9	94.9
	3	73.8	73.1	95.1	94.4	90.4	90.8
	4	65.9	65.0	89.1	94.0	94.7	94.7
	5	64.2	63.6	83.6	91.7	95.2	95.0
	6	65.7	64.6	85.1	92.7	94.9	94.9

Doser, Finley, Banerjee (2023) Ecology

Why use the spatial factor MSOM?

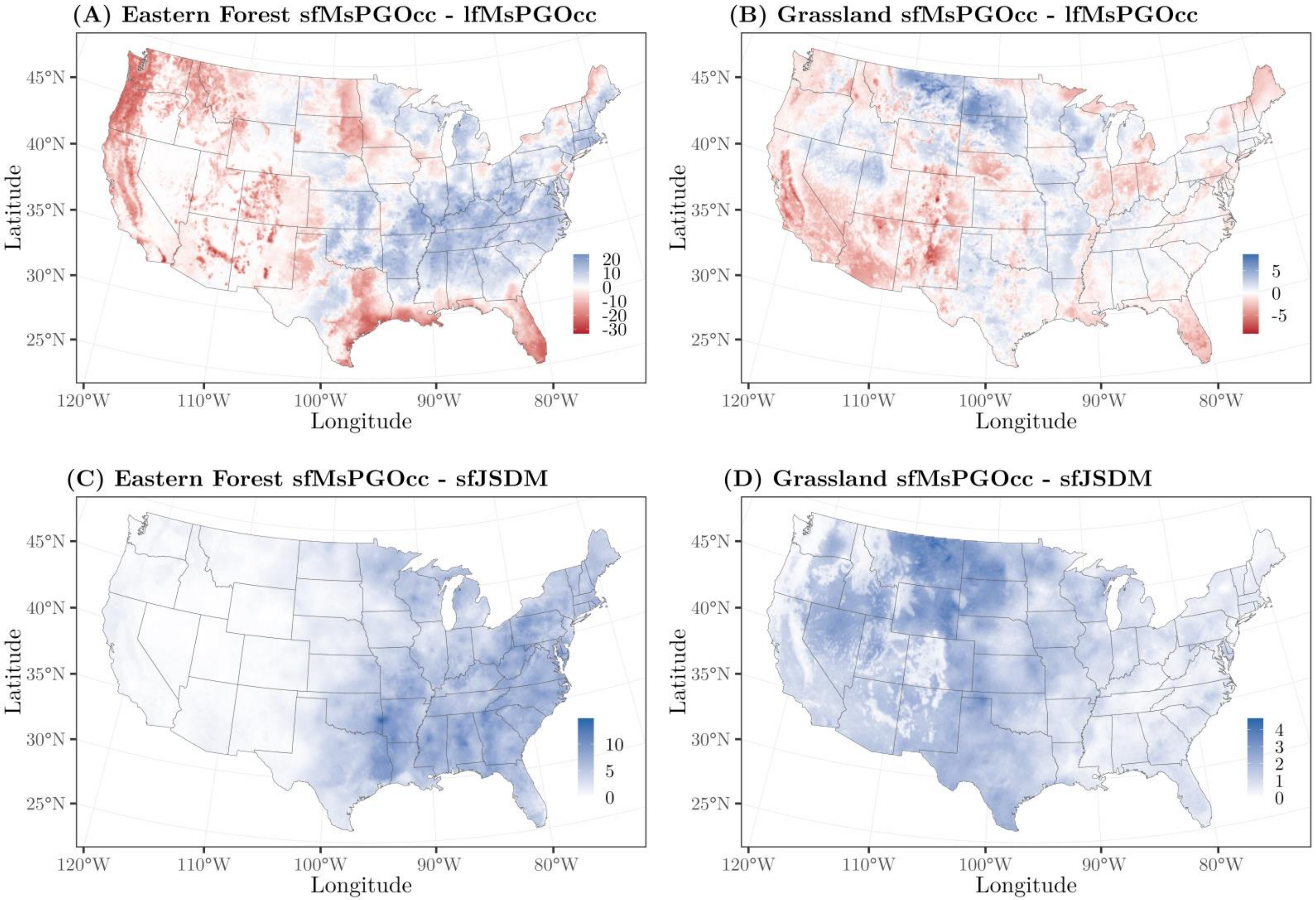
Estimate distributions and biodiversity metrics with big multi-species data sets (ex: 2,619 sites and 98 species)



Doser, Finley, Banerjee (2023) Ecology

Why use the spatial factor MSOM?

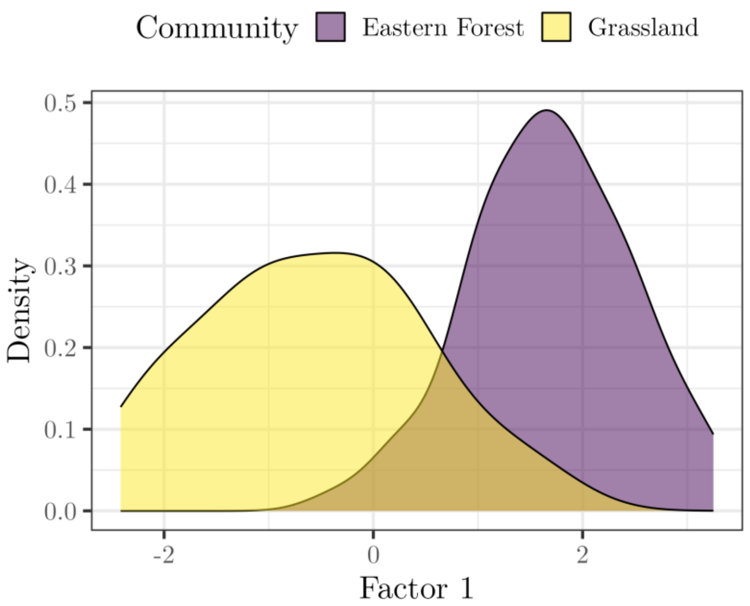
Improved estimates of biodiversity metrics

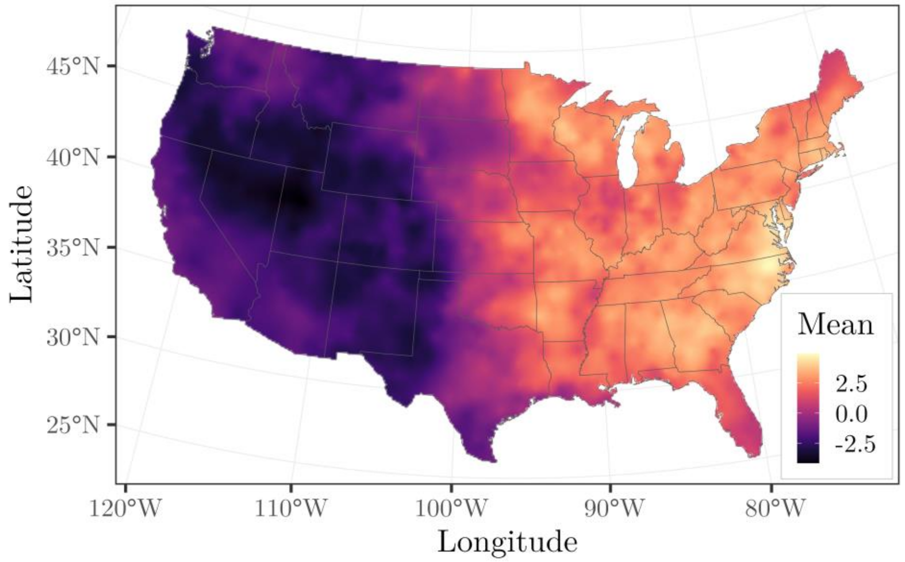


Doser, Finley, Banerjee (2023) Ecology

Why use the spatial factor MSOM?

Model-based ordination





Doser, Finley, Banerjee (2023) Ecology

Fitting spatial factor MSOMs in `spOccupancy`

- `sfMsPGOcc()`
- As with all spatial models, we break up the MCMC iterations into a set of batches, where each batch has a set number of iterations.
- Need to specify number of spatial factors to use

Multi-species, single-season models

spOccupancy function	Species correlations	Spatial autocorrelation	Imperfect detection
lfJSDM	✓		
sfJSDM	✓	✓	
msPGOcc			✓
spMsPGOcc		✓	✓
lfMsPGOcc	✓		✓
sfMsPGOcc	✓	✓	✓

Exercise: Niche partitioning in a simulated plant community

07-plant-spatial-multi-species-occ.R