Hierarchical spatial modelling for applied population and community ecology

Jeffrey W. Doser, Marc Kéry,
Gesa von Hirschheydt

24-27 June 2024

Spatial multi-species N-mixture models

Jeffrey W. Doser

24-27 June 2024

Multi-species N-mixture models

- Many forms of transect surveys or point count surveys collect data on multiple species.

- Given the similarities in data collection protocols with occupancy models, we can extend the multi-species occupancy model to an N-mixture framework.

Biodiversity of man-made open habitats in an underused country: a class of multispecies abundance models for count data

Yuichi Yamaura · J. Andrew Royle · Naoaki Shimada · Seigo Asanuma · Tamotsu Sato · Hisatomo Taki · Shun'ichi Makino

Modelling community dynamics based on species-level abundance models from detection/nondetection data

Yuichi Yamaura[1]*‡, J. Andrew Royle[2], Kouji Kuboi[3]†, Tsuneo Tada[3], Susumu Ikeno[4] and Shun'ichi Makino[1]

Multi-species count data from multiple visits

i = 1, 2, …, I species

j = 1, 2, …, J sites

k = 1, 2, …, K visits

Visit 3

Species	Site 1	Site 2	Site 3	Site 4
A	2	1	0	NA
B	1	0	2	NA
C	4	2	0	NA
D	3	0	0	NA
E	1	1	2	NA
F	0	1	0	NA

Visit 2

Species	Site 1	Site 2		
A	0	NA		
B	1	NA		
C	2	NA		
D	1	NA		
E	0	NA	1	2
F	1	NA	1	2

Visit 1

Species	Site 1	Site 2		
A	2	1		
B	0	1		
C	3	1		
D	1	1		
E	0	2	2	1
F	0	0	0	3

Three types of multi-species N-mixture models in `spAbundance`

1. Multi-species N-mixture models (`msNMix`)
   - o Species are modelled hierarchically (as random effects)
   - o No residual species correlations

Three types of multi-species N-mixture models in `spAbundance`

1. Multi-species N-mixture models (`msNMix`)
    o   Species are modelled hierarchically (as random effects)
    o   No residual species correlations

2. Latent factor multi-species N-mixture models (`lfMsNMix`)
    o Species are modelled hierarchically (as random effects)
    o Residual species correlations

Three types of multi-species N-mixture models in `spAbundance`

1. Multi-species N-mixture models (`msNMix`)
   - o Species are modelled hierarchically (as random effects)
   - o No residual species correlations

2. Latent factor multi-species N-mixture models (`lfMsNMix`)
   - o Species are modelled hierarchically (as random effects)
   - o Residual species correlations

3. Spatial factor multi-species N-mixture models (`sfMsNMix`)
   - o Species are modelled hierarchically (as random effects)
   - o Residual species correlations
   - o Spatial autocorrelation

Why aren't there spatial multi-species models without species correlations in `spAbundance`?

Why aren't there spatial multi-species models without species correlations in `spAbundance`?

- Recall in `spOccupancy` there were two types of spatial multi-species occupancy models: `spMsPGOcc()` and `sfMsPGOcc()`.

Why aren't there spatial multi-species models without species correlations in `spAbundance`?

- Recall in `spOccupancy` there were two types of spatial multi-species occupancy models: `spMsPGOcc()` and `sfMsPGOcc()`.

- `spMsPGOcc()` fits a separate spatial random effect for each species.

Why aren't there spatial multi-species models without species correlations in `spAbundance`?

- Recall in `spOccupancy` there were two types of spatial multi-species occupancy models: `spMsPGOcc()` and `sfMsPGOcc()`.

- `spMsPGOcc()` fits a separate spatial random effect for each species.

- Because algorithms are less efficient for count data than binary data, such an approach becomes very slow for count data models.

Why aren't there spatial multi-species models without species correlations in `spAbundance`?

- Recall in `spOccupancy` there were two types of spatial multi-species occupancy models: `spMsPGOcc()` and `sfMsPGOcc()`.

- `spMsPGOcc()` fits a separate spatial random effect for each species.

- Because algorithms are less efficient for count data than binary data, such an approach becomes very slow for count data models.

- Further, simulations in Doser, Finley, Banerjee (2023) show the spatial factor occupancy model performs equally well as `spMsPGOcc()` even when there aren't residual species correlations.

Multi-species N-mixture models (`msNMix`)

$$N_{i,j} \sim \mathrm{Poisson}(\mu_{i,j})$$

$$\log(\mu_{i,j}) = \beta_{1,i} + \beta_{2,i} \cdot X_{2,j} + \cdots + \beta_{r,i} \cdot X_{r,j}$$

$$\beta_{1,i} \sim \mathrm{Normal}(\mu_{\beta,1}, \tau^2_{\beta,1})$$

$$y_{i,j,k} \sim \mathrm{Binomial}(N_{i,j}, p_{i,j,k})$$

$$\mathrm{logit}(p_{i,j,k}) = \alpha_{1,i} + \alpha_{2,i} \cdot V_{2,j,k} + \cdots + \alpha_{d,i} \cdot V_{d,j,k}$$

$$\alpha_{1,i} \sim \mathrm{Normal}(\mu_{\alpha,1}, \tau^2_{\alpha,1})$$

- Species-level effects are modelled as random intercepts/slopes from a community-level distribution
- Often yields improved precision due to "borrowing strength", particularly for rare species

Latent factor multi-species N-mixture models (`lfMsNMix`)

$$N_{i,j} \sim \text{Poisson}(\mu_{i,j})$$

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

$$\beta_{1,i} \sim \text{Normal}(\mu_{\beta,1}, \tau^2_{\beta,1})$$

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

$$w_{q,j} \sim \text{Normal}(0, 1)$$

$$y_{i,j,k} \sim \text{Binomial}(N_{i,j}, p_{i,j,k})$$

$$\text{logit}(p_{i,j,k}) = \alpha_{1,i} + \alpha_{2,i} \cdot V_{2,j,k} + \cdots + \alpha_{d,i} \cdot V_{d,j,k}$$

$$\alpha_{1,i} \sim \text{Normal}(\mu_{\alpha,1}, \tau^2_{\alpha,1})$$

- Account for species correlations through a set of latent factors.
- Latent factors arise from a standard normal distribution.

Spatial factor multi-species N-mixture models (`sfMsNMix`)

$$N_{i,j} \sim \text{Poisson}(\mu_{i,j})$$

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

$$\beta_{1,i} \sim \text{Normal}(\mu_{\beta,1}, \tau^2_{\beta,1})$$

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

$$\mathbf{w}_q \sim \text{Multivariate Normal}(\mathbf{0}, \tilde{\boldsymbol{C}}(d, \phi_q))$$

$$y_{i,j,k} \sim \text{Binomial}(N_{i,j}, p_{i,j,k})$$

$$\text{logit}(p_{i,j,k}) = \alpha_{1,i} + \alpha_{2,i} \cdot V_{2,j,k} + \cdots + \alpha_{d,i} \cdot V_{d,j,k}$$

$$\alpha_{1,i} \sim \text{Normal}(\mu_{\alpha,1}, \tau^2_{\alpha,1})$$

- Account for species correlations through a set of spatial factors.
- Spatial factors modelled with NNGPs.

Fitting multi-species abundance N-mixture models

- All the same prior constraints and convergence guidelines from multi-species occupancy models and GLMMs apply here.

- Abudance-based joint species distribution model that accounts for imperfect detection.

Fitting overdispersed multi-species N-mixture models is tricky!

- As we discussed with single-species models, multi-species spatial N-mixture models can be very difficult to fit in practice.

- Models may need to run for hundreds of thousands of MCMC iterations, and they still may mix poorly.

- Can be very sensitive to the default initial values.

- What can we do to make this process a bit easier?
  - Fit model for a short period of time to extract initial values for a longer model run.
  - Run multiple chains of the model in parallel by running R script through the terminal (with reasonable starting values based on the previous model run)

Exercise: Estimating abundance of four foliage-gleaning birds

12b-hbef-main-sfMsNMix.R