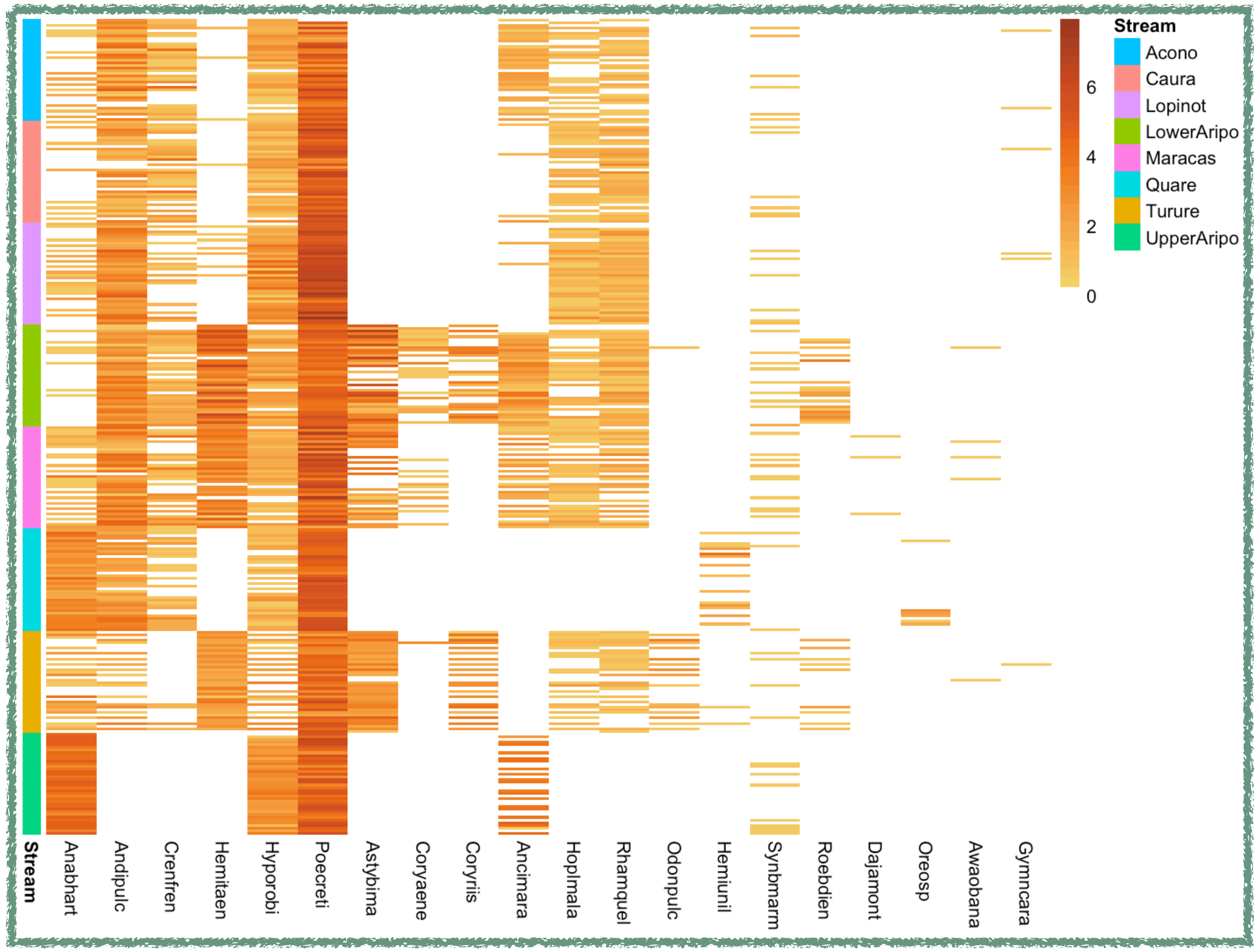
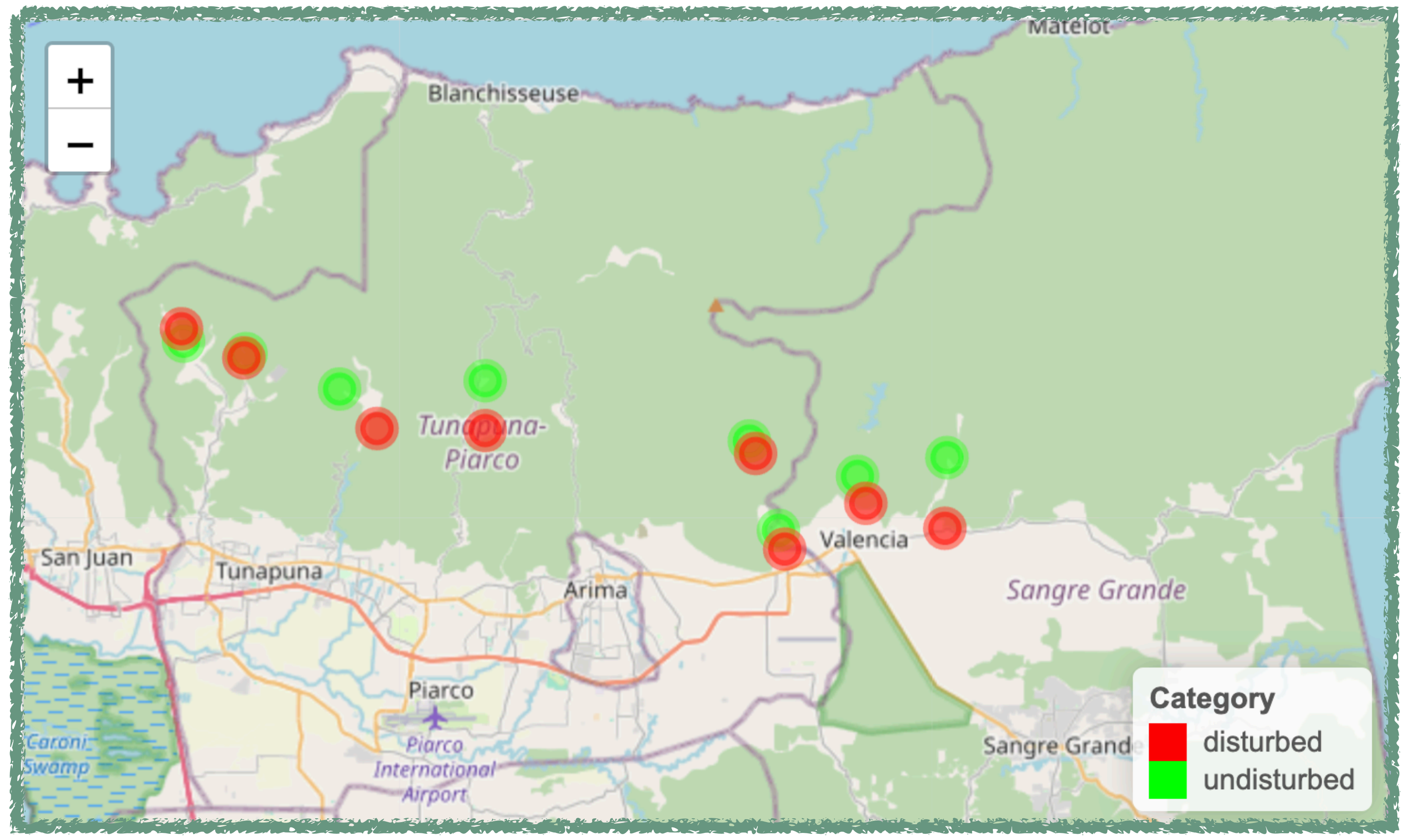
A simulation study to better understand the correspondence between biotic interactions and JSDM-inferred associations in species distribution data.

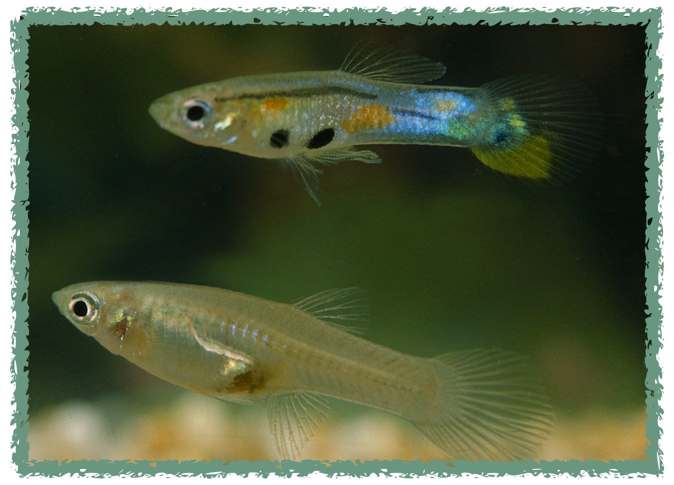
Questioning the relationships between ecological processes and statistical patterns in species distribution data.

Jeanne Tous - Statistiques au sommet de Rochebrune 2026 ⛰️

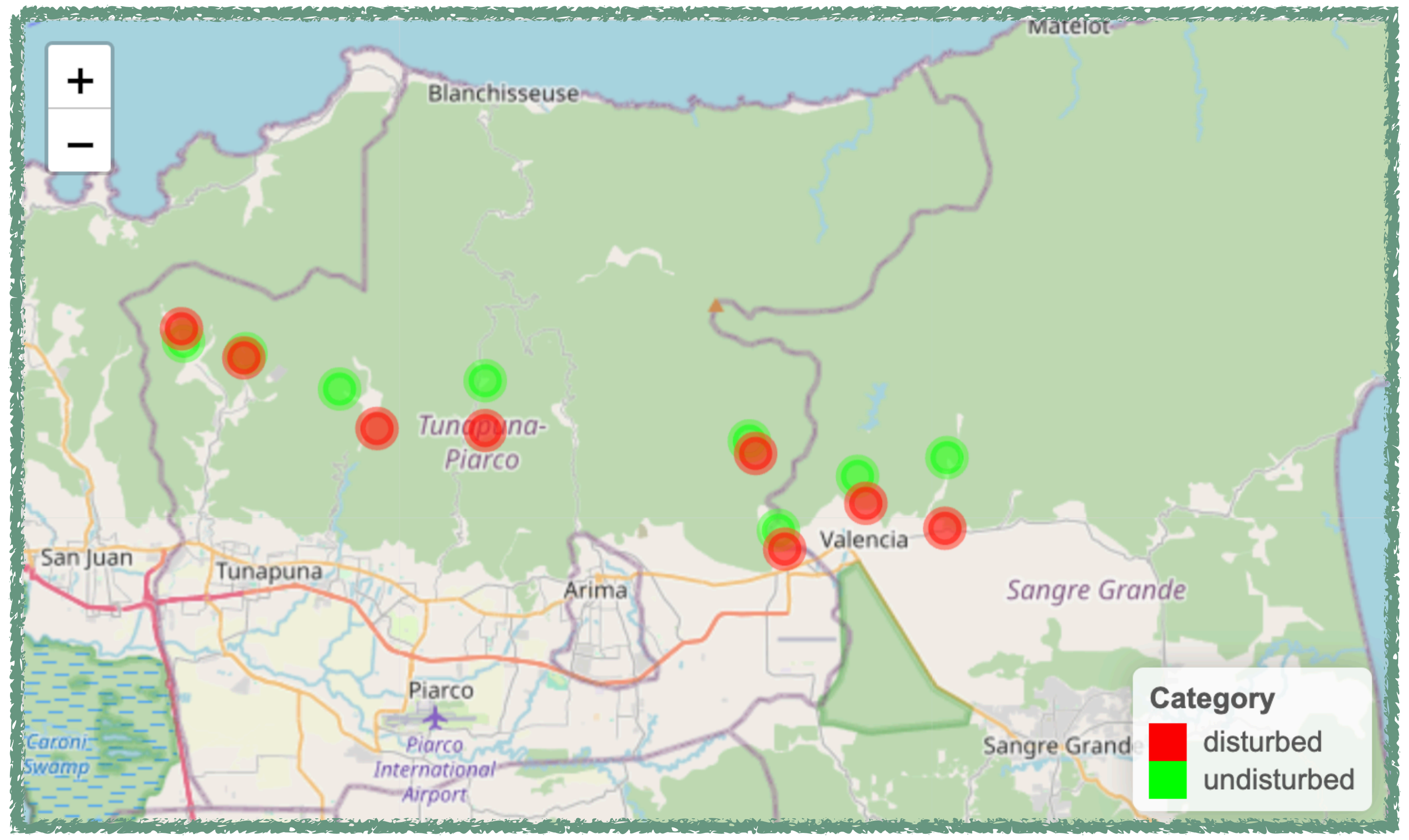
1. Species distribution data

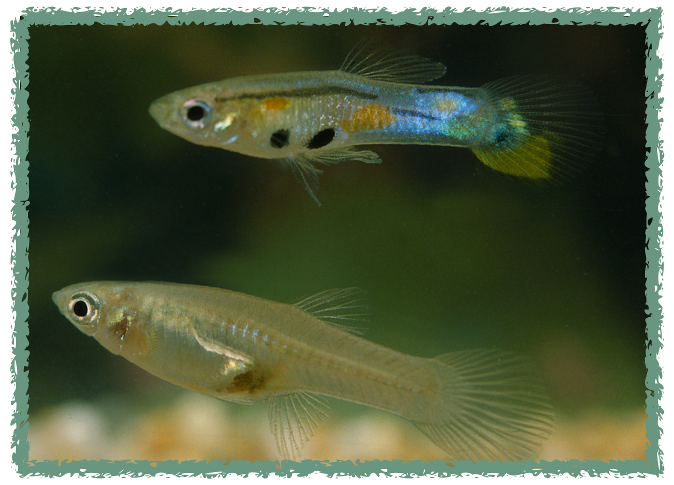


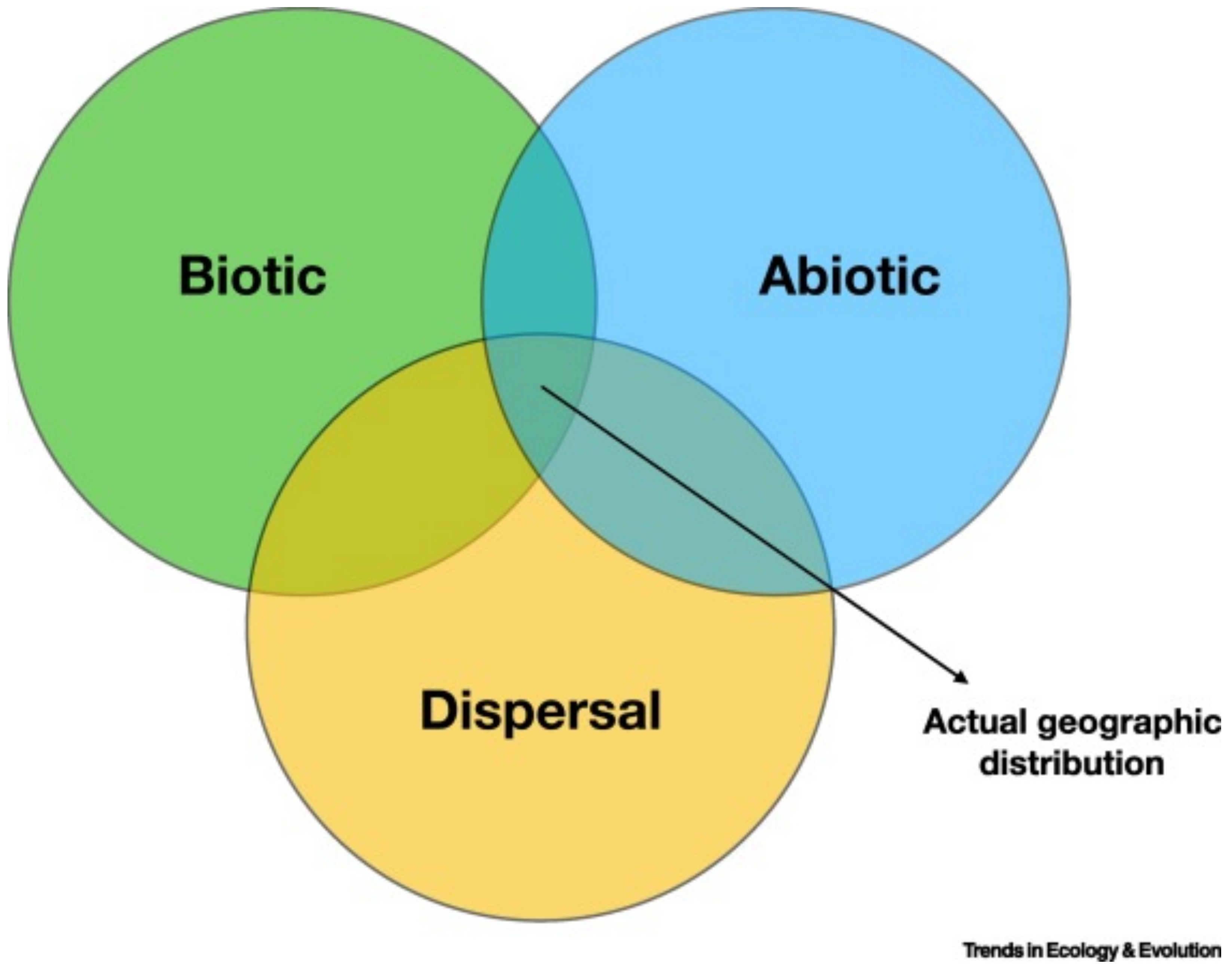




1. Species distribution data







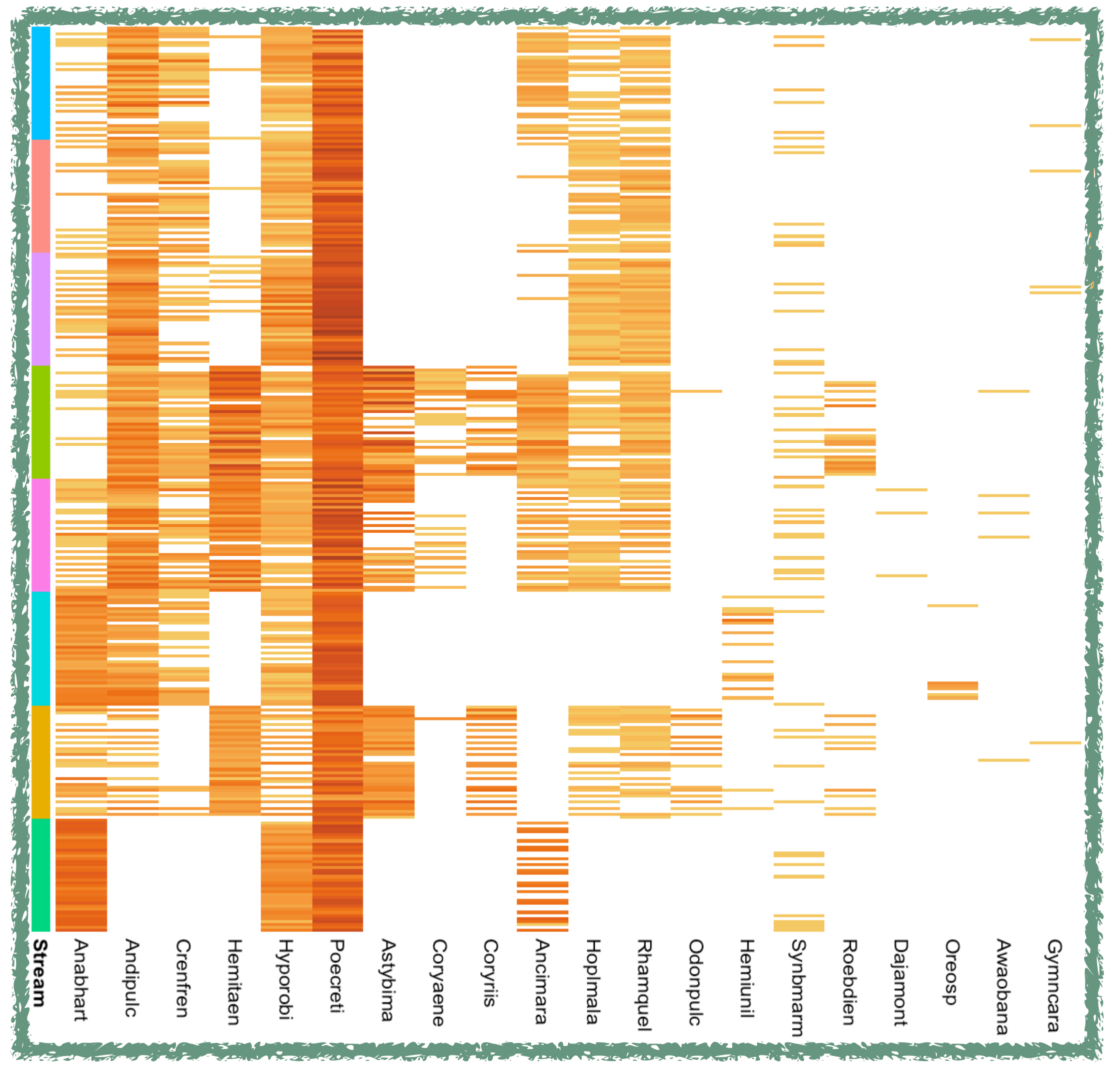
Poggiato et al. [2021]

2. Patterns, processes and mechanisms

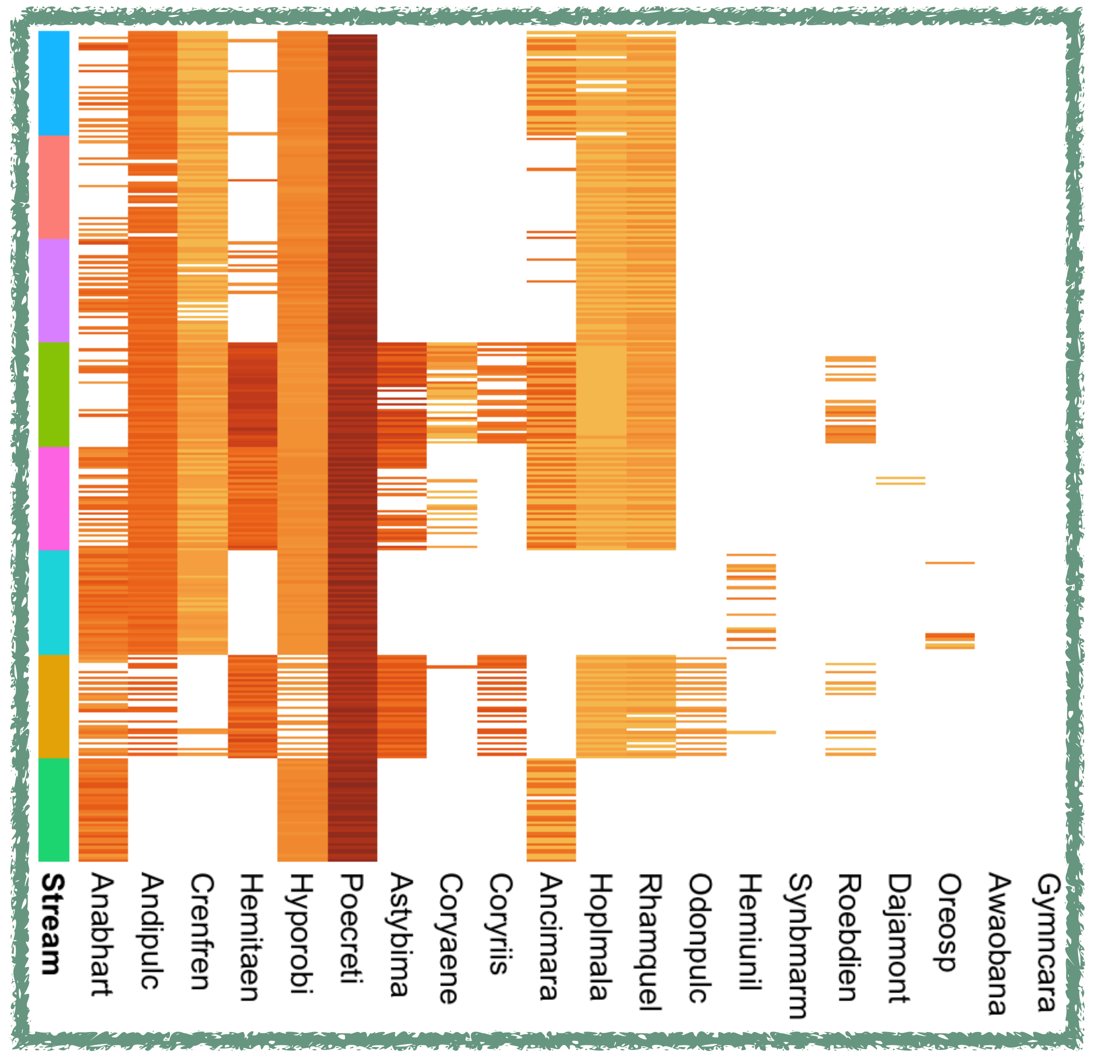
- Anand (1994) : *"patterns are what we perceive, processes describe how these patterns come about, and the mechanisms provide explanations as to why these patterns occur"*

- We often don't have a direct access to ecological processes / mechanisms.
- Old idea in ecology: identify patterns to understand processes.
    - Null models, SDMs, JSDMs…

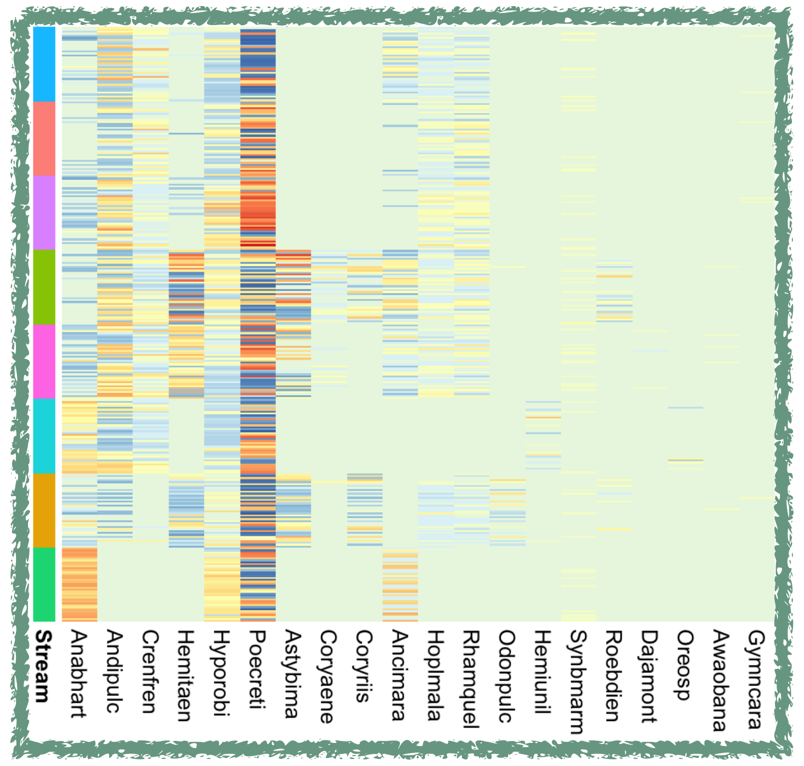
3. Joint Species Distribution Models (JSDMs)

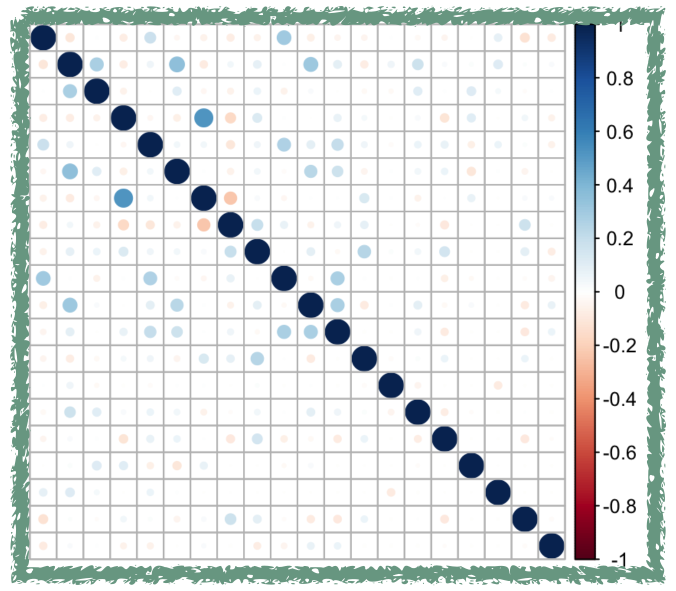


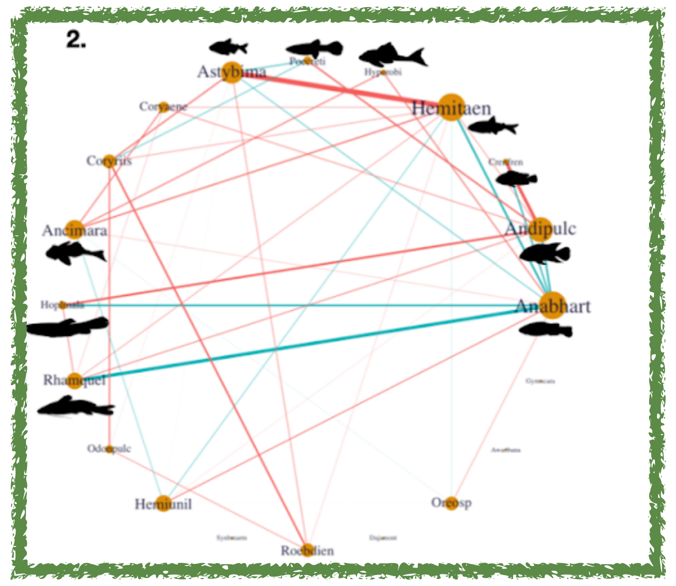
Explain what you can with abiotic covariates



Look at the residuals







3. Joint Species Distribution Models (JSDMs)

Graphical models = possible JSDMs

Ex: PLN-network [Chiquet, Robin and Mariadassou, 2019]

$$Z_i \sim \mathcal{N}(0, \Omega^{-1})$$

$$Y_{ij} | Z_{ij} \sim \mathcal{P}(\exp(X_i B_j + Z_{ij}))$$

3. Joint Species Distribution Models (JSDMs)

Graphical models = possible JSDMs

Ex: PLN-network [Chiquet, Robin and Mariadassou, 2019]

$$Z_i \sim \mathcal{N}(0, \Omega^{-1})$$

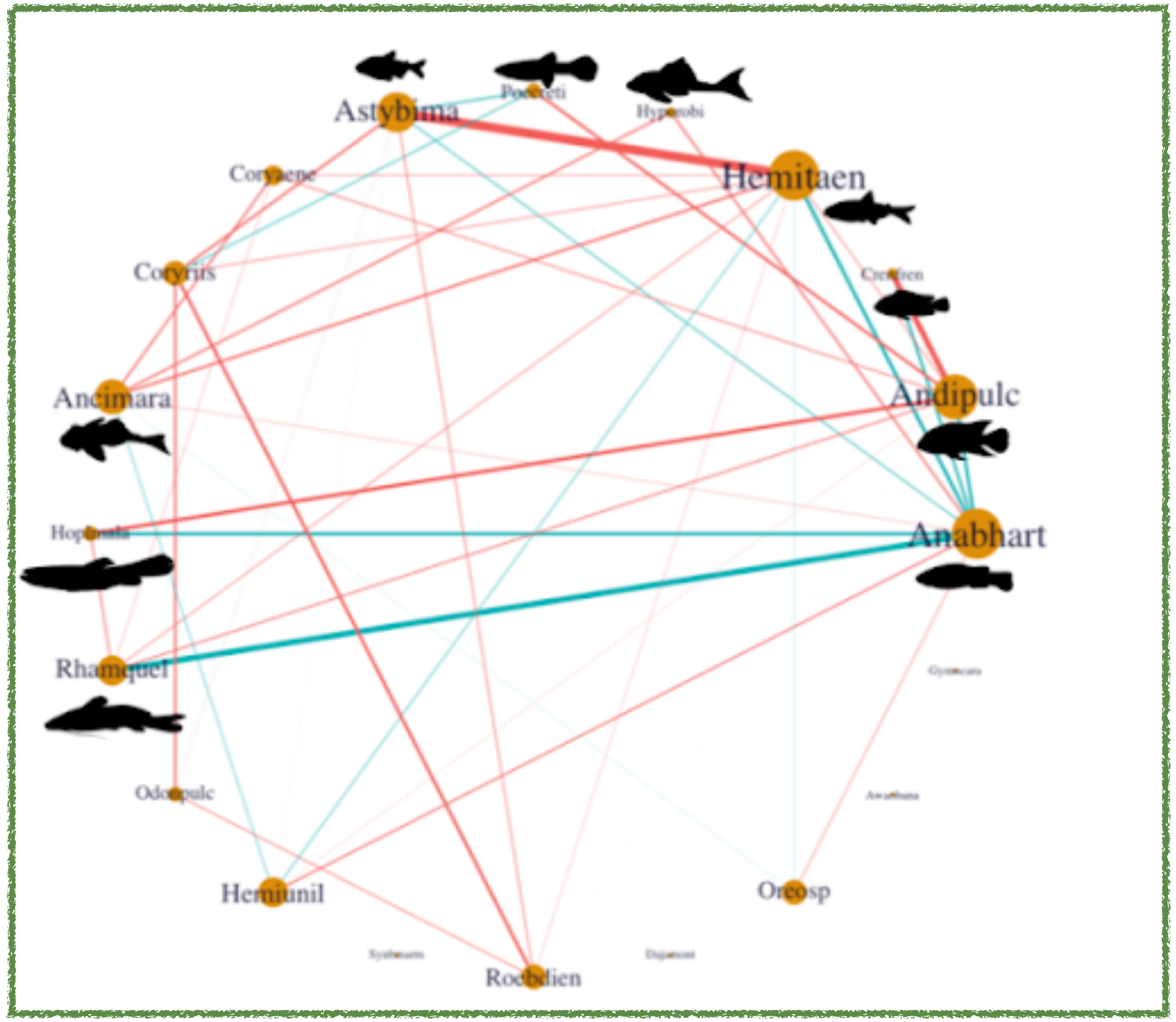
Residual partial correlations

$$Y_{ij} \mid Z_{ij} \sim \mathcal{P}(\exp(X_i B_j + Z_{ij}))$$

Count data

Measured abiotic effects

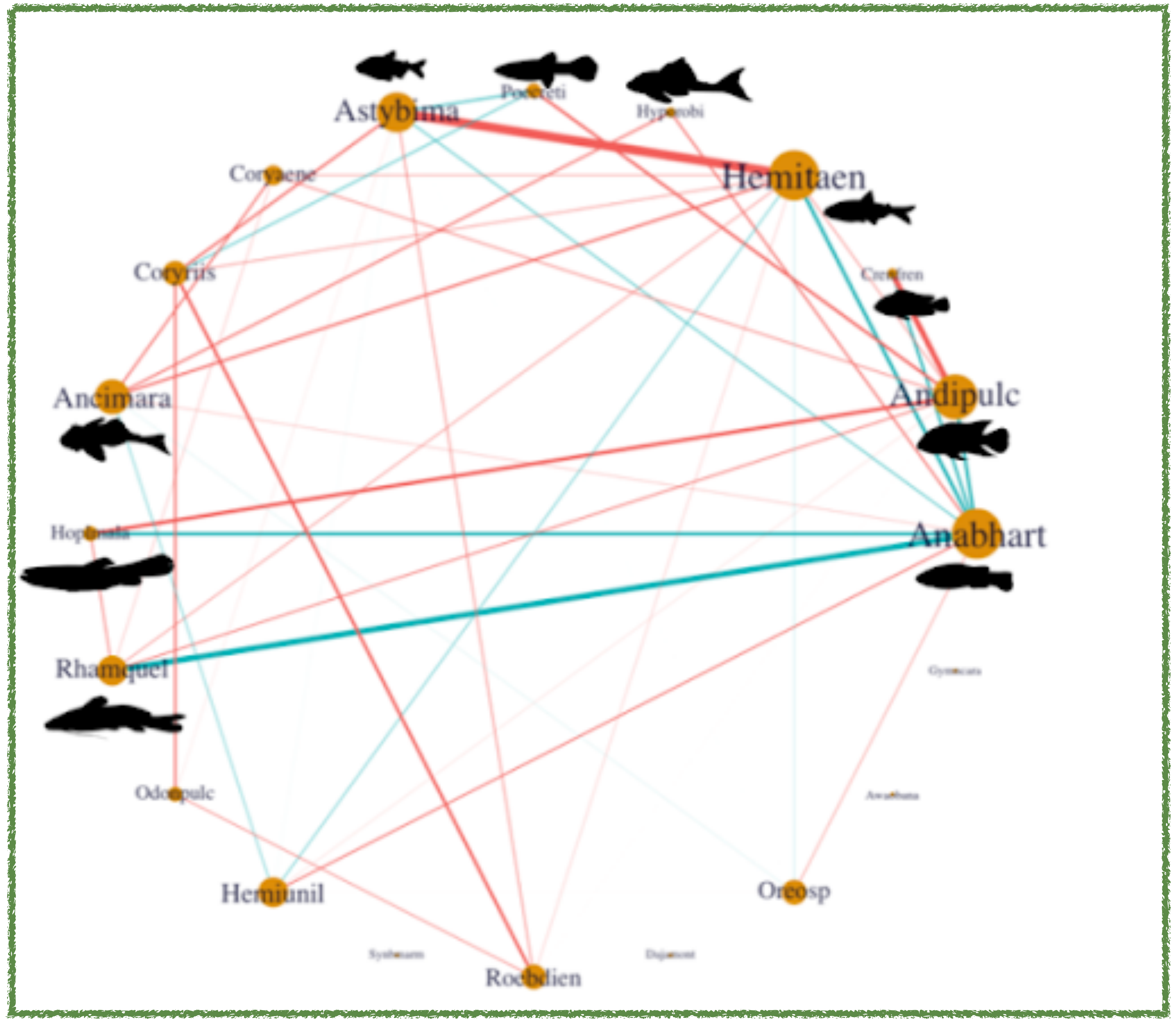
3. Joint Species Distribution Models (JSDMs)



Possible interpretations :

- Model misspecification.
- Missing abiotic covariates.
- Missing species.
- Biotic interactions.

3. Joint Species Distribution Models (JSDMs)



Possible interpretations :

- Model misspecification.
- Missing abiotic covariates.
- Missing species.
- ***Biotic interactions ?***

What if we knew exactly what interactions were involved in the distribution ?

4. A “virtual ecologist” approach

General method

Virtual Ecologist idea (Zurell et al., 2010) :

An “approach where simulated data and observer models are used to mimic real species and how they are ‘virtually’ observed. “

4. A "virtual ecologist" approach

General method

1. Simulate species distribution models with a biotic interactions based model.
2. Run JSDMs over the simulated data.
3. What level of correspondence between biotic interactions (*processes*) and statistical correlation (*patterns*).

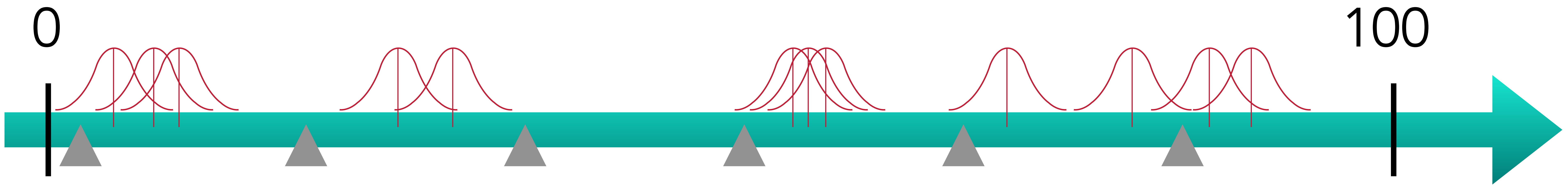
3'. Bonus: how is this influenced by: the interactions type / intensities /densities / symmetry…

4. A “virtual ecologist” approach

The VirtualCom model (Münkemüller and Gallien, 2015)

- Triple filtering: environmental, biotic interactions and reproduction.

- 1D environmental gradient. Sites and species niches over this gradient.

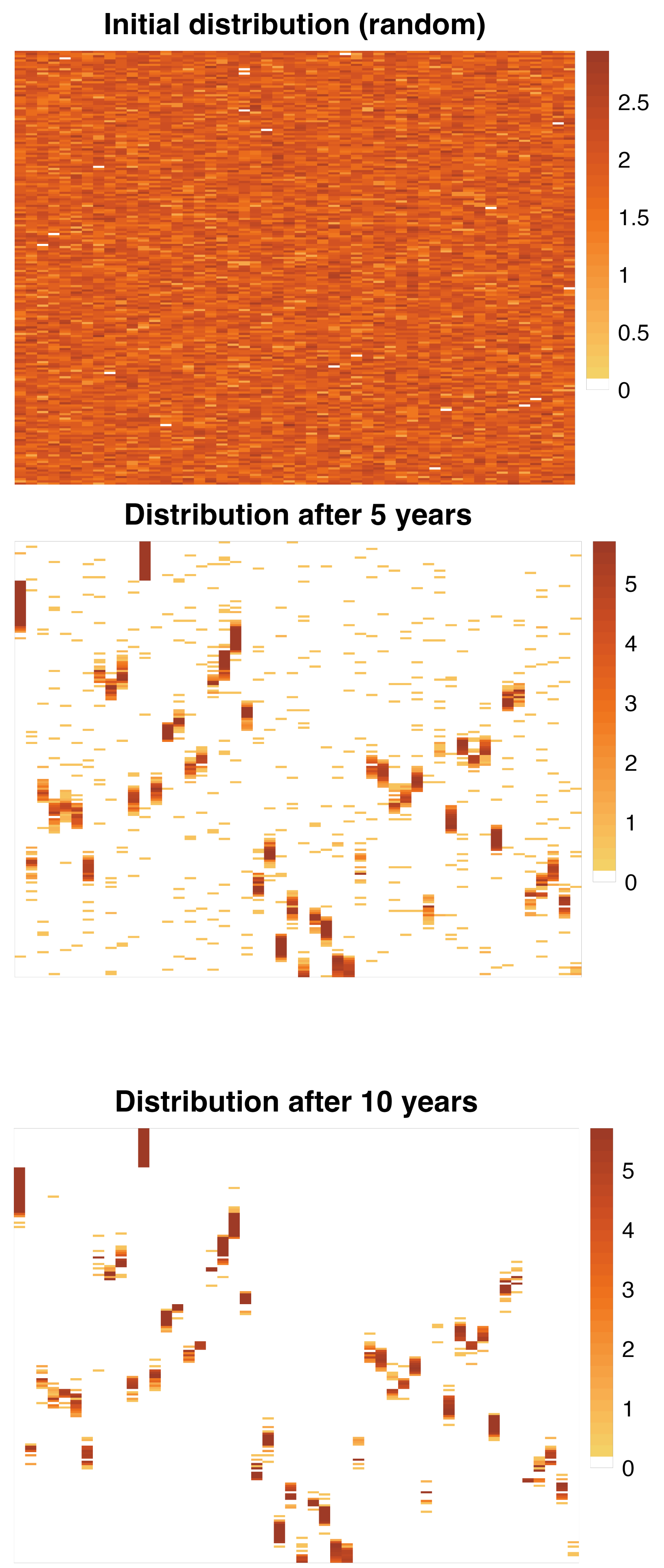


4. A "virtual ecologist" approach

The VirtualCom model (Münkemüller and Gallien, 2015)

For each site *i*:

1. Random initialization: K individuals are drawn from the species pool
2. K individuals are successively drawn and replaced with an individual from species *j* with probability p_{ij}.
3. *Step 2* is repeated over *n years*.



4. A “virtual ecologist” approach

The VirtualCom model (Münkemüller and Gallien, 2015)

Filtering allowed by *p*, resulting from the normalization of:

$$W_{ij} = \exp(\beta_{\text{env}} \log(p_{\text{env},i,j}) + \beta_{\text{abun}} \log(p_{\text{abun},i,j}) + \beta_{\text{comp}} \log(p_{\text{comp},i,j}))$$

Environmental filter $p_{\text{env}} = f(E_i; opt_j, sd_j)/f(E_i; E_i, sd_j)$, with f the density of a Gaussian distribution

Reproduction filter $p_{\text{abun},i,j} = \frac{N_{ij}}{K}$

4. A "virtual ecologist" approach

The VirtualCom model (Münkemüller and Gallien, 2015)

Filtering allowed by *p*, resulting from the normalization of:

Competition / Facilitation filter $$p_{comp} = 1 - \frac{1}{K} \sum_k a_{jk} N_{ik}$$

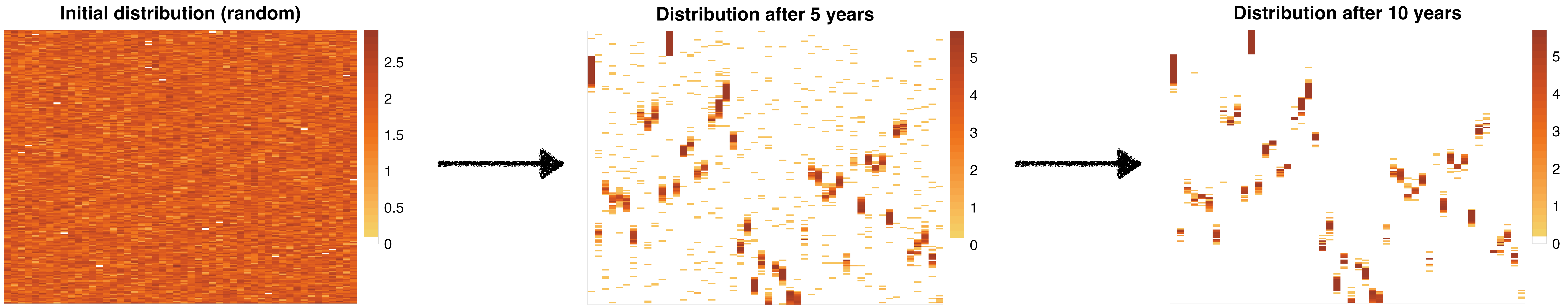
$a_{jk} > 0$ if j and k in competition, $a_{jk} < 0$ if j and k facilitate each other's presence.

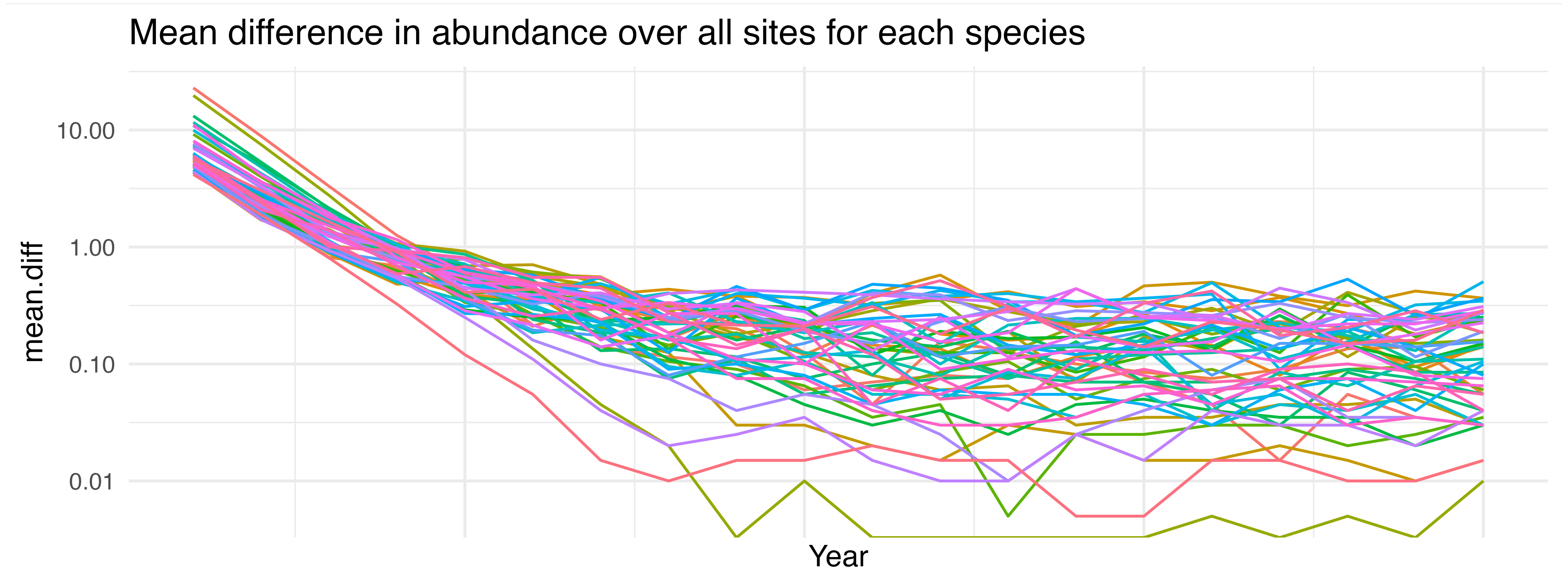
To generate the model we don't use the original approach but instead fix:

- A niche distance threshold beyond which two species can't interact.
- Competition and facilitation densities among the species whose niches are closer than this threshold.

4. A “virtual ecologist” approach

The VirtualCom model (Münkemüller and Gallien, 2015)





4. A "virtual ecologist" approach

Simulation protocol

- Simulate VirtualCom data with 50 species, 200 sites, niche distance threshold of 10.
- Changing interaction intensities & densities.
- We also test asymmetric interactions and sampling effect.

4. A "virtual ecologist" approach

Simulation protocol

- We test the adequation between the intuitive effect of interactions and what they really look like with a *relative abundance index*:

$$\mathrm{RAI}_{jk} = \mathrm{avg}\left(\{\delta_{i,j} | 1 \leq i \leq n\, y_{ij} > 0, y_{ik} > 0\}\right)$$
$$\text{with } \delta_{i,j} = y_{ij} - \bar{y}_j,$$

- We run several penalized methods evaluated in terms of AUC
  - True positive = competition matching negative residual correlations or facilitation matching positive residual correlations.

4. A “virtual ecologist” approach

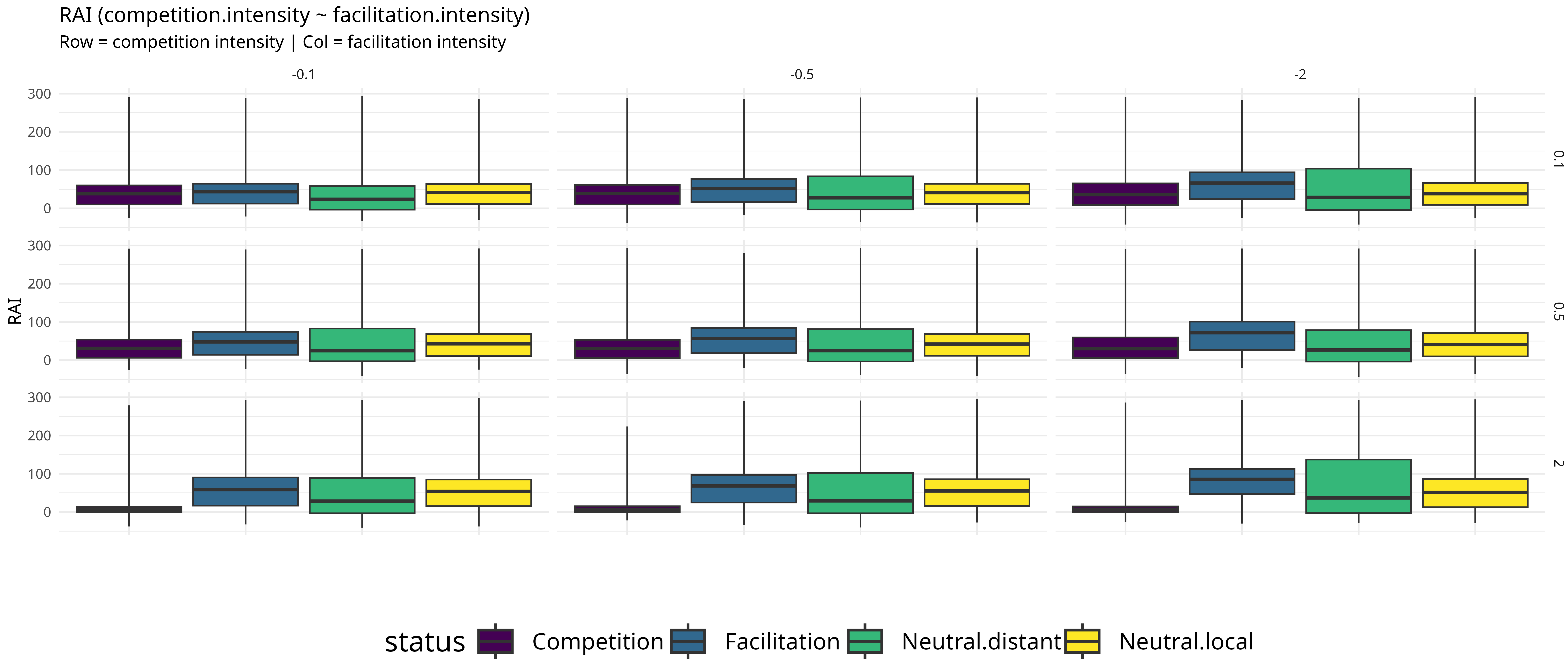
Simulation protocol

We test several JSDMs including a Poisson Log-normal model with:

- No covariate.
- The environmental gradient as a continuous covariate.
- A discretized version of the environmental gradient.
  - Goal = account for the niche modelling in the covariate effect.

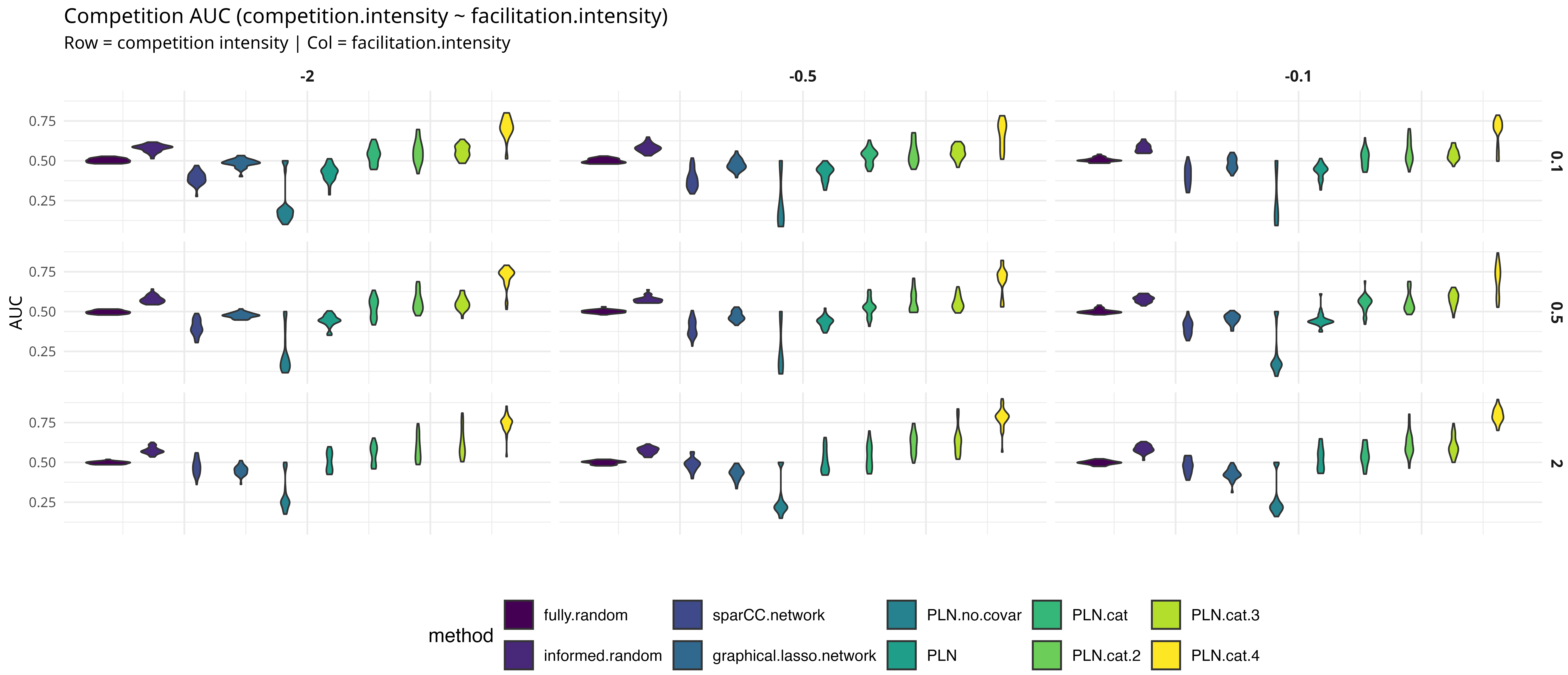
4. A "virtual ecologist" approach

Results: competition



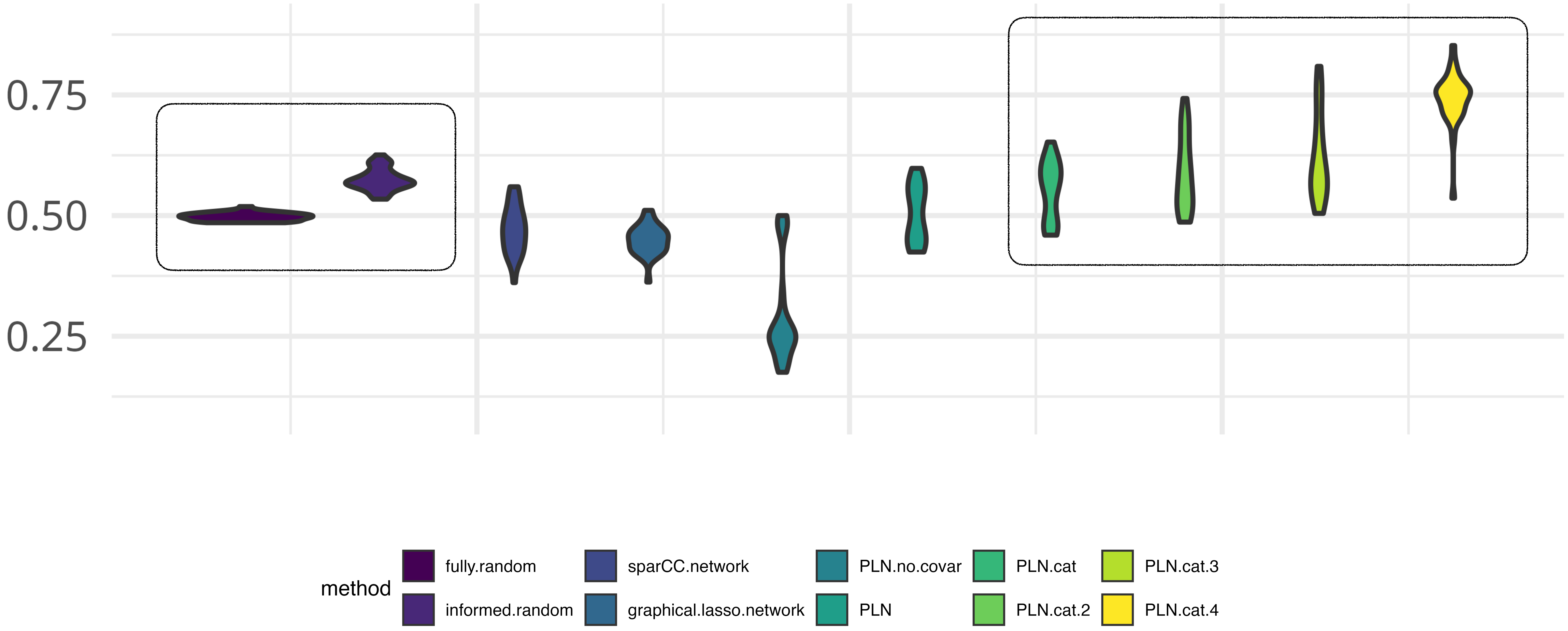
4. A “virtual ecologist” approach

Results: competition



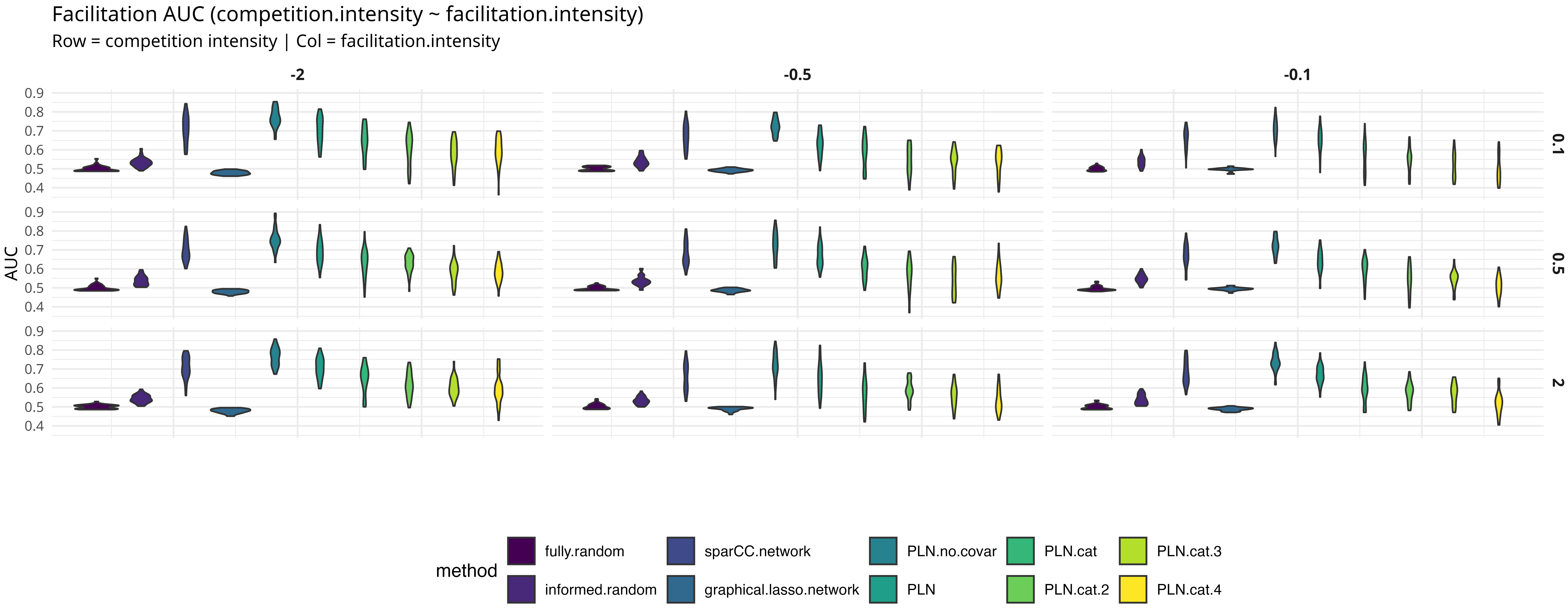
4. A “virtual ecologist” approach

Results: competition



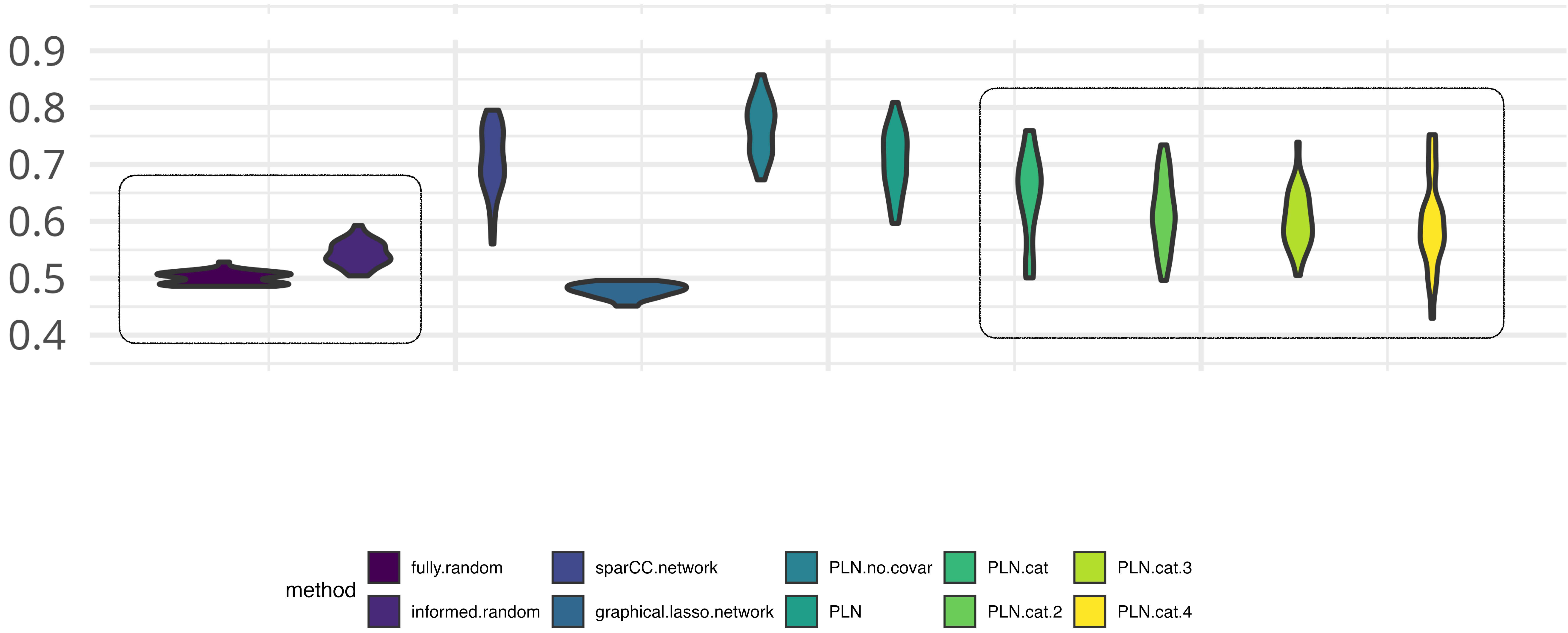
4. A "virtual ecologist" approach

Results: facilitation



4. A "virtual ecologist" approach

Results: facilitation



4. A "virtual ecologist" approach

Interpretation

- Reversed results for competition and facilitation: antagonistic *vs* synergistic effects ?

- Importance of adjusting the modelling of abiotic effects.

- Results robust to changes in intensities and densities.

- Slightly more surprising: they're also robust to asymmetry (competition / neutral)

5. Conclusion

- An original virtual framework to test a type of processes - patterns correspondence.
- Good correspondence for competition.
- Things are more complex for facilitation, that’s probably owing to the synergetic vs antagonistic situation in this framework.
- Underlines the importance of a careful abiotic modelling.

6. Limits & perspectives

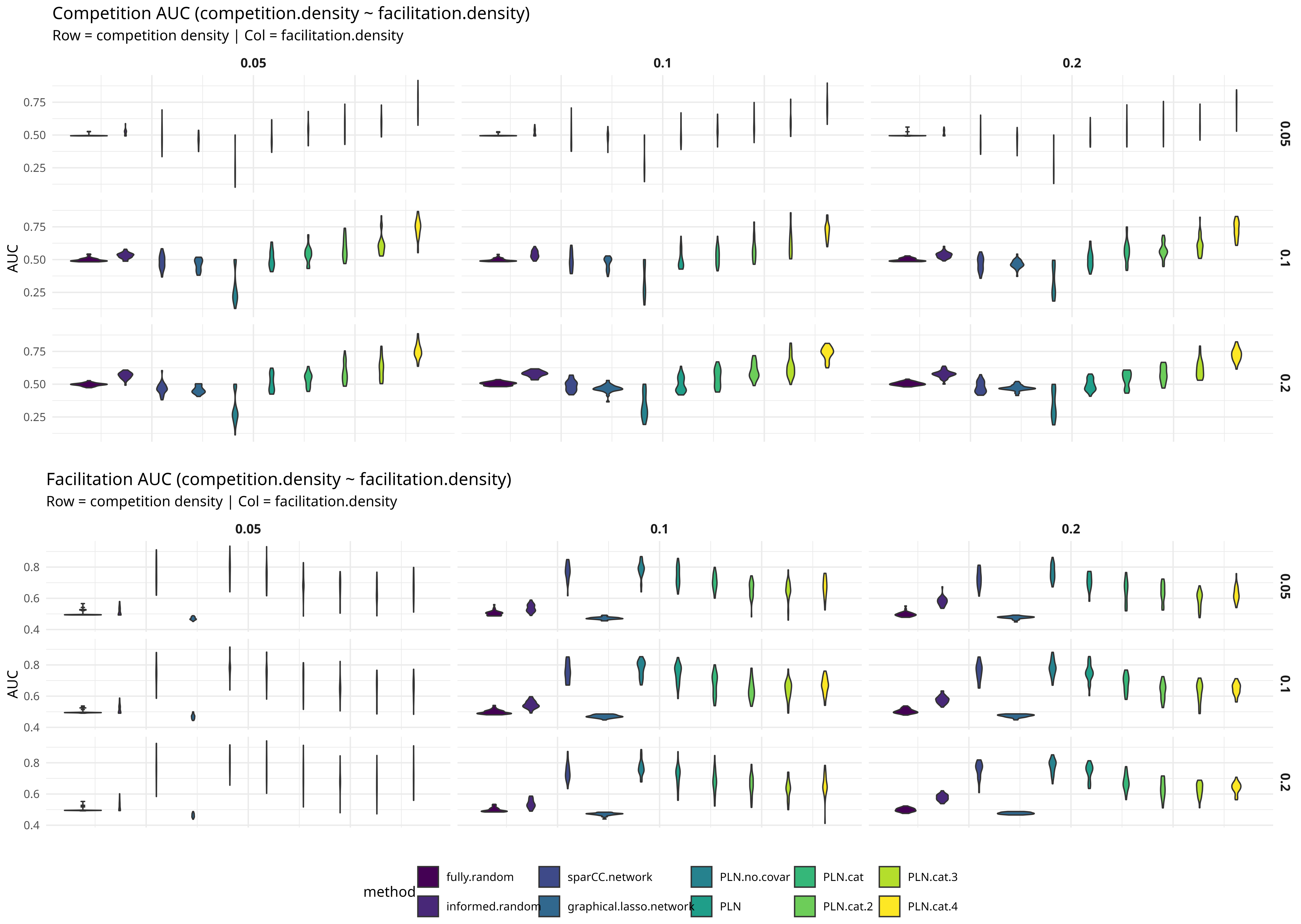
- There are more rigorous ways to model abiotic effects, eg using splines.
- Other interactions?
- Proximity between VirtualCom simulations and real data?
- To what extent is this a circular reasoning? 🤔

Thank you !

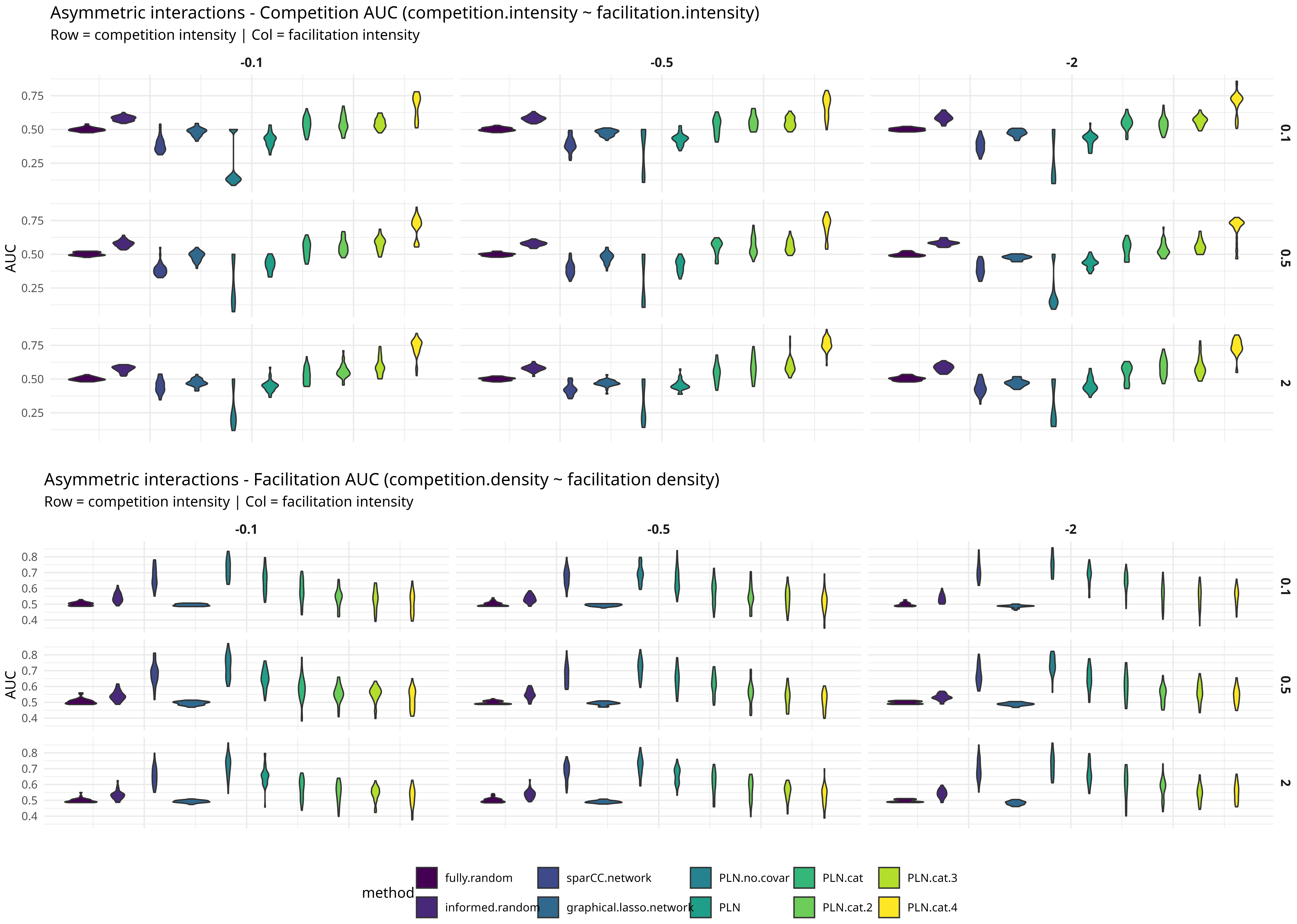
Questions ? Suggestions ?

Remarks ?

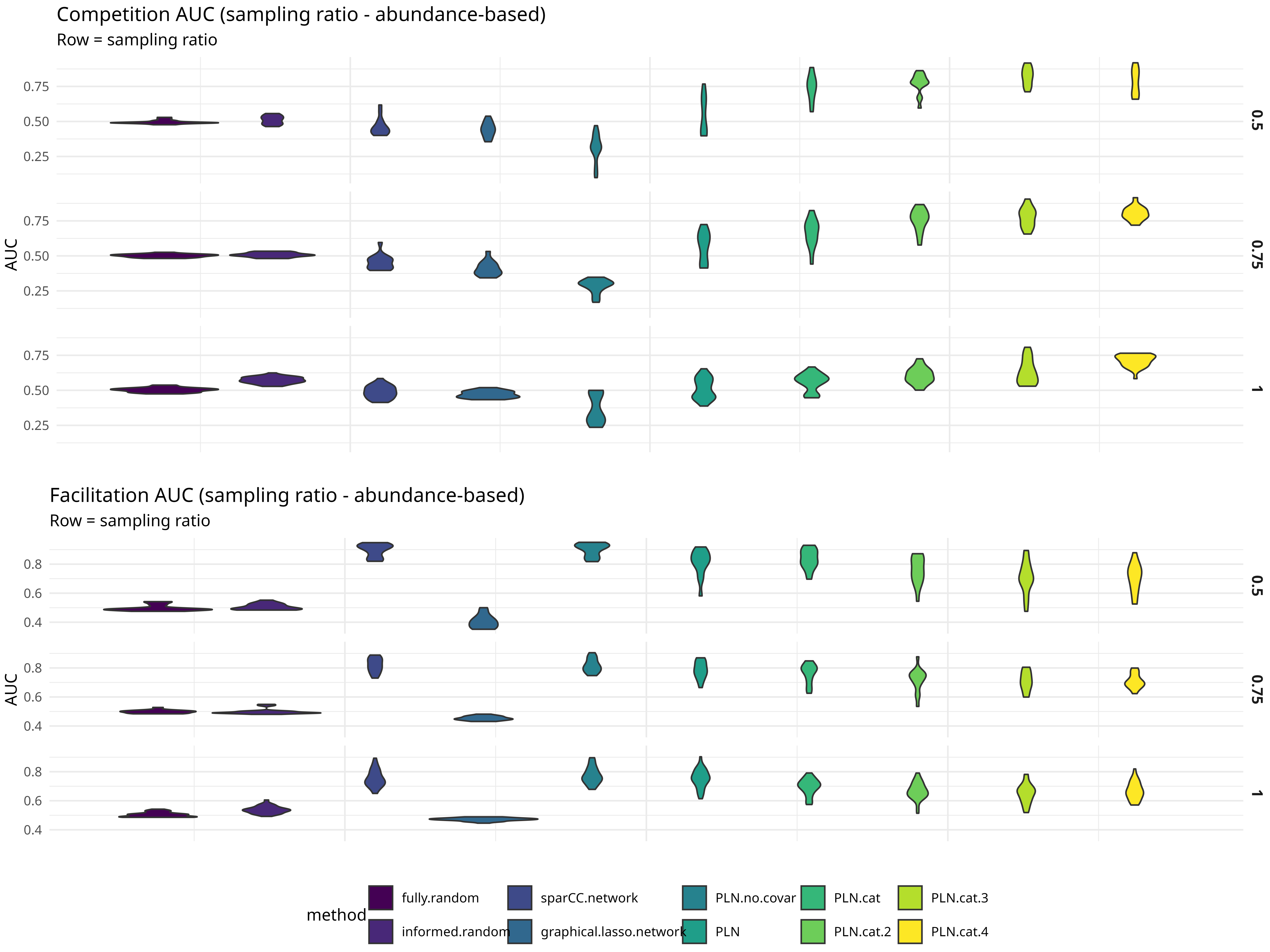
Appendix 1: results for changing densities



Appendix 2: results for asymmetric interactions



Appendix 3: abundance-based sampling effect



Appendix 4: random sampling effect