

Planned and ongoing Hmsc developments

- Scaling up computational efficiency for bigger data
- More flexibility with data models
- New model structures related to ecological and evolutionary processes
- Stability and usability of the R-package Hmsc

Scaling up computational efficiency for bigger data

Published and implemented

- Better MCMC mixing via collapsed Gibbs updaters

In progress

- Efficient HPC and GPU -compatible implementation
- Phylogenetic component scalable to 1000+ species
- Numerically scalable spatio-temporal latent factors

More flexibility with data models

We have:

Linear model:

$$y_{ij} = L_{ij} + \varepsilon_{ij}$$

Probit model:

$$y_{ij} \sim \text{Bernoulli}(\Phi(L_{ij}))$$

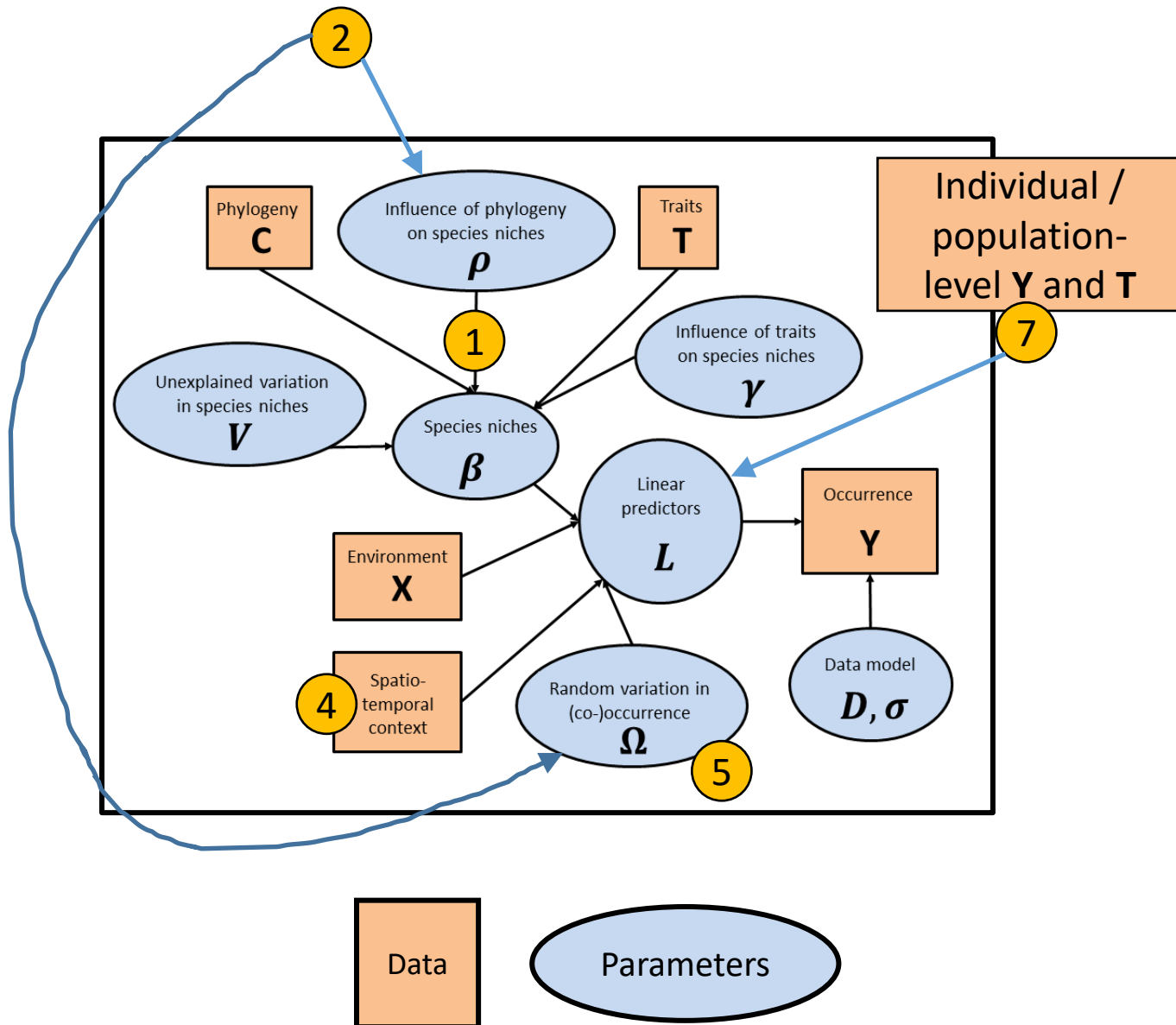
Lognormal Poisson model:

$$y_{ij} \sim \text{Poisson}(\exp(L_{ij} + \varepsilon_{ij}))$$

We need:

- More (and better working) models for (zero-inflated) abundance (count) data, such as negative binomial
- Multinomial data (e.g. sequencing data)
- Point-pattern data (e.g. log Gaussian Cox process)
- Data matrices with uncertainty (especially Y, but maybe also X, T, C)
- Explicit separation of the process model and the observation model

Model structures related to ecological and evolutionary processes



1. More flexible and/or mechanistic models to account for species evolution via phylogeny
2. How traits and phylogenies influence co-occurrence?
3. Informatively communicating massive results
4. Capability to include flexible, but scalable covariance structures for latent factors
5. Network modelling (including possibility of bringing prior information about known interactions and unfeasible interactions).
6. Can 5. be used to disentangle species associations from real interactions?
7. Individual-level perspective – incorporating collected meta-info on studied individuals