The basic idea here is that we want to cluster species regarding how they respond to the environment using presence/absence data.

Let denote the presence (=1) or absence (=0) of species s in location l at time t. We assume that:

if

otherwise

where

In this model, is a species specific intercept, is a vector of covariates, and is a vector containing the regression slopes for cluster c. We use a species-specific intercept because we are not interested in clustering species according to their overall prevalence.

We adopt the following priors

Where implies that

and

Where can take any of the following numbers with equal probability: 0.1, 0.15, 0.2,…, 0.9, 0.95, 1

#------------------------------------------

#------------------------------------------

Full conditional distributions

* For :
* For

This implies that

* For

Where is the number of species assigned to group k (i.e., ).

Therefore

* For

According to Gorur and Rasmussen 2010, we have two situations:

1. If group k already exists
2. If group k does not exist

Let and. This implies that . Therefore:

Taking logs, this becomes:

We can sample this parameter from a categorical distribution with probabilities proportional to that described above

* For

Where

where . This implies that

* For

we can sample this parameter using a multinomial distribution.