

# 1 Materials and Methods

## 1.1 Taxon occurrence information

## 1.2 Geographic provinces

## 1.3 Model specification

Taxon presence was modeled has a hierarchical two-state hidden Markov model (HMM) where the three “process parameters” of gain ( $\gamma$ ), persistance ( $\phi$ ), and observation ( $p$ ). For each province, each of these process parameters were modeled hierarchically so that estimates were allowed to vary over time but in cases of little information those estimates were drawn to the overall mean for that province. The estimates for each province were also estimated hierarchically in relation to each other; this way all estimates were relative to each other. The hierarchical structure of this model helps control for both overfitting and multiple comparisons during posterior analysis [? ? ].

$y_{i,j,t}$  is the observed occurrence of taxon  $i$  in province  $j$  at time  $t$ , where  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, J$ , and  $t = 1, 2, \dots, T$ .  $y = 1$  is occupied while  $y = 0$  is unoccupied.  $z_{i,j,t}$  is the “true” occurrence of taxon  $i$  in province  $j$  at time  $t$ , given the estimate of sampling. Just as with  $y$ ,  $z = 1$  is occupied while  $z = 0$  is unoccupied.

$\phi_{j,t}$  is the probability of surviving, in province  $j$ , from time  $t$  to time  $t + 1$  ( $Pr(z_{t+1} = 1|z_t = 1)$ ).  $\gamma_{j,t}$  is the probability of newly entering province  $j$  at time  $t + 1$  ( $Pr(z_{t+1} = 1|z_t = 0)$ ).  $p_{j,t}$  is the probability of preservation ( $Pr(y = 1|z = 1)$ ) in province  $j$  at time  $t$ .

$\psi$  is probability of sit occupancy/probability of occurrence ( $Pr(z_{i,t=1} = 1)$ ). The first time point is defined in terms of  $\psi$  because there is (assumed) no previous time points.

The parameters  $\phi$ ,  $\gamma$ , and  $p$  are then all defined hierarchically within each province, hierarchical by time bin with the mean of that hierarchical by province.  $\Phi_j$ ,  $\Gamma_j$ , and  $P_j$  are the overall probabilities for province  $j$ .  $M_\phi$ ,  $M_\gamma$ , and  $M_p$  are the overall estimates of survival, origination, and preservation probabilities.

Diversity dependent origination and survival was included as a regression coefficients in the parameterizations of  $\phi_{j,t}$ , and  $\gamma_{j,t}$ . In particular, the diversity of any province can affect the origination and survival probabilities of an province; this is conceptually and mathematically similar to how clade competition was modeled by ? ].

$\alpha$  and  $\beta$  are both  $J \times J$  matrices of regression coefficients, where  $\beta_j$  would be a column vector.  $X$  is a  $J \times T$  matrix where  $X_t$  is a column vector where each element is defined  $X_{j,t} = \sum_{i=1}^N z_{i,j,t}$  (i.e. the sum of the diversity in province  $j$  at time  $t$ ). All regression coefficients (i.e. all elements of the matrices) are given weakly-informative independent normally distributed priors.

And finally, I use independent uniform priors for  $\psi_j$  by province  $j$ :  $\psi_j \sim U(0, 1)$ .

In total, the model can be summarized by the following statements:

$$\begin{aligned}
y_{i,t,j} &\sim \text{Bern}(p_{t,j} z_{i,t,j}) \\
z_{i,t=1,j} &\sim \text{Bern}(\psi_j) \\
z_{i,t,j} &\sim \text{Bern}(\phi_{j,t-1} z_{i,t-1,j} + \gamma_{j,t-1} (1 - z_{i,t-1,j})) \\
\text{logit}(\phi_{j,t}) &\sim \text{N}(\Phi_j + X_{t-1} \beta_j, \sigma_{\phi,j}) \\
\text{logit}(\gamma_{j,t}) &\sim \text{N}(\Gamma_j + X_{t-1} \alpha_j, \sigma_{\gamma,j}) \\
\text{logit}(p_{j,t}) &\sim \text{N}(P_j, \sigma_{p,j}) \\
\Phi_j &\sim \text{N}(M_\phi, \sigma_\Phi) \\
\Gamma_j &\sim \text{N}(M_\gamma, \sigma_\Gamma) \\
P_j &\sim \text{N}(M_p, \sigma_P) \\
\sigma_{\phi,j} &\sim \text{C}^+(1) \\
M_\phi &\sim \text{N}(0, 1) \\
\sigma_\Phi &\sim \text{C}^+(1) \\
\beta &\sim \text{N}(0, 1) \\
\sigma_{\gamma,j} &\sim \text{C}^+(1) \\
M_\gamma &\sim \text{N}(0, 1) \\
\sigma_\Gamma &\sim \text{C}^+(1) \\
\alpha &\sim \text{N}(0, 1) \\
\sigma_{p,j} &\sim \text{C}^+(1) \\
M_p &\sim \text{N}(0, 1) \\
\sigma_P &\sim \text{C}^+(1)
\end{aligned} \tag{1}$$

## 1.4 Posterior inference