1 Materials and Methods

1.1 Taxon occurrence information

1.2 Geographic provinces

1.3 Model specification

Taxon presence was modeled has a hierarchical hidden Markov model (HMM) where the three "process parameters" of gain/newly entering a province (γ), persistance/survival (ϕ), 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 [? ?].

Note that the following model is strongly inspired by the dynamic occupancy model presented in [?].

$y_{i,j,t}$ is the observed occurrence of taxon i in province j at time t, where $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, J$, and $t = 1, 2, \ldots, T$. $y = 1$ 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 observing a true occurrence ($Pr(y = 1 | z = 1)$) in province j at time t.

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

The parameters ϕ, γ, and p are then all defined hierarchically within time bins as samples from the shared mean between provinces. Φ_t, Γ_t, and P_t are

the mean probabilities for a given point in time t. M_ϕ, M_γ, and M_p are the overall mean 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 **?**].

α and β are both $J \times J$ matrices of regression coefficients, where β_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 $X_{j,t} = \Sigma_{i=1}^N z_{i,j,t}$ (i.e. the sum of the diversity in province j at time t). These regression coefficients were given informative priors based on the theory that as diversity increases, origination should decrease and extinction should increase/survival should decrease.

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

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

$$
\begin{aligned}
y_{i,t,j} &\sim \mathrm{Bern}(p_{t,j} z_{i,t,j}) \\
z_{i,t=1,j} &\sim \mathrm{Bern}(\psi_j) \\
z_{i,t,j} &\sim \mathrm{Bern}(\phi_{j,t-1} z_{i,t-1,j} + \gamma_{j,t-1}(1 - z_{i,t-1,j})) \\
\mathrm{logit}(\phi_{j,t}) &\sim \mathrm{N}(\Phi_t + X_{t-1}\beta_{j,-}, \sigma_{\phi,t}) \\
\mathrm{logit}(\gamma_{j,t}) &\sim \mathrm{N}(\Gamma_t + X_{t-1}\alpha_{j,-}, \sigma_{\gamma,t}) \\
\mathrm{logit}(p_{j,t}) &\sim \mathrm{N}(P_t, \sigma_{p,j}) \\
\Phi_i &\sim \mathrm{N}(M_\phi, \sigma_\Phi) \\
\Gamma_i &\sim \mathrm{N}(M_\gamma, \sigma_\Gamma) \\
P_i &\sim \mathrm{N}(M_p, \sigma_P) \\
\sigma_{\phi,i} &\sim \mathrm{C}^+(1) \\
M_\phi &\sim \mathrm{N}(0, 1) \\
\sigma_\Phi &\sim \mathrm{C}^+(1) \\
\beta_{j,k} &\sim \mathrm{N}(-1, 1) \\
\sigma_{\gamma,j} &\sim \mathrm{C}^+(1) \\
M_\gamma &\sim \mathrm{N}(0, 1) \\
\sigma_\Gamma &\sim \mathrm{C}^+(1) \\
\alpha_{i,k} &\sim \mathrm{N}(-1, 1) \\
\sigma_{p,j} &\sim \mathrm{C}^+(1) \\
M_p &\sim \mathrm{N}(0, 1) \\
\sigma_P &\sim \mathrm{C}^+(1)
\end{aligned}
\tag{1}
$$

The joint posterior distribution of the HMM model was approximated using a Gibbs sampling MCMC routine as implemented in the JAGS probabilistic programming language CITATION. Four chains were run for 20000 steps, thined to every 10th sample, ans split evenly between warm-up and sampling phases. Chain sampling convergence was assessed using the $\hat{R}$ statistic with values close to 1 (less than 1.1) indicating approximate convergence [?].

Given the estimate of the joint posterior distribution, some downstream metapopulation summary statistics can be calculated. The first of these is turnover τ defined as the probability that the occurrence of a taxon at time t is new ($Pr(z_{t-1} = 0 | z_t = 1)$) [?]. First, the occupancy probability ψ at times

$t = 2, \ldots, T$ can calculated recursively as

$$\psi_t = \psi_{t-1}\phi_{t-1} + (1 - \psi_{t-1})\gamma_{t-1}. \tag{2}$$

Turnover can then be calculated as

$$\tau_t = \frac{\gamma_{t-1}(1 - \gamma_{t-1}}{\gamma_{t-1}(1 - \psi_{t-1}) + \phi_{t-1}\psi_{t-1}} \tag{3}$$

An additional summary is the growth rate λ [? ?] which is calculated as

$$\lambda_t = \frac{\psi_{t+1}}{\psi_t}. \tag{4}$$