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.

$y_{i,j,t}$ is the observed binary 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 estimated binary 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.

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.

$$
\begin{aligned}
y_{i,t,j} &\sim \mathrm{Bern}(p_{t,j} z_{i,t,j}) \\
z_{i,t=1,j} &\sim \mathrm{Bern}(\psi_j) \\
z_{i,t,j} &\sim \mathrm{Bern}(\phi_{j,t-1} z_{i,t-1,j} + \gamma_{j,t-1}(1 - z_{i,t-1,j}))
\end{aligned}
\tag{1}
$$

The parameters ϕ, γ, and p are then all defined hierarchically within each province, hierarchical by time bin with the mean of that hierarchical by province. Φ_j, Γ_j, and P_j are the overall probabilities for province j. M_{ϕ}, M_{γ}, 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.

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). All regression coefficients (i.e. all elements of the matrices) are given weakly-informative independent normally distributed priors.

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

$$
\begin{aligned}
\text{logit}(\phi_{j,t}) &\sim \mathrm{N}(\Phi_j + \beta_j^T X_{t-1}, \sigma_{\phi,j}) \\
\Phi_j &\sim \mathrm{N}(M_\phi, \sigma_\Phi) \\
\sigma_{\phi,j} &\sim \mathrm{C}^+(1) \\
M_\phi &\sim \mathrm{N}(0, 1) \\
\sigma_\Phi &\sim \mathrm{C}^+(1) \\
\beta &\sim \mathrm{N}(0, 1) \\
\text{logit}(\gamma_{j,t}) &\sim \mathrm{N}(\Gamma_j + \alpha_j^t X_{t-1}, \sigma_{\gamma,j}) \\
\Gamma_j &\sim \mathrm{N}(M_\gamma, \sigma_\Gamma) \\
\sigma_{\gamma,j} &\sim \mathrm{C}^+(1) \\
M_\gamma &\sim \mathrm{N}(0, 1) \\
\sigma_\Gamma &\sim \mathrm{C}^+(1) \\
\alpha &\sim \mathrm{N}(0, 1) \\
\text{logit}(p_{j,t}) &\sim \mathrm{N}(P_j, \sigma_{p,j}) \\
P_j &\sim \mathrm{N}(M_p, \sigma_P) \\
\sigma_{p,j} &\sim \mathrm{C}^+(1) \\
M_p &\sim \mathrm{N}(0, 1) \\
\sigma_P &\sim \mathrm{C}^+(1)
\end{aligned}
\tag{2}
$$