

Materials and Methods

Taxon occurrence information

Taxon occurrence information was downloaded from the Paleobiology Database on May 1st 2016. Occurrences were restricted to all Mammalia between the Maastrichtian and Gelasian stages and the North American continent. These occurrences were then further programatically restricted from the Danian through the Gelasian. The taxonomic and stratigraphic metadata for each occurrence, such as life habit and dietary category, was also downloaded. From this base downloaded some taxa were further excluded by their taxonomic and life habit information. Specifically, both volant and aquatic mammals along with the orders and families associated with those life habits were excluded.

Life habit and dietary categories of all occurrences were adjusted following TABLE.

Life habit is greatly expanded compared to previous studies ([?]).

Instead of a single “ground dwelling” category, species initially classified as “ground dwelling” were then further divided by foot posture following [?]. Otherwise, ground dwelling is overwhelming and uninterestingly, demographically. The classification by footposture gives more precise information about the structure of the regional species pool wrt environmental context and time.

Species mass data was sourced from the Paleobiology Database, SMITH, TOMIYA, ETC., ETC.

Model specification

This model is simply a hidden markov model with an absorbing state. The notation follows the Jolly-Seber capture-mark-recapture model presented in Royle and Dorazio (book).

The observed state of a species at a given time is a function of both its “true” state (present or absent) and the probability observation.

The underlying hidden Markov model used in this study has three characteristic probabilities: probability p of observing a species given that it is present,

Parameter	explanation
p	probability of observation
α_0	average observation probability
α_1	change in average observation probability per change in standard deviation of mass
r_t	difference in average observation probability at time t
σ	standard deviation of values of r_t

Table 1: Observation parameters

Parameter	dimensions	explanation
θ	$N \times T - 1$	probability of species presence
β	$T - 1 \times D$	matrix of species-level regression coefficients
γ	$U \times D$	matrix of group-level regression coefficients
Σ	$D \times D$	covariance matrix between species-level effects over time
Ω	$D \times D$	correlation matrix between species-level effects over time
τ	D	vector of standard deviations of species-level effects over time

Table 2: Presence parameters

probability ϕ of a species surviving from one time to another, and probability π of a species first appearing. Because the questions being addressed in this study are concerned only with species presence, ϕ and π are constrained to be equal and referred to as θ : the probability that a species is present at time t given whatever state that species was in at time $t - 1$.

The probability of observation a species p was modeled as a logistic regression with a time-varying intercept and species mass as a predictor. The effect of species mass on probability of observation was considered constant over time. The parameters associated with this aspect of the model

The probability of species occurrence was modeled as a two-level logistic regression: species-level variation and group-level variation. The species-level of the model included all of species specific covariates (e.g. dietary category) while the group-level of the model included all of the time-bin specific covariates (e.g. global temperature). The effects of the species-level covariates were allowed to vary both over time and in response to the group-level covariates in the form of a varying-intercepts, varying-slopes model CITATION. The parameters associated with this part of the model are presented in Table 2.

Following recommendations from CITATION, CITATION, CITATION all parameters not modeled elsewhere were given weakly informative priors. Weakly informative means that priors do not encode previous information but instead help regularize or weakly constrain posterior estimates. These priors have a concentrated probability density around and near zero; this has the effect of tempering our estimates and help prevent overfitting the model to the data CITATION.

$$\begin{aligned}
y_{i,t} &\sim \text{Bernoulli}(p_{i,t}z_{i,t}) \\
p_{i,t} &= \text{logit}^{-1}(\alpha_0 + \alpha_1 m + r_t) \\
\alpha_0 &\sim \mathcal{N}(0, 1) \\
\alpha_1 &\sim \mathcal{N}(0, 1) \\
r_t &\sim \mathcal{N}(0, \sigma) \\
\sigma &\sim \mathcal{N}^+(1) \\
z_{i,t} &\sim \text{Bernoulli}(\theta_{i,t}) \\
\theta_{i,t} &= \text{logit}^{-1}(x_i\beta_t) \\
\beta_t &\sim \text{MVN}(u_t\gamma, \Sigma) \\
\gamma &\sim \mathcal{N}(0, 1) \\
\Sigma &= \text{diag}(\tau)\Omega\text{diag}(\tau)
\end{aligned} \tag{1}$$

Posterior inference and model adequacy