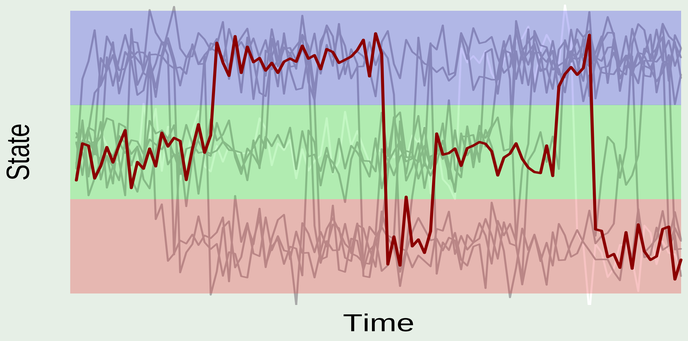
Robust Information from Phylogentic Trees?

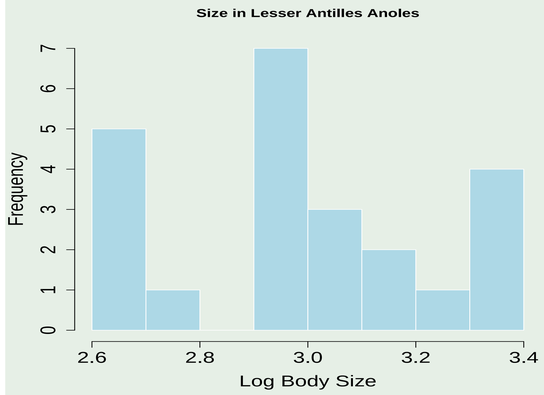
Carl Boettiger, UC Davis

Process dynamics look like:



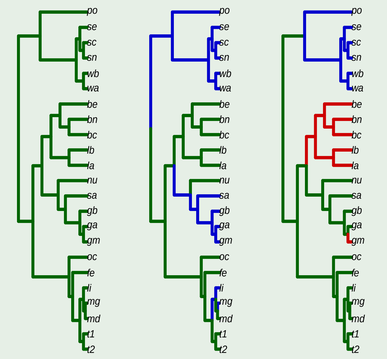
Imagine the evolution of a continuous character trait which may occasionally jump from one optima to another – a large ground-dwelling species becoming smaller tree-dwellers. If we could observe this entire process over many species, it might look like this.

Can we reconstruct it with this. . .

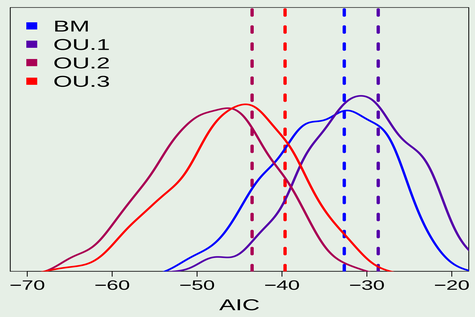


We see only a snapshot in time. Each species as a replicate from this stochastic process. We detect clusters, but can't calculate rates.

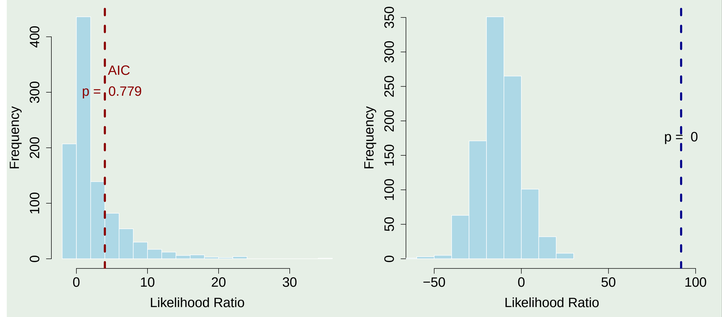
and a Phylogenetic Tree?



Species trait values don't represent independent sample paths. The phylogenetic tree introduces correlations. We will rely on this for temporal information. Colors indicate models with different optima (paintings).

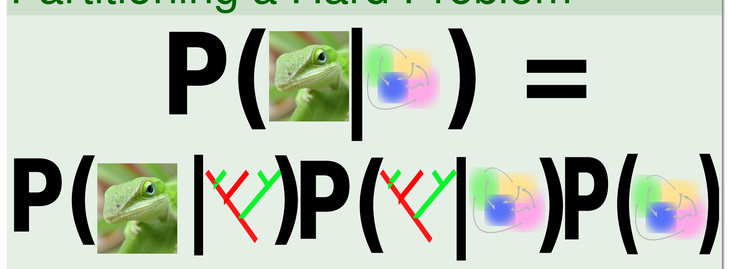


Popular model choice approaches based on AIC fail to reliably distinguish between any number of optima.

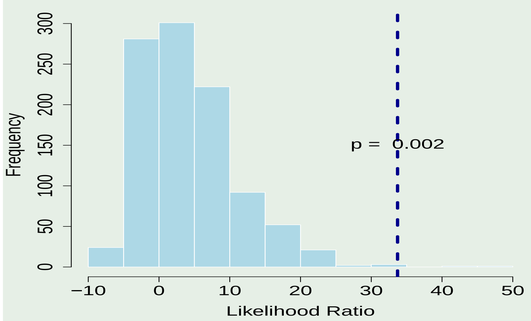


This test provides a more meaningful way to avoid always selecting more complicated models (left), but there is no sensible way to select between non-nested paintings (right).

Partitioning a Hard Problem



Generating paintings from a Markov model rather than as a hypothesis is preferable biologically and statistically. A clever partition of likelihood makes this numerically feasible.



A better way: simulation based parametric bootstrap of likelihood ratios reliably prefers the two optima model over the neutral model.

Data given the painting

$$\mathrm{E}(X_t|X_0) = X_0 e^{-\alpha t} + \theta(1 - e^{-\alpha t})$$

$$V_{ij} = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha s_{ij}})e^{-2\alpha(t-s_{ij})}$$

$$\mathcal{L}(X) = \frac{e^{(X-\mathrm{E}(X))^T V^{-1}(X-\mathrm{E}(X))}}{(2\pi)^{2N}\det V}$$

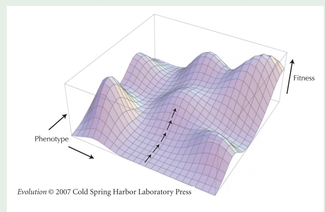
The likelihood of a set of species mean traits is jointly Gaussian even if the model has multiple regimes.

Painting given transition matrix

$$\mathcal{L} = \prod_{ij}\{e^{\mathbb{Q}s_{ij}}\}_{ij}$$

Exponentiating the transition matrix $\mathbb{Q}$ determines the probability of over all possible paths from state i to state j. Taking the product over all branches i, j of the tree gives the probability of the painting given $\mathbb{Q}$

MCMC over paintings, regime parameters & transition rates. Propose from Q and accept steps with probability: $\alpha < \min\left\{\frac{P(x')Q(x^t;x')}{P(x^t)Q(x';x^t)}, 1\right\}$



Summary

- Can infer dynamic evolutionary parameters without fossil history
- Handle uncertainty more robustly than current methods

Acknowledgements

- C. Martin, S. Price, P. Wainwright
- P. Ralph, G. Coop

DOE CSGF

Nature Precedings : doi:10.1038/npre.2010.4614.1 : Posted 3 Jul 2010