

Master thesis proposal

Exploring the power of BiSSE

Krzysztof Bartoszek

October 20, 2017

Background—the BiSSE model

The field of phylogenetic comparative methods aims at understanding trait (like body size, eye colour, presence/absence of tail, compound/simple eyes) evolution on the between species level. From a statistical point of view this implies that the sample (usually some average measurement for each species) cannot be considered to be an independent one. The relationships between the species are described by their phylogeny—a rooted binary tree (in the graph theory sense) that indicates when species diverged from each other. With a given tree one assumes some model for the evolution of the trait, e.g. a multi-state Markov chain, Brownian motion.

Nearly all contemporary methods make a crucial assumption—the evolving trait does not affect the speciation, it only evolves “on top” of the phylogeny. This is of course in stark contrast to any biological intuition—after all something had to change to make the species distinct. The Binary state speciation and Extinction (BiSSE) model [5] was the first (and still is one of the few used) exception to this. BiSSE models a binary trait (two states e.g. compound, simple eyes denoted as 0 and 1) and has six parameters—the speciation rate in state 0 (λ_0), in state 1 (λ_1), the extinction rate in state 0 (μ_0), in state 1 (μ_1), the transition rate from state 0 to 1 (q_{01}) and the transition rate from state 1 to 0 (q_{10}).

Thesis project

Even though the BiSSE model is well established in the biological community it has come under heavy criticism due to its supposed low power to distinguish between models i.e. does a trait affect the branching or not and that to have decent estimates large samples, of at least 300 species are required [1]. On the other hand more recent studies indicate that the situation is not so critical, if one is only interested in some parameters or puts further restrictions on the model [3, 4, 6]. However, none of these studies explored the power and estimation capabilities of the BiSSE model when sets of parameters are set to 0. This especially concerns the extinction rates. It is observed that these are the most difficult to estimate [4, 6], hence removing them completely from the model should significantly improve maximum likelihood estimation. Procedures for the BiSSE model estimation, simulation are available in the `diversitree` R package [2] and also a wrapper function will be provided that makes it possible to estimate parameters with some set to given values.

Goals

The below general goals are for an “ideal” thesis. Depending on the student they will be made more specific in the direction of the student’s interests. In particular the focus of the work will not be on the mathematical models (these will be “provided”) but on implementing and putting together software to do simulations, inference and explore the statistical aspects of the models.

1. Become familiar with the modelling approach in evolutionary biology.
2. Become familiar with model selection techniques.
3. Explore the power and estimation accuracy of BiSSE maximum likelihood estimation under very restricted models.

Data

The topic will be illustrated with simulated data. Furthermore, analyses can be done on datasets attached to various R packages on CRAN.

References

- [1] M. P. Davis, P. E. Midford, and W. Maddison. Exploring power and parameter estimation of the BiSSE method for analyzing species diversification. *BMC Evol. Biol.*, 13:38, 2013.
- [2] R. G. FitzJohn. Quantitative traits and diversification. *Syst. Biol.*, 59: 619–633, 2010.
- [3] A. Gamisch. Notes on the statistical Power of the Binary state speciation and Extinction (BissE) model. *Evol. Bioinf.*, 12:165–174, 2016.
- [4] B. Igic and J. W. Busch. Is self-fertilization an evolutionary dead end? *New Phytol.*, 198:386–397, 2013.
- [5] W. P. Maddison, P. E. Midford, and S. P. Otto. Estimating a binary character’s effect on speciation and extinction. *Syst. Biol.*, 56(5):701–710, 2007.
- [6] M. E. Maliska, M. W. Pennell, and B. J. Swalla. Developmental mode influences diversification in ascidians. *Biol. Lett.*, 9:20130068, 2013.