A novel modelling framework to quantify phylogenetic structure in ecologically-relevant variables

Investigating the evolutionary processes that structure biodiversity distribution is critical, particularly in a time of rapid global change. Phylogenetic information has helped provide a new perspective on species distributions across space and time; determining the extent to which ecologically relevant variables are phylogenetically conserved is important to understanding the causes of biodiversity distribution and its evolutionary and ecological consequences. Current methods to detect phylogenetic structure in ecological variables often rely on strong assumptions of underlying trait models or are performed as ad-hoc analyses, i.e. not as part of a probabilistic model.

We present a new hierarchical Bayesian modelling framework that jointly estimates species-environment associations and phylogenetic correlation in associations. We utilize a latent Gaussian process to link species' ecological niches to phylogenetic structure. The latent process represents our mechanistic understanding of how a given clade evolved. We demonstrate the ability of this model to capture the phylogenetic structure in ecological niche for different models of trait evolution, levels of phylogenetic signal, and varying tree topologies. Further, we explore how Gaussian processes with different kernels in this framework can connect to distinct evolutionary patterns and can therefore be used as a general framework for hypothesis testing.

Estimating phylogenetic correlation in a joint modelling framework leads to robust inference over all parameters of the model, reduces bias in parameter estimation, and produces valid uncertainty estimates. The model we present has the additional advantage of providing pairwise estimation of phylogenetic correlation between species across each ecological variable. In the context of climate change, this is informative of how rapidly a clade might adapt to change across each of its niche axes. This analysis addresses the question of whether ecological variables that are most predictive in species distribution models are evolutionarily labile or conserved. The answer is of significance to the forecasting of biodiversity distribution under climate change.