

Design-based inference about community composition and interactions using multivariate abundance data

David I Warton

School of Mathematics and Statistics and Evolution & Ecology Research Centre
The University of New South Wales
Sydney, Australia
David.Warton@unsw.edu.au

Keywords: bootstrap; Bray-Curtis; generalised linear mixed models; model-based approach; multivariate analysis.

Abstract: Multivariate abundance data are commonly and often erroneously analysed using distance-based algorithms (or techniques related to correspondence analysis), to answer questions concerning change in community composition, or interactions between predictors in their effects on communities. But both composition and interactions are model-based notions and it is shown by example that reliable inference in such situations is not possible without use of a statistical model, and a model that captures key data properties. One key data property that is difficult to capture is correlation between different taxa, because there are typically many more possible pairwise correlations than there are observations in the dataset to estimate them. An alternative approach which can circumvent the issue is using design-based inference, and we will see how block-resampling of residuals constructed via the probability integral transform can ensure valid inference about composition and interactions even when species correlations cannot feasibly be modelled. The approach is quite general, and in the `mvabund` package, it has been implemented in a very flexible framework capable of handling many common types of predictive model. The approach will be illustrated by example, including study of how species traits mediate changes in environmental response across species (the "fourth corner problem").