Introduction to the methods of PRIMER

Application Areas

PRIMER 7 (Plymouth Routines In Multivariate Ecological Research) consists primarily of a wide range of univariate, graphical and multivariate routines for analysing arrays of species-by-samples data from community ecology. Data are typically of abundance, biomass, % area (or line) cover, presence/absence etc, and arise in biological monitoring of environmental impact and more fundamental studies, e.g. of dietary composition. Also catered for are matrices of physical values and chemical concentrations, which are analysed in their own right or in parallel with biological assemblage data, 'explaining' community structure by physico-chemical conditions. The methods of this package make few, if any, assumptions about the form of the data (non-metric ordination and permutation tests are fundamental to the approach) and concentrate on approaches that are straightforward to explain. This robustness makes them widely applicable, leading to greater confidence in interpretation, and the transparency possibly explains why they have been adopted worldwide, particularly in marine science but also in terrestrial and freshwater ecology, forestry, soil science, etc. The statistical methods underlying the software are explained in non-mathematical terms in the accompanying methods manual (Change in Marine Communities, 3rd edition, 2014), which also shows outcomes from many literature studies, e.g. of environmental effects of oil spills, drilling mud disposal and sewage pollution on soft-sediment benthic assemblages, disturbance or climatic effects on coral reef composition or fish communities, more fundamental biodiversity and community ecology patterns, mesocosm studies with multi-species outcomes, etc. Many of the data sets used in the methods manual (abbreviated to CiMC), and all of those used in this User Manual/ Tutorial are available with the installation so that the user can replicate the analyses.

Though the analysis requirements for biological assemblage data are a principal focus, the package is equally applicable (and increasingly being applied) to other data structures which are either multivariate or can be treated as such. These include: multiple biomarkers in ecotoxicology, and their relation to water or tissue concentrations of chemical contaminants; composition of substrate in geology or materials science; morphometric measurements in taxonomy; genetic studies and especially microbial analyses of large numbers of OTUs; signals at multiple wavelengths in remote sensing; even environmental economics, state variables in complex mathematical box models, acute medicine, epidemiology, etc. Univariate measurements which can sometimes be treated more effectively in a multivariate way include particle size analysis for water or sediment samples and size frequency distributions of organisms in cohort studies (the multivariate variables are the discrete particle or organism size classes). Sets of growth curves for individual organisms tracked through time (repeated measures, thus correlated) can also be handled. The unifying feature is that all data sets are reduced to an appropriate triangular matrix representing the resemblance of every pair of samples, in terms of their assemblages, suites of biomarkers, particle size distributions, shape of growth curves, etc. Clustering and ordination techniques are then able to display the relationships among the samples, and permutation tests impose a necessary hypothesis testing structure.

Basic routines

The routines of the package cover: data pre-treatment (transforms, dispersion- and other variable-weighting, assessed by Shade plots); about 50 resemblance measures (now allowing missing data); hierarchical clustering of samples (or species) *via* standard agglomerative and novel divisive and 'flat' techniques; ordination by non-metric (*n*MDS) and metric multidimensional scaling (*m*MDS, *tm* MDS), and principal components (PCA), to summarise patterns in biotic and abiotic samples; permutation-based hypothesis testing (ANOSIM, also in novel ordered form), testing *a priori* group structures of multivariate samples, from different times/locations/treatments, etc; a strong emphasis now on species patterns (novel Coherence curves and Shade plots); linking of multivariate biotic patterns to suites of environmental data or other biotic arrays (BEST, LINKTREE); comparative (Mantel-type) tests of similarities to model structures (RELATE, including novel 2-way forms); second-stage analyses (2STAGE) for 'repeated measures' and comparison of analysis choices (of taxonomic level, transform, resemblance coefficient); suites of diversity indices, dominance plots, SAD curves, species accumulation estimators, taxonomic aggregation, etc., and tests for biodiversity indices based on taxonomic distinctness of species (TAXDTEST); novel region estimates for mean communities from multivariate bootstrapping; and a wide range of other data manipulation and graph types (bar, line, mean, box, scatter, surface, histogram and shade plots), new to PRIMER 7.

Revision #12
Created 12 March 2024 22:34:15 by Arden
Updated 7 January 2025 21:23:20 by Abby Miller