

PRIMER 7 Online Workshop

Outline of Topics

Each **lecture topic** below is followed by a **computer practical** session where participants explore the topic using literature/published datasets.

1	Properties of multivariate data, measures of resemblance (similarity/dissimilarity/distance) for biotic and environmental data types, including shade plots to assess effects of pre-treatment options (standardisation, transformation, normalisation), and guidelines for choosing appropriate options for different data types.
2	Overview of clustering methods, with special focus on hierarchical agglomerative clustering of samples (CLUSTER). Includes discussion of a global test for the presence of any multivariate structure for any given set (or sub-set) of biotic or abiotic samples, using similarity profiles (SIMPROF tests).
3	Ordination (for environmental data) by principal components analysis (PCA).
4	Ordination (for biotic data) by non-metric multi-dimensional scaling (nMDS) and MDS diagnostics (e.g., Shepard diagram, stress, minimum spanning tree, cluster overlay) for assessing its adequacy. Compare and contrast non-metric MDS with metric MDS (mMDS) and threshold-metric MDS (tmMDS).
5	Global hypothesis tests of no agreement between two resemblance matrices (RELATE), e.g., to examine relationship between dissimilarity patterns created using biotic vs environmental data. Creating model matrices to relate biotic (or environmental) structure with linear or cyclical models in space or time.
6	Dispersion weighting to down-weight highly clumped/schooling species having erratic abundances over replicates at the same time/place. Lab session also includes a method for 'fixing' collapsed nMDS plots.
7	Non-parametric multivariate tests for differences among <i>a priori</i> groups of samples using analysis of similarities (ANOSIM , global and pairwise tests). Ordination plots to examine multivariate averages. Bootstrap methods to approximate regions of spread for multivariate group averages in mMDS ordinations.
8	ANOSIM tests for factors with ordered levels and for multi-way designs (up to 3 factors).
9	Relate overall multivariate structures and patterns in biotic data with potential environmental drivers, finding optimal matching subsets (the BEST procedure). Test the null hypothesis of no biotic-environment relationship, allowing for the selection procedure designed to find an optimum match (global BEST test).
10	Link important breaks or shifts in biotic data with potential environmental drivers <i>via</i> non-parametric constrained divisive clustering (LINKTREE). Consider also the unconstrained non-parametric divisive clustering method (UNCTREE).
11	Species' contributions to sample patterns: (a) BVSTEP as a step-wise form of BEST to identify <u>subsets</u> of species that reconstruct multivariate patterns among samples obtained using <u>all</u> species, and (b) species' percentage (SIMPER) contributions to sample similarities <i>within</i> groups and dissimilarities <i>between</i> groups.
12	Measuring relationships among species using an index of association . Identify groups of species (or other variables) that show coherent (highly similar) responses across samples.
13	Diversity measures (DIVERSE); multivariate treatment of multiple indices; dominance plots; taxonomic (or phylogenetic) diversity and distinctness; sampling properties and tests by comparison with regional lists (TAXDTEST).
14	Second-stage analysis (2STAGE) to generate "resemblances among resemblance matrices". Compare patterns obtained using various pre-treatment options (e.g., taxonomic resolutions, transformations, etc.). Consider ANOSIM tests from 2STAGE resemblance matrices to examine the extent to which multivariate patterns (in space and/or time) are "mirrored" across levels of another factor or factors.
15	Any methods that have not arisen in earlier discussion (e.g., further resemblance options: modifying Bray-Curtis for denuded samples; missing data; functional dissimilarity, etc.). Wrap-up of the week with an overview of the PRIMER tools.
16	Break-out sessions for the analysis of 'own data', in consultation with the presenter and other participants.

PRIMER 7 Online Workshop

Provisional Time-Table

The timetable below is a rough guide only. Lectures and labs may flow over or under allotted time-slots, depending on the depth of coverage of specific topics, the number and length of participant-led questions and ensuing discussions, etc. The flow between lectures and labs will be seamless. All times below are given in U.K. Time (GMT/UTC + 0 hrs).

	Monday	Tuesday	Wednesday	Thursday	Friday
Session 1 09:00 – 11:00	(1) Resemblance measures; pre-treatment options	(4) Ordination with nMDS; mMDS	(7) ANOSIM; Bootstrap averages	(11) Species' contributions; BIOENV; SIMPER	(15) Wrap-up; overview of PRIMER v7
Coffee Break 11:00 – 11:30					
Session 2 11:30 – 13:00	(2) CLUSTER; SIMPROF	(4) Ordination with nMDS; mMDS (cont'd)	(8) Ordered and multi-way ANOSIM	(12) Coherent species; SIMPROF	(16) 'Own-data' session
Lunch 13:00 – 14:00					
Session 3 14:00 – 16:00	(2) CLUSTER; SIMPROF (cont'd)	(5) RELATE; seriation or cyclical models	(9) BEST; global test	(13) DIVERSE; dominance plots; TAXDTEST	(16) 'Own-data' session (cont'd)
Coffee Break 16:00 – 16:30					
Session 4 16:30 – 18:00	(3) Ordination with PCA	(6) Dispersion weighting; "fix" nMDS	(10) LINKTREE; UNCTREE	(14) Second-stage analyses; 2STAGE	(16) 'Own-data' session (cont'd)

Throughout, participants will be given real data sets to analyse, but they are also encouraged to bring their own data. These should be in numeric, rectangular arrays, with variables (e.g. species) as rows, samples as columns (or *vice-versa*), in an Excel spreadsheet or text file. Non-numeric information (factors) on each sample are placed below (or to the side of) this table, separated by a blank row (or blank column). There is also a 3-column format (sample label, variable label, non-zero entry) suitable for entry from large record-type databases. Participants will have the opportunity (during the 'own-data' sessions scheduled for Friday) to connect with other participants and to discuss their own data, projects, sampling designs and analyses in direct consultation with the presenter.