PRIMER-e
Empowering Research

OUTLINE of TOPICS – PRIMER WORKSHOP

*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**), comparing assemblage (or environmental) structure with linear or cyclic models in space and time.
6	**Dispersion weighting** to down-weight highly clumped/schooled 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 a priori 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 (**BIOENV**), finding optimal matching subsets (the **BEST** procedure). Test the null hypothesis of no biotic-environment relationship, allowing for the selection procedure (global **BEST** test).
10	Link important breaks or shifts in biotic data with potential environmental drivers via 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 subsets of species that reconstruct multivariate patterns among samples obtained using all species, and (b) species' percentage (**SIMPER**) contributions to sample similarities within groups and dissimilarities between 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; comparisons of samples with regional lists (**TAXDTEST**).
14	Second-stage analysis (**2STAGE**) to generate "resemblances among resemblance matrices". Compare patterns obtained using various pre-treatment options / taxonomic resolutions. ANOSIM tests from 2STAGE matrices to examine how patterns (in space and/or time) are "mirrored" across levels of another factor.
15	Univariate non-parametric tests: Wilcoxon, Mann-Whitney U test, Kruskal-Wallis, Kolmogorov-Smirnov, Tests of Association.
16	Any methods that have not arisen in earlier discussion (e.g., further resemblance options: modifying Bray-Curtis for denuded samples; taxonomic dissimilarity measures, etc.). Wrap-up of the week with an overview of the **PRIMER** tools.
17	Q & A and 'Own-data' analysis session, in consultation with the presenter.

PRIMER-e
Empowering Research

PROVISIONAL TIME-TABLE – PRIMER WORKSHOP

The time-table below is a rough guide. 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.

	Monday	Tuesday	Wednesday	Thursday	Friday
Session 1 08:30 – 10:30	(1) Resemblance measures; pre-treatment options	(4) Ordination with nMDS; mMDS	(7) ANOSIM; Bootstrap averages	(11) Species' contributions; BVSTEP; SIMPER	(15) Univariate Non-parametric Tests
Coffee Break 10:30 – 11:00					
Session 2 11:00 – 12:00	(2) CLUSTER; SIMPROF	(4) Ordination with nMDS; mMDS (cont'd)	(8) Ordered and multi-way ANOSIM	(12) Coherent species; SIMPROF	(16) Wrap-up; overview of PRIMER
Lunch 12:00 – 13:00					
Session 3 13:00 – 15:00	(2) CLUSTER; SIMPROF (cont'd)	(5) RELATE; seriation or cyclical models	(9) BIOENV; BEST; global test	(13) DIVERSE; tax. diversity;TAXDTEST	(17) 'Own-data' session
Coffee Break 15:00 – 15:30					
Session 4 15:30 – 17:15	(3) Ordination with PCA	(6) Dispersion weighting; "fix" nMDS	(10) LINKTREE; UNCTREE	(14) Second-stage analyses; 2STAGE	(17) 'Own-data' session (cont'd)

Throughout, participants will be given real data sets to analyse, but they may also wish to bring/discuss 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.