



Advanced PRIMER 7 / PERMANOVA+ WORKSHOP

in St Petersburg, Florida, USA

Multivariate Analysis in Ecology (& other Sciences)

Presenter: Distinguished Prof. Marti J. Anderson

Venue: Weedon Island Preserve Visitors' Center, St Petersburg, Florida, USA

Dates: March 4th-8th, 2019

PROVISIONAL SCHEDULE

Monday, March 4th

- 08:15-08:30 Introduction
- 08:30-10:15 *Mini-review* – Thumbnail sketch of a typical exploratory multivariate analysis pathway in PRIMER (Data > Transform > Resemblance > Cluster > MDS); the wizards.
- 10:15-10:30 **Coffee break**
- 10:30-12:15 *Get the data in and check it out* – Data types, transformations (including shade plots, histograms, etc.), standardisations, normalisation, overview of dissimilarity measures and their properties, handling of missing data, dispersion weighting, zero-adjusted Bray-Curtis.
- 12:15-13:30 **Lunch break**
- 13:30-15:15 *Cluster analysis and recent extensions* – flexible beta, similarity profiles (SIMPROF) for assessing “significant” splits (Type I) and derived factors.
- 15:15-15:30 **Coffee break**
- 15:30-17:15 *Cluster analysis, continued* – unconstrained divisive method (UNCTREE), non-hierarchical method (KRCLUSTER), constrained clustering (LINKTREE).

Tuesday, March 5th

- 08:15-10:15 *MDS and beyond, part 1* – Shepard plots; new diagnostics and plotting options for non-metric MDS (minimum spanning trees; display of iteration progress and time-series animations; 2d & 3d cluster overlays).
- 10:15-10:30 **Coffee break**
- 10:30-12:15 *MDS and beyond, part 2* – Metric MDS, threshold metric MDS; combined MDS, including a new feature to “Fix collapse”; bootstrap averages for visualising variation in averaged data.
- 12:15-13:30 **Lunch break**
- 13:30-15:15 *Species analyses* – clustering and ordering options for axes of the new shade plots; new testing methods (SIMIPROF Type 3) producing coherent variable curves; segmented bubble plots.
- 15:15-15:30 **Coffee break**
- 15:30-17:15 *Non-parametric tests* – ANOSIM (ordered or unordered factors and up to 3-way designs), RELATE (including tests for seriation and cyclicity), BEST (finding combinations of environmental variables which together create distances to match biotic dissimilarities).

Wednesday, March 6th

- 08:15-10:15 *PERMANOVA, part 1* – ANOVA as a partitioning; extend to a multivariate dissimilarity-based geometric approach, with tests (p -values) by permutation.
- 10:15-10:30 **Coffee break**
- 10:30-12:15 *PERMANOVA, part 2* – Multi-factorial designs – nested vs crossed; fixed vs random; expectations of mean squares; components of variation; contrasts; asymmetrical designs
- 12:15-13:30 **Lunch break**
- 13:30-15:15 *PCO* – (including a brief review of PCA); distances among centroids for multifactorial designs; Monte Carlo P -values in PERMANOVA

- 15:15-15:30 **Coffee break**
- 15:30-16:30 *PERMDISP* – Tests for homogeneity of multivariate dispersions; permutation of residuals; analyses of beta diversity;
- 16:30-17:15 *Other topics in diversity* – Taxonomic diversity, Functional diversity, Phylogenetic diversity, Taxonomic distinctness (AvTD, VarTD), Taxonomic dissimilarities, cophenetic distances.

Thursday, March 7th

- 08:15-10:15 *DISTLM and dbrDA, part 1* – Dissimilarity-based linear models; multivariate multiple regression; constrained ordination.
- 10:15-10:30 **Coffee break**
- 10:30-12:15 *DISTLM and dbrDA, part 2* – model selection criteria (adjusted R², AIC, AICc, BIC) and model selection procedures (forward, backwards, step-wise and “Best” model selection searches).
- 12:15-13:30 **Lunch break**
- 13:30-15:15 *CAP, part 1* – canonical analysis of principal coordinates; discriminant analysis based on distances; classification of new samples using canonical models.
- 15:15-15:30 **Coffee break**
- 15:30-17:15 *CAP, part 2* – Predicting positions along a known gradient; Canonical correlation analysis; Ecosystem health models; monitoring programmes.

Friday, March 8th

- 08:15-10:15 *Synthesis* – overview of methods (when do I use what?); *Gaining multiple insights* - example application of multiple complementary methods learned across the PRIMER/PERMANOVA+ suite of tools, applied to a single ecological dataset.
- 10:15-10:30 **Coffee break**
- 10:30-12:15 Analysis of your own data* using PRIMER and PERMANOVA+, in consultation with the lecturer.
- 12:15-13:30 **Lunch break**
- 13:30-15:15 Analysis of your own data, continued.
- 15:15-15:30 **Coffee break**
- 15:30-17:15 Analysis of your own data, continued.

* Throughout, participants will be given real data sets to analyse, but they may also wish 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 sets of 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 very large arrays. Participants should take the opportunities of all lab sessions and breaks to discuss their own analyses with the lecturer, and not leave all their questions until the final afternoon session!

Some key and well-cited papers on PRIMER and PERMANOVA+ methodology

PRIMER

- Clarke KR (1990) Comparisons of dominance curves. *J Exp Mar Biol Ecol* 138: 143-157
- Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol* 18: 117-143
- Clarke KR (1999) Non-metric multivariate analysis in community-level ecotoxicology. *Environ Toxicol Chem* 18: 118-127
- Clarke KR, Ainsworth M (1993) A method of linking multivariate community structure to environmental variables. *Mar Ecol Prog Ser* 92: 205-219
- Clarke KR, Chapman MG, Somerfield PJ, Needham HR (2006) Dispersion-based weighting of species counts in assemblage analyses. *Mar Ecol Prog Ser* 320: 11-27
- Clarke KR, Gorley RN (2001, 2006, 2015) *PRIMER v5, v6, v7: User manual/tutorial*. PRIMER-E, Plymouth, UK, 91pp, 192pp, 296pp
- Clarke KR, Green RH (1988) Statistical design and analysis for a 'biological effects' study. *Mar Ecol Prog Ser* 46: 213-226
- Clarke KR, Somerfield PJ, Airoidi L, Warwick RM (2006) Exploring interactions by second-stage community analyses. *J Exp Mar Biol Ecol* 338: 179-192
- Clarke KR, Somerfield PJ, Chapman MG (2006) On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages. *J Exp Mar Biol Ecol* 330: 55-80
- Clarke KR, Somerfield PJ, Gorley RN (2008). Exploratory null hypothesis testing for community data: similarity profiles and biota-environment linkage. *J Exp Mar Biol Ecol* 366: 56-69
- Clarke KR, Somerfield PJ, Gorley RN (2016) Clustering in non-parametric multivariate analyses. *J Exp Mar Biol Ecol* 483: 147-155.
- Clarke KR, Tweedley JR, Valesini FJ (2014) Simple shade plots aid better long-term choices of data pre-treatment in multivariate assemblage studies. *J Mar Biol Assoc UK* 94: 1-16
- Clarke KR, Warwick RM (1994, 2001, 2014) *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation*. PRIMER-E, Plymouth, UK. 1st ed: 144pp; 2nd ed: 172pp. 3rd ed: (authors: Clarke KR, Gorley RN, Somerfield PJ, Warwick RM) 260pp
- Clarke KR, Warwick RM (1998) Quantifying structural redundancy in ecological communities. *Oecologia* 113: 278-289
- Clarke KR, Warwick RM (1998) A taxonomic distinctness index and its statistical properties. *J Appl Ecol* 35: 523-531
- Clarke KR, Warwick RM (2001) A further biodiversity index applicable to species lists: variation in taxonomic distinctness. *Mar Ecol Prog Ser* 216: 265-278
- Field JG, Clarke KR, Warwick RM (1982) A practical strategy for analysing multispecies distribution patterns. *Mar Ecol Prog Ser* 8: 37-52
- Somerfield PJ, Clarke KR (1995) Taxonomic levels, in marine community studies, revisited. *Mar Ecol Prog Ser* 127: 113-119
- Somerfield PJ, Clarke KR (2013) Inverse analysis in non-parametric multivariate analyses: distinguishing groups of associated species which covary coherently across samples. *J Exp Mar Biol Ecol* 449: 261-273
- Somerfield PJ, Clarke KR, Olsford F (2002) A comparison of the power of categorical and correlational tests applied to community ecology data from gradient studies. *J Anim Ecol* 71: 581-593
- Warwick RM, Clarke KR (1991) A comparison of some methods for analysing changes in benthic community structure. *J Mar Biol Ass UK* 71: 225-244
- Warwick RM, Clarke KR (1993) Increased variability as a symptom of stress in marine communities. *J Exp Mar Biol Ecol* 172: 215-226
- Warwick RM, Clarke KR (1995) New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. *Mar Ecol Prog Ser* 129: 301-305
- Warwick RM, Clarke KR (1998) Taxonomic distinctness and environmental assessment. *J appl Ecol* 35: 532-543
- Warwick RM, Clarke KR (2001) Practical measures of marine biodiversity based on relatedness of species. *Oceanog Mar Biol Ann Rev* 39: 207-231

PERMANOVA+

- Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26: 32-46
- Anderson MJ (2001) Permutation tests for univariate or multivariate analysis of variance and regression. *Can J Fish Aquat Sci* 58: 626-639
- Anderson MJ (2006) Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* 62: 245-253
- Anderson MJ (2008) Animal-sediment relationships revisited: characterising species' distributions along an environmental gradient using canonical analysis and quantile regression splines. *J Exp Mar Biol Ecol* 366: 16-27
- Anderson MJ (2017) Permutational Multivariate Analysis of Variance (PERMANOVA). Wiley StatsRef: Statistics Reference Online. 1–15. DOI: 10.1002/9781118445112.stat07841
- Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL, Sanders NJ, Cornell HV, Comita LS, Davies KF, Harrison SP, Kraft NJB, Stegen JC, Swenson NG (2011) Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecol Lett* 14: 19-28.
- Anderson MJ, Connell SD, Gillanders BM, Diebel CE, Blom WM, Landers TJ, Saunders JE (2005) Relationships between taxonomic resolution and spatial scales of multivariate variation in kelp holdfast assemblages. *J Anim Ecol* 74: 636-646
- Anderson MJ, Diebel CE, Blom WM, Landers TJ (2005) Consistency and variation in kelp holdfast assemblages: spatial patterns of biodiversity for the major phyla at different taxonomic resolutions. *J Exp Mar Biol Ecol* 320: 35-56
- Anderson MJ, Ellingsen KE, McArdle BH (2006) Multivariate dispersion as a measure of beta diversity. *Ecol Lett* 9: 683-693
- Anderson MJ, Gorley RN, Clarke KR (2008) *PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods*. PRIMER-E: Plymouth, UK, 214pp
- Anderson MJ, Gribble NA (1998) Partitioning the variation among spatial, temporal and environmental components in a multivariate data set. *Aust J Ecol* 23: 158-167
- Anderson MJ, Legendre P (1999) An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. *J Statist Comput Sim* 62: 271-303
- Anderson MJ, Millar RB (2004) Spatial variation and effects of habitat on temperate reef fish assemblages in northeastern New Zealand. *J Exp Mar Biol Ecol* 305(2): 191-221
- Anderson MJ, Robinson J (2003) Generalized discriminant analysis based on distances. *Aust NZ J Stat* 45: 301-318
- Anderson MJ, Robinson J (2001) Permutation tests for linear models. *Aust NZ J Stat* 43: 75-88
- Anderson MJ, Santana-Garcon J (2015) Measures of precision for dissimilarity-based multivariate analysis of ecological communities. *Ecol Lett* 18: 66-73.
- Anderson MJ, ter Braak CJF (2003) Permutation tests for multi-factorial analysis of variance. *J Statist Comput Sim* 73: 85-113
- Anderson MJ, Walsh DCI (2013) What null hypothesis are you testing? PERMANOVA, ANOSIM and the Mantel test in the face of heterogeneous dispersions. *Ecol Monogr* 83: 557-574.
- Anderson MJ, Walsh DCI, Clarke KR, Gorley RN, Guerra-Castro E (2017) Some solutions to the multivariate Behrens-Fisher problem for dissimilarity-based analyses. *Aust NZ J Stat*. 59: 57-79.
- Anderson MJ, Willis TJ (2003) Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology* 84: 511-525
- Legendre P, Anderson MJ (1999) Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecol Monogr* 69: 1-24
- McArdle BH, Anderson MJ (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology* 82: 290-297
- Paul WL, Anderson MJ (2013) Causal modelling with multivariate species data. *J Exp Mar Biol Ecol* 448: 72-84.