

**WORKSHOP on MULTIVARIATE ANALYSIS for
BIOLOGISTS, ECOLOGISTS and ENVIRONMENTAL SCIENTISTS:
Multivariate analysis of complex experimental designs using PERMANOVA+ for PRIMER v6**

13-17 December 2010, LaTrobe University

*Professor Marti J. Anderson
Institute for Information and Mathematical Sciences (IIMS),
Massey University, Albany Campus, Auckland, New Zealand*

Monday, 13 December

- 08:45-10:45 *Lecture: Part 1:* Introduction to the PRIMER environment for implementing PERMANOVA routines. *Part 2:* The nature of multivariate data and its properties; MANOVA, ANOSIM and extending distance-based approaches to more complex experimental designs; Permutational multivariate analysis of variance (PERMANOVA); some changes to underlying assumptions but retaining flexibility and robustness; testing interaction terms for multivariate data; logical choices for pair-wise comparisons
- 10:45-11:00 Coffee break
- 11:00-11:45 *Lecture:* Principal coordinate analysis (PCO) as another ordination technique to accompany direct analyses of dissimilarity matrices, its uses and its relationship with PCA and non-metric MDS
- 11:45-12:45 *Practical:* Introduction to the PERMANOVA+ add-on package to PRIMER v6: PERMANOVA, PCO and interpreting multivariate interactions
- 12:45-14:00 Lunch
- 14:00-15:45 *Lecture:* Permutational tests of homogeneity of multivariate dispersions (PERMDISP). Tests to accompany and help interpret the PERMANOVA tests for differences in location. Tests to examine dispersion issues in their own right, including beta diversity. Considerations regarding the use of different dissimilarity measures when analysing relative within-group dispersions among different groups
- 15:45-16:00 Coffee break
- 16:00-17:45 *Practical:* The use of PERMANOVA, PERMDISP and PCO together for interpreting differences among groups in dispersion and/or location, including tests for differences in beta diversity

Tuesday, 14 December

- 08:45-10:45 *Lecture:* Complex multi-factorial ANOVA experimental designs; fixed vs random factors; nested vs crossed relationships among factors; consequences for the hypothesis being tested and the extent of the inference; consequences for the expected mean squares and construction of appropriate pseudo- F ratios; estimating components of variation; multivariate analogues in PERMANOVA follow the univariate results; permutation tests for complex designs; exchangeable units; Monte Carlo P -values
- 10:45-11:00 Coffee break
- 11:00-12:45 *Practical:* Complex experimental designs using PERMANOVA; getting the model right to begin with matters a lot! Choosing appropriate pair-wise comparisons to do after fitting and analysing the full model; choice of relevant ordination graphics to accompany and interpret analyses
- 12:45-14:00 Lunch
- 14:00-15:00 *Lecture:* Unbalanced designs and designs that include covariates; non-independence of terms in the model and Types of Sums of Squares; consequences for expectations of mean squares; linear combinations of mean squares; tests, interpretations and inferences
- 15:00-15:45 *Practical:* Real examples and practice in the analysis and interpretation of results for unbalanced designs and designs with covariates
- 15:45-16:00 Coffee break
- 16:00-17:00 *Lecture:* Designs for detecting environmental impact; BACI and Beyond BACI; contrasts and asymmetrical designs; designs that lack replication, such as randomized blocks and repeated measures; pooling or removing terms from a model

17:00-17:45 *Practical*: Real examples and practice in analysing designs that lack replication and Beyond BACI asymmetrical designs

Wednesday, 15 December

08:45-10:45 *Lecture*: Analysing the relationship between species assemblage data and environmental variables; multivariate multiple regression; the distance-based linear model (DISTLM); procedures for model fitting (forward selection, backward elimination, step-wise fitting and a 'best' procedure); model selection criteria (R^2 , adjusted R^2 , AIC, AIC_c and BIC); marginal and conditional tests

10:45-11:00 Coffee break

11:00-12:45 *Practical*: Fitting multivariate regression models and model selection using DISTLM

12:45-14:00 Lunch

14:00-15:45 *Lecture*: Visualising regression models in a constrained ordination; distance-based redundancy analysis (dbRDA); the meaning and interpretation of the dbRDA axes; the use of biplot vectors and their interpretation; comparison with PCO

15:45-16:00 Coffee break

16:00-17:45 *Practical*: DISTLM and dbRDA

Thursday, 16 December

08:45-10:45 *Lecture*: Constrained and unconstrained ordination; canonical analysis of principal coordinates (CAP); generalised discriminant analysis based on distances; finding axes through the cloud of points that best discriminate among groups; leave-one-out allocation success

10:45-11:00 Coffee break

11:00-12:45 *Practical*: Constrained and unconstrained ordinations (CAP and PCO) for data with *a priori* groupings, understanding the diagnostics and the results of the analysis

12:45-14:00 Lunch

14:00-15:45 *Lecture*: Canonical correlation analysis using CAP; obtaining models of community change along environmental gradients; models of 'ecosystem health'; interpretation and uses; placement of new points into existing models; predictions; interpretation of biplot vectors; distinguishing the difference between CAP and dbRDA and when to use which one

15:45-16:00 Coffee break

16:00-17:45 *Practical*: Practice in the analysis of community data along environmental gradients and interpretation

Friday, 17 December

08:45-17:45 *Practical*: All day lab session and analysis of 'own data' using PERMANOVA+ and PRIMER v6

Throughout, participants will be given real data sets to analyse in the lab sessions, to exemplify the main points. However, it is anticipated that they will wish to bring some of their own data to the workshop, to analyse during this session, whilst the lecturer is on hand to give advice. Data should be in numeric, rectangular arrays, with variables (e.g. species) as rows, samples as columns, or vice-versa, in Excel or a 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.

Some well-cited and/or recent 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 Progr Ser* 92: 205-219
- Clarke KR, Chapman MG, Somerfield PJ, Needham HR (2006) Dispersion-based weighting of species counts in assemblage analyses. *Mar Ecol Progr Ser* 320: 11-27
- Clarke KR, Green RH (1988) Statistical design and analysis for a 'biological effects' study. *Mar Ecol Progr 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, 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 Progr Ser* 216: 265-278
- Field JG, Clarke KR, Warwick RM (1982) A practical strategy for analysing multispecies distribution patterns. *Mar Ecol Progr Ser* 8: 37-52
- Somerfield PJ, Clarke KR (1995) Taxonomic levels, in marine community studies, revisited. *Mar Ecol Progr Ser* 127: 113-119
- 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 Progr Ser* 129: 301-305
- Warwick RM, Clarke KR (2001) Practical measures of marine biodiversity based on relatedness of species. *Oceanog Mar Biol Annu 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, 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, Ford RB, Feary DA, Honeywill C (2004) Quantitative measures of sedimentation in an estuarine system and its relationship with intertidal soft-sediment infauna. *Mar Ecol Progr Ser* 272: 33-48
- Anderson MJ, Gorley RN, Clarke KR (2008) *PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods*. PRIMER-E: Plymouth, UK, 214 pp.
- 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, ter Braak CJF (2003) Permutation tests for multi-factorial analysis of variance. *J Statist Comput Sim* 73: 85-113
- 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