Appendix 2: Principal literature sources and further reading

A list of some of the core methods papers was given in the Introduction, and the source papers for the data used in examples can be found in Appendix 1. Here we itemize, for each chapter, the source of analyses which repeat those in published literature, and where figures have been redrawn from. Figures or analyses not mentioned can be assumed to originate with this publication. Also sometimes mentioned are historical references to earlier developments of the ideas in that chapter, or other useful background reading.

Chapter 1: Framework. The categorisation here is an extension of that given by Warwick (1988a) . The Frierfjord macrofauna data and analyses (Tables 1.2 & 1.6 and Figs. 1.1, 1.2 & 1.7) are extracted and re-drawn from Bayne, Clarke & Gray (1988) , Gray, Aschan, Carr *et al.* (1988) and Clarke & Green (1988) , the Loch Linnhe macrofauna data (Table 1.4 and Fig. 1.3) from Pearson (1975) , and the ABC curves (Fig. 1.4) from Warwick (1986) . The species abundance distribution for Garroch Head macrofauna (Fig. 1.6) is first found in Pearson, Gray & Johannessen (1983) , and the multivariate linking to environmental variables (Fig. 1.11) in Clarke & Ainsworth (1993) . The 'coherent species curves' (Fig. 1.10) for the Loch Linnhe data are redrawn from Somerfield & Clarke (2013) . The mescosm data from the nutrient enrichment experiment (Table 1.7) and the MDS plot for copepods and nematodes (Fig. 1.12) are extracted and redrawn from Gee, Warwick, Schaanning *et al.* (1985) .

Chapters 2 and 3: Similarity and clustering. These methods originated in the 1950's and 60's (e.g. Florek, Lukaszewicz, Perkal *et al.* (1951) ; Sneath (1957) ; Lance & Williams (1967)). The description here widens that of Field, Clarke & Warwick (1982) , with some points taken from the general texts of Everitt (1980) and Cormack (1971) . The dendrogram of Frierfjord macrofaunal samples (Fig.3.1) is redrawn from Gray, Aschan, Carr *et al.* (1988) , and the zooplankton example (Figs. 3.2 & 3.3) from Collins & Williams (1982) . The SIMPROF test for samples on agglomerative clusters is described in Clarke, Somerfield & Gorley (2008) ; Fig. 3.8 mimics one in Anderson, Gorley & Clarke (2008) , and the other cluster methods (unconstrained divisive and k-R clustering, maximising R) are somewhat new to this publication.

Chapter 4: Ordination by PCA. This is a founding technique of multivariate statistics, see for example Chatfield & Collins (1980) and Everitt (1978) . The MDS from a dosing experiment in the Solbergstrand mesocosms (Fig. 4.2) is from Warwick, Carr, Clarke *et al.* (1988) .

Chapter 5: Ordination by MDS. Non-metric MDS was introduced by Shepard (1962) and Kruskal (1964) ; two standard texts are Kruskal & Wish (1978) and Schiffman, Reynolds & Young (1981) . Here, the exposition parallels that in Field, Clarke & Warwick (1982) and Clarke (1993) ; the Exe nematode graphs (Figs. 5.1, 5.2, 5.4, 5.5) are redrawn from the former. The dosing experiment (Fig. 5.6) is discussed in Warwick, Carr, Clarke *et al.* (1988) . Metric MDS (see Cox & Cox (2001)), not to be confused with the similar, but not identical, PCO ordinations (produced by PERMANOVA+ for example), was also an early introduction but is much less commonly implemented in software. The combining of nMDS and mMDS stress functions bears some relationship to hybrid and semi-strong hybrid scaling methods (Faith, Minchin & Belbin (1987) , Belbin (1991)) but with some important differences in implementation and with a different rationale here (the avoidance of collapsed sub-groups in an MDS plot, and for two nMDS stress functions, the merging of similarities of different types); see footnote on page 5.8.

Chapter 6: Testing. The basic permutation test and simulation of significance levels can be traced to Mantel (1967) and Hope (1968) , respectively. In this context (e.g. Figs. 6.2 & 6.3 and eqt. 6.1) it is described by Clarke & Green (1988) . A fuller discussion of the extension to 2-way nested and crossed ANOSIM tests (including Figs. 6.4 & 6.6) is in Clarke (1993) (with some asymptotic results in Clarke (1988)); the coral analysis (Fig. 6.5) is in Warwick, Clarke & Suharsono (1990) , and the Tasmanian meiofaunal MDS (Fig. 6.7) in Warwick, Clarke & Gee (1990) . The 2-way design without replication (Figs. 6.8-6.12) is tackled in Clarke & Warwick (1994) ; see also Austen & Warwick (1995) . The ordered ANOSIM test is new to this publication, as are the extensions to 3-way crossed/nested designs. Lek, Fairclough, Platell *et al.* (2011) give the 'flattened' 2-way ANOSIM tests for the 3-way crossed example of labrid diets; Fig. 6.15 is redrawn from there. The NZ kelp holdfast data is provided with the PERMANOVA+ software (Anderson, Gorley & Clarke (2008)). Fig. 6.17 is partly extracted from Warwick, Ashman, Brown *et al.* (2002) .

Chapter 7: Species analyses. Clustering on species similarities is given in Field, Clarke & Warwick (1982) for the Exe nematode data; see also Clifford & Stephenson (1975) . SIMPROF test for species ('coherent curves') follows Somerfield & Clarke (2013) ; Figs. 7.1-7.6 are redrawn from there. Shade plots are described in Clarke, Tweedley & Valesini (2014) but have a very long history (see Wilkinson & Friendly (2008)), though there are some novelties in the options outlined here, in terms of combinations of input data, axis ordering, cluster analysis choices, and so on. The SIMPER (similarity percentages) procedure is given in Clarke (1993) , and the 2-way crossed SIMPER first used in Platell, Potter & Clarke (1998) . Simple bubble plots are a staple routine for graphical output but PRIMER 7's segmented bubble plots were first used in Stoffels, Clarke, Rehwinkel *et al.* (2014) and in Purcell, Rushworth, Clarke *et al.* (2014) .

Chapter 8: Univariate/graphical analyses. Pielou (1975) , Heip, Herman & Soetaert (1988) and Magurran (1991) are useful texts, summarising a large literature on a variety of diversity indices and ranked species abundance plots. The diversity examples here (Figs. 8.1 & 8.2) are discussed by Warwick, Platt, Clarke *et al.* (1990) and Warwick, Clarke & Suharsono (1990) respectively, and the Caswell *V* computations (Table 8.1) are from Warwick, Platt, Clarke *et al.* (1990) . The Garroch Head species abundance distributions (Fig. 8.4) are first found in Pearson, Gray & Johannessen (1983) ; Fig. 8.3 is redrawn from Pearson & Blackstock (1984) . Warwick (1986) introduced Abundance–Biomass Comparison curves, and the Loch Linnhe and Garroch Head illustrations (Figs. 8.7 & 8.8) are redrawn from Warwick (1986) and Warwick, Pearson & Ruswahyuni (1987) . The transformed scale and partial dominance curves of Figs. 8.9-8.11 were suggested by Clarke (1990) , which paper also tackles issues of summary statistics (Fig. 8.12, equation 8.7, and as employed in Fig. 8.13) and significance tests for dominance curves (the DOMDIS routine in PRIMER). Use of ANOSIM on distances among curves (growth curves, particle size distributions etc) has been advocated at PRIMER courses for some years and there are now a few examples in the literature. Similarly, the treatment of multiple diversity indices by multivariate methods, to ascertain the true (and limited) dimensionality of information captured, and the consistent (mechanistic) relationships between indices seen in ordination patterns (such as Fig. 8.16), has long been a staple of PRIMER courses, though never specifically published.

Chapter 9: Transformations. The chapter start is an expansion of the discussion in Clarke & Green (1988) ; Fig. 9.1 is recomputed from Warwick, Carr, Clarke *et al.* (1988) . Detailed description of dispersion weighting (DW) is in Clarke, Chapman, Somerfield *et al.* (2006) ; Figs. 9.2, 9.4 of the Fal nematode data (Somerfield, Gee & Warwick (1994a) and Somerfield, Gee & Warwick (1994b)) are redrawn from Clarke, Chapman, Somerfield *et al.* (2006) . The use of shade plots to aid transformation or DW choices is the topic of Clarke, Tweedley & Valesini (2014) . A different form of weighting of variables (by their standard deviation) is described in Hallett, Valesini & Clarke (2012) .

Chapter 10: Aggregation. This description of the effects of changing taxonomic level is based on Warwick (1988b) , from which Figs. 10.2-10.4 and 10.7 are redrawn. Fig. 10.1 is discussed in Gray, Aschan, Carr *et al.* (1988) , Fig. 10.5 and 10.8 in Warwick, Clarke & Suharsono (1990) and Fig. 10.6 in Gray, Clarke, Warwick *et al.* (1990) (or Warwick & Clarke (1993a) , in this categorisation). A methodology for examining the comparative effects on an analysis of choice of taxonomic level (and transform) can be found in Olsgard, Somerfield & Carr (1997) , Olsgard, Somerfield & Carr (1998) , and Olsgard & Somerfield (2000) .

Chapter 11: Linking to environment. For wider reading on this type of ‘canonical’ problem, see Chapter 5 of Jongman, ter Braak & Tongeren (1987) , including ter Braak (1986) 's method of canonical correspondence analysis. The approach here of performing environmental and biotic

analyses separately, and then comparing them, combines that advocated by Field, Clarke & Warwick (1982) : superimposing variables on the biotic MDS, and by Clarke & Ainsworth (1993) : the BIO-ENV program. The data in Table 11.1 is from Pearson & Blackstock (1984) . Fig 11.3 is redrawn from Collins & Williams (1982) and Fig. 11.6 from Field, Clarke & Warwick (1982) ; Figs. 11.7, 11.8, 11.10 and Table 11.2 are from Clarke & Ainsworth (1993) . The global BEST test is given in Clarke, Somerfield & Gorley (2008) , as is the description of linkage trees, the general idea of which (as 'multivariate regression trees') can be found in De'Ath (2002) . The modification to a constrained (2-way) BEST is new to this publication.

Chapter 12: Community experiments. Influential papers and books on field experiments, and causal interpretation from observational studies in general, include Connell (1974) , Hurlbert (1984) , Green (1979) and many papers by A J Underwood, M G Chapman and collaborators, in particular the Underwood (1997) book. Underwood & Peterson (1988) give some thoughts specifically on mesocosm experiments. Lab-based microcosm experiments on community structure, using this analysis approach, are typified by Austen & Somerfield (1997) and Schratzberger & Warwick (1998a) . Figs. 12.2 and 12.3 are redrawn from Warwick, Clarke & Gee (1990) and Figs. 12.5, 12.6 from Gee, Warwick, Schaanning *et al.* (1985) .

Chapter 13: Data requirements. The exposition parallels that in Warwick (1993) but with additional examples. Figs. 13.1-13.3 and 13.8 are redrawn from Warwick (1993) , and earlier found in Colebrook (1986) , Dawson-Shepherd, Warwick, Clarke *et al.* (1992) , Warwick (1988b) and Gray, Aschan, Carr *et al.* (1988) respectively. Fig. 13.4 is redrawn from Warwick, Clarke & Gee (1990) , Fig. 13.5 from Warwick, Platt, Clarke *et al.* (1990) , Fig. 13.6 from Warwick, Clarke & Suharsono (1990) and Fig. 13.7 from Warwick & Clarke (1991) .

Chapter 14: Relative sensitivities. This parallels the earlier sections of Warwick & Clarke (1991) , from which all these figures (except Figs. 14.11 & 14.14) have been redrawn. Primary source versions of the figures can be found as follows: Figs. 14.1-14.3, Gray, Aschan, Carr *et al.* (1988) ; Figs. 14.5-14.7, Warwick, Clarke & Suharsono (1990) ; Figs 14.9-14.10, Dawson-Shepherd, Warwick, Clarke *et al.* (1992)); Figs. 14.11-14.12, Gee & Warwick (1994a) and Gee & Warwick (1994b) ; Figs. 14.14-14.16, Austen & Warwick (1989) .

Chapter 15: Multivariate measures of disturbance and relating to models. The first part on multivariate measures of stress follows the format of Warwick & Clarke (1995a) and Warwick & Clarke (1995b) , and is an amalgamation of ideas from three primary papers: Warwick & Clarke (1993a) on 'meta-analysis' of NE Atlantic macrobenthic studies, Warwick & Clarke (1993b) on the increase in multivariate dispersion under disturbance, and Clarke, Warwick & Brown (1993) on the breakdown of seriation patterns. Figs. 15.1-15.3 and Table 15.1 are redrawn and extracted from

the first reference, Fig. 15.4 and Table 15.2 from the second and Figs. 15.5 & 15.6 and Table 15.5 from the third. The analysis in Table 15.4 is from Warwick, Ashman, Brown *et al.* (2002) . In the second part, the principle of matrix correlations using a Pearson coefficient dates to Mantel (1967) ; RELATE tests are a non-parametric form. The seriation test with replication is discussed in detail by Somerfield, Clarke & Olsgard (2002) , the Tees data is analysed in Warwick, Ashman, Brown *et al.* (2002) , the sea-loch data in Somerfield & Gage (2000) , the Gullfaks Fig. 15.10 is extracted from Somerfield, Clarke & Olsgard (2002) and the Leschenault Fig. 15.12 redrawn from Veale, Tweedley, Clarke *et al.* (2014) .

Chapter 16: Further multivariate comparisons and resemblance measures. The general extension of the Bio-Env approach of Chapter 11, to combinations other than selecting environmental variables to match biotic patterns, is described in Clarke & Warwick (1998a) . This details the forward/backward stepping search algorithm BVStep, and uses it to select subsets of 'influential' species from a biotic matrix. Second-stage MDS was defined by Somerfield & Clarke (1995) and early examples of its use can be found in Olsgard, Somerfield & Carr (1997) and Olsgard, Somerfield & Carr (1998) . Figs. 16.1 to 16.3, and Tables 16.1 and 16.2, are extracted from Clarke & Warwick (1998a) , and Fig. 16.5 from Somerfield & Clarke (1995) . The definition and behaviour of zero-adjusted Bray-Curtis is given by Clarke, Somerfield & Chapman (2006) , and that paper also discusses the relative merits of the resemblance measures covered here and introduces the use of second-stage MDS for comparing coefficients. Figs. 16.7 to 16.10 are a recalculated form of some of the figures of that paper; Fig. 16.11 expands the set of coefficients considered there. The very different use of second-stage analysis to generate 'interaction-type' plots is the subject of . Figs. 16.12 to 16.13 and 16.15 to 16.17 are redrawn from there.

Chapter 17: Taxonomic distinctness measures. Warwick & Clarke (1995b) first defined taxonomic diversity/distinctness. Earlier work, from a conservation perspective, and using different species relatedness properties (such as PD), can be found in, e.g. Faith (1992) , Faith (1994) , Vane-Wright, Humphries & Williams (1991) and Williams, Humphries & Vane-Wright (1991) . The superior sampling properties of average taxonomic distinctness ($\Delta ^ +$), and its testing structure in the case of simple species lists, are given in Clarke & Warwick (1998b) , and applied to UK nematodes by Warwick & Clarke (1998) and Clarke & Warwick (1999) . Variation in taxonomic distinctness ($\Lambda ^ +$) was introduced, and its sampling properties examined, in Clarke & Warwick (2001) , and a review of the area can be found in Warwick & Clarke (2001) , from which Figs. 17.1, 17.2, 17.5, 17.11, 17.12 are redrawn. Fig. 17.3 is discussed in Warwick & Clarke (1995b) , Fig. 17.4 in Warwick, Ashman, Brown *et al.* (2002) , Figs. 17.6, 17.8, 17.9, 17.14, 17.17 in Clarke & Warwick (2001) , Fig. 17.7 in Clarke & Warwick (1998b) and Figs. 17.10, 17.13 in Rogers, Clarke & Reynolds (1999) . Taxonomic dissimilarities are discussed in Clarke, Somerfield & Chapman (2006) , from which the two examples, Fig. 17.19, 17.20 are taken. The measures were first defined in

Clarke & Warwick (1998a) and Izsak & Price (2001) .

Chapter 18: Bootstrap average regions. Bootstrapping univariate data was introduced by Efron (1979) , see also Efron & Tibshirani (1993) . Its specific application to these complex multivariate contexts is new to this publication and might best be treated as experimental, for the moment. Certainly the nominal region coverage probabilities (e.g. 95%) should not be given a formal 95% confidence region interpretation, since some sources of uncertainty are, inevitably, not included in that probability statement – primarily how well the lower-dimensional region represents the higher-dimensional reality.

Revision #36
Created 20 March 2022 08:52:24 by Arden
Updated 20 November 2024 01:32:06 by Abby Miller