

Exploring interactions by second-stage community analyses

K. Robert Clarke ^{a,*}, Paul J. Somerfield ^a, Laura Airoidi ^b, Richard M. Warwick ^a

^a Plymouth Marine Laboratory, Prospect Place, West Hoe, Plymouth PL1 3DH, UK

^b Dipartimento di Biologia Evoluzionistica Sperimentale and Centro Interdipartimentale di Ricerca per le Scienze Ambientali di Ravenna, Università di Bologna, Via S. Alberto 163, I-48100 Ravenna, Italy

Accepted 14 June 2006

Abstract

Many biological data sets, from field observations and manipulative experiments, involve crossed factor designs, analysed in a univariate context by higher-way analyses of variance which partition out ‘main’ and ‘interaction’ effects. Indeed, tests for significance of interactions among factors, such as differing Before–After responses at Control and Impact sites, are the basis of the widely used BACI strategy for detecting impacts in the environment. There are difficulties, however, in generalising simple univariate definitions of interaction, from classic linear models, to the robust, non-parametric multivariate methods that are commonly required in handling assemblage data. The size of an interaction term, and even its existence at all, depends crucially on the measurement scale, so it is fundamentally a parametric construct. Despite this, certain forms of interaction can be examined using non-parametric methods, namely those evidenced by changing assemblage patterns over many time periods, for replicate sites from different experimental conditions (types of ‘Beyond BACI’ design) – or changing multivariate structure over space, at many observed times. Second-stage MDS, which can be thought of as an MDS plot of the pairwise similarities between MDS plots (e.g. of assemblage time trajectories), can be used to illustrate such interactions, and they can be formally tested by second-stage ANOSIM permutation tests. Similarities between (first-stage) multivariate patterns are assessed by rank-based matrix correlations, preserving the fully non-parametric approach common in marine community studies. The method is exemplified using time-series data on corals from Thailand, macrobenthos from Tees Bay, UK, and macroalgae from a complex recolonisation experiment carried out in the Ligurian Sea, Italy. The latter data set is also used to demonstrate how the analysis copes straightforwardly with certain repeated-measures designs.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Community study; Interaction; Multivariate analysis; Non-parametric; Repeated measures designs; Second stage MDS; Spatial patterns; Time series

1. Introduction

1.1. Interactions and non-parametric analysis

Many of the univariate study designs of importance in experimental or observational ecology (Underwood,

1997) have at least two crossed factors (sites and times, treatments and blocks etc.). It is necessary when drawing inference about how the response variable reacts to the different levels of factor A, to separate this from the effects of a further factor, B (and vice-versa). Equally important is to be able to identify and describe situations in which the effects cannot be untangled: the response to factor A differs with the levels of factor B (interaction). Univariate designs, handled by ANOVA, can become

* Corresponding author. Tel.: +44 1752 633100; fax: +44 1752 633101.
E-mail address: krc@pml.ac.uk (K.R. Clarke).

quite complex (Underwood, 1993), with many crossed factors and a hierarchy of main effects, 2-factor interactions, 3-factor interactions, etc. This paper, however, concerns community studies, in which the response is a multivariate matrix of counts, biomass, area cover etc. for a large number of different taxa. For such data, the standard normal linear model assumptions underlying ANOVA are most certainly not met. Widely used analyses here, particularly in marine science, are non-parametric (Clarke, 1993), with simple one- and two-way ‘analysis of similarity’ tests (ANOSIM, Clarke and Green, 1988) replacing ANOVA tests. The study of interactions is, however, problematic in a non-parametric context, even with simple univariate data. It is perhaps too little appreciated how the interpretation of higher-way ANOVA for a single response variable is crucially dependent on the precise measurement scale. The response needs to be parameterised as an additive linear combination of main effects, two-, three- and four-way interaction terms etc., and an additive ‘error’. Making a simple change to the measurement scale, for example recording log respiration instead of respiration itself, or converting a body length into a body volume, can radically alter the linear model. Interactions between experimental factors may appear or disappear, as can be demonstrated as follows.

Imagine a simple 2×2 BACI experiment (Green, 1979), in which the density of a key indicator species is measured Before and After impact, for both Control and Impacted conditions. Suppose that the mean densities are 1 (Before, Control), 10 (Before, Impacted), 10 (After, Control) and 100 (After, Impacted), with little or no variability across replicates within each of the four cells. The density has increased by 90 in the impacted condition after the impact, greatly exceeding a natural temporal difference of 9 for the control condition. Thus density, analysed by two-way ANOVA, will show a strong interaction, indicating an effect of the impact. Now change the measurement scale to $\log_{10}(\text{density})$, i.e. with the 4 cells having means 0, 1, 1, and 2, respectively. The analysis will reach an entirely different conclusion, of no interaction, and thus no impact. Yet $\log_{10}(\text{density})$ and density are monotonically related, i.e. in rank order (non-parametric) terms they are identical sets of readings. The moral is that the partitioning of variation between main effects and multi-way interactions, that we have become so used to from univariate ANOVA, is tied inexorably to linear models with additive errors on the chosen measurement scale. We must, therefore, expect little of this to carry over to the non-linear (non-parameterised, even) structure of non-metric MDS (Kruskal and Wish, 1978) and permutation tests on rank similarity matrices.

1.2. Interactions in multivariate analysis

Can we get anywhere at all with studying what might loosely be termed interactions in such a non-parametric multivariate analysis structure? For example, how can we infer that assemblage temporal patterns have changed, depending on which site is being observed, or that effects of a treatment on communities differ, depending on the site at which it is applied? In an important development, McArdle and Anderson (2001) and Anderson (2001) generalised standard univariate ANOVA models to test interaction effects in multivariate data, using (approximate) permutation tests to avoid grossly unrealistic normality assumptions. Their approach accommodates any similarity measure (more generally, ‘resemblance’ measure) that fits the biological context of a particular application (see, for example, Clarke et al., 2006). Unlike previous attempts at this problem, resemblance coefficients can be accommodated which are not formal distance measures (in the sense of satisfying the ultrametric inequality, Gower and Legendre, 1986). This includes the widely used Bray–Curtis similarity (Bray and Curtis, 1957) and many other ecologically useful coefficients.

The simple example above demonstrates, however, that this approach cannot be non-parametric. It is not a function of the ranks of the (dis)similarities but depends very precisely on the measurement scale (e.g. Bray–Curtis, chi-squared distance, Canberra, etc.) selected for those dissimilarities. Linear models with additive errors are fitted in high-dimensional ‘dissimilarity space’, using the familiar F ratios from ANOVA tables as test statistics. As yet, it is not clear how one interprets the magnitude (and ‘direction’) of such interaction effects, or carries out model diagnostics—assessing the validity of the limited set of linear model assumptions that are being made. Whilst this remains the most promising avenue of exploration for community ecologists needing formal tests for higher-way interaction effects, it is equally important to ask how far non-parametric multivariate approaches (not employing a linear model structure) can be pressed in studying interactions. It is the purpose of this paper to explore one scenario in which such a fully robust, rank-based concept of interaction exists and can sometimes be tested.

2. Method

2.1. Non-parametric approach: second-stage analyses

Underwood (1992), following Stewart-Oaten et al. (1986), reasons that a BACI demonstration of impact, via interaction, becomes more convincing if a series of time

points are collected before and after the impact ('Beyond BACI' designs), so that the temporal change at the impact site, consequent on the impact, can be set in the context of the natural temporal variation of a control site (or preferably a series of control sites, Underwood, 1994). This paper takes a similar, if less formal, viewpoint, in considering the non-normal high-dimensional multivariate data arising from assemblage studies. The key step is the definition and analysis of second-stage resemblance matrices (Fig. 1), introduced by Somerfield and Clarke (1995), which describe relationships between multivariate patterns in conventional, 'first-stage', assemblage resemblance matrices. The schematic diagram (Fig. 1) demonstrates that if a modest number of time points are available at a series of sites, enough to construct a time profile of changing community structure at each site, then a fully non-parametric second-stage analysis of these first-stage multivariate structures provides a plot which can be used to identify anomalous profiles. This leads on to a test for differences between *a priori* defined groups of profiles. Outwith a formal linear model structure, this reflects a more general definition of what is meant by 'interaction' in a multivariate context: the assemblage structure has a different pattern of change through time at different sites. If a series of randomly

chosen sites are examined over a set of times, those sites being divided into replicates of a control and of a putatively impacted condition (the Underwood 'Beyond BACI' context), then in the presence of a real impact we should expect to see a consistent time pattern for the replicates within either of the two 'treatments', but a very different time profile between them. This would lead to a formal statistical test of interaction, and thus of the presence of an impact, using a fully non-parametric, one-way ANOSIM test on the second stage similarities.

Two-way crossed ANOVA designs are symmetric, in the sense that the two factors can usually be interchanged and the analysis steps repeated. For a sites by times study, in which there are a sufficient number of sites to define, for each time point, a multivariate pattern of inter-site relationships based on their assemblages, then the inter-correlations between these site patterns lead to a second-stage MDS of the times. Each (time) point on this plot represents the multivariate assemblage structure of the full set of sites, adjacent points implying that the inter-relationships between the sites is the same at those two time points. Anomalous points on this plot are ones for which the pattern of community differences across the sites is not the same as at other times, again indicating an interaction between time and site.

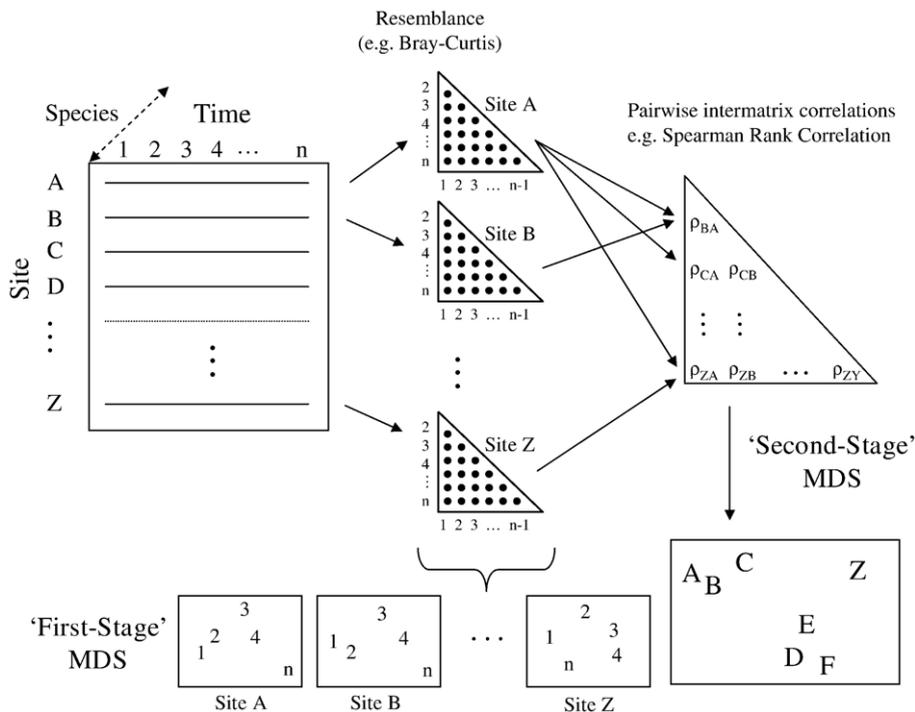


Fig. 1. Schematic diagram of the stages in quantifying and displaying agreement in multivariate pattern between samples collected through time at different sites. The 'second-stage' MDS plot summarises concordance in pattern, with sites at which changes through time are similar (no Site × Time interaction) grouping together.

This basic idea is developed for three diverse, practical community studies in the marine environment (see Section 2.2), corresponding to increasingly structured observational or experimental layouts. Firstly, the spatial structure of a tropical coral community along an onshore–offshore transect is contrasted for a series of different years. Then changes through time in the community structure of soft-sediment macrobenthos are contrasted for a series of sites in Tees Bay, UK. Finally a more structured experimental layout, of clearance and recovery of Mediterranean sublittoral algal communities, is examined. The latter again matches changes in assemblages through time (recolonisation) but this time for differing experimental conditions, namely the different times of year at which clearance was executed. This example also makes another important general point, that this form of second-stage analysis is able to accommodate ‘repeated measures’ designs, which are notoriously difficult to analyse in conventional higher-way ANOVA designs for univariate responses. In this experiment, each patch of cleared rock (replicate) was monitored through time to generate the recolonisation profile—the successive time points therefore constitute repeated measures (and thus serially correlated) data. But the second-stage analyses, testing for differences between treatments (clearance time of year), are performed on the *profile* of community recovery, and use as replicates the profiles from different patches of rock (not the individual time points). The final points on the second-stage MDS plot (Fig. 1) *do*, therefore, correspond to independent (randomly chosen) experimental units, and ANOSIM tests can validly be applied to the second-stage similarities.

2.2. Example data

2.2.1. Reef corals, Phuket, Thailand

The reefs on the south east tip of Ko Phuket are characterised by wide reef flats which extend up to 200 m from the shore, where they terminate in a shallow fore-reef extending to a depth of 5 m. The data used here document a 17-year history of disturbances on these shallow reef flats (Brown et al., 2002). Reef sites established in 1983 have been variously subject to the effects of sedimentation due to dredging of a deep-water port in 1986–7 (Brown et al., 1990; Clarke et al., 1993) and to anomalously low sea levels associated with Indian Ocean dipole events in 1997–98 (Chambers et al., 1999; Webster et al., 1999). Reef flats at three sites (A close to the dredged deep-water port, B slightly further away and C a reference site on the other side of the bay) were monitored almost annually, using a permanently marked

onshore–offshore transect at each. Coral cover was measured by plotless line techniques (Brown et al., 1996). Site A comprises 12×10 m lines, Site B 17×10 m lines and Site C 12×10 m lines, with lines spaced at 10 m intervals perpendicular to the inner shore to outer reef-flat transect.

2.2.2. Soft-sediment macrobenthos, Tees Bay, UK

Abundances of macrobenthic species were monitored twice yearly (March and September) at six discrete survey areas in Tees Bay between 1973 and 1996 (Warwick et al., 2002). Water depths ranged from 15 to 22 m and sediments were very uniform, consisting predominately of very fine sand. A subset of these samples, the September samples from Areas 0, 1, 2, 3 and 4, are used here, since these areas were the only ones sampled every year for nearly the whole period (in fact, 1973–94). In each area, two sites were sampled, with five grabs from each site in the majority of cases, but never less than four. The areas are located at varying distances from the mouth of the industrialised River Tees, with Area 3 just off the river mouth and Area 0 furthest away. During the period of this study, a major change in the North Sea ecosystem occurred, of sufficient magnitude to have been termed a “regime-shift” (Reid et al., 2001a,b). All trophic levels in the system were affected. The causes of the regime-shift are still under debate, but its widespread symptoms are unequivocal. The influence of regional events (North Sea wide regime shift) appeared to dominate any evidence for local effects (Tees estuary outflows) in the observed temporal changes in community structure and composition (Warwick et al., 2002).

2.2.3. Rocky subtidal macroalgae, Ligurian Sea, Italy

Macroalgal colonisation of cleared patches of rock was studied at three stations at depths of 10–13 m on wave-exposed rocky reefs in the Ligurian Sea, south of Livorno, Italy (Airoldi, 2000). Patches measuring 15×15 cm were cleared of all organisms at eight randomly chosen times between October 1995 and September 1996 at three stations. At each station, three replicate patches were cleared at each of three sites, and for the purposes of this study the three replicate patches have been combined. After clearing, percentage cover of each algal species in the experimental patches was scored on six occasions over the following year, using photographic recording methods. Colonisation patterns of individual components (total crust-forming and turf forming algae, and individual dominant erect algal species) were analysed by univariate ANOVA, with evidence found for effects of ‘treatment’ (timing of clearance). These ANOVA analyses use only data from

a single time period after the date of clearance, or the average of a series of such periods, in order to avoid the ‘repeated measures’ problem. In our multivariate analyses of this study, we are able to exploit the full community structure simultaneously, and also the whole recovery profile. We focus on two of Airoldi’s (2000) stations, mainly her Station 2 (Calafuria) but also her Station 1 (Boccale).

2.3. Computations

All calculations use options available in the PRIMER software package (Plymouth Routines In Multivariate Ecological Research, version 6; Clarke and Gorley, 2006).

3. Results and discussion

3.1. Reef corals, Phuket, Thailand

Fig. 2 (left) shows the result of a non-metric MDS ordination on all coral community samples for transect A, using Bray–Curtis similarities calculated from square-root transformed percentage cover values for 60 species. An attempt to pick out the effects of the two factors, year (1983, 86, 87, 88, ..., 2000) and position down the shore (1, 2, 3, ..., 12) has limited success, onshore positions tending to be to the left of the diagram and offshore positions to the right (for clarity, spatial

profiles for only four years are identified). There are, however, far too many points and too much structure for a successful low-dimensional MDS configuration (stress is uncomfortably high, at 0.20). This is commonly the case with two factors, each of which has many levels. Untangling the effects of factor B when studying the levels of factor A was identified at the beginning of the paper as a motivating requirement, and there is, of course, a simple way of achieving this: display the onshore–offshore gradient in coral assemblage structure separately for each year (Fig. 2, right). All stress values are now tolerably low and most (but not all) years are seen to display a clear pattern of serial change. The primary issue of this paper is the elicitation of interaction effects: here they would be manifest as a *different* pattern of change in assemblages down the shore (or no change at all) for some years and not others. The arguments for measuring the degree to which two ordination patterns match, by calculating Spearman rank correlations between the matching elements of their underlying similarity matrices, were articulated by Clarke and Ainsworth (1993), and the principle of matrix correlations dates back to Mantel (1967). Somerfield and Clarke (1995) termed the procedure of computing sets of correlations between pairs of similarity matrices, and then analysing the resulting correlation matrix with multivariate methods, a ‘second-stage analysis’. Olsgard et al. (1997, 1998) and Clarke et al. (2006) used the idea extensively in assessing the relative

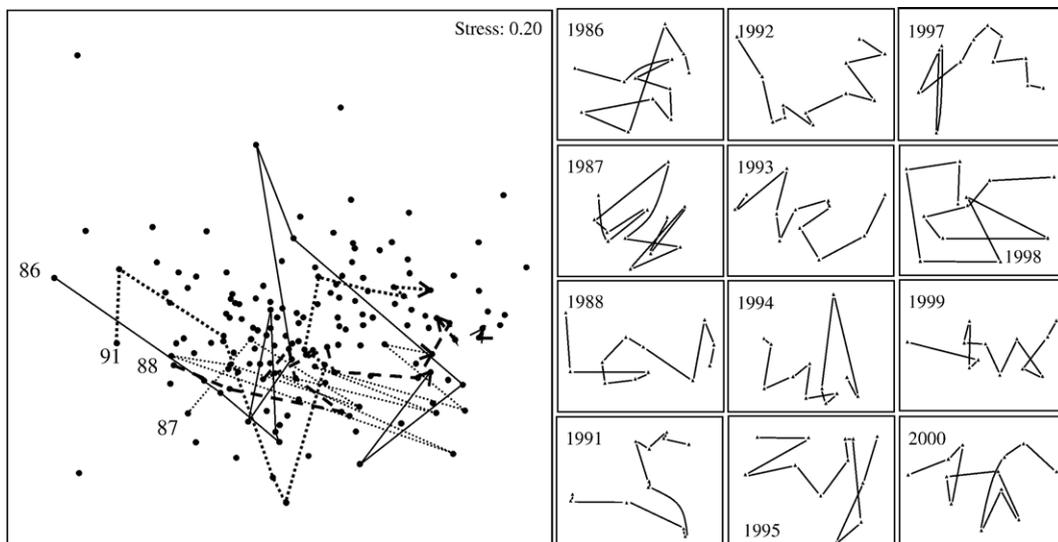


Fig. 2. Ko Phuket corals. Conventional ‘first-stage’ ordination by MDS (on left) of Bray–Curtis similarities between square-root transformed abundances of corals at 12 positions down the shore along Transect A, sampled on 13 occasions between 1983 and 2000. Positions of all samples are indicated by dots. Arrowed lines link samples along the transect in the years 1986, 1987, 1988 and 1991. The small panels (on right) are ‘first-stage’ ordinations of subsets of the data indicating the change in community structure along the transect in individual years (excluding 1983, when low abundance at the innermost station collapses the plot).

effects that choice of taxonomic identification level, data transformation and resemblance measure have on the resulting multivariate description. Here we use the Spearman correlation ρ_s computed between every pair of spatial profiles seen in Fig. 2, to produce a second-stage matrix of year-to-year ‘similarities’, which is subjected to multivariate analysis. This results in a second-stage cluster analysis and second-stage MDS plots (Fig. 3, top and mid). Two features are immediately apparent: the clear distinction of the two years 1986 and 1987 of sedimentation impact from dredging for the deep-water port, and the even clearer separation of 1998, the period of severe desiccation arising when a dipole event in the Indian Ocean lowered sea levels.

There is a minor technical point here concerning the definition of second-stage similarity. The resemblance measure is Spearman correlation but this, in theory, may take values over the range -1 to $+1$, whereas similarities are constrained to be positive (typically lying in the range 0 to 100). The issue did not arise in the very different context in which Olsgard et al. (1997, 1998) used second-stage analysis because the different analysis

choices inevitably led to multivariate patterns with some concordance, and the Spearman correlations were always positive. How best to turn negative correlations into positive similarities depends, to some extent, on the context. When carrying out ordination of variables, for example an MDS of Spearman correlations between environmental variables, it is usually preferable for points close together on the plot to indicate variables that are highly interdependent, whether that dependence is expressed as a positive or negative correlation. Similarity may then simply be defined as $100|\rho_s|$. Here, however, negative Spearman matrix correlations indicate spatial profiles that are so unlike each other that not only is there no relationship between the two multivariate patterns ($\rho_s=0$) but there is a tendency for the patterns to be positively antithetic ($\rho_s<0$), i.e. even less similar than random rearrangements of each other. In this case, similarity is best defined as $100(1+\rho_s)/2$, a value of 50 therefore implying no relation between the spatial profiles, and this is the scale used for the dendrograms of Fig. 3.

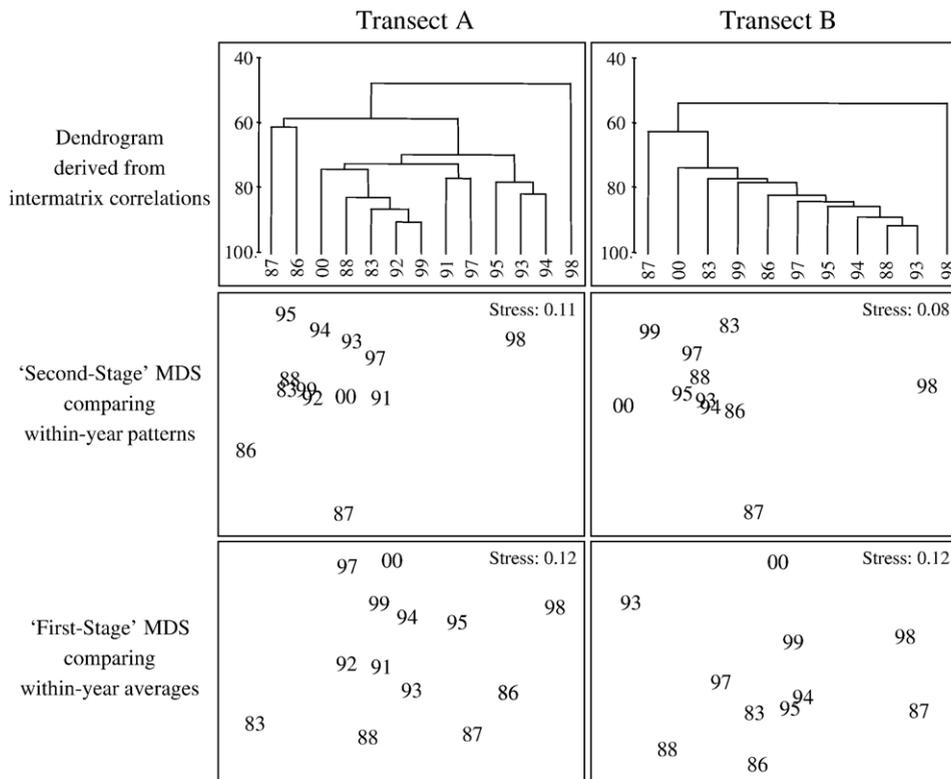


Fig. 3. Ko Phuket corals, Transects A (left) and B (right). Second-stage dendrograms (top) and MDS plots (mid), derived from Spearman correlations (converted to similarities, see text) between first-stage matrices of Bray–Curtis similarity coefficients. The latter are from percentage cover (square-root transformed) of coral species at 12 locations (transect A) and 17 locations (transect B), and the matrices are computed separately for each year. First-stage MDS plots (bottom) are based on Bray–Curtis similarities computed directly between years, using the square-root transformed average cover of each species, over all samples along the transect in each year.

It is also reasonable to ask whether a more simple-minded analysis would have demonstrated the same outcome: would the differences in years 1986, 1987 and in 1998 have been apparent from a conventional ('first-stage') MDS of the coral samples? It is certainly not evident in the first-stage plot containing all samples (Fig. 2), but a natural choice might be to average across the spatial factor and analyse the resulting temporal one. In the MDS from pooled assemblage samples from all positions down the transect (Fig. 3, bottom), i.e. the average community for each of the 13 years (again with square root transformed cover and Bray–Curtis similarities), the differences between years are much less apparent. This is also true when any single transect position is looked at across years.

There is no formal testing structure possible in this case, since the second-stage similarity matrix just consists of 13 samples, namely the 13 years. The conclusions are largely repeated, however, in analyses of data from transect B (Fig. 3, right). In the second-stage analysis, the year of the dipole event (1998) is pulled out as having a different interrelationship of community samples down the onshore–offshore transect compared to other years, as is the time of peak sedimentation from the dredging impact in 1987. (The later arrival of the disturbance to transect B than transect A for this event was noted by Clarke et al., 1993.) Again, the inter-annual variability in average community across the shore (first-stage MDS) does not differentiate these events to anything like the same extent. The reference site for the sedimentation impact, transect C (not shown), does not identify either 1986 or 1987 as distinctive in their spatial profile, but the event of 1998 – an area-wide change – is apparent.

Why is the second-stage MDS more instructive here than the first-stage MDS on average values or on single positions along the transect, and will this always be the case? The answer to the latter question is almost certainly 'no'. Sometimes the main inter-annual differences will be consistent across all sites, and the first-stage plot which averages over the sites will display those year-to-year differences clearly (the 'main effects', in standard statistical terminology). What the second-stage plot does is to *remove* these main-effect differences, because the input to the analysis completely excises all of the similarities between samples in different years, purely utilising similarities between samples *within* each year. The analysis is a purely relative one therefore and concentrates on showing whether the relationships amongst samples within each year differ between years, i.e. whether there are interaction effects. To characterise it more crudely, if there was a perfect serial pattern of

community change down the shore every year, but in some years the endpoint communities for the transect were quite different, the first-stage MDS of all samples in Fig. 2 would consist of a series of roughly parallel lines, and anomalous years would have lines which were well-separated from 'normal' years. First-stage plots would be effective for detecting this difference, and second-stage analyses completely uninformative. If, on the other hand, impacts manifested themselves mainly as changes in the pattern down the shore (because, for example, the impacts affected the deeper samples or mainly the near-shore), then the demonstration of impact would be clearer from the second-stage plot (the 'interactions'), the uninformative variability from the 'main effects' having been removed. This appears to be the case for the current data.

3.2. Soft-sediment macrobenthos, Tees Bay

The second example illustrates another sites \times times study but where the questions are about consistency of temporal patterns across sites (rather than consistency of spatial patterns over time). To be more precise, the hypothesis of interest is whether different areas in Tees Bay (Fig. 4) show different temporal patterns of community change. In particular, do those sites within the mixing zone of contaminant discharges from the Tees estuary (Area 3 and sometimes out to Area 2, Warwick et al., 2002) display a different inter-annual pattern to the coastal areas further removed from this influence, and bracketing it, namely Areas 0, 1 and 4? A strongly common pattern among time trajectories would imply that the dominant factor affecting the soft-sediment macrobenthos was a regional one, e.g. a changing hydrographic regime in the western North Sea. A difference of trajectories between areas would imply localised effects: in particular a contribution to coastal inter-annual variability from the major constructional works and patterns of contaminant discharge in the Tees estuary itself, over that time period (Warwick et al., 2002).

The difficulty of answering such questions by first-stage analyses will immediately be obvious: the 5 areas are widely spaced (10's of km) and though all are of similar depths and sedimentary types, the baseline benthic communities will inevitably differ between areas. In order to examine the time factor, the effect of the 'nuisance' site factor (site differences at one time) needs first to be removed. This is what is achieved by the second-stage procedure, in concentrating solely, as it does, on the similarities among times *within* each site.

Within each of the five areas of Tees Bay analysed here (Fig. 4), only two sites were consistently sampled

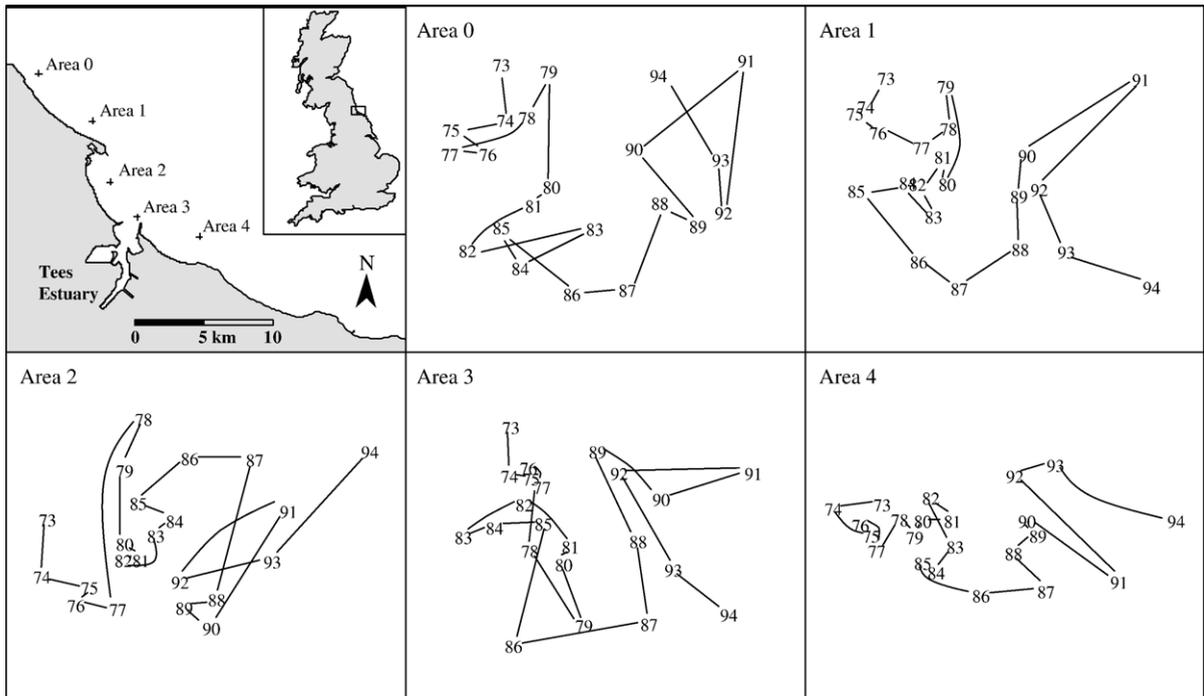


Fig. 4. Tees Bay macrofauna. The map indicates the locations of sampling areas 0, 1, 2, 3 and 4 in relation to the mouth of the Tees estuary. For each area, ‘first-stage’ MDS plots are based on Bray–Curtis similarities calculated from fourth-root transformed abundances. The latter are of averages over all replicates at two sites within each area, the similarities being computed among the years 1973 to 1994.

and replicated each September, over the period 1973–1994. The grab samples at each of these sites (typically 5, occasionally 4) were averaged before analysis. The plots in Fig. 4 show the MDS configurations of the time trajectories for the 5 areas, with data from the two sites in each area being averaged solely for clarity. All 10 sites were retained for the second-stage analysis. Fourth-root transformations were applied to the averaged densities of organisms per 0.1 m², for each of the 199 species observed, and Bray–Curtis similarities between samples input to MDS. The resulting annual series all show a relatively similar pattern, with the ‘regime shift’ in about 1987 (see Section 2.2.2) clearly apparent, superimposed on a gradual linear drift of the assemblage composition through the years.

Spearman rank correlations between the similarity matrices underlying each of these 10 time trajectories are the input to the second-stage analysis. All pairwise correlations are high, in absolute terms, ranging from 0.57 to 0.90. Not surprisingly, given the plots, all of them individually, and the average pairwise correlation, are significantly different from zero using non-parametric Mantel (1967) type permutation tests (the RELATE and ANOSIM2 tests in PRIMER, Clarke and Warwick, 2001). Regional patterns of temporal change are clearly

dominant here, but are there more subtle, local effects on the inter-annual trajectories?

The second-stage MDS plot (Fig. 5) answers this question. Points that are close together on this plot indicate similarity in time trajectories, removing any contribution from similarity in baseline communities at the 10 sites. One thing is immediately apparent: sites within an area (of the order of 1 km apart) appear to have

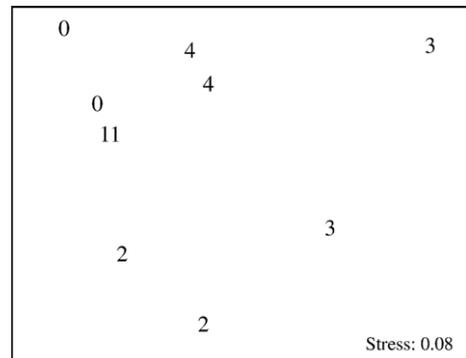


Fig. 5. Tees Bay macrofauna. ‘Second-stage’ MDS ordination. Labels 0–4 represent areas. Within each area data from two sites were analysed. Each symbol represents the pattern of community change at a site through time.

more similar time trajectories than sites in different areas (5 to 20 km apart), implying some degree of local control. To reiterate: this is not simply an expression of the usual spatial pattern that nearby samples have more similar communities than distant ones; it shows instead that the inter-annual dynamics of nearby communities are more similar than distant ones. This is formally testable, because of the structure of two replicate sites within each of the 5 areas. A one-way ANOSIM test (Clarke and Green, 1988) carried out on the second-stage similarity matrix, of the null hypothesis of no area differences in time profiles, gives a test statistic of $R=0.52$ ($p<0.1\%$). The global test has only modest power, there being 945 possible permutations of the area labels, but the observed separation of the groups (5 areas) is larger than for any of the other 944 permutations. Pairwise tests for individual pairs of areas have no power of course, there being only 3 distinct permutations of 2 groups of 2 labels, but the pairwise R statistics are all large, in keeping with the evident separation in the MDS plot. The global ANOSIM test in this context becomes a highly robust test of an interaction effect (time trajectories differ in different areas) which is fully non-parametric, depending only on the rank order of the original Bray–Curtis similarities among times.

A secondary feature of Fig. 5 is that the two areas within the influence of the Tees estuary (Areas 2 and 3) both appear to exhibit greater between-site, within-area variation in their time profiles (site pairs are further apart in the plot). This is not directly testable but is affirmed by examining the second-stage similarities themselves (or the MVDISP routine in PRIMER), i.e. it is not an artefact of the limitations of an ordination in reflecting the real high-dimensional structure of the second-stage similarities. This again has a natural interpretation, in terms of inter-annual changes in areas within the mixing region for estuarine contaminants being under greater local control. This would be expected to result in Areas 2 and 3 being separated from 0, 1 and 4 on Fig. 5, with greater differences between their replicate sites, as is observed.

Finally, effective though the second-stage procedure may be in these two cases, in removing ‘main effect’ differences of one factor in order to express interactions with another factor, it is important not to underestimate the drastic nature of the removal. *All* between-site similarities from the original Bray–Curtis matrix are ignored, so no information is retained about whether the baseline communities at two sites bear any relation to each other. It does not matter, to the second-stage calculation, whether the different sites or areas lie in entirely different parts of the first-stage MDS space, or on top of each other. They may even possess totally different sets of

species, though this would not usually be the case. Neither does it matter whether the inherent variability in community structure across the set of times is much larger for one site than another: absolute variability differences are also removed when matching internal time trajectories. That such variability differences occur here is seen in Fig. 6, which contrasts just two of the areas (0 and 3), showing time trajectories for the two sites for each area on a common first-stage MDS plot. Whilst the pattern of inter-annual change is the same for all four sites, with a general left to right drift and a more discrete shift in the community structure in about 1987 (and perhaps a tendency for greater inter-annual variability after that regime shift), it is also clear that the general year-to-year variability is higher in Area 3 than Area 0. This latter difference is removed in the rescaling inherent in the second-stage procedure, as is clear from the separately plotted time trajectories in Fig. 4. The greater spacing of the two Area 3 sites in the second-stage plot (Fig. 5), compared with that for Area 0 sites, is not therefore accounted for by the difference in scale seen in Fig. 6. It is a more subtle effect than that: a demonstration that the shapes of the time trajectories *differ* between the two sites in Area 3 more than they do in Area 0, possibly reflecting the differing reach of local contaminant or construction events (barrage building, etc.).

3.3. Rocky subtidal macroalgae, Livorno, Italy

The final example is also one of tracking changing community structure through time, so that the second-

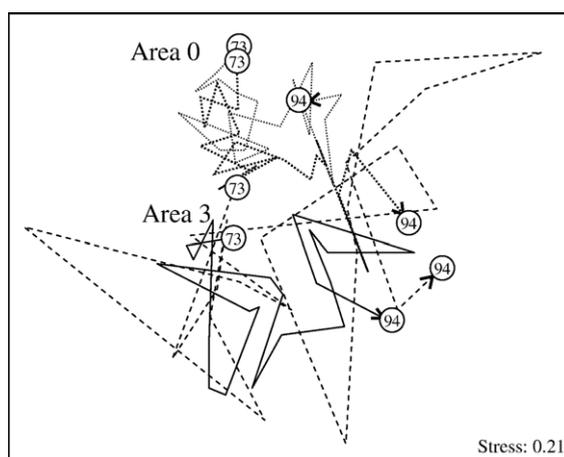


Fig. 6. Tees Bay macrofauna. MDS plot based on Bray–Curtis similarities calculated using fourth-root transformed abundances of macrofauna, at two sites within Area 0 and Area 3, in September of each year between 1973 and 1994. Each sample is the average of abundances of each species in 5 grabs, and consecutive samples from each site are linked by lines, representing the trajectory of community change at an individual site through time.

stage matrix entries are again correlations of time trajectories, for a series of spatial samples. The differences here, however, are that the study is a manipulative experiment in the environment, with a more structured design, and that it contains a non-ignorable element of ‘repeated measures’. At each time point after the initial clearance the same patch of rock surface is re-examined, to measure recolonisation and growth of algal assemblages. Communities at successive time points cannot, therefore, be treated as if independent of each other.

Randomly selected rock patches were cleared of algae in each of 8 starting months between October 1995 and September 1996 (8 ‘treatments’) and recovery monitored at 6 (approximately) bi-monthly periods subsequently. Data were percentage cover, fourth-root transformed and with sample similarities again calculated by Bray–Curtis. The sample structure used here initially is just the samples from the experiment at the Calafuria location, there being three independent, randomly chosen sites monitored through time for each of the 8 treatments. (Data from the three randomly chosen patches at each site, fixed through time, were simply pooled for this example, but there is clearly a lower level of nesting that could have been analysed by second-stage two-way nested ANOSIM).

The hypothesis of interest concerns evidence for a differing pattern of recolonisation and recovery, dependent on the time of year of clearance (treatment). The advantage of a second-stage approach this time is not so much the removal of a spatial difference in starting

community structure, because all plots start from a cleared state. Instead, it is the ability to remove ‘repeated measure’ complications by utilising the full recolonisation sequence as a single, independent replicate in a second-stage ANOSIM test. (Note that, in consultation with K. R. Clarke, Chapman (2003) used the same analysis approach to test succession on defaunated intertidal boulders.)

A first-stage MDS of the 8 treatments and 6 recovery periods, for data averaged over the 3 sites in each treatment (Fig. 7), certainly shows evidence of the recolonisation sequence, broadly left to bottom right and then up the plot. Much more than that is difficult to divine, however. It is not clear that the recovery trajectory evolves with the seasonal change in time of clearance, or even whether the different starting times affect the trajectory at all. Such subtleties are drowned out by the strong changes in algal communities through a recovery cycle.

Fig. 8, on the other hand, focuses in on the time trajectory of recolonisation over the year after clearance, separately for each start date, and is also able to reinstate the separate time profiles for each of the three sites, which make up the averaged points in Fig. 7. The picture now is very much clearer. Characteristically different recovery profiles emerge for each start date, which appear strongly repeatable over the 3 independent (replicate) sites selected for each treatment. For convenience, the profiles for the three sites of each treatment are displayed together on each MDS plot, but the second-stage procedure breaks up all 24 trajectories and

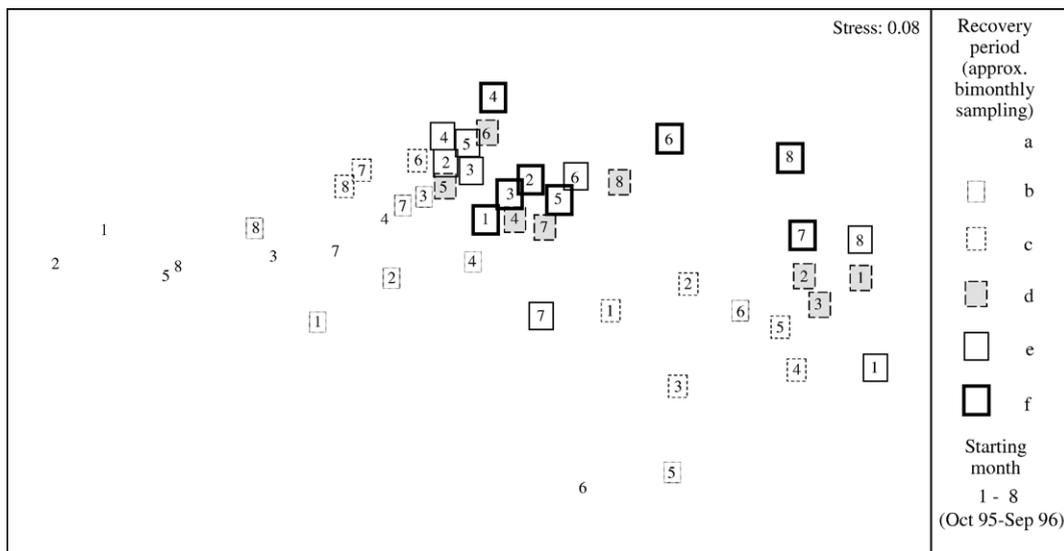


Fig. 7. Livorno algae. ‘First-stage’ MDS plot depicting changes in community structure of algae on sublittoral rocks following experimental clearance at the Calafuria station. Sets of patches were cleared on 8 separate occasions, and each patch non-destructively sampled six times over the following year. Each point represents the average of nine 15×15 cm patches. Bray–Curtis similarities were computed using fourth-root transformed percentage cover data.

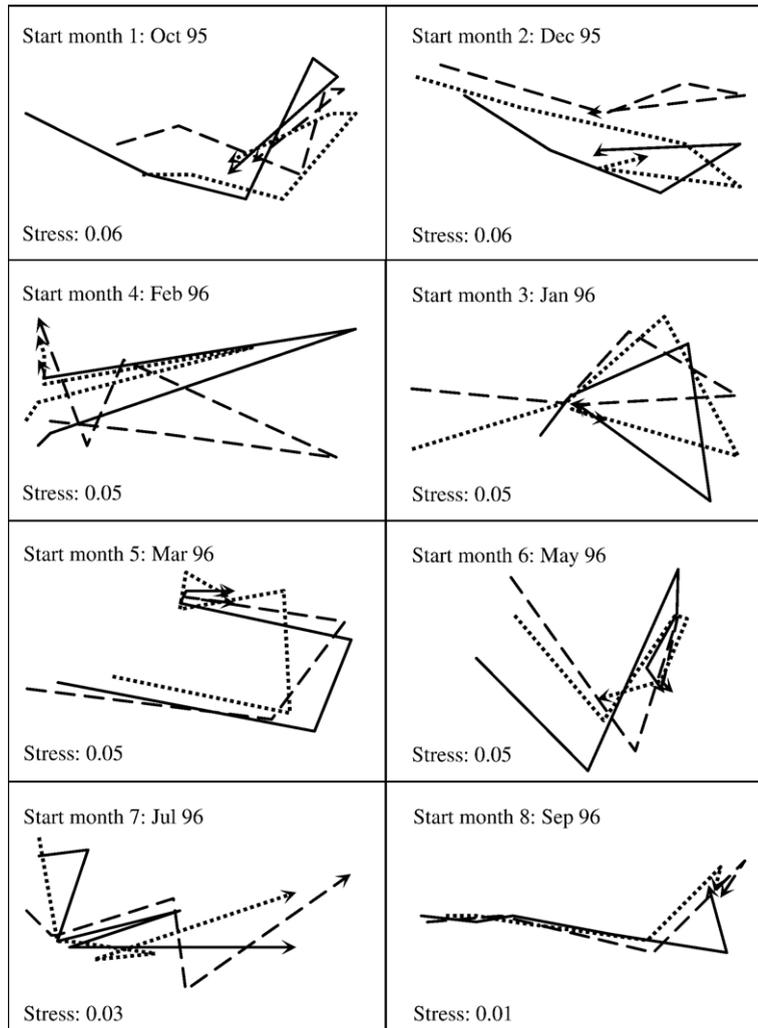


Fig. 8. Livorno algae. Time (=recovery) trajectories for individual sites in the clearance experiment at Calafuria. Each plot is a 'first-stage' MDS based on Bray–Curtis similarities calculated using fourth-root transformed percentage cover of algal species, for the patches cleared in a particular month. Arrowed lines link consecutive times for non-destructive sampling on six occasions over the following 12 months.

calculates the Spearman correlation between the underlying similarity matrices for all $24 \times 23/2 = 276$ pairs. Converting these to similarities (of similarities) and submitting them to ordination analysis generates the second-stage MDS (Fig. 9a). The close matching of profiles amongst sites, and different recovery trajectories across treatments, evident from Fig. 8, lead to the tight replication and clear separation of treatments in Fig. 9a.

In fact, treatment separation is even more clear-cut than it appears in Fig. 9a. The second-stage MDS has a rather poor stress level of 0.17 in two dimensions but a full-dimensional hypothesis test for treatment differences (in recovery profiles) is available, as in the previous example, by one-way ANOSIM on the second-stage similarity matrix. The global R is 0.96 (with the

value of p as small as one cares to make it), and all the pairwise R values are also 1 or very nearly 1, the exceptions being 2 vs. 3: $R=0$ and 2 vs. 6: $R=0.63$.

As with the corals example, the contrast with the first-stage MDS is instructive. The traditional way to avoid the repeated measures problem is to pick a single time point, say the final monitoring period a year after clearance, and to analyse the communities from all treatments and replicates at that time; the samples are then properly independent, if randomly selected to receive treatments in the first place. Analysis of data collected 12 months after the patches were cleared leads to a first-stage MDS (Fig. 9b), again pooling the data from patches in each site so that there are 3 sites (replicates) per treatment. The different treatments (clearance times) are seen to have

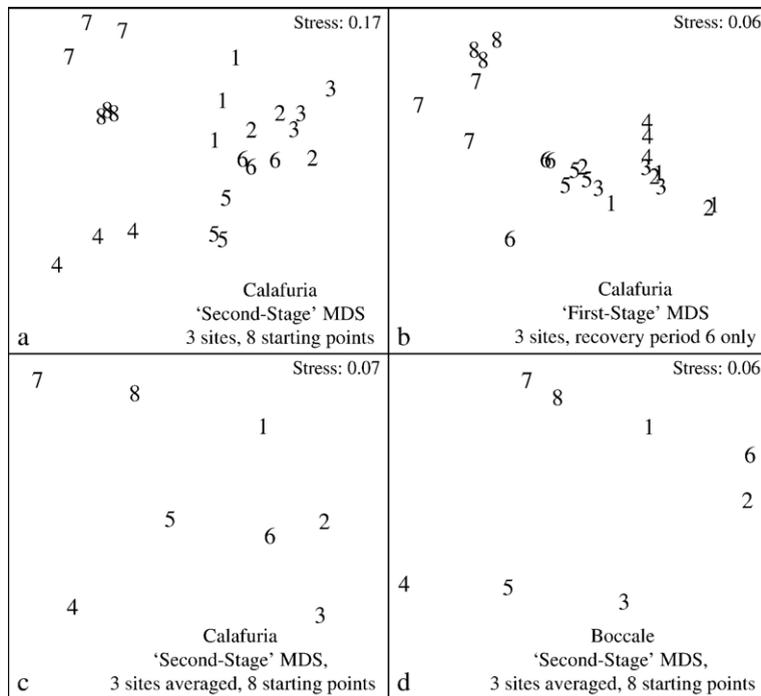


Fig. 9. Ligurian Sea algae. (a) 'Second-stage' MDS plot indicating concordance in multivariate pattern among time trajectories from individual sites at Calafuria, cleared in different months. The trajectories were calculated using the Bray–Curtis measure and fourth-root transformed percentage cover of algal species. Labels indicate the order in which patches were cleared (8 times over 1 year). (b) 'First-stage' MDS based on fourth-root transformed cover of macroalgae in each site, sampled 12 months after the patches were cleared. (c) 'Second-stage' MDS plot as in panel a but with time trajectories averaged over the sites in each 'treatment' (starting month). (d) Second-stage MDS as in panel c but for the repeat of the experiment at Boccale.

differing effects on the final community, though the separations are by no means as clear-cut. The ANOSIM global R is 0.63 (again strongly significant, naturally) but many of the pairwise R values are now much smaller. The point to note here is not just that use of the whole time profile in the second-stage procedure, rather than solely its end point, increases the sensitivity and resolution of the test, but also that the second-stage analysis is displaying subtly different information. Nearby points on the first-stage plot represent final communities that are highly similar, but which may not have arrived at that point by the same 'evolutionary' path; coincident points on the second-stage plot represent parallel evolution of assemblage structure through time, but in which the end points need not be the same. Both types of plot are potentially informative.

Finally, it is a reasonable question to ask if this manipulative experiment in the environment is repeatable (as Underwood has done over a long career, almost uniquely amongst his peers!). Airolidi (2000) did indeed repeat the experiment twice more: at Boccale and

Vasche, two further stations on the Ligurian coast. For two stations, Calafuria and Boccale, we pooled data over sites, as well as patches within sites, so giving 8 treatments by 6 recovery time points, each being an average of 9 patches. These are used to calculate a single recovery trajectory for each of the 8 treatments at each station and the second-stage MDS plots based on correlations between these are shown in Fig. 9c,d. The degree of agreement in the two plots is unequivocal.

This raises a further possibility, though probably of limited application. The degree of agreement between these two second-stage plots can be measured by the Spearman correlation of their underlying second-stage remembrance matrices (of Spearman correlations). It can also be tested for departure from zero (no repeatability of the experiment) by a Mantel-type permutation test (e.g. RELATE in PRIMER). Here $\rho_s = 0.56$, $p < 0.4\%$, so we can decisively reject the hypothesis of no repeatability. There is also a further repeat of the experiment here, at Vasche, and one can envisage still further repetitions. The Spearman correlations between

the second-stage similarity matrices would then constitute a third-stage similarity matrix, leading to a third-stage MDS and further third-stage tests! Although these would answer questions about whether some repeats of the experimental protocol at certain times or places tended to give rather different conclusions (i.e. a higher-way interaction between treatment effects and location of the experiment on recovery profiles), and though easy enough to postulate, a practically useful application of third-stage MDS is unlikely to materialise, since each successive stage involves a further step away from the data of species counts or area cover. It is already likely to be the most serious criticism of second-stage analyses, that in the search for subtle temporal or spatial patterns, they take the user further away from the real data. Just as ANOVA tables are never the final answer to any univariate biological study or experiment, so second-stage analyses should be seen as merely providing evidence for the existence of particular forms of assemblage structuring. The roles of individual species in establishing these multivariate patterns are fundamental and understanding those contributions should continue to be a primary goal of any statistical analysis.

4. Conclusions

- 1) There are difficulties in generalising simple univariate definitions of interaction, from classic linear models, to the robust, *non-parametric* multivariate methods that are commonly used in handling assemblage data.
- 2) Certain forms of interaction can be examined using non-parametric methods, namely those evidenced by changing assemblage patterns over many time periods, for replicate sites from different experimental conditions (types of ‘Beyond BACI’ design)—or changing multivariate structure over space, at several observed times.
- 3) Second-stage MDS plots based on matrix correlations (‘similarities of similarity matrices’) can be used to illustrate such interactions, and they can sometimes be formally tested for by second-stage ANOSIM and RELATE permutation tests.
- 4) The second-stage analyses use (first-stage) similarities which can be defined with *any* resemblance measure appropriate to the biological context, and can therefore preserve the fully non-parametric, high-dimensional generality of related analyses widely used in marine community studies.
- 5) The strategy copes straightforwardly with certain repeated-measures designs which are cumbersome to handle in conventional (univariate) analyses.

Acknowledgments

This work is a contribution to the biodiversity element of the Plymouth Marine Laboratory’s core strategic research programme. It was supported by the UK Natural Environment Research Council (NERC) and the UK Department for Environment, Food and Rural Affairs (DEFRA) through the AMBLE project ME3109. LA was supported by an Assegno di Ricerca of the University of Bologna and her participation instigated whilst working at the Marine Biological Association of the UK, under a LINKECOL Exchange Grant from the European Science Foundation. We are most grateful to Barbara Brown, (retired) professor in tropical marine ecology at the University of Newcastle-upon-Tyne, UK, for permission to use her Phuket coral time series, and to Ross Brown, Nigel Shillabeer and the many other staff of Astra-Zeneca involved in the Tees Bay monitoring programme. KRC and RMW acknowledge their position as honorary fellows of the Plymouth Marine Laboratory, and KRC also his honorary fellowship at the Marine Biological Association of the UK. All authors express their gratitude to Tony Underwood for his friendship, inspiration and support over many years. [SS]

References

- Airoldi, L., 2000. Responses of algae with different life histories to temporal and spatial variability of disturbance in subtidal reefs. *Mar. Ecol. Prog. Ser.* 195, 81–92.
- Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. *Aust. Ecol.* 26, 32–46.
- Bray, J.R., Curtis, J.T., 1957. An ordination of the upland forest communities of Southern Wisconsin. *Ecol. Monogr.* 27, 325–349.
- Brown, B.E., Le Tissier, M.D.A., Scoffin, T.P., Tudhope, A.W., 1990. Evaluation of the environmental impact of dredging on intertidal corals at Ko Phuket, Thailand, using ecological and physiological parameters. *Mar. Ecol. Prog. Ser.* 65, 273–281.
- Brown, B.E., Dunne, R.P., Chansang, H., 1996. Coral bleaching relative to elevated seawater temperature in the Andaman Sea (Indian Ocean) over the last 50 years. *Coral Reefs* 15, 151–152.
- Brown, B.E., Clarke, K.R., Warwick, R.M., 2002. Serial patterns of biodiversity change in corals across shallow reef flats in Ko Phuket, Thailand, due to the effects of local (sedimentation) and regional (climatic) perturbations. *Mar. Biol.* 141, 21–29.
- Chambers, D.P., Tapley, B.D., Stewart, R.H., 1999. Anomalous warming in the Indian ocean coincident with El Nino. *J. Geophys. Res.* 104, 3035–3047.
- Chapman, M.G., 2003. The use of sandstone blocks to test hypotheses about colonization of intertidal boulders. *J. Mar. Biol. Assoc. U. K.* 83, 415–423.
- Clarke, K.R., 1993. Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol.* 18, 117–143.
- Clarke, K.R., Ainsworth, M., 1993. A method of linking multivariate community structure to environmental variables. *Mar. Ecol. Prog. Ser.* 92, 205–219.

- Clarke, K.R., Gorley, R.N., 2006. PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth.
- Clarke, K.R., Green, R.H., 1988. Statistical design and analysis for a 'biological effects' study. *Mar. Ecol. Prog. Ser.* 46, 213–226.
- Clarke, K.R., Warwick, R.M., 2001. *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation*, 2nd ed. PRIMER-E, Plymouth.
- Clarke, K.R., Warwick, R.M., Brown, B.E., 1993. An index showing breakdown of seriation, related to disturbance, in a coral-reef assemblage. *Mar. Ecol. Prog. Ser.* 102, 153–160.
- Clarke, K.R., Somerfield, P.J., Chapman, M.G., 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.
- Gower, J.C., Legendre, P., 1986. Metric and Euclidean properties of dissimilarity coefficients. *J. Classif.* 3, 5–48.
- Green, R.H., 1979. *Sampling Design and Statistical Methods for Experimental Biologists*. Wiley, Chichester.
- Kruskal, J.B., Wish, M., 1978. *Multidimensional Scaling*. Sage Publications, Beverley Hills, California.
- McArdle, B.H., Anderson, M.J., 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology* 82, 290–297.
- Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27, 209–220.
- Olsford, F., Somerfield, P.J., Carr, M.R., 1997. Relationships between taxonomic resolution and data transformations in analyses of a macrobenthic community along an established pollution gradient. *Mar. Ecol. Prog. Ser.* 149, 173–181.
- Olsford, F., Somerfield, P.J., Carr, M.R., 1998. Relationships between taxonomic resolution, macrobenthic community patterns and disturbance. *Mar. Ecol. Prog. Ser.* 172, 25–36.
- Reid, P.C., Borges, M.D., Svendsen, E., 2001a. A regime shift in the North Sea circa 1988 linked to changes in the North Sea horse mackerel fishery. *Fish. Res.* 50, 163–171.
- Reid, P.C., Holliday, N.P., Smyth, T.J., 2001b. Pulses in the eastern margin current and warmer water off the north west European shelf linked to North Sea ecosystem changes. *Mar. Ecol. Prog. Ser.* 215, 283–287.
- Somerfield, P.J., Clarke, K.R., 1995. Taxonomic levels, in marine community studies, revisited. *Mar. Ecol. Prog. Ser.* 127, 113–119.
- Stewart-Oaten, A., Murdoch, W.M., Parker, K.R., 1986. Environmental impact assessment: "pseudo-replication" in time. *Ecology* 67, 929–940.
- Underwood, A.J., 1992. Beyond BACI: the detection of environmental impact on populations in the real, but variable, world. *J. Exp. Mar. Biol. Ecol.* 161, 145–178.
- Underwood, A.J., 1993. The mechanics of spatially replicated sampling programmes to detect environmental impacts in a variable world. *Aust. J. Ecol.* 18, 99–116.
- Underwood, A.J., 1994. On beyond BACI: sampling designs that might reliably detect environmental disturbances. *Ecol. Appl.* 4, 3–15.
- Underwood, A.J., 1997. *Experiments in Ecology: Their Logical Design and Interpretation Using Analysis of Variance*. Cambridge University Press, Cambridge.
- Warwick, R.M., Ashman, C.M., Brown, A.R., Clarke, K.R., Dowell, B., Hart, B., Lewis, R.E., Shillabeer, N., Somerfield, P.J., Tapp, J.F., 2002. Inter-annual changes in the biodiversity and community structure of the macrobenthos in Tees Bay and the Tees estuary, UK, associated with local and regional environmental events. *Mar. Ecol. Prog. Ser.* 234, 1–13.
- Webster, P.J., Moore, A.M., Loschnigg, J.P., Leben, R.R., 1999. Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997–98. *Nature* 401, 356–360.