Oecologia (1998) 117:570–578

A.J. Underwood · M.G. Chapman

A method for analysing spatial scales of variation in composition of assemblages

Received: 14 November 1997 / Accepted: 14 September 1998

Abstract In several areas of research on ecological assemblages, it is useful to be able to analyse patterns of spatial variation at various scales. Multivariate analyses of dissimilarity or similarity in assemblages of species are limited by problems of non-independence caused by repeated use of the sample-units. Where rank-order procedures are used, no comparative quantitative measurements of dissimilarity at different scales are produced. An alternative method is described that uses the sample's average assemblage (or centroid). These estimates are themselves averaged to give centroids for larger spatial scales. Dissimilarities from the centroids at each scale are then calculated using independent replicates for each scale from those in each sample. The dissimilarity measures can then be examined by analysis of variance to detect spatial scales of differences for each sample at every level of a hierarchy of scales. The method is illustrated using data from mangrove forests and rocky shores, involving up to 97 taxonomic groups (species, other taxa). Differences among assemblages at the scales of sites (tens of meters apart) or locations at shores (hundreds of meters apart) were identified. Consequences of different numbers of replicates are discussed, with some potential problems (and their solutions) in application.

Key words Assemblages · Hierarchy · Multivariate analyses · Spatial scale

Introduction

There are pressing reasons for analysing patterns in the composition or structure of assemblages. Numerous intrinsic ecological issues require detailed quantitative understanding of the scales at which there is consistent and predictable variation in the abundances of diverse assemblages. For example, identifying processes that regulate structure and dynamics of interactions between species requires recognition of the scales at which processes operate (Allen and Starr 1982; Bell et al. 1995; Bourget et al. 1994; Dayton and Tegner 1984; Menge and Olson 1990) and therefore quantitative description of spatial and temporal variation.

A second purely ecological issue is the recent upsurge of interest in indirect effects (e.g. Menge 1995; Menge et al. 1994; Wootton 1994). Understanding the web of interactions among species in patchy, heterogeneous habitats must include understanding of how the outcomes vary from place to place and time to time. In addition, there are many examples of considerable variation in abundances of the individual species at a range of spatial scales (e.g. Underwood and Chapman 1996 for intertidal species). Given that these species are involved in the numerous competitive, grazing or predatory interactions which dictate the structure of assemblages, corresponding variation in the structure of assemblages can be predicted.

In addition to such ecological notions, there are more practical considerations requiring detailed knowledge of spatial scales of variation in the structure of assemblages. These relate to continuity and processes creating biogeographical patterns and form the basis for identifying the scales of "managerial units" and thereby for conservation of species, management of marine reserves and recognition of groupings of stocks in multi-species fisheries.

There are, however, many problems in the measurement of spatial or temporal variation in assemblages. Often, the multivariate measures used are measures of dissimilarity (or other distance measures) which incorporate the differences in abundances of all taxa between two replicate sample units (cores, quadrats, whatever) in an assemblage. Some sort of ordination then allows representation of the patterns of dissimilarity or similarity among all the sample units. The

A.J. Underwood (✉) · M.G. Chapman
Centre for Research on Ecological Impacts of Coastal Cities, Marine Ecology Laboratories, A11, University of Sydney, NSW 2006, Australia
e-mail: aju@bio.usyd.edu.au, Fax: +61-2-9351 6713

general strategy is well-known and has been reviewed many times (Clarke 1993; Field et al. 1982; Green 1980; Pielou 1984).

There are two problems with these methods that restrict their usefulness for quantitative analysis of spatial or temporal patterns in multi-species assemblages. First, the sorts of sampling and experimental designs needed to examine, describe or test hypotheses about assemblages are usually complex (e.g. Underwood 1992, 1994 for examples in environmental studies). The multivariate procedures cannot deal with such complexity (Clarke 1993; Green 1979).

Second, the data collected in a sample of some assemblage do not generate independent measurements of dissimilarities, so that replicate units in the sample cannot be used as independent replicates in any analysis (see Clarke 1993). This feature of many types of multivariate measures explains why the designs must be simple.

Where hypotheses and data only concern the composition of assemblages, i.e. the presence/absence of species or other taxonomic groupings, Smith (1989) has already produced a method of analysis. His method produces an analogue of analysis of variance using statistics based on the probabilities of finding each taxon in any pair of sample units. There are, however, numerous studies that need to analyse the patterns of abundance and not just presence/absence of taxa in assemblages.

The present paper describes a procedure based on Bray-Curtis (Bray and Curtis 1957) measures of dissimilarity of abundances of taxa between pairs of replicate units, but the procedure generates relatively independent measures at every spatial (and temporal, although that is not described here) scale and in each location sampled. The method generates replicated univariate data (Bray-Curtis measures) that can be analysed by any usual procedure. This is illustrated here in analyses of assemblages of marine invertebrates and algae on rocky shores and of macro-invertebrates in sediments in mangrove habitats. Provided that the numbers of replicate sample units in the calculations are reasonably large (''reasonably'' being defined in terms of empirical field ecologists), the Bray-Curtis measures used for each component of an experimental design are realistically independent enough to satisfy the requirements of analyses.

The analyses illustrated are examples of tests of hypotheses about spatial scales of variation in the structure of assemblages. The biology of the organisms and their ecology in these coastal habitats are not discussed here; the data are to demonstrate the method and potential uses in ecological studies.

Methods

Data and analysis

The solution to the problem is to take sufficient data, in any site or at any given scale, to acquire replicated multivariate measures of variation in composition of assemblages. Consider the specific example of assemblages of infaunal invertebrates in sediments in mangrove forests. Suppose that there are hypotheses to be tested at various spatial scales (as there often are: Underwood and Chapman 1996; Underwood and Petraitis 1993) or other reasons to need to compare variation in assemblages at the scales of locations (separated by hundreds of meters), sites within locations (separated by tens of meters) and within each site (at scales of meters). For what follows, the actual scales are not important; the proposed analysis can be used for any hierarchical arrangement.

Data on abundances of species (or other, broader taxonomic groupings if that is appropriate for the assemblage being analysed: Clarke 1993; Warwick 1988a,b), are collected in quadrats scattered around each site.

To compare samples from two different sites or locations, a measure of dissimilarity can be calculated to summarize all the differences between abundances of taxa between all pairs of sample units. Bray-Curtis dissimilarity was chosen because it has advantages over other alternatives (Clarke 1993; Faith et al. 1987). Percentage dissimilarity between sample units j and k is:

$$d_{j,k} = 100 \times \frac{\sum_{i=1}^{s} \left| X_{ij} - X_{ik} \right|}{\left(\sum_{i=1}^{s} X_{ij} \right) + \left(\sum_{i=1}^{s} X_{ik} \right)}$$

where X_{ij}, X_{ik} are the abundances of taxon i in sample units j and k, respectively and s is the total number of different taxa found over the two units. When units j and k are from the same sample, $d_{j,k}$ measures variation in the assemblage within the sample; when j and k are from different samples, dissimilarity is between samples. If there are several samples, the units from each sample have to be used numerous times to calculate every possible measure between every pair of samples. The data are not independent for comparisons among samples.

The solution proposed here is to take sufficient replicate units in each sample to be able to use each one only once in any comparisons. To have sufficient replication to make separate comparisons between every possible pair of samples is prohibitive (for the relatively simple designs illustrated here for rocky shores, there would have to be an exorbitant number of quadrats in every sample).

Instead, several sample units (m units) are used from each sample to estimate the average abundances of all taxa in that sample. These average abundances per unit are then used as the average assemblage or ''centroid'' for that sample. Dissimilarities of other units from the centroid measure variation among units in that sample. So, the Bray-Curtis measures are calculated as before, using the abundance of each taxon (X_{ij}) in unit j and the average abundance $\bar{X}_{ij}$ of that taxon calculated from m different units.

The abundance of taxon i in unit l in any sample is $X_{il[k(j)]}$ indicating that unit l is nested in the sample from site k nested in location j. Consider the case of samples from two locations, with two sites in each location and 12 quadrats sampled per site: $m = 12$ quadrats could be used to calculate the centroid for each site $\bar{X}_{ik(j)}$ for taxon i. The centroids from the two sites in each location are averaged to give the centroid for that location $\bar{X}_{ij}$ for taxon i in location j, using $m = 6$ units randomly taken from each of the two sites. Finally, the centroids for each location can be averaged to give the overall centroid $\bar{X}_i$ for taxon i averaged over all quadrats, sites and locations, using $m = 3$ quadrats from every site.

The 12 quadrats in each site can now be used to provide four replicate quadrats to measure Bray-Curtis dissimilarities from the site centroid (between $X_{il[k(j)]}$ and $\bar{X}_{ik(j)}$; $l = 1,...4$ quadrats, $i = 1,...s$ species), four different replicate quadrats to measure dissimilarity between quadrats and the centroid for their location (between $X_{il[k(j)]}$ and $\bar{X}_{ij}$ values) and the remaining four quadrats to measure dissimilarity between quadrats and the overall centroid (between $X_{il[k(j)]}$ and $\bar{X}_i$ values).

The first of these measures only variation in assemblages from quadrat to quadrat in any site. The second includes any variation from site to site. The third set measures variation in the assemblage from quadrat to quadrat in each site, from site to site in each location and from location to location. In general, if there is no variation in assemblages from site to site, the average dissimilarity

of quadrats from their site centroid should be similar to that from the centroid of their location. If there is also no variation among locations in the abundances of organisms in the assemblage, the average dissimilarity of quadrats from their site centroids, from the centroid of their location and from the overall centroid should be similar. In contrast, if there is no variability from site to site, but assemblages differ from location to location, the dissimilarities between quadrats and their site centroid should be equal, on average, to those between quadrats and the centroid of their location. Both sets, will, however, be smaller than the average Bray-Curtis dissimilarities between quadrats and the overall centroid (because these include also the differences among locations).

These dissimilarity measures can therefore be used in an analysis of variance to test the null hypotheses of no differences among sites and no differences among locations. This forms a fixed factor of scale with three levels; dissimilarities among quadrats (i.e. within sites), among quadrats plus sites and among quadrats plus sites plus locations. This factor is orthogonal to the hierarchical spatial structure of locations, sites nested within locations and quadrats nested within sites, allowing a three factor mixed model of scale (fixed) orthogonal to locations (fixed or random depending on the hypothesis), sites within locations (random) and a residual variation among quadrats in each site.

Before considering actual examples from real sets of data, the nature of the analyses are considered, particularly with reference to the nature of independence of the data in such analyses. This further requires consideration of the number of replicate units used to calculate centroids.

Then, analyses are presented for various sets of data from rocky shores and mangrove forests to illustrate the procedure, its interpretation and some practical issues in its use. One set of data was collected from midshore assemblages of epi- and infauna in a mangrove forest and two sets of data were collected from midshore assemblages on rocky intertidal shores. The fauna in the mangrove forest consisted primarily of gastropods, amphipods, isopods, copepods, polychaetes, nematodes, oligochaetes, insect larvae and numerous minor taxa. Similar assemblages were described by Chapman (1998). The assemblages on the rocky shores consisted primarily of encrusting and foliose algae, barnacles, tube worms, limpets and other grazing snails and predatory whelks. Similar assemblages were described in detail in Underwood (1981).

Data set 1

The epifauna and infauna were sampled in each of two sites, approximately 16 m^2 and 40 m apart, in each of two locations (100 m apart) in a mangrove forest in Port Jackson (Sydney). Sixteen 0.1-m^2 quadrats were sampled in each site. In each quadrat, the leaf litter and sediment were collected down to approximately 20 mm depth. This material was sieved through a 1-mm and 500-μm mesh sieve. The two components were preserved in 7% formalin solution and sorted under magnification. Generally, the coarse component was completely examined, but, when there was a large amount, it was subsampled (approximately 30% by wet weight). This gave reliable estimates of numbers of each taxon in the entire sample (Chapman 1998). Fauna were sorted from the fine material using several randomly-chosen subsamples until approximately 30% by volume of the sample was sorted. All invertebrates were counted, identified to different levels of taxonomic resolution [species, morphospecies, families and so on, as described in Chapman (1998) to give 67 taxa]. The numbers of each taxon were adjusted relative to the amount of material sorted to give estimates per quadrat.

Data set 2

Ten quadrats, 50 cm × 50 cm, were sampled in each of two sites, approximately 5 m long and 22 m apart, on each of three shores (hundreds of meters apart) in the Cape Banks Scientific Marine Research Area (Botany Bay). The sites ran alongshore in midshore assemblages. The percentage cover of all algae and sessile animals were estimated from 100 points per quadrat. Species that did not occur under any point were given a nominal percentage cover of 0.5%, i.e they were present but occupied $<1\%$ of the sampled points. All large mobile animals (e.g. gastropods, whelks, starfish) were counted. Small numerous animals (e.g. the gastropods *Littorina unifasciata, Patelloida latistrigata*) were counted in five small subquadrats per quadrat (4% of the area). These numbers were summed, but not multiplied up to the area of the entire quadrat so that the abundances of these extremely numerous animals did not dominate patterns of the entire assemblages (Underwood and Chapman 1998). Overall, there were 98 taxa in this set of data.

Data set 3

The third set of data were from rocky shores at Cape Banks. Percentage covers and abundances of a subset of the above assemblage (26 species) were collected as part of a much larger study. Four shores (separated by about 150 m) were chosen. On each, four sites (separated by about 30 m) were sampled with 16 quadrats. For this set of data, sites were chosen, based on appearance and previous sampling, so that assemblages were likely to differ.

Results

The method: independence, precision and potential for bias

To understand the potential importance in this method of non-independence in the data, it is helpful to draw the analogy with multivariate analysis of variance. Here, the same sample units are used to calculate centroids and then these are averaged at the larger spatial scales (i.e. the higher levels in the hierarchy). These are, of course, the same steps as used in equivalent univariate procedures. Thus, in a nested analysis of variance, sample units are averaged in each site and then averages of all sites in a location are averaged to produce the mean for the location. In the univariate analysis, independence of the sums of squares calculated from these various "centroids" is induced by the usual orthogonal decomposition of quadratic forms (Cochran's theorem, Scheffé 1959). These results cannot carry over in general to the procedure suggested here, but the analogy is sufficiently close to serve as a useful guide. Consider calculation of a centroid from a small number of m replicate units and then picking a random sample of n units to determine dissimilarities from the centroid. These dissimilarities will be correlated. This is straightforward to demonstrate in the univariate case of a set of data X_1, X_2,... X_i,... X_N, of which m are used to generate a mean ($\bar{X}$). A set of n variables (X_1,... X_n) is then chosen to calculate squared distances from the mean ($Y_i = (X_i - \bar{X})^2$) noting that $\bar{X}$ is the mean of the m values (not the set of n). The Y_i values will be correlated because they will depend on how different the mean $\bar{X}^*$ of the X_1,... X_i,... X_n is from $\bar{X}$. The magnitude of correlation in the univariate case, assuming the data are distributed normally, can be shown (K.R. Clarke, personal communication) to be $[1/(m+1)^2]$. In the univariate case, this is negligible pro-

vided m is large. For example, if $m = 4$ units are used to calculate the mean ($\bar{X}$), the correlation among values of squared deviation from the mean (Y_i values) is 0.04 and unlikely to cause serious difficulties for the reliability of the analyses.

What happens with multivariate data, particularly using Bray-Curtis measures (which are non-Euclidean) rather than squared distances is not clear, but should have, at most, similar degrees of correlation when samples are large. In fact, simulations of the sort described below (but not presented here) suggest that the potential correlations of the sort discussed here are, in fact, very small with realistic sizes of samples.

This consideration leads to using the entire sample ($m = N$) to calculate centroids, in order to generate the smallest possible correlation among replicates used to calculate dissimilarities. This would, however, lead to a new potential problem. If centroids for each site are calculated from m sample units and there are, say, two sites in each location, the centroid for that location will be averaged from $2m$ units.

Because the size of sample used to calculate centroids is larger at the larger scales, the correlation among replicates used to calculate dissimilarities will be different for the different scales. Thus, in the univariate case considered earlier, if n units are used to calculate (squared) deviations from the mean of m replicates in a site, their correlation will be of the order of $[1/(m+1)^2]$. If a different n units are used to calculate the deviations of replicates in one site from the mean of the location, their potential correlation will be smaller – $[1/(2m+1)^2]$ – because $2m$ units are used to calculate the mean of the location and centroids calculated from different numbers of replicates have different precision. The estimates of centroids for larger scales are inevitably more precise (being based on larger samples). Dissimilarities from centroids will therefore, all other things being equal, be, on average, larger where centroids are less precisely known. This is illustrated for two sets of simulated data in Fig. 1.

In the first set, 15 "species" were simulated by drawing, for each replicate, 15 random variates from a normal distribution with mean 50 and standard deviation 20. Then, for sizes of sample of 2–20 replicates, the simulated data were used to calculate a centroid, representing increasing precision as m was taken from 2 to 20. Finally, for each simulated centroid, a further sample of 100 replicates was generated to calculate Bray-Curtis dissimilarities (i.e. $n = 100$). Results of three "runs" of this procedure are illustrated in Fig. 1a. Notice that the average dissimilarity declines from 18.6% when $m = 2$ to 16.4% when $m = 20$. A similar decline was evident in each run of the simulation. This was as predicted above.

The second simulation illustrated was identical, except that data were generated from highly skewed distributions by taking the exponential of each variate drawn from the normal distribution used before. Again, there is a decrease with increasing precision used for estimating the centroid. Note also, in each case

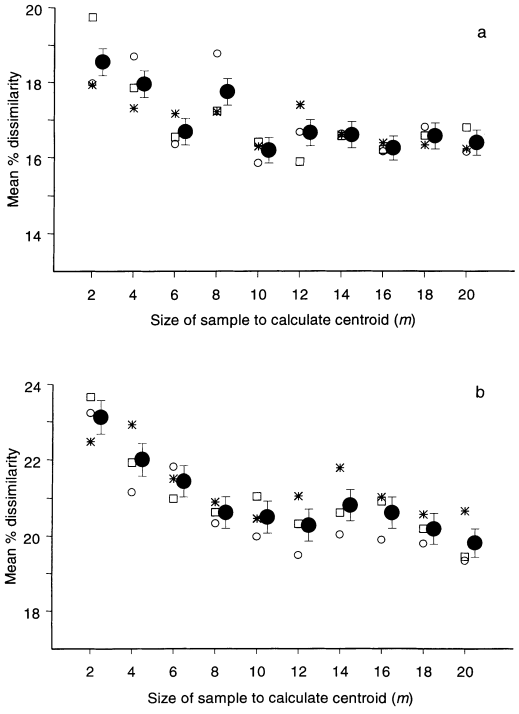

Fig. 1 Imprecision in estimation of centroids can lead to overestimation of Bray-Curtis measures of percentage dissimilarity. Data are from simulations. *Square*, *asterisk*, *open circle* mean % dissimilarity in three runs of the simulation, *filled circle* mean (± average SE) of the three runs, $n = 100$ in each simulation. The size of sample to calculate centroids (m) ranged from 2 to 20. **a** normally distributed data; **b** log-normally distributed data

illustrated, there is little further decline once sizes of samples (m) to estimate centroids reach about 6 or 8.

Nevertheless, if m is fixed, there is a potential bias towards overestimation of the dissimilarity measures for the smallest relative to the largest spatial scale (i.e. quadrats relative to locations or shores) because the latter are estimated with larger samples. In the Methods section, different numbers of quadrats were used to calculate centroids at each level so that all centroids were calculated using the same number of quadrats, or, where this was not possible, the most similar number of quadrats was used for every level.

Data from a mangrove forest

Data from two randomly chosen sites in each of two randomly chosen locations were used to test the hypothesis that there would be greater dissimilarity between locations than between sites. It was further hypothesised that locations and sites would both show greater variations in dissimilarity than occur within sites in each location. All 12 replicate quadrats (m) were used to calculate the centroids in each site. Six were chosen at random in each site to calculate the centroids in each location. Three replicate quadrats were chosen at random in each site to calculate the overall centroid. Thus, $m = 12$ for each site, location and for the overall

centroid. Three independent quadrats (n) were then used to calculate dissimilarities from each of the centroids for each site, location and the overall centroid and analysed in Table 1.

There was a highly significant difference among the scales, with dissimilarities among locations being greater ($P < 0.05$, SNK test) than that among sites and among quadrats (which were on average of similar magnitude, Table 1). There was about 27% dissimilarity among quadrats within each site and about 36% (but not significantly greater) variation in assemblages among sites. In contrast, dissimilarity averaged 52% among locations, an extra 25% compared with variation among quadrats. The procedure gives estimates of variability in assemblages at each scale and identified significantly greater variation among locations. So, these invertebrate assemblages were quite variable in structure at a scale of a few meters (quadrat to quadrat) and at a scale of hundreds of meters (location to location).

Note also that there was significant variation from site to site (Table 1), but no interactions between the factors scale and location nor between scale and sites. Such differences among sites indicate that differences in variability in the assemblages from quadrat to quadrat are greater in some sites than others (i.e. variances are different from site to site). This means that the sites in any location differed in their average dissimilarity from the various centroids, but the differences from the scale of quadrats to that among sites and the differences from the scale of sites to that among locations were similar in all sites (or there would have been significant interactions with scale). The average dissimilarities (averaged over the three scales) are shown for each site in Table 1. These demonstrated differences between sites of about 6%, on average (i.e. 42–43%; 28%–39% in the two locations). Thus, in each location, the two sites differed in how variable from quadrat to quadrat the assemblages were.

Data from rocky shores: three shores, two sites

In the first set of data from rocky shores, three shores, hundreds of meters apart, were sampled. On each shore, two sites were randomly chosen tens of meters apart. Of the ten quadrats from each site, $m = 10$ (i.e. all of them) were used to calculate the centroid for that site. Then six quadrats were randomly chosen from each site to calculate the centroid of that shore (thus, $m = 12$ for each shore). Two quadrats were randomly chosen from each site to calculate the overall centroid ($m = 12$, two from each of two sites on each of three shores). With only ten quadrats from two sites and three shores, it was not possible for m to be identical for all scales. The numbers of units used minimized the difference from scale to scale: $n = 3$ different quadrats were used to calculate dissimilarities from the centroids at each scale, thus keeping the data balanced for analyses. As with the data from mangrove forests, there were no significant interactions between scale and sites or shores. Dissimilarities among quadrats in each site averaged 33% and there was no greater variation among sites (38%). Average dissimilarity at the scale of shores was, however, significantly larger (SNK test at $P < 0.05$) at 48% (Table 2).

In this analysis, there were differences among shores, indicating the greater variance among quadrats on shore 2 (48%) than on shores 1 and 3 (see Table 2). As before, the differences from one scale to another were similar for every site and shore (hence, no interactions in Table 2).

Four shores, four sites

All 16 quadrats were used to calculate the centroids for each site ($m = 16$); four were then chosen at random to

Table 1 Analysis of Bray-Curtis percentage dissimilarities between two randomly chosen sites (tens of meters apart) and in each of two randomly chosen locations (100 m apart) in a mangrove forest; $n = 4$ quadrats for calculating dissimilarities for each scale in each site; $m = 12$ quadrats were used to estimate the centroids in each site, location and overall. Data untransformed (Cochran's $C = 0.29$, $P > 0.05$)

a Analysis of variance

Source of variation	*df*	MS	*F*-ratio	*P*
Among scales = Sc	2	2558.55	28.53	< 0.05
Between locations = L	1	959.75	2.93	> 0.20
Between sites (locations) = S(L)	2	327.82	3.88	< 0.03
Sc × L	2	89.69	1.46	> 0.30
Sc × S(L)	4	61.54	0.73	> 0.50
Residual	36	84.51	–	
Total	47	–		

b Mean values for 3 scales

Scale	Quadrats	Quadrats + Sites	Quadrats + Sites + Locations
Mean (SE); $n = 16$	27.0 (1.8)	35.6 (3.0)	51.9 (3.1)

c Mean values for locations and sites

Location	1		2	
Site	1	2	1	2
Mean (SE); $n = 12$	42.2 (3.7)	43.0 (5.0)	38.9 (3.9)	28.4 (3.3)

Table 2 Analysis of Bray-Curtis percentage dissimilarities between two randomly chosen sites (22 m apart) on each of three randomly chosen rocky shores (hundreds of meters apart); $n = 3$ quadrats for calculating dissimilarities for each scale in each site; $m = 10, 12, 12$ quadrats were used to estimate the centroids in each site, shore and overall. Data untransformed (Cochran's $C = 0.20$, $P > 0.05$)

a Analysis of variance

Source of variation	*df*	MS	*F*-ratio	*P*
Among scales = Sc	2	1032.21	22.20	< 0.01
Among shores = Sh	2	1150.83	25.11	< 0.01
Between sites (shores) = S(Sh)	3	45.83	0.45	> 0.70
Sc × Sh	4	46.49	0.29	> 0.85
Sc × S(Sh)	6	159.58	1.58	> 0.15
Residual	36	100.93	–	
Total	53	–		

b Mean values for 3 scales

Scale	Quadrats	Quadrats + Sites	Quadrats + Sites + shores
Mean (SE); $n = 18$	33.1 (3.2)	37.7 (2.9)	47.9 (2.2)

c Mean values for shores

Shore	1	2	3
Mean (SE); $n = 18$	31.6 (3.0)	47.6 (2.6)	39.4 (2.6)

Table 3 Analysis of Bray-Curtis percentage dissimilarities between four randomly chosen sites (30 m apart) on each of four randomly chosen rocky shores (150 m apart); $n = 5$ quadrats for calculating dissimilarities for each scale in each site; $m = 16, 16, 32$ quadrats for estimating the centroids in each site, shore and overall. Data untransformed (Cochran's $C = 0.12$, $P > 0.05$).

a Analysis of variance

Source of variation	*df*	MS	*F*-ratio	*P*
Among scales = Sc	2	19848.0	12.85	< 0.01
Among shores = Sh	3	1610.4	3.50	> 0.05
Between sites (Shores) = S(Sh)	12	459.6	3.50	< 0.0001
Sc × Sh	6	1544.7	4.16	< 0.005
Sc × S(Sh)	24	371.2	2.82	< 0.0001
Residual	192	131.6	–	
Total	239	–		

b Mean values for 3 scales

Scale	Quadrats	Quadrats + Sites	Quadrats + Sites + Locations
Mean (SE); $n = 80$	30.74 (1.55)	50.09 (1.95)	61.95 (1.52)

c Analysis of rank orders of scales; the three scales (shores, sites, quadrats) were each sampled 16 times (4 shores×4 sites). For each set, the mean dissimilarities for the 3 scales were put in decreasing rank order (Anderson 1959) for analysis. $\chi^2 = 66.75$, 4 *df*, $P < 0.005$

	Scale		
No. of times ranked	Shores	Sites	Quadrats
1 (greatest)	13	3	0
2	3	13	0
3	0	0	16

calculate the centroid for each shore ($m = 16$; four from each of the four sites). Finally, two quadrats were chosen at random from every site to calculate the overall centroid ($m = 32$). It was not possible to have $m = 16$ if there was to be replication of quadrats used from each site. However, m was sufficiently large for any differences from $m = 16$ to $m = 32$ to be trivial. A total of $n = 5$ different quadrats was used to calculate dissimilarities from the site, shore and overall centroids.

In this case, there was significant interaction between the spatial scales and sites within each shore (Table 3), indicating that differences from one scale to another varied according to site and, possibly, shore. There is therefore no point in attempting to examine overall differences between scales – they vary from place to place and can only be interpreted from place to place (Fig. 2). In general, however, there was greater dissimilarity among sites than among quadrats and among shores than among sites. Over all sites and shores, mean dissimilarity was 31% (SE = 1.6, $n = 80$), 50% (SE = 2.0), 62% (SE = 1.5) for quadrats, sites and shores, respectively. Dissimilarity was generally greater among shores than among sites than among quadrats [Fig. 2; Anderson (1959) test, $\chi^2 = 66.75$, 4 *df*; $P < 0.001$, Table 3]. Some sites were obviously more variable than others (dissimilarity among quadrats varied from 15% to 47%; Fig. 2) and some sites were more similar to each other than were others, but, overall, shores differed and sites differed.

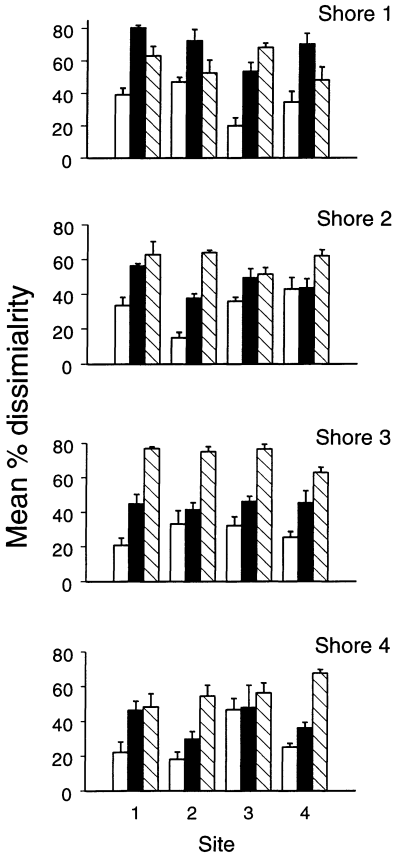

Fig. 2 Mean (SE; $n = 5$ independent replicate quadrats for each spatial scale) Bray-Curtis percentage dissimilarities for samples from four randomly chosen rocky shores, each sampled in four randomly chosen sites. Centroids for each site were calculated from $m = 4$ quadrats. For each site, data are: *open bars* variation among quadrats, *filled bars* variation among quadrats and sites, *shaded bars* variation among quadrats and sites and shores

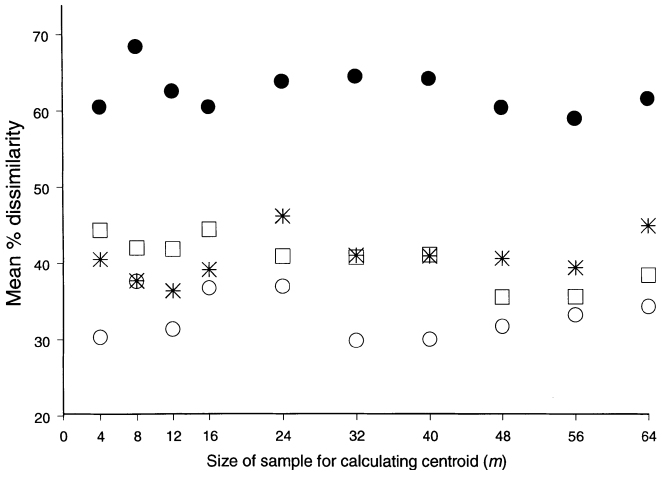

Fig. 3 Mean values of dissimilarities for data from a rocky shore (data set 3). For each of four sites on shore 2, $n = 5$ randomly chosen replicate quadrats were used to calculate dissimilarities from the centroid for the whole shore; the centroid was calculated, in turn, using 1 ($m = 4$), 2 ($m = 8$),...16 ($m = 64$) random quadrats from each site. There is no trend with m. *Open squares*, *asterisks*, *open circles*, *closed circles* are the means of the four sites on the shore

Practical issues about the size of sample

It was earlier pointed out that computer simulations and theory demonstrated greater precision in estimation of centroids would result in bias. To examine the effect of this in practice, the data for one shore (at random, shore 2) were used. In turn, the average centroid for this shore was calculated using one randomly chosen quadrat from each site ($m = 4$), then 2 randomly chosen quadrats ($m = 8$), 3,... all 16 quadrats ($m = 64$). For each centroid, $n = 5$ randomly chosen quadrats were used to calculate dissimilarities from the centroids. These are shown in Fig. 3. Despite the potential problems of bias (see earlier), there was no actual trend in dissimilarities with the number of units (m) used to calculate centroids. This result may strengthen the view that the amount of correlation among dissimilarities is small.

Both issues (potential differences in correlations for estimates at different scales and possible biases due to different precision at different scales) can be solved by using the same total number of replicate units to estimate the centroids at each scale. Thus, more replicates from each sample are used to estimate the centroids in each site than to estimate the centroids for each location or shore than for the overall centroid. This is the procedure used here, as identified in each of the analyses of real data. To keep m (the number of units used to calculate centroids) as large as possible, the entire sample is used to calculate centroids at the smallest scale (i.e. that of each site).

Discussion

The method described has two features that make it potentially useful in various ecological contexts. First, it requires considerably less sampling effort (smaller sizes of samples) than would be the case for equivalent analyses comparing dissimilarities within and among samples using ranked procedures (Clarke 1993). Second, data can be used in any design required by the study, provided that appropriate centroids can be calculated to match the model. In the examples illustrated here, the centroids were for three nested (or hierarchical) spatial scales. The framework we used was that of analysis of variance because of its flexibility and general suitability for ecological experimentation (Underwood 1997).

There is, however, the consideration of potential bias because of increased precision in estimation of centroids with the increasing sizes of samples used for larger scales (or that would be used for larger aggregations in other models for analysis). The degree to which such biases matter depends entirely on the problem being investigated. If the issue is precise estimation of the variation in multivariate data, large samples must be used to estimate centroids.

Situations also exist where random sub-sets of a relatively small sample will not provide a useful or reliable estimate of the average dissimilarity from the centroid for that whole sample. Consider a patchy rocky intertidal habitat where some parts are dominated by algae with their associated fauna and other patches are dominated by barnacles, with a different assemblage of flora and fauna. If a sample of the area has, say, six quadrats in each of the two assemblages, dissimilarity from the centroid of the whole sample will, correctly, be large. If, however, $n = 3$ of the quadrats are chosen for estimating the dissimilarity and, by chance, all happen to be from one of the assemblages (e.g. all happen to be quadrats from barnacle-dominated patches), the dissimilarities will very poorly represent (and will overestimate) that for the whole sample.

Such problems can always exist in any ecological study of heterogeneous habitats where there is insufficient evidence to enable prior stratification, or only small samples are possible. It might therefore be sensible to examine the average dissimilarity of all sample units from the centroid calculated using all the units to provide a check on the validity of sub-sampling.

A further point that must be remembered is that the data are about differences (or similarities if their complements are used) among sample units – not about actual diversity or composition of samples themselves. Thus, relevant hypotheses and their interpretations are in terms of how alike, from sample to sample, are the average dissimilarities among replicates. This is appropriate for many ecological research problems. It is, however, more akin to univariate analyses of variances among replicates, rather than analyses of mean values. This is not a problem in that there are many examples of needs for such methods in cases of univariate data (Bell et al. 1993; Brown et al. 1995; Cliff and Ord 1973; Legendre and Fortin 1989; Schneider 1994; Underwood and Chapman 1996). For the sorts of research programmes that need analyses of spatial or temporal variation, the method used here seems entirely appropriate.

Two final issues need further attention. First, there is often interest in, or a need for, analyses of relationships between structure of multispecies assemblages and physical or chemical variables that may influence structure (e.g. Clarke 1993; Clarke and Ainsworth 1993; Green 1979, 1984; Pielou 1984). The measures described here can be used in a variety of factorial and/or hierarchical designs. It remains to be seen how these can be exploited for exploring correlations and for analyses of covariance with relevant physical and chemical variables.

Finally, all of the currently topical issues about taxonomic resolution and appropriate transformation of data in multivariate analyses are relevant to the methods described here. It is common to calculate measures of dissimilarity based on data that have already been subjected to transformations (e.g. Clarke 1993; Clarke and Green 1988; Pielou 1984). Such transformations serve a number of functions, but are largely used to make measures less influenced by numerically dominant taxa (so transformation to log or double square root reduces influences of very abundant species) or to make rare species equivalent to abundant ones (by transforming all data to presence/absence). Such transformation is done to the data before calculation of centroids and measures of dissimilarity. Note that this is not the same as issues about transformation of data where this is considered appropriate to conform to the assumptions of analysis of variance of the dissimilarity values. That transformation would be on the final Bray-Curtis measures of dissimilarity, if transformation would solve some problem of heterogeneity of variances. Whether that was considered appropriate or not depends on the variances of the replicated measures of dissimilarity and not on whether dissimilarity was calculated for untransformed or transformed or presence-absence data. It is worth noting, in passing, that Bray-Curtis measures in real samples often tend to be approximately normally distributed (authors, unpublished work). It is reasonable to analyse these data untransformed.

The actual structure of data in samples is also manipulated by the choices of taxonomic resolutions used to identify groups of organisms. Much of the important information in the analyses of assemblages can be retained when relatively coarse taxonomic resolutions are applied (Warwick 1988a,b; Warwick and Clarke 1995). Using families or orders (Chapman 1998; Ferraro and Cole 1990; Gray et al. 1990; Herman and Heip 1988; Olsgard and Gray 1995; Somerfield and Clarke 1995) drastically reduces the effort, time and, above all, costs of sampling in many habitats. For some parts of the world and some habitats, taxonomy of species is uncertain, so arbitrarily larger groupings are inevitable. As with transformation of the data, issues of taxonomic resolution must await further investigation.

The method described has some advantages over other methods available in terms of the complexity of designs of studies and sampling that can be used and in terms of relative robustness of (and, for many ecologists, familiarity with) the analytical procedure used. It remains to be seen whether its later uses will identify other advantages, problems unforseen here and, hopefully, new developments based on it. In the interim, it is available as part of the armoury of techniques for examining complex assemblages.

Acknowledgements This study was supported by funds from the Australian Research Council (Special Investigator's Award to A.J.U; Research Centre Funding to A.J.U and M.G.C) and from the Institute of Marine Ecology. We thank J. Harris and V. Mathews for assistance in the preparation of the manuscript and numerous colleagues for help with field-work. As ever, Dr. K.R. Clarke (Plymouth Marine Laboratories) helped prevent serious errors and contributed to our comprehension.

References

Allen TFH, Starr TB (1982) Hierarchy: perspectives for ecological complexities. University of Chicago Press, Chicago

Anderson RL (1959) Use of contingency tables in the analysis of consumer preference studies. Biometrics 15:582–590

Bell G, Lechowicz MJ, Appenzeller A, Chandler M, DeBlois E, Jackson L, Mackenzie B, Preziosi R, Schallenberg M, Tinker N (1993) The spatial structure of the physical environment. Oecologia 96:114–121

Bell SS, Hall MO, Robbins BD (1995) Toward a landscape approach in seagrass beds: using macroalgal accumulation to address questions of scale. Oecologia 104:163–168

Bourget E, Deguise J, Daigle G (1994) Scales of substratum heterogeneity, structural complexity, and the early establishment of a marine epibenthic community. J Exp Mar Biol Ecol 181:31–51

Bray JR, Curtis JT (1957) An ordination of the upland forest communities of southern Wisconsin. Ecol Monogr 27:325–349

Brown JH, Mehlman DW, Stevens GC (1995) Spatial variation in abundance. Ecology 76:2028–2043

Chapman MG (1998) Relationships between spatial patterns of benthic assemblages in mangrove forests using different levels of taxonomic resolution. Mar Ecol Prog Ser 162:71–78

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Aust J Ecol 18:117–143

Clarke KR, Ainsworth M (1993) A method of linking multivariate community structure to environmental variables. Mar Ecol Prog Ser 92:205–219

Clarke KR, Green RH (1988) Statistical design and analysis for a "biological effects" study. Mar Ecol Prog Ser 46:213–226

Cliff AD, Ord JK (1973) Spatial autocorrelation. Pion, London

Dayton PK, Tegner MJ (1984) Catastrophic storms, El Niño, and patch stability in a southern California kelp community. Science 224:283–285

Faith DP, Minchin PR, Belbin L (1987) Compositional dissimilarity as a robust measure of ecological distance. Vegetatio 69:57–68

Ferraro SP, Cole FA (1990) Taxonomic level and sample size sufficient for assessing pollution impacts on the Southern California Bight macrobenthos. Mar Ecol Prog Ser 67:251–262

Field JG, Clarke KR, Warwick RM (1982) A practical strategy for anlysing multispecies distribution patterns. Mar Ecol Prog Ser 8:37–52

Gray JS, Clarke KR, Warwick RM, Hobbs G (1990) Detection of the initial effects of pollution on marine benthos: an example from the Ekofisk and Eldfisk oilfields, North Sea. Mar Ecol Prog Ser 66:285–299

Green RH (1979) Sampling design and statistical methods for environmental biologists. Wiley, Chichester

Green RH (1980) Multivariate approaches in ecology: the assessment of ecological similarity. Annu Rev Ecol Syst 11:1–14

Green RH (1984) Statistical and non-statistical considerations for environmental monitoring studies. Environ Monit Assess 4:293–301

Herman PMJ, Heip C (1988) On the use of meiofauna in ecological monitoring: who needs taxonomy? Mar Poll Bull 19:665–668

Legendre P, Fortin M-J (1989) Spatial pattern and ecological analysis. Vegetatio 80:107–138

Menge BA (1995) Indirect effects in marine rocky intertidal interaction webs: patterns and importance. Ecol Monogr 65:21–74

Menge BA, Olson AM (1990) Role of scale and environmental factors in regulation of community structure. Trends Ecol Evol 5:52–57

Menge BA, Berlow EL, Blanchette CA, Navarrete SA, Yamada SB (1994) The keystone species concept: variation in interaction strength in a rocky intertidal habitat. Ecol Monogr 64:249–286

Olsgard F, Gray JS (1995) A comprehensive analysis of the effects of offshore oil and gas exploration and production on the benthic communities of the Norwegian continental shelf. Mar Ecol Prog Ser 122:277–306

Pielou EC (1984) The interpretation of ecological data: a primer on classification and ordination. Wiley, New York

Scheffé H (1959) The analysis of variance. Wiley, New York

Schneider DC (1994) Quantitative ecology: spatial and temporal scaling. Academic Press, San Diego

Smith W (1989) ANOVA-like similarity analyses using expected species shared. Biometrics 45:873–881

Somerfield PJ, Clarke KR (1995) Taxonomic levels in marine community studies revisited. Mar Ecol Prog Ser 127:113–119

Underwood AJ (1981) Structure of a rocky intertidal community in New South Wales: patterns of vertical distribution and seasonal changes. J Exp Mar Biol Ecol 51:57–85

Underwood AJ (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 AJ (1994) On beyond BACI: sampling designs that might reliably detect environmental disturbances. Ecol Appl 4:3–15

Underwood AJ (1997) Ecological experiments: their logical design and interpretation using analysis of variance. Cambridge University Press, Cambridge

Underwood AJ, Chapman MG (1996) Scales of spatial patterns of distribution of intertidal invertebrates. Oecologia 107:212–224

Underwood AJ, Chapman MG (1998) Spatial analyses of intertidal assemblages on sheltered rocky shores. Aust J Ecol 23:138–157

Underwood AJ, Petraitis PS (1993) Structure of intertidal assemblages in different locations: how can local processes be compared? In: Ricklefs RE, Schluter D (eds) Species diversity in ecological communities: historical and geographical perspectives. University of Chicago Press, Chicago, pp 38–51

Warwick RM (1988a) Analysis of community attributes of the macrobenthos of Frierfjord/Langesundfjord at taxonomic levels higher than species. Mar Ecol Prog Ser 46:167–170

Warwick RM (1988b) The level of taxonomic discrimination required to detect pollution effects on marine benthic communities. Mar Poll Bull 19:259–268

Warwick RM, Clarke KR (1995) New "biodiversity" measures reveal a decrease in taxonomic distinctness with increasing stress. Mar Ecol Prog Ser 129:301–305

Wootton JT (1994) The nature and consequences of indirect effects in ecological communities. Annu Rev Ecol Syst 25:443–466