Vol. 320: 11–27, 2006 | MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser | Published August 29

Dispersion-based weighting of species counts in assemblage analyses

K. R. Clarke[1,2,*], M. G. Chapman[2], P. J. Somerfield[1], H. R. Needham[1]

[1]Plymouth Marine Laboratory, Prospect Place, West Hoe, Plymouth PL1 3DH, UK
[2]Centre for Research on Ecological Impacts of Coastal Cities, Marine Ecology Laboratories A11, University of Sydney, New South Wales 2006, Australia

ABSTRACT: Multivariate analysis of species assemblage data often begins with transformation of abundances, to downweight contributions from dominant taxa to Bray-Curtis dissimilarities computed among samples. Although usually effective, global transformation is a blunt tool: it ignores differences in the variance structure of counts of individual species. Species which are highly spatially clustered should logically be given less weight than those with similar mean abundance but whose replicate counts have the lower variance associated with Poisson distributions (for which individual organisms arrive randomly and independently into the sample). Where replicates are available within sample groups specified *a priori*, this differential downweighting is achieved by dividing the counts for each species by their index of dispersion D, the variance to mean ratio, a clustering ('clumping') measure calculated from replicates within a group, and then averaged across groups. The procedure is justified by assuming a generalised Poisson model for counts, allowing different species to have arbitrarily differing degrees of clustering. Downweighting is applied only where a species shows significant evidence of clumping, this being tested by a powerful, exact permutation test that replaces the standard (large-sample) χ^2 test for $D = 1$, which is often invalid because of low 'expected' frequencies. The resulting dispersion-weighted data matrix has a common (Poisson-like) variance structure across species, but unchanged relative responses of a species in different groups. Transformation may still be needed but now only to downweight consistently abundant species relative to equally consistent but less numerous species, rather than also dealing with erratic counts. Dispersion weighting is shown to be effective in 3 studies that examine: soft-sediment copepods in the metal-polluted Fal estuary, UK; benthic macrofauna of mangrove forests in Bicentennial Park, New South Wales, Australia; and sediment nematodes within and outside seagrass beds in the Yealm estuary, UK. A fourth data set, on macrobenthos from Loch Creran, UK, is added to a comparison of the differing cross-species distributions of the dispersion index.

KEY WORDS: Multivariate analysis · Transformation · Species weighting · Spatial clustering · Dispersion index · Generalised Poisson · Permutation test

INTRODUCTION

This study deals with counts of organisms for a range of species, or higher taxa, in replicate samples from different sites, times and/or treatments (where we use the term 'sample' to denote a single sampling unit such as a grab, quadrat, net haul, etc.). Such data sets are commonly analysed by a multivariate procedure with the following steps:

(1) Transformation of the counts, choosing from options of increasing severity, ranging from no transformation, square root, 4th-root (and, equivalently, $\log(1 + x)$, Clarke & Warwick 2001), to reduction of the counts to presence (1) or absence (0), which can be viewed as the ultimate in severe power transformation, as the exponent goes to zero.

(2) Calculation of similarity between every pair of replicates, within and among the sample groups. This

*Email: krc@pml.ac.uk

frequently uses the Bray-Curtis coefficient (Bray & Curtis 1957), for good reason (the benefits of this choice are articulated by Clarke & Warwick 2001, and Clarke at al. 2006).

(3) Construction of non-metric multi-dimensional scaling plots (MDS, Kruskal 1964) for displaying the patterns among sample groups, and analysis of similarity tests (ANOSIM, Clarke & Green 1988, Clarke 1993) for testing hypotheses about group differences, in relation to variability among replicates within a group. The latter are multivariate permutation procedures carrying out the equivalent hypothesis tests to (inter alia) 1-way ANOVA in univariate statistics.

The initial transformation step is usually justified on the grounds that it reduces the contribution of highly abundant species in relation to less abundant ones in the calculation of the Bray-Curtis measure, with the more severe the transformation, the greater the contribution made by rarer species (Clarke & Green 1988, Clarke & Warwick 2001). This is certainly true. The Bray-Curtis dissimilarity between any 2 samples, 1 and 2, is defined as

$$D_{12}{}^{B-C} = 100 \cdot \frac{\sum_l |y_{l1} - y_{l2}|}{\sum_l (y_{l1} + y_{l2})} \quad (1)$$

where y_{l1} and y_{l2} are the transformed counts of the lth species ($l = 1, 2, \ldots, p$) in Samples 1 and 2, respectively. Species with high abundance clearly dominate both the numerator and denominator of this coefficient, unless a heavy transform reduces the differential between large and small counts. For example, a reduction to presence/absence leaves each species only capable of contributing either '0' or '1', and less-frequently observed species can contribute more to the numerator now than common species, especially if the latter are ubiquitous (so always return $|1 - 1| = 0$ to the top line of Eq. 1).

This is not, however, the last word on the subject of globally-applied transformations. Quite often, a severe transformation is employed not as a subtle way of weighting up the contribution from a modestly abundant but consistently observed species, compared with a more highly abundant but also consistently observed species, but as a simple reaction to a species whose numbers are fluctuating strongly in replicates, from zero to very high counts, and this variability is dominating the analysis. Typically, certainly for soft-sediment marine communities, disturbed environments are characterised by the presence of opportunist species, which are often small-bodied but found in very large and variable numbers because of strongly clustered spatial distributions. The presence of species with highly erratic counts among replicates within a treatment (or site or time) group is not restricted to impact studies, but frequently arises naturally, perhaps as a function of the species' mode of reproduction or from behavioural traits such as schooling. Use of a severe transformation will downweight the large element of 'noise' that this variability can throw into the 'signal' of natural or induced community change, but at quite a high price. Good discriminating species, which have consistent but modest abundance levels among replicates within a group, and subtle but consistent changes between groups, will also feel the full weight of a severe transformation, which will reduce the contribution of their signal. What is needed here is to separate out the statistical and biological reasons for transformation, by first dealing with the statistical problem that some species should be given less weight because of their inherent unreliability in replicates—a sampling problem caused by spatial clumping of the organisms. Only then is it possible to address, by transformation, the biological problem of whether it is better to look for a consistent signal of change in common or rarer species.

This study describes a simple, and fully transparent, means of downweighting species counts that exhibit clumping in replicates. The procedure is fully justified under a general, flexible model of clustering behaviour that is partly distribution-free (a form of generalised Poisson or 'contagious' distribution, Douglas 1979), which is realistic in many ecological contexts. The derived weighting is still likely to be sensible even when the model conditions are not strictly met. As described in this study, suitable applications in practice require 3 conditions: that the data for each species are genuine counts (not densities that have been standardised to some unit volume of water or unit area of sediment etc.); that there are independent replicates within each of the sample groups, so that there is some basis for assessing the within-group variance structure; and that each replicate is of a uniform size (same core diameter, same tow length etc.). It is not required that the number of replicates within each group be the same—though balance in experimental or observational design is always a desirable trait that increases power when carrying out tests for group differences, subsequent to calculation of dispersion-weighted similarities.

METHODS

Generalised Poisson model. The data structure is assumed to be that of a standard 1-way layout, with p species counted for a total of N samples, the latter split into g groups, defined *a priori*. This does not exclude higher-way designs. Groups could be different treatments, sites, times or some combination of these, e.g. a 2-way structure of sites (S) and times (T) would simply

be 'flattened' to a 1-way layout of groups S1T1, S1T2, S1T3, S2T1, S2T2, S2T3, etc. (After the dispersion weighting step, the original higher-way design structure can simply be re-instated for subsequent analysis.) The ith group (i = 1, 2, …, g) contains n_i replicate samples (quadrats, cores, tows etc.) identified by the subscript j (j = 1, 2, …, n_i). Balanced replication would have $n_i = n$, for all i, giving a total of $N = gn$ replicates; for unbalanced replication, $N = \Sigma n_i$. Although clearly more widely applicable than this, it is helpful to think of the organisms of a single species as points located in 2D space, the groups as different sites, and the replicate samples as quadrats 'capturing' these points, with a count of X_{ij} individuals of that species in the jth replicate quadrat of the ith group. This is shown schematically in Fig. 1.

The simple conceptual model for species counts, which motivates the dispersion-weighting procedure, begins by assuming that locations of organisms of different species are independent of each other. This is not to say that correlations calculated between pairs of species across all samples will be zero, because different groups will represent changed environmental conditions, which will cause some species to increase or decrease their mean density in the same way and other species to react in opposite ways. What it is saying is that, within a group (assumed to represent a fixed environmental condition), species do not have competitive interactions or synergies that cause inter-species correlations in the counts across replicates within that group. This is inevitably an over-simplification, but a very necessary one in order to make any sort of theoretical progress. Such inter-species correlations within replicates are usually 'second-order' in relation to the induced correlations from changing conditions along an environmental gradient, or differing treatment regimes in an experimental manipulation. A consequence of this local independence assumption is that we can deal with each species separately, and choose the weighting that is appropriate to its clumping structure (different species can, and will, exhibit different degrees of clustering). There is therefore no need to identify each species by a different subscript in what follows (or in Fig. 1). All notation and formulae refer to the specific, single species under consideration.

If organisms were always randomly located in space then the count X_{ij} from the jth replicate of the ith group would come from a Poisson distribution and, taken across the replicates from a single group, the sample mean and variance would be much the same (a characteristic of the Poisson distribution is that the true mean and variance are identical). Different groups would have different mean densities of organisms, as the species responds to differing conditions, but the variance would always go up or down with the mean density. The ideal would be for all species to display this property because then every 'capture' of an individual is an independent event, and the multivariate analysis would end up giving equal weight to each such event. This ideal, however, is rarely the case in practice. Many species are spatially clumped: they are not captured one at a time but perhaps 5, 50 or even 500 at a time. The variance in replicates is now very much greater than the mean (this is referred to as overdispersion) and such 'erratic' species need to be downweighted in relation to those that are Poisson-distributed. Fig. 1 displays such a situation, with a mild degree of clustering: organisms tend to arrive in clumps of 3 (with some variation around that mean number) but the centres of clumps appear to be more or less randomly distributed in space, with higher mean density of centres for Site 2, lower for Site 1 and least for the final Site (group) g. This is the scenario that needs to be modelled, and it can be achieved by any distribution from the generalised Poisson family, which has 2 components

(1) 'Centres of population' are randomly located in space. The number of centres, C_{ij}, in the jth replicate of the ith group therefore has a Poisson distribution, with mean γ_i say. The density of centres, γ_i, changes with the group i, in response to the changing environment. Some groups will have

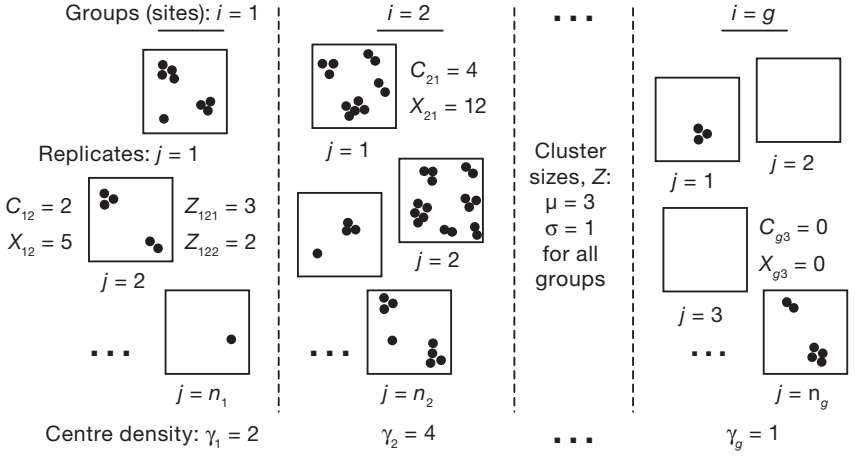

Fig. 1. Schematic of quadrat sampling from a generalised Poisson model for the spatial locations of organisms (●) of a single species. Groups are denoted by i (= 1, …, g), replicate quadrats within a group by j (= 1, …, n_i), and the organism count for the (i, j)th replicate by X_{ij}, consisting of C_{ij} centres with Z_{ijk} organisms at the kth centre (= 1, …, C_{ij}). Mean density of centres per quadrat (γ_i) varies across groups but clustering structure is constant (mean cluster size $\mu \approx 3$, variance $\sigma^2 \approx 1$). The only recorded quantities would be the counts $\{X_{ij}\}$

replicate samples with low, or even zero, density; other groups will represent advantageous conditions for that species and replicates will contain many centres.

(2) Associated with each centre are a number of organisms, Z, from some probability distribution whose form does not need to be specified but which has mean μ and variance σ^2. In effect, this says that for every 'centre' which falls into the replicate, an average of μ organisms are counted. If $\sigma^2 = 0$, then that is exactly what happens: every time we come across 1 organism we capture μ of them. The further special case of $\mu = 1$ gets back to the simple Poisson model for replicate counts. In the general case, the degree of clumping for that species, which is specified by μ (and σ^2), is assumed to remain constant across all groups; the effect of the changing sites, times or experimental conditions is assumed to change the density of this species but not its innate clustering structure.

Under these conditions, Appendix 1 shows that the mean of X_{ij}, the count for a particular species from the jth replicate of the ith group, is $\gamma_i\mu$, and the variance of X_{ij} is $\gamma_i(\mu^2 + \sigma^2)$. The ratio of the variance to the mean, called the 'index of dispersion' D, is therefore $(\mu^2 + \sigma^2)/\mu$, which is not a function of the group i to which the replicate belongs. (The term 'index of dispersion' has a long history in statistics, dating from Fisher at al. 1922; many ecologists will have first met it in the work of Greig-Smith 1952 or the book by Elliott 1971).

Dispersion weighting procedure. The properties of this generalised Poisson model motivate the weighting procedure for any particular species, with the following steps

(1) Calculate the observed index of dispersion, D_i, for each group i, simply as the ratio of the sample variance to the sample mean for the n_i replicates in that group.

(2) Average these indices across all g groups, since they are all estimating the same quantity, to obtain the average index of dispersion:

$$\bar{D} = \frac{\sum_{i=1}^{g}(n_i - 1)D_i}{\sum_{i=1}^{g}(n_i - 1)} \quad (2)$$

This is a weighted average of the dispersions for each group because, in general, the numbers of replicates in each group $\{n_i\}$ may not be equal. The choice of weighting is a natural one, following from the close link between indices of dispersion and chi-squared statistics (see the discussion on testing for the absence of clumping in the Appendix). When replication is balanced, Eq. 2 is an unweighted average, $(\Sigma D_i)/g$, of the estimates from each group, a point returned to in the 'Discussion'.

(3) Use this average index of dispersion to downweight that species, i.e. divide all of its counts, X_{ij}, by $\bar{D}$. The effect of this downweighting is to produce values for that species which are 'Poisson-like', in the sense that their mean and variance, over replicates within a group, are now approximately equal; their over-dispersion has been removed. In fact, Appendix 1 shows that the mean and variance of $X_{ij}^{DW} = X_{ij} / \bar{D}$ are both approximately $\gamma_i / [1 + (\sigma / \mu)^2]$.

Exactly the same procedure is repeated for each species, so that no species should be left with an erratic, over-dispersed distribution among replicates. The downweighted matrix can now be analysed in a standard way, e.g. by calculating Bray-Curtis similarities between all samples and entering this resemblance matrix into multivariate display and testing routines.

There is another small step that could be interpolated in the procedure between Steps (2) and (3). If there is no statistical evidence that a particular species exhibits any over-dispersion at all, then it clearly should not be subject to any downweighting, irrespective of the observed value of $\bar{D}$. This can only happen if $\bar{D}$ is close to 1—its value under the null hypothesis (H_0) of no clumping—or if there are very few replicates and/or very sparse counts; it is not logically satisfactory to divide a species through by $\bar{D} = 1.5$ say, if there is no evidence that this departs from a true dispersion index of 1. What is needed here is an exact hypothesis test for over-dispersion, valid for both large and small counts.

Test for over-dispersion. Under the H_0 that the counts for a species are Poisson, i.e. the dispersion index is $D = 1$, Appendix 1 shows that

$$X^2 = \left[\sum_{i=1}^{g}(n_i - 1)\right] \cdot \bar{D} = \sum_{i=1}^{g}\left[\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2 \Big/ \bar{X}_i\right] \quad (3)$$

has the form of a chi-squared statistic (the standard form of a sum of $(O - E)^2/E$ terms being clearly discernible). Under the usual 'large sample' conditions, it therefore follows a χ^2 distribution on $\Sigma(n_i - 1)$ degrees of freedom (df), and a significance test would reject the H_0 if X^2 were greater than the upper 5% point of χ^2 on $\Sigma(n_i - 1)$ df.

Assuming that the quadrats displayed are the only data (i.e. $g = 3$, $n_1 = 3$, $n_2 = 4$, $n_3 = 4$), the working steps of the χ^2 test and dispersion weighting procedure for the data in Fig. 1 are shown (Table 1). The test suggests that the observed dispersion of 3.17 is large enough to reject the hypothesis that $D = 1$ unambiguously ($p < 0.001$), so that downweighting of the counts X_{ij} would be carried out. The resulting values, X_{ij}^{DW}, are seen to have the 'Poisson-like' property of approximate equality of variance and mean. The counts in this case are more or less large enough for the χ^2 approximation to be reasonable, even though the standard guideline that expected values should mostly be >5 (see Cochran 1954) is not maintained for the counts in Site 3. In practice, the approximation will be much more doubtful for many species in a typical assemblage matrix. In addition to the species that are rare

Table 1. Dispersion weighting procedure for the hypothetical data displayed in Fig. 1. Final column shows both the approximate test of X^2, referred to χ^2 on 8 df, and the exact permutation test of the average of the $\{D_i\}$ dispersion indices, $\bar{D} = (X^2/8)$. $\{X_{ij}^{DW}\}$ are the resulting dispersion weighted values

Groups i =	1			2				3				
Replicates j =	1	2	3	1	2	3	4	1	2	3	4	
Observed counts X_{ij}	8	5	1	12	17	4	8	3	0	0	6	
Expected counts (mean) $\bar{X}_i$		4.67			10.25				2.25			$X^2 = 25.4$ ($p < 0.001$ approx)
Var = $\sum_j (X_{ij} - \bar{X}_i)^2 / (n_i - 1)$		12.33			30.92				8.25			
Dispersion index D_i		2.64			3.02				3.67			$\bar{D} = 3.17$ ($p < 0.001$ exact)
$X_{ij}^{DW} = X_{ij}/\bar{D}$	2.5	1.6	0.3	3.8	5.4	1.3	2.5	0.9	0	0	1.9	
Mean X_i^{DW}		1.47			3.24				0.71			
Var X_i^{DW}		1.23			3.08				0.82			$\bar{D}^{DW} \equiv 1$

everywhere, if the groups correspond to a long baseline of environmental change, giving a high turnover of species, many species will have low counts in replicates from at least one group and thus expected values that are 'too low'.

Permutation tests come to the rescue here, as they do in other aspects of community analysis (Clarke 1993) and in biology in general (Manly 1997). The X^2 form of Eq. (3) is a perfectly valid and sensible test statistic for assessing the hypothesis of clumping of organisms (indeed, tests based on indices of dispersion 'have no serious rivals' in the simple case of replicate quadrat counts from a single group, according to Diggle (1983), and the same is likely to be true of the compound dispersion index introduced here). The problem is simply in determining its exact probability distribution in small samples under the H_0 of no clumping, but this can be derived by permutation. For the data of Table 1, the 14 ind. in Group 1 are independently and randomly allocated to one of the 3 quadrats, mimicking the reality (under H_0) of random arrivals from a known density of organisms. Simultaneously, the 41 ind. in Group 2 are similarly dispersed amongst its 4 replicates, and the 9 ind. from Group 3 amongst its 4 replicates. This is a constrained permutation: obviously organisms are not permuted across the boundaries between groups (a total of 64 ind. allocated randomly to 11 quadrats) because that would only be appropriate to an H_0 specifying common densities across groups, in addition to spatial randomness. A value of X^2 is calculated from Eq. (3) for this re-arrangement and the whole simulation is then repeated, e.g. a total of 999 times. A histogram of the resulting X^2 values gives the null distribution and an upper 5% point, against which the genuine X^2 can be compared. In this case, all 999 simulations give a smaller value than 25.4 for X^2 (3.17 for $\bar{D}$), so the significance level is confirmed at $p < 0.001$ and the hypothesis $D = 1$ rejected.

A simple permutation procedure of this sort is easy to program, and powerful in detecting departures of counts from the Poisson model denoting spatial (or temporal) randomness, even with sparse data, so it is surprising that it has not been more commonly used in ecology. For example, consider just a single group for the moment, in which there are 8 replicate samples giving counts of (0, 0, 0, 0, 4, 0, 0, 0) organisms. The standard X^2 'model goodness-of-fit' test, attempting to fit Poisson probabilities $e^{-\lambda}\lambda^k/k!$ to the frequencies of occurrence of each response (7, 0, 0, 0, 1, 0, 0, … for k = 0, 1, 2, 3, 4, …) cannot even be contemplated with such a small total frequency as 8. The index of dispersion test, which also utilises a X^2 form of test statistic (though conceptually different to the model goodness-of-fit test), has expected values $E_i = 0.5$ in all samples, matching the observed values O_i = (0, 0, 0, 0, 4, 0, 0, 0). This gives $X^2 = 28$, but a $\chi^2_{7\text{df}}$ approximation to the null distribution would be untenable with all $E_i = 0.5$. The permutation test, on the other hand, has no such constraints and shows that this X^2 value is too large to have been obtained by chance ($p < 0.002$); if the density of organisms per quadrat is 0.5, and they are genuinely located randomly and independently, it is too much of a coincidence that the only organisms captured were all in the same quadrat. If this were the only group of samples being analysed, the dispersion weighting procedure would therefore divide all counts by D (= 4) and give the sensible downweighted values of (0, 0, 0, 0, 1, 0, 0, 0) for that species.

Special cases of generalised Poisson model. As presented earlier, the generalised Poisson model is not fully parametric: it is not necessary to specify the distribution of clump sizes (Z) in order for dispersion weighting to be justifiable and operable. Nonetheless, some widely used models are special cases of this general construction, so are worth noting; 5 cases are discussed

(1) ***Poisson distribution:*** If $\mu = 1$, $\sigma^2 = 0$, then there is only 1 organism at each 'centre', so the counts X_{ij} are

Poisson distributed, Po(γ_i) for group i and dispersion $D = 1$. This is the 'ideal' that we aim to approximate in other cases by downweighting.

(2) ***Fixed size clusters:*** If $\mu = m$, $\sigma^2 = 0$, all individuals arrive in clusters of size $Z = m$. The dispersion index is $D = (\mu^2 + \sigma^2)/\mu = m$, and dividing the counts by m exactly restores the Poisson distribution, X_{ij} is Po(γ_i) for group i. Although never a precisely realistic model, in many ways this is the defining model for the approach adopted here. We would like the data to be Poisson distributed, reflecting independent arrivals of organisms into a sample — which is the basis of many ecological models, such as the rarefaction approach of Sanders (1968) and Hurlbert (1971). Some species arrive in ones, but some in tens and some in hundreds. If, for a particular species, every time we see 1 we see 10, then nothing could be more natural than to divide the counts by 10, which restores the Poisson distribution and the independence of arrivals (in effect, 'centres' C, not individuals, are counted). A model in which cluster size Z varies ($\sigma^2 \neq 0$) is more realistic of course, but the same idea carries over. The dispersion index gives the parameter with which to divide by, in order to bring the data back to the Poisson-like property of variance = mean.

(3) ***Negative binomial distribution:*** If the cluster sizes Z are drawn from a log series distribution with probabilities $[\log(1 + \beta)]^{-1}[\beta/(1 + \beta)]^z/z$ for $z \geq 1$ (Fisher et al. 1943), which has $\mu = \beta/\log(1 + \beta)$, $\sigma^2 = \mu[(1 + \beta) - \mu]$, then Quenouille (1949) showed that the resulting counts X_{ij} are negative binomial. Here the mean of X_{ij} is $\gamma_i\beta/\log(1 + \beta)$ and the variance $\gamma_i\beta(1 + \beta)/\log(1 + \beta)$, so the dispersion index $D = 1 + \beta$, and the parameter β reflects the degree of clustering.

(4) ***Neyman type A distribution:*** Neyman (1939). For all these models, the number of 'centres' (C) is assumed to be Poisson, mean γ_i, but if the number of organisms at each centre (Z) is also drawn from a Poisson distribution Po(λ) with probabilities $e^{-\lambda}\lambda^z/z!$ ($z \geq 0$), then $\mu = \lambda$, $\sigma^2 = \lambda$, and the mean and variance of the counts X_{ij} are $\gamma_i\lambda$ and $\gamma_i\lambda(1 + \lambda)$ respectively. This gives a dispersion index of $D = 1 + \lambda$, so λ again reflects the extent of clustering.

(5) ***Pólya-Aeppli distribution:*** (Kendall & Stuart 1963). This is obtained if the cluster size Z has the geometric probabilities $(1 + \tau)^{-1}[\tau/(1 + \tau)]^{z-1}$ ($z \geq 1$), so that $\mu = 1 + \tau$, $\sigma^2 = \tau(1 + \tau)$, and the mean and variance of X_{ij} are $\gamma_i(1 + \tau)$ and $\gamma_i(1 + \tau)(1 + 2\tau)$. This gives $D = 1 + 2\tau$, which is independent of the mean number of centres γ_i, as it must always be in this formulation. In all these cases, by definition, dividing X_{ij} by D restores the property of variance = mean.

RESULTS

Simple numerical example

Return to the hypothetical data of Table 1 but add 2 further species (Table 2a). The counts for Species 2 have the same structure as for Species 1 but are a factor of 10 greater. Clearly, Species 2 has much greater clumping of individuals, reflected in a $\bar{D}$ that is also multiplied by 10. Dispersion weighted values (Table 2b) are now identical for the 2 species, and they are given exactly the same weight in the subsequent analysis. In contrast, Species 3 has precisely the same mean abundance for each of the 3 groups as Species 2, but arrivals are much less clustered. In fact, Species 3 has the same degree of clumping of organisms as Species 1 ($\bar{D}$ = 3.17), so clearly has a substantially larger number of independent 'centres'. After dispersion weighting, it is seen to have much greater impact than the other 2 species (Table 2b), making an order of magnitude greater contribution to the numerator and denominator of the (untransformed) Bray-Curtis coefficient. This is precisely what dispersion weighting sets out to achieve: emphasis is placed on those species (such as Species 3) that are both high in abundance and consistent in replication. In fact, if this were a sub-

Table 2. Dispersion weighting for hypothetical data from 3 species. (a) Counts {X_{ij}}, mean counts in each group and average index of dispersion $\bar{D}$ over the groups. (b) Dispersion weighted values {X_{ij}^{DW}}, and their means in each group

	Group 1			Mean	Group 2				Mean	Group 3				Mean	$\bar{D}$
(a) X_{ij}															
Species 1	8	5	1	4.7	12	17	4	8	10.3	3	0	0	6	2.3	3.17
Species 2	80	50	10	47	120	170	40	80	103	30	0	0	60	23	31.7
Species 3	55	50	35	47	105	125	85	95	103	25	15	15	35	23	3.17
(b) X_{ij}^{DW}															
Species 1	2.5	1.6	0.3	1.5	3.8	5.4	1.3	2.5	3.2	0.9	0	0	1.9	0.7	1
Species 2	2.5	1.6	0.3	1.5	3.8	5.4	1.3	2.5	3.2	0.9	0	0	1.9	0.7	1
Species 3	17.3	15.7	11.0	14.7	33.1	39.4	26.8	29.9	29.0	7.9	4.7	4.7	11.0	7.1	1

set of real data, it might legitimately be argued that Species 3 could be overdominant in an untransformed analysis. A mild transformation such as a square root (applied after dispersion weighting) could therefore be justifiable, appealing to the usual argument about needing to balance contributions of abundant (and consistent) species with rarer (and consistent) ones.

Copepods in creeks of the Fal estuary

Somerfield at al. (1994) present data from 5 creeks of the Fal estuary system, southwest England, characterised by high and varying sediment concentrations of heavy metals—a result of tin and copper mining over hundreds of years. The data used here consist of abundances of 23 infaunal copepod species in 5 replicate meiofauna cores from each creek (7 from the largest creek). Somerfield at al. (1994) showed a significant and substantial difference between creeks in copepod assemblages, using a 1-way analysis of similarities (ANOSIM) test (R = 0.66, p < 0.001) on Bray-Curtis similarities calculated from 4th-root transformed abundances.

Table 3 gives counts for 5 replicates from each of 2 creeks, Mylor and Pill, together with the estimated dispersion indices $\bar{D}$ for 19 species (the omitted species are absent from both creeks) and significance levels (p, as a percentage) of their departure from D = 1 in a 1-sided permutation test based on 999 simulations. MDS ordinations (Fig. 2), based on Bray-Curtis coefficients without and with dispersion weighting (and with no subsequent transformation), show that a greater separation of the 2 creeks is evident under dispersion weighting, as is a tightening up of the variation among replicates for Mylor. This is confirmed by a 1-way ANOSIM test, which takes place in the high-dimensional space of the similarity matrices and not the 2D approximation represented by the ordinations (Clarke & Green 1988, Clarke 1993). The ANOSIM statistics are R = 0.41 for Fig. 2a and R = 0.71 for Fig. 2b, both significant (p < 0.016 and p < 0.008, respectively) but the latter clearly larger, indicating greater distinction of the assemblages.

The reason for the comparative success of dispersion weighting in this case is clear. Of the 3 most abundant species, *Microarthridion fallax*, *Enhydrosoma gariene* and *Platychelipus littoralis*, which will tend to dominate the unweighted analysis, the latter 2 are rather erratic across replicates. The greatest dispersion is found for *P. littoralis*, with abundances ranging from 1 to 88 for Mylor and 12 to 112 for Pill samples. This represents relatively strong over-dispersion, D_M = 35.9 for Mylor and D_P = 36.2 for Pill, giving an average $\bar{D}$ = 36.1. *E. gariene* is similarly over-dispersed, $\bar{D}$ = 27.7. These 2 species will therefore be heavily downweighted in relation to others, allowing the analysis to be steered by less erratic but also abundant (*M. fallax*) or fairly abundant (*Mesochra lilljeborgi*) species, and by highly consistent though low abundance species such as *Stenhelia elizabethae*. Global transformation would, of course, achieve some of the same objectives: the con-

Table 3. Fal estuary copepods. Original counts from 5 replicate meiofauna cores in each of 2 creeks (Mylor, M and Pill, P). Initial columns give the average dispersion index $\bar{D}$, its % significance (p%) in a test of D = 1, and the divisor actually used to downweight each row of the table under the dispersion weighting procedure

Species	$\bar{D}$	p%	Div	M1	M2	M3	M4	M5	P1	P2	P3	P4	P5
Pseudobradya curticorne	8.7	0	8.7	19	2	12	32	13	18	0	6	26	7
Pseudobradya sp. 2	1.5	20.2	1	0	0	2	0	0	0	0	0	0	1
Halectinosoma gothiceps	7.3	0	7.3	0	0	0	0	0	16	6	8	0	0
Tachidius discipes	3.1	0	3.1	6	2	8	0	0	0	0	0	0	2
Microarthridion fallax	8.2	0	8.2	110	88	76	92	58	60	38	38	22	7
Harpacticus flexus	1.5	23.9	1	0	0	0	0	0	2	2	0	0	0
Stenhelia palustris	2.2	2.1	2.2	3	2	8	4	6	12	6	14	2	5
Stenhelia elizabethae	0.5	73.4	1	1	4	4	4	4	0	0	0	0	0
Amphiascoides limicola	3.4	0.1	3.4	1	0	0	0	0	14	14	8	22	0
Robertsonia celtica	5.8	0	5.8	4	24	8	8	0	0	4	6	2	1
Mesochra lilljeborgi	6.6	0	6.6	1	0	6	0	1	10	12	14	34	3
Mesochra pygmaea	2.7	2.3	2.7	0	0	0	0	0	2	0	0	4	0
Enhydrosoma gariene	27.7	0	27.7	19	130	26	44	21	50	34	44	24	6
Enhydrosoma longifurcatum	2.1	2.5	2.1	0	0	2	0	0	10	8	6	4	1
Enhydrosoma propinquum	1.5	20.2	1	1	0	0	0	0	0	0	0	2	0
Platychelipus littoralis	36.1	0	36.1	43	88	26	8	1	12	34	76	112	18
Paronchocamptus curticaudatus	3.6	1.4	3.6	0	0	0	0	0	0	8	8	12	3
Paronychocamptus nanus	1	100	1	0	0	0	0	1	0	0	0	0	0
Asellopsis sp.	1	100	1	1	0	0	0	0	0	0	0	0	0

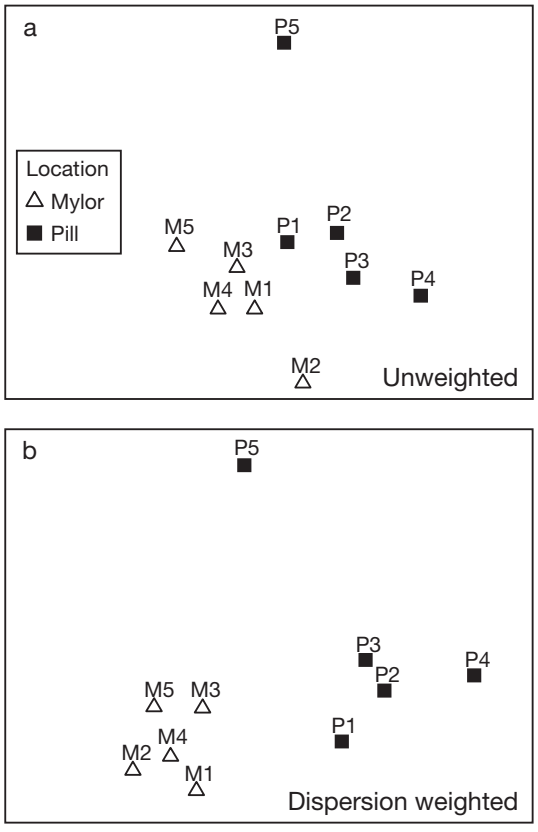

Fig. 2. Fal estuary copepod assemblages from 5 replicates in each of Mylor (M) and Pill (P) creeks. Non-metric multidimensional scaling (MDS) ordinations based on Bray-Curtis similarities from untransformed data: (a) unweighted data (standard analysis); (b) dispersion weighting carried out using divisors from Table 3. Stress values: (a) 0.08; (b) 0.06

tribution of *M. fallax*, *E. gariene* and *P. littoralis* could be removed altogether by a reduction to presence/absence data, since they are present in all samples, instead placing strong emphasis on species like *S. elizabethae* and *Amphiascoides limicola* that are largely present in one creek and not the other. However, this also upweights the relative contribution of rare species, likely to be randomly distributed across the creeks (e.g. *E. propinquum*), which will prove unhelpful in general.

Although it is convenient in Table 3 to consider only a subset of 2 creeks, in order to illustrate the workings of the method on real assemblages without needing to reproduce a complete data matrix, a fairer comparison of the performance of dispersion weighting in this example would use all 27 samples from the 5 creeks. Table 4 contrasts the global ANOSIM *R* statistics, for a test of no differences between any of the creeks, under all 8 combinations of dispersion weighting or no weighting, with no transform, root, 4th-root and presence/absence transformations. This is based, as usual,

Table 4. Fal copepod study for all creeks, showing the effect of dispersion weighting (or not), followed by various transformations of counts prior to calculating Bray-Curtis dissimilarities. Tabulated is the global ANOSIM *R* statistic, larger values indicating greater assemblage differences between creeks ($p < 0.001$ in all cases)

	No transformation	Square root	4th-root	Presence/ Absence
Unweighted	0.465	0.606	0.662	0.640
Dispersion weighted	0.488	0.636	0.675	0.640

on Bray-Curtis dissimilarities computed after any weighting and then transformation has taken place (note, it does not make sense to transform prior to dispersion weighting because the nature of the data as discrete counts is destroyed by transformation). This table shows that the highest value of *R* is achieved by a combination that includes dispersion weighting. In fact, such weighting always improves the separation of the creeks, for any given transformation, because it reduces an element of the 'noise'; however, the effect cannot be claimed to be dramatic. The final column shows that if all the emphasis is ultimately going to be placed on the presence/absence pattern of species occurrence, rather than any quantitative measure, then dispersion weighting is, of course, irrelevant.

Australian mangrove benthic macrofauna

An example showing a more extreme degree of over-dispersion concerns assemblages of 35 macrofaunal taxa from mangrove habitats at 3 locations, denoted A, B and C, approximately 500 m apart, in Bicentennial Park, New South Wales, Australia. From each location 8 replicate quadrats are sampled, and organisms sorted to a mixed taxonomic resolution because of a lack of local taxonomic expertise for many of the groups. Nevertheless, mixed resolution captures many of the important differences in this fauna among habitats (Chapman & Tolhurst 2004) and at different spatial scales (Chapman 1998).

Many of the taxa are highly variable across replicates; for example, at Location C, Oligochaeta counts range from 6 to 509, Sabellidae from 0 to 194 and Nematoda from 0 to 723, and all are sparser at the other sites. Replicate counts for insect larvae vary from 0 to 997 at Location B, from 5 to 131 at A and 0 to 70 at C. This inevitably leads to high values of the average dispersion index ($\bar{D}$ = 100.0, 164.8, 232.1 and 261.1, respectively, for Oligochaeta, Sabellidae, Nematoda and insect larvae). Indeed, the range of $\bar{D}$ values

obtained across taxa could be a useful structural characteristic of an assemblage so defined, and is best illustrated by a cumulative frequency plot (Fig. 3). The continuous line to the right of Fig. 3 records the $\bar{D}$ values for this mangrove data, showing the strong asymmetry in the distribution, with median $\bar{D}$ of 8.0 but maximum of 261. This is contrasted with successively less extreme profiles, namely for the Fal estuary copepods (long dashed line, median $\bar{D}$ of 4.7), and further examples discussed later, where clumping of individuals is much less pronounced (median $\bar{D}$ of 1.5 and 1).

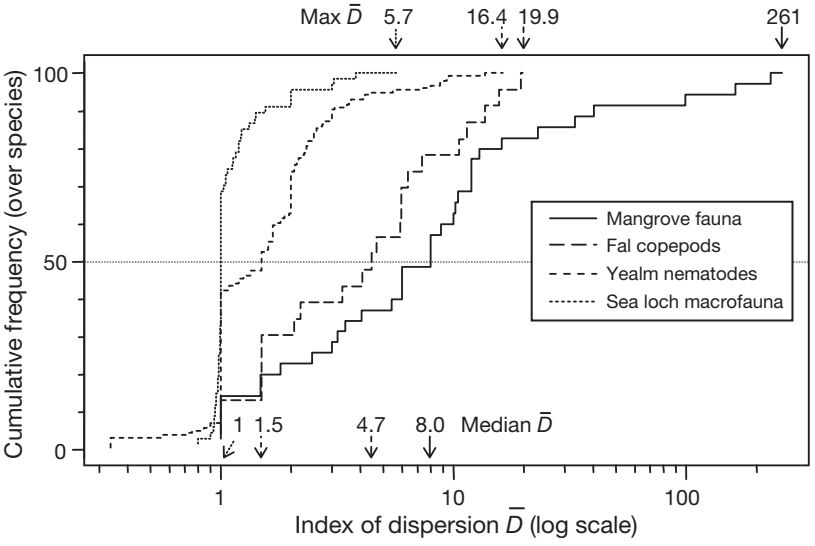

Fig. 3. Cumulative frequency plot ('sample distribution function') for average dispersion indices $\bar{D}$ calculated over all species, in each of 4 real data sets analysed. $\bar{D}$ scale (*x*-axis) is plotted logarithmically, and the *y*-axis (numbers of species with dispersion less than the specified *x* value) is plotted as a percentage of total numbers in the study (35, 23, 156 and 67 species, respectively)

The formal generalised Poisson model represented by Fig. 1 might be thought less convincing for the mangrove data set, where taxa consist of aggregated species counts. Pragmatically, however, the dispersion weighting procedure will still achieve the goal of downweighting counts which are highly erratic across replicates. Thus, Fig. 4 compares the MDS plot from Bray-Curtis similarities based on the standard unweighted, untransformed data matrix with that from dispersion-weighted counts (also untransformed). As with the previous example, the effect of the weighting is apparently to reduce the multivariate variation within Location C, relative to the overall dissimilarities between C, and A and B. The stress levels of the 2D MDS plots are not small in this case, so an ANOSIM test becomes a better arbiter. The *R* statistics for comparing Location C with combined Locations A and B (A & B do not differ) are $R_{\text{C vs. A\&B}} = 0.42$ for Fig. 4a and $R_{\text{C vs. A\&B}} = 0.50$ for Fig. 4b. Both are highly significant ($p < 0.001$ and $p < 0.0001$, respectively), but the latter indicates a greater difference between the assemblages.

More interesting in this case is a similarity percentage analysis (SIMPER, Clarke 1993), Table 5, which breaks down the average dissimilarity between Locations A and B and Location C into its contribution from each species. Using the original data, nematodes and insect larvae are seen to contribute heavily (Table 5a), because for some pairs of replicates from different locations these taxa have very different counts. However, their abundances are not at all consistent among replicates within a group, the wide range of values being remarked on earlier. These counts are therefore heavily compressed by the dispersion weighting divisors of $\bar{D}$ = 232 and 261, whereas oligochaetes are somewhat less so, and the relatively large but slightly less erratic abundances of the mussel *Xenostrobus*

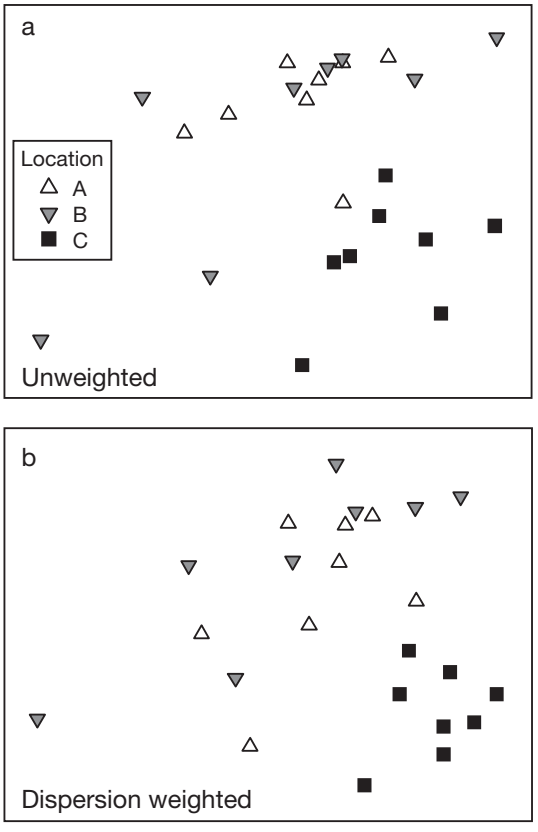

Fig. 4. Australian mangrove benthic macrofauna. MDS ordinations based on Bray-Curtis similarities from 8 untransformed replicate counts in each of 3 locations, using: (a) unweighted data (standard analysis); (b) dispersion weighted matrix. Stress values: (a) 0.11; (b) 0.18

Table 5. Australian mangrove macrofauna, similarity percentage analyses (SIMPER, Clarke 1993). (a) Unweighted, untransformed data: mean counts over all replicates at Locations A and B and at Location C (Columns 1 and 2), for the species making the greatest contribution $\bar{\delta}_i$ (Column 3) between the 2 groups, summing to 75% (Column 5) of the average Bray-Curtis dissimilarity $\bar{\delta}$ = 89.8. Also shown is the ratio of that contribution to its SD (Column 4), low values indicating relative inconsistency in contribution. (b) SIMPER analysis, as in (a) but for dispersion weighted counts, with average dissimilarity between the 2 groups of $\bar{\delta}$ = 96.2

	A & B mean abundance	C mean abundance	Dissimilarity contrib. $\bar{\delta}_i$	$\bar{\delta}_i / SD(\delta_i)$	Cumulative % of $\bar{\delta}$
(a) Unweighted					
Oligochaeta	7.31	175.75	32.13	1.39	35.78
Insect larvae	101.75	25.25	16.74	0.82	54.42
Xenostrobus securis	0.00	40.75	13.17	0.93	69.09
Nematoda	1.06	95.00	6.38	0.49	76.19
(b) Dispersion weighted					
Oligochaeta	0.07	1.76	12.73	1.20	13.23
Xenostrobus securis	0.00	1.22	10.89	1.17	24.55
Nereididae	0.00	1.11	8.95	1.19	33.84
Isopod sp. 1	0.00	0.76	6.10	0.87	40.18
Midges	0.50	0.05	5.29	0.54	45.68
Ostracoda	0.15	0.15	4.86	0.46	50.72
Tanaidacea	0.00	0.67	4.68	1.44	55.58
Arthritica helmsii	0.00	0.59	4.53	0.77	60.29
Amphipod sp. 1	0.00	0.38	3.99	0.51	64.44
Laternula sp.	0.00	0.63	3.51	0.75	68.08
Nephthyidae	0.18	0.27	3.25	0.59	71.46
Insect larvae	0.39	0.10	2.90	0.56	74.47
Salinator fragilis	0.25	0.00	2.86	0.36	77.45

securis are downweighted less severely ($\bar{D}$ = 33). Thus, the SIMPER results for the dispersion weighted case (Table 5b) promote *X. securis* to slightly greater prominence, but insect larvae are demoted to the point where they scarcely feature in this comparison and the Nematoda disappear altogether. To judge by the mean nematode abundance in the 2 location groups (Columns 1 and 2 of Table 5a) this would appear harsh, but it becomes entirely understandable, and the pitfalls of simple averaging become only too apparent, when it is known that the actual counts in the 8 replicates at Location C were 0, 0, 0, 0, 0, 3, 34, 723! Compare this with counts of 0, 4, 5, 6, 16, 17, 27, 40 for the Nereididae, also for Location C replicates, contrasted with complete absence at Locations A and B, and it is clear why this taxon has moved up to third ranking contribution in Table 5b. Note also the larger set of 13 taxa, across a wide taxonomic range, which now contribute to 75% of the average dissimilarity (of 96.2%) between Locations A and B and Location C, whereas in the unweighted analyses only 4 taxa contributed to the 75% threshold and the average dissimilarity was lower (at 89.8%). The erratic contribution from 2 of these taxa (Nematoda and insect larvae) is, of course, the main reason why the dissimilarity between the groups increases after dispersion weighting.

Nematodes in Yealm sea-grass beds

The third data set is an example of a 2-way layout, with a small number of replicates but a large number of species. It examines the differences between free-living nematode compositions in soft sediments, inside and outside seagrass beds (Factor A: 'treatment', with levels in/out). Small cores (internal diameter 19 mm) were taken from 4 different seagrass patches off Cellars Beach in the Yealm estuary, on the southwest coast of England (Factor B: 'blocks', with levels 1 to 4, crossed with Factor A), with a total of 156 nematode species identified and counted. For each block, samples were taken at the corners of a 2 × 2 m square, with corners 1 and 2 in the patch, and 3 and 4 outside the patch; there are therefore only 2 replicates for each treatment × block combination. For the average index of dispersion calculation, this 2-way crossed structure is flattened to a 1-way layout, so the dispersion D_i is computed separately over the 2 replicates for each treatment × block combination. The resulting 8 indices are averaged to obtain $\bar{D}$, using a simple average in this case because replication is balanced. Although the availability of only 2 replicates in each group would seem (and is) a minimal requirement, it is nonetheless viable. $\bar{D}$ is estimated with 8 degrees of freedom, 1 for each treatment × block combination (somewhat analogously to the 8 degrees of freedom for estimating residual error in a univariate ANOVA with this 2-factor design). The test is, in any case, a permutation procedure which is valid for any sample size, so the worst that can happen in situations of relatively low power is that the H_0 is over-protected. That is, D may be set to 1 in some cases where more replicates would have demonstrated some clumping, and a certain degree of downweighting applied. This is no more than a potential loss of efficiency, not in any sense an introduction of bias, since failing to downweight any species at all leads to a perfectly viable analysis (the standard one).

The distribution of $\bar{D}$ values obtained for the 156 species is one of the sample distribution functions shown in Fig. 3 (short-dashed line). Note that for these profiles the actual $\bar{D}$ estimates are used and not the divisors from the dispersion weighting, some of which may be reset to 1 as a result of the hypothesis tests. Nonetheless, several of the estimated $\bar{D}$ values are exactly 1 in this case (and in the final example), corresponding to species represented only by a singleton in one or more of the groups. Use of the calculated $\bar{D}$ values should allow comparison of these cumulative profiles across studies with differing degrees of freedom, without introducing significant bias, since sample means and variances are unbiased for true means and variances whatever the number of replicates employed. The distribution of $\bar{D}$ in this case shows that many species do not depart far from unclustered spatial distributions, the median $\bar{D}$ being only 1.5, and there is only one value in double figures ($\bar{D}$ = 16.4 for *Richtersia inequalis*). In part, this is a reflection of the low densities of all species in what are relatively small-diameter cores.

More importantly, the dispersion weighting does make a real difference to the outcome of the analysis in this case. The 2-way crossed ANOSIM (Clarke 1993) gives $R = 0.50$ ($p < 0.037$) for a test of the hypothesis of 'no treatment effect' (removing block effects) when dispersion weighting is employed, with no subsequent transformation. This compares with $R = 0.31$ ($p < 0.148$) for unweighted, untransformed data, both analyses being based on Bray-Curtis coefficients as usual. Even more dramatically, the hypothesis test for 'no block effects' (removing treatment effects) gives $R = 0.32$ ($p < 0.034$) for dispersion weighted data but only $R = 0.05$ ($p < 0.352$) in the unweighted case. The latter improves a little with transformation, giving $R = 0.20$ ($p < 0.077$) for root-transformed counts and $R = 0.19$ ($p < 0.085$) for 4th-root transformed counts, but neither is significant at the 5% level. It is clear from the corresponding MDS ordinations (Fig. 5) that the dispersion weighting succeeds in displaying an additive signal of block differences (left to right) and treatment differences (bottom to top) more successfully than the standard analysis under any transformation.

Sea-loch macrobenthos in a replication study

Finally, Gage & Coghill (1977) collected a set of 256 contiguous samples of soft-sediment macrobenthos along a single transect at Site C-12 in Loch Creran, Scotland. This is included here not as an example of the application of dispersion weighting to improve a multivariate analysis, but as a further cumulative distribution curve for the dispersion index, one exhibiting

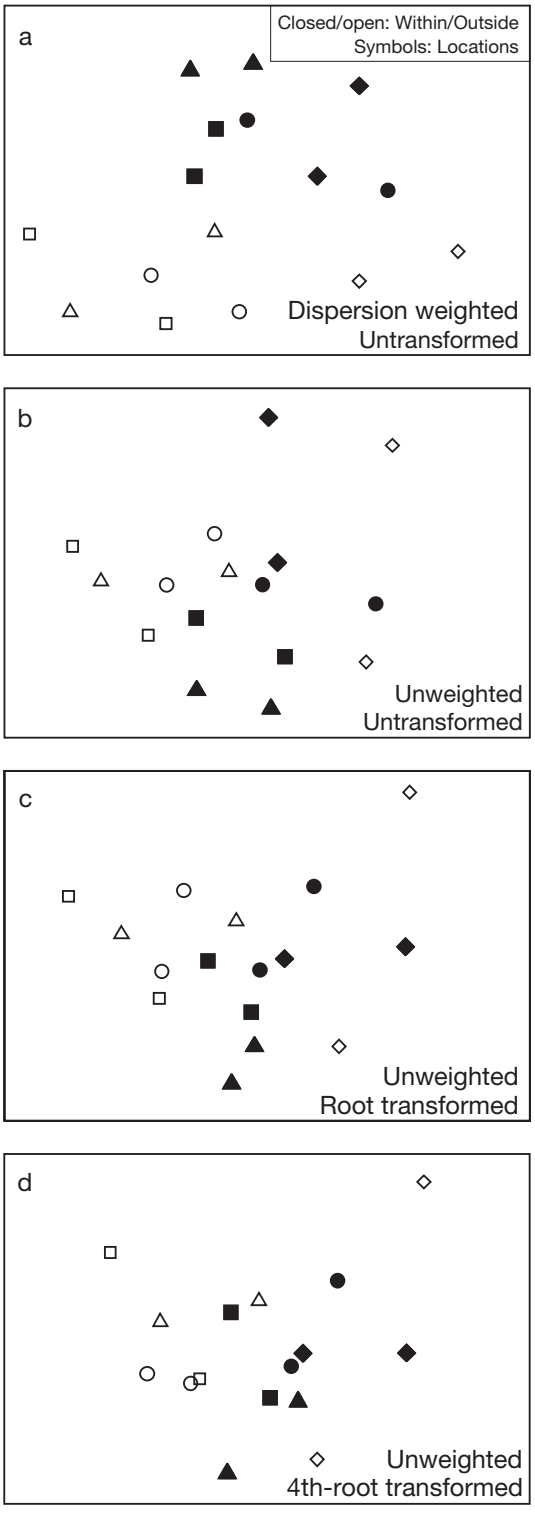

Fig. 5. Yealm estuary nematodes. MDS ordinations of 2 replicate samples from each of 8 conditions: within/outside seagrass patches (closed/open symbols) at 4 locations (different symbol types). Plots are based on Bray-Curtis dissimilarities computed after: (a) dispersion weighting; (b–d) no weighting and with various transformations. Stress values: (a) 0.17; (b–d) 0.16

even less evidence of clumping. There is no group structure here, so all 256 samples are regarded as replicates of a single assemblage, and a single index D is computed for each species. Its sample distribution function is plotted in Fig. 3 (dotted line), and the median dispersion is 1, with 90% of the species having $D < 1.5$ and the largest D being only 5.7. The contrast with the mangrove fauna (continuous line), for example, could not be more marked.

DISCUSSION

Validity of approach

It is argued in this study that weighting each species by its index of dispersion, calculated from replicates, stabilises the quality of information from each species that is input to multivariate routines. Any objection that this is in some way 'interfering with the data' cannot logically be sustained: in the final analysis, all that has happened is that the values for each species have been divided through by a constant. For any particular species, there can have been no change to the relative differences between mean responses for each of the groups. All that has changed is the relative attention given to each species in constructing a similarity coefficient drawn from all p of them. In fact, dispersion weighting will have no effect at all when using a resemblance measure with an explicit species standardisation, designed to equalise the contributions from all species. An example would be the Gower coefficient (Gower 1971, Legendre & Legendre 1998) that divides each count by the range or, in practice, the maximum abundance for that species, so that all data are scaled over the same interval of 0 to 1 (or 0 to 100). However, such coefficients tend to commit the opposite misdemeanour to that which this study tries to correct. Instead of overweighting high abundance species, whose numbers can sometimes be very erratic, they overweight rare species, the range of whose absolute abundances can be very small. Many of the rare species could be more or less random in their occurrence and will tend to add further noise, not signal, when upweighted to have the same range as common species. Coefficients with automatic variable standardisation or normalisation (e.g. normalised Euclidean distance) should be seen as a last resort, in cases where there is no convincing way of balancing the contributions from each of the variables.

Where the latter are not species but environmental variables, a normalised version of Euclidean or Manhattan distance is almost inevitable because of the likely mixture of measurement scales. Here, however, we start with a common count scale and an objective means of deciding which species have counts which are inherently more reliable, so to use a coefficient which removes that knowledge is likely to prove inefficient. Sample standardisation raises a different set of problems: clearly it is not valid to turn each sample into relative compositional data (with all samples adding to 100% across species) prior to dispersion weighting, since the requirement for identically distributed counts over replicates has been compromised. The requirement for uniform-sized replicates was made clear in the 'Introduction'. After weighting, sample standardisation may be justifiable, at least in so far as distance measures that automatically build in such standardisation—Hellinger distance, chi-squared distance etc. (Legendre & Legendre 1998)—are appropriate to the applied context. Indeed, although untested at the moment, it may be that dispersion weighting brings the data closer into line with the spatial Poisson model (multinomial frequencies) that motivates the use of chi-squared distance, since we are more nearly counting 'clumps' now rather than individuals. However, the double standardisation (by both samples and variables) that is inherent in chi-squared distance is likely to remain problematic, because it will continue to overemphasise rare species (Clarke et al. 2006).

Returning to the issue of the degree of 'interference' with the data inherent in dispersion weighting, it cannot be argued that the procedure has any element of subtle selection bias, in the sense of downweighting species that did not prove to be helpful discriminants between groups, and upweighting those whose mean responses did change relatively markedly. This criticism could certainly be levelled at the constrained ordination methods (Canonical Correspondence Analysis, Ter Braak 1986; Canonical Analysis of Principal Co-ordinates, Anderson & Willis 2003), which have to be rather careful to establish that genuine group differences exist before they seek ordination axes that will display those differences to best advantage. They are, in effect, searching to upweight species which show large group differences, and therefore have to be careful that they do not magnify random 'noise' and misrepresent it as 'signal'. This is in sharp contrast with dispersion weighting, which ignores mean differences between groups by concentrating only on the dispersion properties of the replicates within each group, in its construction of the divisor $\bar{D}$. It could perfectly well end up giving greatest weight to a ubiquitous species that had high and consistent counts in all replicates within groups, and showed absolutely no difference, either, in mean count between groups. The fact that (on the limited evidence to date) dispersion weighting does seem to improve significance in ANOSIM tests of group differences is a reflection not of any selection bias, but of the modest increases in power that can

result from down-playing counts with high baseline variability. This is one of the 2 classic ways in univariate statistics of improving power: either increase the number of replicates (which will certainly improve the number of permutations, and hence the power of ANOSIM tests), or reduce the 'error' variance of a replicate. The analogy of the latter is what the dispersion weighting procedure attempts to do.

Generality of application

Although dispersion-weighting has been formally justified in the context of a generalised Poisson model for species counts, with the organisms regarded as points in space (or time), and although this semi-parametric model already has good flexibility to cover a range of practical situations, it is clear that this is not the only context in which weighting by the index of dispersion makes sense. Other structural models could lead to the same conclusion. For example, the counts could come from randomly distributed or fixed grid points on a photographic or remotely-sensed image of a hard-bottom community (or grid points physically placed on the substrate itself). For discrete coral stands of different species (an example of a multivariate 'marked point process', Diggle 1983), the counts for a single image would consist of the number of grid points falling on each species. Certain models, which involve a stationary random process for the coral locations, and distances between sampling points that are large in relation to the size of an individual coral, could lead to an approximate Poisson distribution of counts for each species. 'Over-dispersed' counts would result from individual corals being larger than the grain of the sampling points, so that effectively the same coral is 'captured' several times by the grid points in one image.

This could, of course, lead to high variability in counts from replicate to replicate, with a single large coral contributing nearly all the counts in one image but contributing nothing at all in the next one. Such a species should logically contribute less to defining the constitution of its group of samples than another coral that has the same total area cover but whose individuals are smaller than the sampling grain size, with locations more randomly dispersed, giving more consistent counts in replicates (lower variance for a given mean). Paralleling the generalised Poisson formulation defined earlier, in which the mean number of centres varies across the groups but the 'clumping' structure remains the same, one could envisage a model structure here in which the average numbers of a particular coral species change from group to group but the average size of an individual coral does not. The degree of over-dispersion of counts would then remain the same across groups: the clumping parameter now reflects the number of sampling grid points that intercept the same individual coral, and which are therefore not providing 'independent arrivals'; the analogy with our earlier formulation of a clustered point process is clear. It is outside the scope of this study to formalise this model fully and discuss its implications (and limitations), but the practical consequence is the likely validity and effectiveness of downweighting such grid counts by the dispersion index $\bar{D}$ for each species, calculated from replicate images within a group and then averaged across groups as usual.

Combination of dispersion indices across groups

Needing further discussion is the specification in the generalised Poisson model that, for each species, $\bar{D}$ should be calculated as the unweighted average (given balanced replication) of the separate dispersion indices calculated from each group (Eq. 2). There are actually 2 levels of assumption here.

Firstly it is assumed that, if the model is correct, an unweighted average is an efficient way of combining the estimates $\{D_i\}$ from each group. This is largely sustainable, although it could possibly be improved upon. It follows from the assumption that only the number of 'centres' changes across the groups, and not the average cluster size for a given centre, that the dispersion estimates D_i are estimating the same parameter (they are unbiased). It is also true that they have the same variance in the H_0 case, i.e. when the cluster size is 1, irrespective of the density of centres in a replicate. This is what formally justifies the use of the unweighted average in the test statistic, Eq. (3). Moreover, it can be shown (with more difficulty) that when the true cluster size is >1, and the density of centres high enough, then the indices D_i from each group have equal variance. This again justifies the use of an unweighted average. However, it must be true that for a group in which the density of centres is not large but vanishingly small, so that all the replicate counts for that species are zero, there can be no information about the real cluster size. An optimal average index $\bar{D}$ would probably therefore downweight dispersion estimates from the sparser groups, even though it is not clear at present exactly how that should best be carried out. Note that the calculations of this study employ the weighting given in Eq. (2), i.e. equal weights in the balanced case, except where the replicates of a group are entirely blank for that species, in which case D_i is undefined and given zero weighting.

More fundamentally, the model specifies that the true dispersion indices (and cluster size distributions)

are the same for each group i, so that an averaged D is applicable. This is an explicit assumption and it may not be a very good one in some cases. The effect of an environmental impact on a species might not just be to decrease (or increase) its density but also to change its propensity to form clusters of individuals. For example, instead of a tendency to catch 10 ind. every time 1 is caught, organisms may now arrive singly (or vice-versa). The formal basis for downweighting that species breaks down. This does not mean, however, that when the D_i estimates from each group are variable it becomes undesirable to downweight a species by a calculated average $\bar{D}$. Suppose that the dispersion estimates from 5 groups (with equal replication) are D_i = 10, 1, 20, 5, 14, giving an average of $\bar{D}$ = 10. Downweighting by a factor of 10 seems eminently sensible. It recognises that information about differences between groups from that species is not particularly reliable. To judge by Groups 2 and 4, the species would perhaps be given less emphasis than it merits, whereas Groups 3 and 5 suggest that it be paid even less attention, but on balance it would be right to downweight it in relation to a species which is not overdispersed at all, and an average divisor of 10 will serve the purpose. The alternatives, after all, are either to not selectively downweight anything, however erratic (standard Bray-Curtis), or to completely homogenise the range of variation over groups for all species (normalisation, Gower coefficient etc.). A selective downweighting, for which there is some objective basis and which cannot introduce bias, has to perform at least as well as either of these end-points. What absolutely must not be done, of course, is to downweight the species counts in Groups 1, 3 and 5, and leave them unchanged in Groups 2 and 4. This would simply create differences in mean value between groups that were not there before, and totally compromise the multivariate analysis. It is axiomatic that any differential downweighting of one species in relation to another must be carried out in exactly the same way across all samples, otherwise any susbequent analysis of group differences becomes meaningless.

Upweighting of under-dispersed species

Although the emphasis has been on downweighting certain (clumped) species whilst leaving others (randomly distributed) unaltered, there is no logical reason why a third group of species should not be upweighted. These would be taxa that are distributed with a more uniform spacing than would be expected from random scattering, such as might result from territorial behaviour, for example. The counts in fixed-size quadrats are again not Poisson and are characterised by having an index of dispersion $D < 1$ (variance < mean). They cannot be modelled by a generalised Poisson distribution (this is not completely obvious from Eq. A4 of the Appendix, but it is not difficult to prove that such models always have $D \geq 1$). It is clear, nonetheless, that one could test for $D = 1$ against the alternative $D < 1$ by the same sort of procedure as above. The chi-squared form of statistic (Eq. 3) would now be subject to a 2-sided rather than 1-sided permutation test, and if the H_0 of $D = 1$ is rejected, the species counts could be divided by the observed average across the groups, $\bar{D}$, as usual; $\bar{D} < 1$ in the 'under-dispersed' case, so this would be an upweighting of unifomly-distributed species at the expense of those that are randomly distributed. In practice, this possibility will arise rarely and the adjustment will make little difference, if it was ever justifiable. Certainly for the range of marine communities examined earlier, it is uncommon for the calculated $\bar{D}$ to be much less than 1 and, on a 2-tailed test, it is never significantly so. For the Yealm nematode study, 10 of the 156 species recorded a $\bar{D}$ below 1, with 4 of those as low as 0.33; however, bearing in mind that there are only 2 replicates in each D_i calculation in that case, such figures are readily attainable by chance. (They correspond in this instance to counts of 1 and 2 for the 2 replicates from one of the treatment × block combinations, with no occurrences of that species elsewhere.) It is clear that, in practice, there is complete asymmetry with regard to the occurrence of over- and under-dispersion. Large, sometimes very large, dispersion indices are common, as shown by the cumulative distribution curves of Fig. 3, demonstrating virtually no symmetry around the value $D = 1$ even though plotted on a severe 10-cycle log scale for D. Divisors as large as 100, 164, 232 and 261 are found for the mangrove study, leading to sizeable downweighting of those species. Reciprocally small dispersion values (1/100, 1/164 etc.) are simply never found, with D as low as 0.5 being uncommon in practice. Any upweighting arising from under-dispersion would thus be quite negligible and this complication can safely be ignored.

Other reasons for downweighting species

There can be more than one reason why it might be desirable to carry out differential downweighting of the contributions of some species in relation to others. For example, in a context of volunteer surveys of coral reefs in Belize, Mumby at al. (1996) introduced the weighted Bray-Curtis measure:

$$D_{12}^{WtB-C} = 100 \cdot \frac{\sum_I w_I |y_{I1} - y_{I2}|}{\sum_I w_I (y_{I1} - y_{I2})} \quad (4)$$

which is exactly the form used above (with $w_l \propto 1/\bar{D}_l$) except that in the Mumby study the weights $\{w_l\}$ were determined on the basis of how reliably each species was identified by the volunteer divers. Species often misidentified were downweighted, potentially leading to a more robust analysis. One could also envisage situations in which species were given *a priori* weights according to vulnerability, or to economic or charismatic status, or perhaps to body size or some other estimate of functional importance in the ecosystem, so that the multivariate analysis was directed more towards identifying change that had the greatest consequences for these properties. In this study we have chosen to concentrate on an objective statistical reason for differential weighting, namely the intrinsic reliability of the information provided by each species, but any of these other types of downweighting could be legitimate in particular contexts. Indeed they could even be compounded, provided the temptation of subjective selection bias is avoided ('these 2 species were heavily downweighted because that produced the desired outcome'!)

Another obvious role for differential weighting is in combining fauna whose density has to be assessed in different ways. It is typical in analysis of epibiota of hard substrates in marine contexts to collect a mixture of individual counts of motile organisms with area cover of sessile or colonial biota. This is always troublesome, because values for the different species have to be put on some sort of common scale of measurement before a coefficient such as Bray-Curtis similarity is meaningful. Weighting motile species by the area cover of a typical organism is one possibility, or converting both area cover and single counts to an approximate biomass (or even some approximation to production) is another weighting option. Such reasoning even leads one to query whether the common scale of abundances implicit in standard assemblage matrices really are 'common' across species. Species with high density are very often small-bodied and would therefore get a very different weight in a biomass similarity than an abundance-based coefficient. Fortunately, experience suggests that there is a good deal of robustness in the outcomes of multivariate ordination or ANOSIM tests when manipulating the weighting given to each species. Warwick (1993) commented on this when contrasting specific biomass-based and abundance-based analyses: ordinations were very similar yet the similarity percentages routine (Clarke 1993), which breaks down dissimilarities between groups into their contributions from each species, showed that entirely different species were responsible for the ordination pattern in the two cases. The explanation is that, typically, all parts of an assemblage are responding to a common set of environmental drivers in a common way, so that the same among-sample pattern is obtained for species with higher density as for species with higher biomass. Clarke & Warwick (1998) show the high level of structural redundancy in typical soft-sediment macrobenthic matrices by demonstrating that sample ordinations virtually identical to that from the full community can be generated from 4 or 5 mutually exclusive subsets of the species. They go on to argue that this repetition is part of the reason for the success of multivariate methods in capturing subtle underlying changes in environmental conditions or anthropogenic impacts: it acts as a buffer to the vagaries of sampling fluctuations that tend to dominate any attempt to monitor single populations rather than whole assemblages. However, the structural redundancy inherent in some species × samples arrays does have an obvious corollary: adjusting the weights given to each taxon, as this study does, is not often likely to have dramatic practical consequences for the final analysis.

In conclusion

In some ways, the advantages of the dispersion-weighting approach introduced in this paper are more conceptually satisfying than practically crucial. Exploiting a plausible structure for typical species abundance matrices to downweight erratic species, followed if necessary by mild transformation to balance the commoner and rarer components of the assemblages, seems conceptually preferable to using the 'blunt instrument' of a severe transformation to attempt to solve both problems simultaneously. The fact is, though, that the latter course is often perfectly effective, pragmatically, and has the advantage of being applicable when there are no replicates or the data are not counts (or analogous to counts). The observed robustness of multivariate analyses, discussed above, to major weight changes in the way the species contributions are compounded, means that we should not expect a dispersion-weighting approach to radically alter ordination and test results. Indeed, the improvements shown in the real examples of this paper are sometimes small and subtle, but they are none the less valuable for that. Given the scale of spatio-temporal variability of biotic measurements in some environments, and the difficulties of constructing powerful sampling regimes, any legitimate and objective means to reduce 'noise', without in any way compromising the 'signal', should be welcome.

Acknowledgements. 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); the UK

Department for Environment, Food and Rural Affairs (DEFRA) through projects AE1137, CDEP 84/5/295 and ME3109; the Australian Research Council through its Special Research Centres Programme; and the University of Sydney. We are grateful to the late J. Gage for making paper copies of the contiguous-core data available to us, and thank S. Dashfield for entering them into spreadsheets. We are indebted to R. Gorley for his work on the PRIMER v6 code. This work was instigated when K.R.C. held a visiting professorship at the Centre for Research on Ecological Impacts of Coastal Cities, University of Sydney, and completed under his current position as an honorary fellow of both the Plymouth Marine Laboratory and the Marine Biological Association of the UK.

LITERATURE CITED

Anderson MJ, Willis TJ (2003) Canonical analysis of principal co-ordinates: an ecologically meaningful approach for constrained ordination. Ecology 84:511–525

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

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

Chapman MG, Tolhurst TJ (2004) The relationship between invertebrate assemblages and bio-dependant properties of sediment in urbanized temperate mangrove forests. J Exp Mar Biol Ecol 304:51–73

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

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

Clarke KR, Warwick RM (1998) Quantifying structural redundancy in ecological communities. Oecologia 113:278–289

Clarke KR, Warwick RM (2001) Change in marine communities: an approach to statistical analysis and interpretation, 2nd edn, PRIMER-E, Plymouth

Clarke KR, Somerfield PJ, Chapman MG (2006) On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis measure for denuded assemblages. J Exp Mar Biol Ecol 330: 55–80

Cochran WG (1954) Some methods for strengthening the common χ^2 tests. Biometrics 10:417–451

Diggle PJ (1983) Statistical analysis of spatial point patterns. Academic Press, London

Douglas JB (1979) Analysis with standard contagious distributions. International Co-operative Publishing House, Fairland

Elliott JM (1971) Some methods for the statistical analysis of samples of benthic invertebrates. Freshwater Biological Association Scientific Publication No. 25. Freshwater Biological Association, Ambleside

Fisher RA, Thornton HG, Mackenzie WA (1922) The accuracy of the plating method of estimating the density of bacterial populations, with particular reference to the use of Thornton's agar medium with soil samples. Ann Appl Bot 9: 325–359

Fisher RA, Corbet AS, Williams CB (1943) The relation between the number of species and the number of individuals. J Anim Ecol 12:42

Gage JD, Coghill GG (1977) Studies on the dispersion patterns of Scottish sea loch benthos from contiguous core transects. In: Coull BC (ed) Ecology of marine benthos. University of South Carolina Press, Columbia, SC p 319–337

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27:857–871

Greig-Smith P (1952) The use of random and contiguous quadrats in the study of the structure of plant communities. Ann Bot 16:293–316

Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577–586

Kendall MG, Stuart A (1963) The advanced theory of statistics, Vol 1. Griffin, London

Kruskal JB (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29:1–27

Legendre P, Legendre L (1998) Numerical ecology, 2nd edn. Elsevier, Amsterdam

Manly BJF (1997) Randomization, bootstrap and Monte Carlo methods in biology, 2nd edn. Chapman & Hall, London

Mumby PJ, Clarke KR, Harborne AR (1996) Weighting species abundance estimates for marine resources assessment. Aquat Conserv: Mar Freshw Ecosyst 6:115–120

Neyman J (1939) On a new class of contagious distributions, applicable in entomology and bacteriology. Ann Math Stat 10:35–57

Quenouille MH (1949) A relation between the logarithmic Poisson and negative binomial series. Biometrics 5:162–164

Sanders HL (1968) Marine benthic diversity: a comparative study. Am Nat 102:243–282

Silvey SD (1975) Statistical inference. Chapman & Hall, London

Somerfield PJ, Gee JM, Warwick RM (1994) Soft sediment meiofaunal community structure in relation to a long-term heavy metal gradient in the Fal estuary system. Mar Ecol Prog Ser 105:79–88

Ter Braak CFJ (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67:1167–1179

Warwick RM (1993) Environmental impact studies on marine communities: pragmatical considerations. Aust J Ecol 18: 63–80

Appendix 1. Theoretical properties of dispersion weighting

Dispersion for generalised Poisson model. For the generalised Poisson distribution, the sample that is replicate j ($j = 1, \ldots, n_i$) of group i ($i = 1, \ldots, g$) contains C_{ij} 'centres', with Z_{ijk} organisms at centre C_{ij}, giving count:

$$X_{ij} = \sum_{k=1}^{C_{ij}} Z_{ijk} \quad \text{(A1)}$$

where $C_{ij} \sim \text{Po}(\gamma_i)$, $E(Z_{ijk}) = \mu$, $\text{var}(Z_{ijk}) = \sigma^2$, and the $\{C_{ij}\}$ and $\{Z_{ijk}\}$ are all independent. Using the conditional expectation and variance formulae:

$$E(X_{ij}) = E_{C_{ij}}[E(X_{ij}|C_{ij})] = E[C_{ij}E(Z)] = \gamma_i \mu \quad \text{(A2)}$$

$$\begin{aligned} \text{var}(X_{ij}) &= E_{C_{ij}}[\text{var}(X_{ij}|C_{ij})] + \text{var}_{C_{ij}}[E(X_{ij}|C_{ij})] \\ &= E_{C_{ij}}[C_{ij}\,\text{var}(Z)] + \text{var}_{C_{ij}}[C_{ij}E(Z)] = \gamma_i\sigma^2 + \gamma_i\,\mu^2 \end{aligned} \quad \text{(A3)}$$

$$D = \text{var}(X_{ij}) / E(X_{ij}) = (\mu^2 + \sigma^2)/\mu, \quad \text{for all } i \quad \text{(A4)}$$

Large sample test for $D = 1$. Within group i, the calculated dispersion index is:

$$D_i = \frac{1}{(n_i - 1)} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2 \Big/ \bar{X}_i \ , \ \text{where } \bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij} \quad \text{(A5)}$$

Under the H_0 of only 1 organism at each centre, $\mu = 1$, $\sigma^2 = 0$, $D = 1$ and $X_{ij} \sim \text{Po}(\gamma_i)$. Conditional on their total $\sum_j X_{ij}$, the $\{X_{ij}; j = 1, \ldots, n_i\}$ are jointly multinomial, and $(n_i - 1)D_i$ is seen to have the form of a chi-squared statistic with 'observed' values $O = X_{ij}$ and 'expected' values $E = \bar{X}_i$. Thus, asymptotically:

$$(n_i - 1)D_i \sim \chi^2_{n_i - 1} \quad \text{(A6)}$$

The $\{D_i\}$ are independent, and estimate the same parameter D for all i. They are combined naturally in Eqs. (2) & (3) of the main text. This is the Wald statistic based on the multinomial likelihoods, asymptotically equivalent to the generalised likelihood ratio statistic (e.g. see Silvey 1975). Under the H_0,

$$X^2 = \sum_{i=1}^{g} (n_i - 1)D_i \sim \chi^2_{\sum_i (n_i - 1)} \quad \text{(A7)}$$

because a sum of independent χ^2 distributions has a χ^2 distribution, the degrees of freedom being simply cumulated. This χ^2 distribution is only a 'large-sample' approximation, however, and the main text shows how the same statistic can be tested by permutation in small samples.

Mean and variance of dispersion weighted samples. D is estimated as an average $\bar{D}$ over several groups, from counts in several replicate samples within each group, so its variability can be assumed to be small in relation to that of a single count X_{ij}, for any particular i, j. Thus,

$$X_{ij}^{DW} = X_{ij} / \bar{D} \quad \text{(A8)}$$

is effectively the division of X_{ij} by a constant, and from Eqs. (A2) to (A4):

$$E(X_{ij}^{DW}) \approx \text{E}(X_{ij})/D = \gamma_i \mu / [(\mu^2 + \sigma^2)/\mu] = \gamma_i / [1 + (\sigma/\mu)^2] \quad \text{(A9)}$$

$$\text{var}(X_{ij}^{DW}) \approx \text{var}(X_{ij})/D^2 = \gamma_i (\mu^2 + \sigma^2)/[(\mu^2 + \sigma^2)/\mu]^2 = \gamma_i / [1 + (\sigma/\mu)^2]$$

Of course, if the observed dispersion index $\bar{D}^{DW}$ is calculated for the dispersion weighted X_{ij}^{DW} values defined by Eq. (A8), then $\bar{D}^{DW} \equiv 1$, by definition.

Editorial responsibility: Otto Kinne (Editor-in-Chief), Oldendorf/Luhe, Germany

Submitted: December 12, 2004; Accepted: January 24, 2006
Proofs received from author(s): July 26, 2006