

11.1 What are 'residual' distances?

Rationale

It is sometimes desirable to *remove the effects* of a factor or covariable, and examine 'what is left', i.e., to look at variation in residuals. For example, a factor of primary interest may be statistically significant in a PERMANOVA partitioning of the full model, but its effects may be totally obscured by some other dominant factor(s) or covariate(s) when we look at an ordination of the data.

As seen in [Chapter 10](#), we can construct plots of distances among centroids for main effects and interactions, and these kinds of plots may shed some light on the nature of some minor effects in any given model. Another tool that can help us see effects occurring along minor axes that may not be apparent in unconstrained plots is canonical analysis of principal coordinates ([CAP](#)).

Nevertheless, a dissimilarity-based multivariate analogue to a *residual plot*, from which the variation due to dominant (but perhaps nuisance) factors or covariates has been removed, is desirable ([Anderson \(2017\)](#)).

Construction of residual dissimilarities

Our description here paraphrases [Anderson \(2017\)](#) . Suppose we have a symmetric matrix, $\{\mathbf{D}\}$, of dissimilarities among sampling units, having elements $\{d_{ij}\}$, $i = 1, \dots, N$ and $j = 1, \dots, N$. We have [seen earlier](#) how to derive Gower's matrix, $\{\mathbf{G}\}$ from this (see [Gower \(1966\)](#)), which is also of size $(N \times N)$ and is comprised of elements $\{g_{ij}\}$. The trace (sum of diagonal elements) of matrix $\{\mathbf{G}\}$ is equal to the total sum of squares (total variation) in the space of the chosen dissimilarity measure.[¶]

Let's now suppose that $\{\mathbf{X}\}_r$ is a linear model matrix which contains one or more covariables and/or orthogonal contrasts for one or more factors that we wish to 'remove'. By 'remove', we mean 'on which to condition'. For example, suppose we have a two-factor study design, with factors A and B, and we wish to 'remove' the effects of factor A (e.g., to investigate and hopefully visualise in ordination plots, etc., the effects, if any, of factor B). This amounts to 'taking factor A into account' in our examination of factor B. In such a case, matrix $\{\mathbf{X}\}_r$ will be a full-rank matrix of orthogonal contrasts among the groups in factor A. Now, for any model matrix $\{\mathbf{X}\}_r$, we can construct the usual linear projection ("hat") matrix as:

$$\{\mathbf{H}\}_r = \{\mathbf{X}\}_r [\{\mathbf{X}\}_r' \{\mathbf{X}\}_r]^{-1} \{\mathbf{X}\}_r'$$

We then obtain a 'residualised' Gower matrix directly as

$$\{\mathbf{G}\}^{[R]} = (\{\mathbf{I}\} - \{\mathbf{H}\}_r) \{\mathbf{G}\} (\{\mathbf{I}\} - \{\mathbf{H}\}_r)$$

where $\mathbf{G}^{[R]}$ is an $(N \times N)$ matrix with elements $\{g_{ij}^{[R]}\}$. We note, in passing that $\text{tr}(\mathbf{G}^{[R]})$ is the residual sum of squares after fitting the full model contained in matrix $\mathbf{X}_r$. Once we have this residualised Gower matrix, it is a straightforward back-transformation to arrive at an $(N \times N)$ matrix of *residual distances*, $\mathbf{D}^{[R]}$, with elements:

$$\{d_{ij}^{[R]}\} = \sqrt{g_{ii}^{[R]} - 2g_{ij}^{[R]} + g_{jj}^{[R]}}$$

This residualised distance matrix, $\mathbf{D}^{[R]}$, can then be examined in the usual way *via* ordination methods, such as mMDS, tmMDS or nMDS. The effects of any factors included in $\mathbf{X}_r$ (and/or, the linear relationships in the resemblance space with any covariables in $\mathbf{X}_r$) will be 'removed' from the picture.

The concept of removing a factor

In essence, the removal of a factor is equivalent, conceptually, to superimposing the group centroids onto a common (overall) centroid. For example, let's re-visit the [2-factor crossed study design](#) by [Glasby \(1999\)](#). In this study, the cover of subtidal epifauna was quantified on subtidal settlement plates in response to two crossed factors: Position ('N' = near the seafloor, 'F' = far from the seafloor) and Shade ('S' = shaded, 'C' = procedural control, 'O' = open).

A non-metric MDS plot of these assemblages on the basis of the Bray-Curtis resemblance measure, after a square-root transformation of the raw cover values, is shown in Fig. 11.1 below (this plot is a replica of [Fig. 10.2](#) above).

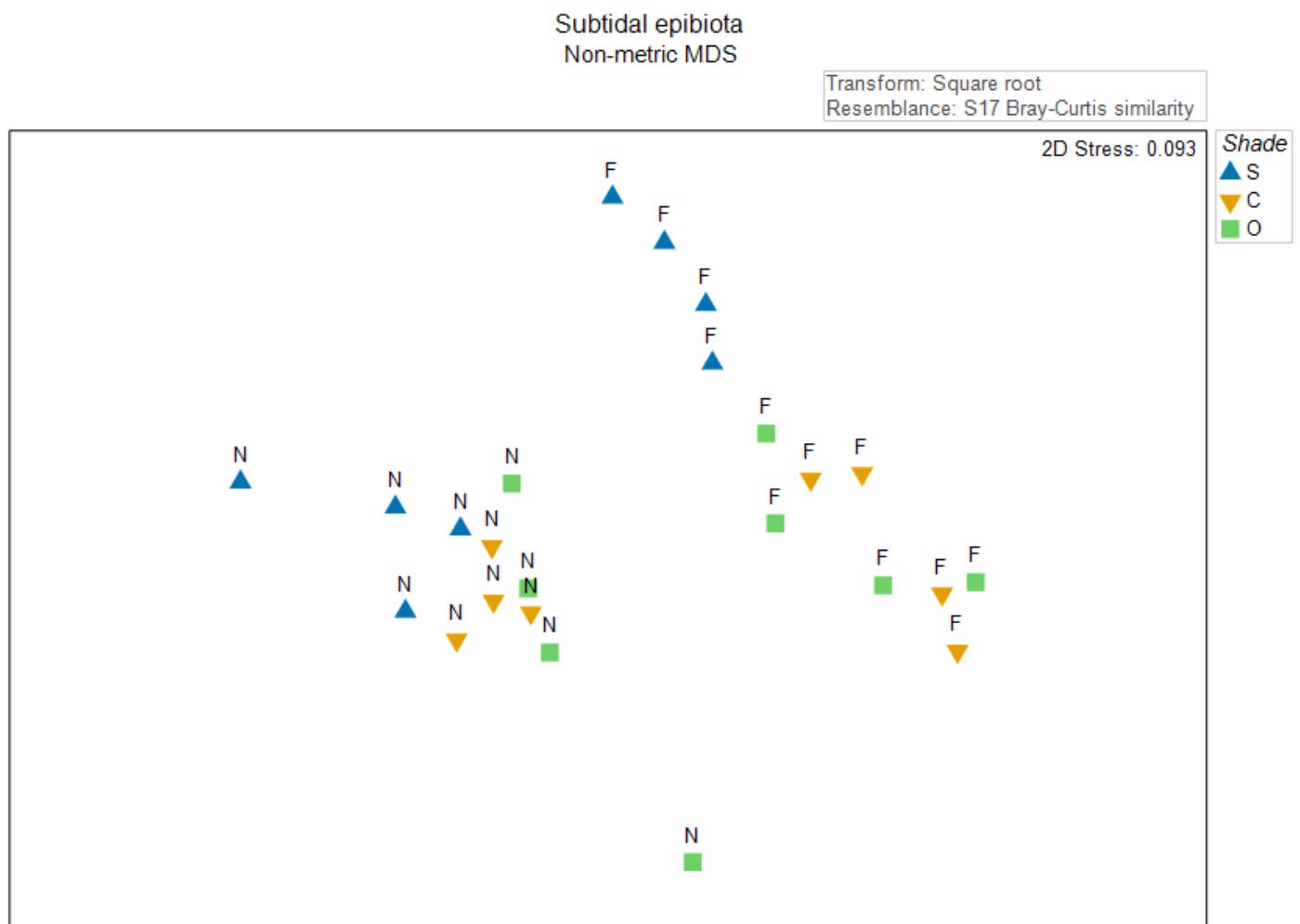


Fig. 11.1. Non-metric MDS of the individual sampling units from the two-factor crossed study design described by [Glasby \(1999\)](#) . Symbols correspond to the factor of 'Shade' ('S' = Shade, 'C' = Procedural Control, 'O' = Open surfaces), and labels correspond to the factor of 'Position' ('N' = Near, 'F' = Far).

The most obvious structuring factor here is Position; there is a large gap between samples that are near ('N') vs far ('F') from the seafloor. Let's suppose that we want to see more clearly the effects of Shade, and so we want to get residual distances after removing[†] the (larger) effect of Position.

Let's use the 2D Euclidean distances in the nMDS plot as a proxy for the full Bray-Curtis space here, just to help visualise what happens when we 'remove' effects. The effects of the Position factor (as represented in the 2D nMDS plot) are shown in Fig. 11.2 below. These are simply the deviations of the individual group centroids ('N' and 'F') from the overall centroid.

Subtidal epibiota Non-metric MDS

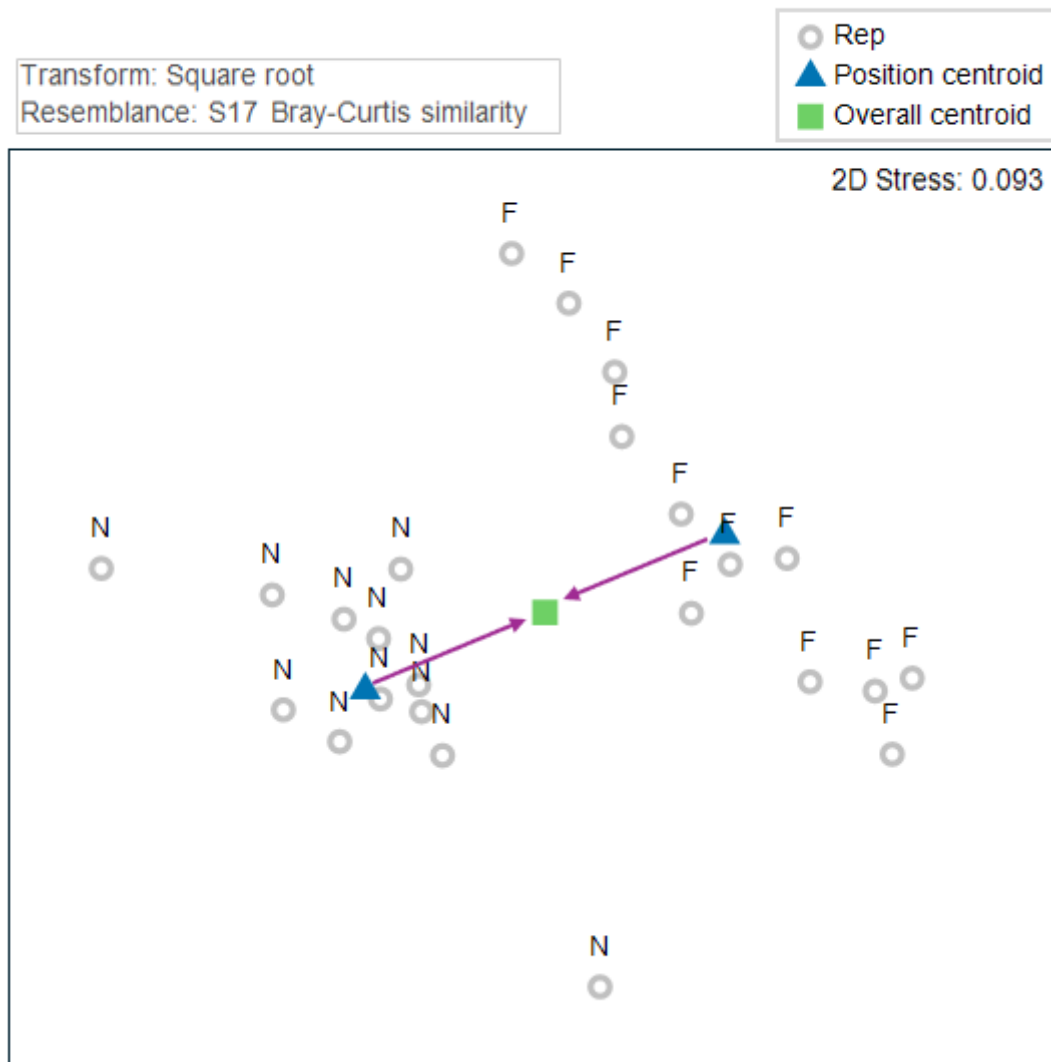


Fig. 11.2. Non-metric MDS precisely as in Fig. 11.1, but here showing only the labels correspond to the factor of 'Position' ('N' = Near, 'F' = Far), along with the centroids for each group (in blue), the overall centroid (in green) and the effects (arrows in a plum colour).

Now, to remove these effects, the steps are as follows:

- (i) Estimate the centroids for each 'Position' group (near and far); these will be calculated as the average of all of the samples along each of nMDS axis 1 and nMDS axis 2, done separately for each group.
- (ii) For each sample within a given group, subtract off the average for that group from its value (score) along nMDS axis 1, and do the same along nMDS axis 2; this operation removes the effect of 'Position' so that both the near and far samples are centered on a common (overall) centroid.
- (iii) Replot the samples anew after this centering operation.

This action of centering the groups onto a common centroid (i.e., of removing the effects of Position), done in the 2d nMDS space, is shown in Fig. 11.3 below.

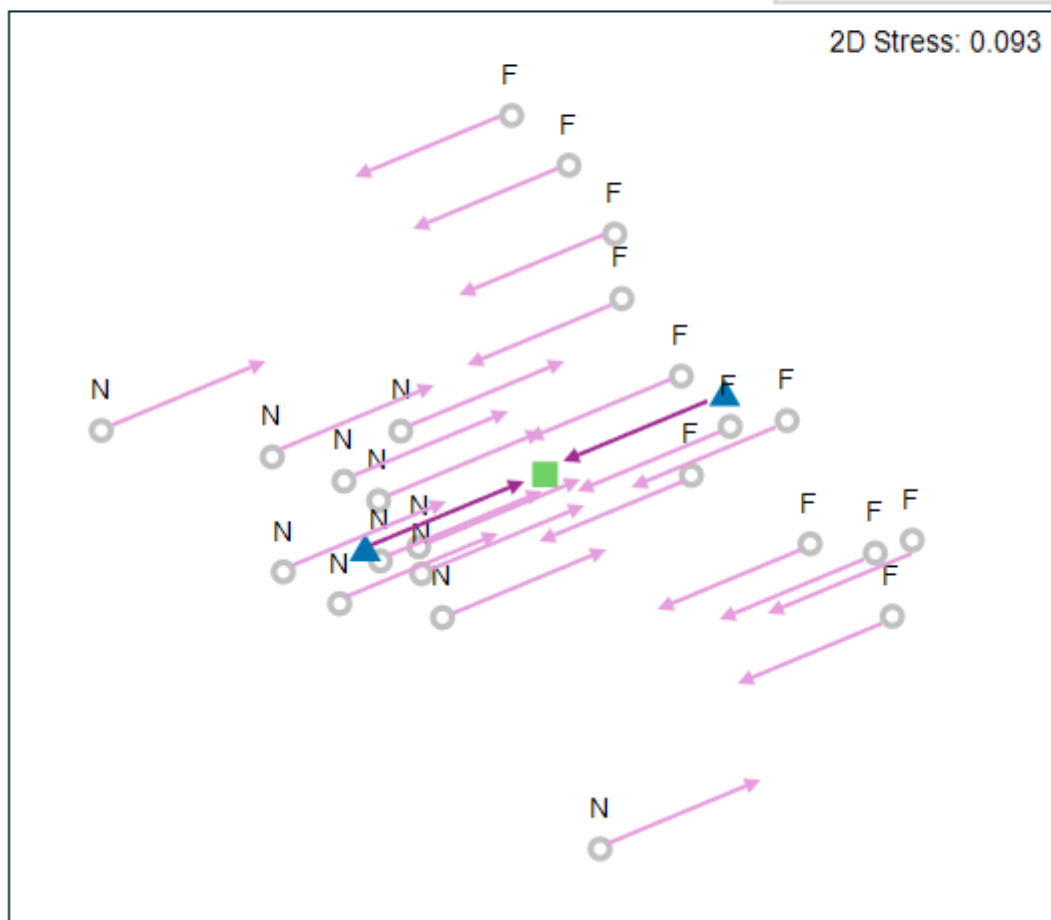
(a)

Subtidal epibiota
Non-metric MDS

Transform: Square root
Resemblance: S17 Bray-Curtis similarity

○ Rep
▲ Position centroid
■ Overall centroid

2D Stress: 0.093



(b)

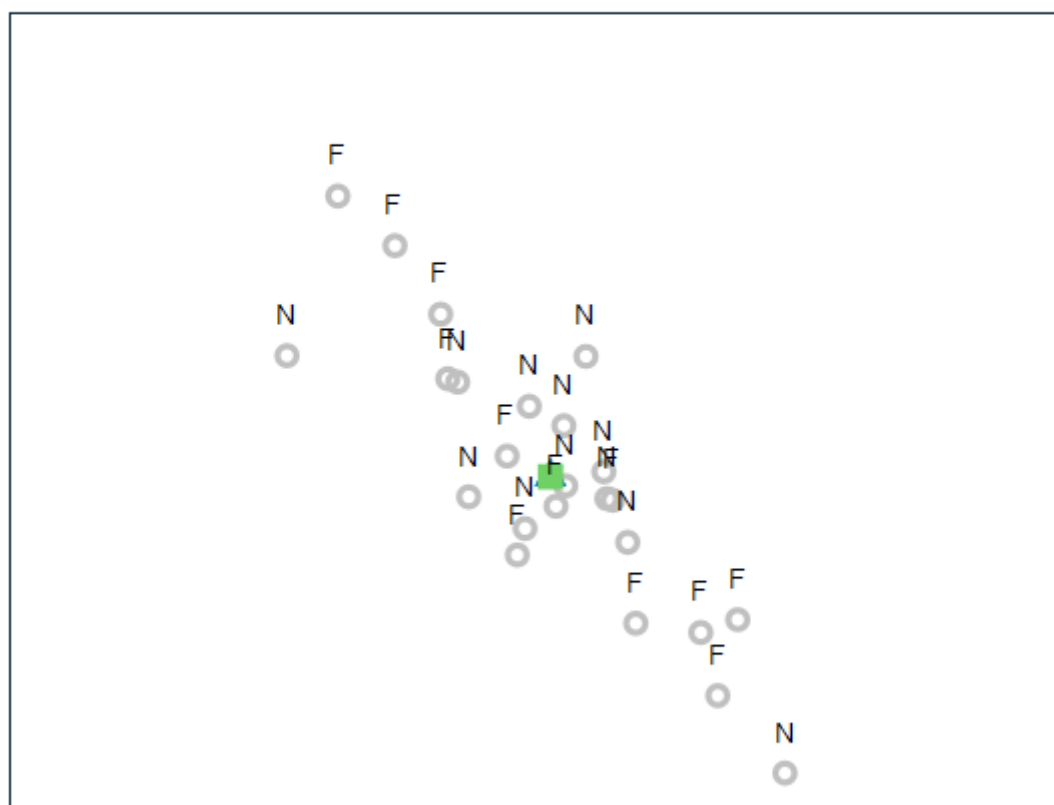


Fig. 11.3. (a) Non-metric MDS precisely as in Fig. 11.2, but here showing also the action of removing the Position effects from each sample unit (arrows in a light plum colour); and (b) the 'residual' plot of all the samples after removing the Position effect in this 2d nMDS space.

If we now put symbols corresponding to the factor of 'Shade', we can see the separation of assemblages in the Shade treatment from those in the Open and Control assemblages very clearly, and without any 'interference' from the effects of 'Position' in our plot (Fig. 11.4).

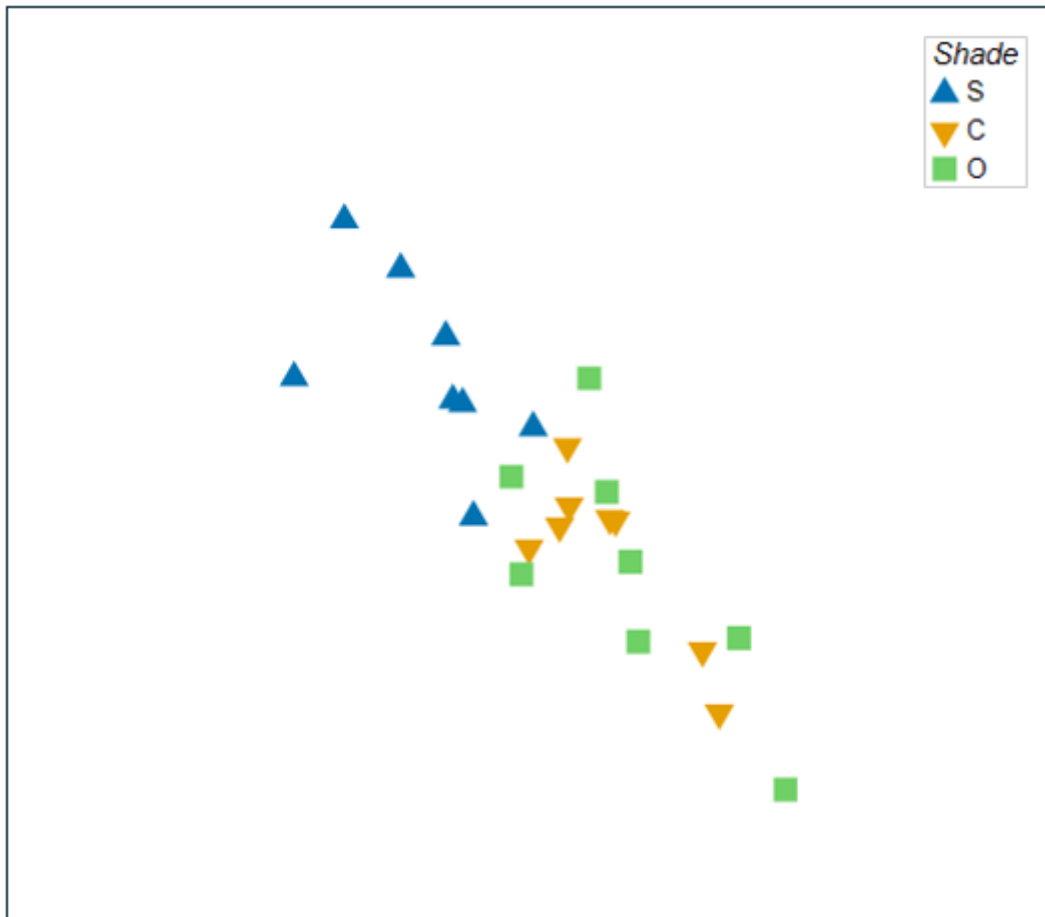
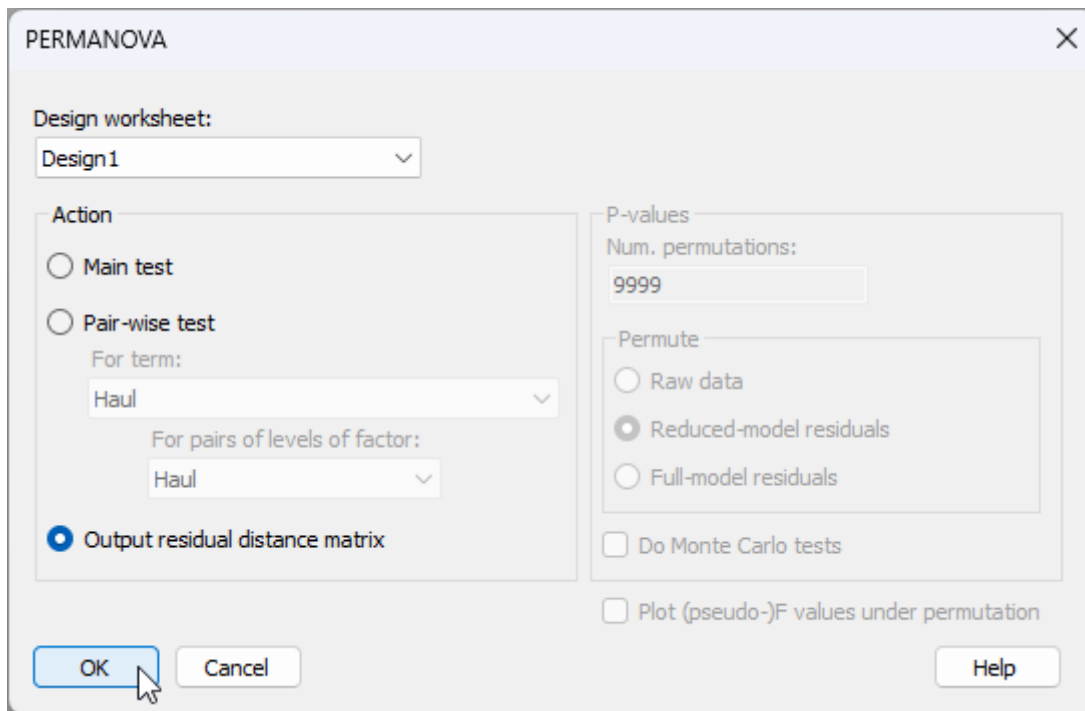


Fig. 11.4. The 'residual' plot of all the samples after removing the Position effect in this 2d nMDS space, precisely as in Fig. 11.3(b), but here showing symbols for the factor of Shade.

The above is a useful exercise to see how the effects of a factor can be removed in a 2D Euclidean nMDS space (Figs. 11.2 - 11.4). However, what we really want here is to perform this 'removing' operation in the full high-dimensional space of the chosen resemblance measure, hence to obtain a *residual distance matrix*, which *then* can be plotted via ordination.

Obtain a residual distance matrix and residual ordination plot

A residual distance matrix can be obtained after removing the effects of factors (and/or covariates) that have been specified in a given PERMANOVA design file by running the **PERMANOVA+ > PERMANOVA...** routine in PRIMER 8 and choosing (under the word 'Action' in the dialog) '\$\bullet\$Output residual distance matrix', as shown below.



The image shows a 'PERMANOVA' dialog box with a light blue title bar and a close button (X) in the top right corner. The dialog is divided into several sections. On the left, under 'Design worksheet:', there is a dropdown menu showing 'Design1'. Below this, the 'Action' section contains three radio buttons: 'Main test', 'Pair-wise test', and 'Output residual distance matrix', which is currently selected. The 'Pair-wise test' option has two sub-dropdowns: 'For term:' set to 'Haul' and 'For pairs of levels of factor:' also set to 'Haul'. On the right, the 'P-values' section includes a text box for 'Num. permutations:' containing the value '9999'. Below this is a 'Permute' section with three radio buttons: 'Raw data', 'Reduced-model residuals' (which is selected), and 'Full-model residuals'. At the bottom of the 'P-values' section are two checkboxes: 'Do Monte Carlo tests' and 'Plot (pseudo-)F values under permutation', both of which are unchecked. At the bottom left of the dialog are 'OK' and 'Cancel' buttons, with a mouse cursor pointing at the 'OK' button. At the bottom right is a 'Help' button.

The samples in that residual distance matrix, once created, can then be plotted *via* an ordination method of choice (typically nMDS) in the usual way.

Thus, for the data from [Glasby \(1999\)](#) , if we obtain a residual distance matrix in this way from a PERMANOVA design file with one factor ('Position'), then run nMDS on those residual distances, the resulting *residual ordination plot* looks like this (Fig. 11.5):

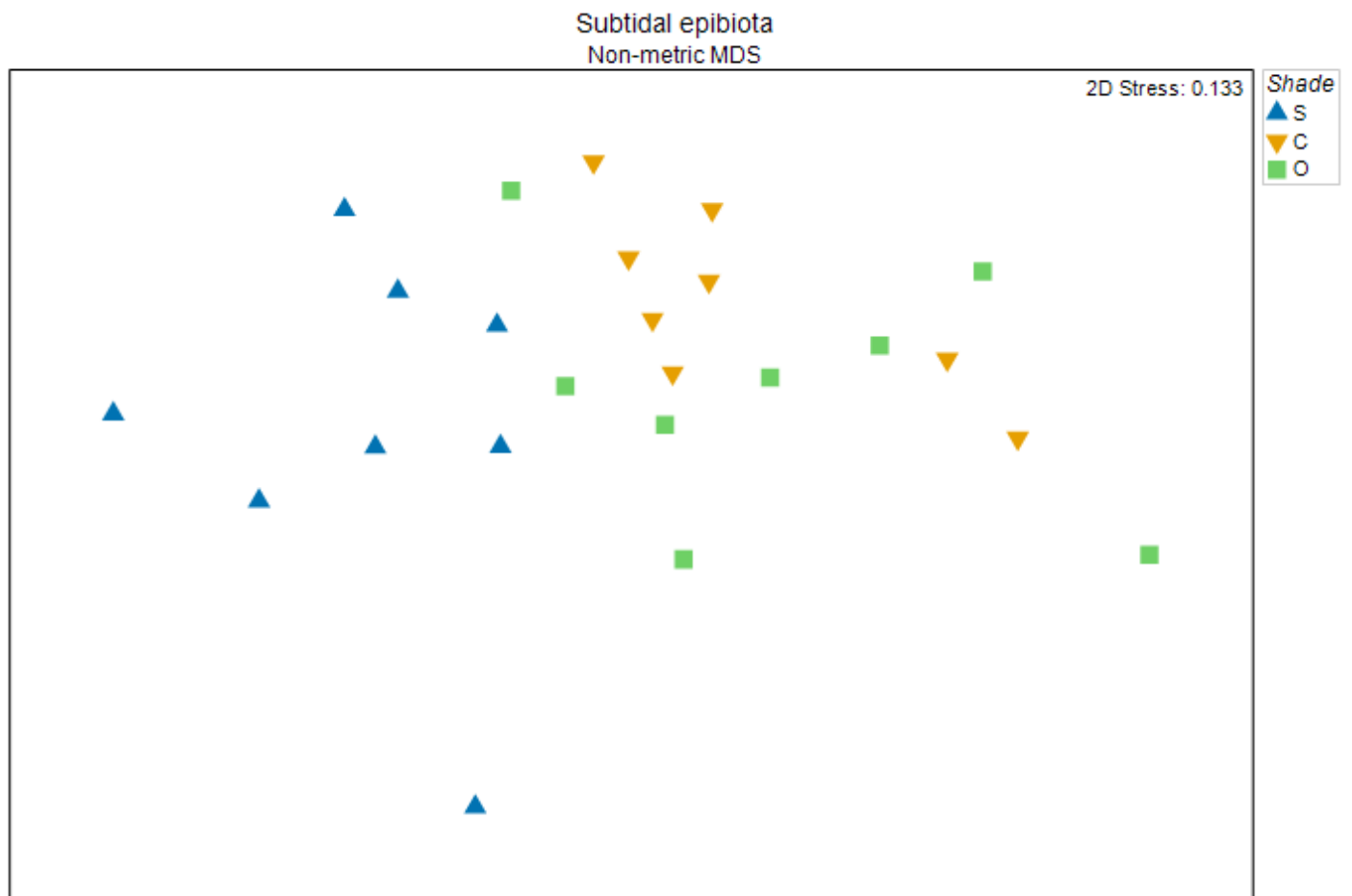


Fig. 11.5. Non-metric MDS plot of a residual distance matrix obtained after removing the effects of the factor 'Position' in the full-dimensional Bray-Curtis space. Symbols correspond to the factor of 'Shade' ('S' = Shade, 'C' = Procedural Control, 'O' = Open surfaces).

This plot shows (with a tolerable level of stress) the inter-sample relationships among multivariate residuals, after removing 'Position' effects in the high-dimensional Bray-Curtis space. The effects of shading on these subtidal epibiotic assemblages are very clearly seen now, indeed.

Removing quantitative (co)variables

It is possible to 'residualise' a distance matrix for a model that contains one or more quantitative covariables, even in the absence of any factors. This is done by fitting specified predictor variables (that is, the covariables you want to *remove*) in **PERMANOVA+ > DistLM...**, and ticking the box '\$\checkmark\$Output residual distance matrix' in the DISTLM dialog (e.g., see below).

DistLM

Predictor variables worksheet: Ekofisk_env_norm

☐ Group variables (indicator)

Select variables... Select groups...

Selection Procedure: ☒ All specified, ☐ Forward, ☐ Backward, ☐ Step-wise, ☐ Best

Selection Criterion: ☒ R², ☐ Adjusted R², ☐ AIC, ☐ AICc, ☐ BIC

☒ Do marginal tests, ☐ Do dbRDA plot, ☐ Variable correlations to worksheet, ☒ Output residual distance matrix

Num. permutations: 9999

Results: Best, Max num of best results: 10, Results detail: Normal

Variable naming: ☐ Number, ☒ Short names, ☐ Full names

OK Cancel Help

In this case, it is the *linear relationships* between the covariable(s) and the distribution of samples in the high-dimensional resemblance space that are 'removed'; of course, non-linear relationships (if any) may remain. The samples in that residual distance matrix can then of course be plotted *via* an ordination method of choice (typically nMDS) in the usual way.

Stress in residual ordination plots

In the example above, we see that the stress of the residual nMDS plot (Fig. 11.5, stress = 0.133) is a bit larger than the stress of the original nMDS configuration (Fig. 11.1, stress = 0.093). This is somewhat to be expected. If there are major structuring forces creating patterns in a multivariate data cloud, like strong effects of a particular factor that 'pulls' sets of samples (e.g., from different groups) away from each other, then the stress will tend to be relatively low, compared to an unstructured (but otherwise similar) multivariate data cloud (i.e., that looks like a 'ball of fuzz'). Consider: it is rather easy to show a clear difference between (say) two groups of samples in 2 dimensions when the rank order relationships among samples are the main (only) thing we are trying to maintain (as in an nMDS).[‡] In contrast, a lack of structuring forces effectively makes it harder to push high-dimensional information into small (2D or 3D) spaces. Thus, when we 'remove' some (one or more) primary factors that do create structure in any particular study and look at 'what is left' in a residual ordination plot, we may well be met with relatively high stress. This is a generalisation, and may not always be the case. It may well depend on how much additional structure (e.g., due to other, more minor factors), still remains. In the example above the difference in stress is fairly modest and poses no difficulties for interpretation.

A note of caution regarding tests of significance

In most cases, a residual distance matrix should *not* be used as input into secondary routines in order to test hypotheses regarding the significance of a factor (or variable) after 'removing' some other terms specified in X_r . The reason for this is that the testing procedures in PRIMER/PERMANOVA+ use permutation methods to obtain p -values. Interestingly, even though you think you have 'removed' the effects of a nuisance factor (say), and so it seems you have already correctly 'conditioned' upon it, it turns out that this 'conditioning' is *not* maintained once you then go and *permute* the data. Essentially, when you have a residual distance matrix, its behaviour under permutation is no longer guaranteed to be independent of the thing(s) you thought you had removed.

Fortunately, when you do a multi-factor PERMANOVA with all terms in the model (including nuisance factors, or anything else you'd like to 'remove' or 'account for'), the conditioning on terms that are *not* being tested is perfectly maintained for every individual test that is done (i.e., for each line of the PERMANOVA output table). The nature of the conditioning done does depend on the [Type of sum of squares](#), but this choice is fully under the control of the end-user and, once chosen, is perfectly maintained throughout all of the tests. Similarly, when you do sequential tests in DISTLM, the conditioning (on all previously fitted terms) is maintained perfectly, including under permutation. The essential trick here is to *re-residualise* for the terms you are trying to condition upon *after every permutation*. For more details, see [Anderson & Legendre \(1999\)](#) and [Anderson & Robinson \(2001\)](#).

Suffice it to say that, generally, it is not wise to run tests on residual distance matrices, but they can be very useful tools for visualising patterns that might be difficult to see in typical ('unconditioned' and 'unconstrained') ordination plots.

[¶]The value of $\text{tr}(\mathbf{G})$ is also equal to the total sum of squared inter-point dissimilarities in $\mathbf{D}$ divided by the number of sampling units, N .

[†]When we say 'removing' the effects, we could also say 'conditioning on' or 'accounting for' those effects.

[‡]The most extreme example of this situation is reflected by the zero stress that is obtained as a 'degenerate solution', and a 'collapsed' nMDS occurs. For example, see [section 5.2 in Change in Marine Communities](#).

Revision #103

Created 16 June 2025 21:17:26 by Marti

Updated 5 December 2025 03:46:04 by Marti