

6.6 Example: one-way PERMANOVA allowing heterogeneity

Let's look now at an example where there is a single factor in the study design, the number of replicates per group is unequal and there is clear heterogeneity in multivariate dispersions among the groups. [Ellingsen & Gray \(2002\)](#) studied the biodiversity of soft-sediment macrobenthic organisms and its relationship with environmental variation over large spatial scales in the North Sea. Samples of soft-sediment macrobenthic organisms were obtained from $N = 101$ sites occurring in five large delineated areas along a transect spanning 15 degrees of latitude (Fig. 6.4). The sample sizes in the five areas were: $n_1 = 16$, $n_2 = 21$, $n_3 = 25$, $n_4 = 19$ and $n_5 = 20$. A total of $p = 809$ taxa were recorded overall, and samples consisted of abundances pooled across five benthic grabs obtained at each site.

Interest lies in comparing the multivariate assemblages of organisms occurring in these five areas. More specifically, we wish to use PERMANOVA to test the null hypothesis:

- H_0 : there are no differences in the centroids of these 5 areas in the space of the Jaccard resemblance measure, allowing for any potential heterogeneity in the within-group dispersions among these areas.

In ecology, the Jaccard measure is directly interpretable as the *percentage of shared species* between every pair of sampling units. When expressed as a dissimilarity, it is often used as a pure measure of *turnover* in studies of *beta diversity* ([Anderson et al. \(2006\)](#)).

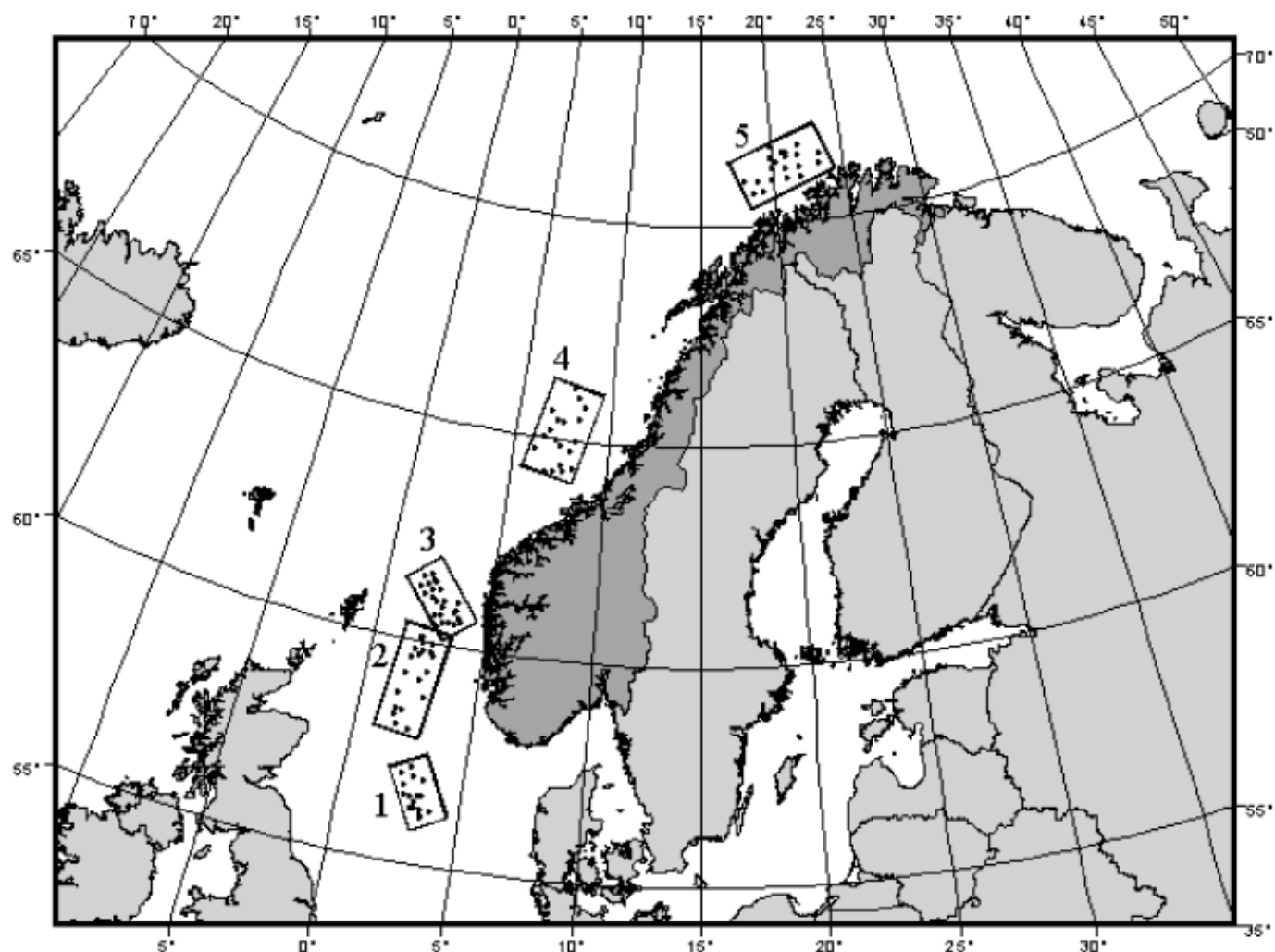


Fig. 6.4. Map showing locations of sites in each of 5 areas in the North Sea from which macrobenthic fauna were sampled (after [Ellingsen & Gray \(2002\)](#)).

Open the data in PRIMER and examine patterns

1. Start running **PRIMER 8**, then click **File > Open...** to open the data file named '[Norway_macrofauna.pri](#)' (found inside the '[Examples_P8 > Norway_macrofauna](#)' folder).

PRIMER

File Edit Select View Wizards Pre-treatment Analyse Plots PERMANOVA+ Tools Window Help

Workspace

Norway_macrofauna

Norway_macrofauna

Norwegian Continental Shelf Macrobenthos Abundance

	Samples							
	1	2	3	4	5	6	7	8
Abludomelita obtus	1	0	0	1	0	0	0	0
Acanthonotozoma d	0	0	0	0	0	0	0	0
Acidostoma nodifer	0	0	0	0	0	0	0	0
Acidostoma spp.	0	0	0	0	0	0	0	0
Ampelisca aequicon	0	0	0	0	0	0	0	0
Ampelisca amblyops	0	0	0	0	0	0	0	0
Ampelisca anomala	0	0	0	0	0	0	0	0
Ampelisca brevicorn	1	1	0	2	2	0	0	0
Ampelisca diadema	0	0	0	0	0	0	0	0
Ampelisca gibba	0	0	0	0	0	0	0	0
Ampelisca macroce	0	1	0	1	0	1	0	0
Ampelisca odontopl	0	0	0	0	0	0	0	0
Ampelisca pusilla	0	0	0	0	0	0	0	0
Ampelisca spinipes	0	0	0	0	0	0	0	0
Ampelisca tenuicom	2	0	0	0	3	1	0	0
Ampelisca typica	0	0	0	0	0	0	0	0
Amphilocheus manu	0	0	0	0	0	0	0	0
Amphilocheus tenui	0	0	0	0	0	0	0	0
Anapagurus laevis	0	0	0	0	0	0	0	0

- Get the resemblance matrix among the sampling units based on the Jaccard measure. Click **Analyse > Resemblance...** > (Measure > $\bullet$ Other) and from the drop-down list choose 'S7 Jaccard'.

Resemblance

Measure

☐ Bray-Curtis similarity

☐ Euclidean distance

☐ Index of association

☒ Other

☒ Similarity ☒ Distance/dissimilarity

☐ Quantitative ☐ P/A

☐ Correlation ☐ Taxonomic P/A

(exc0-0 = excluding joint absences)

S7 Jaccard

Analyse between

☒ Samples

☐ Variables

☐ Add dummy variable

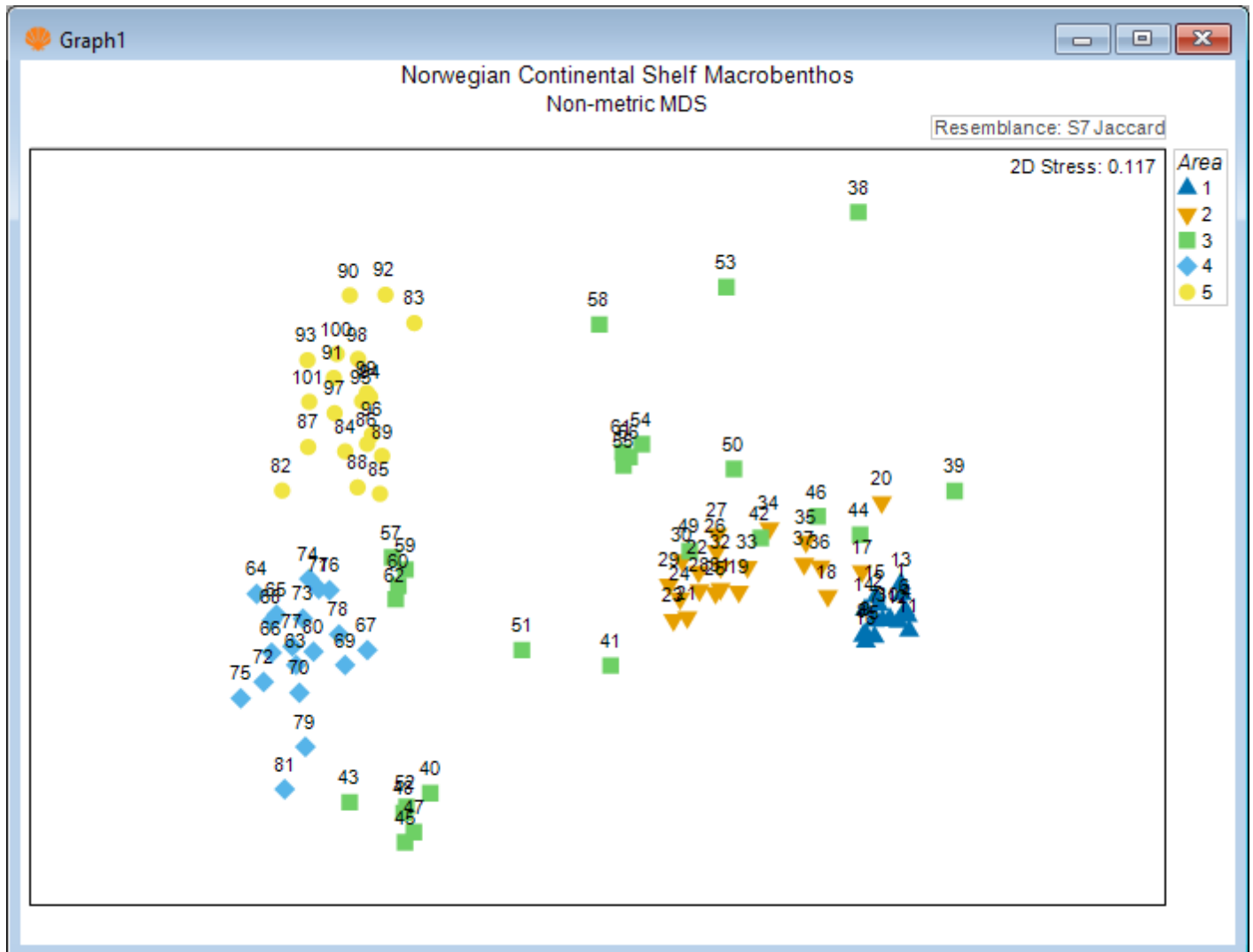
Value:

1

OK Cancel Help

The resulting resemblance matrix will be called 'Resem1'.

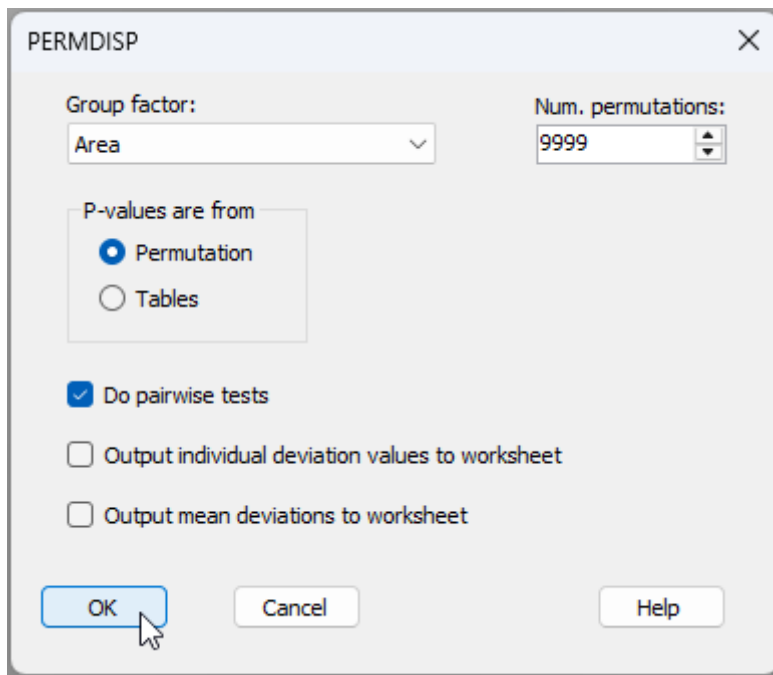
3. To visualise the inter-sample relationships based on the identities of the fauna they contain, obtain a non-metric multi-dimensional scaling (nMDS) ordination plot based on the Jaccard resemblances. From the 'Resem1' similarity matrix, click **Analyse > MDS > Non-metric MDS (nMDS)...**, take the default options and click 'OK'. This will generate the following 2D ordination plot (called 'Graph1'), or a highly similar solution¹:



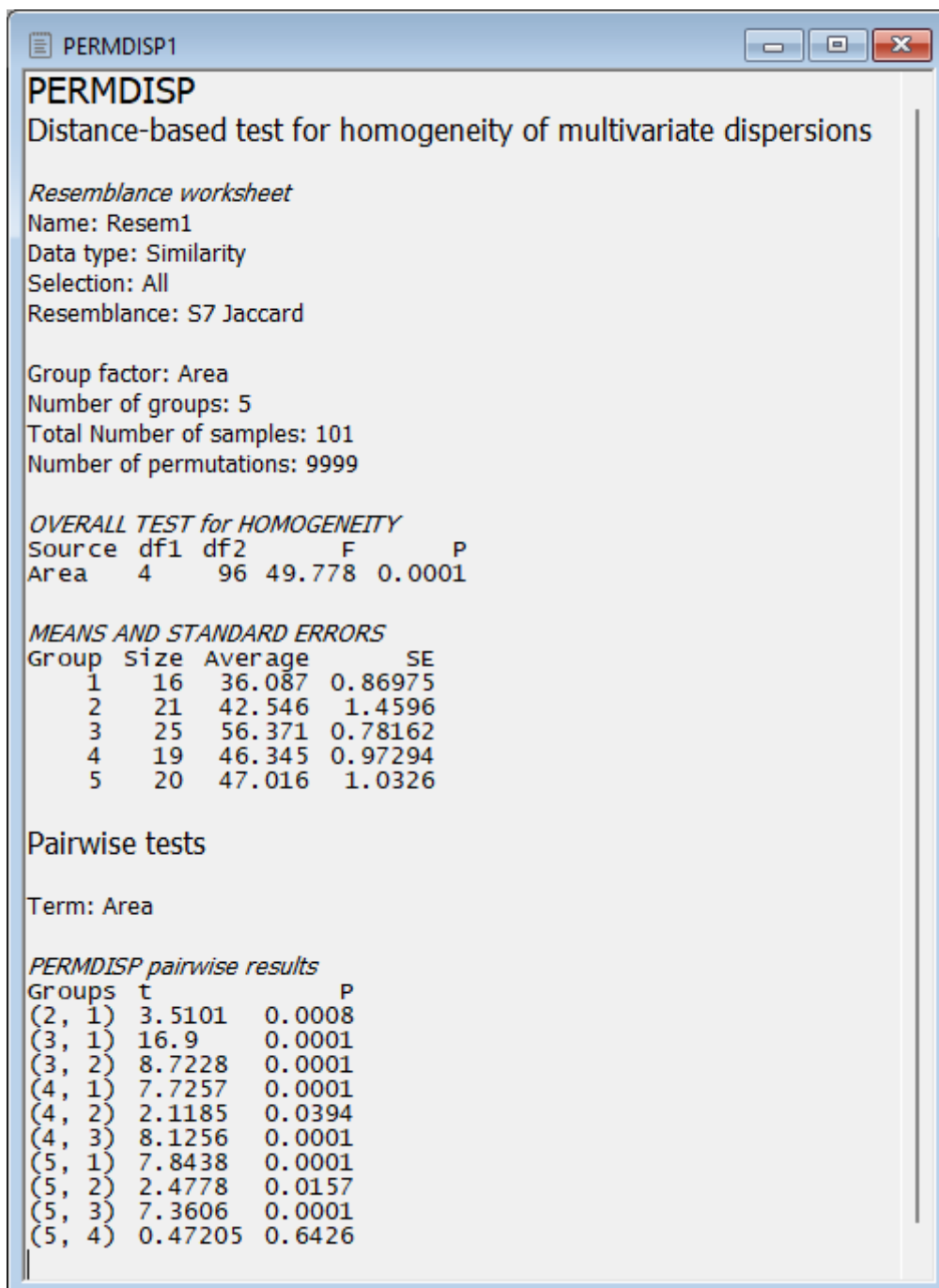
Perhaps the most striking pattern here is the tight clustering of samples from Area 1 (low dispersion) and the very large spread of the samples from Area 3 (high dispersion), compared to Areas 2, 4 and 5.

Test for homogeneity of multivariate dispersions

4. Although it is fairly obvious from the graphic, let's test the null hypothesis of no differences in the within-group dispersions among the five areas using PERMDISP. From the 'Resem1' similarity matrix, click **PERMANOVA+ > PERMDISP...** > Group factor: Area (leaving the rest as their defaults) and click 'OK', as shown below:



The results are given in the file 'PERMDISP1' of the Explorer tree (see below).

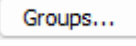


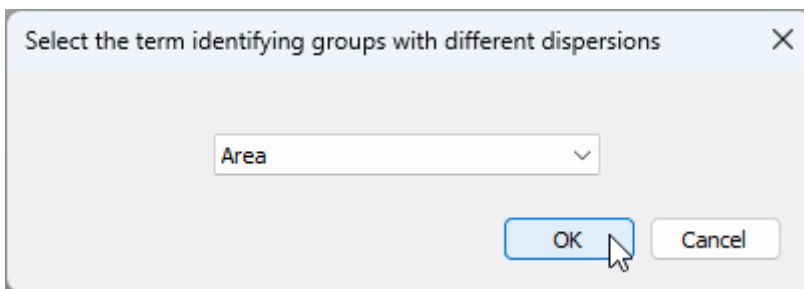
There are clearly highly significant differences in the dispersions among the groups ($F = 49.778$, $P = 0.0001$ with 9999 permutations). The pairwise tests, furthermore, reveal how Area 1 and Area 3 differ significantly from one another and from the other three groups (2, 4 and 5) with respect to their dispersions, reflecting rather directly the patterns of differences in spread we observed in the nMDS plot.

Test for differences in centroids, allowing for heterogeneity

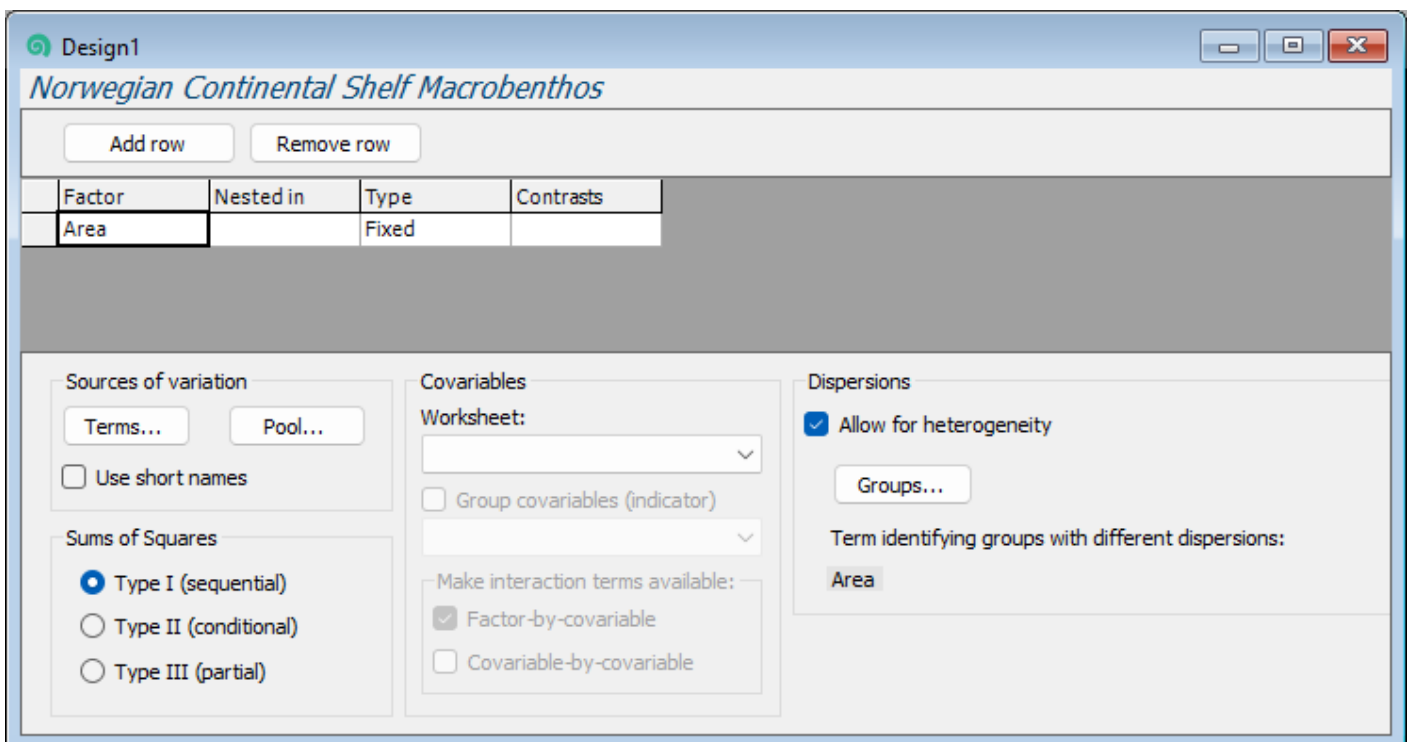
Having observed these dispersion differences among the areas - quite interesting differences in themselves - we now aim to test for differences in centroids, allowing for that heterogeneity.

Running a PERMANOVA requires two steps: (i) setting up a design file; and (ii) running the PERMANOVA analysis on a given data set in response to a specified design.

5. From the 'Resem1' similarity matrix, create the design file by clicking **PERMANOVA+ > Create PERMANOVA Design....** You will see a new item named 'Design1' (symbolised by ) in the Explorer tree. You will need to do the following:
 - Click the white cell in the first column under the word 'Factor' and choose: 'Area' as the sole factor of interest for this design. (Note in passing that the default is to treat this factor as 'Fixed', as shown under the word 'Type' in the third column of the design file, which is fine here).
 - Under 'Dispersions', tick the box ☒ 'Allow for heterogeneity', then click the  button.
 - In the resulting dialog box entitled 'Select the term identifying groups with different dispersions' choose 'Area'.



Your resulting design file should look like this:

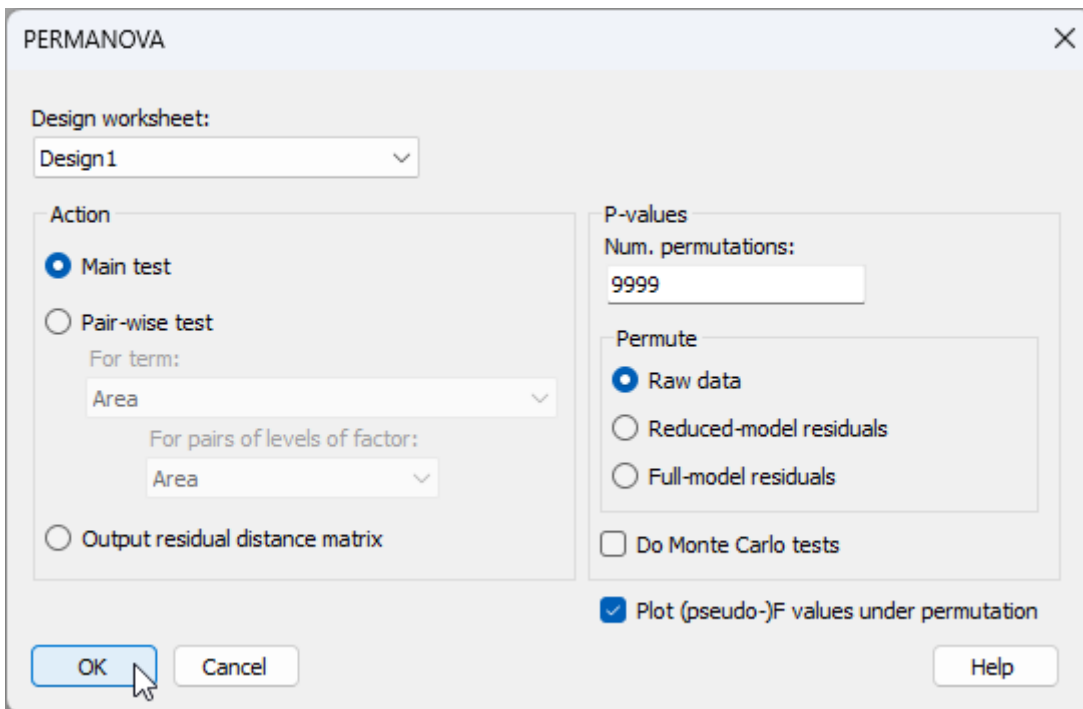


Note that by ticking the box ☒ 'Allow for heterogeneity', we ensure that the PERMANOVA tests will be done using F_2 rather than F_1 for all tests of relevant terms affected by the heterogeneity we have identified using the 'Groups' button.

6. Now that you have created the design file, you are ready to run the analysis itself. We're going to test the null hypothesis of no differences in the centroids among the five areas using PERMANOVA and allowing for heterogeneity in dispersions. Go back to the 'Resem1'

similarity matrix and, from there, click **PERMANOVA+** > **PERMANOVA....** In the PERMANOVA dialog window:

- Under the words 'Design worksheet:' make sure you choose the name of the correct design file, i.e. 'Design1'.
- Under the word 'Action', choose ☒ Main test.
- Under the word 'Permute' choose ☒ Raw data (because there is only one factor here).
- Optionally, you can choose to tick the box to ☒ 'Plot (pseudo-)F values under permutation'. The rest of the items in the dialog can remain as the defaults (see below), then click '**OK**'.



The resulting PERMANOVA output file ('PERMANOVA1', shown below) indicates strong evidence against the null hypothesis of no differences in the centroids among these five areas ($F_{2,13} = 13.51$, $P = 0.0001$ with 9999 permutations). So, not only are there differences in the variability of the assemblages (evidenced by the PERMDISP analysis), there are also clear shifts in centroid evidencing overall turnover in the identities of species across these five areas (also apparent in the nMDS plot above).

PERMANOVA1

PERMANOVA

Permutational MANOVA

Resemblance worksheet

Name: Resem1

Data type: Similarity

Selection: All

Resemblance: S7 Jaccard

PERMANOVA design worksheet

Name: Design1

Title: Norwegian Continental Shelf Macrobenthos

Sums of squares type: Type I (sequential)

Permutation method: Unrestricted permutation of raw data

Number of permutations: 9999

Allow for heterogeneity of within-group dispersions

Terms identifying groups with different dispersions: Area

Factors

Name	Type	Levels
Area	Fixed	5

Excluded terms

No terms excluded

Inestimable terms

No inestimable terms.

PERMANOVA table of results

Source	df	SS	MS	Pseudo-F	P(perm)	Unique perms
Area	4	1.2106E+05	30265	13.51	0.0001	9840
Res	96	2.2548E+05	2348.8			
Total	100	3.4654E+05				

Estimates of components of variation

Source	Estimate	Sq.root
S(Area)	1394.7	37.345
V(Res)	2348.8	48.464

Details of the expected mean squares (EMS) for the model

Source	EMS
Area	$0.2005 * V(Res(5)) + 0.20297 * V(Res(4)) + 0.18812 * V(Res(3)) + 0.19802 * V(Res(2)) + 0.2104 * V(Res(1)) + 20.094 * S(Area)$

Construction of Pseudo-F ratio(s) from mean squares

Source	Numerator	Denominator	Num. df	Den. df
Area	$1 * Area$	$0.2005 * Res(5) + 0.20297 * Res(4) + 0.18812 * Res(3) + 0.19802 * Res(2) + 0.2104 * Res(1)$	4	94.45

Outputs

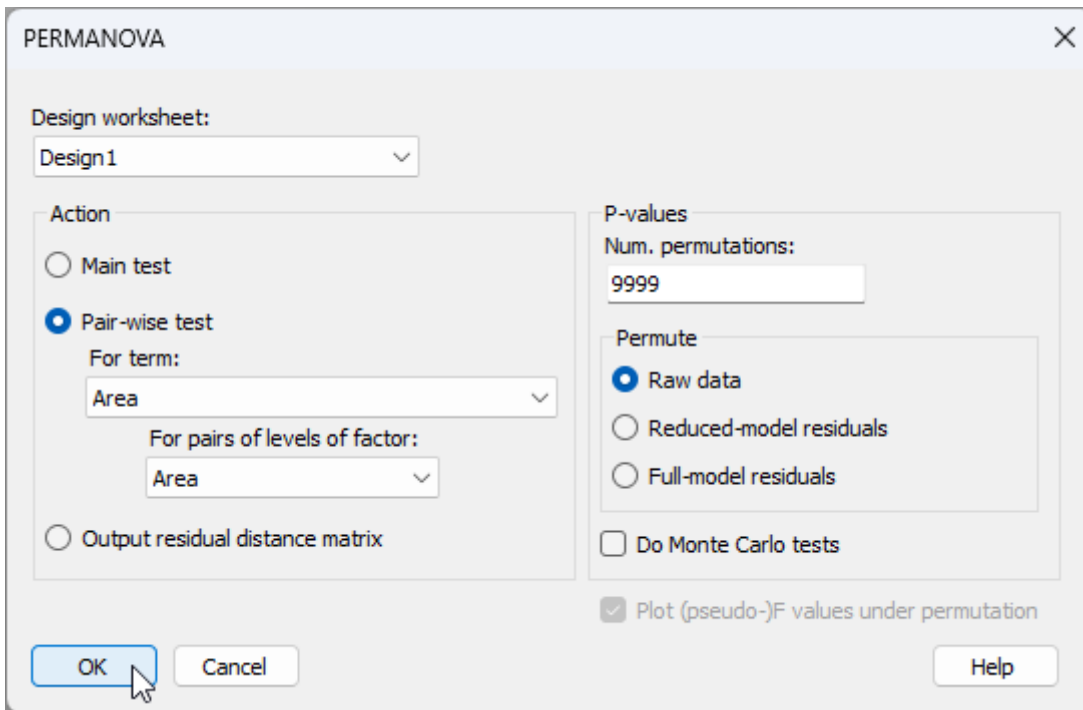
Plot: Graph5

It is worth noticing a couple of things about the PERMANOVA output file that makes it different from what would be obtained if we did not allow for heterogeneity. Specifically:

- The expectation of the mean square for 'Area' includes linear combinations involving *separate individual measures of residual variation* for each of the five areas. In other words, we do not have a single pooled estimate of error variance at work here, but an explicit recognition of the different within-group dispersions.
- The *denominator degrees of freedom* for the test of 'Area', therefore, is not a whole number, but instead this value is drawn directly from the theory for this arising from [the univariate solution to the Behrens-Fisher problem](#) described by [Brown & Forsythe \(1974\)](#) . As previously discussed, this has no direct consequence on the test by permutation using \$F_2\$ for the multivariate setting, but it does point to the fact that the power of the test using \$F_2\$ may well differ from that using \$F_1\$ (which perfectly stands to reason, as they are actually testing different hypotheses).

7. Having found a significant result for the main test in PERMANOVA, it is now desirable to run pair-wise comparisons that also will account for heterogeneity. We simply re-run the PERMANOVA routine from the same resemblance matrix, pointing to the same design file

('Design1'), but this time, under the word 'Action', we choose $\bullet$ Pair-wise test > For term: 'Area' > For pairs of levels of factor: 'Area', like this:



The image shows a 'PERMANOVA' dialog box with the following settings:

- Design worksheet:** Design1
- Action:**
 - ☐ Main test
 - ☒ Pair-wise test
 - For term:** Area
 - For pairs of levels of factor:** Area
 - ☐ Output residual distance matrix
- P-values:**
 - Num. permutations:** 9999
 - Permute:**
 - ☒ Raw data
 - ☐ Reduced-model residuals
 - ☐ Full-model residuals
 - ☐ Do Monte Carlo tests
- ☒ Plot (pseudo-)F values under permutation

Buttons: OK, Cancel, Help

Our results file for the pair-wise tests ('PERMANOVA2') will then appear as follows:

PERMANOVA2

PAIR-WISE TESTS

Term 'Area'

Groups	t	P(perm)	Unique perms
1, 2	3.6287	0.0001	9882
1, 3	3.7141	0.0001	9917
1, 4	5.067	0.0001	9897
1, 5	5.0187	0.0001	9894
2, 3	2.6151	0.0001	9857
2, 4	4.3762	0.0001	9888
2, 5	4.2595	0.0001	9882
3, 4	2.7733	0.0001	9882
3, 5	2.9003	0.0001	9869
4, 5	3.1949	0.0001	9870

Denominators

Groups	Denominator	Den. df
1, 2	0.43243*Res(2) + 0.56757*Res(1)	34.54
1, 3	0.39024*Res(2) + 0.60976*Res(1)	38.97
1, 4	0.45714*Res(2) + 0.54286*Res(1)	32.85
1, 5	0.44444*Res(2) + 0.55556*Res(1)	33.97
2, 3	0.45652*Res(2) + 0.54348*Res(1)	43.66
2, 4	0.525*Res(2) + 0.475*Res(1)	36.77
2, 5	0.5122*Res(2) + 0.4878*Res(1)	38.21
3, 4	0.56818*Res(2) + 0.43182*Res(1)	41.64
3, 5	0.55556*Res(2) + 0.44444*Res(1)	42.87
4, 5	0.48718*Res(2) + 0.51282*Res(1)	36.94

Average Similarity between/within groups

	1	2	3	4	5
1	47.315				
2	25.828	38.336			
3	13.403	19.283	19.262		
4	6.1244	10.071	14.804	32.635	
5	6.6963	11.191	13.4	17.513	31.792

In passing, we can see how using \$F_2\$ on the pairwise tests creates differences from what we would see for a PERMANOVA using \$F_1\$ on these data. These differences essentially mirror what we saw for the main test; namely, there are separate individual measures of residual variation for each area, and there are also non-integer denominator degrees of freedom for each test.

Overall, from this study, we can conclude that there are highly significant differences in the identities of species obtained from each of these five different areas - they clearly contain different sorts of soft-sediment benthic assemblages. This is so despite the very large variation in the assemblages inhabiting different sites sampled from Area 3. Our tests accounted for that.

It is obviously very satisfying to be empowered by the new PERMANOVA routine in PRIMER 8. We can now make statistically rigorous inferences about differences in centroids in the space of a chosen resemblance measure that allows for heterogeneity.

¶If your plot looks different, it is probably because of an arbitrary rotation or perhaps a 'flipping' of the X axis and/or the Y axis. Any nMDS results shown in an ordination diagram are invariant to changes in the signs of the axes and hold equivalent information for interpretation (preserving as they do the rank-order inter-relationships among the points). You can 'flip' either axis by right-clicking anywhere on the plot (to bring up the 'Graph' menu) and then click 'Flip X' and/or 'Flip Y'.

Revision #100

Created 31 March 2025 22:46:11 by Marti

Updated 4 December 2025 01:34:07 by Marti