8. Specify Subject/Whole-plot error in PERMANOVA

8.1 Designs lacking replication

In some cases, experiments are done in a way that lacks replication, often at the smallest spatial or temporal scale in the experimental design, but sometimes at larger scales as well. Examples of designs that lack replication include (but are not limited to):

- randomised blocks
- repeated measures
- split-plots (and split-split-plots, etc.)

Randomised blocks

For clarity on what follows regarding 'lack of replication', let's start by considering a simple *randomised block design*. For example, there might be a = 4 blocks, and within each block, there might be b = 3 treatments randomly allocated to replicates within each block (Fig. 8.1).

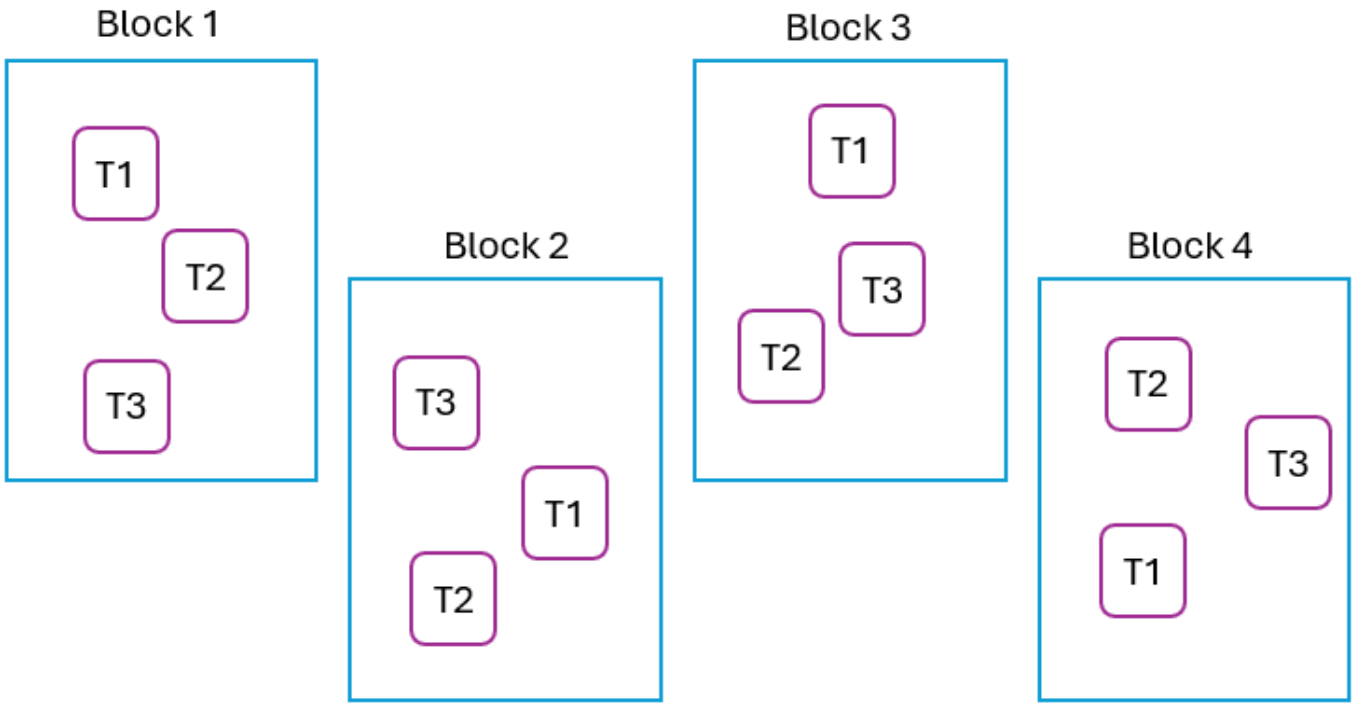


Fig. 8.1. Schematic diagram of a randomised block design, with a = 4 blocks and b = 3 treatments randomly allocated to the three replicates within each block.

The number of replicates in each block is equal to the number of different treatments. Thus, there is no replication of the treatments within any of the blocks; i.e., there is only one sampling unit per treatment per block (n = 1). This means that we cannot estimate any component of variation that might be due to a potential 'Treatment x Block' interaction, as it is inextricably confounded with the residual variation (from one sampling unit to the next). In PRIMER, the PERMANOVA routine will recognise this situation automatically. It will issue a warning to highlight this lack of

replication, but then it *will* permit you to proceed with the analysis anyway and simply exclude the 'Treatment x Block' interaction term from the model.

Repeated measures

A similar thing (lack of replication) often happens with study designs that involve *repeated measures*. Suppose we want to follow the health outcomes for a set of (say) 4 individuals who have taken a drug and another set of 4 individuals that have taken a placebo. We have the factor of 'Treatment' with a = 2 levels (Drug, Placebo), and we monitor the health of our subjects by measuring one or more variables on each individual at a series of time points (time 1, time 2, time 3, ..., time 5). In terms of factors, we therefore also have b = 4 individuals per treatment ('Subject', random and nested in 'Treatment') and c = 5 time-points, *but* we only have one measurement per person per time point (Fig. 8.2).

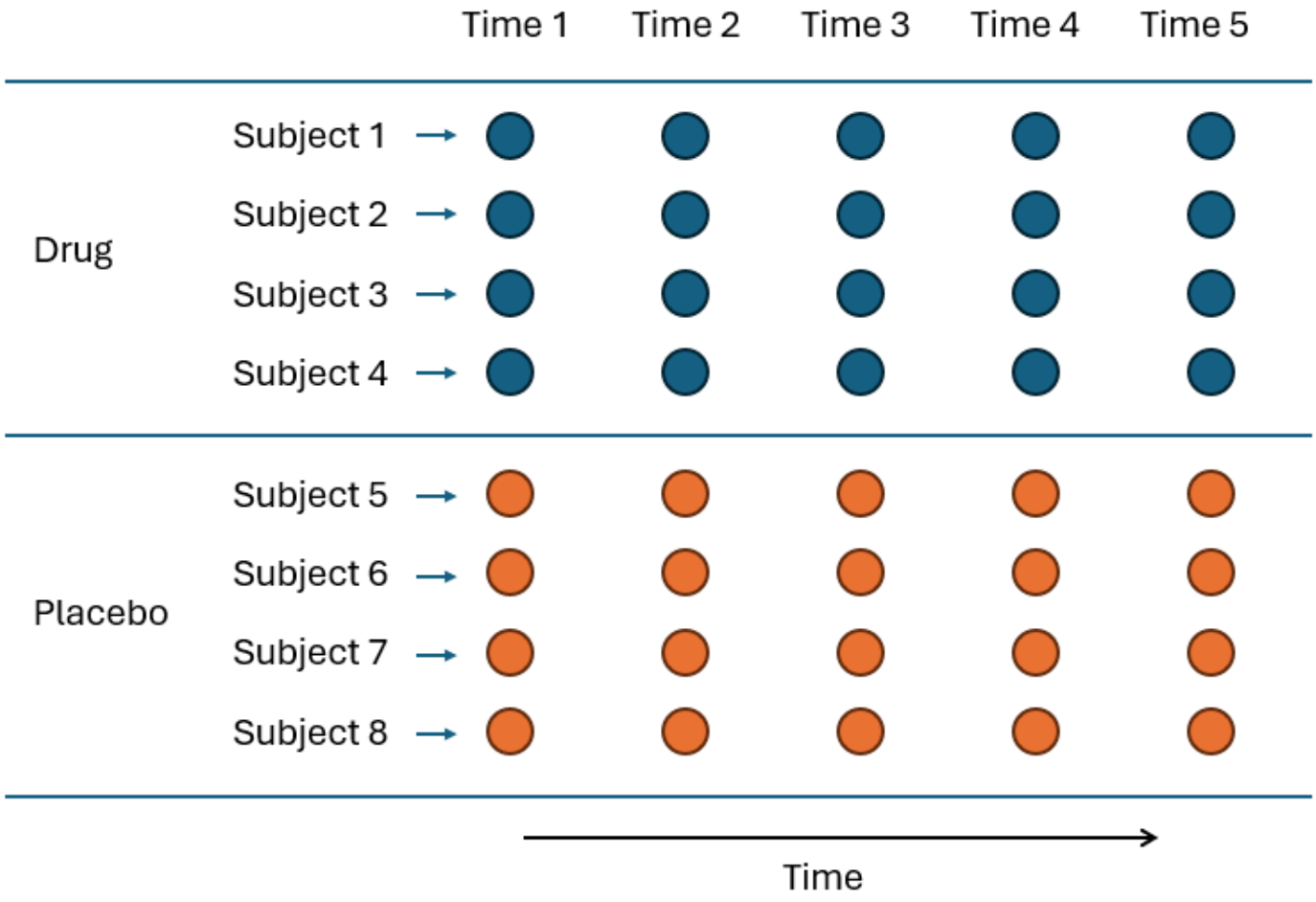


Fig. 8.2. Schematic diagram of a repeated measures design, with a = 2 treatments (Drug and Placebo), administered to each of b = 4 individuals (Subjects) per treatment, who were then each sampled repeatedly through time (t = 1, ..., 5).

We can see there is a similar problem here. The variation due to the potential 'Subject(Treatment) x Time' interaction (if any) is of course inextricably confounded with the residual variation among the sampling units themselves. Once again, just as in the randomised block case, the PERMANOVA routine will recognise this lack of replication and will automatically exclude the 'Subject(Treatment) x Time' term from the model output, after issuing a suitable warning.

Split plots

In some cases, there may be lack of replication not only at the smallest scale, but also at a larger spatial (or temporal) scale in the design. *Split-plot designs* are classically used in agricultural experiments, where the researcher is interested in investigating more than one factor, but perhaps one of these factors occurs at (or must be manipulated or administered at) a larger scale than others. A typical example might be a study of the effects of irrigation and nutrients on the growth of corn (maize). Different irrigation levels might need to be applied to large areas (fields) due to the equipment and logistics involved, while fertilizers providing different nutrient levels can be applied to smaller areas (plots) within the fields having differing levels of irrigation.

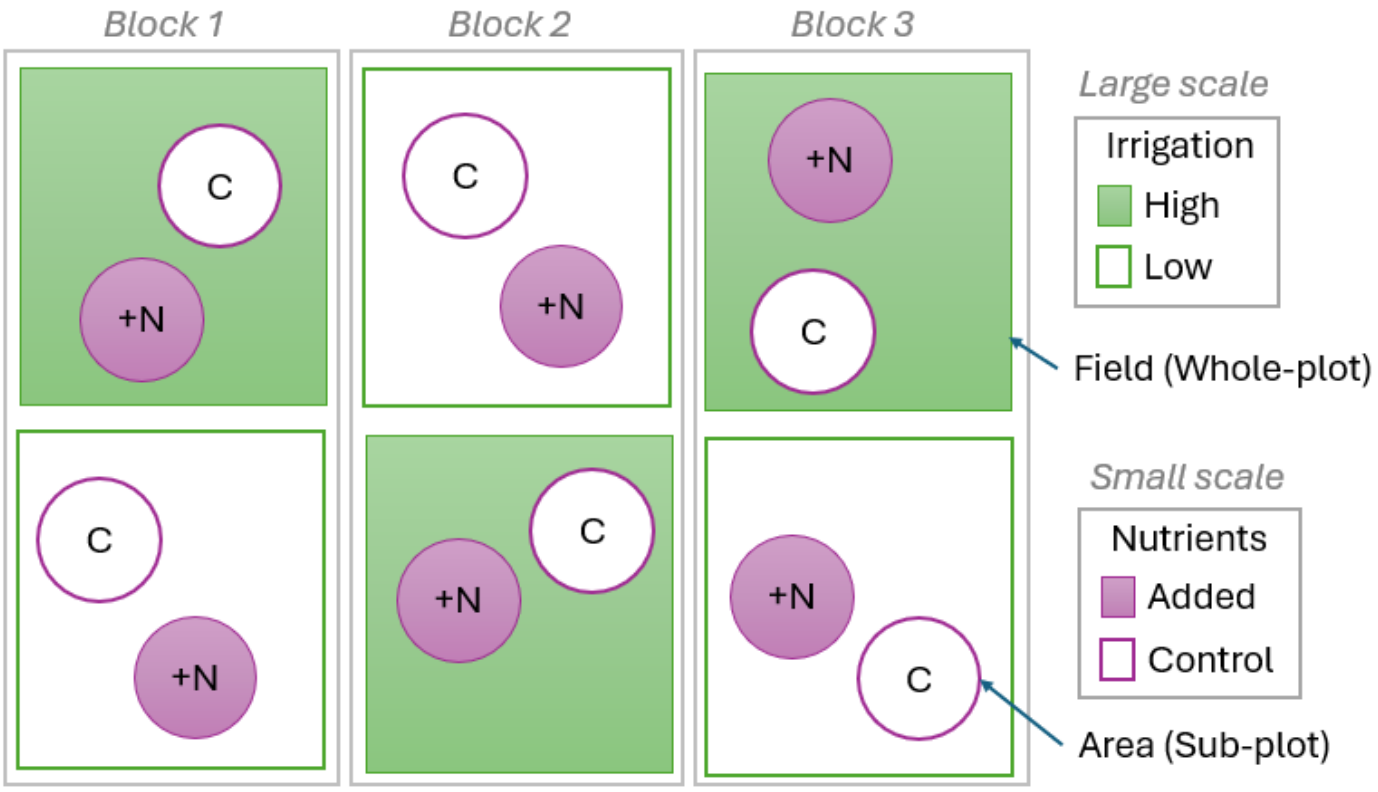


Fig. 8.3. Schematic diagram of a split-plot design, where irrigation levels are administered at a large scale (i.e., to whole-plots), and nutrient levels are administered at a small scale (i.e., to sub-plots).

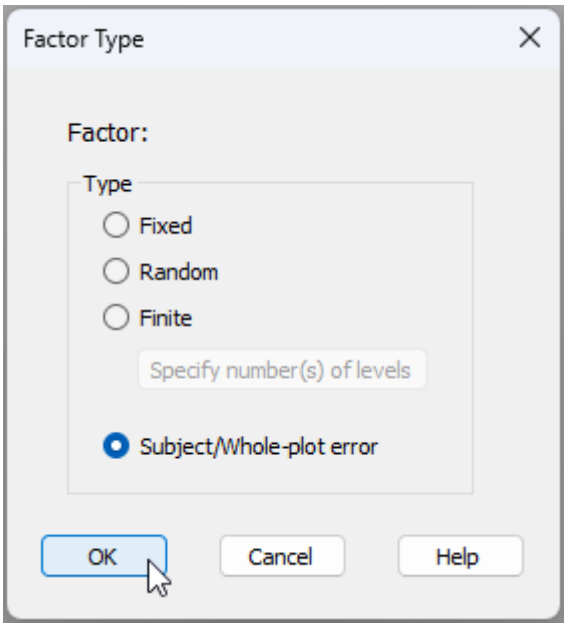
In such cases, we may think of the design as having two different 'error' terms to consider: one at the level of 'whole-plots' (the fields in this example) and one at the level of sub-plots. If we look just at the 'irrigation' factor in this example, it appears precisely like a randomised block design at a large scale: there are 3 blocks and the two irrigation treatments are applied to two large-scale whole-plots in each block. Thus, when we test the effects of irrigation, the variability from one whole-plot to another is the appropriate 'error' term to consider. In turn, when we investigate the effects of nutrients, we clearly need to consider the variability from one sub-plot to another as our 'error' for that test.

A proposed rationale for using a split-plot design is that factors may occur naturally at different scales (e.g., Mead (1988)). Another proposed rationale is that one may already know that factor A (at a large scale) has important effects, and one may be willing to sacrifice information on factor A

to get more precise results for factor B and the interaction A $\times$ B. For more information regarding the assumptions and potential disadvantages of split-plot designs, see Mead (1988) and Underwood (1997) .

New factor type: Subject/Whole-plot

The new PERMANOVA routine in PRIMER 8 caters well to repeated measures and split-plot designs directly. Specifically, there is a new 'Factor Type' called 'Subject/Whole-plot error', as shown below:



This is ideal for situations where entities (plots, individuals, subjects) are sampled multiple times (as in repeated measures), or cases where different treatments are administered at different spatial scales (e.g., to whole-plots and sub-plots in split-plot designs) in a way that lacks replication (at whatever scale). These types of study designs represent special cases where there may be insufficient replication to estimate all of the potential interactions among factors genuinely implied by the full set of factors in the study.

Previous versions of PERMANOVA for PRIMER would handle cases lacking replication within the highest-order cells by issuing a warning and excluding/removing the highest-order interaction (as indicated above for the randomised block and repeated measures cases). This strategy, however, does *not* handle situations where the confounding occurs somewhere else in the study design (e.g., at larger spatial scales), as it typically does for a classical split-plot design. In such cases, historically, when using PERMANOVA (in version 6 or version 7 of PRIMER), the end-user had to figure out which terms (if any) should be excluded from the model, and then would have had to exclude them manually (using the 'Terms' button). Failure to do this could result in certain terms in the model appearing with the words 'No test' in the output.

Basically, PERMANOVA in PRIMER 7 does not directly cater to the situation where you have (effectively) a *nested term that also lacks replication* and that occurs *at a different level* in the design (e.g., at the level of 'whole plots'). The specification of a whole-plot (or subject) type of factor can be thought of like the specification of an additional error term that occurs at a larger spatial scale than the scale of individual replicates.

Thankfully, the new PERMANOVA routine in PRIMER 8 handles these types of factors and designs easily, directly and correctly.

8.2 Example: Split-plot - Woodstock vegetation

The study design

An example of a split-plot design is provided by a study of the effect of fire disturbance and grazers (excluded using fences) on the composition of plant assemblages on the central western slopes of New South Wales in south-eastern Australia (Prober *et al.* (2007)). The design (shown schematically in Fig. 8.4) included the following:

- Blocks (random with r = 4 levels).
- Factor A: Fire frequency (fixed with a = 4 levels: 0 yrs, 2 yrs, 4 yrs, or 8 yrs).
- Whole plots (random and nested within Blocks and Fire frequency, unreplicated).
- Factor B: Fencing (fixed with b = 2 levels: fenced or unfenced).
- Sub-plots (random and nested within all of the above, unreplicated).

The two fencing treatments were randomly allocated to two sub-plots (measuring 5 m $\times$ 5 m) within each fire treatment (whole plots) and there is one of each fire treatment (4 whole plots) randomised within each block. Relative abundances (cover) of higher plant species within each sub-plot were estimated using a point-intercept technique (an 8 mm dowel placed vertically at each of 50 points on a grid across each plot). Although the design was set up at each of two locations (Woodstock and Monteagle) and data were obtained over a number of years (see Prober *et al.* (2007) for more details), we consider here only data from the Woodstock location collected in 2003.

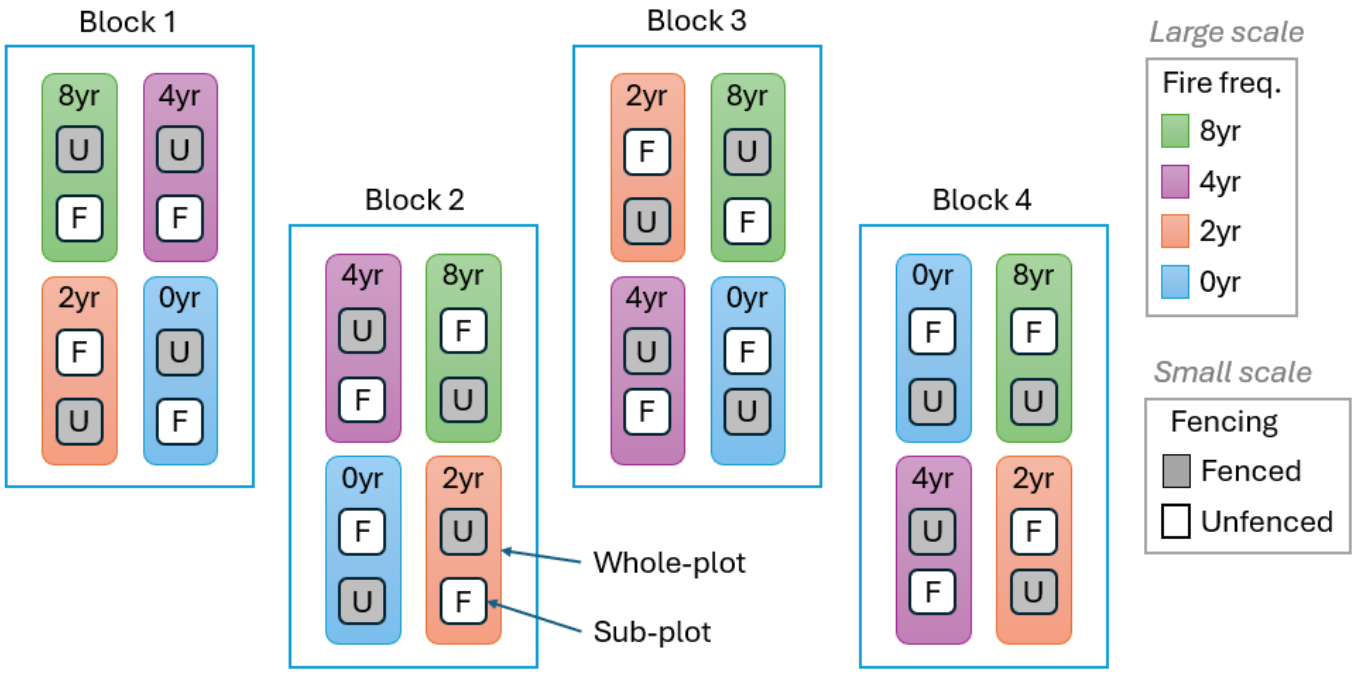
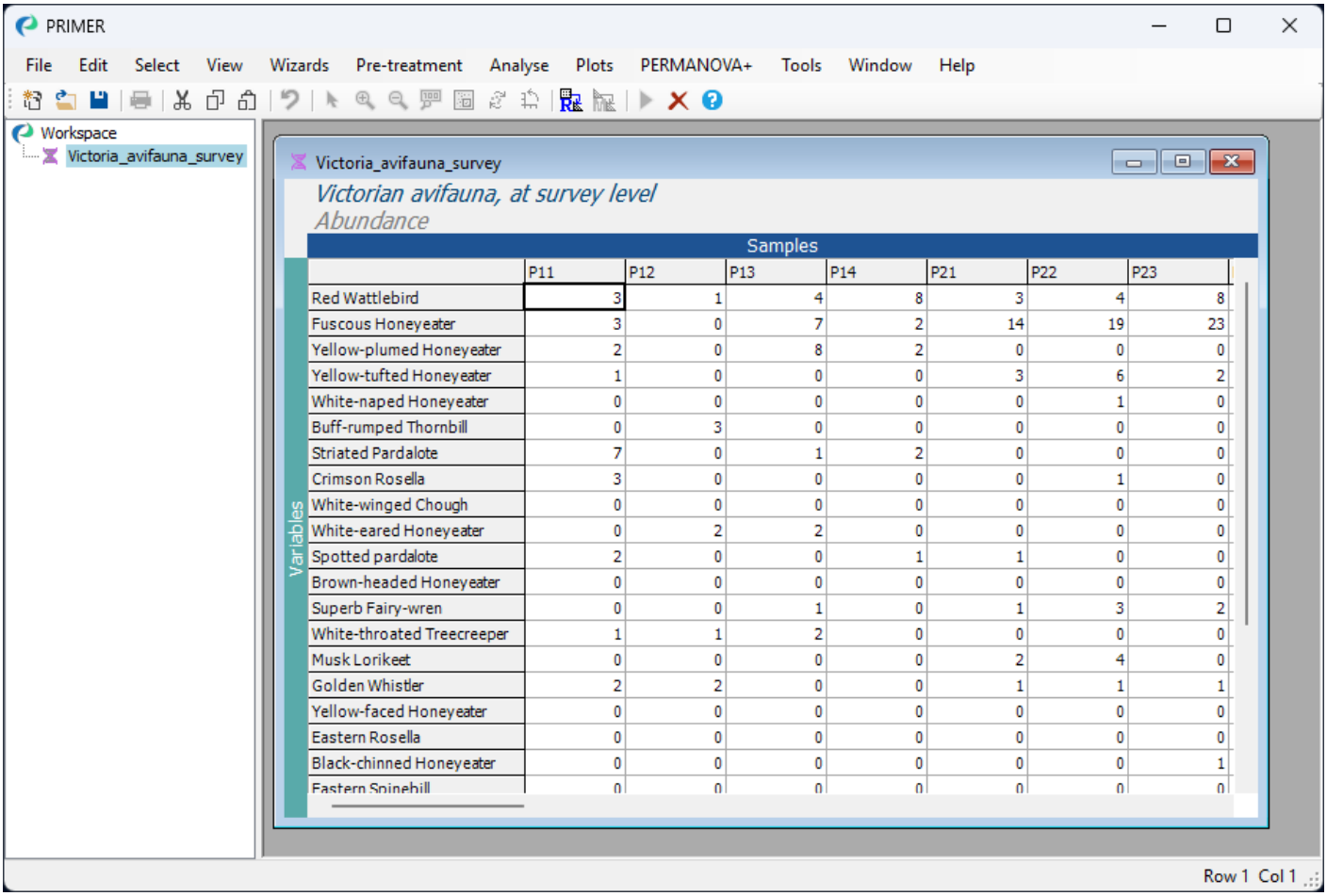


Fig. 8.4. Schematic diagram of the Woodstock split-plot design examining the potential effects of fire frequency and grazers (excluded using fences) on plant assemblages.

For the multivariate analysis, we also will exclude two species from the plant assemblage: *Poa sieberiana* (grey tussock grass) and *Themeda australis* (kangaroo grass), both dominant grasses, from the analysis. These were analysed separately in detail by Prober *et al.* (2007) ; our focus here instead will be on the more subtle potential responses of subsidiary forbs and exotic species.

Input data and select variables

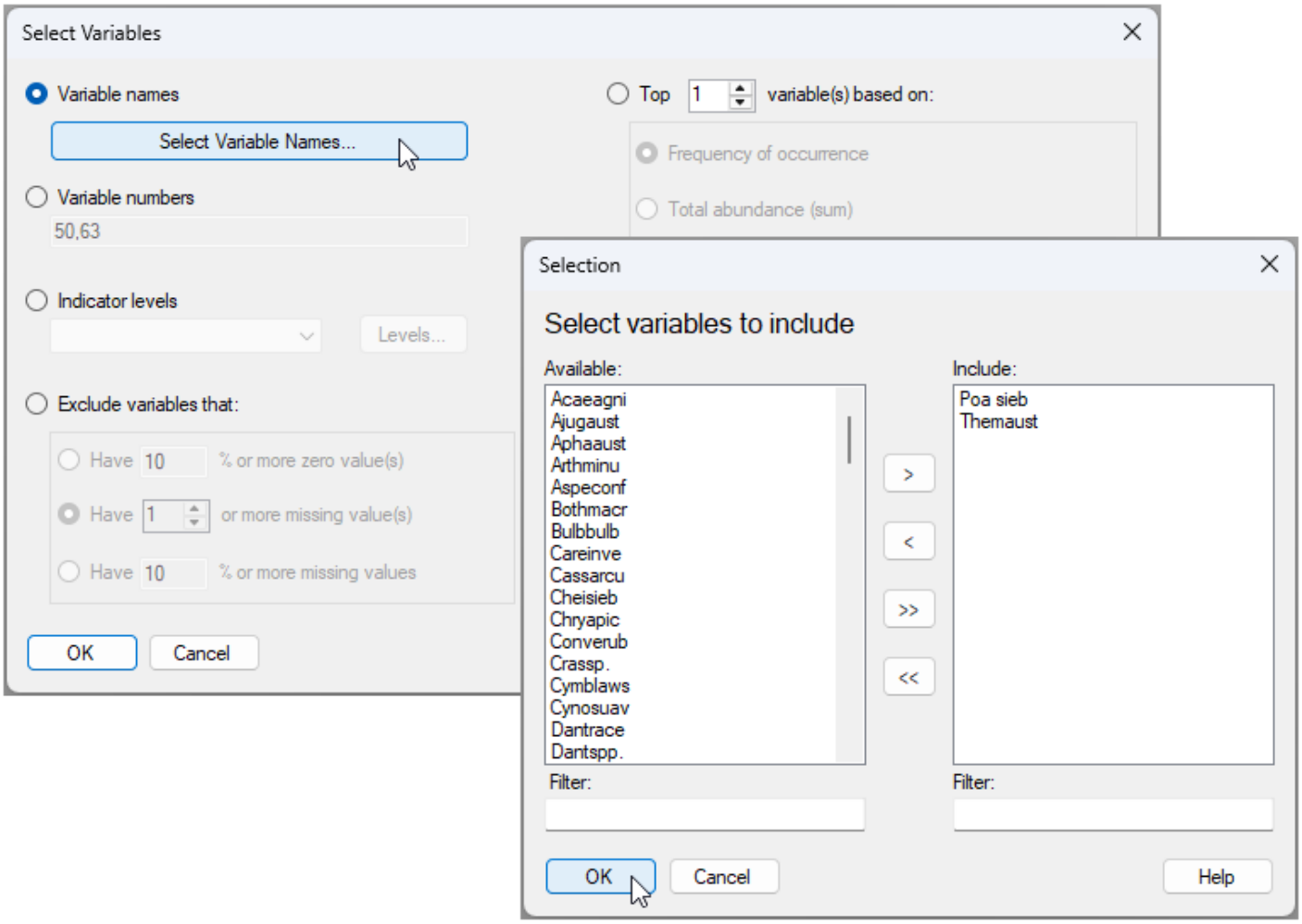
1. Start running **PRIMER 8**, then click **File** > **Open...** to open the data file named ' Woodstock_grassland.pri' (found inside the 'Examples_P8 > Woodstock_grassland' folder).



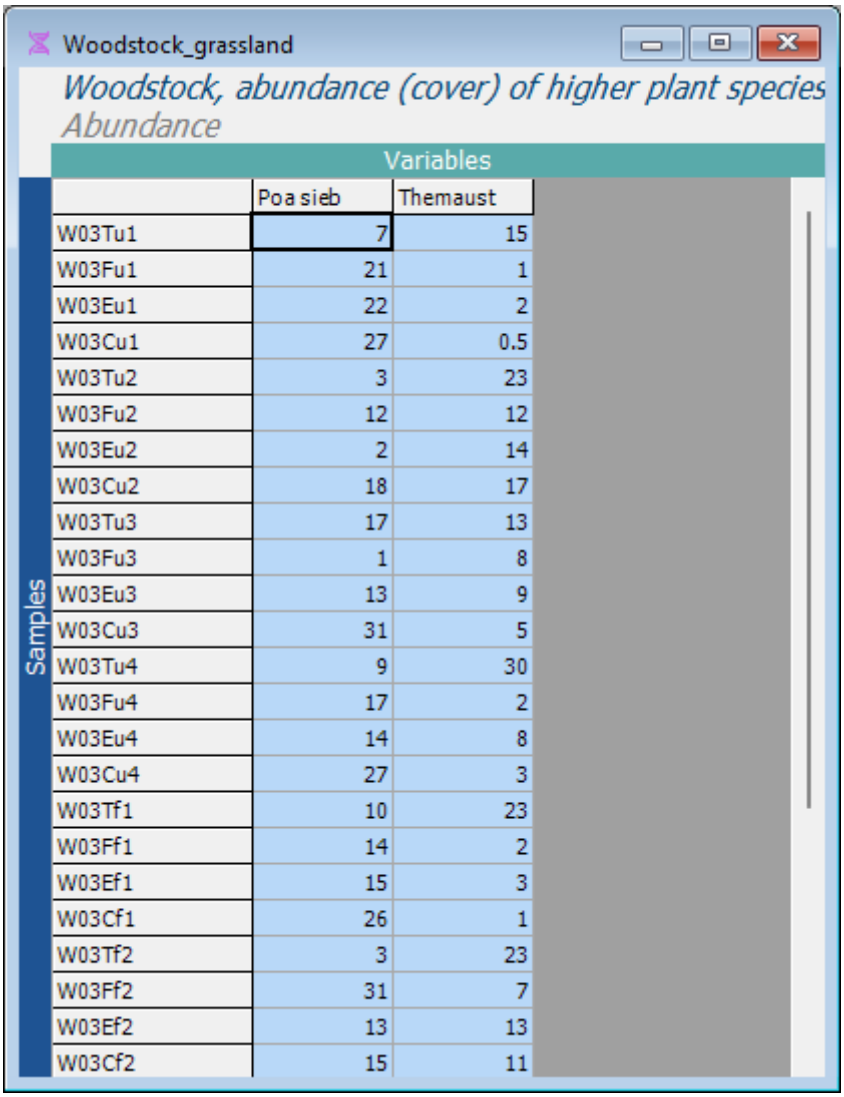
	P11	P12	P13	P14	P21	P22	P23
Red Wattlebird	3	1	4	8	3	4	8
Fuscous Honeyeater	3	0	7	2	14	19	23
Yellow-plumed Honeyeater	2	0	8	2	0	0	0
Yellow-tufted Honeyeater	1	0	0	0	3	6	2
White-naped Honeyeater	0	0	0	0	0	1	0
Buff-rumped Thornbill	0	3	0	0	0	0	0
Striated Pardalote	7	0	1	2	0	0	0
Crimson Rosella	3	0	0	0	0	1	0
White-winged Chough	0	0	0	0	0	0	0
White-eared Honeyeater	0	2	2	0	0	0	0
Spotted pardalote	2	0	0	1	1	0	0
Brown-headed Honeyeater	0	0	0	0	0	0	0
Superb Fairy-wren	0	0	1	0	1	3	2
White-throated Treecreeper	1	1	2	0	0	0	0
Musk Lorikeet	0	0	0	0	2	4	0
Golden Whistler	2	2	0	0	1	1	1
Yellow-faced Honeyeater	0	0	0	0	0	0	0
Eastern Rosella	0	0	0	0	0	0	0
Black-chinned Honeyeater	0	0	0	0	0	0	1
Eastern Spinebill	0	0	0	0	0	0	0

We want to select all variables *except* the two dominant grasses. We will first find these two variables in the data file, highlight them and then *invert* that highlighting in order to select the *remaining* species. The two dominant grasses *Poa sieberiana* and *Themeda australis* are named ' Poa sieb' and 'Themaus' in the data file, respectively.

2. Click **Select** > **Variables...** > ($\bullet$ Variable names), then click on the button 'Select Variable Names...'. In the resulting dialog, start typing 'Poa' in the 'Filter:' box under the 'Available' box on the left. When you find 'Poa sieb', you can click on its name, then click the right arrow > to move this over to the 'Include' box on the right. Repeat the same operation to find and include 'Themaus'. Once both variables you want to select are in the 'Include' box, click '**OK**', then click '**OK**' in the 'Select Variables' dialog window.



The data sheet with only these 2 species selected will look blue in colour, like this:



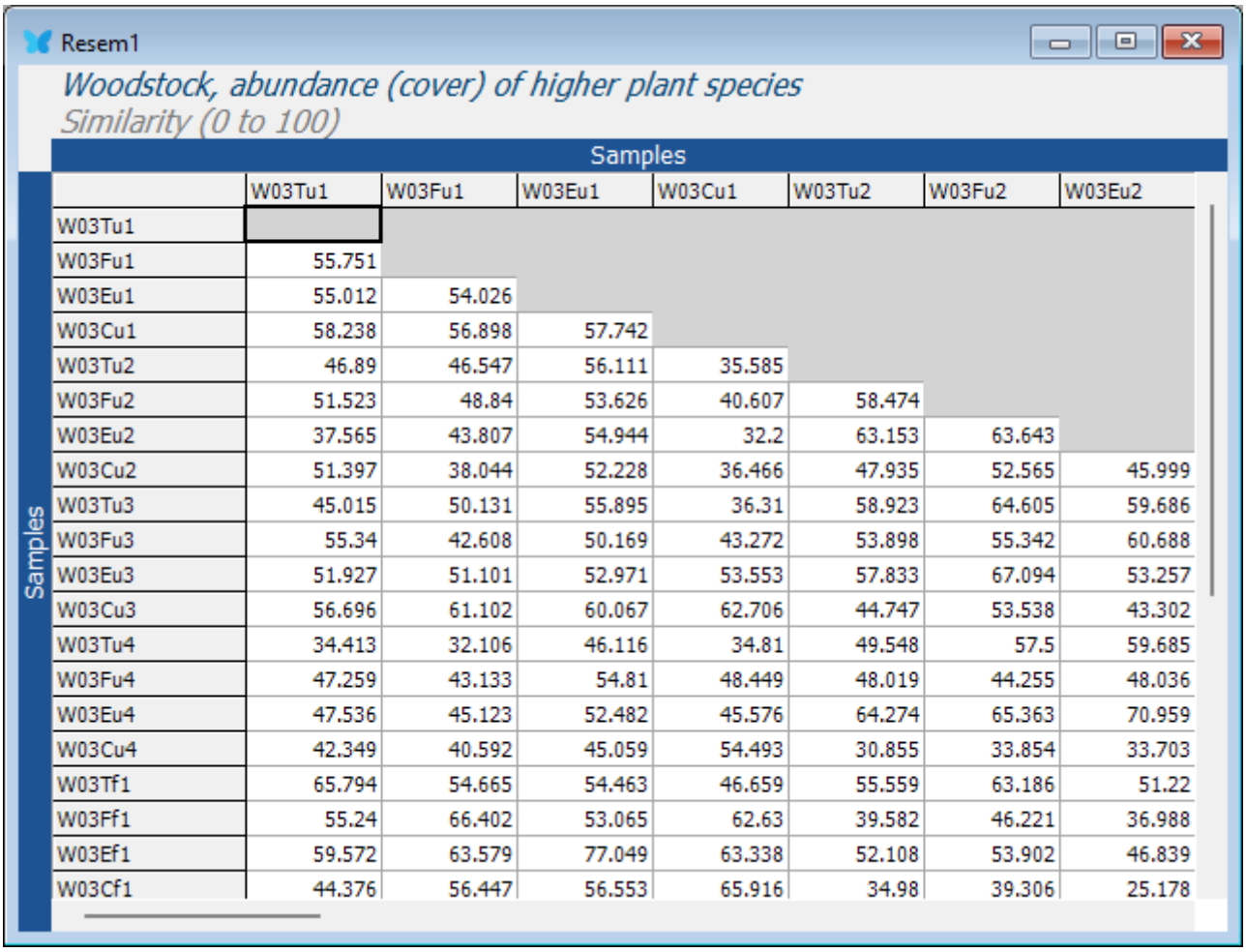
3. From the data sheet having only these 2 species selected, perform the following steps:
   - click **Select** > **All**. This will show the full spreadsheet of all species once again, but now those two species will just be *highlighted* (appearing in orange / pale yellow) within the sheet.
   - click **Edit** > **Invert Highlight**. Now all of the species *except* 'Poa' and 'Themaus' (and all of the rows) will be highlighted.
   - click **Select** > **Highlighted**. Now all of the species *except* 'Poa' and 'Themaus' will be selected (hence blue).
   - click **Tools** > **Duplicate**, to produce a new separate data sheet (which now omits those two grass species) called 'Data1'.

Calculate the resemblance matrix

For analysis of composition of the assemblage, we will apply a square-root transformation, followed by the Bray-Curtis resemblance measure.

4. From the 'Data1' data sheet, click **Pre-treatment** > **Transform(overall)...** and choose 'Square root' from the drop-down menu, then click **OK**. This will generate a new data sheet item containing the transformed data, called 'Data2'.

5. From the square-root transformed data ('Data2'), click **Analyse** > **Resemblance...** and in the 'Resemblance' dialog window, choose (Measure: $\bullet$Bray-Curtis similarity) and (Analyse between: $\bullet$Samples), then click '**OK**'. This will yield a resemblance matrix item in the explorer tree, called 'Resem1'.



Resem1

Woodstock, abundance (cover) of higher plant species

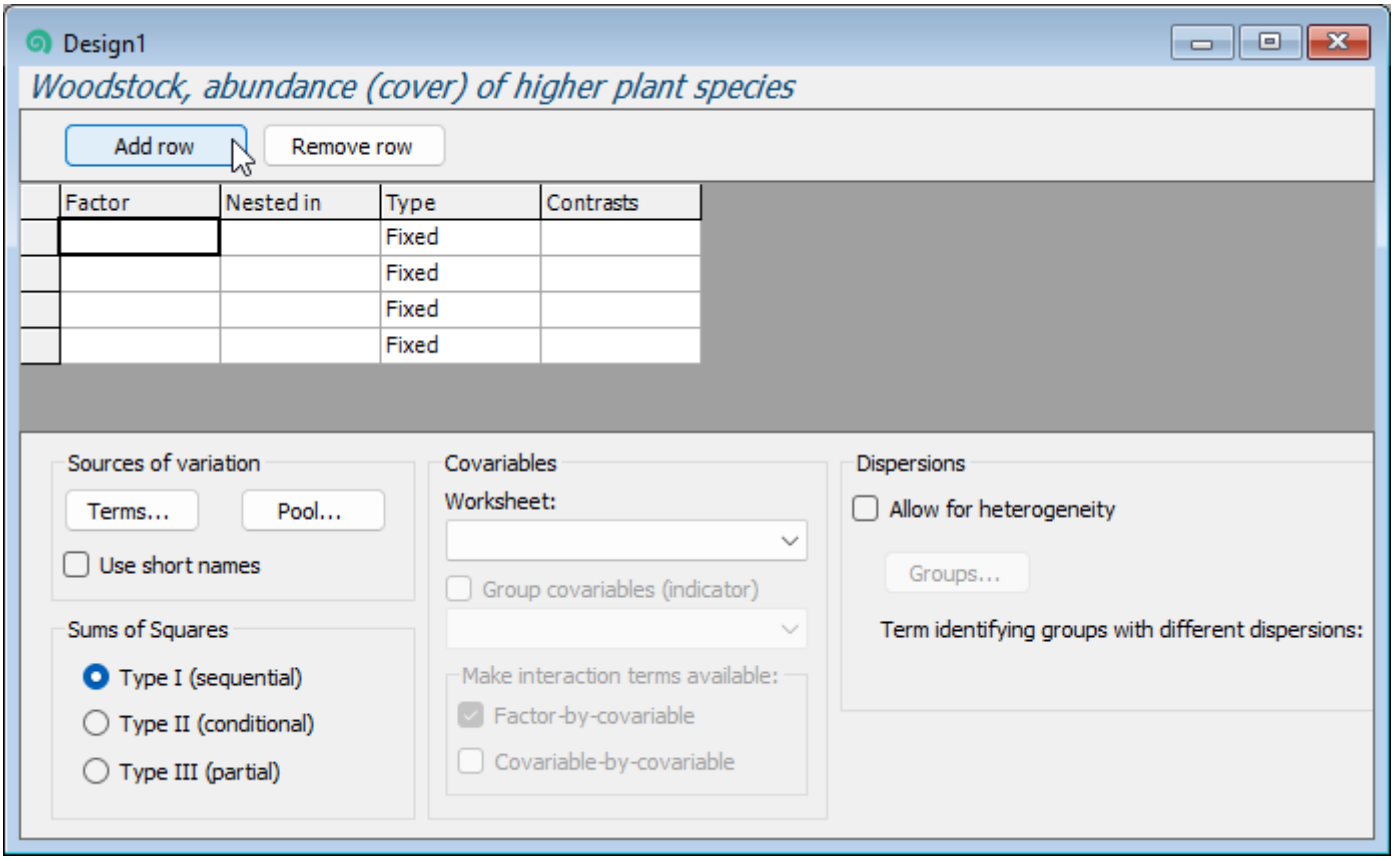
Similarity (0 to 100)

Samples	W03Tu1	W03Fu1	W03Eu1	W03Cu1	W03Tu2	W03Fu2	W03Eu2
W03Tu1							
W03Fu1	55.751						
W03Eu1	55.012	54.026					
W03Cu1	58.238	56.898	57.742				
W03Tu2	46.89	46.547	56.111	35.585			
W03Fu2	51.523	48.84	53.626	40.607	58.474		
W03Eu2	37.565	43.807	54.944	32.2	63.153	63.643	
W03Cu2	51.397	38.044	52.228	36.466	47.935	52.565	45.999
W03Tu3	45.015	50.131	55.895	36.31	58.923	64.605	59.686
W03Fu3	55.34	42.608	50.169	43.272	53.898	55.342	60.688
W03Eu3	51.927	51.101	52.971	53.553	57.833	67.094	53.257
W03Cu3	56.696	61.102	60.067	62.706	44.747	53.538	43.302
W03Tu4	34.413	32.106	46.116	34.81	49.548	57.5	59.685
W03Fu4	47.259	43.133	54.81	48.449	48.019	44.255	48.036
W03Eu4	47.536	45.123	52.482	45.576	64.274	65.363	70.959
W03Cu4	42.349	40.592	45.059	54.493	30.855	33.854	33.703
W03Tf1	65.794	54.665	54.463	46.659	55.559	63.186	51.22
W03Ff1	55.24	66.402	53.065	62.63	39.582	46.221	36.988
W03Ef1	59.572	63.579	77.049	63.338	52.108	53.902	46.839
W03Cf1	44.376	56.447	56.553	65.916	34.98	39.306	25.178

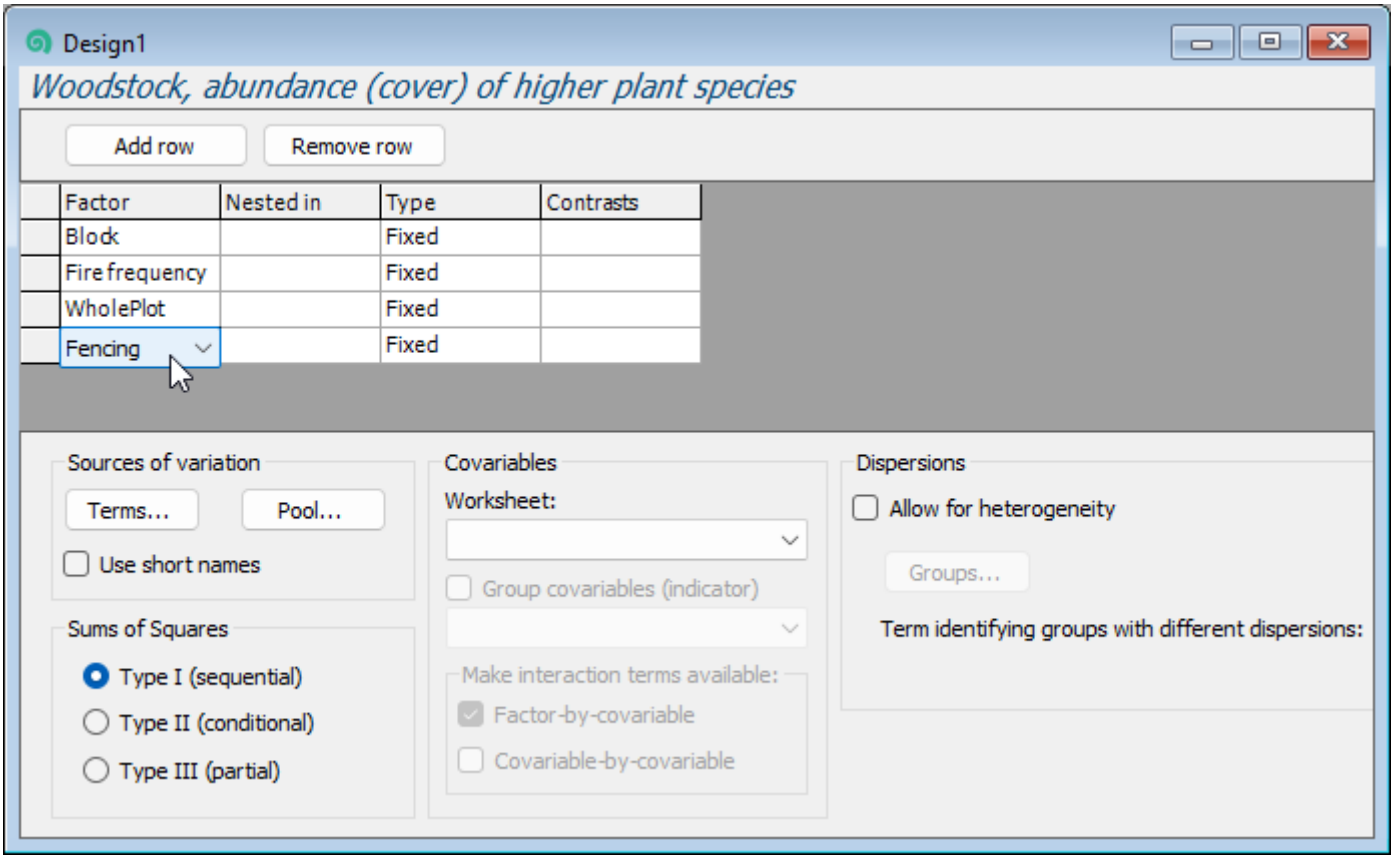
Create the design file

Recall that running a PERMANOVA will require us first to set up a design file in accordance with our study's experimental/sampling design.

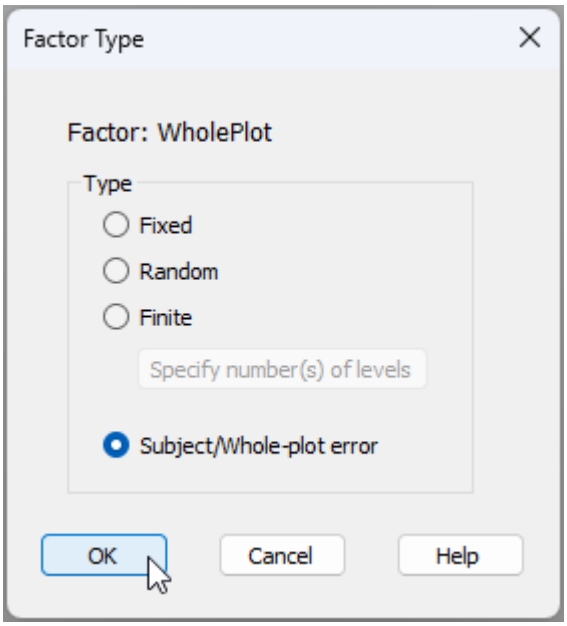
6. From the 'Resem1' matrix, click **PERMANOVA+** > **Create PERMANOVA Design...**, then, in the resulting design file (called 'Design1'), click three times on the button to add a row (Add row) so that there are four rows in total, as shown below.



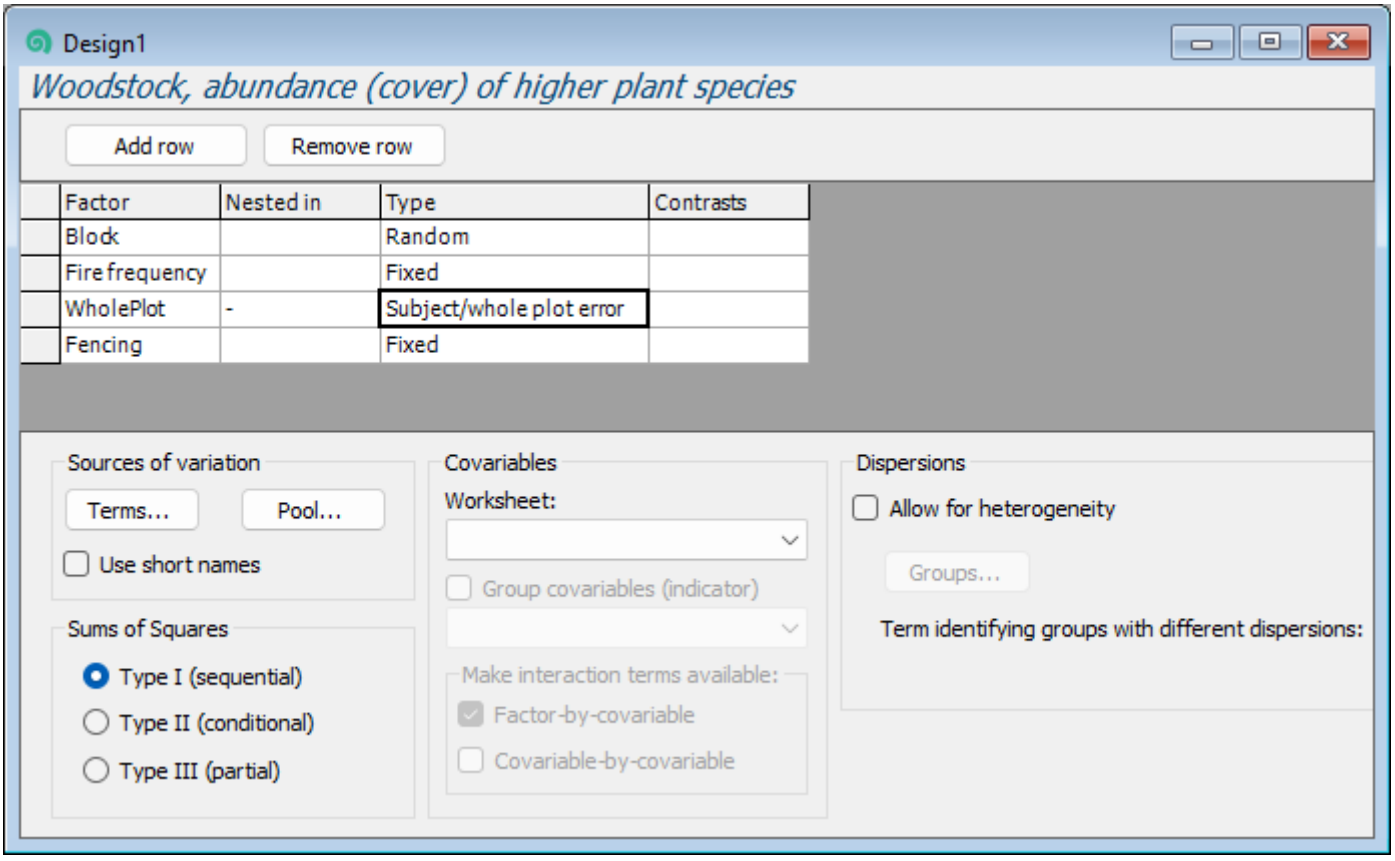
7. **Choose the name of the factor you want in each row** - In this design file (called 'Design1'), double-click inside each cell of the first column (headed 'Factor') in order to choose the following factors so that they occur sequentially in rows 1, 2, 3 and 4, respectively: row 1 = 'Block', row 2 = 'Fire frequency', row 3 = 'WholePlot', and row 4 = ' Fencing', as shown below:



8. **Input the factor types** - Next, in column 3, we need to specify the 'Type' of each factor. By default, everything begins by being listed as 'Fixed'. To change this for any given factor, we double-click on the word 'Fixed' in its respective row and the 'Factor Type' dialog will pop up. For the present design, we are happy to treat the factors of 'Fire frequency' and 'Fencing' as 'Fixed', but we need to specify carefully that:
   - 'Block' is of Type '$\bullet$ Random', and
   - 'WholePlot' is of Type '$\bullet$ Subject/Whole-plot error', like so:



Having done that, the final correct design file (called 'Design1') for this split-plot study will look like this:

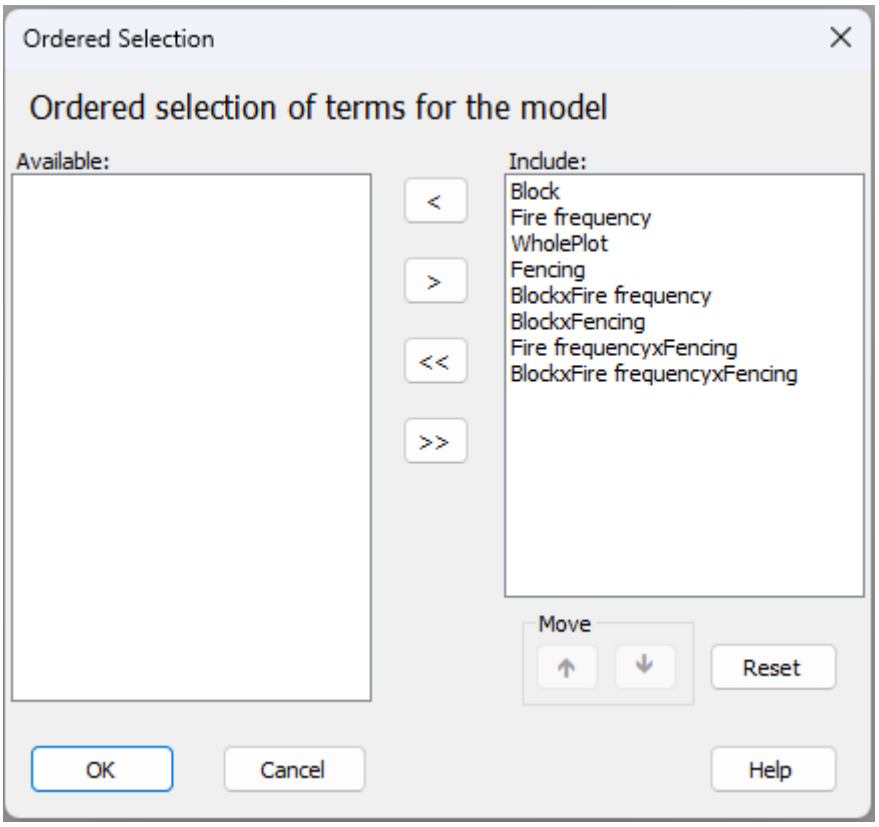


Further design considerations

Before we run the analysis, it is worth pointing out a couple of things about this design file.

First, you will notice that there is a dash ('-') in the 'Nested in' column for the factor of 'WholePlot'. This is because whole-plots, having been identified as such, are deemed effectively to contribute an 'error' into the study design, which means they are necessarily nested within all of the broad-scale factors (in the present case, the broad-scale factors are 'Block' and 'Fire frequency'). Having specified that 'WholePlots' are of Type 'Subject/Whole-plot error', it is not necessary to also articulate this nested structure as well. The dash ('-') merely indicates this column has been 'taken care of' internally, so-to-speak, for the whole-plot factor.

Second, if you click on the 'Terms...' button (Terms...), you will see (on the right-hand side, in a box under the word 'Include:') a list of all of the terms included (by default) in the full PERMANOVA model implied by the design that you have specified.¶ In the present case, it looks like this:



What do we know about this design already?

- (i). The 'Block$\times$Fire frequency' interaction term is not actually estimable, because there is no replication of the Fire-frequency treatments in each block.
- (ii). The 'Block$\times$Fire frequency$\times$Fencing' interaction term is also inestimable, because there is no replication of the two fencing treatments in each whole-plot.
- (iii). The 'Block$\times$Fencing' interaction term would not typically be included in a classical analysis of a split-plot design like this. This is primarily because blocks are set up at large scales and the fencing treatment is applied at small scales.

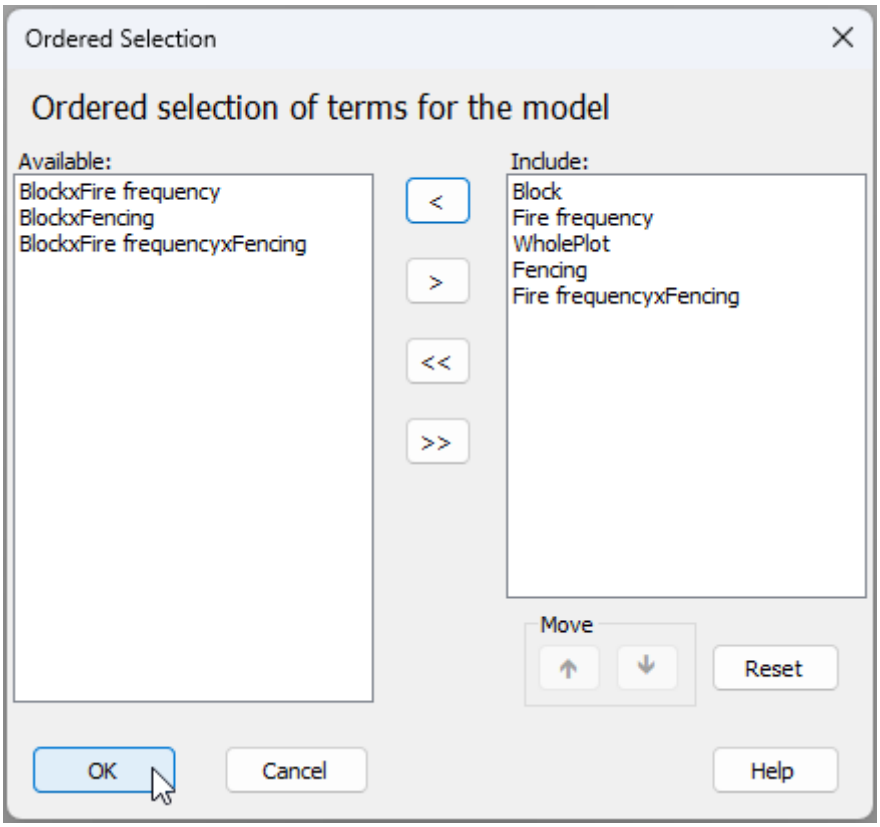
It turns out that, classically, *none* of the interactions involving the factor of 'Block' would be included in the ANOVA partitioning for this split-plot design. The sources of variation and degrees of freedom for this example, according to a traditional split-plot ANOVA partitioning, are shown in Table 8.1

Table 8.1. Sources of variation and degrees of freedom ($\textit{df}$) for a classical split-plot design, where factor A = 'Fire frequency' and factor B = 'Fencing', as per the Woodstock example. Note that 'Whole-plot total' is not an additional source of variation, but rather corresponds to the total variation at the broad scale, i.e., among all of the whole plots, which may be relevant for terms in the top half of the table only. (The line corresponding to 'Whole-plot total' can simply be omitted).

Source	**$\textit{df}$**
Block	$(r-1) = 3$
Fire frequency	$(a-1)=3$
Whole-plot error	$(r-1)(a-1) = 9$
Whole-plot total	$(ra-1) = 15$
----------	----------
Fencing	$(b-1) = 1$
Fire frequency$\times$Fencing	$(a-1)(b-1) = 3$
Sub-plot error	$a(b-1)(r-1) = 12$
Total	$abr - 1 = 31$

If we run the PERMANOVA directly, without manually removing any of the interactions involving the factor of 'Block', then both (i) 'Block$\times$Fire frequency' and (ii) 'Block$\times$Fire frequency$\times$Fencing' interaction terms will be removed automatically in the PERMANOVA run anyway, as these two terms are inestimable. However (in this example), the (iii) 'Block$\times$Fencing' interaction term will remain in the model, simply because it is (technically) estimable.

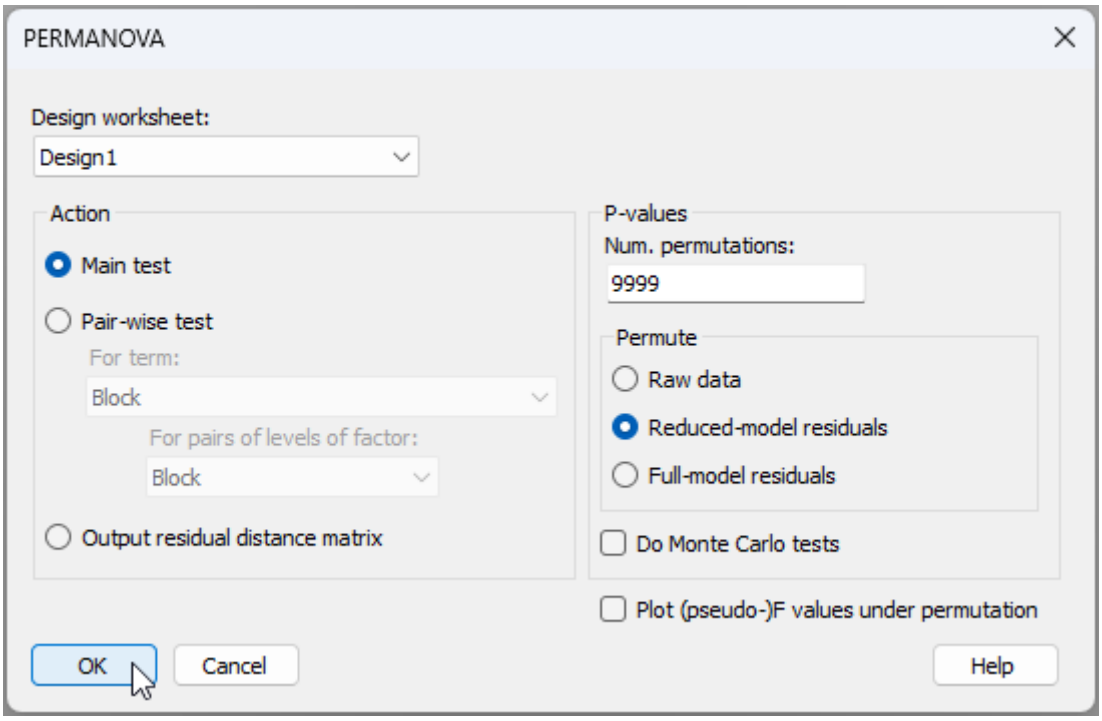
9. **Remove unwanted interaction terms** *(optional)* - If you desire the classical split-plot design, you may wish to manually remove all of the interaction terms involving the factor of 'Block'.‡ To do this, click on the 'Terms' button in the design file and (sequentially) click on each of the terms that you wish to omit from the model, then click on the left arrow (<) to move it over to the 'Available' column. The result (if you do choose to remove all of the interactions involving the factor of 'Block') will look like this:



Click '**OK**' and now the design file is all ready for the split-plot analysis.

Run the PERMANOVA

10. From the original resemblance matrix 'Resem1', click **PERMANOVA+** > **PERMANOVA...**. Ensure that the design worksheet is 'Design1', leave all other defaults, and click '**OK**'.



The resulting output file (called 'PERMANOVA1') looks like this:

PERMANOVA1

PERMANOVA table of results

Source	df	SS	MS	Pseudo-F	P(perm)	Unique perms
Block	3	9891.5	3297.2	2.8394	0.0005	9905
Fire frequency	3	7244.2	2414.7	2.0794	0.0071	9914
WholePlot	9	10451	1161.2	2.0014	0.0001	9835
Fencing	1	840.46	840.46	1.4485	0.1349	9931
Fire frequencyxFencing	3	2252.1	750.69	1.2938	0.1136	9872
Res	12	6962.6	580.22			
Total	31	37642				

Estimates of components of variation

Source	Estimate	Sq.root
V(Block)	266.99	16.34
S(Fire frequency)	156.69	12.517
V(WholePlot)	290.51	17.044
S(Fencing)	16.265	4.033
S(Fire frequencyxFencing)	42.617	6.5282
V(Res)	580.22	24.088

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

Source	EMS
Block	1*V(Res) + 2*V(WholePlot) + 8*V(Block)
Fire frequency	1*V(Res) + 2*V(WholePlot) + 8*S(Fire frequency)
WholePlot	1*V(Res) + 2*V(WholePlot)
Fencing	1*V(Res) + 16*S(Fencing)
Fire frequencyxFencing	1*V(Res) + 4*S(Fire frequencyxFencing)

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

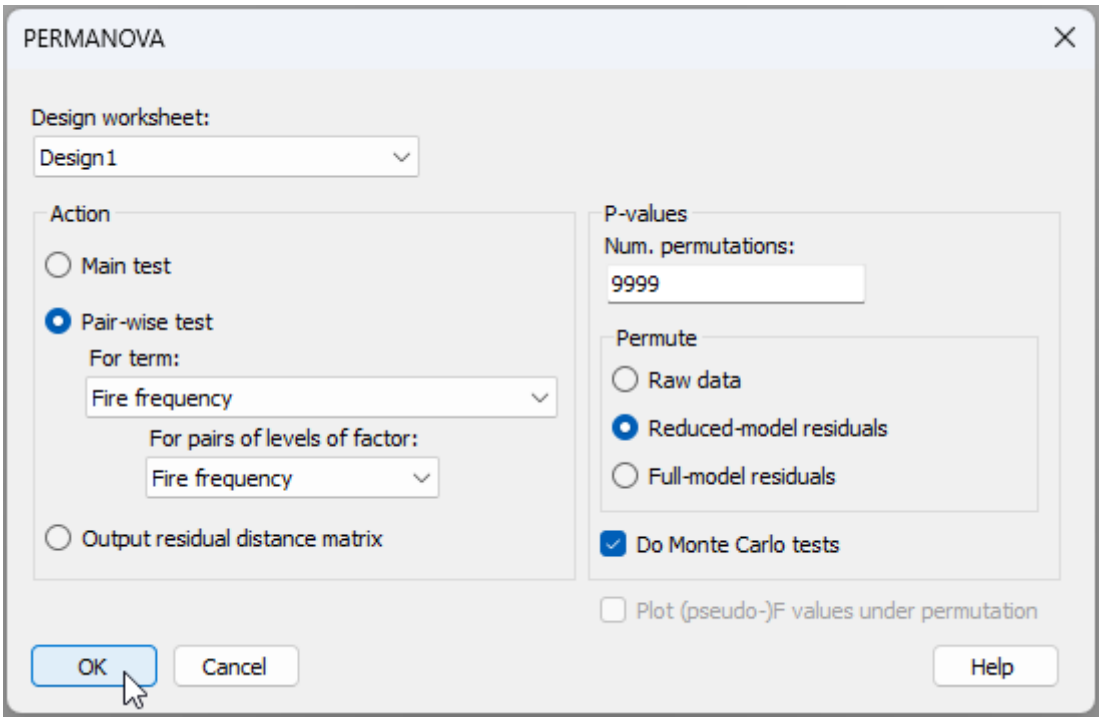
Source	Numerator	Denominator	Num.df	Den.df
Block	1*Block	1*WholePlot	3	9
Fire frequency	1*Fire frequency	1*WholePlot	3	9
WholePlot	1*WholePlot	1*Res	9	12
Fencing	1*Fencing	1*Res	1	12
Fire frequencyxFencing	1*Fire frequencyxFencing	1*Res	3	12

There is a significant effect of fire frequency on the composition of these plant assemblages ($F_{3,9}$ = 2.08, P < 0.01). There was not, however, a significant effect of fencing to exclude grazers ($F_{1,12}$ = 1.45, P > 0.10), nor did fencing treatments change the overall effect of fire frequency (the Fire frequency$\times$Fencing interaction term was not significant; $F_{3,12}$ = 1.29, P > 0.10).

Run pair-wise comparisons

A natural next step might be to examine pairwise comparisons among the whole-plots having different fire frequencies.

11. To run pairwise tests, start from the original resemblance matrix ('Resem1') and click **PERMANOVA+** > **PERMANOVA...**. Ensuring, once again, that the design worksheet is 'Design1', under 'Action', choose: ($\bullet$Pair-wise test > (For term: Fire frequency > (For pairs of levels of factor: Fire frequency))). You might also (optionally) tick the option to '$\checkmark$Do Monte Carlo tests', simply because there are not very many whole plots to permute at the large spatial scale. Leaving everything else as the defaults, click '**OK**', *viz*:



Results of the pairwise comparisons are shown below ('PERMANOVA2'):

PERMANOVA2

PAIR-WISE TESTS

Term 'Fire frequency'

Groups	t	P(perm)	Unique perms	P(MC)
2yr, 4yr	1.511	0.0951	425	0.0776
2yr, 8yr	0.87175	0.6859	425	0.6186
2yr, nil	1.7834	0.0666	425	0.0393
4yr, 8yr	1.2337	0.2234	425	0.2292
4yr, nil	1.4125	0.1385	425	0.103
8yr, nil	1.5732	0.081	425	0.0766

Denominators

Groups	Denominator	Den.df
2yr, 4yr	1*WholePlot	3
2yr, 8yr	1*WholePlot	3
2yr, nil	1*WholePlot	3
4yr, 8yr	1*WholePlot	3
4yr, nil	1*WholePlot	3
8yr, nil	1*WholePlot	3

Average Similarity between/within groups

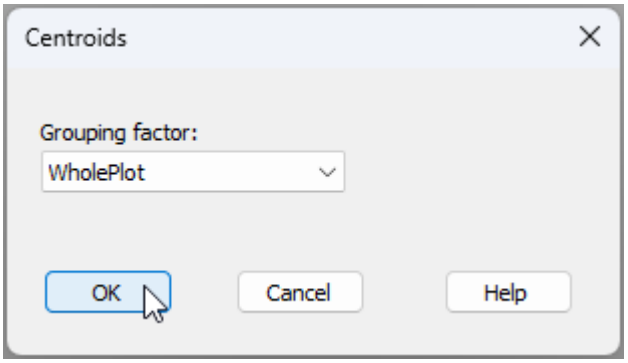
	2yr	4yr	8yr	nil
2yr	54.969			
4yr	50.554	52.879		
8yr	56.802	53.429	58.372	
nil	45.174	49.279	49.563	51.08

Despite the significant overall test, individual treatment levels are not detected as being significantly different from one another (at the 0.05 significance level) in the pair-wise tests.[†] This sort of thing can happen, as the overall (omnibus) F-test can have greater power, given its larger denominator degrees of freedom, compared to the individual pair-wise tests.

Ordination of whole-plot centroids

Given that fire frequency is assessed at the spatial scale of entire whole-plots (and fencing had no detectable effects), we might consider creating an ordination plot of (say) the whole-plot centroids, as a way to visualise the results.

12. From the original resemblance matrix ('Resem1'), we can calculate distances among whole-plot centroids by clicking **PERMANOVA+** > **Distance Among Centroids...**, and choosing the 'Grouping factor' as 'WholePlot', as shown below:



This will produce a new resemblance matrix of Bray-Curtis resemblances among the centroids for the 16 whole-plots, caled 'Resem2':

Resem2

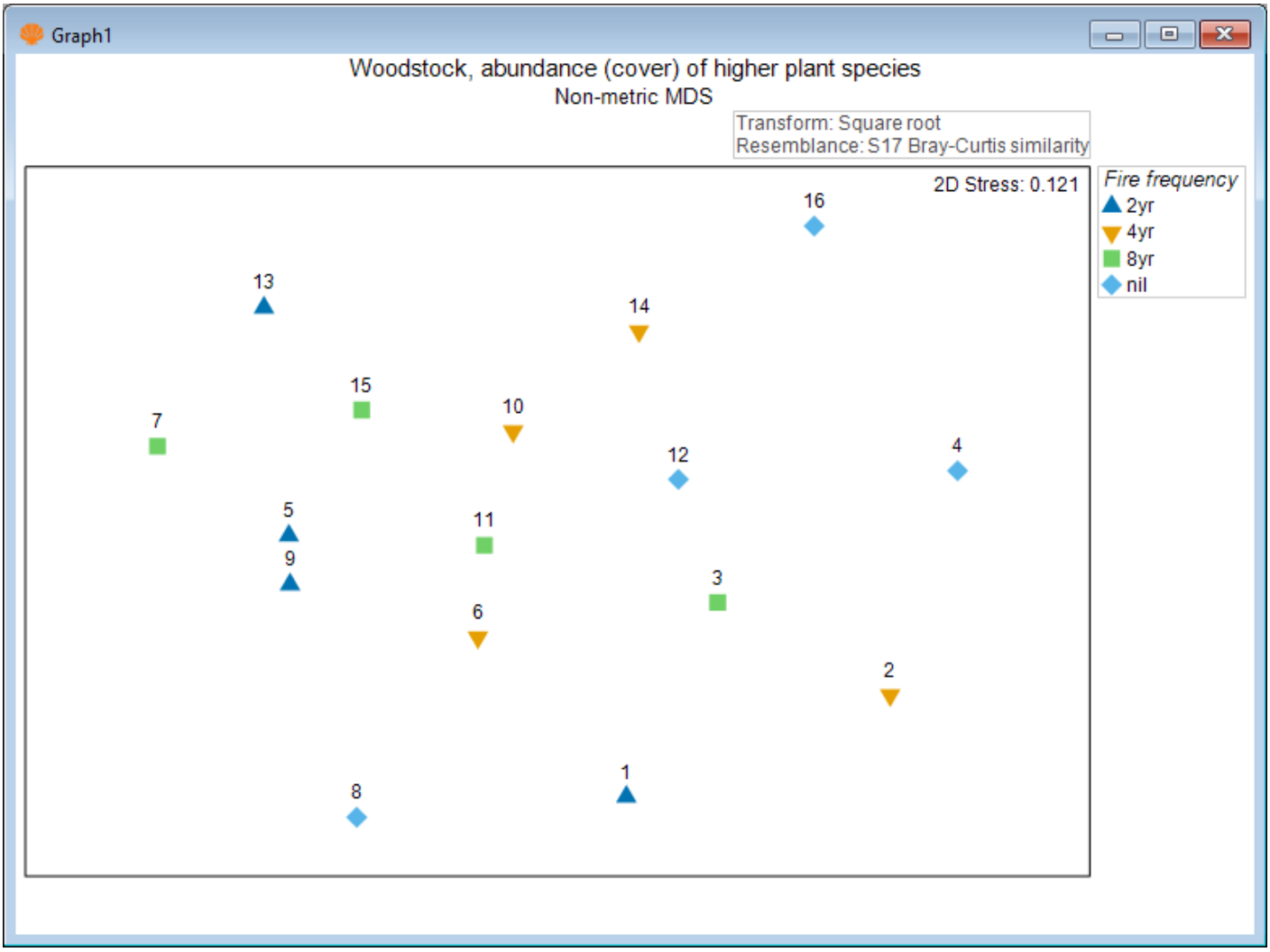
Woodstock, abundance (cover) of higher plant species

Distance (0 to inf)

Samples	1	2	3	4	5	6	7	8
1								
2	37.448							
3	38.464	36.389						
4	46.072	33.744	36.254					
5	40.759	53.7	40.684	59.238				
6	37.993	41.159	39.558	49.225	34.674			
7	47.504	58.289	45.027	69.604	31.007	42.437		
8	39.109	53.232	42.214	57.13	42.879	41.725	42.812	
9	44.613	51.441	41.341	61.577	27.1	27.316	33.311	
10	39.426	45.637	38.622	42.61	39.444	38.942	41.598	
11	41.879	44.814	40.589	44.294	36.349	25.615	46.569	
12	40.061	38.411	35.594	34.491	38.202	35.743	48.856	
13	53.017	61.707	48.642	60.503	39.645	44.075	31.607	
14	39.798	40.044	36.716	41.334	45.319	39.731	45.831	
15	46.565	56.677	43.661	54.301	32.332	32.442	30.861	
16	50.298	49.922	40.889	36.55	57.081	53.356	60.298	

13. From the 'Resem2' matrix, produce a non-metric MDS plot by clicking **Analyse** > **MDS** > **Non-metric MDS (nMDS)...**. Take all the defaults in the 'Non Metric MDS' dialog and click '**OK**'. The resulting 2-dimensional nMDS plot (an item called 'Graph1' in the Explorer

tree) is shown below.



As indicated by the pair-wise tests, we do see that most of the assemblages corresponding to unburned whole-plots (light blue diamond-shaped symbols) tend to occur towards the left-hand side of the ordination plot, whereas most of the assemblages corresponding to whole-plots burned at the highest frequency of every 2 years (the dark blue triangle-shaped symbols) tend to occur towards the right-hand side. However, this is not a super clear split, and there is rather high variation among whole plots within any of these treatments. The patterns on this ordination are quite consistent with the rather marginal pairwise test results that we obtained using PERMANOVA. In other words, from a practical point of view (omitting the two dominant grasses) there may only be somewhat minor differences in these plant assemblages caused by fire frequency, if any, at least for the time-scales examined in this study.

¶*Interaction terms are implied between any factors that are crossed with one another. By 'implied', we mean that they are conceivable, although they may not be estimable in all cases. That will depend on there being adequate replication. An interaction is not implied, however, between a given factor and another factor within which it is nested. For example, If we have Factor A and Factor B nested wtihin Factor A, then the only two terms in this model are A and B(A). Factor B clearly cannot interact with Factor A; therefore no A$\times$B interaction term is implied.*

[†]*The closest we come to statistical significance at the 0.05 level is for the comparison of unburned plots (0 years) with plots burned most frequently, i.e., every 2 years. In that case, we have t = 1.78 and the Monte Carlo p-value is 0.038, although the permutation p-value is 0.063, so this is a pretty marginal result.*

[‡]*An alternative method to achieve the same result here would be to specify the 'Block' factor also as being of Type 'Subject/Whole-plot error'. This will automatically remove all interactions of the 'Block' factor with any other term(s) in the model.*

8.3 Example: Repeated measures - Victorian avifauna

The study design

An example of a repeated-measures sampling design (Fig. 8.5) is provided in a study of Victorian avifauna by Mac Nally & Timewell (2005) . The data consist of counts of p = 27 nectarivorous bird species at each of eight sites having different levels of flowering intensity within the Rushworth State Forest in Victoria, Australia. One pair of sites (S1, S2) had heavy flowering ('good' sites), another pair (S3, S4) had intermediate flowering ('medium' sites), and a third pair (S5, S6) had relatively little flowering ('poor' sites). Two sites (S7, S8) were near the good sites (called 'adjacent' sites) and these were also sampled to explore the potential for 'spill-over' effects. Sampling of the bird assemblages was done using a strip transect method and each of the 8 sites was sampled repeatedly at four different time points.

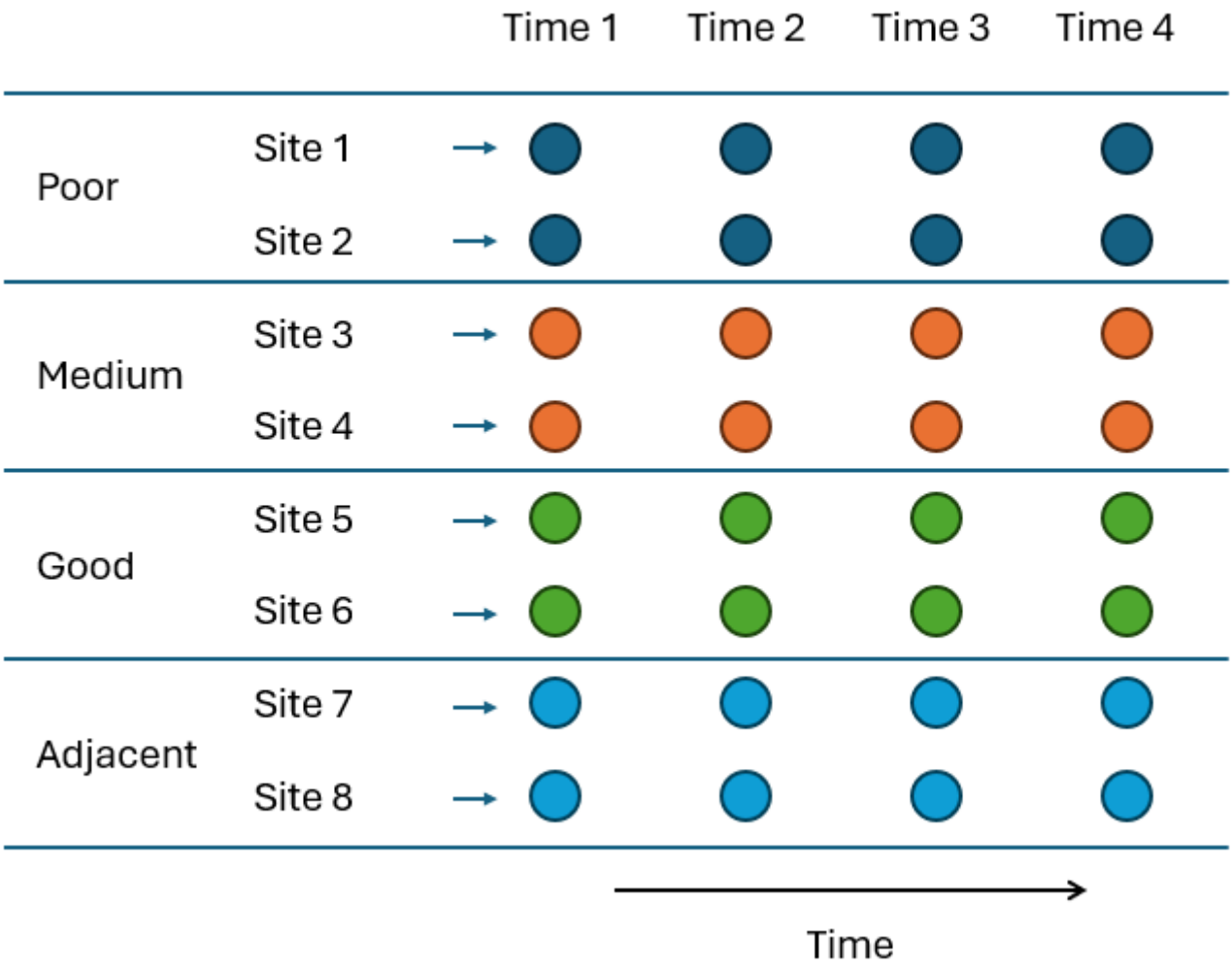
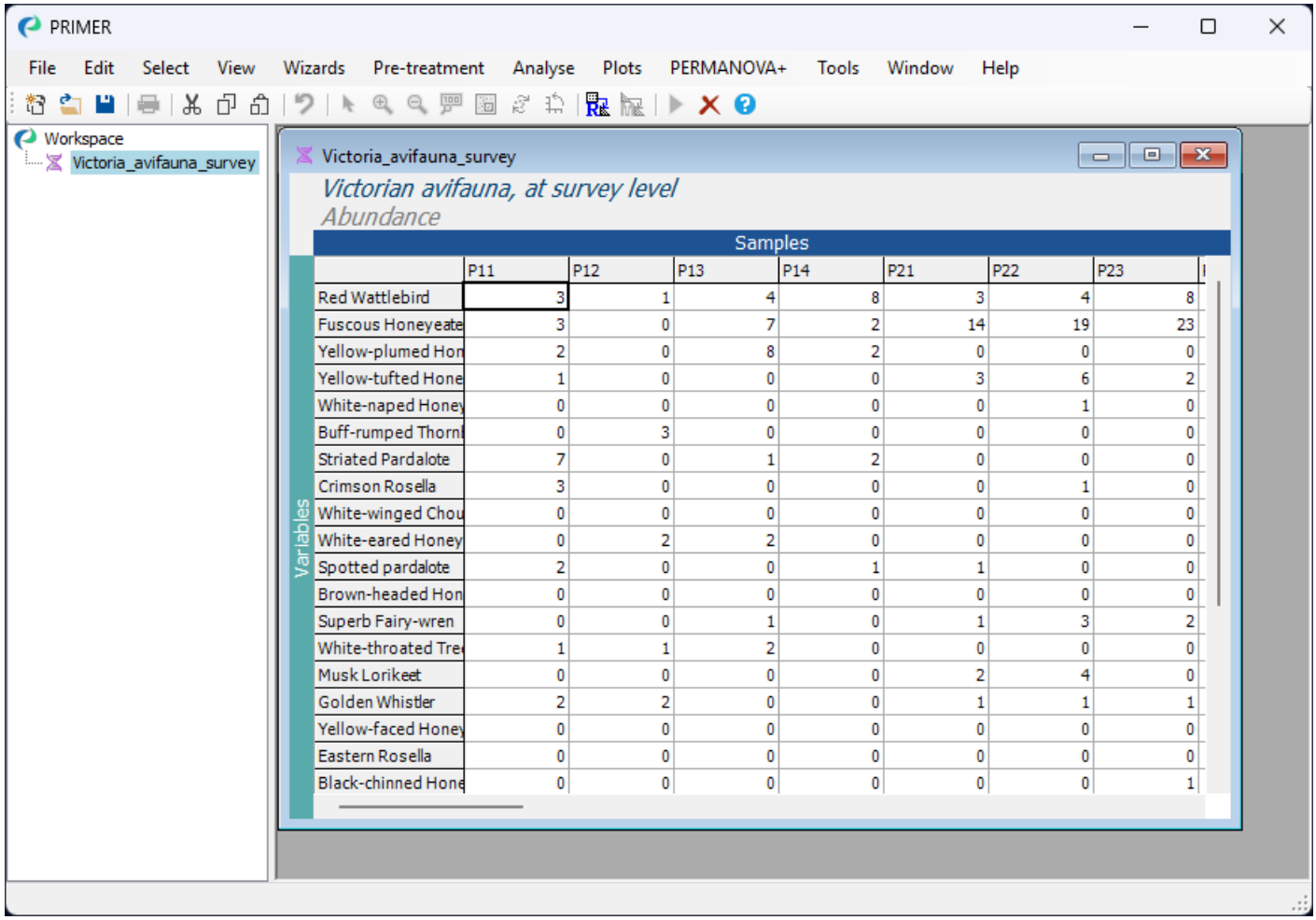


Fig. 8.5 Schematic diagram of the repeated-measures sampling design to study bird assemblages in response to flowering intensity by Mac Nally & Timewell (2005) .

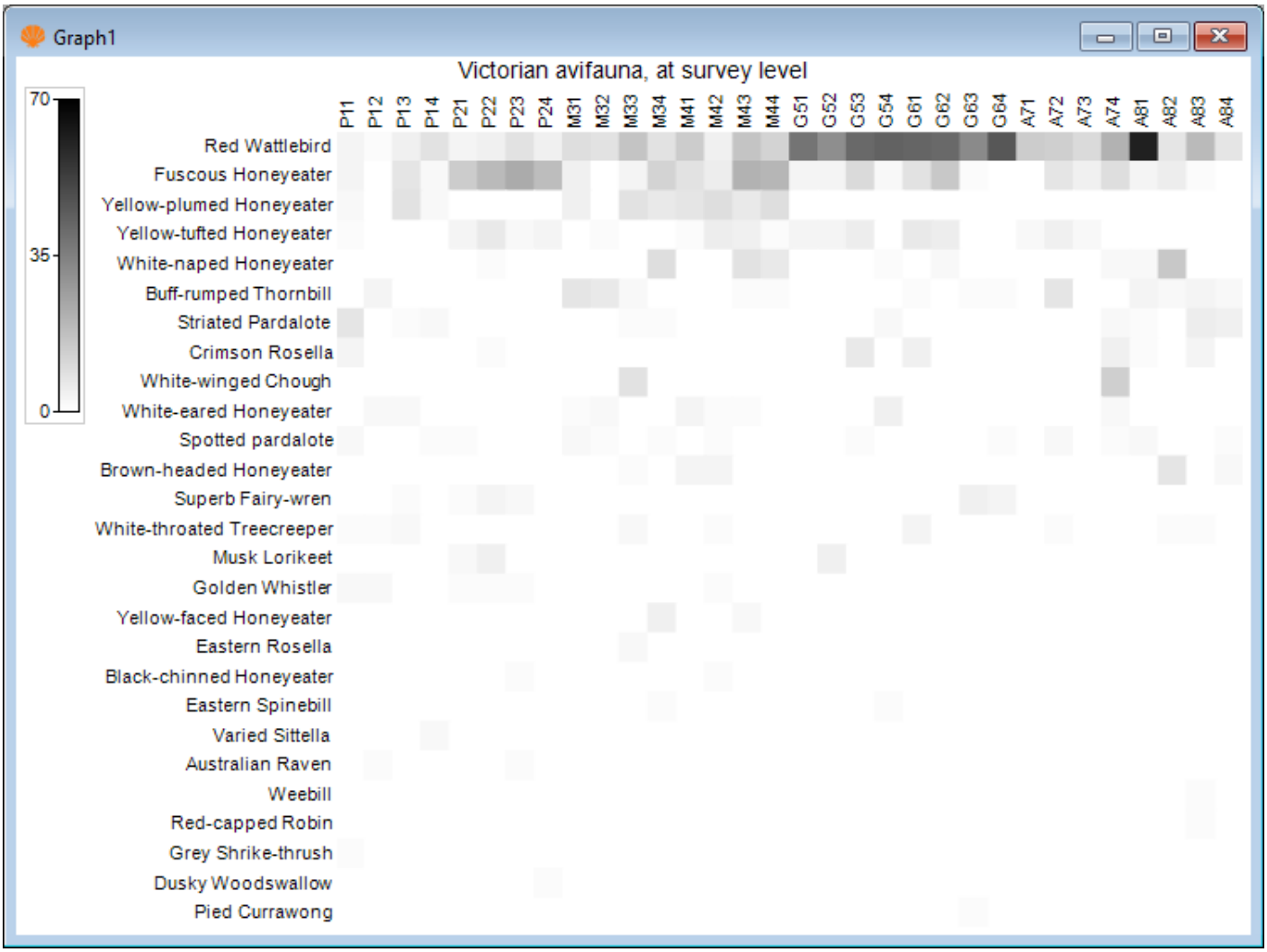
Input data and select a pre-treatment option

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

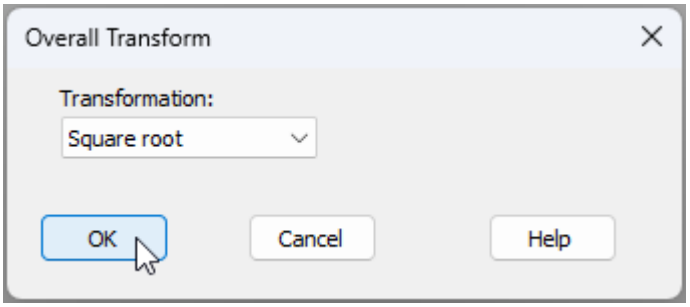


	P11	P12	P13	P14	P21	P22	P23
Red Wattlebird	3	1	4	8	3	4	8
Fuscous Honeyeate	3	0	7	2	14	19	23
Yellow-plumed Hon	2	0	8	2	0	0	0
Yellow-tufted Hone	1	0	0	0	3	6	2
White-naped Honey	0	0	0	0	0	1	0
Buff-rumped Thornb	0	3	0	0	0	0	0
Striated Pardalote	7	0	1	2	0	0	0
Crimson Rosella	3	0	0	0	0	1	0
White-winged Chou	0	0	0	0	0	0	0
White-eared Honey	0	2	2	0	0	0	0
Spotted pardalote	2	0	0	1	1	0	0
Brown-headed Hon	0	0	0	0	0	0	0
Superb Fairy-wren	0	0	1	0	1	3	2
White-throated Tre	1	1	2	0	0	0	0
Musk Lorikeet	0	0	0	0	2	4	0
Golden Whistler	2	2	0	0	1	1	1
Yellow-faced Honey	0	0	0	0	0	0	0
Eastern Rosella	0	0	0	0	0	0	0
Black-chinned Hone	0	0	0	0	0	0	1

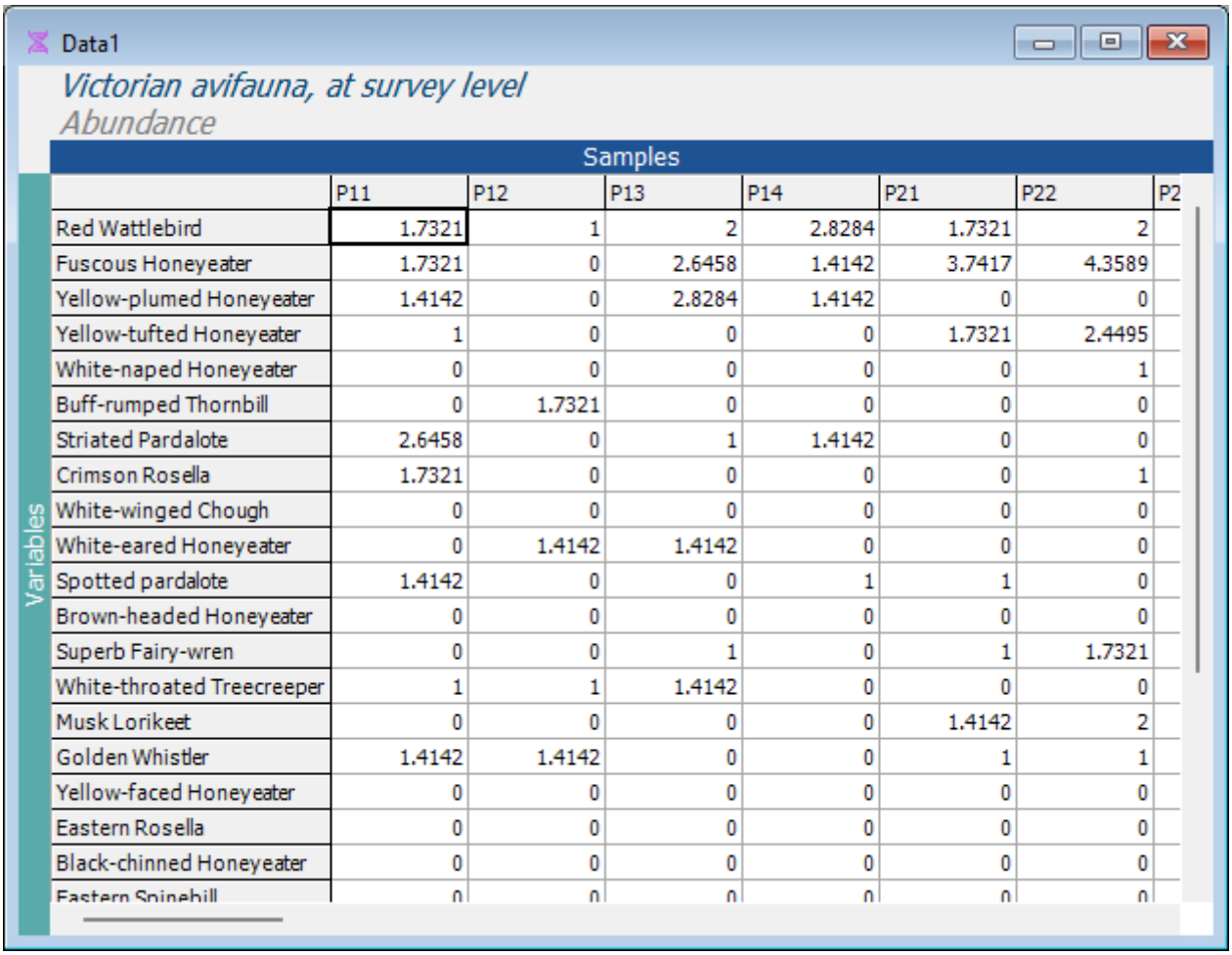
2. From the 'Victoria_avifauna_survey' data sheet in the Explorer tree inside PRIMER, click **Plots** > **Shade Plot**. It is evident from the resulting shade plot ('Graph1) that there is a lot of 'white space', and the range of abundance values is from 0 to 70. Thus a mild (e.g., square-root) transformation would be a sensible choice here as a pre-treatment option, to balance the relative importance of abundant *vs* rare taxa in the calculation of the resemblance measure.



3. From the 'Victoria_avifauna_survey' data sheet, click **Pre-treatment** > **Transform(overall)...**, and in the 'Overall Transform' dialog window, choose 'Transformation: Square root from the drop-down menu, then click '**OK**'.



The square-root transformed data are now provided in a data sheet called 'Data1':



Data1

Victorian avifauna, at survey level

Abundance

Samples

Variables

	P11	P12	P13	P14	P21	P22	P2
Red Wattlebird	1.7321	1	2	2.8284	1.7321	2	
Fuscous Honeyeater	1.7321	0	2.6458	1.4142	3.7417	4.3589	
Yellow-plumed Honeyeater	1.4142	0	2.8284	1.4142	0	0	
Yellow-tufted Honeyeater	1	0	0	0	1.7321	2.4495	
White-naped Honeyeater	0	0	0	0	0	1	
Buff-rumped Thornbill	0	1.7321	0	0	0	0	
Striated Pardalote	2.6458	0	1	1.4142	0	0	
Crimson Rosella	1.7321	0	0	0	0	1	
White-winged Chough	0	0	0	0	0	0	
White-eared Honeyeater	0	1.4142	1.4142	0	0	0	
Spotted pardalote	1.4142	0	0	1	1	0	
Brown-headed Honeyeater	0	0	0	0	0	0	
Superb Fairy-wren	0	0	1	0	1	1.7321	
White-throated Treecreeper	1	1	1.4142	0	0	0	
Musk Lorikeet	0	0	0	0	1.4142	2	
Golden Whistler	1.4142	1.4142	0	0	1	1	
Yellow-faced Honeyeater	0	0	0	0	0	0	
Eastern Rosella	0	0	0	0	0	0	
Black-chinned Honeyeater	0	0	0	0	0	0	
Eastern Spinebill	0	0	0	0	0	0	

4. From the square-root transformed data ('Data1'), click **Plots** > **Shade Plot**, and you will see in the resulting shade plot graphic (called 'Graph2' in the Explorer tree) a more even distribution of abundance information across all bird taxa (less white space), with transformed abundance values now ranging from 0-8.

Calculate dissimilarities and visualise patterns

Now we are ready to calculate dissimilarities (or similarities) based on the transformed data and use this to visualise patterns of relationships among the sampling units, based on the fauna they contain, using ordination methods.

5. From the square-root transformed data ('Data1'), click **Analyse** > **Resemblance...** and in the 'Resemblance' dialog window, choose (Measure: $\bullet$Bray-Curtis similarity) and (Analyse between: $\bullet$Samples), then click **OK**. This will yield a resemblance matrix item in the Explorer tree, called 'Resem1'.

Resem1

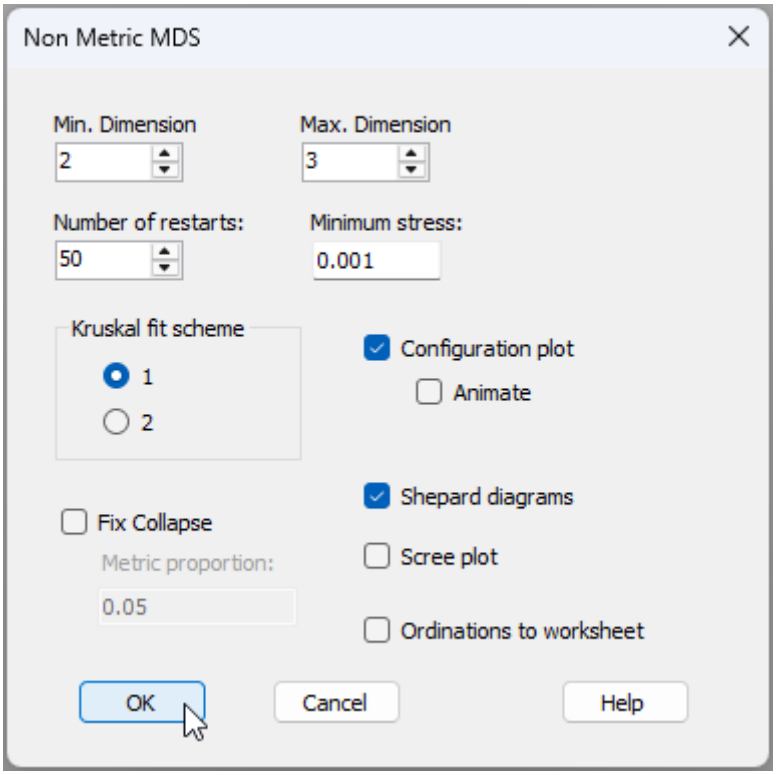
Victorian avifauna, at survey level

Similarity (0 to 100)

Samples

Samples	P11	P12	P13	P14	P21	P22	P23	P24	M31
P11									
P12	30.154								
P13	50.23	34.377							
P14	56.774	11.733	53.502						
P21	48.412	20.855	44.96	39.291					
P22	42.215	17.315	40.554	27.286	78.202				
P23	38.295	28.554	43.841	36.992	70.897	70.274			
P24	37.109	12.095	43.669	36.99	69.977	65.059	68.281		
M31	46.363	38.042	57.465	61.794	39.967	28.985	37.851	38.032	
M32	31.393	51.023	32.524	42.123	36.747	24.76	34.576	33.962	70.1
M33	42.049	27.106	59.963	49.098	23.685	22.501	29.343	28.054	54.8
M34	43.224	8.2296	55.743	58.388	43.693	40.047	41.68	42.495	54.
M41	40.872	22.732	67.012	48.839	43.957	39.893	49.069	51.455	62.
M42	56.751	33.663	63.595	45.376	55.349	47.077	51.592	47.406	56.6
M43	34.217	22.359	51.273	39.338	46.647	53.763	53.932	56.459	57.4
M44	38.584	17.432	53.988	45.49	47.943	54.057	57.566	59.462	58.2
G51	36.127	11.635	34.034	44.393	48.908	43.419	51.771	58.744	43.6

6. From the Bray-Curtis resemblance matrix ('Resem1'), click **Analyse** > **MDS** > **Non-metric MDS (nMDS)...**. Leaving all of the defaults in the 'Non Metric MDS' dialog window, click '**OK**'.



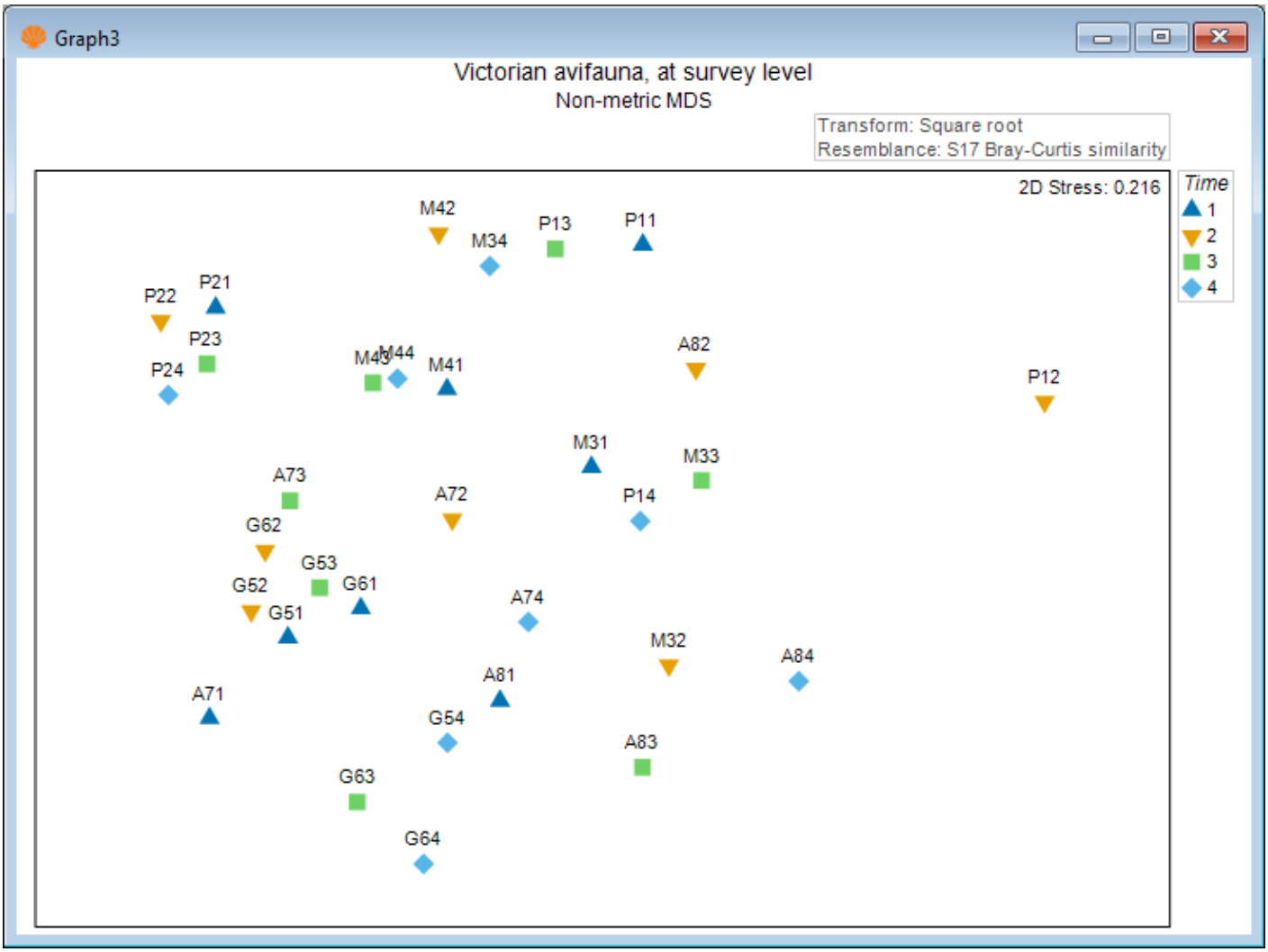
You will see a new item called 'MultiPlot1' in the Explorer tree, which contains 4 different graphics. Click on the '+' symbol next to the word 'MultiPlot1' in the Explorer tree, like so:

and this will "unfurl" (in the Explorer tree) to show all of the graphics in the MultiPlot:

- 'Graph3' (the 2D MDS plot),
- 'Graph4' (the Shepard diagram for the 2D MDS plot),
- 'Graph5' (the 3D MDS plot), and
- 'Graph6' (the Shepard diagram for the 3D MDS plot).

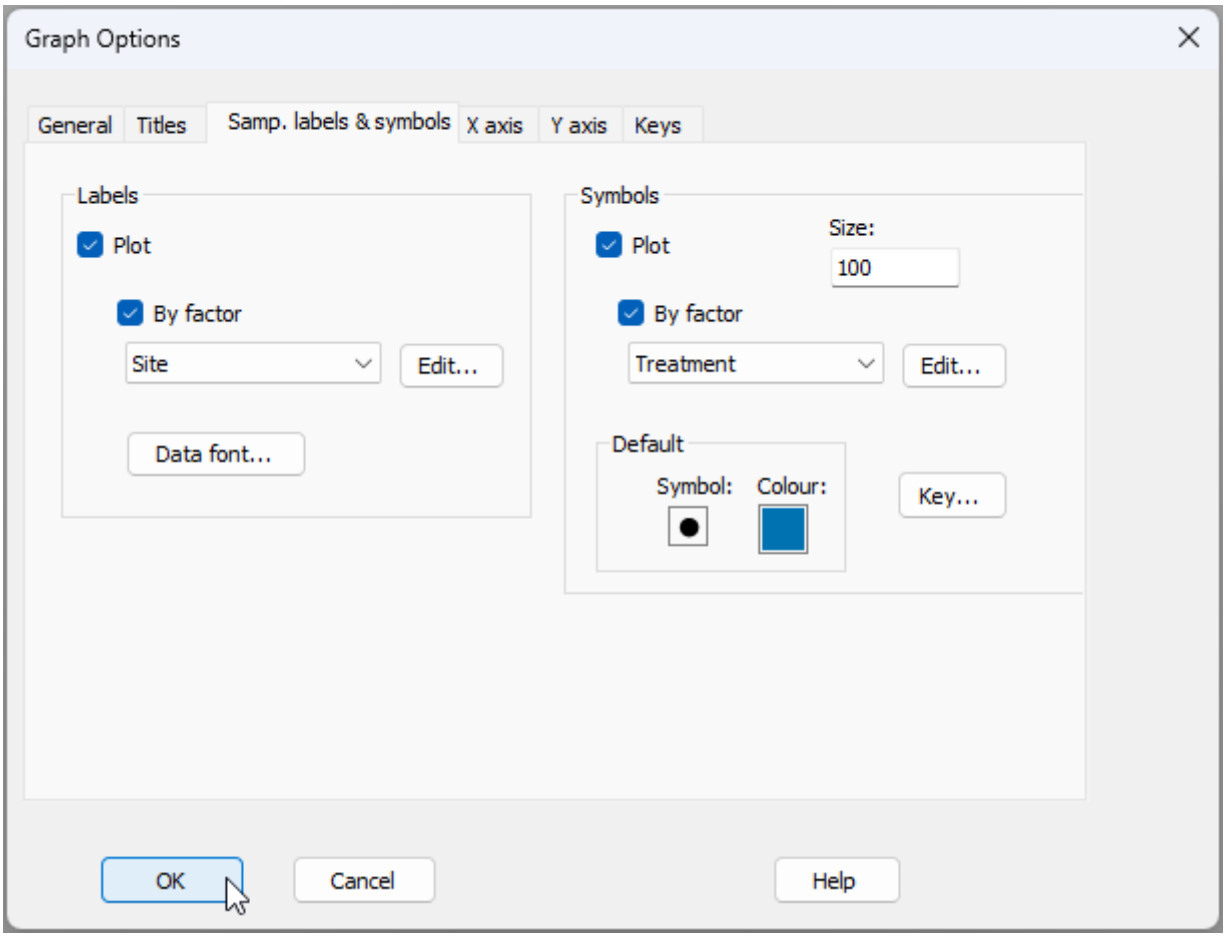
You can also simply click any of the individual graphics within the 'MultiPlot1' graphic itself to see it on its own.

7. Click on 'Graph3' to see the 2D MDS plot. The default plot in the output here looks a bit messy at first:

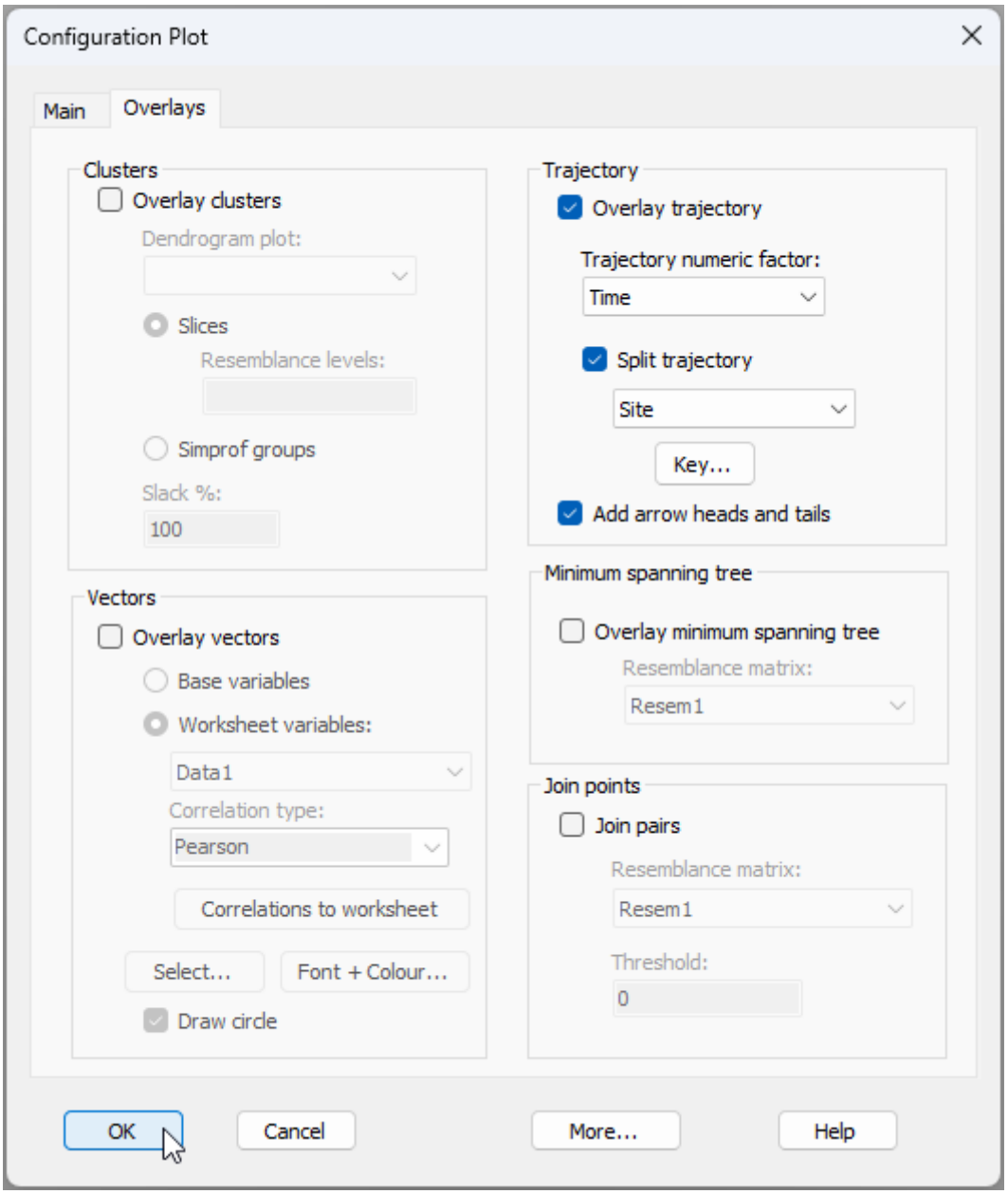


We will do two things (visually) to this graphic to make it easier to see the potential effects of important factors in this study.

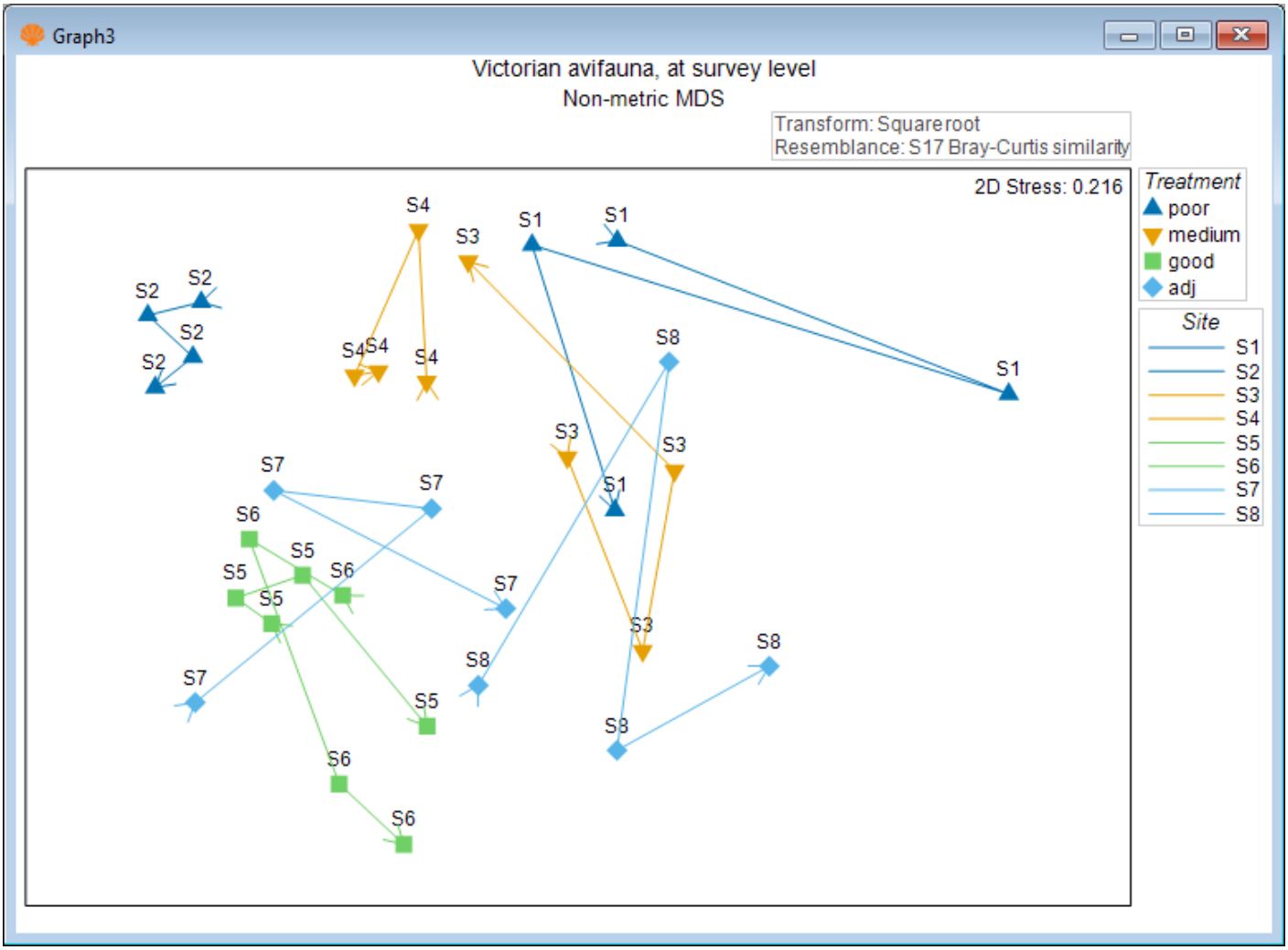
First, we will *change the labels and symbols* so that they correspond to spatial factors of interest. Click **Graph** > **Sample Labels & Symbols...** and change the **Labels** to reflect the factor of 'Site', and change the **Symbols** to reflect the factor of **Treatment**, as shown below, then click **OK**.



Second, we will *superimpose a trajectory to connect observations from the same site through time*. Click **Graph** > **Special**, then click on the 'Overlays' tab in the 'Configuration Plot' dialog, and under the word 'Trajectory', tick the box to $\checkmark$Overlay trajectory > Trajectory numeric factor: Time and $\checkmark$Split trajectory by Site, as shown below. Note that you can (optionally) click on the 'Key' button (Key...) here to make the colours of individual Site trajectory lines match their Treatment colour, then click **OK**.



The resulting 2D nMDS graphic will look like this:



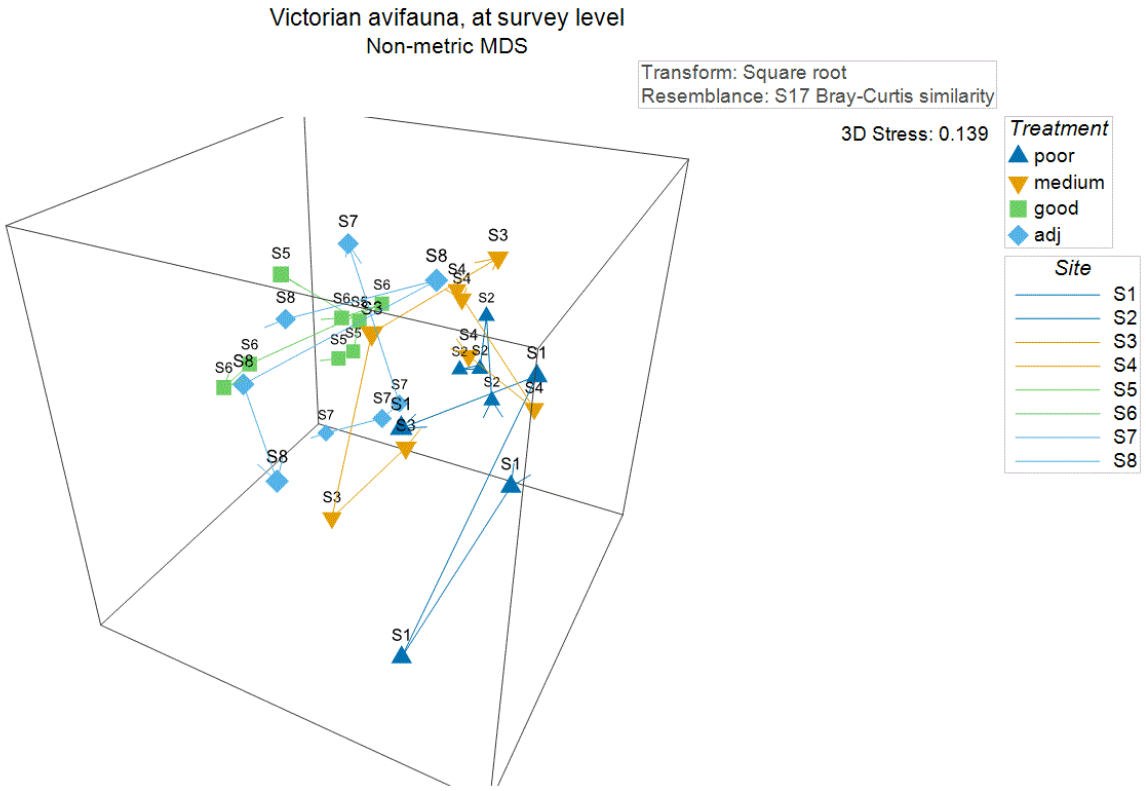
We can see here that there is a general pattern of gradual change in bird assemblages as you go from the 'good' (high flowering intensity) sites (in green) through the 'adjacent' (light blue) and 'medium' (orange) sites, towards the 'poor' sites (in dark blue); i.e., a sort of gradient from the lower left of the nMDS plot to the upper right. However, there is also quite a substantial difference in the bird assemblages seen at the two 'poor' sites (S1 *vs* S2), with S2 not really sitting in the appropriate place along the observed gradient. The variation through time (the overall length of each trajectory) also seems to differ somewhat across the sites.

An important thing to note here also is the relatively high stress of 0.216. This suggests that we should not read too much into any of the patterns we see on this 2D plot. Indeed, as the stress is higher than the usual 'rule-of-thumb' cut-off for interpretability of nMDS plots generally (stress = 0.20), we should really consider looking at the 3-dimensional nMDS ordination here instead.

8. To look at the 3D MDS plot, click on 'Graph5' and start by doing the same two operations we did on the 2D plot, i.e., change the labels and symbols to correspond to 'Site' and ' Treatment', respectively, then superimpose trajectories through time for each site (and adjust the Key for the 'Site' factor so that the colour of the line for each site corresponds to the colour of the treatment to which it belongs).

It is difficult to get a real sense of the patterns being shown in a 3D plot unless you see it in motion. From 'Graph5', click **Graph** > **Spin**, and you will see a series of buttons at the top of the graphic that look like this:

Hit the 'play' button ( ) and the graphic will start to spin. The blue slider permits you to increase or decrease the speed of the spin and you can change the angle of your view of the 3D image by clicking and dragging your mouse on it. You can also hit the red dot to record the spinning action, as shown in the animated $*$.gif image below (click on the image below to enlarge your view and see this as it would appear within PRIMER).

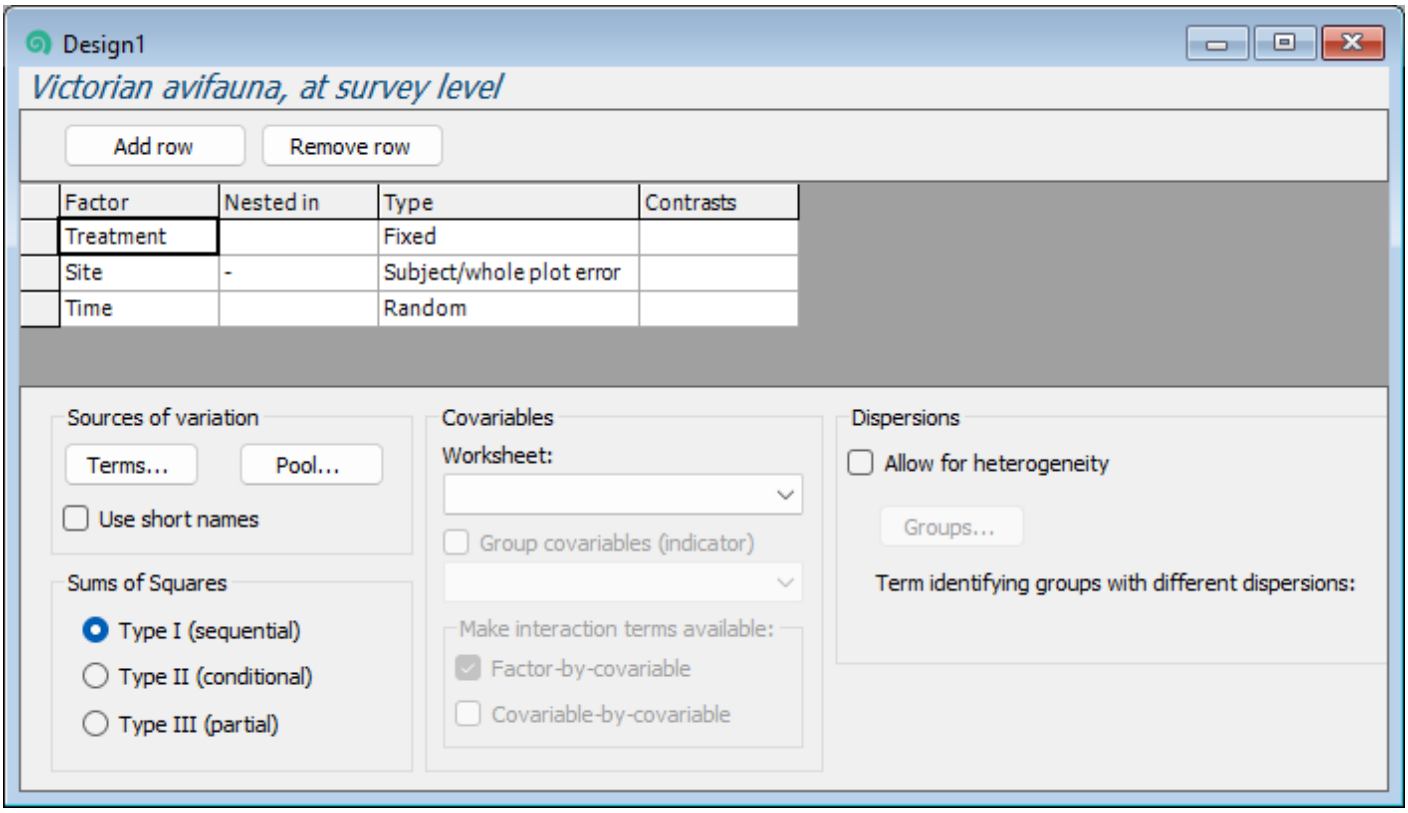


The patterns we see here are similar to what we saw in the 2D plot. There is a suggestion of a gradual shift in the bird assemblages with changes in flowering intensity, but there are also apparently substantial differences between sites of a similar flowering intensity, particularly between the two 'poor' sites.

Run the PERMANOVA

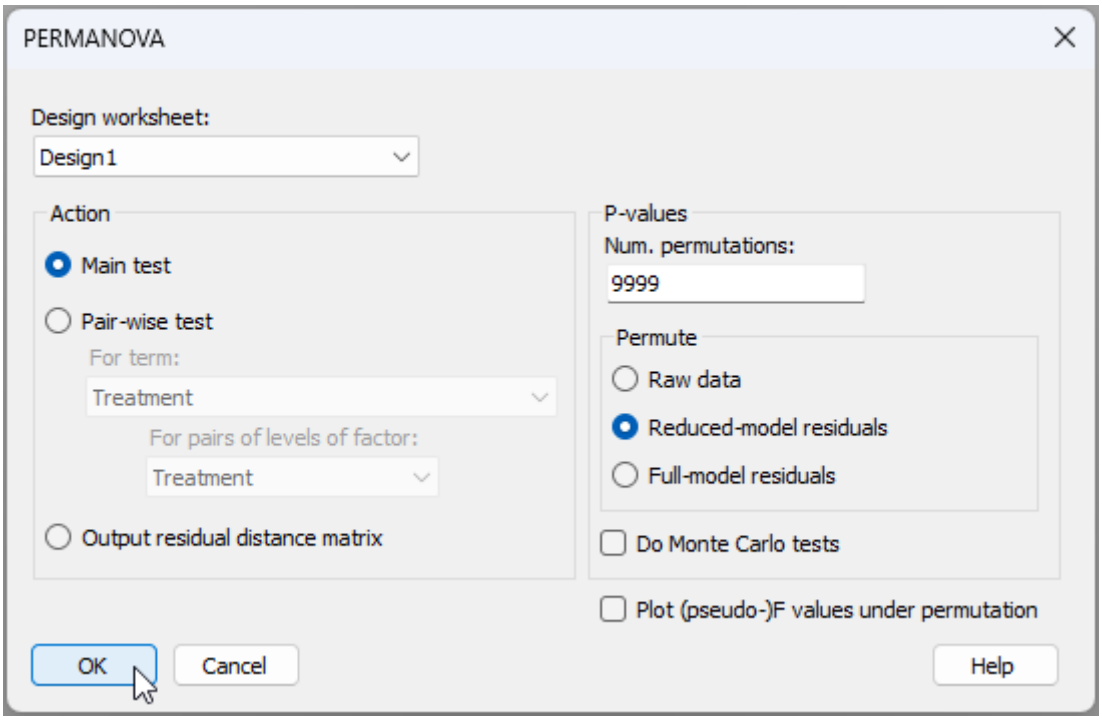
We need to create a design file, then run the PERMANOVA model on these data to formally test hypotheses regarding the potential effects of these factors.

10. From the Bray-Curtis resemblance matrix ('Resem1'), click **PERMANOVA+** > **Create PERMANOVA Design...**. Click the 'Add row' button (Add row) until there are (in this case) three rows (one for each factor): 'Treatment', 'Site' and then 'Time'. Next, we shall nominate 'Treatment' as a 'Fixed' factor, and 'Time' as a 'Random' factor. Note that individual *sites* are the specific items that are *repeatedly sampled*. Therefore, we need to specify 'Site' as a factor of type 'Subject/whole plot error'. The resulting Design file (called 'Design1') will look like this:



Now we are ready to run the analysis.

11. From the Bray-Curtis resemblance matrix ('Resem1'), click **PERMANOVA+** > **PERMANOVA**, take all of the defaults in the dialog and click '**OK**'.



The resulting PERMANOVA output file ('PERMANOVA1') will look like this:

PERMANOVA1

PERMANOVA
Permutational MANOVA

Resemblance worksheet
Name: Resem1
Data type: Similarity
Selection: All
Transform: Square root
Resemblance: S17 Bray-Curtis similarity

PERMANOVA design worksheet
Name: Design1
Title: Victorian avifauna, at survey level

Sums of squares type: Type I (sequential)
Permutation method: Permutation of residuals under a reduced model
Number of permutations: 9999

Factors

Name	Type	Levels
Treatment	Fixed	4
Site	Subject/whole plot error	8
Time	Random	4

Excluded terms
No terms excluded

Inestimable terms
No inestimable terms.

PERMANOVA table of results

Source	df	SS	MS	Pseudo-F	P(perm)	Unique perms
Treatment	3	14959	4986.4	2.0206	0.068	105
Site	4	9871	2467.8	2.3715	0.0071	9931
Time	3	4512.1	1504	1.4453	0.1417	9919
TreatmentxTime	9	7793.9	865.99	0.83221	0.7386	9886
Res	12	12487	1040.6			
Total	31	49623				

There is significant variability from site to site *within* each of the different flowering intensity treatments ($F_{(4,12)}$ = 2.37, P = 0.0071). Over and above this site-to-site variation, however, there was only weak evidence of any treatment effects ($F_{(3,4)}$ = 2.02, P = 0.068).

These results make perfect sense by reference to the patterns seen in the nMDS plot(s), where we could discern a pattern of change in bird community structure with differences in flowering intensity, but also saw substantial site-to-site variation.

We might consider analysing the data again but without any pre-treatment transformation, in order to emphasise changes in the more abundant birds rather more (*try it!*). Another idea would be to analyse some other aspect or variable associated with these bird assemblages in response to the study design, such as species richness, or the univariate abundances of one or more dominant species (such as the *Red wattlebird*).

Ultimately, we might consider sampling a larger number of sites to increase the power of the test for Treatment effects (4 degrees of freedom in our denominator was not really very many).

Choosing sites to be as similar as possible to one another in other respects (environmentally) in order to reduce site-to-site variation would also be wise.