Change to active sheet for BEST

In what is one of the very few examples of ‘moving the furniture around’ between PRIMER 7 and earlier versions, the active window for a run of **Analyse>BEST** is no longer the data matrix of (usually abiotic) variables, from which selections are made to best match to a fixed resemblance matrix (usually of assemblage pattern), but the converse, i.e. **Analyse>BEST** is run from an active sheet which is the fixed resemblance matrix, and a secondary matrix supplied which is the (abiotic) data matrix from which variables are selected. There are several good reasons for this switch, the most compelling of which is that it allows much greater consistency in the way PRIMER decides which samples to use in an analysis when there is only a partial match of sample labels between the two data sets. The consistent rule now, throughout PRIMER (and PERMANOVA+), is that the active matrix (in its currently selected form, if any selections are in place) determines that sample set. Any secondary matrix supplied to the routine (here, the abiotic variables) are treated as a ‘look-up’ table from which the required set of samples is extracted. Thus, the environmental matrix can (and sometimes will) cover a much wider range of sites than are utilised in the current community samples – they might for example be interpolations from some physico-chemical or remote-sensing model for the whole region. What is required is that BEST can find all (biotic) resemblance sample labels in the (abiotic) data matrix, otherwise an error is returned – with the usual relaxation of strict label matching that if the two matrices have exactly the same number of samples, i.e. BEST will ask the user if it should should proceed on the assumption that the samples are in the same order. [Other benefits from switching the active matrix for BEST include consistency with the DISTLM routine in PERMANOVA+, which is the semi-parametric equivalent to the non-parametric BEST program, and a close multivariate analogue of (univariate) multiple linear regression. In standard statistical thinking, the *response* here is the community sample, thought of as subject to sampling/ spatio-temporal variability and that is regressed on the observed values of the *explanatory* variables (the latter considered fixed, under a conditionality argument). PERMANOVA+ routines thus start from the response, as given by the community resemblances – which are always the active sheet – and explanatory variables (in DISTLM), covariates (in PERMANOVA) etc are always secondary.]

Revision #1
Created 3 October 2024 01:14:53 by Arden
Updated 3 October 2024 01:17:54 by Arden