Limiting the number of combinations

The final area of the main BEST dialog, headed Permutations, which carries out the *global BEST test* for statistical significance of the best matching combination of variables, is deferred until later in this section. The Next > button, under (Method•BIOENV), gives a dialog with a single entry, a choice of (Max num of trial variables/groups:), for which 5 is the default. This limits the search to $\le$5 (abiotic) variables at a time, and this maximum number should be increased, where feasible. A default of the total number of input variables is not used because the number of combinations of these in an exhaustive search could be very large: for *p* variables there are c = 2^p – 1 combinations, and a practically realistic limit therefore has to be about *p* = 17 (giving c $\approx$ 100,000). The context for Section 13 is the matching of subsets of environmental variables to assemblage patterns. Quite often, the number of abiotic variables is then <17 or, if not, the number should probably be pruned before running BEST – so only a full search (BIOENV) will be illustrated now. BVSTEP could be run in much the same way on a larger set, but the reason this is likely to prove unattractive is that, with so many abiotic variables, it is inevitable that they will be strongly inter-correlated. There are then a plethora of equally good solutions and a rather unfocussed interpretation. Deletion of all but one of a highly mutually-correlated set of variables and/or prior reduction to one representative of each different type of environmental variable, may be desirable, just as in multiple linear regression (see the discussion in Chapter 11 of CiMC). In some of the other applications – e.g. when the data matrix is of species variables and *a priori* selection defeats the point of the analysis – the stepwise form (Method•BVSTEP) will be essential, and such an example is seen in Section 14.

Revision #3
Created 3 October 2024 22:30:51 by Arden
Updated 12 February 2025 01:31:18 by Abby Miller