Non-metric, non-linear, non-additive

The **Analyse>Cluster>LINKTREE** routine has a number of features that are designed to mesh to the PRIMER approach. Firstly, as seen for unconstrained UNCTREE clustering (Section 6), each successive split of the biotic samples into two groups (of potentially unequal size) maximises the ANOSIM R statistic (Chapter 6, CiMC). An ANOSIM test is not carried out, of course (that would be totally invalid since the same data would be used to define the groups as to test them!) but R has a general role as a non-parametric measure of multivariate difference between groups (in high-d), rather than just as a test statistic. Unlike the much more computationally intensive UNCTREE, not all possible binary divisions are permitted (there are $\sim$2^{16} possibilities for just the initial split of the 17 lagoon sites, which is why UNCTREE needs an iterative search algorithm!). In fact LINKTREE can simply examine all splits that correspond to a threshold condition on an abiotic variable (so for 3 variables there are at most 3$\times$16=48 ways to divide 17 samples into two groups). Secondly, the procedure is truly *non-metric*, not just on the community resemblance matrix but also on the abiotic variables. A (monotonic) transform of the environmental variables can make no difference to the outcome of LINKTREE, since all that is being used is how a criterion like In-N<965 or >1380 splits up the samples (again there are no samples with In-N between 965 and 1380). That division is unchanged under transformation, just becoming log(In-N)<log(965) or >log(1390) for example. Thirdly, and more subtly, the way the different abiotic variables are combined in the partitioning of the biotic samples is clearly *non-linear* but is also *non-additive*. In contrast, BEST is non-metric and can certainly accommodate non-linear responses of the assemblages to driving environmental variables, but does make an implicit assumption that their effects are additive. For example, if high PO$_4$ were to be an important variable in separating the diatom communities but only in low salinity environments (with equally large variation in PO$_4$ having no effect on the biota in high salinities), then this would clearly degrade the BEST match (ρ). Such interactions are one explanation for the failure to get a good match, along with several others: high sampling ‘noise’, failure to measure the important abiotic variables, communities structured by competition not external driving variables, etc. However, LINKTREE attempts only local explanations – rather than holistic ones in the way BEST does – and is clearly capable of showing, for example, that PO$_4$ is important for structuring low-salinity groups but not high-salinity ones (with similar PO$_4$ ranges). A big disadvantage with the local (piecemeal) explanations offered by LINKTREE is that many abiotic inequalities will explain the same assemblage divisions, unless the environmental variable set is initially drastically pruned. An advantage is that it is geared towards prediction, and not just interpretation.

Revision #5
Created 7 October 2024 21:49:42 by Arden
Updated 12 February 2025 03:14:15 by Abby Miller