Environment-type data

PRIMER uses the term *environmental variables* as a shorthand for a wide variety of data types (including biological data!), extending well beyond the archetypal case of physical or chemical measurements made on the environment surrounding an assemblage sample. Environment-type variables can also include matrices of biomarker responses (biochemical, sub-cellular or whole body health indicators from individual organisms, Section 4), morphometric measurements on individuals (perhaps with the aim of separating putative species), PSA data (size-class spectra for soil/sediment/water particulates, Section 4), organism body-size distributions, etc. The unifying factor for these disparate examples is that: a) they all give rise to multivariate arrays of variables by samples which can be analysed by the methods in PRIMER; b) the criteria which lead to use of community-type similarity measures such as Bray-Curtis are not appropriate (e.g. always positive entries, with many zeros and zero playing a special role – joint absences carrying no information, samples with no species in common having zero similarity – and always a common measurement scale across variables, of abundance, biomass, % cover etc.). Instead, resemblance between samples of environment-type variables is better described by standard distance measures such as Euclidean distance (Section 5), where zero plays no special role (e.g. zero temperature, but on what scale?), where negative values can occur (indeed will occur if normalising different scales to common units, Section 4), and where positive similarity is always inferred if two samples have the same value of a variable, even a zero value (e.g. neither sample has a detectable PCB or Hg level, neither sample has particles > size *x*, etc.). The key message here is that whole assemblage data is different, and requires the specialised methods that are at the core of PRIMER (biological similarity coefficients, non-metric MDS plots, non-parametric ANOSIM tests etc.), environmental-type data is more standard and is often (after individual transformations and normalisation) best treated by the more classic approaches of Euclidean distances and Principal Components (PCA) ordination. The derivation and purpose of PCA is covered in detail in Chapter 4 of CiMC.

Revision #5
Created 26 September 2024 22:54:12 by Arden
Updated 11 February 2025 23:06:14 by Abby Miller