11.7 Concluding remarks

For this chapter as a whole, two final points need to be made. The topic of experimental and field survey design for ecologists is a large one, addressed to some extent in the accompanying PERMANOVA+ manual (Anderson, Gorley & Clarke (2008))[¶], but this is a problematic area for all multivariate techniques because of the difficulty of specifying an explicit alternative hypothesis to the null hypothesis of, for example, no link of an assemblage to abiotic variables. A specified alternative is required to define *power* of statistical procedures but there are a myriad of ways in which individual species can react, even to a single environmental variable (some increase along an abiotic gradient, some decrease, some increase then decrease, others change little etc), *any* combination of which, for *each* of the variables, will be inferred as a biotic-abiotic link. Formal power calculations, analogous to those for simple univariate regression (e.g. Bayne, Clarke & Moore (1981)), are a non-starter, and simulation from observed alternatives to the null conditions are the only possible approach (see, for example, Somerfield, Clarke & Olsgard (2002)). However, in the context of linking biotic and abiotic patterns, it is intuitively clear that this has the greatest prospect of success if there are a moderately large number of sample conditions, and the closest possible matching of environmental with biological data. In the case of a number of replicates from each of a number of sites, this could imply that the biotic replicates would each have a closely-matched environmental replicate. Without matching of biotic and abiotic samples none of the methods of this chapter could be used, so data from the two sources will always need averaging up to the lowest common denominator, giving a one-to-one match of ‘response’ and ‘explanatory’ samples.

Another lesson of the Fal estuary nematode study and the Garroch Head example of Fig 11.9 is the difficulty of drawing conclusions about *causal* variables from *any* observational study. In the Garroch Head case, four of the abiotic variables were so highly correlated with each other that it was desirable to omit all but one of them from the computations. There may sometimes be good external reasons for retaining a particular member of the set but, in general, one of them is chosen arbitrarily as a *proxy* for the rest (e.g. in the Garroch Head data, %C was a proxy for the highly inter-correlated set %C, Cu, Zn, Pb). If that variable does appear to be linked to the biotic pattern then any member of the subset could be implicated, of course. More importantly, there cannot be a definitive *causal* implication here, since each retained variable is also a proxy for any potentially causal variable which correlates highly with it, but remains *unmeasured*. Clearly, in an environmental impact study, a design in which the main pollution gradient (e.g. chemical) is highly correlated with variations in some natural environmental measures (e.g. salinity, sediment structure), cannot be very informative, whether the latter variables are measured or not. A desirable strategy, particularly for the non-parametric multivariate analyses considered here, is to limit the influence of important natural variables by attempting to select sites which have the same environmental conditions but a range of contaminant impacts (including control sites[†] of course). Even then[§], in a purely observational study one can never entirely escape the stricture that any apparent change in community, with changing pollution impact, could be the result of an unmeasured and unconsidered natural variable with which the contaminant levels happen to

correlate. Such issues of causality motivate the following chapter on experimental approaches.

¶ *Green (1979) also provides some useful guidelines, mainly on field observational studies, and Underwood (1997) concentrates on design of field manipulative experiments; both books are largely concerned with univariate data but many of the core issues are common to all analyses.*

† *Note the plurality; Underwood (1992) argues persuasively that impact is best established against a baseline of site-to-site variability in control conditions.*

§ *And in spite of impressive modern work on causal models that bring a much-needed sense of discipline to the selection of abiotic variables and prior modelling of causal links among variables and responses, see Paul & Anderson (2013) .*

Revision #8
Created 6 March 2022 11:21:16 by Arden
Updated 6 November 2024 05:50:53 by Abby Miller