6.14 Three-way ANOSIM designs

Table 6.4 details all viable combinations of 3 factors, A, B, C, in crossed/nested form, ordered/unordered, and with/without replication at the lowest level. *Fully crossed designs* are denoted *A$\times$B$\times$C*, e.g. locations (A) each examined at the same set of times (B) and for the same set of depths (C) ¶.

***Table 6.4. 3-way ANOSIM** (global) test statistics, for crossed and nested designs, with unordered or ordered factors, and with or without replication at the lowest level of the design. Also given are the existence (or not) of pairwise tests, details of the test constructions (making reference to test numbers in Table 6.3) and examples of contexts in which they might be employed.*

No.	Type of design	Factors	Factor levels ordered?	Repli-cates?	Statistic used	Pairwise test?	Construction of statistic	Examples
3a	3-way crossed	AxBxC	A,B,C unordered	Yes	A,B,C: $\bar{R}$	Yes	As two-way crossed test 2a, combining pairs of factors in turn [1]	A: location, B: time, C: habitat
3b	3-way crossed	AxBxC	A,B,C unordered	No	A,B,C: ρ_{av}	No	As 2b, e.g. comparing resemblance matrices of A across BxC levels [1]	As above but no reps (or pooled)
3c	3-way crossed	AxBxC	A,B unordered C ordered	Yes/No	A,B: $\bar{R}$ C: $\bar{R}^{Oc}/\bar{R}^{Os}$	Yes/No	A,B: as test 3a/3b C: as test 2c/2d, collapsing A,B to single factor AxB [1]	A: location, B: time, C: depth range with/without reps in AxBxC cells
3d	3-way nested, C within B within A	C(B(A))	A,B,C unordered	Yes	A: R B,C: $\bar{R}$	A: Yes B,C: No	A,B: as test 2g of B nested in A, using levels of C as replicates [2] C: as test 2g for C nested in all the B levels (at all A levels)	A: region, B: location, C: site, with replicate samples at each site
3e	3-way nested, C within B within A	C(B(A))	A,B,C unordered	No	A: R, B: $\bar{R}$, C: -	A:Yes, B:No, C: -	A,B: exactly as for test 3d (except no averaging of C level reps needed) C: no basis for a test	A: region, B: location, C: site, with (e.g.) one pooled sample at each site
3f	3-way nested, C within B within A	C(B(A))	A,B unordered C ordered	Yes/No	A: R, B: $\bar{R}$ C: $\bar{R}^{Oc}/\bar{R}^{Os}$	A: Yes B,C: No	A,B: as test 2g of B nested in A, using C levels (/single C values) as reps [3] C: as tests 2k/2l but for 'B' read C and for 'A' read B(A), B levels in all A	A: location, B: shore, C: along shore transect, reps(or not) at transect pts.
3g	3-way nested, C within B within A	C(B(A))	A unordered B ordered, C either	Yes/No	A,C: above B: $\bar{R}^{Oc}$	A: Yes B,C: No	A,C: as the relevant tests in 3d-3f B: as 2k for B within A, using levels of C (/single C values) as reps [4]	A: sea region, B: transect of sites, C: random days at each site (rep trawls)
3h	3-way, C nested in AxB	C(AxB)	A,B,C ordered or unordered	Yes or No	Various	A,B: Yes C: No	A,B: as for tests 2a,c,e with C levels as reps (averaged where needed) [5] C: as for B in nested cases 2g,h,k,l, but for 'A' read all combinations AxB	A: location, B: season, C: different sites/days in each AxB (& rep cores)
3i	3-way, B crossed with C(A) (i.e. only C is nested in A)	BxC(A)	A,B,C unordered	Yes	A: $\bar{R}$ B: $\bar{R}$ C: $\bar{R}$	A: Yes B: Yes C: No	A: average C levels (on resemblances, note 2), test A as 2a, for AxB [6] B: standard 2-way crossed test 2a for B across all levels of C (over all A) C: standard 2-way nested test 2g for C within all combined levels AxB	A: location, B: time, C: same random sites in location returned to each time, with replicate samples at sites
3j	3-way, B crossed with C(A)	BxC(A)	A,B,C unordered	No	A: $\bar{R}$ B: $\bar{\rho}_{av}$ C: $\bar{\rho}_{av}$	A: Yes B: No C: No	A: as test 2a, for AxB, using C as reps but constrained perms (as note 6) B: ρ_{av} statistic (2b) for B patterns across C levels in each A, then meaned[7] C: converse ρ_{av} of C patterns matched across B levels, ρ_{av} meaned over A[8]	A: location, B: time, C: same random sites in location returned to each time for single sample (or pooled sample)
3k	3-way, B crossed with C(A)	BxC(A)	A unordered B unordered C ordered	Yes/No	A: $\bar{R}$ B: $\bar{R}$ /$\bar{\rho}_{av}$ C: $\bar{R}^{Oc}/\bar{R}^{Os}$	A: Yes B: Yes/No C: No	A: as test 3i/3j B: as test 3i/3j C: as tests 2k/2l, but for 'B' read C, for 'A' read all combinations of AxB	A: location, B: time, C: same (representative) transect of sites in location returned to each time
3l	3-way, B crossed with C(A)	BxC(A)	B ordered A,C ordered or unordered	Yes/No	B: $\bar{R}^{Oc}/\bar{R}^{Os}$ A,C: above or below	B: Yes/No A:Yes, C: No	B: as crossed tests 2c/2d for B, but for 'A' read all levels of C (over all A) A,C: as the relevant tests for A,C in 3i-3k,3m	A: location, B: yearly time trend, C: same random sites in each location visited each year
3m	3-way, B crossed with C(A)	BxC(A)	A ordered B,C ordered or unordered	Yes/No	A: $\bar{R}^{Oc}$ B,C: as above	A:Yes B: Yes/No C: No	A: as crossed test 2e for A across B, with C reps (if present) averaged as in note 2, then tested using block-constrained permutations, as note 6 B,C: as the relevant tests for B,C in 3i-3l	A: latitudinal region, B: yearly trend, C: same transect of sites in region each year (+reps) - A,B,C all ordered

[1] Test for A uses average of 1-way R (for A) across all levels of B and C in combination (BxC), then B v (AxC) and C v (AxB). Same idea for 3b (use test 2b), and if two of the factors are ordered still use 3a, b or c.
[2] Starts from ranked resemblances of C reps, then averaged and re-ranked (twice for A test). Or (eg if unsure of quality of C reps) test A & B by averaging C reps in data matrix and using 2g on A and B(A).
[3] C levels (averaged where needed, as in note 2) are assumed representative replicates of B(A) condition.
[4] If A ordered (whether B,C are or not), it changes nothing except the test of A, which is now as 2m.
[5] Similar comments as for note 2 apply, about whether it may be better sometimes to average replicates of C externally, on the data matrix, then calculate resemblances and submit to the 2-way crossed cases for AxB.
[6] But with one important difference: there is a new (for PRIMER) block-constrained permutation test here under the null, with values across B for each C level being permuted as a batch across C(A) and A levels. A common structure is A: locations, C: sites (nested in A), B: period, all sites visited in each period. Test for A uses sites as replicates but keeps the periods for each site together under permutation across locations.
[7] This is a new doubly-averaged statistic $\bar{\rho}_{av}$, matching patterns in B over the C levels for each A level (the usual ρ_{av}), then averaging ρ_{av} over A levels. Permutations are the usual random ordering of B for each C(A).
[8] E.g. ρ_{av} calculated of matching relationships among sites for different periods, separately for each location, then ρ_{av} averaged over locations. Standard permutation of sites within all levels of location x period.

With a fully symmetric design like this (cases 3a-c in Table 6.4), the idea is to test each factor in turn (A, say), by 'flattening/collapsing' the other two into a single factor (B$\times$C) whose levels are all the possible combinations of levels of B and C; the test for A from the relevant 2-way crossed design is then carried out. E.g. the global test for time effects (B removing A$\times$C) will only compare those different times at the same depth and location, and will then average those time-comparison statistics across all depth by location levels. Whichever of the definitions $\overline{R}$ / $\overline{R}^{Oc}$ / $\overline{R}^{Os}$ / ρ_{av} is used, the three global statistics (A removing B$\times$C, B removing A$\times$C, C removing A$\times$B) can be directly compared to gauge relative importance of A, B & C.

The *fully nested design C(B(A))*, e.g. area (C) nested in site (B), nested in location (A), cases 3d-g, can also be handled by repeated application of the 2-way case. This tests the lowest factor (C) inside the levels of the next highest (B), then averaging (in some form, see later) the replicate level, so that levels of C are now replicates for a test of B, then averaging the levels of C so that B levels are the replicates for a test of A.

Another straightforward possibility is *C(A$\times$B)*, 3h, in which C is nested in all combinations of A and B, e.g. multiple sites (C) are chosen from all combinations of location (A) and habitat type (B), in a case where all habitat types are found at each location, with replication (or not) at each site. The test for C uses the A$\times$B 'flattened' factor at the top level of a 2-way nested design, and tests for A and B are exactly as for the 2-way crossed design but, if replicates exist, averaging them (again, in some form) to utilise the levels of C as replicates for the crossed A and B tests.

The only other practical combination, and one which is quite frequently encountered, is *B$\times$C(A)*, 3i-m, in which only C is nested in A, and B is crossed with C, e.g. multiple sites (C) are identified at locations (A), and the *same sites* are returned to in each of a number of seasons (B), with (or without) genuine replicate day/area samples taken at each site in each season. Here there are one or two new issues of principle and these are illustrated in more detail later.

¶ *One of the commonest mistakes made by people new to ANOVA-type designs (whether in ANOSIM or PERMANOVA) is to assume here that depth is a nested factor in location, since the differing depth samples are all taken at the same location. But they are the same depths (or depth ranges) across locations, hence one can remove the location effect when studying depth and the depth effect when studying location, which is the whole point and power of a crossed design.*

Revision #9
Created 24 February 2022 11:29:04 by Arden
Updated 30 October 2024 00:50:09 by Abby Miller