Aquatic Ecology
Electronic Supplementary Materials

Response Variable selection in Principal Response Curves using permutation testing

Electronic Supplement 2: Data Generation

Nadia J. Vendrig · Lia Hemerik · Cajo J.F. ter Braak

Data Generation

The simulation experiment used in this paper is a simplified version of the Pyrifos experiment, explained in detail by van Wijngaarden et al (1996) and Van den Brink et al (1996). We simulated three treatment levels, each applied to the same number of ditches. Abundances were recorded on five occasions and only after the treatment was applied.

Sensible estimates for effect-size and covariance structure were based on results of PRC performed on a subset of the Pyrifos data set containing only the 21 RVs with less than 40% zeros. The structural part of the simulated data consisted of the sum of a_{tk} and the product of c_{dt} and b_k. As input a_{tk}, we used a random sample from the predicted values in the control group. As input c_{dt} we used the c_{dt} estimates of the first five sample times after treatment of the second and highest dose. As input b_k for the effect-RVs, we used -1, 1, 2, and 3 which reflected the range of estimated b_ks.

The random part of the data was estimated using a multivariate normal distribution which requires an input-covariance structure. To obtain Pyrifos-like covariance structure estimates, we first calculated residuals by subtracting fitted values from observed values. For *e.g.* species covariance structure, we estimated covariance between residuals obtained at the same time in the same ditch for the different species; and subtracted the mean covariance for that combination of time and ditch. Covariance structures for time and ditches were calculated in a similar fashion. As we assume ditches are independent, the variance-covariance matrix for ditches contained only variance estimated on the diagonal. Based on the observations, the variance-covariance matrix for time was assumed to have a first-order heterogeneous structure with ρ=0.4. We obtained only 21 estimates for species covariance-structures whilst we needed many more for the simulation study. To ensure that the variance-covariance matrix of the simulated data set, like that of the Pyrifos data set, contained clusters of species, we assigned a parent-species from the Pyrifos data set to

Nadia J. Vendrig
Biometris, Wageningen UR, P.O Box 16, 6700 AA Wageningen, The Netherlands
Tel.: +31-317-483574
E-mail: nadia.vendrig@wur.nl

Lia Hemerik
Biometris, Wageningen UR, P.O. Box 16, 6700 AA Wageningen, The Netherlands
Tel.: +31-317-482083
E-mail: lia.hemerik@wur.nl

Cajo J.F. Ter Braak
Biometris, Wageningen UR, P.O. Box 16, 6700 AA Wageningen, The Netherlands
Tel.: +31-317-480803
E-mail: cajo.terbraak@wur.nl

every species in the simulation data set. The correlation structure of the species is the same as that of the parent-species. For species with the same parent species, we set the correlation to 125% of the maximum correlation observed between species. In the covariance-matrix constructed from the correlation matrix, diagonals were increased by 40% to reduce probability of the matrix not being positive definite. The make.positive.definite function in the corpcor package (Schäfer et al, 2015) was used to ensure all matrices were positive definite. Variance-Covariance matrices for species and time were combined into one data set Variance-Covariance matrix and used to generate error estimates per ditch (mvrnorm function in the MASS package (Venables and Ripley, 2002)).

The range of the data simulated using the method described above was wider than that in the Pyrifos data set. Therefore, the estimates were multiplied by a factor to have the same maximum value as in the Pyrifos data set. Thereafter, the estimates were transformed back to abundance scale ($0.1 * (\exp(x))$) and used as expected values for a Poisson distribution.

References

Schäfer J, Opgen-Rhein R, Zuber V, Ahdesmäki M, Silva APD, Strimmer K (2015) corpcor: Efficient Estimation of Covariance and (Partial) Correlation

Van den Brink PJ, Van Wijngaarden RP, Lucassen WG, Brock TC, Leeuwangh P (1996) Effects of the insecticide Dursban® 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: II. Invertebrate community responses and recovery. Environ Toxicol Chem 15(7):1143–1153, DOI 10.1002/etc.5620150719

Venables WN, Ripley BD (2002) Modern Applied Statistics with S, 4th edn. Springer, New York

van Wijngaarden RPA, van den Brink PJ, Crum SJH, Brock TCM, Leeuwangh P, Oude Voshaar JH (1996) Effects of the insecticide dursban® 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: I. Comparison of short-term toxicity between the laboratory and the field. Environ Toxicol Chem 15(7):1133–1142, DOI 10.1002/etc.5620150718