The following supplement accompanies the article

Environment, not characteristics of individual algal rafts, affects composition of rafting invertebrate assemblages in Irish coastal waters

E. Clarkin[1], C. A. Maggs[1], A. L. Allcock[2], M. P. Johnson[2,*]

[1]Queen's University Belfast, School of Biological Sciences, Medical Biology Centre, 97 Lisburn Road, Belfast, BT9 7BL, UK
[2]Ryan Institute and School of Natural Sciences, National University of Ireland, Galway, University Road, Galway, Ireland

*Email: mark.johnson@nuigalway.ie

Marine Ecology Progress Series 470: 31–40 (2012)

Supplement. Multivariate regression trees for the assemblage structure of rafts

Multivariate regression trees (De'ath 2002) were used to attempt to link the structure in the invertebrate dissimilarities between rafts to environmental or raft-specific variables. This approach seeks to partition the variation to maximise the variance explained at each step. Models are evaluated on the cross-validated error: calculated by randomly splitting the data into two and fitting models to one half, while estimating the prediction error on the unused data. Cross-validated relative error close to zero indicates a good model. Values above 1 indicate poor models with little predictive power.

The appropriate size of tree can be extimated by resampling to find the mean cross validated error at each succesive partition of the data. Mulitvariate regression trees were fitted using mvpart in R (R Development Core Team 2008).

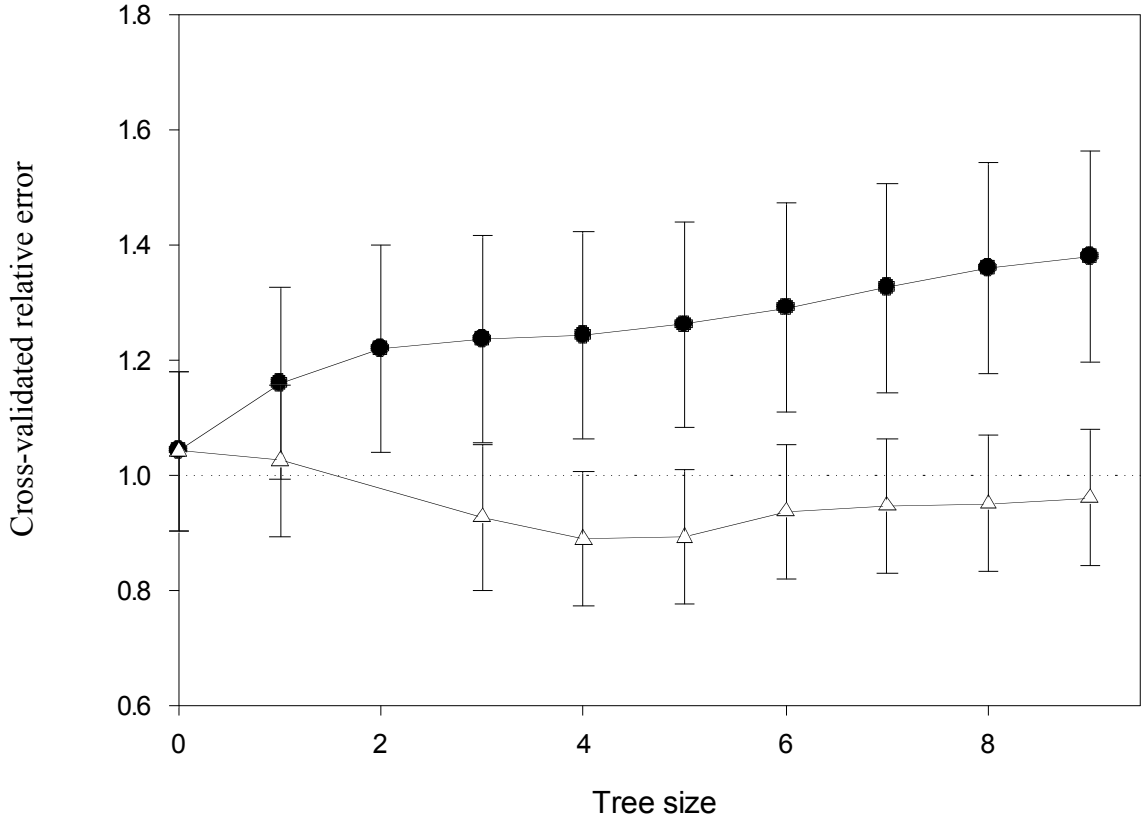

Fig. S1. Cross-validated error for regression trees of different complexity. Solid circles are models using algal biomass by species. Open triangles are models using the same location-related variables as Fig. 3b in the main article: sea temperature (Temp), salinity, chlorophyll fluorescence (Chl) and dissolved oxygen concentration (DO), Irish Grid Easting (Easting) and Irish Grid Northing (Northing). Error bars are SE, generated by 1000 separate randomizations

The algal data could not produce a useful model of the variation in invertebrate assemblage structure. Cross-validated errors all exceeded 1 and errors only increased with model complexity. In contrast, the location-specific environmental data had a model with minimum predictive error suggested with 4 variables. The regression tree for environmental variables is shown in Fig. S2.

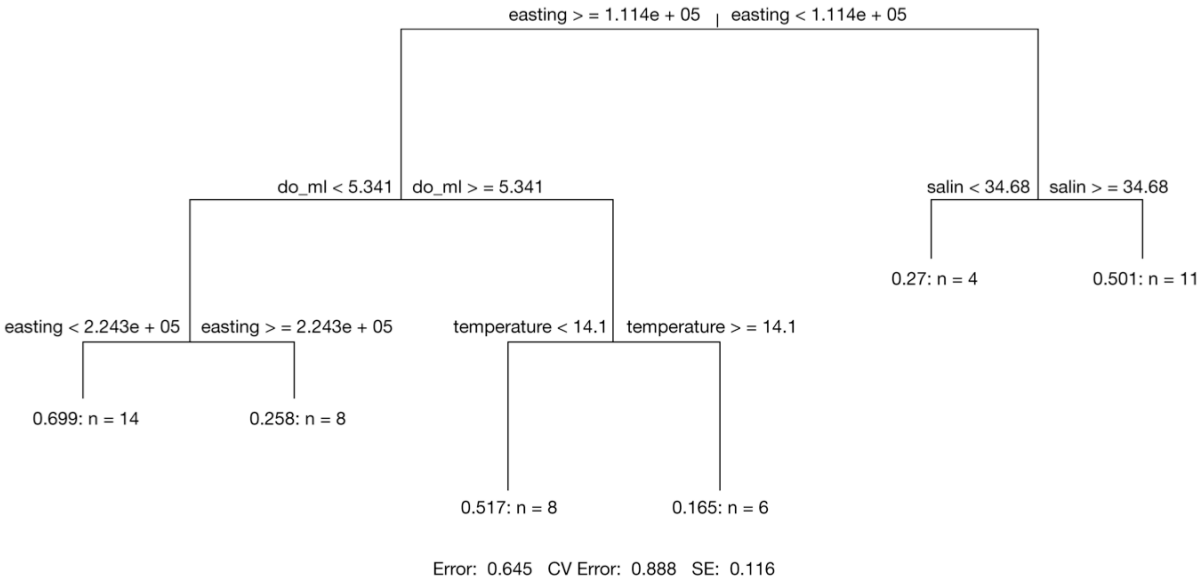

Fig. S2. Multivariate regression tree partitioning variance in invertebrate species composition (Simpson's dissimilarity) with chosen environmental predictor variables: easting, sea surface temperature, dissolved oxygen (do_ml) and salinity (salin). Number of rafts in each group is shown as n=. The tree explains 36% of the variability among rafts, with a cross-validated error of 0.88

LITERATURE CITED

De'ath G (2002) Multivariate regression trees: a new technique for modelling species–environment relationships. Ecology 83:1105-1117

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. www.R-project.org