Biogeosciences Discuss., 10, C883–C889, 2013
www.biogeosciences-discuss.net/10/C883/2013/

Biogeosciences
Discussions

Interactive comment on "Natural variability in hard bottom communities and possible drivers assessed by a time-series study in the SW Baltic Sea: know the noise to detect the change" *by* M. Wahl et al.

Anonymous Referee #1

Received and published: 9 April 2013

General comments

The study analyses the spatio-temporal variation in the structure of 12-mo-old encrusting assemblages developing on artificial substrata in the Baltic Sea. Seven locations, separated by 10s of Km, have been sampled for six year. The authors analyse taxonomic and functional diversity, in addition to environmental variables in an attempt to explain the emerging mesoscale patterns. Finally, a confidence interval (CI) of natural variability or "noise" (multivariate dissimilarity at lag 1) and the use of PRIMER's routine "RELATE" are proposed to assess the bounders of natural variability and direc-

tional changes in community structure, respectively. I think that the research questions are pertinent for the field of community ecology and that the results of the analyses contribute substantially to them. Also, including particle distribution modeling into the analysis of temporal variability of benthic assemblages is a interesting idea. I had, however, some problems to find support to the utility of the proposed CI, given the limited short duration of the time series. In addition, the manuscript would benefit from improving some details throughout the text, especially along the statistical analysis section.

The manuscript would also benefit from including a prediction. I understand that the aim was to describe patterns in biotic and abiotic variability in order to suggest hypotheses and predictions (i.e. a more inductive scientific approach), but there should be sufficient information out there about biotic and abiotic processes in the Baltic sea that may help to draw a couple of sensible predictions. One or few predictions may help the reader to better grasp the main conclusion of the manuscript.

As expressed above, I had some troubles to buy the CI proposed as confidence range of natural variability. The idea is very good—i.e., to provide a quantitative threshold of impact—, but I have the impression that the time span of the observations is too short to "sample" a comprehensive amount of variation. In addition, i feel that the authors do not provide enough methodological information regarding the calculation of the CI (see specific comments, below); If the CI is one of the highlights of the paper, then I would devote a bit more details regarding how it was estimated.

The last highlight of the abstract is the presentation of a statistical procedure to tease apart a signal from background noise in community structure. I was very excited regarding this proposition, and I expected to find a new way to analyse data. However, after reading the manuscript I was a bit disappointed, as there was no novel statistical procedure and the proposed analysis was basically the already available RELATE routine implemented in PRIMER. In addition, the discussion does not provide any analysis of the advantages of RELATE over other method (e.g. Mantel test or formal time-series

analyses). Therefore, I would suggest to downplay the statistical procedure and focus the abstract more on the results of the study.

Specific comments

2968, 13: “functional characteristics”. Please, provide a brief definition for this concept and functionality early in the abstract or introduction.

2969, 2: Positive effects of redundancy (=insurance) on stability. OK, but note that the effect of biodiversity on stability depends on the degree of asynchrony among species’ abundance temporal fluctuations. If species’ fluctuations are too synchronous, then compensation of lost species within functional groups is unlikely [e.g. Loreau and Mazancourt 2008)].

2969, 8: “This hypothesis...” Placed here, this sentence sounds like the aim of the study is to test the diversity-stability hypothesis.

2969, 26 and so on: This paragraph is a bit lose, maybe it could be combined with the last paragraph of page 2970.

1971, 3: Experimental assemblages—how these assemblages compare with those developing on natural substata? Are the experimental assemblages representative enough to estimate parameters of the natural community?

2975, 19-23: The description of the PCA is a bit complicated. For example, I do not know how to calculate the covariance of a “distribution”. Please try to rephrase it in order to make it clearer. In addition, PCA is a ordination method based on correlations. Were the assumption of correlation (e.g. linearity and normal distributions) assessed?

2976, 4: Variables with different units were analysed together with Bray-Curtis dissimilarites. Were these variables standardised before the analysis?

2976, 16-21: Use of ANOVA and null-hypothesis testing in general. Spatial autocorrelation among observations may results in a lack of independence among residuals,

leading to an underestimation of standard errors of estimators and overestimation of significance of the calculated statistic. Therefore, classical (frequentist) null hypothesis testing is not the appropriate way to analyse spatial data (e.g. Burnham and Anderson 2002). Legendre (1993) offers extensions of linear models that account for spatial autocorrelation by partitioning the variance in the dependent variable between the locality factor and environmental variables, for example. You could also explore the use of Mantel tests to assess the spatial variation in community structure. Of course, all this makes sense only after presenting the corresponding scientific hypotheses. Please also indicate if the factor time was included in these ANOVA, and if so, indicate how did you deal with temporal autocorrelation of observations.

2976, 24: Use of PCA. Again, it is necessary to indicate whether the assumptions of correlations were assessed. Alternatively, you might explore the use of Canonical Analysis of Principal coordinates (CAP, Anderson 2004), which is based on distance measures (e.g. Euclidean distance or Bray-Curtis) and therefore does not have restrictive assumptions such as linearity and normality in the distribution of errors.

2977, 9: "...species abundance data were averaged..." I think this procedure is inappropriate, because the mean of raw abundance data of individual species does not necessarily resemble the average assemblage across time (McArdle and Anderson, 2001). In addition, lack of normality can increase bias in the estimation of the mean. It would be advisable to generate the Bray-Curtis matrix, estimate the centroids of each cell (interaction between time and site), and then use them in the trend analysis.

2977, 12-14: Use of SIMPER. Maybe I missed something, but it seems that the PCA described above (2976, 24) and this SIMPER analysis have the same aim; that is, to identify the species that explain most of multivariate dissimilarities. If these analyses are redundant, then I would leave only one of them.

2977, 15-21: This is a good idea. As expressed above, however, use of classical hypothesis testing on spatial, autocorrelated data may be problematic. You should

either adjust your model to account for autocorrelation (e.g. a mixed model) or maybe use a different approach (see above). On the other hand, the use of these species, and not others, should be justified. Were them identified with SIMPER?

2977, 21-23: This is a bit confusing. First, you say that data were not stratified by site or year, but then you state that data were averaged across replicates at a given site and year. Please clarify what was actually done.

2978, 5-8: Albeit this is not strictly a description of method, I think it is nice to provide this information in order to refresh the reader.

2978, 12-15: Model selection method. Using significance (I guess you're using alpha=0.05) as threshold for retaining a term can be problematic because of the uncontrolled error Type I error caused by multiple hypothesis tests. You may want to use AIC to identify the appropriate model. Since you use R, you may want to check the meifly::fitall() function in order to run all possible models.

2979, 27 and so on: Species trends described here are hard to identify in the Fig. 4. As far I remember, you ran a PCA on abundance data. You could improve this analysis by including lat-long in a RDA (I would prefer a dbRDA a.k.a. CAP - see above) and show a biplot of species and lat-long.

2980, 19 and so on: Was year included in the ANOVA model, or are you considering the years as replicates? The reader could better know what was done if you include the degrees of freedom in the description of the results.

2982, 7-15: Interannual variability. It is unclear how this is presented in Fig. 8. Did you calculate an average lag-1 dissimilarity or was this value estimated by resampling? About the CIs, if the CI is the highlight of the manuscript, then it should be better described and defended in the text. For example, were the CIs obtained by means of resampling (e.g. bootstrap)?

2983, 10: What do you mean with "total abiotic variability"? Are you talking about the

Bray-Curtis dissimilarity matrix calculated from environmental data?

2983, 21 and so on: Effect of temperature on compositional variability. Since you are analysing spatial patterns, without the possibility of isolate any factor, it is risky to assume causality in regressions.

2983, 26: Please explain what do you refer to with "warm winters" here and along the discussion.

2984, 6: Prevalence corresponds to the proportion of observations that have a condition. This is different from percentage cover, as the latter is estimated for each observation.

2985, 15 and so on: This is a good point. If you have in mind the hypothesis of "presence of redundancy", you may want to fit a non-linear, asymptotic function between taxonomic and functional richness; a good fit would indicate redundancy.

2986, 12-27: Another interesting point. Compositional variability decreased with decreasing species richness. I would add that compositional variability is predicted to be negatively related with aggregate variability. Aggregate variability (i.e. the variability in aggregate, community-level properties like total biomass) was the metric that decreased with species richness in the study of Valdivia and Molis (2009). Therefore, the relationship between diversity on stability depends on the organisation level that we observe (e.g. Micheli et al 1999).

2987, 28 and so on: Please, present and defend the definition of "noise" earlier in the manuscript. I say this because noise can have different ecological meanings, depending on the degree of temporal autocorrelation in the variable.

2988, 25-27: I would replace this brief paragraph with a conclusive, take-home message of your work.

References: I have not checked for consistency between the text and reference list.

Technical corrections

2971, 5: Replace “which” with “that”.

2978, 7: I would suggest to replace the comma (,) with a semicolon (;).

2978, 12-20: I would swap the description of the maximal (full, isn’t?) model and how you selected the minimal adequate (reduced, isn’t?) model, placing first the full, and then the selection criteria.

2980, 13: Shouldn’t be **6** years?

2981, 10 and along the result section: 7a, b should be in uppercase, according to the figs.

2983, 14: I would suggest to combine Figs. 3 and 9B into one figure.

2984, 14 and along the discussion: “Explains”. Please, use past tense in the discussion when referring to the results of the present study.

2984, 25: “*Polysiphonia* suffers...” I am no sure if this is correct.

2985, 11: What is a “trophic situation”? I would use “food availability” or alike.