Biogeosciences Discuss., 4, S2477–S2481, 2008
www.biogeosciences-discuss.net/4/S2477/2008/

Biogeosciences Discussions

Interactive comment on "Dynamics of dimethylsulphoniopropionate and dimethylsulphide under different CO_2 concentrations during a mesocosm experiment" *by* M. Vogt et al.

M. Vogt et al.

Received and published: 21 January 2008

Response to referee #1 for manuscript bgd-4-3673-3699

General comments:

We thank referee #1 for his/her comments to our work and are pleased to hear s/he recommends our manuscript for publication. We would like to stress that referee #1's concerns regarding the accuracy of our paper writing are almost entirely limited to

problems with the figure captions and not with the manuscript. We have carefully corrected the figure captions and apologise for the oversight.

Specific comments:

page 3682 The significance levels for all tests have been added in the manuscript, together with a brief description of ANOVA's standard output variables. All statistical quantities were named according to standard naming conventions and are shown in italic font (df, F, σ), as opposed to the letters in regular font designating the 3 treatments (FF, F, P). However, the latter problem is resolved, as referee #2 suggested we rename our treatments, in order to be in agreement with the naming conventions for the other special issue manuscripts, so "P" became "1xCO_2", F "2xCO_2", and FF became "3xCO_2", so that treatment labels and statistical parameters can no longer be confused.

same page Wingenter et al. (2007) present the absolute differences in the integrated means between the groups and specifies error ranges with a variable confidence interval. We cite: "The time integrated average amount DMS was 26% ($\pm$ 10% having a confidence interval (CI) of $\approx$ 90%) and 18% ($\pm$ 10% or $\approx$ 80% CI) higher in the 2x and 3xCO_2 mesocosms, respectively (days 0-17)." Wingenter et al (2007) do not specify the statistical method used to derive these results in their manuscript. In a recent personal communication, the authors clarified that they used a Student's t-test to derive their results.

The difference between our results and those by Wingenter does not arise at the results level, but from differences in the interpretation based on statistical test results:

BGD
4, S2477–S2481, 2008

Firstly, we report very similar absolute differences of the integrated means (26% and 14%; days 0-22), but reject the significance of the difference at the 95% confidence level, a level generally recommended for use in bio-statistical applications (Cowles and Davis, 1982 and references therein). Hence, the acceptance of a difference at the confidence level of 90% and 80% with only 3 replicates for each treatment is associated with a high probability of committing type I errors (rejecting the null hypothesis ("no difference between populations") when in fact it should be retained). In the literature, significance is generally rejected at such low levels. In addition, varying confidence intervals should not be used to compare data.

Secondly, comparing the means of 3 populations in pairs of 2 with respect to a fixed factor (CO_2), as done in Wingenter et al. (2007), is statistically incorrect (see e.g. Zar, 1996, p. 177-178). Such a procedure introduces additional type I errors during the statistical analysis, i.e. the null hypothesis (no differences between populations) is rejected when in fact it is true. We use One-way ANOVA, a test designed to compare the means of 3 and more independent populations, which avoids the introduction of such errors. It assesses the ratio of the within-group variation to the between-group variations and in our case does not show a significant difference between the 3 populations at the 95% confidence level ($df = 2, F = 1.799, \sigma = 0.244, p = 0.05$).

For every decision based on statistical test results there is always a chance of committing type I (see above) or type II errors (retaining the null hypothesis when in fact it is should be rejected), a property inherent in every probabilistic theory. What we report is that we "do not have enough evidence to reject the null hypothesis" rather than we "know 100% for sure that there were no differences". The test we use does not detect significant differences, which is partly due to the low sample size (3x3 measurements) and partly due to the large spread between the 3 "replicates" of every treatment.

BGD
4, S2477–S2481, 2008

Interactive Comment

We have carefully chosen the statistical method we think is correct and we apply a stringent significance level for our decision-making. Thus, we are convinced that our approach is valid and that our results are solid within the chosen boundary conditions. For the two reasons specified above, we consider Wingenter et al.'s interpretation of their statistical results as problematic and we explain the origin of the discrepancy between their and our results carefully in the revised manuscript.

page 3685 These are indeed the means of the ratios and this error has been corrected in the revised manuscript.

page 3886-3887 Repetitions have been eliminated.

page 3691 We think that it is indeed possible to derived information about the effect of ocean acidification on marine ecosystems from studying the DMS cycle. For example, if we are able to separate the effect of physiological changes in phytoplankton on DMS patterns from differences caused by changes in bacterial activity, then this would give us information about which compartment of the marine ecosystem are likely to be predominantly impacted by changes in pH. Of course, this implies a careful study of the ecosystem and its functioning in order to understand the DMS cycle, which will in turn offer explanations why it is produced and which biological compartments exert most control on the patterns we see. Understanding pH induced changes of DMS concentration patterns will reduce the uncertainty for DMS model predictions. If pH were a major control on DMS, then this dependence would need to be included in future predictions of the effect of climate change on the DMS cycle. The CLAW hypothesis would then not only have to focus on temperature and light changes in the ocean, but would have to include the pH related changes in DMS production.

BGD

4, S2477–S2481, 2008

Interactive Comment

In order to avoid confusion, we have removed the last 2 sentences, as we did not consider them to be central to the main argument.

page 3696 ff All figure captions have been corrected.

References:

Cowles, S. M. and Davis, C., "On the origins of the .05 level of statistical significance", Amer. Psychol. 37, 553-558, 1982.

Wingenter et al., Geophys. Res Lett., 34, L05710, doi:10.1029/2006GL028139, 2007.

Zar, J. H., "Biostatistical Analysis", Fourth Edition, Prentice-Hall International, Inc., ISBN 0-13-082390-2, 1999.

Interactive comment on Biogeosciences Discuss., 4, 3673, 2007.