Biogeosciences Discuss., 6, C1427–C1430, 2009
www.biogeosciences-discuss.net/6/C1427/2009/

Biogeosciences Discussions

Interactive comment on "Statistical validation of a 3-D bio-physical model of the western North Atlantic" *by* M. K. Lehmann et al.

Anonymous Referee #2

Received and published: 3 August 2009

The paper describes the results of a high-resolution regional bio-physical model of the northeastern North American Shelf. The authors set out to quantitatively investigate an often-made statement, i.e. that increased ecological model complexity leads to an improved representation of ecological gradients. They use a number of statistical measures to show that a model with two phytoplankton functional types outperforms a simpler model with a single phytoplankton class. Overall, the tools are presented very well, and the results of the complex model are shown in some detail. The study is timely and the topic is certainly suitable for Biogeosciences. I have, however, a few major concerns with the way the results are presented and interpreted. These should be addressed carefully in a suitably revised version of the manuscript.

1. A major concern I have is that the comparison of simple and complex model may

not be fair. The authors seem to be biased in favor of their more complex model. Often they state good agreement when, in fact, the model is worse than climatology (in terms of their ME measure). They select only a subset of the subregions shown in Fig., which cover only about half of the model domain. Inspection of Fig.7 suggests, that the selected areas are those in which the model agrees best with the data. In all other areas the model chlorophyll may be worse than climatology. This may be the case for the simple model as well, but excluding these areas from the analysis does not give much confidence in the results. At least the Taylor plots (Fig.11,13 - why are the results of this study different in the two plots of the same property? Different years?) and Fig. 14 should show all model domains to allow for a complete model intercomparison.

2. Another issue is that the different ecological models seem to be embedded into different physical environments. Presumably, the physics of the most recent model is rather better than worse of that of the earlier publications. It remains to be shown that differences in the physics are not responsible for the differences in simulated chlorophyll.

3. Presumaby, the more complex model is more complex in many aspects (e.g., different process descriptions, different stoichiometric assumptions, different parameters, different number of phytoplankton functional types). This not described in sufficient detail. Is it possible to identify the model aspect that explains most of the improvement when going to the more complex model?

4. Data sets: As ocean color data are used, which are often compromised by clouds, the monthly mean of the data is usually not a true monthly mean because it consists of a few cloud-free images only. Comparing the "correct" monthly mean of the model with the incorrect mean of the data may adversely affect the results. A straightforward test would be to sample the model on where and when data are available. This ensures that you compare equal properties.

specific points:

page 5667, line 23. Does the Mellor-Yamada parameterization account for tidal mixing? Is the barotropic tide converted into baroclinic tides that may have sufficient vertical shear to be "seen" by the turbulence closure scheme?

page 5672, l.20: why "exceptionally" well? Compared to which "average" well?

page 5672, l.23: Here you say that the model performs well in tidally mixed areas. Is this because there is realistic tidal mixing in the model? Is it? If not, is there good agreement for the wrong reasons?

page 5673, l.1-2. As far as I know, ocean color algorithms are problematic everywhere. Can you give a quantitative error estimate?

page 5674. l.15-16./ Fig.4. There seems to be a systematic underestimation of low chlorophyll values during all seasons. Is this a systematic model deficiency or a problem with the data?

page 5674, l.23: What is the reason for the low SST bias? Too much diapycnal mixing? Presumably, this should correlate with a high bias in nutrient supply. Does the model need special tuning to still reproduce chlorophyll quite well?

page 5675, l.5: I am not convinced that the "mesoscale" argument is valid. Please show this explicitly. As you analyze monthly means, mesoscale features will to some extent be smoothed before the data analysis. Does any of the maps show mesoscale features? Perhaps it's more related to the different sampling of satellite data and model output? This should be checked!

page 5676, l.14: "...in much of the study region" is actually less than half of the study region.... If the areas are small enough, this would be expected for a pure random model as well. Can you show that your model is better than a random model?

page 5677, l.13ff; "near zero" with respect to what? Doesn't make sense without a scale! The Sargasso Sea seems to be dominated by negative ME values, i.e. the model is worse than climatology here. This should be stated clearly. Again, the

mesoscale-eddy argument is not convincing. Please show that this really applies! The statement that "chlorophyll dynamics are matched very well" (l.27) is a bit overoptimistic, given that simulated chlorophyll is worse than climatology in at least one subregion.

Interactive comment on Biogeosciences Discuss., 6, 5661, 2009.