**Biogeosciences
Discussions**

Interactive
Comment

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

M. K. Lehmann et al.

katja.fennel@dal.ca

Below we repeat the reviewer's comments with our responses interspersed in ***bold-italics***.

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

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

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.

Response: As we have pointed out in our response to reviewer 1, the primary objective of our paper was not to compare a simple and a complex model, but to implement a particular complex model (that has been used successfully in a coarse North Atlantic domain before) and to assess the model results in a statistically rigorous manner (this is done in detail in our manuscript). We included the comparison with the simpler model in Figures 12, 13 and 14 because we feel that this is a useful addition to our study and we suggest likely reasons for the improvement in model-data agreement in the oligotrophic region. We don't intend to make general statements about more complex models performing better than simpler models and modified the first sentence in the Conclusions to make it explicit that we are comparing two specific models. This sentence now reads as follows:

"The implementation of the size-structured model by Lima and Doney (2004) improved the agreement between model-predicted fields of surface chlorophyll with chlorophyll estimates from the SeaWiFS satellite compared to the model by

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Fennel et al. (2006, 2008)...”

We display only a subset of the regions for display in Figures 3, 5 and 6 simply because of space considerations. We have prepared the analog to Fig. 5 for all the other regions, which we will add to our paper as Online Supplement (also included with this response). The figure shows that model and data agree within one standard deviation everywhere, except for winter blooms in the southern MAB where the model underestimates peak chlorophyll concentrations. We also produced a comprehensive version of Fig. 14 that includes all model regions and will also be included in the Online Supplement (also included in this response). We would also like to point out that the whole domain is shown in the comparisons in Figures 2, 4, 7 and 8.

The Taylor plots in Figures 11 and 13 differ because Figure 13 includes results only from 2004 and 2005. We now state this explicitly in the caption.

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.

Response: The model results of the simpler model that we compare the present model against are from Fennel et al. (2008) (this is stated explicitly in our manuscript, e.g. in the captions of Figures 12, 13 and 14) and uses the latest physics. It is true that the model physics has improved from Fennel et al. (2006) to Fennel et al. (2008) and this is stated explicitly in the latter publication.

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?

Response: We agree that it would be interesting to know which particular aspect results in most of the improvement, but our study is not designed or intended to do this. One would likely need a data-assimilative framework like the one by Friedrichs et al. (2007) in order to do such an analysis rigorously. However, we would like to point out that optimization of biological parameters through variational data assimilation has not been undertaken for 3D models (not feasible at present because of the high computational cost). A detailed discussion of differences between the two models is not within the intended scope of our paper (note that both models are described in detail in the published literature). Our main objectives are described above in our first response (see also response to Reviewer 1).

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.

Response: This test was performed. We append a redrawn Figure 5 in which the calculation of monthly subarea means of chlorophyll concentration uses only pixels, which contain a satellite-derived chlorophyll value. The differences to the original time series in Figure 5 are minute and do not affect our conclusions.

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?

Response: Our model includes tidal mixing as stated on page 5667, line 15-17: "We [...] specify free-surface and depth-averaged velocity boundary conditions

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

with external values defined by HYCOM NCODA plus M2 tidal harmonics from an ADCIRC simulation of the western Atlantic (Luettich et al. 1992)."

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

Response: We replaced "exceptionally" with "very".

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?

Response: Yes, because of realistic tidal mixing in the model (also see response above).

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

Response: We modified the sentence starting on p. 5672, line 24 to the end of the paragraph as follows:

"However, it should be noted that algorithms deriving chlorophyll from ocean color are known to be problematic in optically complex near-shore waters and may overestimate true chlorophyll concentrations by as much as 100

O'Reilly, J.E. and 21 co-authors, 2000: Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4. In: O'Reilly, J.E. and 24 co-authors, 2000: SeaWiFS postlaunch calibration and validation analyses, part 3. NASA Tech. Memo. 2000-206892, Vol. 11, S.B. Hooker and E.R. Firestone, Eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 9-23.

Yoder J. A., Schollaert S. E., and O'Reilly J. E.: Climatological Phytoplankton Chlorophyll and Sea Surface Temperature Patterns in Continental Shelf and Slope Waters off the Northeast U.S. Coast, Limnol. Oceanogr., 47, 672-682, 2002.

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?

Response: The only systematic underestimation we can discern in Fig. 4 is directly at the southwestern boundary where the Gulf Stream enters. This is a boundary artifact. We added the following text to our manuscript on page 5672: "A boundary artifact is noticeable at the southwestern boundary where the Gulf Stream enters the model domain and simulated chlorophyll concentrations are lower than observed."

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?

Response: The SST bias is slight. We have no reason to believe this is due to a problem with diapycnal mixing and no reason to suspect that this correlates with a high bias in nutrient supply. We would like to point out that SST is reproduced extremely well by the model (e.g. Fig.s 2 and 3).

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!

Response: Since the same sampling is used for calculation of monthly means in Figures 2 and 3 as in the calculation of RMSE we don't think that the calculation of monthly means can explain this. We feel that a mismatch of features is a plausible explanation and suggest it as such. We do not state this as fact. An explicit analysis of RMSE and other metrics for pattern comparison of satellite and model data is in preparation as separate paper by a subset of the authors.

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?

Response: We replaced "in much of the study region" with "in a large fraction of the study region". We feel that the latter phrase is a sensible characterization of the yellow area in the right panel of Figure 7. Although we did not show explicitly that our model performs better than a purely random model we feel it is highly likely.

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.

Response: ME is a dimensionless quantity; ME of zero is defined as is described on page 5671, lines 11 – 14. As requested we replaced the phrase "negative ME is most frequent in the Sargasso Sea" with "negative ME for SST dominates in the Sargasso Sea". We removed reference to chlorophyll (this was a typo). See comment above concerning the mesoscale variability argument.
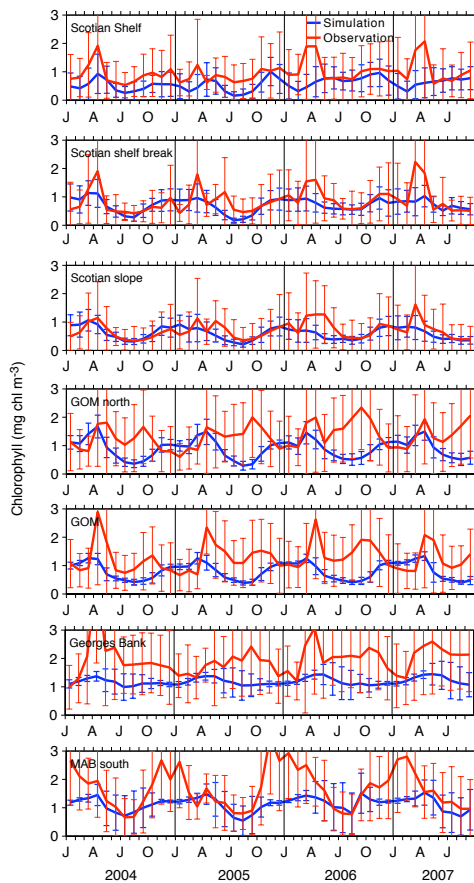
Interactive comment on Biogeosciences Discuss., 6, 5661, 2009.

Fig. 1. An extended version of Fig. 5.

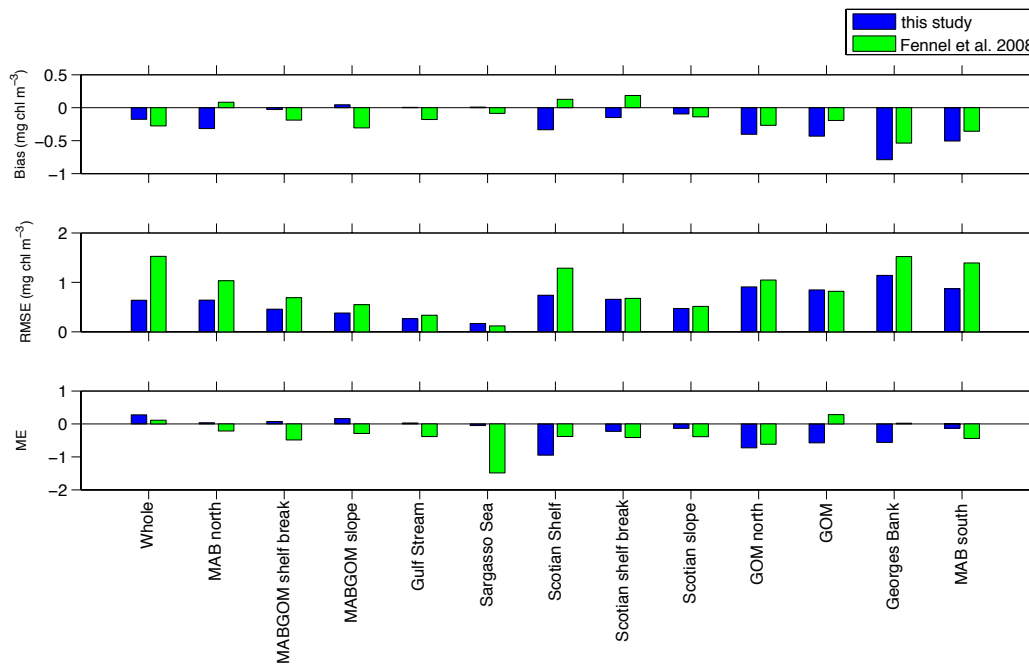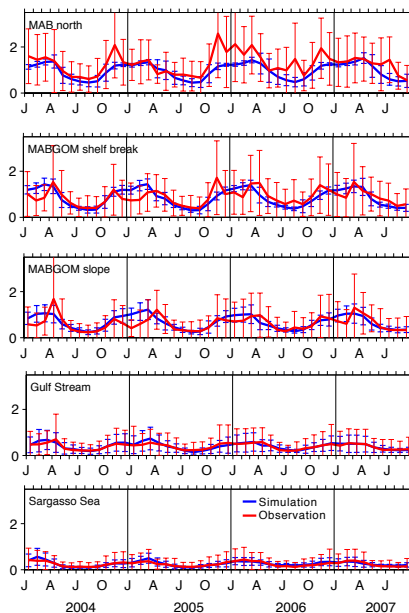**Fig. 2.** An extended version of Fig. 14.

Fig. 3. Another version of Fig. 5, which now only includes model-pixels that correspond to cloud-free satellite images.