

## ***Interactive comment on “BFM17 v1.0: Reduced-Order Biogeochemical Flux Model for Upper Ocean Biophysical Simulations” by Katherine M. Smith et al.***

**Anonymous Referee #2**

Received and published: 21 October 2020

Review on “BFM17 v1.0: Reduced-Order Biogeochemical Flux Model for Upper Ocean Biophysical Simulations” by Katherine M. Smith, Skyler Kern, Peter E. Hamlington, Marco Zavatarelli, Nadia Pinardi, Emily F. Klee, and Kyle E. Niemeyer

**Summary:** The presented manuscript compares two model versions of a pelagic biogeochemical model, embedded in a 1-dimensional ocean model. The first model version refers to the BFM-model with 56 state variables (BFM56), while a second, reduced version (BFM17) includes only 17 state variables for the biogeochemistry. The authors compare the results of both model versions to climatological observations at the sites Bermuda Atlantic Timeseries (BATS) and Bermuda Testbed Mooring (BTM).

C1

**Major comments:** The manuscript addresses the important, so far unresolved question on the necessary complexity of pelagic biogeochemical models. The topic is per se interesting. Still I have some major concerns:

(1) BFM17 excludes some processes relative to BFM56. This reduces by construction the applicability of the newly developed model (e.g., coastal shallow regions might not be represented so well anymore, due to the lack of the sediment processes, while simulating global phosphate distributions might still work very well). I would thus appreciate a clear statement for which purposes the new model was designed. Also, the overall aim of the presented results did not get entirely clear to me: is the idea of BFM17 to mimic/resemble certain processes of BFM56 or is BFM17 a model on its own, which might perform even better than BFM56 relative to the observations?

(2) In my eyes the study would need a better illustration that the presented model is up to the tasks it is designed for – or at least state-of-the-art. Currently, both model versions are compared to climatological, station based in-situ observations close to Bermuda. I am surprised that both model versions don't do better (e.g., for oxygen the explained variances are in the order of 5-15%, cf. Fig.7). The study might get more convincing if the authors could at least outline what goes wrong.

(3) When reading the model description, I found it rather difficult to keep overview what exactly was done. I would suggest to summarize the main model assumptions somewhere in the beginning of the model description (the equations might as well be collectively listed in an appendix). E.g. it did not become clear to me how the reduction of state variables was motivated and how the new state variables and parameter settings were chosen.

(4) I agree in general with the approach to compare both model versions initially in a low-dimensional environment, which is less costly to run than a full-fledged ocean model. Unfortunately, the authors do not take advantage of such a computationally cheap setup, e.g., by testing the sensitivities of both models towards environmental

C2

changes in idealized experiments. Such an approach would be feasible without too much effort and could explore how much BFM17 resembles BFM56 under various conditions (preferably also in well fertilized regions).

I am, however, sceptic about comparing the 1-dimensional model results to the observations. First, the comparison captures only one very special region – while I assume that the model is rather designed to answer questions on a global scale and I regard a 3-dimensional ocean component as more appropriate. Secondly, I fear that ocean mixing might not be realistically represented and that the restoring below 150m might induce spurious effects in the 1-dimensional model. In the real ocean, regions with intense sinking of organic matter and upwelling of nutrients are often spatially decoupled. How I understood the presented setup, the 1-dimensional model needs to balance both processes in a single box and thus induces some eddy driven transports (line 504ff + line 515ff), to reach some kind of steady-state (line 526). The latter seems rather arbitrary to me (please correct me in case I got things wrong). My scepticism is supported by the rather poor representation of oxygen in both model versions (because oceanic oxygen concentrations close to the surface are strongly impacted by water temperature and mixing). Also, I am puzzled that both models deviate strongly from the observations even at 150m – i.e. close to the prescribed boundary conditions at depth (cf., Fig. 6).

Minor comments:

Line 1: I find the expression “reduced-order model” in the presented context a bit confusing. The authors seem to use this term whenever a certain (non-defined) number of state variables is undercut. It might make sense to define the expression somewhere.

Line 7: I expect large differences in the tropical Pacific when disregarding iron in BFM17 (unless the top-down control is too strong). Depending on the region of interest the authors might consider to use at least a fixed, prescribed iron concentrations.

Line 79: The “accuracy” certainly depends on the underlying measure. What is re-

C3

garded as important will certainly depend on the selected region and the purpose of the model.

Line 83ff: So which parameters were calibrated and how was this done?

Line 135ff: I agree that presenting all equations is important. Still I find it impossible to keep overview in the presented form. The rationale behind the construction of the new BFM17-model does not at all become clear to me.

Line 380ff: I find it very hard to understand the physical setup and would appreciate if the main assumptions were summarized in the beginning of this Subsection. Currently, I get lost in too many details. Also, it should be clearly stated that only the upper 150m are simulated. What is the maximum mixed layer depth?

Line 520: I would suggest to describe initial values and boundary conditions in relation to the description of the physical setup.

Figure 2 What is the purpose of this figure? Why are not both model versions depicted and what do these experiments add to the conclusions?

Figure 3+4: Again – the purpose of showing these figures does not become clear to me (e.g., Fig.4 vs. Fig.6?)

Figure 6: I see substantial differences between both model versions – even in the base currency “phosphate”. How do the authors rate which differences matter? Why are the models so different from the observations at 150m (close to the bottom boundary conditions)?

Figure 7: Compared to the huge differences between the reference model and the observations, both model versions seem indeed rather similar (while in fact I regard these differences as substantial). I would suggest to add a more intuitive representation of the differences between both model versions in the corresponding text, e.g., % relative to the reference model.

C4

Table 8 + line 568ff: I like the idea behind this comparison to illustrate that the presented model is state-of-the-art. Still comparability needs to be ensured. Do all cited models refer to the same sampling in time (please specify) and state variables? Otherwise these numbers are not comparable. Also, the underlying physical model setups should be listed as well. It is much harder to simulate a specific timeseries in a 3D context than to be able to tune the model at a single location (looking at Figure 1 and the tables, I would expect at least 40 poorly known parameters to match 12 moments in time – this should work perfectly at least for one state variable at one depth level).

Line 605ff (Conclusion): I would suggest to add a clear statement for which purpose the model BFM17 was designed and what it could be used for. Subsequently, the study would in my eyes benefit from an outline why the authors think that the model is up-to-the-task, or at least state-of-the-art.

Line 615: As outlined before, I do not agree on the generalized, conclusive statement “the model does well”.

~

---

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2020-134>, 2020.