

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

Katherine M. Smith et al.

peh@colorado.edu

Received and published: 17 December 2020

Response to Reviewer 2

We appreciate the constructive and insightful comments from the reviewer. A number of changes have been made to the text at the reviewer's request and we feel that the revised paper is now substantially clearer and stronger as a result. Detailed replies to each of the reviewer's points (in blue italics) are provided below.

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

C1

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).

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.

Major Comments:

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.

We agree with the reviewer that the applicability of BFM17 is reduced in comparison to BFM56 and, in order to more clearly state the range of applicability of the model, we will state in the revised manuscript that “BFM17 is a pelagic model intended for upper-thermocline, open-ocean, non-iron or silicate limited, oligotrophic regions”. With respect to the purpose of the model, we will clarify that “BFM17 has been developed for the primary intended use within high-resolution, high-fidelity idealized models, such as a Large Eddy Simulation (LES), for process, parameterization, and optimal parameter estimation studies (BFM56 is far too costly for this kind of application). Thus we do not validate its efficacy as a global BGC model (though its relative agreement with BFM56 suggests its potential to do quite well as such) and provide the caveat that it is missing potentially important processes for such an application.”

C2

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?

We agree with the reviewer that this was not entirely clear and we will add the following text to the revised manuscript: "Additionally, we later compare to the original BFM56, not in an effort to demonstrate that BFM17 is superior and thus should be used in place of BFM56 in global applications, but rather to demonstrate that, although it is reduced in complexity, BFM17 is equally appropriate for use in seasonal process, parameterization, and optimal parameter estimation studies for which a more complex model such as BFM56 may be too computationally expensive."

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.

We agree with the reviewer that not all presented fields are reproduced by BFM17 and BFM56 with great accuracy (e.g., oxygen), and we will clarify that the exact reproduction of all fields was not our goal with BFM17. In the revised manuscript, we will state that "... oxygen is historically a difficult field to reproduce in BGC models of any complexity. It is hypothesized that this is due in part to a lack of physical processes being represented in mixing parameterizations. For example, our model struggles with oxygen, in part, because the second-order mixing scheme of Mellor and Yamada (1982) lacks sufficient resolution of the winter mixing using just the monthly mean temperature and salinity. However, since it is often not included or presented at all in models of similar complexity to BFM17 (i.e.

C3

models reduced enough to reasonably couple to a high-fidelity, high-resolution physical model), studies that explore this hypothesis have been historically difficult to undertake. Thus, we include oxygen in BFM17 and present our results here to illustrate this exact point and lend motivation to developing and using a model such as BFM17 to study what physical processes are missing from mixing parameterizations and how they can be better represented."

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.

We agree that our previous model description was potentially difficult to follow. We will thus move the detailed equations and OD test to appendices in the revised manuscript. This will help to focus the reader in the main model description section on the model assumptions and choice of state variables.

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 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).

We agree with the reviewer that this kind of comparison would be computationally cheap and easy to perform. However, since BFM17 was primarily developed for use in high-resolution, high-fidelity physical models (something BFM56 cannot be

C4

used for because of its cost) and not for use as a global or column model, this comparison is beyond the scope of the model description and baseline demonstration that are the primary objectives of the present paper. Nevertheless, we do agree with the reviewer that this is an important direction for future research, and we will clarify this point in the revised manuscript with the following text: “It should be noted that the primary focus in the present study is to demonstrate the efficacy of BFM17 as an accurate BGC model in high-resolution, high-fidelity upper-ocean process, parameterization, or optimal parameter estimation studies, by comparing to results from observations and BFM56; as such, here we only consider one open-ocean location (i.e., the Sargasso Sea) and make the assumption that its ability to reproduce important and difficult key behaviors at this location, such as the initial spring bloom and subsequent subsurface chlorophyll maxima, are indications of its efficacy for use as a process study model. However, the correspondence between BFM17 and the more general BFM56 provides confidence that the reduced model will also prove effective at modeling other ocean locations and conditions, and exploring the range of applicability of BFM17 remains a subject of future research.”

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.

We apologise for any confusion we might have caused by not clearly stating that BFM17 was not developed for use as a global model. We will clearly state this in the revised manuscript: “BFM17 has been developed for the primary intended use within high-resolution, high-fidelity for process, parameterization, and optimal parameter estimation studies (BFM56 is far too costly for this kind of application). Thus we do not validate its efficacy as a global BGC model (although its relative

C5

agreement with BFM56 suggests its potential to do quite well as such) and note that it is missing potentially important processes for such an application, which we elaborate on shortly.”

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 504 + line 515), 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).

We agree with the reviewer that any 1D column model is unlikely to realistically represent all processes, particularly those associated with mixing and spatially decoupled processes. However, the 1D physical model is only used here for validation purposes, and a primary reason for developing BFM17 is to explore issues such as the effects of mixing and temperature in the surface ocean on oxygen or how “patchy” (or spatially decoupled) nutrient upwelling effects the distribution and fate of a phytoplankton bloom using more physically realistic fully non-hydrostatic 3D simulations. We will elaborate on this in the text of the revised manuscript through the following additions: “Dissolved oxygen is of particular interest as it is historically difficult to get correct using BGC models of any complexity. It is hypothesized that this is in part due to missing physical processes in the mixing parameterizations used in global and column models. A primary goal of developing BFM17 is to use it within high-resolution, high-fidelity physical models to understand what these physical processes are and how to properly account for

C6

them in mixing parameterizations. Dissolved and particulate organic matter, each with their own partitions of carbon, nitrogen, and phosphate, are also included to account for nutrient recycling and carbon export due to particle sinking. Another primary goal of developing BFM17 is to explore how spatially decoupled (or “patchy”) processes, such as the sinking of organic matter and the subsequent upwelling of multiple recycled nutrients, not just nitrate, effect the fate and distribution of a phytoplankton bloom.”

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).

We agree with the reviewer that the nutrient profiles (particularly nitrate) deviate strongly from the observations at depth. Initially, in order to represent the recycling of PON and surface layer nitrogen replenishment from depth, we included a bottom boundary condition for ammonium based on the gradient in PON observations (direct observations of ammonium at the BATS location were lacking). However, as the reviewer points out, this process can be spatially decoupled and thus the assumptions we made associated with this boundary condition are potentially unsupported. In light of the fact that we do not know the exact relationship between the sinking PON and up-welled ammonium, we have now chosen to set this boundary condition to zero (consistent with the idea that most of the nitrogen flux from depth is occurring as nitrate flux). The nutrient results for both models now more closely match observations, both overall and at depth. The new correlation coefficients and *rms* errors in parentheses are now as follows for BFM17: Chlorophyll = 0.63 (0.08), Oxygen = 0.37 (31.18), Nitrate = 0.94 (0.22), Phosphate = 0.91 (0.01), PON = 0.85 (0.15), NPP = 0.93 (0.26). For BFM56, they are: Chlorophyll = 0.60 (0.10), Oxygen = 0.18 (21.84), Nitrate = 0.93 (0.16), Phosphate = 0.93 (0.005), PON = 0.85 (0.11), NPP = 0.87 (0.63).

C7

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.

We agree that this expression is confusing. We will remove this description and replace it with “upper-thermocline, open-ocean” throughout the revised manuscript in order to be clear as to how exactly the model is “reduced”.

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.

We agree with the reviewer that BFM17 may not perform well in regions where iron-limitation is important, such as the tropical Pacific and the Southern Ocean. We believe that the following altered statements in the revised manuscript convey that BFM17 was not developed, nor has it been qualified, for used in such regions: “To reduce the overall computational cost and to focus on upper-thermocline, open-ocean, and non-iron or silicate limited conditions...” and “BFM17 is a pelagic model intended for upper-thermocline, open-ocean, non-iron or silicate limited, oligotrophic regions...”. We will also note that “Iron is omitted from BFM17, limiting the applicability of the model in regions where iron components are important, such as the Southern Ocean and the tropical Pacific. Thus, if used in such a region, at least a fixed concentration of iron might be needed (though this has not been method has not been validated within BFM17).”

Line 79: The “accuracy” certainly depends on the underlying measure. What is regarded as important will certainly depend on the selected region and the purpose of the model.

C8

We agree that “accuracy” is a relative term that depends on the intended goal of the model. We will thus adjust this statement in the revised paper to read: “It should be noted that the primary focus in the present study is to demonstrate the efficacy of using BFM17 as an accurate BGC model for high-resolution, high-fidelity upper-ocean process or parameterization studies, by comparing to results from observations and BFM56; as such, here we only consider one open-ocean location (i.e., the Sargasso Sea) and make the assumption that its ability to reproduce important and difficult key behaviors at this location, such as the initial spring bloom and subsequent subsurface chlorophyll maxima, are indications of its efficacy for use as a process study model.”

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

We will provide in the revised manuscript an explanation of how the parameter values were determined. Briefly, the values used in this paper were based largely on the values established by Vichi *et al.* (2003, 2007, 2013) for BFM56. We are also currently pursuing computational optimization of these parameters to minimize model error when compared to the BATS observational data; this future work will be mentioned in the Conclusions section of the manuscript.

Line 135: 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.

We agree with the reviewer that this section was rather long and confusing. We will move the detailed equations and OD test to appendices at the end of the manuscript. This section will thus be focused on just the assumptions and choices of state variables for clarity with references to the appendices for further detail.

Line 380: I find it very hard to understand the physical setup and would appreciate if

C9

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?

We will summarize the physical setup assumptions and clearly state that only the upper 150m are simulated in the revised manuscript: “In order to focus on the upper-thermocline, open-ocean regime BFM17 was developed for, the physical model only extends 150 m in depth and diagnostically calculates diffusivity terms based upon proscribed temperature and salinity profiles from BATS. While a 1D physical model is unlikely to resolve all processes relevant for biogeochemistry in the upper-thermocline, we have made additions, such as large-scale general circulation and mesoscale eddy vertical velocities and relaxation bottom boundary conditions for nutrient upwelling, to better represent missing processes.” Additionally, we will also state that the maximum mixed layer depth from the climatology is approximately 125 m based upon a $\Delta 0.02 \text{ kg/m}^3$ criteria in Section 4.2.

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

Initial and boundary conditions are described in Section 4.4; the great length of the previous model description may have obscured these definitions, and by moving many of the model equations and OD tests to the appendices, the specification of initial and boundary conditions should now be more readily apparent to the reader.

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?

We again agree with the reviewer that differences in how the physics are represented in a 1D model can greatly affect the results. We thus include the OD

C10

test and this figure to show that, without a specific physical model, BFM17 reproduces a stable and well-behaved seasonal cycle similar to that observed in the BATS region. We will also include the following qualifications in the revised manuscript: “As an initial test of BFM17 (without the influence of any particular physical model) . . .” and “. . . indicating that a self-consistent and stable seasonal cycle with reasonable ecosystem values can be attained by the model, regardless of its coupling to a physical model.” We do not include results from BFM56 in this figure as it is already a validated, established model [see the cited papers by Vichi *et al.* (2003, 2007, 2013)] and is not the focus of this manuscript. It should also be noted that this figure will be moved to an appendix at the end of the revised manuscript, de-emphasizing its overall importance, in line with the reviewer’s comments.

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

Figure 3 shows the monthly input forcing profiles used in the 1D coupled model. Given that they are profiles derived here in this study (and thus not used in previous studies that could easily be referenced) and they are more easily described in figure format, rather than in text, we feel that Figure 3 is an important figure that should remain in the manuscript. Figure 4 shows the target climatological fields we aim to reproduce with the models. However, given that the top row of Figure 6 shows this same information, we agree that Figure 4 is redundant and will take it out of the revised manuscript.

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)?

C11

We again agree with the reviewer that the nutrient profiles deviate strongly from the observations at depth. This was due to some assumptions we made (described in detail in an earlier response) for a particular bottom boundary condition that we now feel are unsupported. We have changed this (again, details described in an earlier response) and the models now more closely reproduce the observations at depth, with improved correlation coefficients and *rms* errors. Additionally, we do not believe that we rate which differences matter, but rather present the results for all comparisons and allow the reader to make that decision for themselves based upon their intended application or study focus.

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.

We agree with the reviewer that in comparison to the observations, the two models seem rather similar. To quantitatively support this conclusion, following the reviewer’s suggestion we will present in the revised manuscript correlation coefficients and *rms* error values for BFM17 in comparison to BFM56 (similar to what was done for the BATS target fields) for each of the presented target fields. The correlations between the two models are as follows: Chlorophyll = 0.85, Oxygen = 0.56, Nitrate = 0.99, Phosphate = 0.99, PON = 0.95, NPP = 0.97.

Table 8 + line 568: 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

C12

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).

We agree with the reviewer that whether these models were comparable was not clearly stated. We will elaborate on each of the three models in the revised manuscript through the following additional text: “All models used climatological monthly mean forcing from the BATS region and reported climatological monthly means for their results. Care was taken to ensure that the same variable definition was compared between all models. Ayata *et al.* (2013) used a similar 1D physical model as was used here, while Spitz *et al.* (2001) and Fasham (1990) used a time-dependent box model of the upper-ocean mixed layer. As such, correlations and RMS error values for comparison to Ayata *et al.* (2013) are computed over the entire domain (Ayata *et al.* (2013) calculated their metrics over the top 168 m of their domain) and for comparison to Spitz *et al.* (2001) and Fasham (1990) were calculated only within the mixed layer (defined as the depth at which the density is 0.02 kg/m³ greater than the surface density) and are shown as separate columns in Table 3.”

Line 605 (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.

We agree that the original manuscript could have made this more clear. We now provide clear statements both at the beginning and in the conclusion of the revised manuscript that state this: “BFM17 was developed primarily for use within high-resolution, high-fidelity 3D physical models, such as LES, for process, parameterization, and optimal parameter estimation studies, an application for which its more complex counterpart BFM56 would be much too costly for.” and “BFM17

C13

is now of a size that it can be efficiently integrated in high-resolution, high-fidelity 3D simulations of the upper ocean, and future work will examine model results in this context”

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

We agree that this statement was overly general, and so we will provide a more detailed and nuanced statement that the model accurately captures the subsurface chlorophyll maximum and bloom intensity observed in the BATS data, and that results from BFM17 are similar to those from BFM56.

C14