

Interactive comment on “A perturbed biogeochemistry model ensemble evaluated against in situ and satellite observations” by Prima Anugerahanti et al.

Anonymous Referee #2

Received and published: 24 April 2018

The authors present a study in which an ensemble of biogeochemical 1D ocean models is generated by using different functional forms for key model pathways. The goal is to highlight the structural uncertainty in biogeochemical models as well as the benefits of the ensemble in fitting various datasets.

general comments:

The manuscript attempts to show two aspects: (1) there is a high level of structural uncertainty in biogeochemical models and (2) the uncertainty can be exploited to better fit a range of different observations. In my opinion, the authors succeed in providing evidence for first aspect but I have doubts about the second: all comparisons of the

C1

ensemble are based on a default run that does not seem to perform very well. Other studies have shown that 1D models with the same parameter values do not perform well across multiple locations but here the same parameter values appear to be used across all stations. Have the parameters of the default run been optimized to fit the datasets used in this study? The results of the default run can have knock-on effects on the ensemble: in multiple parts of the manuscript the authors note that when there is a large bias between the model (ensemble) and the observation, that the ensemble spread is too low when really other model aspects may be to blame for the bias. In other words, problems with the parametrization, the physical model, or the 1D nature of the model cannot be explained by structural uncertainty in the biogeochemical model.

When looking at Figure 1, I noticed that the linear function in (c) provides a bad fit to the other functions and that all functions are shown on a log scale. I am wondering if a log-transformation has also been used in the function fitting exercise in Sections 2.1-2.3? If not, I would recommend that this should at least be tried as the procedure could otherwise overemphasize the fit at high tracer concentrations which may explain the slope of the linear function.

Parts of the manuscript are difficult to understand and especially the introduction contains a large number of long run-on sentences. I have highlighted some of these instances in my specific comments below but I would recommend that the authors read through the manuscript again carefully and revise some sections for clarity.

specific comments:

abstract:

I1: "mathematical structure": What exactly does this refer to? The model formulation? I would suggest to rephrase or an improved explanation.

I3: "intermediately complex BGC model" -> "BGC model of intermediate complexity"

I9: "using phytoplankton phenology (...) and other statistical measures": phytoplankton

C2

phenology is not a statistical measure.

I11: Is this the range found in the ensemble (as opposed to e.g. different coastal stations)? Please make this explicit.

I14: "the errors are mostly reduced": This is not clear: model misfit with respect to the in situ obs is smaller for the ensemble mean/median than the model with standard parameters? I suggest to rephrase.

I15: Here a narrow spread is reported, a few lines above a "large" spread was described.

p2:

I7: This reads like the forecasting systems are having an impact on ocean biogeochemistry. The climate change aspect of the sentence reads like a repeat of sentence in line 2. Please revise for clarity.

I12: Even NPZ models represent "several" processes. Please be more precise.

I16: There can be spatial variability without iron!

I29: "only small perturbations are usually produced even with large variations in parameter values" This is a very strong statement and very much depends on what a "large variation" entails. Perhaps weaken the statement and just make the point that structural uncertainty is often larger than parametric?

p3:

I13: "linear density-dependent mortality produces the most significant differences when applied to diatoms": What exactly does this mean? Please revise.

I18: "However, not all processes give significantly different model outputs." The next sentence seems to imply that the differences maybe due to very similar inputs, can this effect thus really be attributed to the process?

C3

I22: "However, it is still unclear what will happen if formulations of all the core processes [...] are perturbed together." The preceding sentence is very general and I would say it is quite clear that the perturbations of all core processes would also "give rise to different effects". I would suggest to rephrase.

p4:

I3: "using all possible functional combinations": Given that there can be an infinite amount of different functional forms, I would suggest to rephrase this sentence. (Later on it becomes clear that only a few functional forms are considered.)

I22: Mention right away that Table 1 contains the equations for all functions.

I29: Mention that "T" is temperature here.

I32: "the default": Is this U_1?

p5:

I4: "The small microzooplankton": this makes it sound like there are small and large microzooplankton. Use something like "The small zooplankton category consists of microzooplankton..."

I5: Is "non-diatoms" referring to the "smaller phytoplankton" in the previous sentence?

I8: This is the third time Michaelis-Menten and Holling type II are mentioned together.

I9: "II" -> "III"

I9: Why say "hereafter G_1/G_2" when "Holling type II/III" is used throughout the text?

I19: Was the shape of the curves adjusted again? If so, how?

I29: What is a "distinct trend" here?

I30: It is not clear to me how the linear function was made to match the others. Figure 1(c) seems to suggest something went wrong. Or are large values here simply

C4

overemphasized in the fit?

p6:

l31: How long is the spin-up period for the runs?

p7:

l9: Why this lengthy comment about physical data assimilation? Is the capping done to remove the perceived negative influence of the physical data assimilation? What about rapid shifts in mixed layer depth which is also an input of the model, may also be affected by physical data assimilation and may also drastically change nutrient concentrations in the model. It is also not quite clear how the mixed layer depth influences the 1D model.

l26: It would be good to mention these locations the first time the stations are introduced.

Sec 2.5.2: Here the description is confusing, it goes from initial conditions to validation data, back to initial conditions and then to validation data.

p8:

l8: "one of MarMOT's test stations" What exactly is this test station?

p9:

l13: "These have been done at the five oceanographic stations which can be classified into three regional types:" This has been mentioned before.

l21: Mention PAP.

p11:

l4: How well does NRR work with a significant bias?

l10: "these members use functional combinations ..." The notation for the combinations

C5

is not clear here

Table 1: It does not make sense to call μ 's the maximum rates here.

Fig 1: Use "U_1" etc. here.

Fig 7: A better description of the x and y axes are needed. Why do b,d,f and h have no y-axis? Use the same color scale across all stations. Same comment applies to Fig. 8 where the font becomes too small.

Interactive comment on Biogeosciences Discuss., <https://doi.org/10.5194/bg-2018-136>, 2018.

C6