Biogeosciences Discuss., referee comment RC1
https://doi.org/10.5194/bg-2022-7-RC1, 2022


# Comment on bg-2022-7

Vassilios Vervatis (Referee)

Referee comment on "Quantifying biological carbon pump pathways with a data-constrained mechanistic model ensemble approach" by Michael R. Stukel et al., Biogeosciences Discuss., https://doi.org/10.5194/bg-2022-7-RC1, 2022

Title: Quantifying biological carbon pump pathways with a data-constrained mechanistic model ensemble approach.

by Stukel et al.,

General comments:

The study investigates the pathways of biological carbon pump, performing ensemble simulations of biogeochemical model parameterizations, constrained by data assimilation with the use of several data types obtained from Lagrangian experiments.

The ms. is well written and well structured, being very informative for the processes controlling BCP pathways. The idea of using an ensemble-based approach to quantify model parameter uncertainties and constrain them by data assimilation is innovative and the general approach is meaningful.

I am not an expert on the various aspects of biogeochemical model parameterizations, but I understand the most important feedbacks between the different compartments of the BGC model and the importance of the physical forcing. In this work, there are some assumptions that can be considered as simplifications (e.g., 1D model, physical forcing, length of simulations etc.), but in my opinion there are all justifiable and there are other novelties that compensate for the study approximations.

Overall, I find the ms. worthy of publication in Biogeosciences after minor revisions. Please find below a list of comments that I would like the authors to address.

Specific comments:

1) P6, L183 and L188. Vertical eddy diffusivity is varying with depth or is set constant? Please clarify.

2) P6, L196. Which model variables, in addition to the euphotic zone, you could have simulated? Please clarify why those variables were excluded from the simulation (e.g., computational cost?) and explain how this may affect model uncertainty in relation to other error assumptions.

3) P7, L237 and P8, L252-264. In the context of data assimilation, observational errors are often considered as a combination of instrument and representativity errors, the latter usually being the most important of all. The authors here quantify observational errors as the standard deviation of their measurements and/or the instrument error; if I understood correctly, representativity errors are not considered here. Are these errors relevant in terms of magnitude with observation representativity errors?

4) I am confused with the threshold limit "detlim" referred as "experimental detection limit". How this threshold is defined? I see that the "detlim" depends on indeces i,j,k and that k-index is not an option for the observations; why? I think the authors should provide more explanations regarding the "detlim" threshold, because the cost function decrease (after several iterations) largely depends on this (at least this is what I understand from the definitions of J(p) and error_i,j,k at the end of page 7).

5) Overall, in the data assimilation Section 2.4, it is not clear to me which model variables consist the control vector e.g., is it the same with the model state vector described in Table 1 (or not)? Please clarify.

6) P18, L669-670 *"our work shows that very different parameter sets can result in similar cost function values, despite generating distinctly different model outputs"*. This is an interesting result, but what does it means exactly (especially here where the cost function is different wrt variational approaches)? Please elaborate.

7) P19, L690-692 *"A further study (Anugerahanti et al., 2020) simultaneously perturbed physical circulation fields and the biogeochemical model and found that results were most sensitive to variability in the biological model"*. Vervatis et al., (2021a) and (2021b) performed ensemble simulations, using a 3D high-resolution ocean physics and

biogeochemical coupled model, to investigate unresolved scales and processes, perturbing (1) only ocean physics, (2) only BGC sources and sinks, and (3) both physics and BGC simultaneously, and found that uncertainties in physical forcing and parameterizations have larger impact on chlorophyll spread (and other BGC variables) than uncertainties in ecosystem sources and sinks. Moreover, this had an impact on increment analysis correction and on empirical consistency between model-data misfits, using various datasets (e.g., SST, SLA, total CHL and/or class-based PFTs). I think part of this information would improve the quality of the paper. This is merely a suggestion and I leave it up to the authors to decide if it is relevant to their work.

Minor comments:

1) P1, L26. Please avoid acronyms in the abstract e.g., CCE.

2) P7, L250. Do you mean N_O,i,j instead of N_M,i,j?

Best regards,

V. Vervatis

References:

Vervatis, D. V., P. De Mey-Frémaux, N. Ayoub, J. Karagiorgos, M. Ghantous, M. Kailas, C.-E. Testut and S. Sofianos, 2021: Assessment of a regional physical-biogeochemical stochastic ocean model. Part 1: Ensemble generation, Ocean Modelling, 160, 101781, https://doi.org/10.1016/j.ocemod.2021.101781.

Vervatis, D. V., P. De Mey-Frémaux, N. Ayoub, J. Karagiorgos, S. Ciavatta, R.J.W. Brewin and S. Sofianos, 2021: Assessment of a regional physical-biogeochemical stochastic ocean model. Part 2: Empirical consistency, Ocean Modelling, 160, 101770, https://doi.org/10.1016/j.ocemod.2021.101770.