

Interactive comment on “Biological data assimilation for parameter estimation of a phytoplankton functional type model for the western North Pacific” by Yasuhiro Hoshiba et al.

Anonymous Referee #2

Received and published: 26 June 2017

Parameters estimation in biogeochemical models is a critical challenge and assimilation of data is an important tool to better constrain these parameters. This manuscript is thus dealing with a topic of relevance. It is clearly written and presented, figures are of good quality. I have however important criticisms concerning the methodology used in the manuscript. Mainly, I find that the paper does not provide enough convincing arguments that the way data assimilation is used here clearly improves model performances. The use of data assimilation requires a thorough quantitative assessment of model performances before assimilation and after assimilation and this assessment is insufficient here. Only the seasonal cycle of monthly averaged surface concentration of phytoplankton is presented and broadly compared with observation and this is

not sufficient. I would recommend that the authors strongly improve that point. Then, I would like that the authors comment of the propagation of errors that may happen when 1) interpolating a global chlorophyll product to their grid (is there not a regional product available for the region?), 2) converting global chlorophyll to PFT, 3) converting chlorophyll to nitrogen values. The paper does not convincingly demonstrate that model performances are really improved thanks to the assimilation of such data (conversely looking at e.g. figure 3, we have the feeling that data assimilation degrades model performances see my details comments below). The readers would be better convinced by the gain obtained by the assimilation of such data if the authors presents a thorough error assessment computing error statistics like bias, RMS,... Therefore, I find that the paper cannot be accepted for publication in its present form and would require a substantial revision.

Detailed comments: Line 16-17: The authors say " The approach used a one-dimensional emulator that referenced satellite data". Please clarify what you mean by "emulator" and "reference to satellite data". I guess that you mean that the estimation of parameters is based on a 1D simulation using satellite data for constraining biogeochemical parameters.

Lines 17-18: The comparison with other models that do not assimilate data is only limited to the model used in this paper and cannot be considered as a general feature. Please reformulate.

Line 33-35: "Physiological parameters have often been tuned up empirically and arbitrarily, although ecosystem models have recently added more parameters to increase the number of prognostic and diagnostic variables". It is not clear here why there is an opposition of the two parts of the sentence (use of although). The fact that the number of models parameters increases makes it more and more difficult to tune them automatically. On the other, I agree that ecosystem models are more and more complex with the aim to increase their reliability but the increased number of uncontrolled parameters may make them less realistic.

Line 36 : Please clarify what you mean by “a reasonable estimate of the physiological parameters” – Line 40: set and not set?

Line 42: I find that the use of “This algorithm” may be confusing. I guess that here the authors refer to the model NSI-NEM and so I would use “model” instead of “algorithm”.

Line 45: I guess that the sentence is incomplete. I would add something like ‘based on ‘ before following the . . .

Lines 44-49: This part would be better placed in the section on materials and methods

Line 50 : “and the Parametres-optimized approach” this is vague I would say and a micro genetic algorithm to optimize..

Lines 55-60: I would suggest to add on Figure 1 that shows the region of interest, a schematic representation of the main circulation features (i.e. Kuroshio, Oyashio currents, + transport of iron from the sea of Okhotsk).

Line 62: It is not clear whether this paper is the one that force NSI-NEM by a 3D circulation fields or whether the novelty lies in the fact that the physics was of better quality than before due to data assimilation. Please clarify.

Line 62: what is the “data assimilated physical field”? Model results with data assimilation. I miss in the introduction a clear description of the objectives of the paper. It is mentioned that the region presents a high variability (and so models have limited capacities in presenting the physical and biogeochemical characteristics of the region) but then the authors focused on monthly averaged values. This is confusing. I also recommend a brief description of the sections of the manuscript.

Line 69 : please clarify which physical fields you are using. I guess that you mean OGCM model outputs.

Line 88: What do you mean by “employed by those”?

Line 113: What about lateral inputs/exports?

[Printer-friendly version](#)[Discussion paper](#)

Line 114-115: Usually the set of parameters on which the optimization efforts are concentrated are selected based on sensitivity/identifiability studies. Here this is not really clear how they are selected. (how many parameters are in the model?)

Lines 130-131: Please comment on the assigned weights in the cost function. $0.1 \mu\text{mol/l}$ seems to be quite low.

Lines 149- +Figure 3: The performances of the model at the two stations with data assimilation are not convincing at all. The model without data assimilation has better performances than the model with data assimilation. The authors mention “The seasonal variations of PS in the Parameter-optimized case (dashed lines) for the two stations simulated by the satellite data (solid lines) were more accurate than those in the Default case (dotted lines).” Looking at Figure 3b, this is not what we can see.

Figure 4: I would recommend to compute error metrics like bias, RMS, .. in order to quantify more objectively model data misfit and to assess whether the assimilation of data really helps to improve model performances.

Figure 1: please explain why you have different colors for the boxes?

Interactive comment on Ocean Sci. Discuss., <https://doi.org/10.5194/os-2017-39>, 2017.

Printer-friendly version

Discussion paper

