Biogeosciences Discuss., 6, C1185–C1194, 2009
www.biogeosciences-discuss.net/6/C1185/2009/

Biogeosciences Discussions

Interactive comment on "Skill assessment of the PELAGOS global ocean biogeochemistry model over the period 1980–2000" *by* M. Vichi and S. Masina

M. Vichi

vichi@bo.ingv.it

Received and published: 14 July 2009

1 Changes in the paper structure

We understand the Referee's concerns about the value of this paper for BG, which we believe was not immediately captured due to our incomplete description of the aims in the Introduction and in the Discussion. We have now restructured the paper rewriting the first section and explaining the objectives of the work more clearly. The paper is now divided in two major parts. The first part (Sec. 3) is related to the objective assessment using the various data bases. The second part (Sec. 4) defines the acceptability

thresholds of performance indicators and discuss the major biases and scientific findings of the objective validation. Based on the comments of all the referees, we focused the discussion on three topics: the analysis of major biases (Sec. 4.1); the role of DOC in the estimates of primary production through models and in situ observations (Sec. 4.2); the implications on the metabolic balance of the ocean derived from model results (Sec. 4.3). We removed the discussion on the variability of primary production in the equatorial Pacific (former Sec. 4.3) since, as pointed out by the Referee, it required a more thorough analysis of the driving processes.

2 Answers to general comments

OBGCMs are nowadays included in many climate models and used for the projection of changes under future scenario conditions. Many of these models are being employed for making projections without any specific assessment of their capabilities, besides visual subjective comparisons. Face validity is the first necessary step of the validation process, but without quantification there is no way forward for model improvements. The added value of using objective measures to qualify model skills is in our opinion straightforward in biogeochemical modelling. With the usage of objective metrics it is possible to assess whether the mathematical (and numerical) transposition of a set of conceptual mechanistic relationships is qualified to describe the observable reality, either highlighting sistematic biases or pointing to features that the model is able to robustly simulate. For instance, Carr et al. (2006) demonstrated that most satellite-based primary production models were affected by bias errors and this lead to significant improvements (Friedrichs et al., 2009).

There is a corollary aspect of using objective assessment for global scale models. Scoring model performance is very much useful in the context of model development. These values can be used as benchmarks to check the efficacy of a new component

addition, parameterization changes or newly available data. The heuristic nature of biogeochemical models implies that model formulations are parameterized on specific datasets or derived from general considerations on ecosystem functioning. Testing their genericity against a set of benchmarks is therefore one possible way forward for building more robust formulations.

This is for instance the outcome of the comparison with the ClimPP dataset presented in the former Sec. 4 (now Sec. 2.3). It is known that a considerable fraction of primary production may be lost directly as dissolved organic carbon in nutrient-deplete conditions (Ogawa and Tanoue, 2003). Our results indicate that considering this fraction when comparing with in situ primary production estimates considerably improves the results. This occurs because our model of primary production simulates the different carbon pathways, but it has implications also for other models aiming at the estimation of net ecosystem productions, because by neglecting this fraction they may underestimate the flow of carbon through the food web. As requested by the other reviewers, we have now put these considerations more clearly in the discussion section.

Concerning the limitations of the assessment exercise to chlorophyll data and NPP (though we assessed microbial biomass and production in a wider sense) and not focusing on export production, we have to remark that data on export production cannot be directly obtained. Carbon export has been quantified mostly by means of empirical data-model, with calculations based on algorithms starting with satellite-based primary production and continuing with conversion of primary production to sinking particle flux (e.g. Dunne et al., 2007). Therefore, primary production is the first and foremost parameter to assess before moving on to carbon export adding further uncertainties. We added this consideration partly in the introduction and in the new discussion section. There are few regions where export parameters have been directly estimated through the available measurements, as done for instance at BATS and HOT JGOFS stations (Brix et al., 2006). We compared model results with these estimates and we now moved this comparison from former Sec. 4.2 (where it was indeed mixed with other

results and not sufficiently clear) to the discussion section on the metabolic rate of the ocean (now Sec. 4.3). Metabolic balance of the surface ocean and export production may be seen as synonims, therefore Sec. 4.3 is actually dedicated to the assessment of export production as simulated by the model.

3 Specific comments

We are particularly grateful to the reviewer for the detailed specific comments on the text and we have included all the suggested corrections in the revised version. Here follow the answers to the ones that require more explanations:

- *p3519 l 18 how can 'higher-than-observed' variability be caused by 'too low summer concentrations' (when Chl concentrations should be high).* Rephrased as follows: The northern hemisphere higher-than-observed variability, particularly evident in the North Atlantic, is caused by too fast decrease of surface biomass after the spring bloom, a feature that has not improved from the climatological model results (Vichi et al., 2007b).
- *l 22 how can a spring bloom be driven by the sudden starting of stratification? Isn't it the increasing light availability that's crucial?* It is crucial only if the MLD is sufficiently shallow to use light favourably and thus carbon production exceeds respiration losses.
- *l 21 which resulted greatly underestimated ? what is meant?* This sentence has been removed and the comments on the relationship between MLD and phytoplankton have been further detailed in the Discussion section.
- *p3521 l 12 change Fig 1 to Fig. 5 . This comment is not clear. We mean Fig. 1 since we describe the difference between chl concentration and primary*

production. We rephrased the sentence as:
"There is a good agreement in the spatial distribution of maxima, especially in the location of the frontal maximum in the Antarctic Circumpolar Current (AACC). This good result is obtained in spite of the large positive bias in the annual mean chl value (Fig. 1 and previous section)."

- *p3522 l 6-7 The comparison . . . can be considered an assessment. . . what is meant here?*
  Rewritten to: "The comparison of model results with satellite-derived primary production is a valid assessment only if satellite-based NPP models are good in reproducing in situ observations."

- *p3524 l 15/16 this sentence is incomplete.* Now rewritten to: "NPP is instead much better than chlorophyll and in line with the results of the other PPARR3 models as further shown below. The NPP2 estimate of PP improves all the performance indices (Table 1): for instance, the bias is much reduced with respect to NPP1 and consequently the total $RMSD$."

- *p3525 l 2 'either the highest or lowest than' does not make sense*. Rewritten: "Two additional artificial data points that represent the worst cases have been added. They are obtained by combining the worst scores from all the models and by taking the highest and the lowest standard deviation values. "

- *p3526 l 26 from Fig 8 and 9 there is a lag of at least 2-3 months between MLD and PP – this should be discussed.* Data points are not uniformely distributed in time, therefore a lag correlation cannot be applied in this case since it would be biased by data clustering within each month. It is however possible to test this correlation in model results using the monthly mean values of NPP and MLD. We verified the existence of significant correlations at BATS in the interval 0-2 months (peak at 1 month) as correctly pointed out by the reviewer (Table 1). A

higher correlation is also found at HOT with a one month lag, although the correlation is generally lower than at BATS. The following sentences have been added in the revised text in Sec. 3.4.1 and Sec. 3.4.2, respectively:
"The lag-correlation analysis performed on model results shows a peak at 1 month and remain larger than 0.4 for an interval of about 3 months. This implies that production starts when the mixed layer is still deep and the peak of production is reached after the onset of stratification."
"The MLD evolution is well-reproduced by the model (Fig. 9) , but there is a less clear relationship with NPP as seen for instance at BATS ($r = 0.46$ both in observations and model data; the lag-correlation in model data is higher with 1 month lag, $r = 0.57$)."

- *p3527 l 4-6 this sentence is unclear.* This was releated to the discussion above. Rewritten to: "MLD is visually well predicted by the model although the scatter plot (Fig. 9b) is not as significant as for NPP due to the wintertime bias. The underestimation of NPP during winter is likely due to the underestimation of MLD since the misfits are linearly correlated ($r = 0.56$). The model simulates the NPP inter-annual variability quite well, particularly when linked to distinct physical features. This occurs for instance during the low-production event of winter 1994 when the observed MLD is shallower than other years and the model is able to simulate it correctly. "

- *l 8 'the model is able to partly recovery the export' - a) unclear b) how is export coming into play?* This sentence has been removed in the revised text. Export rates are now better explained in the new Sec 3.3.

- *p3528 l 4 rephrase 'to bring MEF higher than 0' l 5 which has a no bias – a or no?* Rephrased to: "There is a small linear phase correlation (confidence interval 0.11-0.46, p<0.01) which is likely caused by the presence of a weak seasonal signal both in data and model, although the MEF index is still close to 0 confirming

that the model can only capture the mean value." Also removed the "a": BP has no bias at all

- *p3528 l 26/27 by the filter. leading to . . . ??* This sentence was incomplete and has been removed. It is now rewritten in the new Sec. 4.2 that discusses the DOC issue more in detail.

- *p3534 l 19 ESMs 'solve' the carbon cycle – rephrase again here is a reference to carbon sequestration, but this is never discussed in the ms. so the following statements are a bit misleading.*
  This sentence has been completely rephrased, following the new structure of the paper, and the meaning of carbon sequestration in the context of the work has been explained: "The aims were twofold. Firstly, to evaluate the performance of the model under current climate conditions in view of its usage in climate change scenario simulations in the context of Earth System Models (ESM). The focus was thus on the production of organic carbon and its transformation along the microbial food web. Secondly, based on additional comparisons with measured basin-scale carbon exchange rates in the Atlantic, we computed the efficiency of the surface net community production taken as a proxy for the biological pump."

- *p3538 l 4 on the other hand is close to ?* If the index is close to 0, the model is as good a predictor as the data mean. This implies that the model correctly reproduces the mean but that the simulated variability is lower than observed.

- *p3551 Fig 2 is there really an overlap of 20deg? NH and tropical look fairly similar while SH is quite different.* We corrected the latitudinal ranges since they were indeed wrong. There is however no overlapping between the regions. The tropics are similar to the summer signal of each hemisphere; the difference with the Southern Hemisphere is larger because this region extends to 60^{o}S where the chl bias is higher.

- *p3553 Fig 4 is the mean annual NPP (sum) or the annual mean NPP shown? Numbers are similar to Fig 5, where zonal and annual means are Shown so this implies annual means are shown in 4a.* The referee is right: the figure shows the annual mean NPP. The caption has been corrected.

Station	-1	0	1	2	3
BATS	0.13	0.65	0.91	0.83	0.56
HOT	0.14	0.46	0.57	0.37	0.13

Table 1. Lag linear correlations (in months) between simulated NPP and MLD at BATS and HOT.

References

Brix, H., Gruber, N., Karl, D. M., and Bates, N. R.: On the relationships between primary, net community, and export production in subtropical gyres, Deep-Sea Res. Pt. II, 53, 698–717, 2006.

Carr, M. E., Friedrichs, M. A. M., Schmeltz, M., Aita, M. N., Antoine, D., Arrigo, K. R., Asanuma, I., Aumont, O., Barber, R., Behrenfeld, M., Bidigare, R., Buitenhuis, E. T., Campbell, J., Ciotti, A., Dierssen, H., Dowell, M., Dunne, J., Esaias, W., Gentili, B., Gregg, W., Groom, S., Hoepffner, N., Ishizaka, J., Kameda, T., Le Quere, C., Lohrenz, S., Marra, J., Melin, F., Moore, K., Morel, A., Reddy, T. E., Ryan, J., Scardi, M., Smyth, T., Turpie, K., Tilstone, G., Waters, K., and Yamanaka, Y.: A comparison of global estimates of marine primary production from ocean color, Deep-Sea Res. Pt. II, 53, 741–770, 2006.

Dunne, J. P., Sarmiento, J. L., and Gnanadesikan, A.: A synthesis of global particle export from the surface ocean and cycling through the ocean interior and on the seafloor, Glob. Biogeochem. Cy., 21, GB4006, doi:10.1029/2006GB002907, 2007.

Friedrichs, M. A. M., Carr, M.-E., Barber, R. T., Scardi, M., Antoine, D., Armstrong, R. A., Asanuma, I., Behrenfeld, M. J., Buitenhuis, E. T., Chai, F., Christian, J. R., Ciotti, A. M., Doney, S. C., Dowell, M., Dunne, J., Gentili, B., Gregg, W., Hoepffner, N., Ishizaka, J., Kameda, T., Lima, I., Marra, J., Melin, F., Moore, J. K., Morel, A., O'Malley, R. T., O'Reilly, J., Saba, V. S., Schmeltz, M., Smyth, T. J., Tjiputra, J., Waters, K., Westberry, T. K., and Winguth, A.: Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean, J. Mar. Sys., 76, 113–133, doi:10.1016/j.jmarsys.2008.05.010, http://www.sciencedirect.com/science/article/B6VF5-4SMDYY6-8/2/5d74d6d829a3d2b9874a5b1e3ec1fd91, 2009.

Ogawa, H. and Tanoue, E.: Dissolved organic matter in oceanic waters, J. Oceanogr., 59, 129–147, 2003.

Interactive comment on Biogeosciences Discuss., 6, 3511, 2009.