Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-173-RC2, 2016

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "Calibrating a global three-dimensional biogeochemical ocean model (MOPS-1.0)" *by* Iris Kriest et al.

M. Butenschön (Referee)

momm@pml.ac.uk

Received and published: 21 October 2016

1 Referee report for "Calibrating a global three-dimensional biogeochemical ocean model" by Kriest, I. et al., GMD-discussions, 2016

1.1 General Comments

The authors address a highly relevant topic to the modelling of global biogeochemical ocean cycles and marine ecosystem modelling in general, i.e. the optiomisitation of the parameterisation of these models on the base of external constraints in the from of observations of the marine environment. They present a generalised optimisation framework based on state-of-the art components, such as the Transport Matrix Method

Printer-friendly version

Discussion paper

to efficiently resolve the advection-diffusion process of ocean tracers and the Estimation of Distribution Algorithms used in the optimisation score offering a highly efficient and generalised tool to the scientific community to address this topic (and others). The authors are to be complemented in their effort in building this system and are discussing the strengths and weaknesses of their method in substantial detail, but before recommanding the work for publication in Geoscientific Model Development, I'd like to see the following points addressed.

1.1.1 Main points

- It would be nice to see a test of the optimisation against standard test cases (e.g. Lennart-Jones clusters or similar) in terms of convergence and efficiency with respect to other optimisation methods? The test cases given serve well as an illustration of the procedure, but not as a benchmark. (Maybe some more benchmarks are given in the cited literature that can be referred to?)

Page 5, line 21 f.: While the normale distribution is a sensible choice, I wonder if "... is considered to provide the best search diversity..." reflects the authors oppinion (in which case this should be made clearer by rephrasing) or general consensus (in that case: are there any references?). In addition, a lot of biogeochemical parameters will not be valid for negative values, so a truely symmetric probability density function is unlikely. The Gaussian assumption may still be good enough for the "relevant part" of the parameter space, i.e. the area within the bounding constraints, but maybe this point deserves some consideration.

Page 5, line 26 f.: How are the total number of samples and the number of samples to be replaced chosen? How do these choices affect the performance?

Page 6, line 18 f.: Again, what is the choice for the weight factors and how does it affect performance?

Page 7, line 1 f.: What does the c_mu factor mean for the performance? I'd expect it to slow convergence down. In that case, a discussion of trade-offs between using more samples or a higher c_mu would be interesting.

Section 2.5: How is the choice of parameters to be optimised motivated? Especially with respect to surface vs. deep processes and the focus of optimisation on the deep reservoire.

Page 11, lines 1-5: Might be worth loosing a couple of words on what kind of configuration/set-up MOPS-RemHigh is. Do I understand this well that the TWIN experiments evualates TMM+MOPS against nutrients fields from a MOPS-RemHigh?

Page 11, line 21: Until this point I wasn't sure if the simulations actually did run the full 3000 years for each parameter set candidate or if they used a "collective" spin-up. I'm glad the former is the case, but it might be worth making this point clearer before (section 2.2).

Page 12, line 14 f.: Also the global mean of phosphorus in the model is unconstrained, as there's is no constraint on the total amount of biomass in the current optimisation framework as far as I can gather. It is true that there are no global sources or sinks for phoshorus in the model, but that doens't mean that there's no error in the total amount. It just doesn't change during the optimisation (as long as the total phosphorus initial condition isn't included in the optimisation parameters).

Page 15, line 30: I thought that the main aspect of the issue of OBS-WIDE was not so much that it was trapped in a local minimum, but that it ended up in an area of the parameter space that yielded unfeasible results on the base of observational constraints not used in the optimisation (e.g. the grazing rates). So in principle, the minumum found may even be global (even if in this case it is not, looking at the misfit values achieved), but the resulting biogeochemical fields and fluxes are unreasonable. To me that is in an important difference, highlighting the fact that the automated optimisation process is not guaranteed to deliver acceptable results, but may still require expert

judgement as additional validation, as long as the observations used in the optimisation process are not sufficient to fully constrain the ecosystem functions modelled. (In fact the misbehaviour might in principle fall into a part of the modelled food-web that is not sufficiently constrained to demonstrate misbehaviour quantitatively, in particular for more complex models).

Page 16, line 6 f.: "Increasing the population size . . . ." Reiterating the previous comment, it is not guaranteed that an optimal solution that is judged unfeasible on the base of data or knowledge that is not used in the optimisation process, represents only a local minimum and not a global one. Specifically, there is no reason why this minimum should have higher misfit values than any other minimum within the set of other optimal solutions that deliver reasonable results. So there's actually no guarantee that increasing the population size would help.

Figure 1: This figure is hard to understand and needs to be explained better in the caption in order to be useful, e.g. what are the letters? what do the different box shapes (circles vs. squares) represent?

1.1.2 Minor Comments

Page 4, lines 22 ff.: Might be worth explaining exploitation vs exploration for readers less familiar with the subject of optimasation.

Page 5, line 4: Here we use . . .

Page 5, line 13: therefore

Page 5, line 24: ". . . a some what misuse. . . ", please rephrase to "a misuse to some degree" or similar.

Page 5, line 30: towards

Page 8, line 15: What is mu_eff?

Page 8, line 18: Where does this damping parameter appear from?

Page 10, line 4: See the information. . .

Page 10, line 5: . . . , e.g.

Page 11, lines 2-5: dissolved inorganic oxygen

Page 11, lines 8-13: How are the parameter bounds chosen?

Page 14, line 27: What is “it”?

Table 2 heading: I can’t find lambda in the table, so non need to specify it here.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-173, 2016.