Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2017-133-RC1, 2017

Geoscientific Model Development Discussions
Open Access

Interactive comment on "Error assessment of biogeochemical models by lower bound methods" *by* Volkmar Sauerland et al.

Anonymous Referee #1

Received and published: 24 July 2017

General Comments

I reviewed the paper from the perspective of someone with experience in development and application of coupled physical-biogeochemical models, and employing common metrics for evaluating their skill. The manuscript describes an approach to determine a lower bound for the observation-model fisfit using a surrogate non-parametric model, which can then be used for termination of parameter fitting algorithms. The proposed approach is tested against synthetic data and albeit simple, real observation data. As the assets of the manuscript, the objective of developing a technique for reduction of the computational costs of model calibration efforts is certainly worthwile, and the proposed method for meeting this objective, insofar as I understand, seems promising. However, I have mainly 3 problems with the manuscript in its current form. First, the extent to which the work relates to previous work is unclear. Second, the description of the proposed algorithm is at times too technical and the flow of logic is not always crisp. Third, I am not convinced about the applicability of the proposed approach in real world. I recommend therefore a re-review of the paper after major revisions. I would be willing to review the revised paper.

Specific Comments

1) Link with previous work is unclear: in particular, the introduction section is very poorly written. The first few general paragraphs are littered with no-content sentences (eg., P2,L6) followed by multiple citations (in that specific case 11!, I am not going to list all such instances). Such mass-citations do not help at all. Be specific with the arguments and supporting references, and expand when needed. For making general remarks, refer to a recent review paper if possible. I also noticed that many of these needless references are self-citations: at some point I had to count, to find out that 13 references out of a total of 39 are co-authored by at least one of the authors of the manuscript. It is difficult to believe that this is a respectful account of the earlier work. Then, once the text approaches to the subject matter with the second sentence in P2,L20, literally NO reference is cited for the rest of the section. As a concrete example, after finding it difficult to follow the methods section (see below), I had no idea what to read to bring myself one step closer to the subject of the paper.

2) The method section is difficult to follow: I see 3 separate problems. a) The general outline of the proposed algorithm (hinted only at the very end of the manuscript, in the conclusion section) is unclear. Perhaps a schematic illustration would help. b) The logic behind the choice of model properties is not clear. These properties (e.g., monotonicity, number of extremes and steepness of gradients) might be theoretically sufficient to describe any given trajectory, but such an argument needs justification either by a reference or a proof. Or if it just a claim that most trajectories produced commonly by bgc models can be described with these properties, then this needs to be explicitly communicated as such, and what other behavior that will be left out needs

to be discussed c) The mathematical notation and technical jargon is indeed required for precision, but the non-mathematician (like myself) would find verbal expansions helpful. In addition, for explaining the terms like convex optimization problem, quadratic and dynamic programming, and order of computational effort, an information box might be helpful.

3) Evidence for the applicability of the approach in real world is missing: in section 3.1.2, the intuition is confirmed that failing to specify properties suitable to the underlying parametric model leads to a relatively low error ratio, thus, potentially to false lower bounds. But then, in a real world case, given that such error ratios will not be available (as the motivation for the approach is to avoid fitting the parametric model at the first place), how should the suitable set of properties be determined? As a 'real world' example in section 3.2, authors use a climatological seasonal cycle of phytoplankton biomass and state that 'for more complex problems, number of extremes should be iteratively increased until the lower bound does not increase anymore'. Illustration of exactly such an exercise is I believe critical for providing evidence to the applicability of the proposed approach in real world, because such complexity in ecosystems research is the rule, and not the exception. In specific, I would be curious to see how the method is applied to an observation set that contains a long-term trend (e.g., due to eutrophication), inter-annual variability (e.g., due to meteorological extremes), and seasonal cycles. It is stated (P14, L10) that the data at the Bornholm station used for Fig.6 is available for the period 1962-2009, which I guess will display such a mixture of signals.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2017-133, 2017.