

Response to reviewer comments “Parameter sensitivity and identifiability for a biogeochemical model of hypoxia in the northern Gulf of Mexico”, Beck MW, Lehrter JC, Lowe LL, Jarvis BM.

We thank the editor and reviewer for providing thoughtful comments on our manuscript. Responses to these comments are shown in italics. Page and paragraph numbers refer to the revised manuscript.

Editor’s comments:

The authors executed the work on parameter sensitivity and identifiability for a biogeochemical model of hypoxia in the northern Gulf of Mexico. The authors did not provide proper conceptual diagram of the model.

The previous conceptual diagram was for CGEM that included both the biogeochemical components of FishTank and hydrodynamic components. A new figure was created that includes only the FishTank components.

The model is not presented in detail, it is necessary to give more details on original model.

We have added Table 1 that includes a full list and description of parameters that were evaluated for the sensitivity analysis. This table includes short and long descriptions, units, and starting values.

Full details of the model are provided as an appendix in Lehrter et al. 2017. We have not included the content here because it includes nearly twenty pages which we do not wish to duplicate. We refer the reader to the appendix twice in section 2.1, which includes a new sentence in the first paragraph: ‘A full description of the model structure, equations, and parameters is described in appendices A-F in Lehrter et al. (2017).’

Discussion part is unnecessary long, lot of introduction in the discussion. This part needs major revisions and complete restructure.

The discussion was shortened and introductory material was removed. Specifically, the first paragraph of the discussion was removed and the first paragraph of sec. 4.1 was shortened. This reduced the length by over one page.

The sensitivity analysis was adequately implemented but further clarification is needed for how the water quality model was applied in Weeks Bay.

We present a thorough recalibration of the model

In view of the above comments and reports of reviewers, your manuscript has been evaluated and you are informed to resubmit the manuscript after proper and thorough revisions in accordance with the comments of editor/reviewers. Review reports are

appended below. Thank you for your submission in Ecological Modelling.

Reviewer 1:

The authors present the sensitivity analysis of a 0-Dimensional biochemical model implemented in Weeks Bay Alabama. The authors evaluated the sensitivity of the model predictions of DO, ammonium, nitrate, chl-a and irradiance to perturbations in several model parameters. In general, the theoretical basis of the sensitivity analysis is very interesting and mathematically sound and the use of the collinearity index (Eq. 3) to explore linear correlations between parameters seems to be very useful to help identify the most important parameters controlling the model performance (in terms of goodness of fit to observed datasets). The paper can be accepted pending major revisions.

My main concern with the manuscript is that while the application of the sensitivity analysis to the FishTank model seems ok, the application of the FishTank model to Weeks Bay, AL is weak and not convincing. First of all, a correct representation of the advective and dispersive transport is critical to be able to capture the fate and transport of water quality in an estuary. Without a correct representation of transport it is almost impossible to be able to obtain a predictive water quality model. The FishTank model in this case neglects all modes of transport in Weeks Bay, AL and therefore assume the estuary is like a bathtub. This is obviously incorrect and represents a clear limitation of the paper.

Calibration the model to the Weeks Bay data was meant to demonstrate application of the parameter selection heuristics to improve insight into model behavior. We agree with the reviewer that this application is not entirely appropriate given the potential hydrodynamic influences on the observed data. However, we still argue that unit-testing is an efficient means of improving larger models that are computationally intensive. To address the concern of the reviewer and to maintain our emphasis on unit-testing, we have chosen a more appropriate dataset, described below. In addition, we have added this sentence to the the final paragraph of the introduction to emphasize the benefits of our zero-dimensional model.

'In general, the analysis makes a case for unit-testing with lower-dimensional models, particularly when computational limitations preclude the use of conventional optimization algorithms with larger three-dimensional models.'

The new dataset was

For the discussion:

*'However, calibration to most parameter sets showed little improvement in model fit despite relatively large changes in parameter values. This result underscores the challenge of model optimization when considering many parameters. Although we demonstrated model sensitivity to individual parameters, the interactive effects are also important. For example, optimization to zooplankton parameters showed almost no change in model **rmse!** (**rmse!**)*

*despite a relatively large sensitivity of the model to individual parameters. A more comprehensive approach that considers the interaction of zooplankton parameters with others, such as phytoplankton, is warranted given the structural dependencies between parameter categories. Indeed, the reduction in **rmse!** for selection heuristics that weighted the parameter categories equally showed the second largest reduction in model error.'*

For the sensitivity analysis the authors manipulated the observations of DO in the estuary filtering the tidal impacts to be able to implement the FishTank model. However, there is no evidence of the resulting time series of DO or of the other variables analyzed. I am concerned that while removing the DO variations due to tidal cycles, the authors could have also removed the daily variation of DO which can be associated to diurnal cycles of sunlight and temperature. If this is the case, then the DO calibration is based almost on synthetic and unrealistic DO data. The authors must include a section showing the data that was used for the calibration analysis. This section should include a location figure showing Weeks Bay, AL, the stations used to extract the original datasets, plots of the unprocessed and processed time series of DO, ammonium, nitrate, chl-a and irradiance as well as plots showing the processed data against the calibrated model outputs for each variable. The authors must also include a table with the calibrated parameter values. This information is standard in all modeling studies.

Please also describe what are the forcing conditions of the FishTank model. There are two tributary rivers in Weeks Bay and one outlet condition. How is each one simulated in your FishTank model?

Other comments:

Introduction. I suggest the authors to include also references of couple hydrodynamic and water quality studies available in the region. For example, the work of Camacho et al. (2014a) in St. Louis Bay, MS and part of Mississippi Sound is a relevant citation that is missing in this investigation. The work of Dortch et al. (2007) in Mississippi Sound should also be cited.

Citations for Dortch et al. 2007 and Camacho et al. 2014a were added to the first paragraph.

Sensitivity analyses have also been conducted in the past in the region. For example, Camacho et al (2014b) conducted an uncertainty and sensitivity analysis of a hydrodynamic model of Weeks Bay Alabama. I suggest the authors to include these references. The sensitivity estimates provided by Camacho et al. (2014b) seem to be based on the same equation Eq. (1) included in this manuscript. Can the authors briefly indicate (perhaps one or two sentences) if both methods are the same?

Yes, this approach is similarly described by Camacho et al. 2014b. The following text was added to paragraph one in sec. 2.2: 'This approach is described as a First Order Variance Analysis that linearly propagates changes from model parameters to model predictions

(Camacho et al. 2014b).’

lines 24 - 39 page 3. This part of the text is full of vague statements. Please be more explicit. What are the “characteristics expected from the output” and the “information represented by the structural components”. What do you mean with “Given that these characteristics cannot be simultaneously achieved, models are developed in partial dependence of reality and theoretical constructs, completely separate from both, or dependent on one or the other”.

This section was revised as follows:

‘The development of a model represents a tradeoff between achieving predictive accuracy and a realistic representation of environmental processes.... Given that these characteristics cannot be simultaneously achieved, models are developed that balance predictive accuracy with environmental realism, often favoring one at the expense of the other.’

lines 24 - 39 page 3. I think the reference Levins (1966) is extremely old and must be removed.

This citation was referenced in Ganju et al. 2016, from which this information was obtained. We have not removed the citation because we feel that age does not make a publication inadequate. However, we have added the citation to Ganju et al. 2016 as a more current reference.

It is interesting to see that the authors constantly use the term “precision” to denote goodness of fit. In general, precision is used to denote the number of digits of a number. Note that a number can be precise but inaccurate. I suggest the authors to revise the text and replace “precision” with a better term to denote goodness of fit. e.g. level of agreement between observations and simulations.

We have found that precision is commonly used in the literature to describe goodness of fit, although we understand the distinction made by the reviewer. Considering an example where observed data represent a biased sample of an environmental process (e.g., infrequent or uneven sampling of an event that varies over time), model calibration to these data can only possibly achieve precision. Although practitioners would not knowingly use observed data that are biased for model calibration, we are certain that many ecological datasets are incomplete representations of actual processes. We argue then that model calibration in many cases is better described as seeking precision, although in practice it is stated as predictive accuracy given an incomplete knowledge of the observed data relative to the true environmental process. Regardless, we have replaced precision with “goodness of fit” or “fit” as a more generic term.

Line 9 page 10. Selection or Selected?

Changed to: ‘Heuristics for parameter selection also recognized...’

Line 4 page 12. what are the default parameter conditions?. You should include a table with the default parameter values and parameter values under 50% modified conditions.

A new table (Table 1) was added for the full parameter set including names, long description, units, values, and the 50% increase in values for the sensitivity analyses.

Lines 29-30 page 16. Please include the units in the reported RMSEs. mmol O₂?

RMSE units are reported as mmol o₂ m⁻³ and was added to the text.