Title....

Marcus W. Beck[1], John C. Lehrter[1]

[1]*USEPA National Health and Environmental Effects Research Laboratory*
Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561
Phone: 850-934-2480, Fax: 850-934-2401
Emails: beck.marcus@epa.gov, lehrter.john@epa.gov

Version Date: Wed Jun 8 16:40:27 2016 -0500

1

Abstract

Bio-geo-chemical models are useful tools in environmental sciences that can guide management and policy-making. Significant time and resources are spent developing these models in system-specific contexts. The optimization of model parameters to maximize precision, including transferability of these models to different systems, are fundamental concerns in the development and application of these tools. This study describes quantitative limitations of coupled hydrodynamic-ecological modelling by contrasting numeric and ecological certainty with a systematic framework for characterizing parameter uncertainty. We evaluate a simple bio-geo-chemical model that is the 1-dimensional unit of a larger spatio-temporal model of hypoxia on the Louisiana Shelf of Gulf of Mexico as an example. Results from analysis of the 1D model are used to infer larger trends in dissolved oxygen dynamics over time, having implications for understanding factors that contribute to environmental conditions that are detrimental to aquatic resources. In particular, we focus on issues of parameter identifiability using global sensitivity analyses to provide quantitative descriptions of numerical constraints on model precision. We argue that quantitative and ecological certainty in model calibration are often at odds and the practitioner must explicitly choose model components to optimize given tradeoffs between the two. We further conclude that numerically optimal parameter sets are often small subsets of the complete parameter set for a model because of redundancies in the unique effects of paramater perturbations on model output. As a result, we demonstrate that use of a model for inference into ecological mechanisms of observed or predicted changes in environmental condition can be potentially misguided in the absence of quantitative descriptions of identifiability. Although these concerns have been expressed in the literature, they are rarely explicitly addressed or included in model evaluations. In addition to immediate implications for regional models, we provide a framework for describing the effects of parameter uncertainty and identifiability that can be applied to similar models to better inform environmental management.

1 Introduction

1. Simulation/biogeochemical/process-based models overview, contrast with statistical models

2. What models seek to provide - generality, precision, realism Levins (1966), there is a tradeoff so models are 1) developed in partial independence and dependence on the world

2

and theory, 2) function autonomously from both, or 3) represent both at the same time, from Morrison and Morgan (1999), cited in Ganju et al. (2016). This is similar to the bias-variance tradeoff for statistical models, e.g., overparameterization of a model makes it very biased as it fits the data (the world) exactly

3. How is model performance/uncertainty evaluated regarding what they should provide - structural, observational, parameter Beck (1987)?

4. Parameter uncertainy as low-hanging fruit - can do post-hoc and from inner to outer level of complexity

5. Challenges related to uncertainty - similar to degrees of freedom, identifiability definition from Brun et al. (2001) and need to evaluate identifiability Fasham et al. (2006)

6. objectives of study, definition of 'optimum'

   (a) What parameters have the greatest influence on dissolved oxygen as identified by local and global sensitivity analyses?

   (b) What is the 'parameter space' (from sensitivity analysis) and 'identifiability' of the model parameters?

   (c) Can we get a better estimate of identifiability other than 10-15 according to Brun et al. (2001), Soetaert and Petzoldt (2010)?

   (d) Can we reconcile differences in numeric uncertainty and ecological uncertainty, e.g., given a finite parameter space, what is an appropriate way to choose parameters given question of interest?

2   *Methods*

2.1   Model description

GGEM, FishTank, parameters and parameter categories, how they compare to other models

2.2 Local sensitivity analysis

Methods from Soetaert and Petzoldt (2010)

Sensitivity of the model to parameter changes is evaluated using the difference in model results before and after perturbing each parameter value. Each parameter is perturbed by the same value as a percentage of the whole. The default value (as 1 + 1e-8 proportion, default value for 'tiny' argument in 'sensFun'). A sensitivity value S is estimated for each time step i given a set value for parameter j as:

$$S_{ij} = \frac{\partial y_i}{\partial \Theta_j} \cdot \frac{w_{\Theta_j}}{w_{y_i}} \tag{1}$$

where the estimate depends on the change in the predicted value for response variable y divided by the change in the parameter Θ_j multiplied by the quotient of scaling factors w for each. The estimates are unitless such that their relative magnitudes provide a comparison for model sensitivity to parameter changes that may vary in scale. The FME package summarizes sensitivity as $L1$ and $L2$ across the time series:

$$L1 = \sum |S_{ij}|/n \tag{2}$$

$$L2 = \sqrt{\sum \left(S_{ij}^2 \right) |/n} \tag{3}$$

The mean, minimum, and maximum S_{ij} values for each parameter are also reported. In general, positive mean sensitivity estimates generally indicate that a parameter has a positive effect on the model results for a given increase in the parameter. However, the effect can change over time so a plot of the 'sensitivity function' should be viewed which shows the difference from the values of the response variable before and after changing a parameter value. Note that the perturbation factor in 'sensFun' depends on the default value of each parameter such perturbations are not consistent for parameter values less than or greater than the perturbation factor. This can produce results that are inconsistent for different levels of perturbation. A custom function was keeps the perturbation factors constant regardless of the magnitude of the default values. Sensitiviy for each parameter using the custom function are estimated using the above

equations from the FME package. All parameters that are considered 'sensitive' should be further
evaluated by plotting the predictions before and after perturbation and across the total range of the
parameter.

Each parameter was perturbed by 50% using 'sensfun' to identify model sensitivity.
Parameters that produced different model results from the default values are shown below, by
category.

The parameters that have the greatest effect on the model by category (optics, organics,
phytoplankton, zooplankton) are as follows.

Plotting the raw values from the sensitivity analysis provides a visual assessment of
changes.

2.3   Global sensitivity and identifiability

Identifiability describes the ability to estimate a parameter in relation to variation among
the remaining parameters. A parameter is identifiable if all parameters within the set can be
uniquely estimated based on the observed data. Parameters that are unidentifiable typically
produce similar model outputs for a given relative perturbation, i.e., the effect of altering one
parameter can be undone by altering one or more other parameters. Model calibration will not
converge for parameters sets that are unidentifiable. Identifiability is estimated from the minimum
eigenvector of the cross-product of a model's sensitivity matrix:

$$\gamma = \frac{1}{\sqrt{\min\left(\text{EV}[\hat{S}^{\intercal}\hat{S}]\right)}} \tag{4}$$

where γ ranges from one to infinity for perfectly identifiable (orthogonal) or unidentifiable
(perfectly collinear) results for a set of parameters in the sensitivity matrix S. Identifiability can
be estimated for any combination of parameters for a model. Values less than 10-15 indicate
parameters are generally identifiable. The FME package provides the 'collin' function to estimate
identifiability of output from the 'sensFun'. For reasons described above, a custom function was
created to determine identifiability from model residuals, where the residauls were based on a
relative perturbation of the model parameters. A comparison of results of identifiability using the
custom functions and those from the FME package were generally in agreement. Standard output
from the identifiability function is shown below, where identifiability is displayed as a matrix with

5

increasing combinations of subsets of the parameters.

The identifiability functions evaluate the ability to identify all subsets from pairwise to all parameters in the subset. As an example, the identifiability of pairwise combinations for each of the categories are shown below.

Pairwise identifiability of the top two most sensitive parameters in each category is shown, first as a table showing all unique combinations from two to all parameters and second as a figure showing pairwise combinations.

2.4 Identifying subsets and testing calibration

Model calibration must consider the competing objectives of parameter inclusion and identifability. Estimating parameters for novel datasets increases model complexity and the ability to identify parameters decreases with the inclusion of parameters to estimate. The above analyses indicated that estimating all of the parameters was impossible because of redundancies in the the model output. Calibration to novel datasets should balance the competing objectives of identifiability while including parameters for which model output is most sensitive. The plots below show different scenarios of parameter inclusion for model calibration. Reading each plot from left to right can be interpreted as including additional parameters, where each parameter is ranked by relative sensitivity. The inset in each plot shows the identifiability of including parameters, up to a maximum where additional inclusion exceeds the identifability threshold of fifteen. The scenarios for including parameters all begin with the parameters that have the greatest effect on the model output. The inclusion of additional parameters depends on the scenario. The first scenario selected parameters by decreasing sensitivity within each category (i.e., four separate models calibrated for optics, organics, phytoplankton, or zooplankton), the second scenario selected parameters by sensitivity regardless of category, and the third scenario selects parameters by sensitivity with equal representation between categories. Boxes in each plot represent a unique parameter and the numbers in each box represent the relative rank of a parameter's sensitivity to the model output within each category.

The following plots show identifiability with the addition of parameters.

3 Results

4 Discussion

Questions specific to GOM - what initial conditions are important? How many phytoplankton groups do we need (e.g., related to structural uncertainty)?

How does the assimilation of additional parameters (e.g., other state variables) during calibration influence the conclusions?

How does uncertainty translate to what a model should provide (generality v precision)? The first step - find out what can be optimized but then do not overfit....

Need to talk about Flynn (2005)

References

Beck MB. 1987. Water quality modeling: A review of the analysis of uncertainty. Water Resources Research, 23(8):1393–1442.

Brun R, Reichert P, Künsch HR. 2001. Practical identifiability analysis of large environmental simulation models. Water Resources Research, 37(4):1015–1030.

Fasham MJR, Flynn KJ, Pondaven P, Anderson TR, Boyd PW. 2006. Development of a robust marine ecosystem model to predict the role of iron in biogeochemical cycles: A comparison of results for iron-replete and iron-limited areas, and the SOIREE iron-enrichment experiment. Deep-Sea Research I, 53:333–366.

Flynn KJ. 2005. Castles built on sand: dysfunctionality in plankton models and the inadequacy of dialogue between biologists and modellers. Journal of Plankton Research, 27(12):1205–1210.

Ganju NK, Brush MJ, Rashleigh B, Aretxabaleta AL, del Barrio P, Grear JS, Harris LA, Lake SJ, McCardell G, O'Donnell J, Ralston DK, Signell RP, Testa JM, Vaudrey JMP. 2016. Progress and challenges in coupled hydrodynamic-ecological estuarine modeling. Estuaries and Coasts, 39(2):311–332.

Levins R. 1966. The strategy of model building in population biology. American Scientist, 54(4):421–431.

Morrison M, Morgan MS. 1999. Models as mediating agents. In: Morgan MS, Morrison M, editors, Models as Mediators, page 401. Cambridge University Press, Cambridge.

Soetaert K, Petzoldt T. 2010. Inverse modelling, sensitivity, and Monte Carlo analysis in R using package FME. Journal of Statistical Software, 33(3):1–28.
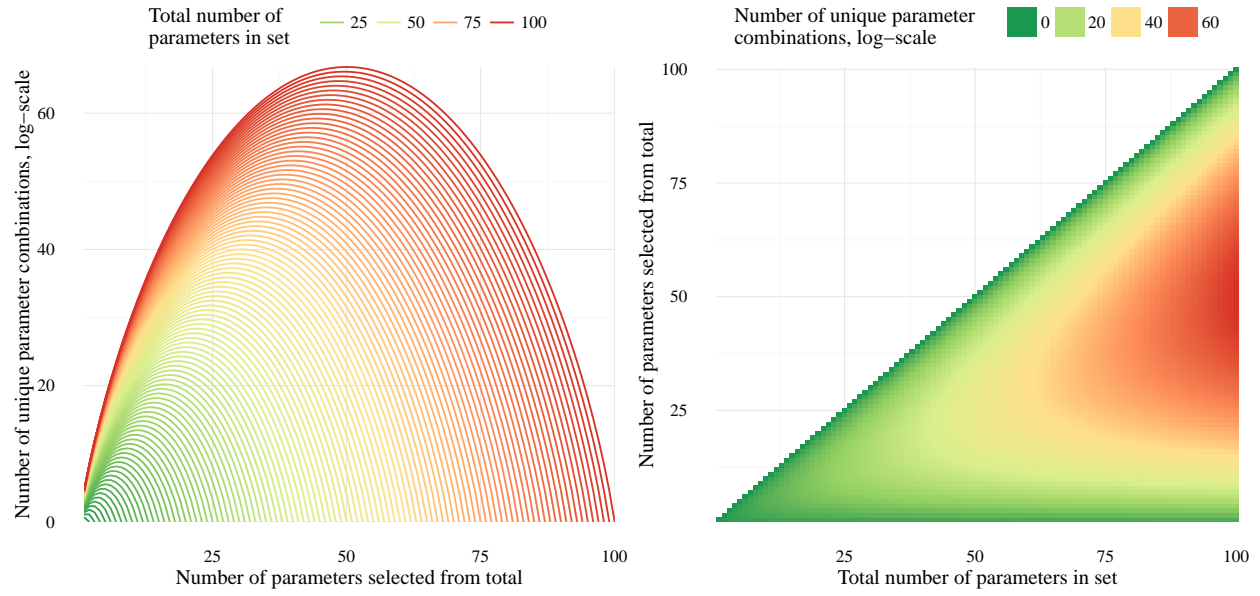
Fig. 1: Number of unique combinations that can occur by selecting parameters for optimization from different parameter sets. The number of combinations are shown for increasing numbers of selected parameters from the total in the set, where 100 parameter sets are shown each with one through 100 total parameters. For example, there are six unique parameter combinations that result from choosing two parameters from a set. Note that the number of unique combinations is shown as the natural-log. The same information is viewed differently in each plot.