Ecological Modelling 363 (2017) 17–30

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Parameter sensitivity and identifiability for a biogeochemical model of hypoxia in the northern Gulf of Mexico☆

Marcus W. Beck[a,*], John C. Lehrter[b], Lisa L. Lowe[c], Brandon M. Jarvis[a]

[a] USEPA National Health and Environmental Effects Research Laboratory, Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL 32561, United States
[b] University of South Alabama, Dauphin Island Sea Lab, Dauphin Island, AL 36528, United States
[c] CSRA Inc., supporting the USEPA, Research Triangle Park, NC 27709, United States

ARTICLE INFO

Article history:
Received 10 April 2017
Received in revised form 18 August 2017
Accepted 19 August 2017
Available online 5 September 2017

Keywords:
Coastal General Ecosystem Model (CGEM)
Gulf of Mexico (GOM)
Hypoxia
Identifiability
Sensitivity

ABSTRACT

Local sensitivity analyses and identifiable parameter subsets were used to describe numerical constraints of a hypoxia model for bottom waters of the northern Gulf of Mexico. The sensitivity of state variables differed considerably with parameter changes, although most variables were responsive to changes in parameters that influenced planktonic growth rates and less sensitive to physical or chemical parameters. Variation in sensitivity had a direct correspondence with identifiability, such that only small subsets of the complete parameter set had unique effects on the model output. Selecting parameters by decreasing sensitivity demonstrated that only eight of 51 total parameters had a sufficiently unique effect on model output for accurate calibration. As a result, parameter selection heuristics were used to identify parameters for model calibration that depended on combined effects on output, relative sensitivity of each parameter, and ecological categories for the biogeochemical equations. The calibrated zero-dimensional (0-D) unit of the hypoxia model had improved fit to the observed data if sensitive phytoplankton parameters were included in an identifiable subset. Extension of results to a three-dimensional grid of the Gulf of Mexico showed that sensitive parameters for the (0-D) model translated to non-trivial changes in the areal estimates of hypoxia.

Published by Elsevier B.V.

1. Introduction

Hypoxia formation in bottom waters of coastal oceans occurs primarily from excess nutrient inputs from land-based sources (Justíc et al., 1987; Diaz and Rosenberg, 1995; Howarth et al., 1996). These events are detrimental to aquatic organisms and have significant negative effects on economic resources derived from coastal ecosystems (Lipton and Hicks, 2003; Diaz and Rosenberg, 2011). An understanding of the biological, physical, and chemical processes that influence hypoxic areas is a critical concern for mitigating and preventing these negative impacts. Numerical ecosystem models are important tools that synthesize knowledge of ecosystem processes that contribute to hypoxia formation and for predicting the effects of proposed management activities or future scenarios (Scavia et al., 2004; Hagy and Murrell, 2007; Camacho et al., 2014; Pauer et al., 2016). Unlike statistical models with more generic structures, simulation and process-based models include explicit descriptions of relevant processes that are constrained by empirical or observational data relevant to the system of interest (e.g., Omlin et al., 2001b; Eldridge and Roelke, 2010). These models are often coupled with hydrodynamic grids to provide spatially-explicit representations of patterns in three dimensions (Warner et al., 2005; Dortch et al., 2007; Zhao et al., 2010; Ganju et al., 2016). Combined hydrodynamic and bio-geo-chemical models have been developed specifically to describe hypoxic conditions on the Louisiana continental shelf (LCS) in the northern Gulf of Mexico (GOM) (Fennel et al., 2013; Obenour et al., 2015; Pauer et al., 2016; Lehrter et al., 2017). This area drains a significant portion of the continental United States through the Mississippi-Atchafalaya River Basin (MARB) and is the second largest hypoxic area in the world (Rabalais et al., 2002). Understanding processes that contribute to the frequency and duration of hypoxic events remains a critical research goal for the region.

The development of a model represents a tradeoff between achieving predictive accuracy and a realistic representation of environmental processes. An ideal model is sufficiently generalizable across systems, provides results that are accurate given the inputs, and includes components that are valid descriptions of actual processes (Levins, 1966; Ganju et al., 2016). Given that

☆ *Acronyms:* chlorophyll a (chl-a); Coastal General Ecosystem Model (CGEM); dissolved oxygen (O_2); dissolved organic matter (DOM); Gulf of Mexico (GOM); Louisiana continental shelf (LCS); Mississippi-Atchafalaya River Basin (MARB); particulate organic matter (POM); root mean squared error (RMSE); zero-dimensional (0-D).

* Corresponding author.
E-mail addresses: beck.marcus@epa.gov (M.W. Beck), jlehrter@disl.org (J.C. Lehrter), lowe.lisa@epa.gov (L.L. Lowe), jarvis.brandon@epa.gov (B.M. Jarvis).

http://dx.doi.org/10.1016/j.ecolmodel.2017.08.020
0304-3800/Published by Elsevier B.V.

these characteristics cannot be simultaneously achieved, models are developed to balance predictive accuracy with environmental realism or often favor one at the expense of the other (Morrison and Morgan, 1999; Ganju et al., 2016). These challenges are analagous to the well-known bias-variance tradeoff in statistical models that balances the competing objectives of over- and under-fitting to an observed dataset. Process-based models are more commonly imbalanced between reality and theory, such that most are over-parameterized in an attempt to completely describe reality (Denman, 2003; Nossent and Bauwens, 2012; Petrucci and Bonhomme, 2014). Quantitative limitations of over-parameterization are analagous to degrees of freedom in standard statistical models as free parameters cannot be numerically estimated when constrained to an observed dataset (Kirchner, 2006). More importantly, over-parameterization can limit use across systems outside of the data domain and impose uncertainty in model predictions as realistic values for every variable may not be known or inaccurately applied from existing studies (Durand et al., 2002; Refsgaard et al., 2007; Wade et al., 2008).

Model fit can be evaluated relative to the effects of initial conditions or the observed data used for calibration, changes in parameter values, or variation in the structural components (i.e., observational, parameter, or structural uncertainty) (Beck, 1987). Evaluating effects of parameter changes is by far the most common and simplest approach. Although sensitivity analyses should be integrated with model development, parameters are often evaluated post-hoc as a form of 'damage control' for further calibration. This approach is sometimes called inverse modeling where results from sensitivity analyses are used to guide calibration or fit of the developed model to observations (Soetaert and Petzoldt, 2010, or confronting models with data, *sensu* Hilborn and Mangel, 1997). Parameter sensitivity analysis combined with inverse modeling necessarily involves questions of parameter 'identifiability'. Redundancies in parameter effects lead to unidentifiable models where calibration is empirically impossible (i.e., standard algorithms will not converge) or parameter values may be non-unique leading to the right answer for the wrong reason (Kirchner, 2006). Unidentifiable parameter sets have effects on model output that can be undone or compensated for by alteration of other parameters. Identifiability issues are not foreign to hypoxia or eutrophication models (Omlin et al., 2001a; Estrada and Diaz, 2010; Mateus and Franz, 2015), although there is a clear need for greater integration of these concepts in practice (Fasham et al., 2006).

This study describes a sensitivity and identifiability analysis of a (0-D) unit of a larger spatial-temporal model of hypoxia dynamics on the LCS. 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 objectives were to provide a statistical approach that demonstrates numerical limitations of parameter sets for model calibration and provide a framework for selecting parameters within the identifiability constraints. The specific goals were to (1) identify the parameters that have the greatest influence on state variables using local sensitivity analysis, (2) quantify the identifiability of subsets of the total parameter space based on sensitivity, and (3) provide a set of heuristics for choosing parameters based on sensitivity, identifiability, and parameter categories. The (0-D) model was calibrated with selected parameter subsets to demonstrate use of the selection heuristics for optimization. Sensitive parameters for the (0-D) model were also varied in the larger 3-dimensional model to demonstrate scalability of the results. In addition to dissolved oxygen (O_2), other state variables that were evaluated included ammonium, chlorophyll *a* (chl-*a*), irradiance, nitrate, POM, DOM, and phosphorus. In general, we provide empirical results to support the assumption that models are generally over-parameterized and only a finite and smaller subset of the complete parameter set can be optimized.

2. Materials and methods

2.1. Model description

Hypoxic events, defined as <2 mg L^{-1} of O_2 (<64 mmol m^{-3}), occur seasonally in bottom waters in the northern GOM. The hypoxic area averages 15,540 km^2 annually (1993–2015) with minimum concentrations observed from late spring to early fall. Seasonal variation is strongly related to carbon and nutrient export from the MARB (Lohrenz et al., 2008; Bianchi et al., 2010), whereas hydrologic variation, currents, and wind patterns can affect vertical salinity gradients that contribute to hypoxia formation (Wiseman et al., 1997; Paerl et al., 1998; Obenour et al., 2015). The Coastal General Ecosystem Model (CGEM) was developed to describe hypoxia dynamics on the LCS and includes elements from the Navy Coastal Ocean Model (Martin, 2000) for hydrodynamics and a biogeochemical model with multiple plankton groups, water-column metabolism, and sediment diagenesis (Eldridge and Roelke, 2010). The hydrodynamic component of CGEM provides a spatially-explicit description of hypoxia using an orthogonal grid with an approximate horizontal resolution of 1.9 km^2 and twenty equally-spaced vertical sigma layers on the shelf. The biogeochemical component includes equations for 36 state variables including six phytoplankton groups (with nitrogen and phosphorus quotas for each), two zooplankton groups, nitrate, ammonium, phosphate, dissolved inorganic carbon, oxygen, silica, and multiple variables for dissolved and particulate organic matter from different sources.

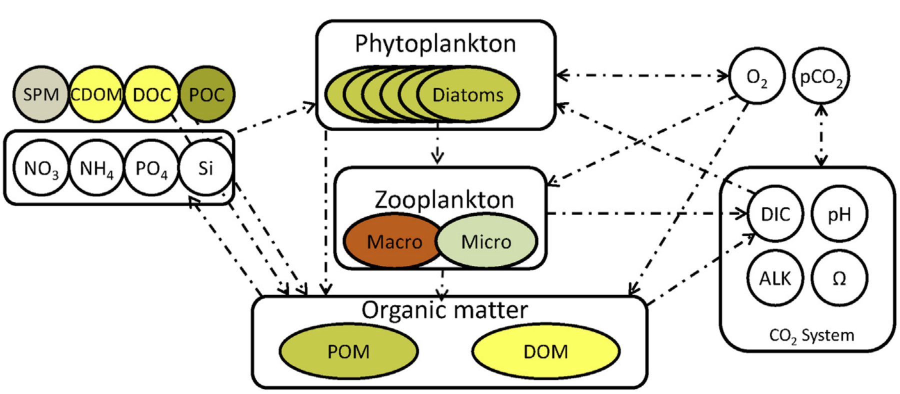

Fig. 1. Conceptual representation of biogeochemical components included in FishTank (complete equations in Eldridge and Roelke, 2010, appendices A–F in Lehrter et al., 2017). The Coastal General Ecosystem Model couples FishTank with a hydrodynamic model that includes advection, mixing, and dispersion.