Applied Ocean Research 47 (2014) 138–153

Contents lists available at ScienceDirect

Applied Ocean Research

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

Uncertainty analysis of estuarine hydrodynamic models: an evaluation of input data uncertainty in the weeks bay estuary, alabama

René A. Camacho[a,b,*], James L. Martin[b], Jairo Diaz-Ramirez[c], William McAnally[b], Hugo Rodriguez[d], Peter Suscy[e], Song Zhang[f]

[a] Tetra Tech, Inc., 2110 Powers Ferry Road SE, Suite 202, Atlanta, GA, USA
[b] Department of Civil and Environmental Engineering, Mississippi State University, USA
[c] River Research Center, Alcorn State University, USA
[d] Tetra Tech, Inc., USA
[e] St. Johns River Water Management District, FL, USA
[f] Department of Computer Science and Engineering, Mississippi State University, USA

ARTICLE INFO

Article history:
Received 15 September 2013
Received in revised form 7 February 2014
Accepted 21 April 2014
Available online 17 May 2014

Keywords:
Estuarine hydrodynamics
Uncertainty analysis
FOVA
Monte Carlo Uncertainty
Skill assessment

ABSTRACT

Uncertainty analyses are necessary to identify, evaluate, and report the main sources of errors in modeling studies and their impacts on the model predictions. Although uncertainty analysis has been a subject of increased interest by the water resources community during the last couple of decades, in estuarine hydrodynamic modeling this is an emerging topic that requires more research. Some of the most relevant problems remaining in the practice include the identification of the principal sources of errors affecting the model predictions, and the identification of effective and computationally feasible methodologies for their quantification. This investigation evaluates the impacts of input data errors on the predictions of a 3D hydrodynamic model of the Weeks Bay estuary, Alabama. The uncertainty analysis is performed using the First Order Variance Analysis (FOVA) and the results compared to a standard Monte Carlo Uncertainty Analysis (MCUA). A procedure to implement a skill assessment as a fundamental component of the FOVA method is presented. The uncertainty analyses are performed temporally as well as spatially distributed over the model domain. The results indicate that the uncertainty in a prognostic variable is not homogeneously distributed over the computational domain, and that there are areas prone to a higher or lower uncertainty. The identification of these areas is relevant for the design of data collection plans intended to improve the confidence in the model results. The comparison of the methods indicates that both are effective to provide uncertainty estimates, although FOVA tends to overestimate the predictions obtained by MCUA. In general this overestimation can be considered as a conservative estimation of the uncertainty given the existence of other sources of errors more complex to evaluate.

1. Introduction

Uncertainty analysis is defined as the procedures and strategies implemented to identify, evaluate and report the main sources of errors in a modeling application and their impacts on the model predictions. Model uncertainty arises as a consequence of errors in the model structure (e.g. oversimplification of processes or errors in the set of equations that define a model), errors in the input data due to random or systematic errors during the collection of data, and parametric errors caused by the difficulty of identifying physically representative parameter values valid at the temporal and spatial scales of the model [27,41].

During the last years uncertainty analysis has received special attention (particularly in hydrological, ecological and climate modeling) and several strategies have been developed and implemented to investigate the impacts of different types of error on the predictions of numerical models [14,32,35,39]. The existing strategies are generally classified into analytical and approximate methods [40], although only these later have applicability for most models used in the practice. The approximate methods include moment based methods such as the First Order Variance Analysis (FOVA) and the Advanced First Order Variance Analysis (AFOVA) [3,28,44], Probabilistic Point Estimate Methods (PPEMs) such as the Rosenblueth and Harr methods [19,40], and Monte Carlo based methods such as Bayesian analysis (BA), Markov Chain Monte Carlo (MC^2) and the Generalized Likelihood Uncertainty Estimation method, GLUE (e.g. [2,11,37]).

Despite the importance of uncertainty in the modeling practice and the availability of methods for its quantification, uncertainty analysis is an uncommon practice in hydrodynamic investigations. This is partially explained because even in small scale hydrodynamic models, the computational burden (some estuarine models require several hours or a day to perform a simulation) can make unpractical the implementation of some of the most widely used strategies for uncertainty analysis (e.g. Monte Carlo simulations, Bayesian Monte Carlo analysis, GLUE). As a result, in estuarine modeling uncertainty analysis remains as an emerging topic where some of the most important problems include (a) the identification of the most relevant sources

* Corresponding author. Tel.: +1 770 850 949.
E-mail address: rene.camachorincon@tetratech.com (R.A. Camacho).

0141-1187/$ – see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apor.2014.04.004

of uncertainty, and (b) the evaluation of suitable strategies for the quantification of their impacts on the model predictions. Recent investigations have attempted to fill these gaps by showing the important effects of model structure, input data, and parametric uncertainty in several estuarine models and by introducing FOVA as an effective and parsimonious strategy to perform uncertainty analyses [3,7,23,24,36]. The simplicity of implementation makes FOVA particularly attractive for complex hydrodynamic models, but more research is necessary to determine if the uncertainty estimates from FOVA are consistent with those derived from other reference methods such as Monte Carlo simulations.

The objective of this investigation is to address this later issue by evaluating the impacts of input data uncertainty on the predictions of the hydrodynamic model of the Weeks Bay estuary, Alabama (USA). We implement FOVA as an alternative strategy to traditional Monte Carlo Uncertainty Analyses (MCUA), and investigate the benefits and limitations of both methods for practical purposes. We also investigate in detail, based on temporal and spatial analyses, how the errors in the specification of the water surface elevations at the open boundary, the model bathymetry, and the freshwater flows, impact the capacity of the model to reproduce observations of water surface elevations (WSE), current velocities, and salinities within the estuary. Finally, we illustrate how a skill assessment can be used as a procedural component of the FOVA method.

The paper is organized as follows: Section 2 describes the FOVA and MCUA methods, the study area, the hydrodynamic model development and the criteria for the model evaluation. Section 3 discusses the sources and magnitude of the errors in our input datasets. Section 4 presents a sensitivity analysis based on FOVA estimates. Section 5 presents the results of the uncertainty analyses including the comparison between methods. Section 6 presents a discussion of the results, and finally, Section 7 presents the conclusions of the investigation.

2. Methods

2.1. Uncertainty analysis

2.1.1. First Order Variance Analysis (FOVA)

The First Order Variance Analysis (FOVA) provides estimates of model uncertainty by linearly propagating the errors from input data or model parameters to the model predictions. The method is relevant because it also provides sensitivity estimates and its implementation is computationally straightforward. Only few hydrodynamic modeling studies have implemented FOVA for the evaluation of input data or parametric uncertainty (e.g. [3,36]), and although further research is necessary to evaluate the limitations and benefits of the method compared to alternative strategies, the existing studies show that the method is effective in: (a) identifying the most relevant input variables or parameters contributing to the output uncertainty, (b) evaluating the degree of sensitivity of the model's output to any variable under analysis, and (c) quantifying uncertainty on model predictions.

The fundamental idea behind FOVA is to construct a Taylor series expansion truncated at the first order term of the function $F(\mathbf{X})$ that predicts the evolution of a given output variable $\mathbf{Y}=F(\mathbf{X})$ [3,29,44]. In the context of this investigation F can be interpreted as the hydrodynamic model, $\mathbf{X}$ a vector containing the input variables $x_1, x_2, \ldots, x_n$ (e.g. wind speed, model bathymetry, and freshwater flows) and $\mathbf{Y}$ the vector describing a set of predicted variables $y_1, y_2, \ldots, y_n$ (e.g. *WSE* and flow velocities). The expansion of the function $F(\mathbf{X})$ is performed around the model predictions $f=F(\mathbf{X}_0)$ resulting from a set of unperturbed or mean values of the input variables $\mathbf{X}_0=(x_{1o}, x_{2o}, \ldots, x_{no})$. The resulting expansion is given by,

$$F(X)=f(x_{1o}, x_{2o}, \ldots, x_{po})+\sum_{i=1}^{p}\left.\frac{\partial F}{\partial x_i}\right|_{x_i=x_{io}}(x_i-x_{io}) \quad (1)$$

where p is the number of input variables or parameters under study; x_{io} is the value of the ith input variable at the expansion or unperturbed point o; and $\partial F/\partial x_i$ is the local change of the model predictions $\mathbf{Y}$ due to changes in the input variable x_i. The expected value and variance of the model predictions $\mathbf{Y}$ are estimated based on Eq. (1) as [3],

$$E[Y]=E[F(X)]=f(x_{1e}, x_{2e}, \ldots, x_{pe}) \quad (2)$$

$$Var(Y)=\sigma_Y^2=\sum_{i=1}^{p}\left[\left.\frac{\partial F}{\partial x_i}\right|_{x_i=xo}\right]^2 E(x_i-x_{io})^2 + 2\sum_{i=1}^{p}\sum_{j=1, j\neq i}^{p}\left[\left.\frac{\partial F}{\partial x_i}\right|_{x_i=x_{io}}\left.\frac{\partial F}{\partial x_j}\right|_{x_j=x_{jo}}\right] E[(x_i-x_{io})(x_j-x_{jo})] \quad (3)$$

If the input variables $(x_{1o}, x_{2o}, \ldots, x_{no})$ are statistically independent, then Eq. (3) is simplified to:

$$Var(Y)=\sigma_Y^2=\sum_{i=1}^{p}\left[\left.\frac{\partial F}{\partial x_i}\right|_{x_i=x_{io}}\sigma_i\right]^2 \quad (4)$$

where σ_Y^2 is the variance associated with the model predictions (i.e. output uncertainty), and σ_i^2 is the variance of the ith input variable (i.e. input uncertainty). The term $\partial F/\partial x_i|_{x_i=x_{io}}$ quantifies the change in the output variable of interest as a result of a perturbation of the input variable x_i from the unperturbed point x_o. The derivative term ∂F can be evaluated numerically using a simple difference scheme. For example, using forward differencing $\partial F=F(x_{io}+\Delta x_i)-F(x_{io})$, where $F(x_{io}+\Delta x_i)$ is the model prediction obtained after perturbing the input variable x_i in a magnitude equivalent to Δx_i from the unperturbed or mean value x_{io}. Note that Eq. (4) can be expressed alternatively in terms of coefficients of variation if the standard deviations in this equation are normalized with the mean or unperturbed values of their corresponding variables (i.e. $CV_Y \equiv \sigma_Y/Y_o$, $CV_{x_i} \equiv \sigma_i/x_{io}$). In this case Eq. (4) would be expressed as

$$CV_Y^2=\sum_{i=1}^{p}[(DSC_i^{F_{x_i=x_o}}) * CV_{x_i}]^2 \quad (5)$$

where DSC_i is given by [3]

$$DSC_i^{F_{x_i=x_o}} \equiv \left[\frac{\partial F/\partial x_i}{F(x_i)/x_i}\right]_{x_i=x_o} \quad (6)$$

and is defined as the dimensionless sensitivity coefficient, which quantifies the relative importance of changes in the input variable x_i on the model predictions.

The complete procedure for implementing FOVA can be summarized in five steps: (1) define the input and output variables of interest, (2) define the mean or unperturbed values of the input and output variables, (3) define the magnitude of the perturbation for each input variable and compute the DSC by using Eq. (6). (4) Estimate the coefficient of variation CV_{x_i} (or alternatively the standard deviation σ_i) for each input variable (i.e. estimate the degree of uncertainty in each input variable), (5) propagate the uncertainty in the input variables to the output variables using Eqs. (4) or (5) (i.e. compute the CV (Eq. (5)) or standard deviation σ_i (Eq. (4)) of the output variable of interest). Step four can also be implemented in parallel or interchangeably with step two or three.

In this study, the input variables of interest were the water surface elevations in the open boundary (η_{ob}), the model bathymetry (M_{btm}), and the freshwater inflows (Q_f), and the output variables of interest were the model predictions of water surface elevation (WSE), current speeds (U_s), and water salinity (W_{sal}) (Step one). For step two we identified the input uncertainty based on the instruments and procedures used to collect the input data. For step three we used a skill assessment in order to set up the mean values of the output variables in the "calibrated" values (highest agreement achieved between model predictions and observations). Note also that the skill assessment is useful to evaluate the quality of the input