Journal of Hydrology 565 (2018) 623–635

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Uncertainty and its propagation estimation for an integrated water system model: An experiment from water quantity to quality simulations

Yongyong Zhang[a,*], Quanxi Shao[b]

[a] Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[b] CSIRO, Data 61, Private Bag 5, Wembley, Western Australia 6913, Australia

ARTICLE INFO

This manuscript was handled by G. Syme, Editor-in-Chief, with the assistance of Jesús Mateo-Lázaro, Associate Editor

Keywords:
Uncertainty propagation
Model performance
Bootstrap resampling
SCE-UA auto-calibration
HEQM

ABSTRACT

Multiple uncertainty sources directly cause inaccurate simulations for water related processes in complicated integrated models, as such models include many interactive modules. A majority of existing studies focus on the uncertainties of parameter and model structure, and their effects on the model performance for a single process (e.g., hydrological cycle or water quality). However, comprehensive uncertainties of different modules and their propagations are poorly understood, particularly for the integrated water system model. This study proposes a framework of uncertainty and its propagation estimation for integrated water system model (HEQM) by coupling the Bootstrap resampling method and SCE-UA auto-calibration technique. Parameter and structure uncertainties of both hydrological cycle and water quality modules are estimated, including final distributions of parameters and simulation uncertainty intervals. Additionally, the effect of uncertainty propagation of hydrological parameters is investigated. Results show that: (1) HEQM simulates daily hydrograph very well with the coefficient of efficiency of 0.81, and also simulates the daily concentrations of ammonia nitrogen satisfactorily with the coefficient of efficiency of 0.50 by auto-calibration in the case study area; (2) The final ranges of all interested hydrological parameters are reduced obviously, and all the parameter distributions are well-defined and show skew. The uncertainty intervals of runoff simulation at the 95% confidence level bracket 18.7% of all the runoff observations due to uncertainties of parameter, and 86.0% due to both parameter and module structure, respectively; (3) The uncertainty propagation of hydrological parameters changes the optimal values of 37.5% of interested water quality parameters, but does not obviously change the water quality simulations which match well with the prior simulations throughout the period and bracket only 1.7% of observations at the 95% confidence level. Due to the further introduction of module structure uncertainties, 94.8% of observations are bracketed, only except the extreme high and low water quality concentrations; (4) The uncertainty of water quality parameters contributes 12.1% of total water quality simulations at the 95% confidence level. The figure increases to 21.0% and 92.0% if the uncertainty propagation of hydrological parameters, structure uncertainties of water quality module are considered, respectively. Therefore, although the parameter uncertainty and its propagation contribute a certain proportion of the whole simulation uncertainties, the module structure itself is the primary uncertainty source for the integrated water system model (HEQM), particular for the water quality modules.

1. Introduction

Linkages, interconnections and interdependencies of water cycle have been gradually recognized at basin or global scales due to the rapid growth of environmental science and the constantly emerging water issues (e.g., drought, flooding, erosion, pollution and ecological degradation) (GWSP, 2005). Integrated consideration or simulation of multiple water related processes become a new trend along with further explorations of interaction mechanisms among multiple processes, rapid developments of computer facilities and observation techniques of multiple data sources (Paola et al., 2006; Zhang et al., 2016a,b). Many successful model integrations have been implemented with different objectives in the earth system studies. For example, land surface models (e.g., VIC-Variable Infiltration Capacity; Liang et al., 1994) could be coupled with hydrological models to reveal interactions and feedbacks between atmosphere and hydrology at large scale. Similarly,

* Corresponding author.
E-mail address: Zhangyy003@igsnrr.ac.cn (Y. Zhang).

https://doi.org/10.1016/j.jhydrol.2018.08.070
Received 22 May 2018; Received in revised form 27 August 2018; Accepted 28 August 2018
Available online 31 August 2018
0022-1694/

hydrological models (e.g., SWAT-Soil and Water Assessment Tool; Arnold et al., 1998) could be coupled with soil erosion or biogeochemical processes to reveal precipitation-induced losses of water, soil and nutrient from lands to river networks. Hydrological models are also able to be couple with hydrodynamic and water quality models of water bodies (i.e., rivers or lakes) to capture the migrations of waterborne variables with high spatial and temporal resolutions, such as EFDC (Environmental Fluid Dynamic Code) (Hamrick, 1992). Ecosystem models (EPIC- Erosion/Productivity Impact Calculator; Sharpley and Williams, 1990; DNDC: DeNitrification/DeComposition; Li et al., 1992) could also be coupled with evapotranspiration model and soil biogeochemical model to reveal vegetation growth processes with considerations of nutrient and water stresses. Furthermore, integrated water system model is proposed and usually formed as a chain of water related modules to capture the interactions and feedbacks among physical, biological and geochemical processes, as well as the impacts of water-related human activities (GWSP, 2005), such as CLM series (Community Land Model) (Dai et al., 2003).

However, not all the processes are physically interpreted by mathematical equations due to the current insufficient knowledge (Willems, 2008). Empirical conceptualization and theoretical simplification are usually adopted and are easy to introduce potential uncertainties of parameters and model structures into the complicated models (Todini, 2007; Freni et al., 2009). For the integrated models that are made up of many modules, multiple uncertainty sources of upstream modules are transferred to the downstream modules as inputs (Freni et al., 2008). Therefore, along with the increasing of coupled modules, multiple uncertainty sources from different modules not only affect the simulation performance of their own modules, but also might be accumulated and propagated to subsequent modules and thus distort their performances and even the whole model. For example, in the applications of integrated water quantity and quality models, water quality simulation performance is not usually satisfying using the step-by-step calibration approach even though the water quality modules are well formulated, which is probably caused by the uncertainty or error propagation from upstream modules (Zhang et al., 2016b). It is critical to investigate the estimations of uncertainty sources of different modules and their propagation, as well as their effects on simulation performance.

Large majorities of studies are implemented to identify model uncertainty sources, and assess their effects on model performance. The identified uncertainty sources mainly include model parameter uncertainty (Beven and Binley, 1992; Bates and Campbell, 2001; Beven and Freer, 2001; Yang et al., 2007; Li et al., 2010a,b; Leta et al., 2015) and model structure uncertainty (Beven and Binley, 1992; Refsgaard et al., 2006; Li et al., 2010a). Most of the existing techniques are categorized into two classes, i.e., (1) the frequentist approach with model calibration techniques (e.g., Shuffled Complex Evolution: SCE-UA, Particle Swarm Optimization: PSO) which is advantageous to be implemented without timing consumption and the representative techniques are GLUE procedure (Generalized Likelihood Uncertainty Estimation) (Beven and Binley, 1992; Beven and Freer, 2001) and Parameter Solution (Duan et al., 1992), both of which are based on subjective determination of generalized likelihood measures between simulations and observations, Sequential Uncertainty Fltting algorithm (SUFI-2) with global sampling techniques based on multi-criteria thresholds of model calibration (Abbaspour et al., 2007), and Bootstrap resampling with recalibration (Li et al., 2010b; Novic et al., 2018); (2) the classical Bayesian theorem with sampling techniques (e.g., Markov Chain Monte Carlo, Latin hypercube) based on the observations and prior information of model parameters (Bates and Campbell, 2001; Engeland and Gottschalk, 2002; Montanari and Brath, 2004; Yang et al., 2007), which is robust and widely used to estimate the reliable uncertainties of model parameters. However, all of these studies are model-specific (Engeland and Gottschalk, 2002; Gallagher and Doherty, 2007), and only focus on the inherent uncertainties of single process model or model structure, e.g., hydrological models (Beven and Binley, 1992; Bates and Campbell, 2001; Engeland and Gottschalk, 2002; Montanari and Brath, 2004; Li et al., 2010a,b; Shao et al., 2014; Arsenault et al., 2015) and water quality models (Beck, 1987; Freni et al., 2008; Jia et al., 2018; Novic et al., 2018). The comprehensive uncertainty estimations in the integrated model of multiple processes are still deficient.

Furthermore, most of existing studies about uncertainty propagation analysis are limited to the impact assessments of observation quality on model performances which are induced by potential errors from sampling, instrument, laboratory, observation method or algorithm (Harmel et al., 2006; Xu et al., 2006; Leta et al., 2015). The related uncertainties include uncertainties associated with model inputs (e.g., climate variables, geographic information) (Crosetto et al., 2001; Gabellani et al., 2007; Shao et al., 2012; Yen et al., 2014; Novic et al., 2018) and observations used for the model calibration (e.g., flow regime, water quality variables) (Harmel et al., 2006; Shao et al., 2014; Yen et al., 2014). Both of two uncertainty sources are propagated and probably distort the probability distributions of model parameters, and thus disturb the capability of integrated models to portray the real world, particularly for the simulations of subsequent modules. For example, Harmel et al. (2006) examined that the probable uncertainties for observed variables ranged from 6% to 19% for streamflow, from 11% to 100% for NH_4-N, from 11% to 104% for total nitrogen (TN), and from 8% to 110% for total phosphorous (TP), among which the contributions of sample collection, preservation storage and laboratory analysis were from 4% to 48%, from 2% to 16% and from 5% to 21%, respectively. However, only a few studies are reported about the uncertainty propagation investigation among different modules, particularly for integrated water system models.

The main purpose of this study is to comprehensively assess the uncertainties of multiple modules and their propagations among different modules for complicated integrated water system models. As a typical integrated water system model, HEQM (Hydrological, Ecological and water Quality Model) is adopted to investigate the effect of multiple uncertainty sources on parameter distributions and simulation performances of both hydrological and water quality modules. The specific objectives are to: (1) propose a comprehensive assessment of uncertainty sources and their propagations for HEQM by using Bootstrap resampling with SCE-UA optimization technique; (2) estimate the probability distributions of hydrological parameters and uncertainty intervals of runoff simulation caused by uncertainties of parameters and module structures; (3) estimate the probability distributions of water quality parameters and uncertainty intervals of water quality simulation caused by parameter uncertainty propagation of hydrological cycle module; (4) estimate the probability distributions of water quality parameters and uncertainty intervals of water quality simulation caused by uncertainties of parameter and module structure. This study is expected to extend the scope of model uncertainty analysis, and assist modelers in further improvements and calibrations of complicated integrated models of multiple processes.

2. Models and methodology

2.1. Integrated water system model (HEQM)

HEQM is an integrated water system model proposed by Zhang et al. (2016a) in order to investigate hydrological cycle processes, its accompanied biogeochemical and water quality processes as well as their interactions at catchment scale. The main water related processes are mathematically described by hydrological cycle module (HCM), soil erosion module (SEM), overland water quality module (OQM), water quality module in water bodies (WQM), crop growth module (CGM), soil biochemical module (SBM) and dams regulation module (DRM) (Fig. S1 in the Supplementary material). Furthermore, a parameter analysis tool (PAT) is provided to conveniently conduct the parameter sensitivity analysis, model calibration and performance assessment. All