Environmental Modelling & Software 51 (2014) 84–93

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Watershed model parameter estimation and uncertainty in data-limited environments

André Fonseca [a,*], Daniel P. Ames [b], Ping Yang [c], Cidália Botelho [a], Rui Boaventura [a], Vítor Vilar [a,*]

[a] LSRE – Laboratory of Separation and Reaction Engineering – Associate Laboratory – LSRE/LCM, Department of Chemical Engineering, Faculty of Engineering of the University of Porto, Porto, Portugal
[b] Department of Civil and Environmental Engineering, Brigham Young University, Provo, UT, USA
[c] Texas Institute for Applied Environmental Research, Tarleton State University, Stephenville, TX, USA

ARTICLE INFO

Article history:
Received 3 June 2013
Received in revised form
25 September 2013
Accepted 27 September 2013
Available online

Keywords:
HSPF
GLUE
Monte Carlo simulations
Watershed modeling
Uncertainty assessment
Sensitivity analysis

ABSTRACT

Parameter uncertainty and sensitivity for a watershed-scale simulation model in Portugal were explored to identify the most critical model parameters in terms of model calibration and prediction. The research is intended to help provide guidance regarding allocation of limited data collection and model parameterization resources for modelers working in any data and resource limited environment. The watershed-scale hydrology and water quality simulation model, Hydrologic Simulation Program – FORTRAN (HSPF), was used to predict the hydrology of Lis River basin in Portugal. The model was calibrated for a 5-year period 1985–1989 and validated for a 4-year period 2003–2006. Agreement between simulated and observed streamflow data was satisfactory considering the performance measures such as Nash–Sutcliffe efficiency (E), deviation runoff (Dv) and coefficient of determination (R^2). The Generalized Likelihood Uncertainty Estimation (GLUE) method was used to establish uncertainty bounds for the simulated flow using the Nash–Sutcliffe coefficient as a performance likelihood measure. Sensitivity analysis results indicate that runoff estimations are most sensitive to parameters related to climate conditions, soil and land use. These results state that even though climate conditions are generally most significant in water balance modeling, attention should also focus on land use characteristics as well. Specifically with respect to HSPF, the two most sensitive parameters, INFILT and LZSN, are both directly dependent on soil and land use characteristics.

Software and data availability

Name BASINS 4.0 (Better Assessment Science Integrating point & Non-point Sources) with a non-proprietary, open source, free GIS system, MapWindow (www.MapWindow.org)
Developer U.S. EPA with AquaTerra Consultants and Idaho State University
Contact http://www.aquaterra.com/contact/index.php
Availability and cost The software is available for free download at USEPA (United States Environmental Protection Agency) website. Mapwindow is an open source programmable GIS (VB, C++, .NET, and Active X controls) that supports manipulation, analysis, and viewing of geospatial data and associated attribute data in several GIS data formats.
Name GLUEWIN
Developer EEMC – Euro-area Economy Modelling Centre
Contact Riccardo.Girardi@jrc.it
Availability and Cost GLUEWIN is a code designed for analyzing the output of Monte Carlo runs when empirical observations of the model output are available and implements the combination of GSA and GLUE methodologies. The software can be obtained for free at EEMC website (http://eemc.jrc.ec.europa.eu/Software-GLUEWIN.htm).

* Corresponding authors. Tel.: +351 918257824; fax: +351 225081674.
E-mail addresses: andrerdf@gmail.com (A. Fonseca), vilar@fe.up.pt (V. Vilar).

1364-8152/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envsoft.2013.09.023

1. Introduction

Like many similar government initiatives throughout the world, the European Union Water Framework Directive (WFD) was established to restore and protect both surface and ground water with ambitious goals to be met by a target date of 2015 (2000/60/EC, 2000). Watershed modeling software can be used to help scientists and watershed managers to meet these goals by simulating the effect on water quality and quantity considering different water management strategies, types of land use and climate change.

Many commercial and open source watershed simulation models are available and much care and consideration needs to be employed when choosing a model for application in a particular watershed (Borah and Bera, 2004). Regardless of the model chosen, even greater attention must be given to the task of "populating" the selected model with appropriate and physically meaningful model parameters that accurately characterize the surficial landscape, subsurface geology, atmospheric conditions, and other constraints affecting the storage and flux of water and contaminants through the environment (Donigian, 2002; Doherty and Johnston, 2003).

Because many watershed simulation model parameters are difficult or impossible to measure in the natural world, parameters must often be estimated or otherwise evaluated from secondary information sources and hence are typically laden with notable degrees of uncertainty (Gallagher and Doherty, 2006). This paper presents a watershed modeling study in the Lis River basin (Portugal) with the express purpose of identifying those parameters in the selected watershed model, whose accurate characterization is most critical for successful model application. The study includes a complete model parameterization and calibration effort combined with parameter uncertainty estimation techniques. Great attention is given to accurate characterization of those parameters for the success of the modeling effort. The results provided here can be used to inform other modelers in data and resource limited situations as to which parameters warrant the greatest resource-allocation and technical attention – allowing for the more efficient use of limited resources.

The hydrological model, Hydrologic Simulation Program FORTRAN (HSPF) was used in this study. HSPF is based on the original Stanford Watershed Model IV (Crawford and Linsley, 1966) and is a consolidation of three previously developed models: Agricultural Runoff Management Model (ARM) (Donigian and Davis, 1978), Non-point Source Runoff Model (NPS) (Donigian and Crawford, 1976) and Hydrological Simulation Program (HSP) including HSP Quality (Donigian et al., 1991, 1995; Hydrocomp, 1977).

HSPF is a semi-distributed model that simulates water and contaminant transport through spatially distributed, physically homogenous areas within a watershed called Hydrologic Response Units (HRUs). HRUs are presumed to hydrologically respond similarly to given meteorological inputs (precipitation, potential evapotranspiration and temperature). In this way, HSPF can simulate the hydrological, hydraulic and water quality processes on pervious and impervious land surfaces, in soil profiles and in streams and well-mixed impoundments on a continuous basis (Bicknell et al., 2001).

A graphical user interface for HSPF is included in the free software, BASINS, developed and distributed by the United States Environmental Protection Agency (http://water.epa.gov/scitech/datait/models/basins/index.cfm). BASINS is built on the open source geographic information system (GIS) MapWindow GIS (Ames et al., 2008). The use of BASINS to develop HSPF models has been reported in several studies (Bergman et al., 2002; Carrubba, 2000; Lian et al., 2007; Lowe and Doscher, 2003; J. Zhang et al., 2009).

Calibration of HSPF is an iterative procedure of parameter evaluation, as a result of comparing simulated against observed values of interest. Since HSPF uses a large number of parameters that can be adjusted to represent the physical environment, an expert system, HSPExp is available to assist modelers with hydrology calibration. Typically a dozen or less parameters are used in most studies. HSPExp advises the user on which parameters can be meaningfully adjusted to reduce simulation error while providing explanations regarding the modifications (Lumb et al., 1994). However, there are limitations to HSPF, such as limited spatial definition (finite element analysis model), limited to non-tidal freshwater systems and extensive data requirements (i.e. meteorological and most important gaging stations of interest in the watershed).

Model validation is necessary in any model application and is an extension of the calibration process. Its purpose is to assure that the calibrated model properly assesses all the variables and conditions which can affect model results and, the ability to predict the behavior for periods separate from the calibration. Model credibility is based on the ability of a single set of parameters to represent the entire range of observed data. If a single parameter set can reasonably represent wide range of events, then this is a form of validation (Donigian, 2002).

Sensitivity and uncertainty analysis of model parameters is conventionally considered to be one of the primary steps in the development and evaluation of models (Sudheer et al., 2011; Jakeman et al., 2006). Over the past decade it has become widely accepted that hydrological models with numerous parameters are likely to produce equally acceptable predictions for multiple different parameter sets (Hope et al., 2004) and a unique "best" parameter set cannot necessarily be found in the parameter space (Christiaens and Feyen, 2001). A structured method to quantify model uncertainty, Generalized Likelihood Uncertainty Estimation (GLUE), proposed by K. Beven and Binley (1992) establishes uncertainty bounds for the simulated value using parameter sets that are determined to be acceptable based on a performance likelihood measure. The implementation of this method requires the user to make a number of subjective decisions such as the threshold value of the likelihood measure to classify the model as acceptable or unacceptable (K. Beven and Binley, 1992; K. J. Beven, 2001). With GLUE, model parameters are sampled from distributions, typically with independent uniform or normal distributions for each parameter. The model is then run with each parameter set, generating multiple sets of model output, which are used to generate uncertainty intervals for model predictions. The generated model parameters are grouped in two categories: behavioral, sets of model parameters that produce results consistent with the observations and non-behavioral, results that contradict the observations (Spear and Hornberger, 1980).

As discussed previously the model HSPF, requires extensive parameterization to simulate hydrology fluxes due to both uncertainty in modeled processes and observation errors (Ratto et al., 2001).

The remainder of this paper includes a discussion of the methods employed, and results related to the hydrology calibration parameters with focus on assessing the sensitivity of the model to selected parameters and determination of parameter priority in the model.

2. Methods

2.1. Study area

The Lis River basin is one of the most important natural resources of the Leiria region in Portugal, with a population rich in fish species and much sought for sport