Environmental Modelling & Software 68 (2015) 129–146

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Assessment of the different sources of uncertainty in a SWAT model of the River Senne (Belgium)

Olkeba Tolessa Leta [a,*], Jiri Nossent [a], Carlos Velez [a], Narayan Kumar Shrestha [a], Ann van Griensven [a,b], Willy Bauwens [a]

[a] Department of Hydrology and Hydraulic Engineering, Earth System Sciences Group, Vrije Universiteit Brussel (VUB), Brussels, Belgium
[b] UNESCO-IHE Institute for Water Education, Core of Hydrology and Water Resources, The Netherlands

ARTICLE INFO

Article history:
Received 4 March 2014
Received in revised form
27 December 2014
Accepted 16 February 2015
Available online 7 March 2015

Keywords:
Uncertainty assessment
Rainfall uncertainty
Stream flow
SWAT
DREAM
River Senne

ABSTRACT

Although rainfall input uncertainties are widely identified as being a key factor in hydrological models, the rainfall uncertainty is typically not included in the parameter identification and model output uncertainty analysis of complex distributed models such as SWAT and in maritime climate zones. This paper presents a methodology to assess the uncertainty of semi-distributed hydrological models by including, in addition to a list of model parameters, additional unknown factors in the calibration algorithm to account for the rainfall uncertainty (using multiplication factors for each separately identified rainfall event) and for the heteroscedastic nature of the errors of the stream flow. We used the Differential Evolution Adaptive Metropolis algorithm (DREAM$_{(zs)}$) to infer the parameter posterior distributions and the output uncertainties of a SWAT model of the River Senne (Belgium). Explicitly considering heteroscedasticity and rainfall uncertainty leads to more realistic parameter values, better representation of water balance components and prediction uncertainty intervals.

Published by Elsevier Ltd.

1. Introduction

River basin simulators are useful tools to support the decision making processes in view of an integrated river basin management. As a result of the advances in techniques for data acquisition and measurement, the use of spatially distributed, physically-based and thus complex river basin simulators has gained increased attention. However, most of the parameters of such simulators cannot be measured directly and therefore need to be estimated by model parameter optimization techniques (Laloy et al., 2010; Laloy and Vrugt, 2012; Nossent, 2012; Sorooshian and Dracup, 1980). Traditionally, model parameter optimization has been practiced through minimizing the difference between the observations and the model simulations (e.g. the sum of squared residuals). This approach primarily focuses on the parameter uncertainty, thereby neglecting input, calibration data and model structure uncertainties (Ajami et al., 2007; Arnold et al., 2012; Di Baldassarre and Montanari, 2009; Götzinger and Bárdossy, 2008; Gupta et al., 2012; McMillan et al., 2011; Minasny et al., 2011; Refsgaard et al., 2007; Renard et al., 2010; Thyer et al., 2009; Vrugt et al., 2009a, 2008; Yang et al., 2007).

In hydrological modeling, most of the catchment processes are conceptualized based on the (limited) physical understanding of these processes. This inherently results in a simplification of the actual processes and of their spatial variability (Abbaspour et al., 2007). In addition, the hydrological simulators require (uncertain) input data (e.g. rainfall) and they contain conceptual parameters that need to be estimated through a calibration process. Parameter estimation remains, however, uncertain and the calibrated parameter values do not necessarily correspond to the physical reality (Abbaspour et al., 2007; Götzinger and Bárdossy, 2008; Yang et al., 2008). The latter is due to random or systematic errors in the initial conditions (e.g. soil moisture content), in the model inputs, in the observed output data that are used for the model calibration (e.g. stream flow) and to errors due to the incomplete or biased model structure.

In the field of hydrology, rainfall uncertainty typically dominates the uncertainty of the input data (Ajami et al., 2007; Renard et al., 2009, 2011; Vrugt et al., 2008). The rainfall uncertainty is mainly due to measurement and sampling errors and to the spatial and temporal variability of the rainfall (Götzinger and Bárdossy, 2008; Laloy et al., 2010; McMillan et al., 2011; Thyer et al., 2009). Even though it remains a challenging task to address the rainfall

* Corresponding author. Tel.: +32 2 629 3027; fax: +32 3 205 2437.
E-mail address: olketole@yahoo.com (O.T. Leta).

http://dx.doi.org/10.1016/j.envsoft.2015.02.010
1364-8152/Published by Elsevier Ltd.

uncertainty, some promising techniques have emerged in this field. One of them consists of the use of multiplicative rainfall uncertainty factors at storm event scale (Kavetski et al., 2006a, b; Thyer et al., 2009; Vrugt et al., 2009a, 2008) or at daily time scale (Thyer et al., 2009). Recently, a suitable representation of the rainfall uncertainty through the use of rainfall multiplicative factors has been validated against experimental evidence by McMillan et al. (2011). However, the method has so far only been tested on lumped simulators and not for more complex (semi-) distributed simulators (Kavetski et al., 2006b; Thyer et al., 2009; Vrugt et al., 2009a, 2008). Additionally, the feasibility and the applicability of such methods have not been demonstrated in maritime climate regions where rainfall is very frequent and persists for longer periods. For such situations, the large number of rainfall multipliers might indeed lead to an over-parameterization of the model and to high computational demands.

In general, neglecting different sources of uncertainty during the model optimization may result in model outputs that cannot consistently represent the observations (Ajami et al., 2007). More importantly, the residuals between the observations and the model outputs can be characterized by a significant variation in bias (non-stationarity), variance (heteroscedasticity) and serial dependence (autocorrelation) under different hydrologic conditions (Ajami et al., 2007; Schoups and Vrugt, 2010; Sorooshian et al., 1983; Vrugt et al., 2005; Yen et al., 2014).

In this study, the rainfall multiplication method has been applied for a complex, highly parameterized, distributed model in a maritime region, using the Soil and Water Assessment Tool (SWAT). Considering the high rainfall frequency in the region and the multi-parameter of (semi-) distributed model, we expected that the existing methods for rainfall multiplication factors might lead to high-dimensionality problems during the model calibration (due to the fact that many additional parameters have to be assessed). Consequently, we adapted the latter methods in order to reduce the number of rainfall multipliers. Hereto, we propose a method that still applies rainfall multipliers, but that minimizes the number of factors by focusing on significant, independent rainfall events, as determined by an analysis of the observed stream flow time series.

In addition to the input and parameter uncertainties, also the uncertainty on the observed output variable used for the calibration – i.e. the stream flow – was considered. Stream flow data are not error-free, mainly as a result of the uncertainty in the stage-discharge rating curve (Domeneghetti et al., 2012; McMillan et al., 2010; Renard et al., 2011; Thyer et al., 2009).

In what follows, we use a traditional approach for model calibration and uncertainty analysis, whereby only model parameter uncertainty is considered, and compare it with an approach whereby we explicitly consider the rainfall uncertainty and the heteroscedastic nature of the error of the model outputs. For the latter, the rainfall uncertainty is represented in the calibration process by unknown independent rainfall multipliers and the error heteroscedastic behavior of the output errors is represented by statistical parameters that need to be calibrated. To infer and assess all these sources of uncertainty, we applied the Differential Evolution Adaptive Metropolis algorithm ($\text{DREAM}_{(\text{zs})}$) (Laloy et al., 2012; Laloy and Vrugt, 2012; Vrugt et al., 2009b). $\text{DREAM}_{(\text{zs})}$ is a Markov Chain Monte Carlo (MCMC) sampler (Bates and Campbell, 2001; Kuczera and Parent, 1998) that uses sampling from past states to select candidate points for the individual chains.

This paper concerns the uncertainty analysis (UA) for a SWAT model of the River Senne in Belgium. The objectives of the paper are:

a) to evaluate the feasibility of applying rainfall correction factors for independent rainfall events for a semi-distributed, physically-based environmental model in a region with a maritime climate;
b) to assess the different sources of uncertainty in the SWAT model of the River Senne and their impact on the parameter distributions;
c) to assess the impact of considering or neglecting different sources of uncertainty on the predictive uncertainty of the simulated stream flows.

2. Materials and methods

2.1. The SWAT simulator

The Soil and Water Assessment Tool (SWAT) is a physically-based, semi-distributed, hydrologic simulator that operates on different time steps at the basin-scale. SWAT was originally developed to simulate the impact of watershed management on water, sediment, nutrients and agricultural and chemical yields (Arnold et al., 1998). Moreover, SWAT can model complex watersheds with varying land use, weather, soils, topography and management conditions over a long period of time. A watershed is divided into a number of sub-basins that have homogeneous climatic conditions (Van Liew et al., 2005). Sub-basins are further sub-divided into hydrological response units (HRUs), based on a homogenous combination of land use, soil type and slope class (Arnold et al., 2011).

SWAT assesses a water balance by considering precipitation, evapotranspiration, surface runoff, interflow, return flow and deep groundwater losses (Neitsch et al., 2011). The simulator offers a choice to use a modification of the Soil Conservation Service Curve Number (SCS-CN) method (USDA-SCS, 1986), which determines the surface runoff based on the area's hydrologic soil group, land use, treatment, and antecedent moisture content for each HRUs or, alternatively if sub-daily precipitation data are provided, the Green and Ampt method (Green and Ampt, 1911) as modified by Mein and Larson (Mein and Larson, 1973). The percolation through each soil layer is estimated using a storage routing techniques (Arnold et al., 1995). SWAT offers three options to estimate the potential evapotranspiration (PET) from climatic data: the Penman-Monteith method (Monteith, 1965), the Hargreaves method (Hargreaves et al., 1985) and the Priestley-Taylor method (Priestley and Taylor, 1972). The PET can also be read in from a file if measured time series of PET data are available. River routing can be performed by the variable storage method (Williams, 1969) or by the Muskingum method (Chow, 1959). In this paper, we used the SCS-CN method for surface runoff simulations, the Penman-Monteith method for the PET estimation and the Muskingum routing method for the daily stream flow routing.

2.2. The DREAM algorithm

Various Bayesian algorithms exist for model optimization and uncertainty quantifications, e.g. the Generalized Likelihood Uncertainty Estimation, GLUE (Beven and Binley, 1992; Blasone et al., 2008; Feyen et al., 2008; Freer et al., 1996; Jin et al., 2010; Rojas et al., 2010), the Bayesian Total Error Analysis, BATEA (Kavetski et al., 2006b, c; Kuczera et al., 2006; Renard et al., 2011; Thyer et al., 2009), the Bayesian Model Averaging technique, BMA (Ajami et al., 2007; Duan et al., 2007; Raftery et al., 1997; Rings et al., 2012; Vrugt et al., 2006; Vrugt and Robinson, 2007) and the Differential Evolution Adaptive Metropolis, DREAM (Laloy and Vrugt, 2012; Schoups and Vrugt, 2010; Vrugt et al., 2008, 2009b). In this paper, the latter method has been selected to explicitly account for the different sources of uncertainty and to approximate uncertainty bands based on the posterior probability density function (pdf) of the considered parameters.

The DREAM algorithm is based on Markov Chain Monte Carlo (MCMC) sampling schemes (Bates and Campbell, 2001; Kuczera and Parent, 1998) and employs Bayesian updating for maximizing a likelihood function. It was primarily designed for simultaneous model parameter optimization and uncertainty analysis (UA) for high dimensional problems, and is particularly suited for parallel computing. For population evolution, Differential Evolution (DE) is applied in combination with a Metropolis selection rule to define whether the current candidate replaces the parent or not. The algorithm is mainly developed based on the Differential Evolution-Markov Chain (DE-MC) method (ter Braak, 2006) and the Shuffled Complex Evolution Metropolis (SCEM) algorithm (Vrugt et al., 2003). The main adaptation with regard to DE-MC is the addition of self-adaptive randomized subspace sampling. DREAM can also be seen as an adaptation of the SCEM algorithm, whereby DREAM provides a formal proof of convergence, allows for parallel computing and is able to deal with multimodal distributions (Nossent, 2012).

DREAM starts with an initial population of points to strategically sample the space of potential solutions. It searches for the posterior distributions of parameter values by maximizing the posterior parameter distribution using Bayesian updating, while simultaneously providing estimates of the parameter uncertainty. The posterior distributions of the considered parameters can be obtained after the run converges to its stable posterior distribution (Laloy et al., 2010). The convergence can be monitored using the $\hat{R}$ criterion proposed by Gelman and Rubin (1992). The