Environmental Modelling & Software 88 (2017) 22–34

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Combined analysis of time-varying sensitivity and identifiability indices to diagnose the response of a complex environmental model

Mehdi Ghasemizade [a, b, *], Gabriele Baroni [c, d], Karim Abbaspour [b], Mario Schirmer [b, e]

[a] University of California, Davis, 1 Shields Avenue, Dept. LAWR, USA
[b] Eawag: Swiss Federal Institute of Aquatic Science and Technology, Ueberlandstrasse, 133, CH-8600 Dübendorf, Switzerland
[c] UFZ - Helmholtz Centre for Environmental Research, Permoserstrasse 15, 04318 Leipzig, Germany
[d] Institute of Earth and Environmental Sciences, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany
[e] University of Neuchâtel, Centre of Hydrogeology and Geothermics (CHYN), Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

ARTICLE INFO

Article history:
Received 29 November 2015
Received in revised form
29 October 2016
Accepted 30 October 2016

Keywords:
Temporal sensitivity
Identifiability
Preferential flow
HydroGeoSphere
Output uncertainty

ABSTRACT

Sensitivity and identifiability analyses are common diagnostic tools to address over-parametrization in complex environmental models, but a combined application of the two analyses is rarely conducted. In this study, we performed a temporal global sensitivity analysis using the variance-based method of Sobol' and a temporal identifiability analysis of model parameters using the dynamic identifiability method (DYNIA). We discuss the relationship between the two analyses with a focus on parameter identification and output uncertainty reduction. The hydrological model HydroGeoSphere was used to simulate daily evapotranspiration, water content, and seepage at the lysimeter scale. We found that identifiability of a parameter does not necessarily reduce output uncertainty. It was also found that the information from the main and total effects (main Sobol' sensitivity indices) is required to allow uncertainty reduction in the model output. Overall, the study highlights the role of combined temporal diagnostic tools for improving our understanding of model behavior.

1. Introduction

The advances in computer science have supported the development and use of integrated and complex environmental models. In principle, the advantage of such models is that they provide a detailed description of the system, including the spatial and temporal incorporation of the relevant processes. Therefore these models offer a variety of possibilities in scenario analysis and decision support systems (De Lange et al., 2014). However, the applicability of such models has found some limitations (Beven, 2006). Among others, parametrization of complex environmental models is recognized as a crucial step for a proper model application. On the one hand, a large number of parameters usually required by these complex models cannot always be measured directly. On the other hand, inverse modeling to determine these specific parameters could be hindered by the limitation in the availability of data. Under these conditions the available observations do not provide sufficient information for the identification of the model parameters and therefore compromise model performance. For these reasons, there remains a need for diagnostic methods (Gupta et al., 2008; Matott et al., 2009) to link model formulation to its consequent impacts on process-level behavior in order to inform model selection, calibration, and interpretation (Herman et al., 2013). Diagnostic methods explore the input-output response of the models and thus give a better overview of the model behavior.

Sensitivity analysis (SA) is a diagnostic method which can be used for exploring uncertainty within complex parameter spaces and interpreting model behavior in the context of the system being modeled (Franchini et al., 1996; Hall et al., 2005; Herman et al., 2013). Among many, global sensitivity analysis (GSA) methods are preferred since they study the effects of input variations on the outputs in the entire allowable ranges of the input space. Sobol' analysis is a global and variance-based method which is independent of model linearity and model monotonicity (Saltelli et al., 2010). It separates the contributions of individual parameters as well as the interacting contributions of the parameters to the output variance. In this paper, we address two main purposes of

* Corresponding author. University of California, Davis, 1 Shields Avenue, Dept. LAWR, USA.
E-mail addresses: mghasemizade@ucdavis.edu, mehdighasemizade@ymail.com (M. Ghasemizade), gabriele.baroni@ufz.de (G. Baroni), Karim.Abbaspour@eawag.ch (K. Abbaspour), mario.schirmer@eawag.ch (M. Schirmer).

http://dx.doi.org/10.1016/j.envsoft.2016.10.011
1364-8152/

Sobol' analysis. The first one helps to identify the parameters that are not important (low/insensitive parameters) to model outputs. The aim of this application is to reduce the number of free-varying and less influential parameters by assigning constant values to them, as these parameters have a minimum potential to reduce the output uncertainty (factor fixing). For example, van Werkhoven et al. (2009) applied Sobol' analysis in order to identify insensitive parameters that had an overall sensitivity level less than a certain threshold. Although they reduced the burden of the calibration process by excluding insensitive parameters in the calibration (assigning fixed values to them), their simulation results had essentially the same predictive performance as opposed to the case where all the parameters were included in calibration. For this reason, the second application of Sobol' analysis which is known as factor prioritization is to identify the important parameters (sensitive) to the model output response or an objective function. The focus of factor prioritization is on those parameters that have the potential to maximally reduce the output uncertainty and to improve the performance of the model output (Abebe et al., 2010; Saltelli et al., 2006). As a diagnostic tool, Sobol' analysis has been used temporarily to identify the model components which control the performance under different conditions (Herman et al., 2013; Massmann and Holzmann, 2012). Temporal application of Sobol' can help diagnose to what extent and at what sensitivity level the parameters can impact the model performance in different time periods (Garambois et al., 2013; Massmann et al., 2014) and therefore can assist in identifying the dominant processes and conditions for deriving a more accurate estimation of the model parameters (Guse et al., 2014; Reusser and Zehe, 2011).

Another relevant diagnostic tool used to explore the input-output response of a model is the so-called identifiability analysis (IA). Various definitions of parameter identifiability can be found in the literature (e.g., Brun et al., 2001; Matott et al., 2009) and therefore, several methods have been developed to address parameter identifiability (Bastidas et al., 1999; Doherty and Hunt, 2009; Wagener et al., 2003). However, the goal of some of these methods overlaps with the optimization approaches (Shin et al., 2015). In this study, identifiability is defined as the capability to constrain the range of variation of parameters for a given set of available observations. The dynamic identifiability analysis (DYNIA) method developed by Wagener et al. (2003) is a method specifically developed for the identifiability analysis. It has been widely used since it is relatively immune for the effects of model nonlinearity on parameter estimates and also due to its ease of use. The advantage of the DYNIA method, in comparison to the other competitive identifiability approaches, is that it avoids the aggregation of model residuals into an objective function calculated for the entire simulation time. In fact, DYNIA does not let specific modes of hydrological simulations dominate the individual response modes, because it calculates the objective function for running windows that are shorter than the entire simulation time. This capability of DYNIA enables a modeler to determine the time periods where narrow range of parameter values can be obtained. For example, Abebe et al. (2010) applied DYNIA analysis on a rainfall-runoff model and found a systematic dependence between some model parameters and the state of soil moisture, indicating a calibration scheme with state variable, where possible, may help find the representative values of a given catchment.

While sensitivity and identifiability analyses are relevant, they answer different questions. However, in some studies, identifiability of model parameters has been discussed in terms of parameter sensitivity, assuming that the chances of parameter identifiability increase as the parameter becomes more sensitive (Hartmann et al., 2013; Kelleher et al., 2013; Shin et al., 2013). Cibin et al. (2010) evaluated the identifiability of SWAT model parameters visually and showed that high parameter sensitivity does not necessarily lead to parameter identifiability. Recently, Pianosi and Wagener (2016) presented a SA and IA showing the relative importance of parameter uncertainty versus data uncertainty. Therefore, the main objective of our study is to investigate whether there is a clear relationship between identifiability and sensitivity analyses. In addition, as the two analyses can be accomplished with the same inputs, and both of the analyses show different aspects of the input-output space, we chose their combined use for diagnosis analysis. To that end, we applied the numerical and physically-based model HydroGeoSphere (HGS) (Therrien et al., 2010) to simulate the daily water balance components (i.e. drainage, evapotranspiration and soil moisture) in the framework of a weighing lysimeter. The model has 28 parameters and consists of four soil layers as well as a preferential flow component. Due to the advantageous properties of the diagnostic Sobol' analysis, a temporal sensitivity analysis is conducted to analyze the temporal dynamics of parameter sensitivity. This analysis can help to diagnose the time periods where specific model components such as preferential flow dominate the simulation response. It also helps to apportion the uncertainty in the outputs of the model to the individual contribution of the model parameters as well as to the interaction of the parameters. In the second step, we run a temporal identifiability analysis to distinguish the time periods where parameter values can be constrained based on the available observations, i.e. free drainage (lysimeter discharge), evapotranspiration and soil moisture. In the end, we discuss the relationship between the results obtained from the temporal SA and IA and show how these two analyses provide complimentary information.

2. Methods

2.1. Experimental site and data-set

The lysimeter of our study belongs to the category of large weighing lysimeters and is located in the Rietholzbach catchment in northeast Switzerland. The mean annual precipitation, evapotranspiration, and temperature are 1450 mm, 560 mm and 7.1 °C, respectively (Seneviratne et al., 2012). The lysimeter is 2.5 m deep and has been back filled with gleyic cambisol soil from the surrounding area. The vegetation on top of the lysimeter is grass and represents the surrounding area. Free draining seepage, actual evapotranspiration, and water content are continuously measured since 1976. Soil moisture content is measured with time domain reflectometry (TDR) at depths of 5, 15, 25, 55 and 80 cm. The lysimeter is located close to a weather station where precipitation, net radiation, temperature and wind speed are measured continuously. Precipitation is measured with heated tipping buckets located at 1.5 m above ground and at the ground level. We used the above ground gauge data because it is less biased, particularly in winter. For more details on the lysimeter and errors of the measuring devices we refer to Seneviratne et al. (2012) and Ghasemizade et al. (2015).

2.2. Model set-up

The HGS model has demonstrated good capability in reproducing the main components of the water balance under different conditions (Li et al., 2008; Rozemeijer et al., 2010; Zhu et al., 2012). HGS simulates evapotranspiration based on the method of Kristensen and Jensen (1975) and matrix flow based on Richards equation. Due to the existence of preferential flow in the Rietholzbach lysimeter (Menzel and Demuth, 1993; Vitvar and Balderer, 1997), the preferential flow component was included in our modeling framework. We applied the method of dual