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

Corresponding author:

Mehdi Ghasemizade^{1,2}, mghasemizade@ucdavis.edu; mehdi.ghasemizade@eawag.ch;
mehdighasemizade@ymail.com

Co-authors:

Gabriele Baroni^{3,4}, gabriele.baroni@ufz.de

Karim Abbaspour², Karim.Abbaspour@eawag.ch

Mario Schirmer², mario.schirmer@eawag.ch

Affiliations:

¹ University of California, Davis, 1 Shields Avenue, Dept. LAWR, USA.

² Eawag: Swiss Federal Institute of Aquatic Science and Technology, Ueberlandstrasse 133, CH-8600 Dübendorf, Switzerland.

³ UFZ - Helmholtz Centre for Environmental Research, Permoserstrasse 15, 04318 Leipzig, Germany

⁴ Institute of Earth and Environmental Sciences, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

Abstract

Sensitivity and identifiability analyses are common diagnostic tools to address the problem of over-parametrization in complex environmental models, but 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 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 both Sobol' sensitivity indices (main and interaction effects) are 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.

Key words: Temporal sensitivity; Identifiability; Preferential flow; HydroGeoSphere; Output uncertainty

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). Tang et al. (2007) compared state of the art in sensitivity analysis and found that Sobol' analysis, which is a global variance-based sensitivity analysis (GVSA) method, yielded more robust sensitivity ranking. Sobol' analysis is independent of model linearity and model monotonicity (Saltelli et al., 2010) and studies the effects of input variations on the outputs in the entire allowable ranges of the input space. It separates the contributions of individual parameter as well as their interacting contributions to the output variance. In general, Sobol' analysis has two main purposes. The first one helps to identify the parameters that are not important (insensitive) to model outputs. The aim of this application is to reduce the number of free-varying and less influential parameters by removing them or assigning constant values to them, as these parameters have 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 from the calibration, their model results had essentially the same predictive performance as the non-fixed parameter case. 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 more accurate estimations (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). The idea in this study is to identify those parameters whose range of variations can be constrained via given 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 the ease of use of its employed algorithm. 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 and commonly used for the study of model behaviors, the two analyses are rarely used combined, particularly temporally for diagnostic analysis. In most of the studies, identifiability of model parameters has been discussed in terms of their 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 model parameters qualitatively (visually) and showed that sensitivity and identifiability analysis provide complementary information to model behavior.

This paper aims at a better evaluation of the relationship and dependency between the two analyses, i.e., SA and IA. In this study, we applied the numerical and physically-based model HydroGeoSphere (HGS) (Therrien et al., 2010) to simulate the daily water balance components in the framework of a weighing lysimeter. The model has 28 parameters and consists of four soil layers in addition to the preferential flow component. Due to the advantageous properties of the diagnostic methods such as Sobol' analysis, a first objective of this research is to conduct a temporal sensitivity analysis to analyze the temporal dynamics of parameter sensitivity to the model output variables. This analysis can help to determine 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 interaction of the parameters as well as individual contribution of the model parameters. In the second step, we run a temporal identifiability analysis to extract the maximum information content from available observations i.e., free drainage lysimeter discharge, evapotranspiration and soil moisture. The focus of this analysis is twofold: (1) to determine the time periods where model parameters are highly identifiable; (2) to investigate whether there is a dependency between soil moisture as a state variable and the time periods where model parameters are highly identifiable. 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. The main innovations of this work are that the same outputs used for SA were used for IA and a combined analysis of temporal IA and SA was performed.

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 heating 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 permeability, developed by Gerke and van Genuchten (1993), to include preferential flow simulation. The simulation of flow in the unsaturated zone was carried out based on the formulation of the retention and the unsaturated hydraulic conductivity curves of van Genuchten and Mualem (Mualem, 1976; van Genuchten, 1980). Descriptions of the main equations relevant for the present study are reported in Appendix A.

The first two layers (represented by TDR measurements at the depths of 5 and 15 cm) were merged, due to the similarity of their water content data. Therefore, the lysimeter domain was divided into four different soil layers at the depths of 10, 25, 55 and 80 cm from the top (surface) of the lysimeter, each represented by the measurements of soil moisture data with TDRs. It was assumed that the bottom soil layer of the lysimeter (80-250 cm) is homogenous and has the same parameter values of the fourth layer. The lysimeter depth was discretized vertically into 125 layers, each having a thickness of 2 cm, for numerical simulation of the 1D flow. The sides of the domain were all assigned as no flow boundaries, while the bottom boundary was assumed to be a free drainage boundary. Daily precipitation and potential evapotranspiration (ET_p), calculated based on the method of Penman (1948), were extracted from meteorological data available at the experimental site. These time series were imposed as top boundary conditions. The grass on the top of the lysimeter is cut four times per year from June to October with an average time interval of 45 days. In order to define the leaf area index (LAI) for interception and transpiration calculations, we assumed that LAI grows linearly from its minimum to maximum during each cutting period. The range of variability for LAI of the lysimeter grass was extracted from Scurlock et al. (2001). It was also assumed that no overland flow occurs on the top of the lysimeter, due to being surrounded by an edge on the top. The study was conducted from April 2012 to November 2013, and was

intended to simulate the daily variability of evapotranspiration, lysimeter discharge (from this point on discharge) and average-profile water content (from this point on water content) of the four soil layers. A warm-up period of 100 days was applied to eliminate the sensitivity of model outputs to the initial conditions.

2.3 Temporal and global sensitivity analysis

The temporal and global sensitivity analysis considered in this study is based on the variance-based approach proposed by Sobol' (2001) and further developed by Saltelli et al. (2010). The method decomposes the model output variance into variances of individual model parameters as well as the variance of interactions among model parameters. The variance decomposition can be written as follows:

$$v = \sum_i v_i + \sum_{i < j} v_{ij} + \sum_{i < j < m} v_{ijm} + \dots v_{12..k} \quad (1)$$

where

$$v_i = v\left(E\left(Y\left|X_i = x_i^*\right.\right)\right) \quad (2)$$

$$v_{ij} = v\left(E\left(Y\left|X_i = x_i^*, X_j = x_j^*\right.\right)\right) - v\left(E\left(Y\left|X_i = x_i^*\right.\right)\right) - v\left(E\left(Y\left|X_j = x_j^*\right.\right)\right) \quad (3)$$

In the above formulas, v is the variance over all the possible values of X_i , Y is the model output, X_i denotes an input factor of the model, and $E\left(Y\left|X_i = x_i^*\right.\right)$ denotes the expectation of Y conditional on X_i having a fixed value equal to x_i^* . Similarly the higher order variances can be computed. However, due to the high computational demands, it is not very practical to calculate the higher order variances individually i.e., two by two, three by three, etc. Instead, two indices, named main effect/first order index (S_i) and total effect/total order index (S_{Ti}), are calculated for each model parameter as follows:

$$S_i = \frac{v_{X_i}(E_{X_{\sim i}}(Y|X_i))}{v(Y)} \quad (4)$$

$$S_{Ti} = \frac{E_{X_{\sim i}}(v_{X_i}(Y|X_{\sim i}))}{v(Y)} \quad (5)$$

where $X_{\sim i}$ denotes the array of all inputs except X_i , v and E denote variance and expectation operators, respectively. The main effect represents the average reduction in the variance of the model output when the input factor (X_i) is fully known and fixed. In other words, it shows the individual contribution of parameter i to the total variance of model output. The total order index includes all the interactions that X_i has with other parameters as well as its main effect and can be interpreted as the expected variance that would be left if all

parameters, except for X_i , are fixed. S_i and S_{Ti} both vary in the range of zero to one and their difference indicates the ratio of the contribution to the total output variance that rises due to the interaction of the parameter i with other parameters.

Sensitivity indices in equations (4) and (5) can be estimated using Monte Carlo numerical integration methods. One of the most used approach is the method proposed by Sobol' (2001) and further developed by Saltelli (2002) and Saltelli et al. (2010). This method requires global and particular sampling approaches from the parameter space. Latin hypercube sampling (LHS) and Sobol' sequence random sampling methods are very popular due to their efficient stratification properties (Nossent et al., 2011; Zhang et al., 2013). Burhenne et al. (2011) compared standard sampling procedures for simulation applications and concluded that sampling based on Sobol' sequences performs better. We, therefore, used a quasi-random Sobol' sequence of parameter values with uniform distributions in this work. The analysis includes 28 parameters representing preferential flow, soil matrix characteristics, and evapotranspiration. The range of parameter variabilities and their descriptions are all depicted in Table 1. Ranges in Table 1 were carefully defined from an understanding of the system, observed/measured data and from values used in previous studies. For more detailed information on the selection of parameter ranges, readers are referred to Ghasemizade et al. (2015) and other references cited there. Total number of simulations required for Sobol' computation is $(k+2) \times n$, where k is the number of parameters and n is the sample size. Consequently, in the current study, 33000 model simulations were performed for the SA (28 parameters and 1100 samples each).

The Sobol' indices can be computed directly based on the variance of the model outputs. However, due to the availability of observed data in our study, the analysis was performed based on the root mean square error (RMSE) between observation and simulation values. In addition, the analysis is conducted based on a running window to diagnose possible temporal behavior in the parameters response. Generally the window size depends upon the quality of input and observed data as well as the length of the period over which the parameters are influential. In view of this, the size of windows for each dataset varies and was determined as 5, 35 and 47 days for evapotranspiration, water content, and discharge, respectively. The application of these timescales explores the model responses at a moderate temporal resolution without a focus on exceptional events or periods. The computations were performed within the R environment, by using the sensitivity package (Pujol et al., 2015).

203 Table 1. Parameters of the model

Param. No.	Anotation	Description	Lower Limit	Upper limit	Units
1	k_{sm}	Saturated hydraulic conductivity of macropores	42	1150	cm^{-1}
2	w_f	Percentage of macropores existence	0.01	0.2	(-)
3	α_w	Transfer coefficient of matrix flow with macropores	13×10^{-6}	24×10^{-2}	d^{-2}
4	α	parameter of water retention curve for macropores	0.01	0.15	cm^{-1}
5	n	parameter of water retention curve for macropores	1.5	4	(-)
6	ϕ	Porosity of macropores	0.35	0.6	(-)
7	k_{s1}	Saturated hydraulic conductivity of matrix for layer 1	0.5	10	cmd^{-1}
8	k_{s2}	Saturated hydraulic conductivity of matrix for layer 2	0.5	10	cmd^{-1}
9	k_{s3}	Saturated hydraulic conductivity of matrix for layer 3	0.5	10	cmd^{-1}
10	k_{s4}	Saturated hydraulic conductivity of matrix for layer 4	0.5	10	cmd^{-1}
11	θ_{s1}	Porosity of matrix for layer 1	0.35	0.6	(-)
12	θ_{s2}	Porosity of matrix for layer 2	0.35	0.6	(-)
13	θ_{s3}	Porosity of matrix for layer 3	0.35	0.6	(-)
14	θ_{s4}	Porosity of matrix for layer 4	0.35	0.6	(-)
15	α_1	Parameter of water retention curve for layer 1	0.001	0.07	cm^{-1}
16	α_2	Parameter of water retention curve for layer 2	0.001	0.07	cm^{-1}
17	α_3	Parameter of water retention curve for layer 3	0.001	0.07	cm^{-1}
18	α_4	Parameter of water retention curve for layer 4	0.001	0.07	cm^{-1}
19	n_1	Parameter of water retention curve for layer 1	1.2	2	(-)
20	n_2	Parameter of water retention curve for layer 2	1.2	2	(-)
21	n_3	Parameter of water retention curve for layer 3	1.2	2	(-)
22	n_4	Parameter of water retention curve for layer 4	1.2	2	(-)
23	C_1	Fitting parameter of transpiration	0.01	1.0	(-)
24	C_2	Fitting parameter of transpiration	0.01	0.5	(-)
25	C_3	Fitting parameter of transpiration	1	30	cmd^{-1}
26	Fc	Saturation at field capacity	0.3	0.7	(-)
27	rd	Root depth of vegetation	1	100	cm
28	ed	Evaporation depth	1	100	cm

204

205 It is worth noting that sample-based methods of sensitivity analysis are all prone to uncertainty due to the
206 limitation in the number of samples (Pappenberger et al., 2006). This issue may lower the reliability of SA
207 results and therefore needs to be addressed carefully before making any interpretations. In order to analyze the
208 adequacy of the sample size, we increased the number of samples for each parameter by increments of 50.
209 Bootstrapping (Rindskopf, 1997) was applied at each increment to compute the 95% confidence intervals. Figure
210 1a-b indicate the convergence analyses of the main (top left panel) and total effects (top right panel) of
211 parameter 3 (α_w) to the RMSE of lysimeter discharge for a specific day in addition to the confidence bounds at
212 each increment. The results show that despite the relatively low number of samples, the indices converge to a
213 certain value and the confidence intervals obtained with the maximum sample size are relatively low (0.13 and
214 0.09 for the main and total effect, respectively). Similar results apply to the other parameters of the model. In

order to sum up these results (28 parameters for three outputs) the confidence intervals of the sensitivity indices using random sub-sets of the maximum sample size (1100) were used for the overall assessment of the analysis (Sarrazin et al., 2016). Figure 1c-d show the box plots of variations in confidence bounds for the main (bottom left panel) and total effects (bottom right panel). The plots show all the confidence intervals are below the value of 0.14. Overall, the values are relatively low and indices are considered reliable for the purpose of the present study.

Figure 1. Convergence of the main (a) and total (b) effects for parameter 3 (α_w) with increasing sample size and 95% confidence bounds to running RMSE of lysimeter discharge. Boxplots of the 95% confidence intervals of main effects (c) and total effects (d) for the maximum sample size (1100) to different outputs. Dis, ET and WC represent discharge, evapotranspiration and water content, respectively.

2.4 Temporal identifiability analysis

Temporal identifiability analysis (IA) was performed based on the DYNIA approach (Wagener et al., 2003). The method requires Monte-Carlo simulations of the model. We used the same simulations and running windows, which were employed for SA, to calculate the RMSE of daily observed versus simulated values. At each time step, the best top 10% of simulations were selected. As a result, the 90% confidence intervals of the parameters,

which lead to the 10% selected simulations were calculated based on their cumulative probability distribution. Where the parameters have narrower and non-uniform probability distributions, they have higher identifiability. We further transformed the distributions into the so-called information content. This is calculated as one minus the normalized width of the 90% confidence interval with respect to the original parameter range at each time step and its formulation follows as:

$$IC(t) = 1 - \frac{P_{95\%} - P_{5\%}}{\Delta p} \quad (6)$$

with $P_{95\%}$ and $P_{5\%}$, respectively, the lower and upper confidence interval of the obtained marginal parameter distribution and Δp the initial parameter range (upper-lower limit) as given in Table 1. Information content varies in the range of zero to one and therefore facilitates the comparison of results obtained from the sensitivity analysis. IA was performed in the Matlab environment based on the SAFE toolbox (Pianosi et al., 2015).

3 Results

3.1 Global sensitivity analysis

Sensitivity analyses results are shown via level plots (Figures 2, 3 and 4). The values on the figures vary from zero to one and the intensity of the shading at each time step shows the magnitude of the sensitivity index of the corresponding parameter in representing the RMSE values. Figure 2 indicates the main and interaction effects (i.e., total effect minus main effect) of the model parameters on discharge. To analyze the temporal change in sensitivity indices versus the change in the status of the system, the water content in the lysimeter is also shown (Figure 2a). For the comparison, water content was smoothed based on a moving average of 47 days, equal to the window size used for computing the sensitivity indices for the discharge. Results show that parameter 3 (α_w), which represents the preferential flow has the highest main effect, in comparison to other parameters (Figure 2b). This finding is in line with Ghasemizade et al. (2015) who found, based on the comparison of three model concepts, the necessity of including a preferential flow component for daily simulation of seepage (lysimeter discharge). Figure 2b also indicates that this parameter has the highest main effect on the lysimeter discharge at average water content values (November to May) and drier conditions (July and August). This point suggests that soil moisture has an influence on the contribution of preferential flow to discharge over different hydrologic conditions and corroborates the findings of Graham and Lin (2011) who found that initial soil moisture had a clear influence on preferential flow initiation.

Figure 2. Time-varying daily sensitivity of the discharge RMSE metric to model parameters within the simulation period, i.e., July 2012 to November 2013; (a) smoothed average water content with a window of 47 days. The dotted line indicates the average of the water content for the hydrologic year of 2012-2013; (b) main effect of model parameters; (c) interaction effect (i.e., total effect minus main effect) of model parameters; PF, SM and ET show groups of preferential flow, soil matrix and evapotranspiration parameters, respectively.

Figure 2c implies a correlation between the interacting effects of parameters 3 and 4 (α_w, α), representing preferential flow and the smoothed variation of water content, respectively, since the interaction effects drop to their lowest values when water content declines to its minimum (July and August) and retain their highest values when water content rises above the average value (March to the beginning of June and September to October). The present results also suggest the difficulties to calibrate the preferential flow parameters (α_w, α) at near saturation values due to the high interaction effects of preferential flow parameters with the other parameters.

This finding is also supported by the reported correlations of dual permeability parameters based on Bayesian inverse modeling (Laloy et al., 2010) and Bayesian uncertainty analysis of dual permeability parameters (Arora et al., 2012).

Parameters 10 (k_{s4}) and 18 (α_4) and 22 (n_4), which represent the matrix flow in the bottom layer (layer 4), are the second most influential parameters of the model in terms of their main and interaction contributions to the variance of the objective function. Similar to the parameters discussed before, this means that the bottom soil layer parameters i.e., 10, 18 and 22 play an important role in explaining the variance of the model but a correct calibration could be difficult due to their interactions with the other parameters. Another apparent issue in Figure 2 is that in contrary to the minimal main effects of the parameters during wetter conditions (water contents higher than 0.5), interaction effects increase to their maximum values during these conditions. However, not all the parameters show a clear dependency on soil moisture suggesting that additional environmental conditions (e.g., evapotranspiration rate) should be considered to explain this dynamic. The time dependency of parameter interactions as a function of change in hydrologic conditions highlights the value of time-varying sensitivity analysis from variable model behavior. We also note that care has to be taken when interpreting results of time steps at the beginning and the end of time series since the time windows used tends to decrease (Wagner et al., 2003).

Figure 3 indicates the results of the temporal sensitivity analysis of simulated daily ET versus observed values. It is not surprising that the parameters 23 and 24 (C_1 , C_2), which are the fitting parameters of ET calculations in HGS, contribute the most to the variance of the RMSE with their dominating main effects (Figure 3b). However, they show a different temporal dynamic for different water content conditions. The main effect of parameter 24 (C_2) becomes higher during average soil moisture conditions (i.e., close to the annual average of the water content), while the main effect of parameter 23 (C_1) becomes distinguished under relatively drier conditions (less than annual average water content) and in wetting periods. In addition, it is interesting to note that for the specific conditions of our simulations the parameter 25 (C_3) is not relevant. This parameter regulates the stress effect on the transpiration rate and this suggest that these conditions do not occur at the lysimeter.

The more important issue on the sensitivity of the ET values is noted when the interactions of the parameters are considered (Figure 3c). In this case, the results show that the majority of the parameters become highly interacting, specifically the parameters related to the characterization of the soil matrix. In addition, there is not a

clear pattern to show the time/moisture dependent interaction of the parameters suggesting how other environmental conditions (e.g., evapotranspiration rate) should be considered to explain this dynamic.

Figure 3. Time-varying daily sensitivity of the evapotranspiration RMSE metric to model parameters within the simulation period, i.e., July 2012 to November 2013; (a) smoothed average water content with a window of 5 days. The dotted line indicates the average of water content for the hydrologic year of 2012-2013; (b) main effect of model parameters; (c) interaction effect (i.e., total effect minus main effect) of model parameters; PF, SM and ET show groups of preferential flow, soil matrix and evapotranspiration parameters, respectively.

The temporal sensitivity of the model parameters of simulated water content versus observed values with a window size of 35 days are displayed in Figure 4. It is clearly visible in Figure 4b that parameter 10 (k_{s4}) has the largest individual contribution to the variance of the water content. In the second place, porosity values of layers 1 to 4 i.e., parameters 11, 12, 13 and 14 ($\theta_{s1}, \theta_{s2}, \theta_{s3}, \theta_{s4}$) have the highest contributions to the variation

of water content. Moreover, the temporal dynamics of the indices is relatively low and not well correlated to the soil moisture dynamics. For these reasons, the results suggest the relative importance of the parameters in the whole simulation period. Considering the interacting effect, it is noteworthy that the interactions of these four parameters ($\theta_{S1}, \theta_{S2}, \theta_{S3}, \theta_{S4}$) with other parameters are very close to zero (Figure 4c), implying their independent and crucial role in affecting the water content. It is also clear in Figure 4c that similar to the other investigated outputs, i.e., discharge and ET, parameters 4, 9, 10, 15, 16, 17, 18 and 22 ($\alpha, k_{S3}, k_{S4}, \alpha_1, \alpha_2, \alpha_3, \alpha_4, n_4$) are affecting the variation of water content mainly through their interactions with other parameters. This issue can impair identification of these parameters and therefore necessitates the inclusion of other datasets, such as matric potential measurements, for model calibration (Schelle et al., 2012).

Figure 4. Time-varying daily sensitivity of the water content RMSE metric to model parameters within the simulation period, i.e., July 2012 to November 2013; (a) smoothed average water content with a window of 35 days. The dotted line indicates the average of water content for the hydrologic year of 2012-2013; (b) main effect of model parameters;

(c) interaction effect (i.e., total effect minus main effect) of model parameters; PF, SM and ET show groups of preferential flow, soil matrix and evapotranspiration parameters, respectively.

3.2 Temporal identifiability analysis

The information content of all the 28 parameters, based on the different studied outputs, is shown in Figure 5. Similar to the results presented for the sensitivity analysis, the values vary from zero to one and the intensity of the shading at each time step shows the magnitude of the information content of the corresponding parameter in representing the observed values. For comparison, also the smoothed water content values (averaged every 10 days to show the trend of variation) are shown on the top (Figure 5a).

Figure 5. (a) Smoothed average water content with a window of 10 days for the simulation period, i.e., July 2012 to November 2013 in addition to the annual average of the water content (the dotted line) for the hydrologic year of

2012-2013. Time-varying daily information content of the (b) discharge RMSE, (c) evapotranspiration RMSE and (d) water content RMSE metric for model parameters. PF, SM and ET show groups of preferential flow, soil matrix and evapotranspiration parameters, respectively.

Figure 5b indicates the variable information content of the model parameters versus the lysimeter discharge. The results show a decrease of the information content of parameter 3 (α_w) when the soil is relatively dry. This suggests that soil moisture has a controlling effect on the identifiability of parameter 3 (α_w), which represents the preferential flow. In particular, the information content of parameter 3 (α_w) reaches its maximum under normal conditions (close to annual average water content), decreases to lower values as the water content exceeds the annual average of the water content and finally reaches to its minimum value (near zero) in dry conditions. The results suggest a threshold-dependent behavior, implying that when the water content reaches near saturation, the interaction between matrix and macropores decreases. The smaller values of parameter 3 (α_w), which explicitly represents the mass exchange between the matrix and macropores, within the narrow confidence bounds in near saturation conditions supports our argument.

The identifiability analysis of model parameters versus evapotranspiration dataset is shown in Figure 5c. It is discernable that parameters 23 and 24 (C_1 , C_2) have the highest variability in information content in comparison to the other parameters. We observe that the information content for these two parameters stays up during wetting up periods and goes down in drier conditions. It is also clear that parameter 3 (α_w), similar to its identifiability versus discharge, has higher identifiability during wetting up periods. Information content of model parameters versus water content (Figure 5d) is slightly different than the other information content plots. While porosity values of matrix flow in layers 1 to 4 (θ_{s1} , θ_{s2} , θ_{s3} , θ_{s4}) and coefficient of exchange between matrix and macrospores (α_w) have the highest information content values in comparison to other parameters, their values do not change much over the entire simulation time. It can be interpreted that the average water content may not provide detailed information on the dynamics of the system and on when the maximum amount of information can be extracted from the data. In fact, it gives an average identifiability of the parameters. It can also be said that using water content data leads to much better identification of porosity values of the top layer where the maximum fluctuation in water content happens. Overall, the comparison of Figure 5b to 5d reveals the fact that relatively few parameters are identifiable in comparison to the total number of parameters i.e., 28.

3.3 Relationship between sensitivity and identifiability analyses

The analyses presented so far show the complex behavior of the model parameters in comparison to the different outputs investigated. The sensitivity analysis (Figures 2, 3 and 4) shows that the sensitive parameters with high

main effects swap when different outputs (discharge, evapotranspiration, water content) are analyzed. For example, when the sensitivity of parameters to discharge are analyzed, parameter 3 holds the highest main effect. However, it is parameters 23 and 24 that retain the highest main effects, where evapotranspiration is analyzed. This seems to contradict the results obtained with the identifiability analysis where, independent of the outputs, almost the same parameters i.e., 3, 10-14, 23 and 24, were found to have relatively high information content (Figure 5b-d).

To further explore this behavior, sensitivity indices and information content for all the considered outputs obtained at each time step are plotted in Figure 6. Figure 6a shows main effects versus information content (identifiability values). We observe that the values are mainly concentrated below the 1:1 line. The spread of the values suggests that high identifiability does not necessarily guarantee a reduction in output uncertainty, as there are some parameters with high identifiability (greater than 0.8) and almost zero main effects. On the other hand, we observe in Figure 6a that high sensitive parameters (greater than 0.1) do not show a clear relationship with their identifiability counterparts. Therefore, it can be interpreted from Figure 6a that SA (the main effect) cannot replace IA.

Figure 6. Information content of discharge, evapotranspiration and water content versus (a) main effects (b) interaction effects for all the model parameters for the entire simulation period, i.e., July 2012 to November 2013.

Figure 6b indicates the relationship between information content and interaction effect (i.e., total effect – main effect) of the model parameters at all the time steps for all the model outputs. It is noteworthy that an inverse type of relationship, in comparison to the one discussed above, is apparent meaning that when the interaction is high, the information content is generally low. Thus, a parameter that shows high interactions (more than 0.8) with other parameters in explaining the variance of the model output is likely unidentifiable. On the other hand,

when the interaction is low, the parameters do not have a clear correlation with the information content (i.e., they can be either identifiable or unidentifiable). In that sense, the results show how low interaction is a necessary but not sufficient condition for a parameter to be identifiable.

Overall, the comparison between sensitivity and identifiability analysis shows that the two analyses are not interchangeable. The difference between SA and IA can be attributed to the different parameter space analyzed for each individual analysis. While global sensitivity analysis explores the entire input space, identifiability analysis filters the input space based on additional information (selecting the best parameter sets based on an objective function, here RMSE). In this context, it is worth noting that while SA helps distinguishing the important parameters, it does not guarantee whether the parameters can be identified via the available observations and therefore improve the performance of the model. On the contrary, identifiability analysis indicates the capability to find the optimum (constrained) parameter values based on an objective function but it does not guarantee the identification of the important (or not important) parameters based on the available information. For these reasons, the results of this study highlight how the use of combined diagnostic tools can be a promising approach for better understanding of the input-output space of complex environmental models (Bouda et al., 2014; Pappenberger et al., 2008).

4 Discussion

Based on the analyses conducted, it is interesting to note that only few parameters (e.g., α_w , C_I , C_2) meet all the requirements (i.e., high main effects, relatively low interactions and high identifiability) for the reduction of the output uncertainty. Besides, these requirements occur only under some specific conditions (i.e., water content conditions). Moreover, also relatively few parameters (k_s , w_f , φ , θ_{S1} , θ_{S2} , θ_{S3} , θ_{S4} , $C3$) have in general low interaction effects, suggesting that most of the parameters are important for the simulation of the processes considered. For these reasons, the study shows, based on the available data at the lysimeter scale, how complex and integrated models, such as HGS, are attractive solutions to reproduce complex behavior of the environmental system but they have the severe difficulties of parametrization, leading to their reduced predictive capabilities.

Nonetheless, the combined application of diagnostic tools suggests several possible strategies that can be considered for further improvements in the application of complex models. The results indicate for example that different parameters are sensitive to different outputs. This issue suggests the necessity of including multiple datasets for parameter calibration. Similarly, different objective functions could be integrated. In this study, RMSE was used as the only metric since the focus of our study was to highlight the relationship between

identifiability and sensitivity analyses. Other performance metrics could be used to explore specific region of the model performance (Gupta et al., 2008; Wagener et al., 2009; Zhang et al., 2013). Moreover, the time dependency of parameter interactions as a function of change in hydrologic conditions highlights the value of time-varying sensitivity analysis from variable model behavior.

In the present study, while the length of running-time windows varied with regard to the considered outputs, all the parameters had the same window size for each analyzed output. Massmann et al. (2014) used different time scales for individual parameters of a rainfall-runoff model in order to analyze the temporal parameter sensitivities. They found out that the window sizes over which the time-varying sensitivity is best integrated generally varies for different parameters. The fact that there is no optimal window size has some practical implications when planning measurement campaigns since the length of the measurements needs to be sufficient for being able to identify the processes and parameters of interest.

Finally, the temporal dynamics of the sensitivity indices (main and interaction effects) as well as the temporal information content of the parameters were compared in this study to the variation in soil moisture. The indices for several parameters show a good correlation with the soil moisture. However, this was not always the case indicating that additional and interacting factors could control the behavior of each parameter (i.e., evapotranspiration rate).

5 Summary and conclusion

This paper investigates the relationship between temporal sensitivity and identifiability of model parameters and explores the contribution of each analysis in providing a framework for reducing the dimensionality of model parameters as well as the reduction in output uncertainty. For this purpose, we applied the HydroGeoSphere model to represent a complex and nonlinear system in which the links between formulation and behavior are difficult to discern a priori. Our system comprised a weighing lysimeter with high-quality data which were used to analyze the bottom discharge, evapotranspiration and average-profile water content in the system. We analyzed the temporal sensitivity of model parameters versus each output, using the Sobol' approach (Sobol', 2001). Results indicate that the sensitive parameters and the values of sensitivity indices change significantly due to changes in hydrologic conditions, indicated by the variation in water content, and the outputs considered. In contrast to temporal sensitivity results, identifiability analysis, conducted based on the DYNIA approach (Wagener et al., 2003), indicated that while the identifiability values of model parameters vary notably over time, the likely identifiable parameters remain the same, independent of what output is considered. Finally,

analyzing sensitivity indices versus identifiability values provides a more general understanding of their relationship. Firstly, it is shown how both sensitivity indices should be considered to identify the relevant parameters that can be further used to reduce the uncertainty in the model output. In the second place, we showed how identifiability of the parameters is a necessary but not a sufficient condition for a parameter in order to reduce the uncertainty in the model output. Thirdly, the temporal analyses showed that for many time periods, there were not simultaneous significant sensitivities and high information content to adjust model parameters to more (high) certain values. For this reason, our analyses suggest the difficulties to parametrize this type of model based on these settings. Further improvements could be considered based on different objective function (instead of only RMSE) and different window sizes. Overall, our study indicated the use of complementary temporal diagnostic tools as a useful approach for a better understanding of the behavior and the applicability of complex environmental models.

Acknowledgements

This study was partially financed by the Competence Center Environment and Sustainability (CCES) of the ETH domain in the framework of the RECORD Catchment (Coupled Ecological, Hydrological and Social Dynamics in Restored and Channelized Corridors of a River at the Catchment Scale) project. In addition, this work was supported by the Swiss National Science Foundation (SNF grant 200021-129735 “Alpine Hydrogeology”). We would like to thank Sonia Senevirante, Martin Hirschi and Dominik Michel (Land-Climate Dynamics group, Institute for Atmospheric and Climate Science IAC, ETH Zurich) for providing the data and the discussions over the lysimeter measurements. We are also grateful for the comments of three anonymous reviewers.

Appendix A

HydroGeoSphere (HGS) is a three-dimensional surface-subsurface physically-based numerical model (Therrien et al., 2010). The dual-permeability simulation approach that was employed for preferential flow simulation in HGS is formulated as follows:

$$\frac{\partial \theta_f}{\partial t} = \frac{\partial}{\partial z} \left(k_f \frac{\partial h_f}{\partial z} - k_f \right) - \frac{\Gamma_w}{w_f} \quad (\text{A.1})$$

$$\frac{\partial \theta_m}{\partial t} = \frac{\partial}{\partial z} \left(k_m \frac{\partial h_m}{\partial z} - k_m \right) + \frac{\Gamma_w}{1 - w_f} \quad (\text{A.2})$$

where subscript m denotes matrix and subscript f denotes fracture (macropores). Also, θ is the volumetric water content, w_f is the ratio of volume of macro-pores to the total pore systems, k is hydraulic conductivity, h is the

hydraulic head and Γ_w is the water exchange between the two pore systems and is formulated based on a first-order exchange term in the following form :

$$\Gamma_w = \alpha_w (h_f - h_m) \quad (\text{A.3})$$

in which α_w is a first-order mass transfer coefficient. This term can be described for well-defined geometries of the pores after the approach of Gerke and van Genuchten (1993) given by :

$$\alpha_w = \frac{\beta}{a^2} \gamma \times (k_{sa}) \quad (\text{A.4})$$

where a is the characteristic length of the soil aggregate, β is a geometry dependent shape factor, γ is an empirical coefficient and k_{sa} is the saturated hydraulic conductivity of the fracture/matrix interface. The soil water retention and hydraulic conductivity functions are described as in Mualem (1976) and van Genuchten (1980) as follows:

$$S_w = S_{wr} + (1 - S_{wr}) \left[1 + |\alpha h|^n \right]^{-m} \quad (\text{A.5})$$

$$K_r = S_e^{\ell_p} \left[1 - \left(1 - S_e^{\frac{1}{m}} \right)^m \right]^2 \quad (\text{A.6})$$

$$m = 1 - \frac{1}{n} \quad (\text{A.7})$$

$$S_e = \frac{(S_w - S_{wr})}{(1 - S_{wr})} \quad (\text{A.8})$$

$$S_w = \frac{\theta}{\theta_s} \quad (\text{A.9})$$

Where α , n and ℓ_p (was kept constant equal to 0.5) are fitting parameters, S_e is effective saturation, S_{wr} is residual water saturation, S_w is the water saturation, θ_s is the saturated water content and K_r is the relative permeability.

HGS calculates interception as follows: (A.10)

$$I_{\max} = C_{\text{int}} LAI$$

where LAI is the leaf area index and C_{int} is the canopy storage parameter. For each time increment Δt , the actual interception storage (S_{int}) is calculated as follows.

$$S_{\text{int}}^* = \min(S_{\text{int}}^{\max}, S_{\text{int}}^0 + P_p \Delta t) \quad (\text{A.11})$$

$$E_{can}\Delta t = \min(S_{int}^*, ET_p \Delta t) \quad (A.12)$$

$$S_{int} = S_{int}^* - E_{can}\Delta t \quad (A.13)$$

where S_{int}^0 and S_{int}^* are the previous time and intermediate values of S_{int} . The canopy evaporation is E_{can} and ET_p is the potential evapotranspiration. Evapotranspiration is estimated in HGS as the combination of plant transpiration and of evaporation from both surface and subsurface domains. Transpiration occurs from the root zone and is calculated with the following relationship:

$$T_p = f_1(LAI) * f_2(\theta) * RDF * [ET_p - E_{can}] \quad (A.14)$$

f_1 is a vegetation function which relates transpiration to leaf area index as follows:

$$f_1(LAI) = \max\{0, \min[1, (c_2 + c_1 LAI)]\} \quad (A.15)$$

C_1 and C_2 are dimensionless fitting parameters. The features of vegetation root are summarized in RDF which is a root distribution function and is described as:

$$RDF = \frac{\int_{z_1}^{z_2} rf(z) dz}{\int_0^{L_r} rf(z) dz} \quad (A.16)$$

In equation (A.16) $rf(z)$ is the root extraction function, L_r is the effective root length and z is the depth coordinate from soil surface. The moisture content function (f_2) in equation (A.14) relates transpiration to the soil moisture and is expressed as:

$$\left[\begin{array}{l} 0 \text{ if } 0 \leq \theta \leq \theta_{wp} \\ f_3 \text{ if } \theta_{wp} \leq \theta \leq \theta_{fc} \\ 1 \text{ if } \theta_{fc} \leq \theta \leq \theta_o \\ f_4 \text{ if } \theta_o \leq \theta \leq \theta_{an} \\ 0 \text{ if } \theta_{an} \leq \theta \end{array} \right] \quad (A.17)$$

$$f_3 = 1 - \left[\frac{\theta_{fc} - \theta}{\theta_{fc} - \theta_{wp}} \right]^{c_3} \quad (A.18)$$

$$f_4 = 1 - \left[\frac{\theta_{an} - \theta}{\theta_{an} - \theta_0} \right]^{C_3} \quad (A.19)$$

Where C_3 is a fitting parameter, θ_{wp} is the moisture content at wilting point, θ_{fc} is the moisture content at field capacity, θ_0 is the moisture content at the oxic limit and θ_{an} is the soil moisture content at the anoxic limit. The logic supporting equation (A.17) is that when the soil moisture is below wilting point, transpiration is zero; transpiration then increases to a maximum when moisture content reaches field capacity. This level of transpiration is maintained until soil moisture increases to the oxic limit. From the oxic limit on, transpiration decreases until soil moisture hits the anoxic limit where transpiration equals zero. In fact, when soil moisture exceeds the anoxic limit, the roots become inactive due to lack of aeration.

Evaporation occurs along with transpiration, resulting from the energy that penetrates in the vegetation cover and is described as follows:

$$E_s = \alpha^* (E_p - E_{can}) [1 - f_1(LAI)] EDF \quad (A.20)$$

where EDF is the evaporation density function and is assumed to be effective from the soil surface to a prescribed extinction depth. Similar to transpiration, evaporation from the soil surface and subsurface soil is dependent on the availability of moisture content. α^* in equation (A.20) is a moisture dependent coefficient which varies between zero and one and is expressed as:

$$\left[\begin{array}{l} 0 \text{ if } \theta < \theta_{e2} \\ \frac{\theta - \theta_{e2}}{\theta_{e1} - \theta_{e2}} \text{ if } \theta_{e2} \leq \theta \leq \theta_{e1} \\ 1 \text{ if } \theta > \theta_{e1} \end{array} \right] \quad (A.21)$$

where θ_{e1} is the moisture content at the end of the energy-limiting stage (above which full evaporation can occur) and θ_{e2} is the limiting moisture content below which evaporation is zero.

References

- Abebe, N.A., Ogden, F.L., Pradhan, N.R., 2010. Sensitivity and uncertainty analysis of the conceptual HBV rainfall-runoff model: Implications for parameter estimation. *J Hydrol*, 389(3-4): 301-310. DOI:10.1016/j.jhydrol.2010.06.007
- Arora, B., Mohanty, B.P., McGuire, J.T., 2012. Uncertainty in dual permeability model parameters for structured soils. *Water Resour Res*, 48. DOI:10.1029/2011wr010500
- Bastidas, L.A., Gupta, H.V., Sorooshian, S., Shuttleworth, W.J., Yang, Z.L., 1999. Sensitivity analysis of a land surface scheme using multicriteria methods. *J Geophys Res-Atmos*, 104(D16): 19481-19490. DOI:10.1029/1999jd900155
- Beven, K., 2006. A manifesto for the equifinality thesis. *J Hydrol*, 320(1-2): 18-36. DOI:10.1016/j.jhydrol.2005.07.007
- Bouda, M. et al., 2014. Implementation of an automatic calibration procedure for HYDROTEL based on prior OAT sensitivity and complementary identifiability analysis. *Hydrological Processes*, 28(12): 3947-3961. DOI:10.1002/hyp.9882
- Brun, R., Reichert, P., Kunsch, H.R., 2001. Practical identifiability analysis of large environmental simulation models. *Water Resour Res*, 37(4): 1015-1030. DOI:10.1029/2000wr900350
- Burhenne, S., Jacob, D., Henze, G., 2011. SAMPLING BASED ON SOBOLO SEQUENCES FOR MONTE CARLO TECHNIQUES APPLIED TO BUILDING SIMULATIONS, 12th Conference of International Building Performance Simulation Association, Sydney.
- Cibin, R., Sudheer, K.P., Chaubey, I., 2010. Sensitivity and identifiability of stream flow generation parameters of the SWAT model. *Hydrological Processes*, 24(9): 1133-1148. DOI:10.1002/hyp.7568
- De Lange, W.J. et al., 2014. An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument. *Environ Modell Softw*, 59: 98-108. DOI:10.1016/j.envsoft.2014.05.009
- Doherty, J., Hunt, R.J., 2009. Two statistics for evaluating parameter identifiability and error reduction. *J Hydrol*, 366(1-4): 119-127. DOI:10.1016/j.jhydrol.2008.12.018
- Franchini, M., Wendling, J., Obled, C., Todini, E., 1996. Physical interpretation and sensitivity analysis of the TOPMODEL. *J Hydrol*, 175(1-4): 293-338. DOI:10.1016/S0022-1694(96)80015-1
- Garambois, P.A., Roux, H., Larnier, K., Castaings, W., Dartus, D., 2013. Characterization of process-oriented hydrologic model behavior with temporal sensitivity analysis for flash floods in Mediterranean catchments. *Hydrol Earth Syst Sc*, 17(6): 2305-2322. DOI:10.5194/hess-17-2305-2013
- Gerke, H.H., van Genuchten, M.T., 1993. A Dual-Porosity Model for Simulating the Preferential Movement of Water and Solutes in Structured Porous-Media. *Water Resour Res*, 29(2): 305-319. DOI:10.1029/92wr02339
- Ghasemizade, M., Moeck, C., Schirmer, M., 2015. The effect of model complexity in simulating unsaturated zone flow processes on recharge estimation at varying time scales. *J Hydrol*, 529: 1173-1184. DOI:10.1016/j.jhydrol.2015.09.027
- Graham, C.B., Lin, H.S., 2011. Controls and Frequency of Preferential Flow Occurrence: A 175-Event Analysis. *Vadose Zone J*, 10(3): 816-831. DOI:10.2136/vzj2010.0119
- Gupta, H.V., Wagener, T., Liu, Y.Q., 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes*, 22(18): 3802-3813. DOI:10.1002/hyp.6989
- Guse, B., Reusser, D.E., Fohrer, N., 2014. How to improve the representation of hydrological processes in SWAT for a lowland catchment - temporal analysis of parameter sensitivity and model performance. *Hydrological Processes*, 28(4): 2651-2670. DOI:10.1002/hyp.9777
- Hall, J.W., Tarantola, S., Bates, P.D., Horritt, M.S., 2005. Distributed sensitivity analysis of flood inundation model calibration. *J Hydraul Eng-Asce*, 131(2): 117-126. DOI:10.1061/(ASCE)0733-9429(2005)131:2(117)

- Hartmann, A. et al., 2013. Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resour Res*, 49(6): 3345-3358. DOI:10.1002/wrcr.20229
- Herman, J.D., Reed, P.M., Wagener, T., 2013. Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior. *Water Resour Res*, 49(3): 1400-1414. DOI:10.1002/wrcr.20124
- Kelleher, C. et al., 2013. Identifiability of transient storage model parameters along a mountain stream. *Water Resour Res*, 49(9): 5290-5306. DOI:10.1002/wrcr.20413
- Kristensen, K.J., Jensen, S.E., 1975. A model for estimating actual evapotranspiration from potential evapotranspiration. *Nord Hydrol*, 6: 170-188. DOI:10.2166/nh.1975.012
- Laloy, E., Weynants, M., Bieters, C.L., Vanclooster, M., Javaux, M., 2010. How efficient are one-dimensional models to reproduce the hydrodynamic behavior of structured soils subjected to multi-step outflow experiments? *J Hydrol*, 393(1-2): 37-52. DOI:10.1016/j.jhydrol.2010.02.017
- Li, Q. et al., 2008. Simulating the multi-seasonal response of a large-scale watershed with a 3D physically-based hydrologic model. *J Hydrol*, 357(3-4): 317-336. DOI:10.1016/j.jhydrol.2008.05.024
- Massmann, C., Holzmann, H., 2012. Analysis of the behavior of a rainfall-runoff model using three global sensitivity analysis methods evaluated at different temporal scales. *J Hydrol*, 475: 97-110. DOI:10.1016/j.jhydrol.2012.09.026
- Massmann, C., Wagener, T., Holzmann, H., 2014. A new approach to visualizing time-varying sensitivity indices for environmental model diagnostics across evaluation time-scales. *Environ Modell Softw*, 51: 190-194. DOI:10.1016/j.envsoft.2013.09.033
- Matott, L.S., Babendreier, J.E., Purucker, S.T., 2009. Evaluating uncertainty in integrated environmental models: A review of concepts and tools. *Water Resour Res*, 45. DOI:10.1029/2008wr007301
- Menzel, L., Demuth, N., 1993. Tracerhydrologische Untersuchungen am Lysimeter Rietholzbach [in German], Berichte und Skripten, Rep. 52,, ETH Zürich, 27 pp.
- Mualem, Y., 1976. New Model for Predicting Hydraulic Conductivity of Unsaturated Porous-Media. *Water Resour Res*, 12(3): 513-522. DOI:10.1029/Wr012i003p00513
- Nossent, J., Elsen, P., Bauwens, W., 2011. Sobol' sensitivity analysis of a complex environmental model. *Environ Modell Softw*, 26(12): 1515-1525. DOI:10.1016/j.envsoft.2011.08.010
- Pappenberger, F., Beven, K.J., Ratto, M., Matgen, P., 2008. Multi-method global sensitivity analysis of flood inundation models. *Adv Water Resour*, 31(1): 1-14. DOI:10.1016/j.advwatres.2007.04.009
- Pappenberger, F., Iorgulescu, I., Beven, K.J., 2006. Sensitivity analysis based on regional splits and regression trees (SARS-RT). *Environ Modell Softw*, 21(7): 976-990. DOI:10.1016/j.envsoft.2005.04.010
- Penman, H.L., 1948. Natural Evaporation from Open Water, Bare Soil and Grass. *Proc R Soc Lon Ser-A*, 193(1032): 120-&. DOI:DOI 10.1098/rspa.1948.0037
- Pianosi, F., Sarrazin, F., Wagener, T., 2015. A Matlab toolbox for Global Sensitivity Analysis. *Environ Modell Softw*, 70: 80-85. DOI:10.1016/j.envsoft.2015.04.009
- Pujol, G., Iooss, B., Janon, A., 2015. sensitivity package.
- Reusser, D.E., Zehe, E., 2011. Inferring model structural deficits by analyzing temporal dynamics of model performance and parameter sensitivity. *Water Resour Res*, 47. DOI:10.1029/2010wr009946
- Rindskopf, D., 1997. An introduction to the bootstrap - Efron, B., Tibshirani, R.J. *J Educ Behav Stat*, 22(2): 245-245.
- Rozemeijer, J.C. et al., 2010. Integrated modeling of groundwater-surface water interactions in a tile-drained agricultural field: The importance of directly measured flow route contributions. *Water Resour Res*, 46. DOI:10.1029/2010wr009155
- Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. *Comput Phys Commun*, 145(2): 280-297. DOI:DOI 10.1016/S0010-4655(02)00280-1

- Saltelli, A. et al., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comput Phys Commun*, 181(2): 259-270. DOI:10.1016/j.cpc.2009.09.018
- Saltelli, A. et al., 2006. Sensitivity analysis practices: Strategies for model-based inference. *Reliab Eng Syst Safe*, 91(10-11): 1109-1125. DOI:10.1016/j.ress.2005.11.014
- Sarrazin, F., Pianosi, F., Wagener, T., 2016. Global Sensitivity Analysis of environmental models: Convergence and validation. *Environ Modell Softw*, 79: 135-152. DOI:10.1016/j.envsoft.2016.02.005
- Schelle, H., Iden, S.C., Fank, J., Durner, W., 2012. Inverse Estimation of Soil Hydraulic and Root Distribution Parameters from Lysimeter Data. *Vadose Zone J*, 11(4). DOI:10.2136/Vzj2011.0169
- Scurlock, J.M.O., Asner, G.P., Gower, S.T., 2001. Worldwide Historical Estimates of Leaf Area Index, 1932-2000. ORNL Technical Memorandum ORNL/TM-2001/268. Oak Ridge National Laboratory, Oak Ridge, Tenn.
- Seneviratne, S.I. et al., 2012. Swiss prealpine Rietholzbach research catchment and lysimeter: 32 year time series and 2003 drought event. *Water Resour Res*, 48. DOI:10.1029/2011wr011749
- Shin, M.J., Guillaume, J.H.A., Croke, B.F.W., Jakeman, A.J., 2013. Addressing ten questions about conceptual rainfall-runoff models with global sensitivity analyses in R. *J Hydrol*, 503: 135-152. DOI:10.1016/j.jhydrol.2013.08.047
- Shin, M.J., Guillaume, J.H.A., Croke, B.F.W., Jakeman, A.J., 2015. A review of foundational methods for checking the structural identifiability of models: Results for rainfall-runoff. *J Hydrol*, 520: 1-16. DOI:10.1016/j.jhydrol.2014.11.040
- Sobol', I.M., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math Comput Simulat*, 55(1-3): 271-280. DOI:10.1016/S0378-4754(00)00270-6
- Tang, Y., Reed, P., Wagener, T., van Werkhoven, K., 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrol Earth Syst Sc*, 11(2): 793-817.
- Therrien, R., McLaren, R.G., Sudicky, E.A., Panday, S.M., 2010. HydroGeoSphere. A Three-dimensional Numerical Model Describing Fully-integrated Subsurface and Surface Flow and Solute Transport.
- van Genuchten, M.T., 1980. A Closed-Form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Sci Soc Am J*, 44(5): 892-898. DOI:10.2136/sssaj1980.03615995004400050002x
- van Werkhoven, K., Wagener, T., Reed, P., Tang, Y., 2009. Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. *Adv Water Resour*, 32(8): 1154-1169. DOI:10.1016/j.advwatres.2009.03.002
- Vitvar, T., Balderer, W., 1997. Estimation of mean water residence times and runoff generation by O-18 measurements in a pre-Alpine catchment (Rietholzbach, eastern Switzerland). *Appl Geochem*, 12(6): 787-796. DOI:10.1016/S0883-2927(97)00045-0
- Wagener, T., McIntyre, N., Lees, M.J., Wheater, H.S., Gupta, H.V., 2003. Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis. *Hydrological Processes*, 17(2): 455-476. DOI:10.1002/hyp.1135
- Wagener, T., van Werkhoven, K., Reed, P., Tang, Y., 2009. Multiobjective sensitivity analysis to understand the information content in streamflow observations for distributed watershed modeling. *Water Resour Res*, 45. DOI:10.1029/2008wr007347
- Zhang, C., Chu, J.G., Fu, G.T., 2013. Sobol's sensitivity analysis for a distributed hydrological model of Yichun River Basin, China. *J Hydrol*, 480: 58-68. DOI:10.1016/j.jhydrol.2012.12.005
- Zhu, Y., Shi, L.S., Lin, L., Yang, J.Z., Ye, M., 2012. A fully coupled numerical modeling for regional unsaturated-saturated water flow. *J Hydrol*, 475: 188-203. DOI:10.1016/j.jhydrol.2012.09.048