UNIVERSITÉ DE NEUCHÂTEL

CHYN
Centre d'hydrogéologie
et de géothermie

eawag
aquatic research

Ph.D. Thesis

MODEL COMPLEXITY AND DIAGNOSTIC-TOOL BASED ANALYSES OF INTEGRATED AND PHYSICALLY BASED MODELS

by

Mehdi Ghasemizade

presented to the Faculty of Science of the University of Neuchâtel to satisfy the requirements of the degree of Doctor of Philosophy in Science

Supervisory committee:

Prof. Dr. Mario Schirmer, University of Neuchâtel (Director of the thesis)
Prof. Dr. Daniel Hunkeler, University of Neuchâtel
Prof. Dr. Jan Seibert, University of Zürich
Prof. Dr. Marc Walther, Technische Universität Dresden, Germany

Defended:

12.08.2016

Faculté des sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel - Suisse
Tél: + 41 (0)32 718 2100
E-mail: secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Mehdi Ghasemizade

Titre:

"Model Complexity and Diagnostic-tool Based Analyses of Integrated and Physically Based Models"

sur le rapport des membres du jury composé comme suit:

- Prof. Mario Schirmer, directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Daniel Hunkeler, co-directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Jan Seibert, Université de Zürich, Suisse
- Prof. Marc Walther, Technische Universität Dresden, Allemagne

Neuchâtel, le 24 juin 2016

Le Doyen, Prof. B. Colbois

B. Colbois.

Acknowledgments

The completion of this Ph.D. thesis would not have been possible without the support of colleagues, friends and family. First and foremost, I thank my supervisor, *Mario Schirmer*. He gave me the opportunity to do my Ph.D. thesis in his group, supported me in all the situations, kept his office door always open and welcomed any discussion. I am grateful to *Daniel Hunkeler* for accepting the responsibility of being my co-director and providing the opportunity for me to do my Ph.D. at the CHYN in Neuchatel university. I also thank *Jan Seibert* and *Marc Walther* for their commitment as external examiners.

I gratefully acknowledge the funding source that made my Ph.D. work possible. I was awarded a full scholarship by the Iranian ministry of science and technology. I would also like to take this opportunity to thank Aquanty company and its president and CEO Steven Berg for his timely hints on the usage and updates of HydroGeoSphere code.

A big THANK YOU to the Hydrogeology Group and the department of water resources and drinking water in Eawag. You all contributed to this thesis in some way and created an enjoyable working environment at Eawag. I also thank Daniel Pellanda who did his best to provide the optimum computing facilities beside bringing humor and fun at the right times. I am also grateful to Karim Abbaspour for keeping his door open all the times for fruitful modeling discussions.

My time at Eawag was made enjoyable in large part due to the many outdoor and sport activities. I am grateful for the time spent playing in Empa/Eawag volleyball club with friendly people (Abdolla, Sarah, Franziska, Denise, Bahare, Rolf, Franz and my good friend Manouchehr), for our memorable trips into the mountains with Hamid, Omid, Zahra, Federica, Stefano, Romina and Alina, and lastly for Friday gatherings of the Iranian community.

Finally, I would like to thank my *family*, especially my *parents*, *my sister* and my *brother-in-law (Ehsan)* for all their love and encouragement. Mom, thanks a lot for all your worships and praying you did for me.

Abstract

The proper management of water resources nowadays is a critical issue. In that sense, accurate measurement of water balance components is a prerequisite for the proper management of water resources since one cannot manage what one cannot measure. Due to the difficulty in direct measurements of some of the water balance components such as deep percolation, simulation models are applied. Recent increases in computational power have motivated the application of more complex models of coupled environmental processes. These models, however, require outnumbered parameters, which lead to the problem of over-parameterization, meaning that many different parameter sets can lead to identical fits to the observed data. Therefore, this study explores the application of integrated and physically-based model HydroGeoSphere (HGS) in the framework of a weighing lysimeter in north-east of Switzerland to pursue: I) comparing the performance of different levels of complexity (in terms of the number of parameters) for simulating daily water balance components (actual evapotranspiration, water content, and lysimeter discharge) where three model concepts were introduced; II) addressing the output uncertainty of each concept at different time scales; III) application of a global and temporal sensitivity analysis as a diagnostic tool to address how individual parameters of the model as well as their interactions can affect the output uncertainty; VI) using a time-varying identifiability analysis method to investigate when the maximum amount of information about model parameters can be derived, considering the available data. The results of the study indicated that the most complex concept outperformed the other simpler concepts in reproducing the daily water balance components based on the performance metrics of R^2 and RMSE. However, the ideal required level of complexity, when considered in terms of output uncertainty, was shown to be dependent on the time scales of the simulated outputs. Exploring the results of the sensitivity analysis revealed that the individual effects of model parameters as well as their interaction effects on model outputs are required to be analyzed simultaneously to allow for the reduction in output uncertainty. The identifiability analysis indicated that identifiability is a necessary but not sufficient condition for a parameter to allow for reduction in the model output uncertainty. Overall our research indicated that, based on the available data at the lysimeter scale, complex and integrated models, such as HGS, are attractive solutions to reproduce complex features of the system but they have the severe difficulties of parametrization, leading to their reduced predictive capabilities.

Keywords: Physically based, HydroGeoSphere, Identifiability, DYNIA, Model complexity, Prediction uncertainty, Preferential flow, Matrix flow, Temporal sensitivity, SOBOL', Lysimeter , Recharge, Evapotranspiration, Water content, Rietholzbach

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

1.1 Introduction

The proper management of water resources nowadays is a critical issue, not only in lands which suffer from water scarcity but also in countries which are exposed to the risk of high flows. In that sense, accurate measurement of water balance components is a prerequisite for the proper management of water resources since one cannot manage what one cannot measure. Water balance estimates strengthen water management decision-making, by assessing and improving the validity of visions, scenarios and strategies. Therefore, water balance estimation is an important tool to assess the current status and trends in water resources availability in an area over a specific period of time. The water balance equation is a simplistic mass balance equation, in which the difference between inputs and outputs is equivalent to the change in storage of water in the system. The water balance for a given time interval depends upon existing system storage (soil water content) and fluxes from the sides, top (precipitation, runoff and evapotranspiration) and bottom (deep percolation) boundaries of the model domain. In the following, a brief review of the importance and the estimation of those components of a water balance equation which are difficult or cannot be directly measured is given.

Water storage has been shown to be a controlling factor in generating high flows and sustaining base flows in shallow groundwater areas such as headwater catchments. For example, Rinderer et al. (2015) found in a pre-alpine monitoring site in Switzerland that antecedent conditions (storage) were among controlling factors in the response time of groundwater, preceding the peak of streamflow. Penna et al. (2015) found in a similar study in a headwater catchment in the Italian Alps that independent from the geology and land cover settings, the jointly effect of storage in the unsaturated zone and precipitation amount played the dominant role in triggering the piezometric response and total catchment runoff. On the other hand, the influence of water storage in sustaining the base flows have been addressed by many (e.g., Blumstock et al., 2015; Hilberts et al., 2007; Matonse and Kroll, 2013). Tetzlaff and Soulsby (2008) used isotopic and hydro-chemical data and showed the significance of storage in pre-alpine headwater catchments on the quantity and quality of the base flow waters.

The change in water storage in a model domain is highly non-linear and has been shown to be highly interacting with other water balance components such as evapotranspiration (Bowling et al., 2003). Evapotranspiration is the sum of evaporation and plant transpiration from the Earth's land surface to the atmosphere. The common approach to estimate actual evapotranspiration (ETa) where direct measurement methods such as eddy covariance are not available is to calculate potential evapotranspiration (ETp). It is the amount of water that would be evaporated and transpired if there were sufficient water available. In that sense, ETp will be adjusted to derive ETa based on the available water and cropped plant. However, this method is uncertain due to the existence of over 50 methods to calculate ETp (Thompson et al., 2014). Vazquez and Feyen (2003) evaluated the effect of three different methods for estimating potential evapotranspiration on effective parameters and performance of the MIKE SHE-code. They found that model performances were comparable and the best model performance was obtained by using the higher ETp values. Bae et al. (2011) used three alternative semi-distributed models and different ETp methods to simulate climate change scenarios in central South Korea. Their results showed that the different ETp methods impacted runoff changes, with the magnitude of ETp-related differences varying between hydrological models and season.

Deep percolation (DP) or recharge, which is the amount of water that infiltrates into the ground, passes the root zone and finally reaches the water table, is the other component of the water budget. DP drives many of the hydrological processes (Bakker et al., 2013) and may provide benefits including: recharging the aquifers, delaying return flow to the streams, diluting contaminants from other sources such as septic tanks. Estimation of DP is often concomitant with uncertainty due to the fact that it is very difficult and costly to measure it directly (Lee et al., 2007). DP is usually estimated via indirect methods such as variations of river streamflow (Combalicer et al., 2008), fluctuation of the water table (Marechal et al., 2006), analytical soil water balance models (Rodriguez-Iturbe et al., 1999) and numerical modeling using Richards’ equation (Carrera-Hernandez et al., 2012).

1.2 Problem description

Due to the difficulty in direct measurements of water balance components, simulation models as useful tools are applied to simulate soil water balance processes (Soldevilla-Martinez et al., 2014; Stumpp and Maloszewski, 2010). With regard to the interaction of the water balance components and the fact that the components themselves are influenced by various factors such as heterogeneity, sub layering, preferential flow paths in addition to hydrodynamic parameters such as field capacity, unsaturated conductivity, antecedent moisture and pore connectivity (Augenstein et al., 2015; Morbidelli et al., 2011; Morbidelli et al., 2014), mechanistic and physically-based models seem to be promising for simulation purposes (e.g., Bolger et al., 2011; Lafond et al., 2014; Rahim et al., 2012). Such models incorporate the affecting factors on water balance components into one framework and help to better understand the involved processes and the time when specific processes become dominant (e.g., Ameli et al., 2015; Cornelissen et al., 2014; Frei et al., 2010). Preferential flow, for example, is one of the processes that the dominant controls on its initiation and its interaction with initial soil moisture are poorly understood (Merdun et al., 2008). In macroporous soils, higher antecedent soil moisture generally increases the depth to which macropore flow penetrates as well as increasing total percolated volume, (Granovsky et al., 1994; Jarvis, 2007). Graham and Lin (2011) analyzed 175 events and found out that initial soil moisture (storage status) was clearly a control on preferential flow initiation. In contrast, Merdun et al. (2008) reported that preferential flow was more evident when soil was initially dry compared to two wetter treatments. Shipitalo and Edwards (1996) also found the relative contribution of macropores to pesticide transport was greatest when the soil was dry and decreased as the soil became wetter. Nimmo (2012) reviewed preferential flow occurrences and observations in unsaturated conditions and suggested the need for models which do not imply wetter- faster concept. Therefore, one of the main objectives of this thesis is to apply a physically based model and investigate the initiation of preferential flow, its relevance to antecedent moisture condition and its effects on the estimation of evapotranspiration, deep percolation and storage change. Nonetheless, one should note that the outputs of mechanistic and physically-based models could be very uncertain due to the problem of non-uniqueness or parameter equifinality (Beven, 2006). Non-uniqueness happens when model parameters cannot be estimated uniquely and therefore different sets of parameters values lead to similar values of model performance criteria. There are a number of factors that cause the non-uniqueness problem, including the interactions and correlations among the parameters being optimized simultaneously, and the insufficient information content of experimental data used for calibration. Therefore, the other two objectives of this thesis are to; i) investigate different levels of complexity, implying the optimized number of parameters required to represent vadose zone processes and avoid equifinality ii) evaluate the worth of different observation datasets in constraining model parameters and therefore reducing their uncertainty.

In order to accomplish the objectives of this thesis, high quality data of water balance components are required. It goes without saying that an accurate estimation requires accurate measured data for validation. In that sense, weighing lysimeters are appropriate experimental facilities for accurate measurement of soil moisture change, evapotranspiration and deep percolation. Lysimeters have been widely used in hydrological and water balance studies due to:

Measurements: All measurements have limitations in accuracy and one can just expect the measured data to be the least-biased. Weighing lysimeters provide the opportunity to measure the water balance components such as actual evapotranspiration with high accuracy. In that sense, actual evapotranspiration based on lysimeters measurements have been employed as reference data to evaluate established methods and develop new formulations for estimating actual evapotranspiration (Kashyap and Panda, 2001; Liu and Luo, 2010). In the sense of recharge estimations, lysimeter data have been used as reference to validate recharge estimation methods (e.g., Soldevilla-Martinez et al., 2014; von Freyberg et al., 2015).

Scale: Lysimeters can reproduce field-like conditions. These conditions can be discussed in terms of the atmospheric boundary conditions and heterogeneity versus homogeneity among soil particles. It is well-accepted in the hydro(geo)logy community that the established formulations, such as Richard's equation for the simulation of flow in variably saturated medium fail at the field scales (Beven and Germann, 2013; Gerke and Kohne, 2004; Kohne et al., 2006). The reason is that such formulations have been developed under well-controlled boundary conditions in the lab and assume homogeneity among soil particles. Applications of such equations at the lysimeter scale help to understand the limitations of the established formulations under transient conditions and where heterogeneity among soil particles exist. Also, due to the rather small area of a lysimeter in comparison to a catchment, the spatial variation in the precipitation as the model input is negligible.

Wide applicability: The information about evapotranspiration and seepage values which can be derived from lysimeters have wide applicability for large-scale management of water. For example, Seneviratne et al. (2012) showed that the lysimeter seepage and catchment-wide discharge at monthly scale had a linear correlation of 0.91 based on thirty one years of recorded data. Lysimeters also increase the capability for the cross comparison of results between sites.

In spite of the above-mentioned advantages of using lysimeters in hydrological studies, one should also note the limitations of lysimeters for water balance and hydrological studies. For example, although unsaturated zone drainage from the weighing lysimeters provides the most direct measure of potential recharge, it does not incorporate spatial variability that is contained in watershed-wide estimates of net recharge. Also, the walls of lysimeter casings prevent the latter movement of water from or to the surrounding area. This issue becomes important in the periods of perched water tables. Last but not least is the measurement of runoff which may occur on the top of lysimeters. Usually lysimeters are built, including the one used in this research, with an edge on the top which does not let the collected water on the upper-lying areas to be diverted around the lysimeter.

1.3 Structure of the thesis

This PhD thesis consists of five chapters which address the research objectives stated above. The first chapter starts with introducing the importance of water budget estimation in managing water resources. It continues with a brief review of the methods that are applied to estimate those components of a water balance equation which cannot be or are difficult to measure directly. Consequently, mathematical models as useful tools to simulate water balance components are introduced. In that sense, an elaboration on the advantages and limitations of mechanistic and physically based models is given. To alleviate the limitations of the applicability of such models, lysimeters and their applications in water balance studies are introduced. The chapter comes to an end after outlining a brief review of the contents of the chapters in this thesis.

Chapter two describes the role of subsurface flow dynamics in generating runoff and sustaining the base flow of rivers in shallow groundwater catchments. Application of the tracers methods and simulation models for identification of subsurface flow mechanisms and quantification of subsurface flow contribution to flow generations are addressed. Advantages and limitations of integrated and physically-based models for simulating the water balance components, including subsurface flow dynamics are highlighted in this chapter.

The third Chapter focuses on comparing different levels of complexity for simulating the water balance components in a weighing lysimeter. Three conceptual models, each representing one level of complexity are introduced. In the following, it is described how a targeted calibration of these three models were accomplished and how the calibrated models were compared in terms of their performances. In the end, an assessment of the predictive capability of the three models in reproducing deep percolation at the time scales of rainfall events, months, seasons and years are presented.

Chapter four addresses the application of sensitivity and identifiability analyses as two diagnostic tools for better understanding of the complex models behaviors. The main objectives in this chapter are to; i) perform a temporal sensitivity analysis (TSA) to study how the uncertainty in the model output can be apportioned to different inputs, ii) carry out a temporal identifiability analysis (TIA) of model parameters to extract the maximum information content from available observations, and iii) discuss the relationship between TSA and TIA results.

Last chapter comprises two sections. In the first section, a short summary as well as the key findings of the research are described. In the second part, the limitations of the study in addition to some recommendations for future research are presented.

Chapter2

Subsurface flow contribution in the hydrological cycle: Lessons learned and challenges ahead - A review

Published in Environmental Earth Sciences Journal

Ghasemizade, M., Schirmer, M., 2013. Subsurface flow contribution in the hydrological cycle: lessons learned and challenges ahead-a review. Environmental Earth Sciences 69(2) 707-718.

Abstract

Subsurface flow to maintain base flow and its contribution to high flow is of high significance. The high contribution of subsurface flow to stream flow has usually been determined based on the application of tracer methods. However, there are some studies that challenge tracer test applications. These studies have shown that tracer test applications lead to a high percentage of subsurface flow contribution since advection and dispersion effects are not individually considered in the mass balance equation. On the other hand, there is not yet a broad consensus of the responsible mechanisms that justify high contributions of underground water to river flows. In this paper, we focus on the contribution of subsurface flow to high flows, although a brief description of their role in low flows is included. We discuss different suggested mechanisms, considering their applicability, strengths and inadequacies. Also, the application of tracer experiments is elaborated. Finally, the challenges of modeling surface/subsurface flow interactions are addressed, followed by a short description of our future targets.

2.1 Introduction

Despite the fact that groundwater and surface water are often hydraulically interconnected, they are traditionally considered as two separate systems and are analyzed independently. Such a separation is partly due to the belief that groundwater movement has a much larger timescale than that of free surface water movement, and partly due to the difficulties in measuring and modeling their interactions. There exist extensive hydrodynamic models, with different levels of complexity that treat the surface and subsurface flows independently. Nevertheless, the importance of considering the surface water and groundwater as a single body has become an increasing necessity, in terms of both high flows/peak flows/floods and low flows/base flow (Liang et al., 2007; Weill et al., 2011; Winter et al., 1998). We note that regional/trans-boundary deep groundwater flow is not the focal point of this paper, particularly when we discuss high flows. In fact, the focus is on hillslope areas where groundwater table is shallow. In these areas the unsaturated zone controls the separation of rainfall into surface runoff and infiltration during a rainfall event.

2.1.1 Base flow and low flow

Streams can originate from different sources. The main sources are glaciers, overland flow due to precipitation and subsurface (groundwater) flow. Among these, the latter is the least variable source (Winter, 2007) and therefore the role it plays in terms of sustainability should be considered carefully. This is especially true when groundwater provides a storage mechanism that can help to potentially mitigate negative effects of climate warming on the availability of water resources and maintaining river base flows.

Base flow is defined as the component of flow in a river which is not the direct consequence of the rainfall event but is considered as the outflow of the groundwater reservoir feeding the river during the rainless period (Frohlich et al., 1994). Nevertheless, base flow is typically investigated in the context of rainfall runoff studies in which it is separated from generated stream flow during precipitation. Regarding the importance of base flow in maintaining sustainability, few studies have investigated the involving mechanisms which generate stream flow during inter-storm/seasonal base flow periods (e.g. Kish et al., 2010; Payn et al., 2012). These mechanisms become important when the object is determining base flow (low flow indices) in ungauged catchments (sites). In recent years, problems of droughts have focused attention on base flow periods and the processes sustaining water resources for both human consumption and ecosystem needs during dry spells (Jones et al., 2006b; Lehner et al., 2006).

Nonetheless, base flows are often viewed as rather "dull", static periods compared with more "exciting" flood events. Furthermore, the processes contributing to low flows are often considered to be "simply" groundwater discharges to surface waters. Also, in most cases base flow separation has been accomplished during a rainfall runoff simulation that does not help understanding base flow processes seasonally, particularly when evapotranspiration is high. Additionally, a given system or reach may be losing during high flow/river stage but become gaining as flow declines and the hydraulic gradient shifts toward the channel. Studies have shown that such two-way exchange does occur and that it can impact riparian groundwater and stream flow chemical composition long after floodwaters recede (Baillie et al., 2007; Squillace, 1996; Whitaker, 2000).

In many cases, the majority of stream flow discharge during low flow periods is derived from groundwater storage releases (Smakhtin, 2001). Low flow, as it was defined by the international glossary of hydrology (WMO, 1974) is the "flow of water in a stream during prolonged dry weather". So, considering groundwater resources as reservoirs that could maintain sustainability as well as knowing how these reservoirs are operating are of great significance. The percentage contribution from groundwater to streams has been reported as high as 60 % by Liu et al. (2004), greater than 75 % by Clow et al. (2003) and up to 80–100 % for snowmelt in three high elevation basins by Huth et al. (2004) [For more examples of the role of groundwater in maintaining base flow, readers are referred to Winter (2007)]. Using a multiple linear regression equation to predict seasonal low flows in Selwyn River in New Zealand, McKerchar and Schmidt (2007) concluded that low flows decreased at a rate of about 32 L/s per year over the 22 years of recording. They attributed this decrease to groundwater abstraction and emphasized as well the role that groundwater could play in maintaining low flow.

To avoid seemingly different interpretations in sustaining stream flow, a distinction should be made between the water that is stored in the soil and moves through the phreatic zone (inter-flow or through-flow) and deep groundwater. Although there is rich literature on the importance of soil in sustaining base flow seasonally, it is not well documented how soil water interacts with base flow. Maybe the research done by Edlefsen and Bodman (1941), was one of the earliest in the context of soil water dependent base flow. They showed in a plot scale, which was soaked to a depth of 7 m by irrigation and sealed to prevent evaporation, that drainage was continuous over a period of 832 days. Nixon and Lawless (1960) calculated from moisture measurements the downward movement of approximately 28.5 cm of previously stored soil moisture (soil-water) from a 6 m profile of sandy soil during a 6-month dry season. They concluded that slow drainage from unsaturated soil may contribute significantly to groundwater recharge. Remson et al. (1960) indicated through their studies of an intermediate zone at Seabrook, New Jersey, USA, that downward gradients of hydraulic head produced slow but continuous rates of drainage even during the season of evapotranspiration.

Recent studies at the mesoscale (ca. >100 km^2) have shown that different parts of catchment landscapes can have markedly contrasting roles in low flow generation (Orr and Carling, 2006; Peters et al., 2006). The aggregated effects of such spatial variation in catchment characteristics are often unclear. For example, using geochemical tracers and hydrometric data, Tetzlaff and Soulsby (2008) showed for a 1849 km^2 watershed in Scotland that periods of base flow were very dynamic for sub-catchments of the watershed, based on different reactions of sub catchments to isolated small rainfall events. The issue of diurnal variability in low flows is clearly an issue that warrants further study in order to identify the process controls (Wondzell et al., 2007). Also, there are a few studies which have investigated the nature of interacting controls on low flow generation mechanisms in larger river systems (> 1000 km^2). Due to the usual absence of major aquifers in montane headwaters, they are not considered as large contributors of base flow. Therefore, attentions are often shifted to larger groundwater resources in lowland areas as the assumed sources of base flows. According to Shaman et al. (2004) the two limiting factors for lack of enough large-scale studies on controlling factors of low flow generation mechanisms are: 1) absence of tools that allow processes to be extrapolated from point scales to larger catchment scales; 2) downstream increasing anthropogenic impacts in larger catchments and thus, masking natural variability. Tetzlaff and Soulsby (2008) stated that the role of headwater on groundwater in maintaining sustainable downstream low flow is not well recognized in the UK. They also emphasized that base flow generating mechanisms are more complex than what is believed.

Based on what has been explained above, it is clear that further research is needed to understand how base flows sustain water supplies and aquatic ecosystems, if appropriate management is sought to protect these catchment services from environmental change. We believe that better understanding of the interacting controls on low flow generation mechanisms can lead to better management of limited water resources.

2.1.2 High flow

The exerted role of subsurface flow has been shown to be of key importance in runoff generation. Pinder and Jones (1969) were among the first scientists who showed the influential contribution of groundwater in runoff through employing a mass balance equation for solutes. They showed that the groundwater component of runoff varied from 32 to 42 percent for three sub basins in the US. To many, it might seem that groundwater movement speed is not fast enough to contribute to runoff generation, but it has been shown, through numerical and experimental studies, that subsurface flow can transmit water at rates sufficient to contribute to storm flow (Fiori et al., 2007; Freeze, 1972; Harr, 1977; Pierson, 1980). Wenninger et al. (2004) showed that subsurface contribution was about 80% during a double peak flood event.

It should be mentioned that when the term subsurface flow is used, it could be the old water (pre-event) already stored in the catchment or new water (event water) that moves underground due to precipitation. Whether the subsurface flow contribution is dominated by old water or new water is still challenging due to different research results. For example, on one hand, Cloke et al. (2006) indicated that pre-event water played a minor role in runoff generation and just in a small number of cases high proportions of old water were observed at the outflow. On the other hand, applying a series of two-dimensional (2D) numerical simulations, Fiori and Russo (2007) concluded that the principal mechanism for stream flow generation in rainfall runoff processes is subsurface flow along the soil-bedrock interface combined with groundwater ridging in the vicinity of the hillslope base. In fact, they determined pre-event water as the dominant discharge contributor to stream flow. This topic is discussed in detail in Section Mechanisms.

It is generally agreed that once rain falls on the land surface, the unsaturated zone controls the separation of rainfall into surface runoff and infiltration. However, how and when the unsaturated zone starts to play this role is under intensive research. Some theories have been suggested from which three of them have been widely accepted. They are subsurface storm flow, variably saturated subsurface flow and partly saturated subsurface flow. These conceptualizations of runoff generation are discussed in details in Section Mechanisms. Generally, it is agreed that if the dominant mechanism is determined or observed, the way for estimating flood features in ungauged catchments is paved. In practical engineering, dominant mechanism or physics-based applications are rarely pursued. Instead, engineers apply a probability distribution model for estimating rare flood events for designing flood control structures. Although this approach is easy to use and may result in good estimations, particularly in catchments which have long flow records, it assumes that future events are similar to those previously observed (stationarity). Also, this method is ill suited to address hydrologic responses to climate or/and land use changes. In summary, knowing peak flow generation mechanisms can lead to estimations which make sense physically and could also be applied in ungauged catchments as well as catchments in which long records of flow data do not exist.

With respect to the studies of high flows, there are two different kinds of challenges. On the one hand, different theories have been suggested to explain the physical responsible mechanisms that convert the subsurface flow into stream discharge (Cloke et al., 2006; Mcdonnell, 1990; Weiler and Naef, 2003). On the other hand, there are studies that challenge the standard application of mass balance equations, which are used as a basis to estimate subsurface flow contribution to stream flow. These equations are believed to lump the advective and dispersive/diffusive fluxes and thereby affect the interpretation of data (Chanat and Hornberger, 2003; Jones et al., 2006a; Park et al., 2011). In the following, we review the two above-mentioned challenges individually and address the research needs in these areas.

2.2 Mechanisms

Subsurface storm flow is defined as “the water that infiltrates through the ground surface, flows laterally toward the stream as unsaturated flow or shallow perched saturated flow and enters the stream through a seepage face that is above the stream flow level and below the line that the water table intersects the bank river” (Freeze, 1974). Freeze (1974) described the terms “interflow” and “base flow” as part of the stream hydrograph that can be attributed to lateral inflow from the subsurface storm flow and groundwater flow, respectively. He divided the responsible mechanisms for *runoff* generation in an arbitrary classification into two categories: overland flow and subsurface storm flow.

The concept of runoff generation due to overland flow was first discussed by Horton (1933). He showed through some observations and empirical infiltration curves, that runoff happens if the rainfall intensity exceeds the infiltration capacity. Rubin (1966) showed that if unsaturated soil properties, initial soil moisture conditions, and rainfall intensity are known, the infiltration curves can be predicted. He identified rainfall rates greater than the saturated hydraulic conductivity and rainfall duration longer than the time required for soil to become saturated at the surface, as necessary conditions for overland flow generation. However, Freeze (1974) challenged the Hortonian runoff generation mechanism as the dominant mechanism. He inferred that two conditions are required in order to accept Horton concept as a runoff generating mechanism: 1) overland flow is generated when soil becomes saturated from above (the surface) by rainfall; 2) the runoff processes described by Horton are dominated in arid or semi-arid regions where rainfall intensity exceeds soil infiltration rates. Intensive studies in the beginning of the 1970's, particularly in humid vegetated areas, showed that Horton's concept could not justify runoff generation since rainfall intensity did/could not exceed infiltration rate in many cases. For example, in regions with sandy or gravelly soils, rainfall could not surpass infiltration rate, yet nearby stream flows increased [the reader is referred to papers by Rawitz et al. (1970) and Hills (1971)]. The overwhelming conclusion of all those studies was that overland flow was a rare occurrence in time and space in humid vegetated basins. So, the incapability/inadequacy of Horton's concept in describing runoff processes led to two other theories named "partial area contribution" concept (Betson, 1964), and "variable source area/variable saturated flow (VSF)" concept (Dahlke et al., 2012; Hewlett, 1974; Hewlett and Hibbert, 1963). Partial area contribution theory was based on regular overland flow contributions of some fixed parts of the watershed, whereas the concept of VSF assumed an expanding channel network wherein the channels reach out to tap the subsurface flow systems which have overridden their capacity to transmit water beneath the surface (Freeze 1974). The two major differences between these two theories are: 1) contracting/expanding areas in VSF concept are not fixed parts as they are in partial area theory; 2) partial area concept assumes that saturation starts from above, whereas in VSF theory saturation initiates from below.

2.2.1 Capillary fringe

Although the theory of subsurface flow was discussed as one of the likely dominant mechanisms of stream flow generation in early works of Hewlett and Hibbert (1963), and Whipkey (1965), the theory did not get support from researchers until late 70's and early 80's due to lack of enough evidence. Sklash and Farvolden (1979) showed through field observations, isotope applications and computer simulations that rapid increase in hydraulic head near streams caused groundwater ridging and was therefore responsible for rapid contributions of soil water to stream flow. Later, Gillham (1984) did a point-scale field experiment in which he showed the effect of the capillary fringe on water table fluctuations. He indicated that constant specific-yield-based prediction of a recharge value led to a number that was about 30 times away from reality. He then concluded that considering specific yield as a constant value to calculate recharge amounts results in tremendous errors, especially in areas where the water table is close to the ground surface. Therefore, he suggested the specific yield to be determined based on water content-pressure head relation (water retention curves) and the depth to the water table. He then expressed the idea of capillarity and specified that near-zero specific yield values are present in capillary fringe. To show the effectiveness of the capillary fringe theory on subsurface contribution, Abdul and Gillham (1984); Abdul and Gillham (1989) designed lab and field experiments. Within their lab experiment, they designed a box 140 cm long, 8 cm wide, 120 cm high and packed it with medium fine sand in a way that the top right level of the sand stood at 108 cm and the left bottom was kept at the level of 80 cm. Throughout the experiment, they maintained the water table at three different depths and applied rainfall at two different (high and low) intensities. Using chloride as a tracer, their experiment results indicated that the discharge of pre-event water to the pipe at the bottom of the slope proceeded event water, especially at early times of stream flow. They attributed the rapid movement of subsurface flow in the box to the capillary effect. Abdul and Gillham (1989) also conducted a field experiment in an area of 18 m × 90 m in a shallow sandy aquifer at Canadian Forces Base Borden, Ontario, Canada. Based on their short interval water table measurements in their heavily instrumented site, they attributed the sharp rise of the water table in the vicinity of the man-made channel, flowing through the middle of the catchment, to capillarity. Their conclusion was very critical as they wrote "the temporal and spatial variations in the hydraulic-head and water table responses can *only* be explained by invoking the principles of the capillary fringe".

Jayatilaka and Gillham (1996) argued that capillarity is a key factor in controlling dynamics of near stream flow and that incorporation of capillary fringe effects in models could improve the representation of runoff processes as well as their enhanced predictive accuracy. Based on this work, they developed their own model named HECNAR. The model was based on the perception that a watershed can be divided into three zones based on their respective storage characteristics. Zone 1 was the area which extended up to a point in which the water table depth equaled the capillary fringe height. Zone 2 was considered the area where soil moisture was between field capacity and residual moisture, independent of the water table depth. Finally, the moisture deficient area, due to evapotranspirational losses, was named zone 3. The assumptions that were made to approximate the physical system included isotropy and homogeneity of porous media, neglecting interception and depressional storages, and ignored water loss owing to evapotranspiration as it was assumed to be small within the duration of an event. They believed that "HECNAR incorporates the high discharge of subsurface water to the stream as a result of increased hydraulic gradient toward the stream".

McDonnell and Buttle (1998) challenged Jayatilaka and Gillham (1996), regarding the capillary-fringe-induced groundwater-ridging as the major mechanism of pre-event contributions to streams in near stream environments. They suggested alternative mechanisms such as preferential flow. In fact, they based their criticism on the observation of rapid water table responses in the absence of a capillary fringe. We also think that the assumption "water loss due to evapotranspiration could be neglected" in HECNAR contradicts the definition of zone 3. McDonnell and Buttle (1998) inferred that the widespread applicability of groundwater ridging mechanism remains uncertain as rapid pre-event contributions to storm flow can originate from a range of hydrological processes. Moreover, they were confident that a conceptual paradox exists since the capillary fringe height of a soil is usually inversely related to its hydraulic conductivity. Therefore, the greater the tendency for capillary fringe rise, the less likely that rapid Darcian flux of groundwater can occur even with steepened hydraulic gradients in the near stream zone (McDonnell and Buttle, 1998; Zaltsberg, 1986). Cloke et al. (2006) took the laboratory experiment of Abdul and Gillham (1989) to validate the hypothesis of capillary fringe effect on pre-event contributions within a 2D finite element numerical model. They showed that while the ridge has not yet reached the surface, Darcian velocity vectors move away from, rather than toward, the channel to fill the area of storage in the unsaturated zone. In fact, they indicated through their simulation results that the ridge formation was not responsible for the pre-event contribution to the stream as the pre-event contribution started to begin when the surface pressure head equaled to zero. Afterwards, they showed the low proportion of pre-event water contribution to stream discharge, which was hypothesized to be due to groundwater ridging in specific conditions of the Abdul and Gillham (1984) laboratory experiment. They varied some influential variables and carried out a set of numerical simulations to look for evidence of groundwater ridging mechanism and pre-event contributions in other conditions. The variables which they varied were, initial water table depth, rainfall intensity, slope, saturated hydraulic conductivity, capillary fringe height, and volume of the sand box. It is beyond the scope of this paper to discuss the effects of individual and interrelated variables, however, the main findings of their numerical experiments were as follow:

1) Rainfall intensity was the most sensitive variable which influenced the portion of pre-event contribution, though its effect in ridge formation was limited to high hydraulically conductive areas where the capillary fringe did not reach the ground surface.

2) Whereas the capillary fringe was seen to be a controlling factor in ridge development, it had little effect on the pre-event water contribution.

3) Initial water table height had the maximum effect on both ridge development and domination of pre-event water discharge.

Park et al. (2011) also applied a numerical model to a simple catchment and concluded that capillarity cannot lead to enough mechanical flow. Based on the above discussion, groundwater ridging (capillarity), which has been debated over the last three decades, could not be relied on as an influential mechanism to explain subsurface flow contribution to runoff generation. We briefly review two other widely expected mechanisms in the following.

2.2.2 Pressure wave translatory flow

The mechanism is very analogues to variable saturated flow as it suggests that some subsurface layers will be saturated temporally and will extend in area and volume across slopes or large parts of catchments. Compared to the VSF mechanism, however, pressure wave translatory flow will initiate when continuous hydraulic connection is established across slopes and elevation zones, and thus individual groundwater bodies link together (Becker, 2005). Burt and Butcher (1985) provided evidence to show the applicability of this mechanism by observing groundwater level fluctuation in a densely instrumented 1.4 ha hillslope in UK. They observed that as soon as previously disconnected groundwater bodies at bedrock interface merged and formed a continuous saturation layer across the slope, a secondary rise in stream flow occurred. Similar observations were reported in other catchments (Bazemore et al., 1994; Becker, 2005; Kirnbauer and Haas, 1998; Torres et al., 1998). Although the mechanism seems to be logical and it makes sense physically, experimental evidence on this kind of subsurface runoff and the conditions that control it are poorly understood. Also, quantifying different components of this perceptual model has not been widely done. For the most recent applications of pressure wave theory, readers are referred to Vidon (2012).

2.2.3 Transmissivity feedback

The mechanism is based on the idea that saturated hydraulic conductivity decreases as depth increases. In fact, transmissivity feedback is a special case of translatory flow where shallow groundwater displacement is enhanced by a decrease in saturated hydraulic conductivity with depth (Uhlenbrook and Hoeg, 2003). This mechanism was first introduced by Bishop (1991) and since then it has been widely applied in the field of hillslope runoff generation. Cloke et al. (2006), for example, incorporated the method in a numerical experiment to test its applicability in explaining high amounts of observed pre-event water. They concluded that even though the water table levels rose rapidly, less stored (old/pre-event water) water was enabled as discharge due to decreased hydraulic conductivity (potential water movement). Bishop et al. (2011) described runoff response and quantified total water storage, flow paths, and vertical distribution of lateral flow in a catchment of 6300 m^2, using the principles of the transmissivity feedback runoff generation mechanism. [For more applications of transmissivity in runoff generation, readers are referred to Kendall et al. (1999); Laudon et al. (2004); Detty and McGuire (2010)].

This variety of interacting processes, found in different environments, makes the estimation of how water enters the stream at a given site problematic without field investigations. We strongly believe that there is not yet a broad consensus on how subsurface flow contributes to stream flow, even in one specific catchment or site. It goes without saying that first-order controls in one catchment may not be controlling factors in other catchments, depending on variation in geology, soil properties, rainfall features (duration and intensity), geometry, land use, etc. [for a review of how above-mentioned factors may affect stream flow generation, readers are referred to Bachmair and Weiler (2011)]. It seems that state variables are promising for generalization to similar catchments. Weiler and McDonnell (2004) argued that documenting idiosyncrasies of new hillslope environments should be replaced with defining generalizable appropriate state variables in different environments. They believe that if this shifting occurs, major experiments and excavations done in a specific hillslope/catchment will have transference value to a neighboring environment as a variety of properties change. Weiler and McDonnell (2004) developed a numerical physically-based model, named HillVi, and explored the variation of drainable porosity as first-order control in hillslope hydrology. They tested their hypothesis (assuming drainable porosity as a first-order control) for a virtual hillslope by application of their model to simulate flow and transport for two different drainable porosity values while keeping other parameters and inputs constant. They concluded that drainable porosity can explain spatial and temporal variation of subsurface flow, saturation depth, tracer movement and its concentration as well.

2.3 Numerical analyses of tracer applications

McGuire et al. (2007) argued that tracer experiments and their resulting breakthrough curves can be counted on as additional data sources which reflect the complexity of physical processes into one signal, like a hydrograph, as well as integrating flow heterogeneity and thus as tools that can constrain parameterization and reduce model uncertainty. Tracers can also delineate the origin of water (Chen et al., 2012). Nevertheless, the incorrect judgment based on their applications could end in misleading results.

In principle, the contribution of pre-event water can be derived based on the results of tracer data, which are interpreted using mass balance equations. The assumption that has been implicitly put into mass balance equations is that hydrodynamic mixing processes (such as mixing of pre-event and event water) are adequately accounted for in the calculation of the volumetric subsurface flow contribution (Jones et al., 2006a). To determine the proportion of event water and pre-event water with application of conservative tracers, it is very common to first sample subsurface water and rainwater to know their respective tracer signatures and then take multiple samples in the stream at regular time intervals during the storm and for a while after it has ended. Afterwards, based on different ratios of concentrations in the stream water and unit hydrograph, the above mentioned proportion will be calculated. The key point about the hydrograph separation done this way is that it can only differentiate sources of water (event/pre-event) and cannot separate between water pathways (Jones et al. 2006a). In fact, there should be a clear distinction between temporal water sources (event/pre-event or old/new) and water flow pathways (overland or subsurface saturated/unsaturated). Renaud et al. (2007) define pre-event water as the water that is stored in a catchment prior to the beginning of a rainfall event. It is very important to note that pre-event water can follow different pathways to contribute to stream flow. Buttle (1994) accounted groundwater as only one out of six processes that can deliver pre-event water. In summary, there seem to be a necessity to scrutinize the efficiency of mass balance equation applications in order to better estimate the percentage of pre-event contribution.

VanderKwaak (1999) applied a finite element method to simulate the rainfall runoff experiment of Abdul and Gillham (1989) relying on a tracer-based separation method similar to that used by Abdul and Gillham (1989). He found significant discrepancy between model results (subsurface contribution) which were obtained when tracer (bromide) concentrations at the outlet were entered into mass balance equations and when nodal tracer fluxes were summed. Whereas he did not explicitly separate advective tracer contributions from dispersive/diffusive contributions to total solute fluxes entering the channel at each time step, he suggested that the discrepancy could have occurred due to dispersive/diffusive mixing processes at the surface subsurface interface. In light of the factors that can affect the strength of hydrodynamic mixing, Jones et al. (2006a) introduced mechanical dispersion, molecular diffusion, and rainfall intensity/duration as the influential factors. They conducted numerical experiments to compare the computed Darcian-based groundwater fluxes contributing to stream flow with estimates of those contributions based on trace-based separations. They found that contributions calculated based on the above two mentioned methods were significantly different. They attributed the difference to the hydrodynamic dispersion of event and pre-event water tracers. It was featured in their study that hydrodynamic mixing processes can dramatically affect estimates of pre-event water contributions based on tracer-based separation method, as well as demonstrating that the actual amount of groundwater contribution was smaller than tracer-based estimated amount even if the mixing processes were weak. Jones et al. (2006a) showed through their numerical simulations that event and unsaturated zone pre-event waters mix with each other by means of dispersive/diffusive processes before discharging into the channel. To further demonstrate the impact of dispersive/diffusive mixing processes on traditional based hydrograph separation, they assessed the influence of subsurface longitudinal dispersion, rainfall/intensity duration and multiple sequential rainfall events. Having increased the value of the dispersion coefficient, they observed a noticeable increase in the estimate of tracer-based pre-event contributions. In contrast, they decreased the coefficient to near zero. Then, the subsurface contribution minimally declined in comparison to the base case. They stated that even though the effect of mechanical mixing was eliminated, molecular diffusion can strongly influence the mixing process. To indicate the influence of rainfall intensity/duration, they set two scenarios. In the first scenario, they increased rainfall intensity and decreased the duration and in the second one they did just the opposite. In both scenarios the volume of rainfall was maintained equal to the base case amount. They concluded that increased rainfall intensity leads to less tracer-based pre-event contribution, as event and pre-event waters have less time to hydro-dynamically mix before being transmitted to the channel. The converse argument was also made regarding the effect of decreased intensity. Finally, they subjected the system to multiple sequential rainfalls separated by a three-day recovery period. They observed that subsurface contribution decreased as it was expected. They attributed the decline to less mixturing of pre-event and event water as progressively more pre-event water would discharge from the system.

It should be noted that the challenging relationship of capillary fringe and pre-event contribution to stream flow was not clearly and explicitly discussed in Jones et al. (2006a). However, Park et al. (2011) later showed that the capillary fringe can accelerate the mixing of event and pre-event water parcels. Renaud et al. (2007) criticized Jones et al. (2006a) for not distinguishing between temporal sources and mechanical carriers of water contributions to stream flow. This issue was then discussed by Park et al. (2011) stating that the tracer technique for hydrograph separation to deduce the temporal origins of water entering a stream is influenced by pure mechanical flow processes. Also, Renaud et al. (2007) challenged Jones et al. for ignoring kinematic dispersion in water molecules as a potential source of error in estimating the pre-event contribution. Therefore, Renaud et al. (2007) stressed that diffusion and dispersion coefficients for water molecules themselves should be accounted for in modeling, in order to represent their travel through the subsurface, as well as parameterizing them based on site characteristics and tracer properties. Park et al. (2011) clarified the arguments of Renaud et al. (2007) and Sudicky et al. (2007) by showing that the "tracer technique for hydrograph separation to deduce the temporal origins of water entering a stream is influenced not only by pure mechanical flow processes, but also by mixing processes induced by potential chemical gradients".

Using the fully surface/subsurface integrated model of HydroGeoSphere (HGS), Park et al. (2011) analyzed the relationship between the spatial and temporal origins of storm flow in the stream as well as looking into how precipitation influences the flow in the catchment. To accomplish that, two cross sections, parallel (A) and perpendicular (B) to the stream, of a simplified virtual catchment were assumed. To maintain simplicity, they ignored evaporation and transpiration and assumed uniformity and isotropy of hydraulic properties. Regarding the simulation in plane (A), they observed that pre-event discharge increased far greater than the mechanical subsurface flow component as rainfall intensity augmented. They ascribed the strong pre-event stream discharge, often interpreted based on conventional tracer-based hydrograph separations, to added effects of diffusion and mechanical dispersion. As it is generally accepted (e.g., McDonnell 1990; Weiler and Naef 2003) that considering macropores in porous media can explain the high contribution of pre-event water, Park et al. (2011) applied a dual-permeability approach through attributing 1% of the bulk volume a high hydraulic conductivity value to test this hypothesis. While they considered an arbitrary value of 1% as the simulated bulk volume occupied by macropores, results showed that mechanical contribution of subsurface flow increased, whereas their contribution diminished as rainfall intensity rose again due to further mixing. They concluded that "compared to single continuum simulation cases, pre-event water contributes more to the total stream discharge because of enhanced mechanical input of water and because of the enhanced dispersive input to the stream induced by macropores". In plane (B), perpendicular to the stream, while pre-event unsaturated discharge ratio to saturated portion incremented due to increase in rainfall intensity, the ratio of increase in exfiltration values (mechanical mechanism) did not reconcile. These results led them to the conclusion that "capillary fringe groundwater ridging may not generate enough mechanical flow for observed pre-event discharge, but it may accelerate mixing processes such that more pre-event water discharges to the stream". They reached the same conclusion in plane (B) as in plane (A) saying that pre-event water contribution by mechanical flow processes to the stream discharge is limited without dispersion. Results of dual-continuum simulation in plane (B) were similar to those derived for plane (A).

2.4 Modeling

Models as useful hypothesis testing tools enable us to study combinations of conditions which have not yet been encountered in field studies or cannot be replicated at field scale (Johansson, 1985). Recently, physically-based models have been vastly utilized to simulate short-term (event-based) and long-term interactions of subsurface surface flow on the premise that such models can account for internal processes and complexities (James et al., 2010; Jones et al., 2008; Mirus et al., 2011) and hence could be applied in ungauged catchments. Assuming the above mentioned assumption is true, the question that quickly follows is: why such models are calibrated? McDonnell et al. (2007) answer this question quoting "…models based on current theories rely on calibration to account for our lack of knowledge of the spatial heterogeneities in landscape properties and to compensate for the lack of understanding of actual processes and process interactions".

With respect to process understanding, around three decades ago, Dooge (1986) published a paper titled "Looking for hydrologic laws" and asked for new visions in the science of hydrology. Dooge suggested a new framework for developing new theories including: 1) searching for new macroscale laws; 2) developing scaling relations across watershed scales, and 3) upscaling from small-scale theories. It is surprising that after about 26 years, Dooge's suggestions have not been fully pursued. Meanwhile, it is worth mentioning that some alternative concepts such as Representative Elementary Watersheds (Reggiani et al., 1998; Reggiani et al., 2000; Zhang and Savenije, 2005) or Hydrological Response Units (Flugel, 1995; Kouwen et al., 1993; Viviroli et al., 2009) have been introduced. Darcy- Richard's equation, which is the foundation of many physically-based models, as a subgrid-scale parameterization approach is often consistent with the point-scale measurements (tensiometers, TDR etc.), particularly in soils where flow is dominated by matrix flow. Darcy-Richard's equation often breaks down at larger scales or in soils dominated by preferential flow (Weiler and Naef, 2003). Furthermore, spatial discretization of Richard's equation is another issue whose limits have been addressed in many papers (Downer and Ogden, 2003; van Dam and Feddes, 2000). Vogel and Ippisch (2008) showed critical spatial discretization length at unit gradient in a typical sand is about 5 cm. They stated that if spatial discretization goes beyond its critical limit, convergence of the solver and accuracy of the solution would be influenced.

Type of data we currently collect is another issue that hinders the ideal application of 3-D physically-based models in catchment scale as such data cannot fully characterize the catchment (Loague et al., 2005). It is now widely accepted in the hydrologic community that laboratory data or even data collected at individual points in the field are of limited value in parameterizing large-scale modeling (Doherty and Christensen, 2011; James et al., 2010). Recent approaches, such as combined application of geophysical and hydrogeological data to delineate subsurface heterogeneity (Doro et al. 2013) should be pursued and developed. Therefore, as long as new methods or new devices to account for macro-scale processes are not established, applications of highly parameterized 3-D physically-based models are not promising.

Doherty and Christensen (2011) appreciated the value of micro scale physics-based simulations as they wrote "complex numerical models have the advantages of allowing representation of complex processes and heterogeneous system property distributions inasmuch as these are understood at any particular study site". On the other hand, they challenged application of complex models due to their long run times, occasional numerical instability, and analysis of their predictive uncertainty. There is a broad consensus that such heavily parameterized models lead to high predictive uncertainty (Beven, 2000). One more issue that is addressed well by Brunner et al. (2012) is the worth of observation data in identification of parameters (parameter identifiability) and predictive uncertainty. One conceptual relationship between model complexity, data availability and predictive performance is illustrated in Figure 2-1. As indicated in Figure 2-1, for a limited to moderate amount of data, increasing model complexity leads to reduction in model performance. Having reviewed the misleading large-scale application of available physically-based models, we strongly believe that such models that consider internal dynamic processes are valuable learning tools provided that the uncertainty is reduced. Loague et al. (2006) argue that 3-D physically-based models provide foundations for understanding coupled systems at hillslope scale, give new understanding and prompt new experiments. Bredehoeft (2010) introduces models as tools to organize our thinking. He emphasizes that by writing "For me the model is not an end in itself, but rather a powerful tool that organizes my thinking and my engineering judgment".

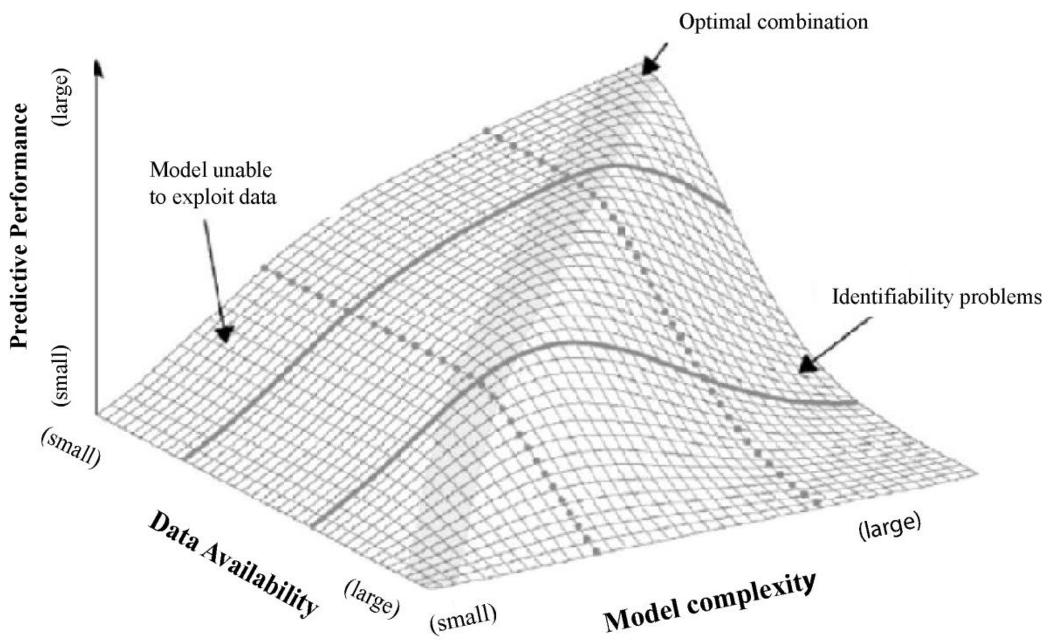

Figure 2-1. Relationship between model complexity, data availability and model performance from Grayson and Blöschl (2001).

We fully agree with the statement which was first suggested by Ebel and Loague (2006) and then was emphasized by James et al. (2010) saying "… the value of physically-based simulation of hillslope and small catchment response will be the examination of their failure to replicate experimental observations and the changes it will bring about to the models themselves".

2.5 Concluding remarks and challenges ahead

In summary, new approaches should not rely on calibration, but rather on systematic learning from observed data, and on increased understanding and search for new hydrologic theories through embracing new organizing principles behind watershed behavior that are derived from our sister disciplines (McDonnell et al., 2007). As Cloke et al. (2006) point out in their paper, most field environments have complex geometries which are very different from lab experiments. Water table topography is a clear example. Moreover, we believe that much of the research in the field of modeling hillslope hydrology needs revision. It should be noted that modeling micro scale (lab experiments) can be beneficial. However, what matters to decision makers is the potential application of hydrology in solving practical problems at catchment or watershed scales. This issue requires more test cases including experimental data sets from lab scale to real world catchments (Grathwohl et al. 2013). Regarding that, transit time distribution, for example, has been shown to be promising in representing integrated responses of diverse flow pathways in hillslope and catchment scale and thus connecting process complexity with model simplification (Doherty and Christensen, 2011; McGuire et al., 2007). Since many catchments and large-scale applications are concerned with water quality aspects such as acidification (Stoddard et al., 1999), cumulative effects (Sidle and Hornbeck, 1991) and nutrient cycling (Creed and Band, 1998), the age or transit time of water offers a link to water quality since the contact time in the subsurface largely controls stream chemical composition, revealing information about the storage, flow pathways and sources of water in a single measure (McGuire and McDonnell, 2006; McGuire et al., 2007).

Acknowledgements

This study was 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”). The first author is holding a PhD Scholarship from the Iranian Ministry of Science and Technology. We wish to thank Jana von Freyberg for discussions. We also appreciate the beneficial comments of the reviewers for their timely comments which helped this paper to sound more consistent and clear.

Chapter 3

The effect of model complexity in simulating unsaturated zone flow processes on recharge estimation at varying time scales

Published in Journal of Hydrology

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. Journal of Hydrology 529 1173-1184.

Abstract

Recent increases in computational power have led to the development of more advanced physically-based models which can handle a wide range of environmental processes. Although these models are very useful for increasing our understanding of unsaturated zone flow processes, their outputs usually contain high uncertainty, particularly when the level of complexity is not supported by observations. In this context, the aim of this paper is to compare the performance of three different model conceptualizations of a shallow unsaturated soil zone using the physically-based model HydroGeoSphere (HGS). To accomplish this task, we simulated actual evapotranspiration (ET), water content (WC) and discharge (D) from a weighing lysimeter for each of the conceptual models. Conceptual Model 1 considers the lysimeter as a homogeneous zone with matrix flow, while conceptual Model 2 has an added preferential flow component. Conceptual Model 3 includes layered heterogeneity in addition to the matrix and preferential flow components. The results indicated that the model performance in reproducing daily ET, WC and D improves when we move from simple models to more complex models. A comparison between event-based, monthly, seasonal and yearly time scales indicates that the simplest conceptual model is not reliable for reproducing event-based discharges. However, it can compete with more complex models at annual scales, although the uncertainty bound for the simple model is very high. While increasing complexity from the simplest to the more complex model leads to lower uncertainty bounds and more reliable values of the lysimeter discharge at monthly and seasonal time scales, uncertainty bounds became larger when complexity increased in the most complex model. This is related to a higher number of unknown model parameters in the calibration which are not supported by the available observation datasets.

3.1 Introduction

Recent increases in computational power have motivated the scientific community to develop and apply more complex models of coupled environmental processes. Hydrological models, for example, have become more sophisticated in terms of the simulated unsaturated zone flow processes. Although applications of such complex models may in theory lead to increased knowledge into the governing processes, they have their own limitations such as over-parameterization when applied in practice (Perrin et al., 2001). On the other hand, simple models have proved their efficiency and applicability in hydrology, particularly in operational contexts such as rainfall-runoff (Birkel et al., 2010; Brauer et al., 2013).

Physically-based numerical simulation models are valuable tools which can be applied with a range of complexity, depending on the available data. These models have proven to be useful due to the integration of multiple hydrological processes (Brunner et al., 2012; Schwarzel et al., 2006), their applicability for areas which suffer from lack of long-term data (Bolger et al., 2011) and where the measured values of their required parameters are available (Mirus et al., 2011). However, these types of physically-based models have also been criticized for their complexity (Beven, 1993; Beven, 2006), including problems of over-parameterization and equifinality (non-uniqueness). In fact, as processes are simulated with greater detail, the number of parameters in physically-based models will accordingly increase. This issue is exacerbated when physically-based models are employed at large scales such as catchments where limited available data does not support the embedded complexity and therefore can have an adverse effect on prediction uncertainty (Uhlenbrook et al., 1999; van der Perk, 1997).

To overcome this limitation and fill the gap between the laboratory- and field-scale application of physically-based models, weighing lysimeters can provide a valuable framework to estimate recharge. It is worth noting that lysimeters with free-drainage boundary conditions do not fully represent the real field-scale processes leading to recharge because they are disconnected from the water table and therefore capillary fringe effects are neglected. However, Abdou and Flury (2004) showed that the differences are minimal unless strong vertical structure exists among soil particles. Seneviratne et al. (2012) have also shown that the monthly streamflow is highly correlated to the monthly lysimeter discharge. The fact that initial and boundary conditions are well-known in a lysimeter makes them ideally suited for a variety of purposes such as estimating soil hydraulic parameters (Abaspour et al., 1999; Durner et al., 2008), measurement of water uptake and root distribution (Schelle et al., 2013), and for estimating recharge (Kendy et al., 2003; Selle et al., 2011), which is the focus of the current study. It should be noted that we use the term “lysimeter discharge” in this text to represent the aquifer recharge. This assumption seems reasonable since the bottom of the lysimeter is within half a meter from the average water table which is measured in a nearby piezometer. Many of the published lysimeter studies, however, have neglected preferential flow while its importance under natural conditions is now well accepted (Anderson et al., 1997; Beven and Germann, 2013; Weiler and Fluhler, 2004; Zheng and Gorelick, 2003). For instance, Selle et al. (2011) showed that neglecting preferential flow in numerical simulations can bias the predictions of soil moisture and deep percolation.

Preferential flow is one of the processes that, when applied in physically-based models, can greatly increase the number of model parameters. Among the many approaches that have been suggested for preferential flow modeling (e.g., Beven and Germann (2013); Kohne et al. (2009); Simunek et al. (2003)), dual permeability (DP), which assumes that water can flow in the matrix as well as in macropores, is one approach that has been widely applied (Kohne et al., 2006; Kordilla et al., 2012; Wang et al., 2014). The DP approach can simulate peaks due to flow through macropores, as well as base flow and recessions arising from matrix flow. Although attractive, employing the DP concept to represent preferential flow in vadose zone modeling requires additional parameters, such as the fraction of macropores which cannot typically be measured in the field and therefore calibration is required.

While considerable effort has been devoted to seeking simplification strategies involving parameter and/or process lumping in environmental simulations (Shen et al., 2013; Touhami et al., 2013; Zhu and Sun, 2009), few studies have explored the effects of simplification on a model’s predictive performance (Doherty and Welter, 2010; Watson et al., 2013). This issue is addressed in the current study and is, to the best of our knowledge, the first time compared temporally. We applied three different conceptual models (representing different levels of complexity) and compare them in the framework of a weighing lysimeter simulation, using the physically-based model HydroGeoSphere (HGS) (Therrien et al., 2010). Output uncertainty of each conceptual model was minimized through calibration of the parameters against actual evapotranspiration, water content and discharge. The main objective of our study is to evaluate the performance of different conceptual models in replicating the lysimeter discharge for a varying range of time scales. Specifically, we focused our research on whether process simplification and parameter reduction result in acceptable estimations of lysimeter discharge values.

In this paper, we first describe the features of our research lysimeter. We then briefly explain how DP, flow interception (Appendix 3-A), evapotranspiration (Appendix 3-A) and uncertainty analysis are considered in our research. Subsequently, we introduce the three different model complexities and model parameterization. We then show the performance of the three different conceptual models and discuss uncertainty of the predictions for different time scales.

3.2 Data and methods

3.2.1 Rietholzbach lysimeter

The lysimeter of our study is a free drainage lysimeter located in the Rietholzbach research catchment (Seneviratne et al., 2012), which is a pre-alpine headwater catchment of the Thur river basin in northeastern Switzerland. The average annual sums of precipitation (measured at the meteorological station at the same site) and actual evapotranspiration (measured with the lysimeter) in Rietholzbach are around 1,450 and 560 mm, respectively (based on data from 1976–2007, Ewen et al. (2011)). The lysimeter belongs to the category of large lysimeters with a diameter of 2 m and a depth of 2.5 m, and is equipped with TDR sensors for measuring water content (soil moisture). These sensors are located at depths of 5, 15, 25, 35, 55, 80, and 110 cm from the top of the lysimeter. However, due to inconsistency in soil moisture data, the water content values at depths of 35 and 110 cm were not included in the simulations. The lysimeter has been back-filled with gleyic cambisol (clay loam) in 1974 with surrounding soil. Whether the soil is layered or structured is unknown. In fact, this is one of the reasons we included different levels of complexity in our modeling experiment. The surface of the lysimeter is covered with grass to reflect the surrounding natural conditions of the vegetation (structure, cutting, fertilization). Using a dye tracer, Menzel and Demuth (1993) found preferential flow in the Rietholzbach lysimeter, confirmed by Vitvar et al. (1999) by application of ^{18}O isotopes. However, to the knowledge of the authors, no quantitative modeling of this fast flow in the Rietholzbach lysimeter has been reported.

3.2.2 Model setup

In order to evaluate different levels of complexity and the effect of increasing model parameters on predictive uncertainty, three different conceptual models were analyzed within the framework of the HGS model. In conceptual model 1 (C1) with 7 parameters, it was assumed that the lysimeter consists of only matrix flow without any explicit component to represent flow in macropores (Figure 3-1a). However, in order to allow fast flow, the matrix hydraulic conductivity was given a wider range of variation during calibration (Elci and Molz, 2009). Conceptual model 2 (C2) with 13 parameters was set similar to C1 with the same calibration dataset, with the only difference being that DP was added to C2 in order to include preferential flow (Figure 3-1b). This model represents mainly the effect of adding complexity through including *additional processes*. Conceptual model 3 (C3) includes 25 parameters and was distinguished from C2 by adding heterogeneity consisting of four layers of soil, each represented by soil moisture data at different depths (Figure 3-1c). Model C3 mainly represents the effect of adding complexity through including *additional parameters*.

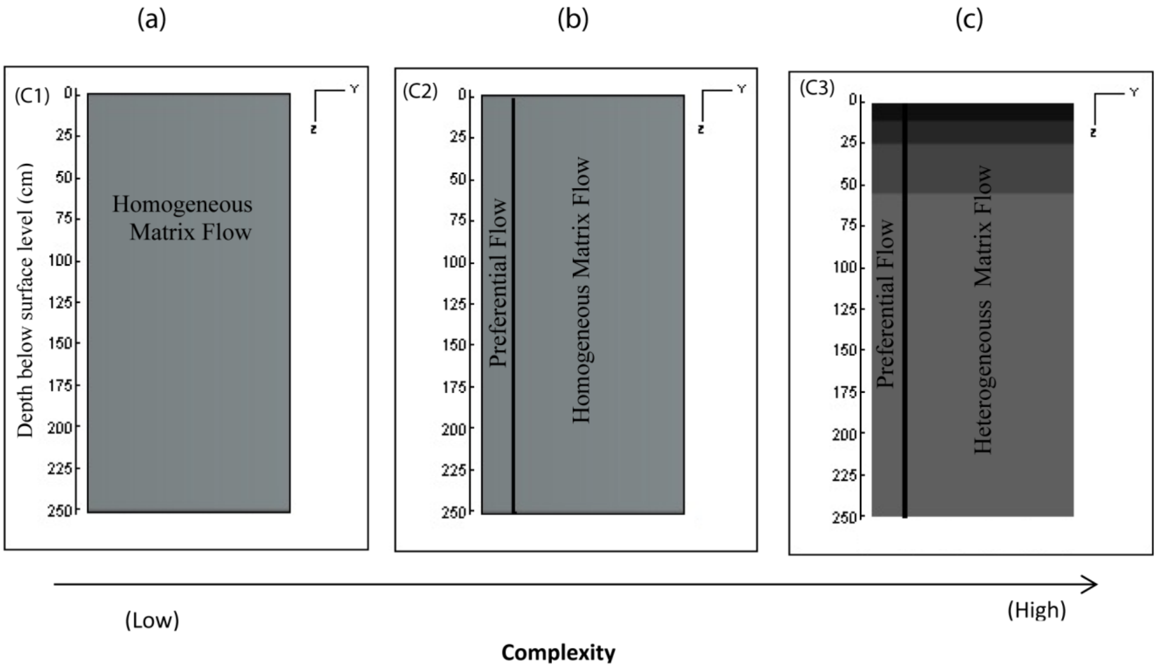

Figure 3-1. Conceptual model setups with different levels of complexity in a vertical cross-section of the lysimeter. (C1) is conceptual model 1 with 7 parameters and only consists homogeneous matrix flow, (C2) is conceptual model 2 with 13 parameters with an explicit preferential flow component and (C3) is conceptual model 3 with 25 parameters. It includes layering (heterogeneity) and an explicit preferential flow component.

We assumed no overland flow forms on the top of the lysimeter and that all the rainfall reaching the ground infiltrates into the lysimeter storage due to the existence of a 10 cm steel edge surrounding the top of the lysimeter. Daily ETp (calculated based on the method of (Penman, 1948)) and daily rainfall, measured in the meteo station at the same site, were applied as variable flux boundaries on the top. The bottom boundary was represented by a seepage face, with no flow when the boundary was unsaturated and using a fixed atmospheric pressure head once saturation was reached. The sides of the domain were all assigned no-flow boundaries due to isolation of the lysimeter container. The lysimeter was discretized vertically into 125 layers, each having a thickness of 2 cm. This fine discretization was necessary to obtain a numerical solution with reasonable accuracy (Vogel and Ippisch, 2008).

3.2.3 Dual Permeability

The dual-permeability simulation approach was first developed by Gerke and van Genuchten (1993a), which assumes that Richard's equation is valid and applicable for both micro- and macro-pore systems. The formulation can be written as:

$$\frac{\partial \theta_f}{\partial t} = \frac{\partial}{\partial z}\left(k_f \frac{\partial h_f}{\partial z} - k_f \right) - \frac{\Gamma_w}{w_f} \tag{3.1}$$

$$\frac{\partial \theta_m}{\partial t} = \frac{\partial}{\partial z}\left(k_m \frac{\partial h_m}{\partial z} - k_m \right) + \frac{\Gamma_w}{1 - w_f} \tag{3.2}$$

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 the 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 :

$$\Gamma_w = \alpha_w (h_f - h_m) \tag{3.3}$$

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 (1993a) given by :

$$\alpha_w = \frac{\beta}{a^2} \gamma k_{sa} \tag{3.4}$$

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:

$$S_w = S_{wr} + \left(1 - S_{wr}\right)\left[1 + |\alpha h|^n\right]^{-m} \tag{3.5}$$

$$K_r = S_e^{(\ell_p)}\left[1 - \left(1 - S_e^{\frac{1}{m}}\right)^m\right]^2 \tag{3.6}$$

$$m = 1 - \frac{1}{n} \tag{3.7}$$

$$S_e = \frac{(S_w - S_{wr})}{(1 - S_{wr})} \tag{3.8}$$

$$S_w = \frac{\theta}{\theta_S} \tag{3.9}$$

where α , n and ℓ_p are fitting parameters, S_e is effective saturation, S_{wr} is residual water saturation, S_w is water saturation, θ_s is saturated water content and K_r is relative permeability.

3.2.4 Model Parameterization

To facilitate parameter tracking, the parameterization was divided into two separate tables. Table 3-1 indicates the range of variability for the evapotranspiration parameters, which were calibrated, as well as the assigned values of the parameters which were not calibrated within the simulation process. The range of parameters for the matrix and macro-pore flow continua is shown in Table 3-2.

Table 3-1. Evapotranspiration parameters.

Parameter	Lower Limit	Upper limit	Reference	Simulation value
c_{int} (*cm*) see Eq. (3-A.2)	0.005	0.01	Andersen et al. (2002)	0.005
c_1 see Eq. (3-A.7)	0.01	1.0	Vazquez and Feyen (2003)	*
c_2 see Eq. (3-A.7)	0.01	0.5	Vazquez and Feyen (2003)	*
c_3 (cmd^{-1}) see Eq. (3-A.10)	1	30	Kristensen and Jensen	*
Lr (cm) see Eq. (3-A.8)	20	100	http://www.soilandhealth.org	50
LAI see Eq. (3-A.7)	0.1	2.5	Scurlock et al. [2001]	See Appendix 3-A
θ_{fc} see Eq. (3-A.10)	0.36		Twarakavi et al. (2009)	0.36
θ_{wp} see Eq. (3-A.10)	$\varphi \times 0.2$		Panday and Huyakorn	$\varphi \times 0.2$
θ_0 see Eq. (3-A.11)	$\varphi \times 0.76$		Panday and Huyakorn	$\varphi \times 0.76$
θ_{an} see Eq. (3-A.9)	$\varphi \times 0.9$		Panday and Huyakorn	$\varphi \times 0.9$
θ_{e1} see Eq. (3-A.13)	$\varphi \times 0.32$		Panday and Huyakorn	$\varphi \times 0.32$
θ_{e2} see Eq. (3-A.13)	$\varphi \times 0.2$		Panday and Huyakorn	$\varphi \times 0.2$
RDF (Lr) see Eq. (3-A.8)	Close to cubic decay with depth		Li et al. (2008)	Cubic decay with depth
EDF see Eq. (3-A.12)	Close to cubic decay with depth		Li et al. (2008)	Cubic decay with depth

Note: φ is the porosity

**indicate the fitting/calibrated parameters*

Table 3-2. Unsaturated flow parameters.

Model Parameter	Lower bound	Upper bound	Reference
$\alpha_f (cm^{-1})$ see Eq. (3.5)	0.01	0.15	Carsel and Parrish (1988)
$\alpha_m (cm^{-1})$ see Eq. (3.5)	0.001	0.07	Carsel and Parrish (1988)
n_f see Eq. (3.5)	1.5	4	Carsel and Parrish (1988)
n_m see Eq. (3.5)	1.2	2	Carsel and Parrish (1988)
θ_{sf} see Eq. (3.9)	0.35	0.6	Weiler (2001)
θ_{sm} see Eq. (3.9)	0.35	0.6	Weiler (2001)
w_f see Eq. (3.1,3.2)	0.01	0.1	Weiler (2001)
$k_f (cmd^{-1})$ see Eq. (3.1, 3.2, 3.10)	42	1150	Weiler (2001)
$k_m (cmd^{-1})$ see Eq. (3.1, 3.2, 3.10)	0.5	10	Weiler (2001)
α_w (d^{-2}) see Eq. (3.3, 3.4)	13×10^{-6}	24×10^{-2}	(Gerke and van Genuchten, 1993b; Kohne et al., 2006)

Note: *f* and *m* subscripts represent fracture and matrix, respectively.

It was assumed that saturated soil moisture equals porosity.

The saturated hydraulic conductivity of the fracture–matrix interface (k_{sa} in equation 3.4), was taken to be 10^2 to 10^4 times less than the saturated conductivity of the matrix, based on the work of Gerke and van Genuchten (1993a). Since the parameters β, a, γ in equation (3.4) are difficult to define at the field scale, these parameters as well as k_{sa} were lumped into a single parameter. The parameters describing the shape factor of the interface were assumed to be identical to those of the matrix (Kordilla et al., 2012). The residual water content was set to zero for the macro-pore continuum (Kohne et al., 2006) and set equal to 0.07 for the matrix and the interface based on measurements by Mittelbach et al. (2012) at a nearby site. It should be noted that these values do not represent exactly the lysimeter soil and are slightly different. However, they do not adversely affect the description of water flow in the relatively wet water content range evident in the lysimeter. In order to define the upper and lower bounds for the saturated hydraulic conductivity of the dual continuum, the equation suggested by Kohne et al. (2002) was applied:

$$k_f = \frac{k_o - k_m(1 - w_f)}{w_f} \tag{3.10}$$

where k_m and w_f represent hydraulic conductivity of matrix and the ratio of volume of macro-pores to the total pore systems, respectively. k_o in Equation (3.10) is the measured value of the saturated hydraulic conductivity based on soil samples in the laboratory or field.

The value for this parameter was reported by Mittelbach et al. (2012) to vary from 6 to 12 cm d^{-1}, measured within 10 meters of the lysimeter at the same field site. The range of variability for k_f was calculated with the assumption that the macropore fraction varies from 0.01 to 0.1 as the upper and lower bounds, respectively, based on field observations.

3.2.5 Calibration Approach

In order to provide valid initial conditions for the simulation, the model was run for the first 100 days, i.e., March 31st , 2012 to July 8th, 2012 as a warm-up period without any calibration. Following the warm-up period, all three conceptual models were calibrated over the period from July 9th, 2012, to March 30th, 2013, against daily observed evapotranspiration (lysimeter-measured values), lysimeter discharge and water content time series to minimize bias. The chosen calibration period covers a time series with dry and wet conditions in order to include high information content which should lead to a robust parameter estimation for all model concepts. However, we cannot fully make assurance that applying the models under conditions different from the one used for calibration will give exactly the same results. This effect has already been demonstrated in many studies (Coron et al., 2012; Merz et al., 2011; Seiller et al., 2012; Vaze et al., 2010). Using a split sample test could be used to evaluate the model performance for different situations than the calibration period but is not the scope of the paper.

Since the observed water content data at the depths of 5 and 15 cm were very similar, the average was used as the soil moisture data at the depth of 10 cm in order to represent the first layer in C3. The soil moisture content values in the remaining layers were calibrated against observed water content data at depths of 25, 55, and 80 cm. Mertens et al. (2006) demonstrated that using an average moisture content in the calibration enables to identify reliable retention and hydraulic conductivity curves, although local fluctuations at different depths are neglected. Therefore, the average of the observed water content data at depths of 10, 25, 55 and 80 cm were used in the calibration in order to simulate the profile-averaged water content for conceptual models C1 and C2. The evapotranspiration and lysimeter discharge datasets were common observations used for calibrating each conceptual model (C1 to C3). The calibration was carried out using the parameter estimation package PEST [Doherty 2010]. The multi-component objective function which was minimized in the calibration was as follows:

$$\phi = \min\left[\sum_{i=1}^{n} w_i \left(Do_i - Dm_i\right)^2 + \sum_{i=1}^{n} v_i \left(Eo_i - Em_i\right)^2 + \sum_{j=1}^{m}\sum_{i=1}^{n} u_i \left(Wo_{ji} - Wm_{ji}\right)^2\right] \tag{3.11}$$

where the indices of *o* and *m* represent observed and simulated values, respectively. The variables w_i, v_i, u_i are weights that were assigned to lysimeter discharge (*D*), actual evapotranspiration (*E*) and water content observations (*W*), respectively, reflecting their measurement errors. These values were set equal to the inverse of the measurement error for each observation group. The measurement error (precision) of the lysimeter's electronic scale (for measuring storage change) is $\pm$ 100 grams, which corresponds to a water column of approximately 0.032 mm (Seneviratne et al., 2012). The resolution of the seepage tipping bucket is 50 ml, which corresponds to 0.016 mm of the water column. It was assumed that the uncertainty of the lysimeter discharge measurement is also just due to the precision of the measurement device and therefore is not variable over time. TDR sensors have been reported (based on their manuals) to have a measurement error of 2%, which was assumed equal to the coefficient of variation. In order to maintain consistency for different groups of observations, we assumed that measurement errors have normal distributions and are all within one standard deviation from the average. Finally, the weights were assigned to observation groups according to the methodology of Hill and Tiedeman (2007) which correlates the weights with the inverse of measurement error, assumed to be equal to one standard deviation. We note that measurement error of the measuring devices, assumed to be constant over time, was considered as the only source of output uncertainty in this research.

In order to validate the model calibrations, we ran all conceptual models with the calibrated parameters (Table 3-3) from March 31st, 2013 to November 22nd, 2013. To assess the performance of each model, the coefficient of determination (R^2) and centered pattern root mean square error (RMSE) (Taylor, 2001) were calculated. Further, the ratio between the standard deviation of the simulations and the respective observations were explored in order to assess the transient variability between the models and observations. The reason for using more than one statistics is that assessing the performance of a model based on a single statistic can be misleading (Bennett et al., 2013). In order to show all the evaluation metrics in one plot and make the comparison easier, Taylor plots were applied (Figure 3-3). The radial distance from the origin indicates the ratio of the relative standard deviation between the simulated and observed values. If the simulation results correspond perfectly with the observed values, one would expect the simulations to lie on the intersection of a unit-circle (ratio of 1) and the horizontal axis (correlation of 1). Simulations below the unit-circle indicate that the model under-estimates the observed values while those above suggest that the model has a tendency toward over-estimation. Also, the distance of each point on the plot from the point named "Ref" on the horizontal axis indicates a centered pattern RMSE (Taylor, 2001).

3.2.6 Uncertainty Analysis

Following the model performance assessment, uncertainty analyses of lysimeter discharge predictions at different temporal scales were carried out using the PEST package (Doherty et al., 2010). Direct application of Bayes' theorem for computing posterior parameter and predictive probability distributions can be demanding and numerically intensive, especially when applied to physically-based models. In order to avoid such issues and still be able to deal with uncertainty, linearity can be assumed (Brunner et al., 2012; Fienen et al., 2010). Where there is a linear relationship between model parameters and its outputs, equation (3.12) holds :

$$h = Zp + \varepsilon \tag{3.12}$$

where p is the vector of model parameters, Z is the functioning model matrix, h is the vector of outputs and ε is the noise (measurement error) in observed values. In the discussion that follows, the noise is assumed to be normally distributed and there is no auto correlation between them. The noise associated with h, independent of its source, is characterized by a known covariance matrix $C(\varepsilon)$. If we let the scalar s represent a required prediction of the model and the sensitivity of s to all the parameters be encapsulated in the vector y and ignore offsets, equation (3.13) holds :

$$s = y^T p \tag{3.13}$$

Posterior probability distribution of s can be computed as follows based on the work of Fienen et al. (2010) :

$$\sigma^2_{(s/h)} = y^T C(p)y - y^T C(p)Z^T \left[ZC(p)Z^T + C(\varepsilon)\right]^{-1} ZC(p)y \tag{3.14}$$

where $C(p)$ is the prior parameter covariance matrix. A major feature of equation (3.14) is that neither the absolute values of parameters p nor observations h are required. Formulation of equation (3.14) reveals that the only requirements are the sensitivities of new model outputs to model parameters as well as knowledge of the noise of the dataset (Brunner et al., 2012; James et al., 2009; Moeck et al., 2015).

3.3 Results

3.3.1 Calibration and Validation

In order to assess the performance of different conceptual models in simulating evapotranspiration, soil moisture and water discharge, each conceptual model was calibrated individually as described in Section 3.2.5. Table 3-3 includes the individual values of the calibrated parameters for each conceptual model. Figure 3-2 indicates the simulation results of lysimeter discharge, evapotranspiration and average-profile water content variations versus observed values for models C1 to C3 based on calibrated parameters.

Table 3-3. Calibrated values of the parameters used for the three different conceptual models.

Scenario No.	Layer No.	Matrix				Macropores						Evapotranspiration		
		$\alpha(cm^{-1})\times10^{-3}$	n	θ_s	$k_s(cmd^{-1})$	$\alpha(cm^{-1})\times10^{-3}$	n	$k_s(cmd^{-1})$	θ_s	w_f	α_w	C_1	C_2	$C_3\left(cmd^{-1}\right)$
1		3.09	1.33	0.35	5.3	-	-	-	-	-	-	0.24	0.06	1
2		1.58	1.8	0.46	1.7	21.5	2.4	694	0.42	0.01	1.7e-3	0.11	0.06	30
3	1	2.72	3.00	0.58	0.94	10.2	1.5	852	0.6	0.01	7.8e-4	0.09	0.06	29.8
	2	2.99	2.66	0.53	3.5									
	3	2.02	2.25	0.47	3.3									
	4	1.00	2.4	0.39	1.0									

The evapotranspiration parameters are described in Appendix 3-A. The flow parameters are the same as those given in equations (3.1) to (3.9).

Figure 3-2 shows that the model performance improves for all water balance components when more complexity is added. Nevertheless, some spikes in the observed lysimeter discharge cannot be well represented by any of the model setups in C1 to C3. March 8th provides an example and is indicated on Figure 3-2 by a vertical arrow. We attribute this error to neglecting the thawing of frozen water among soil particles, which can cause a sudden release of water. This assumption seems to be valid because the average daily temperature increases gradually from -2.9 °C to 6.4 °C (from March 4th to March 8th). The effect of thawing-freezing cycles in releasing water from the lysimeter is not included in our modeling context as these processes are beyond the scope of this study [interested readers are referred to Mohanty et al. (2014) for more information on freeze-thaw cycle effects on preferential flows]. ET measurements in winter, based on lysimeter weight change, are also noteworthy. While the ET graph (Figure 3-2b) shows that increasing model complexity leads to better ET simulation, particularly during summer, there are some spikes in the observed ET in early winter 2013 which do not seem satisfactorily represented by any of the conceptual models. Considering high values of albedo (measured at the meteorological station at the same site) at these specific dates suggests a snow cover. It has been shown by Gurtz et al. (2003) that winter ETa values, which are measured in the Rietholzbach lysimeter, are problematic (much higher than potential ET). They assigned this problem to a snow bridge in the space between the lysimeter and surrounding area that can perturb correct measurements of lysimeter weight.

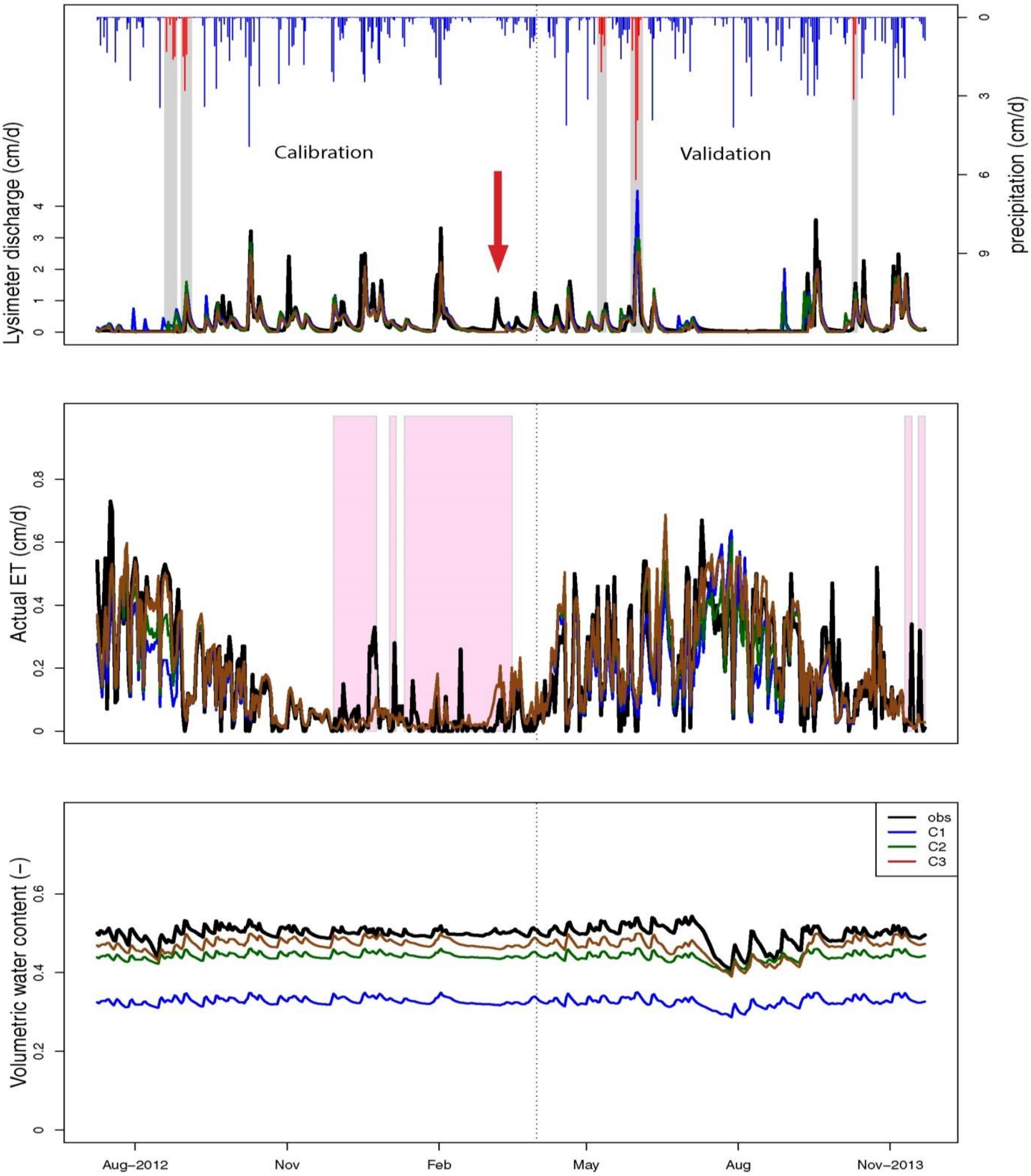

Figure 3-2. Daily observed versus simulated lysimeter discharge (top), evapotranspiration (middle) and average water content (bottom) from 0 to 80 cm below lysimeter surface for all the three conceptual models (C1, C2, C3) for calibration (June 9th, 2012 to March 30th, 2013) and validation (March 31st, 2013 to November 22nd, 2013) periods. The selected events (highlighted in gray) indicate the event-based lysimeter discharge simulations for each conceptual model. Snow cover periods are highlighted in pink. The vertical red arrow indicates the March 8th event (refer to the text for more explanation).

It is also apparent on Figure 3-2c that profile-averaged water contents for all the three scenarios are underestimated. Nevertheless, the agreement between observed and simulated values is good based on R^2 shown in Taylor plots (Figure 3-3). These plots are employed to compare the model performances in simulating the water balance components for all the three scenarios in a more robust way (Figure 3-3).

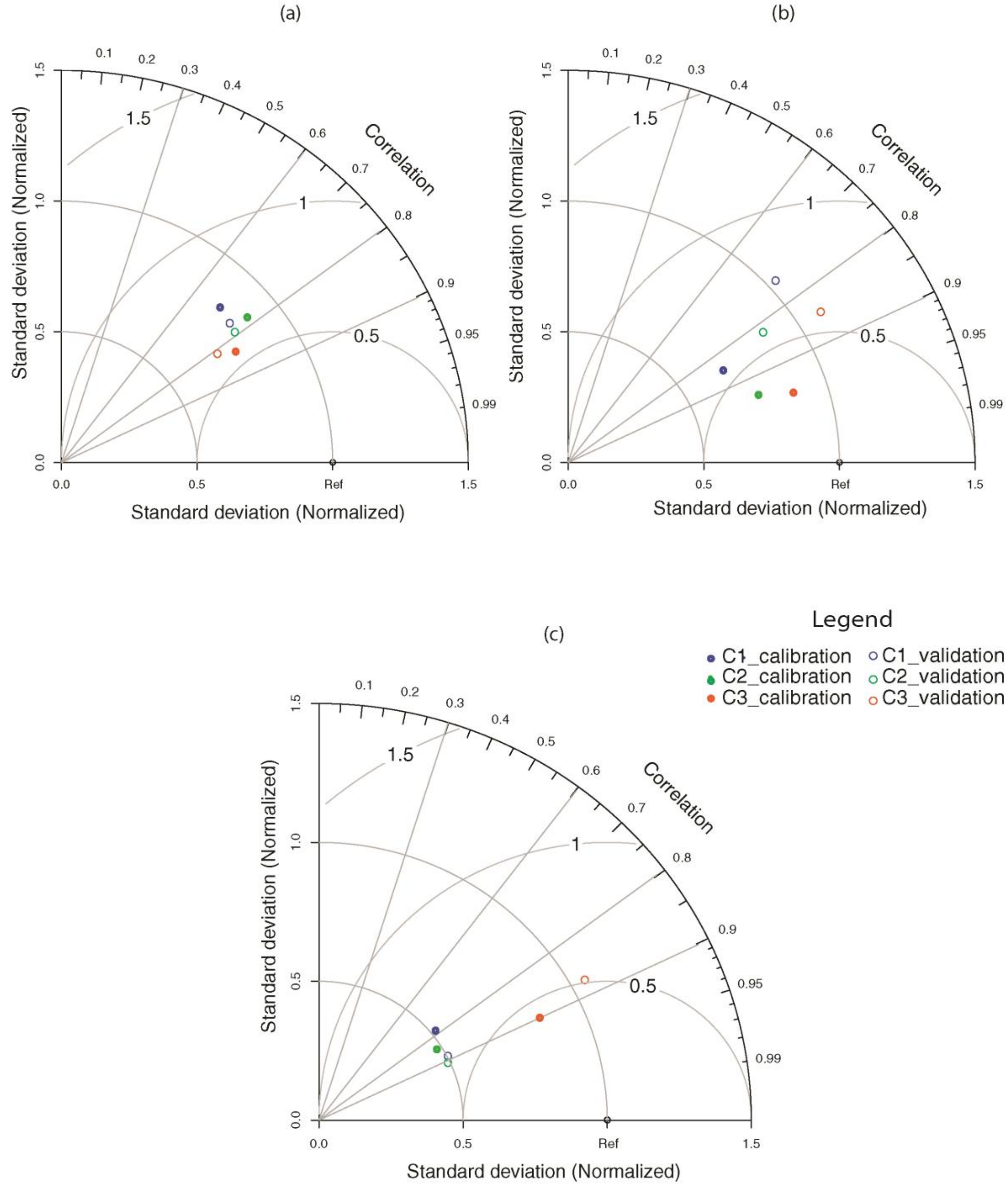

Figure 3-3. Taylor plots for lysimeter discharge (a), evapotranspiration (b) and water content (c) for conceptual models C1 to C3. Different colors indicate different models. The results are shown for calibration (filled circles) and validation (empty circles).

As indicated on Figure 3-3a, all model setups underestimate overall lysimeter discharge. This issue is more pronounced in winter and early spring (Figure 3-2) and could be attributed to the biased precipitation which is part of the input for all models. Gurtz et al. (2003) showed that precipitation (liquid and solid) measurements with the conventional gauge at a standard height in the Rietholzbach are concomitant with large errors due to wind speed effects. They indicated that the monthly correction factors (always positive) are high in late fall, over the entire winter and early spring, starting from +5% in October for rainfall and rising up to +62% in March for snow, with the minimum values in summer and early fall. Therefore it is apparent that the simulated discharge results are likely to be more erroneous (underestimated) in winter than in summer due to inaccurate model input. Figure 3-2c indicates that the WC is continuously underestimated for all the models (the reason for the underestimation will be discussed in more detail in section 3.4). However, the discrepancy between observed and simulated values decrease when we move from model C1 to C3. Generally, it is apparent on the Taylor plots in Figure 3-3 that for all the simulated water balance components, model performance improves as the complexity increases.

3.3.2 Uncertainty Analysis and the Effect of Time Scales

Figure 3-2a shows some highlighted non-snow rainfall events (red with gray bar) which were selected in order to compare the performance of all three conceptual models in reproducing the corresponding discharge values of the events. Figure 3-4 displays the total simulated discharge extending from one day before the event until one day after the event with the 95% confidence interval of each simulated value versus corresponding observed value. Based on assumptions of the normal distribution, the 95% confidence interval is within 1.95 variance of the mean. Therefore, the variances calculated by equation (3.14) were multiplied by 1.95.

It is apparent from Figure 3-4 that model C1 cannot reproduce any of the discharge values within the uncertainty bounds. Figure 3-4 also shows that model C3 represents event based discharge values better than C2 due to i) the simulated values in C3 are closer to observed values except for one event, and ii) the 95% confidence intervals of model C3 simulations include three out of the five observed values, while C2 includes only one. It seems that more complexity is required in order to increase the model performance for event simulations.

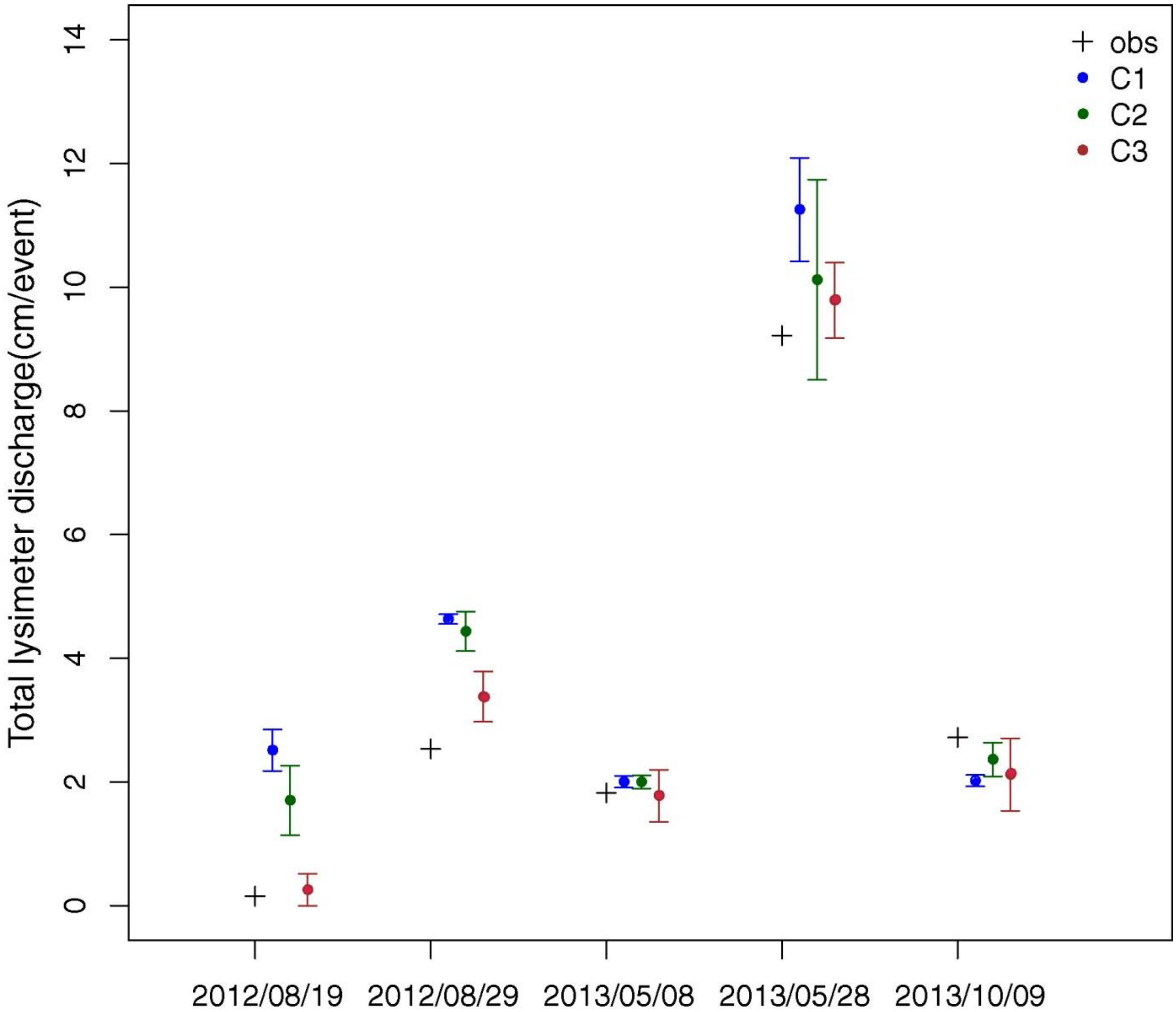

Figure 3-4. Event-based observed versus simulated lysimeter discharge for conceptual models C1, C2 and C3. The events (shown with grey panels on Figure 3.2a) are selected in a way to include varying rainfall intensities. Error bars indicate the 95% uncertainty bounds.

In order to validate the effect of increasing complexity at different time resolutions, we ran all the three calibrated models from October 2002 to June 2005 based on daily inputs, and we calculated the monthly and seasonal discharges. Figure 3-5 displays the total simulated discharge values plus the 95% confidence intervals for each month and season for all the three conceptual models. It is worth noting that while model complexity generally increases from C1 to C2, uncertainty bounds shrink both monthly and seasonally. However, this trend stops when additional complexity is included. In fact, the bounds start to become larger, particularly in summer and fall, when we move from C2 to C3. On the other hand, the lysimeter discharge values are all underestimated in winter and spring for all the models. A discussion about this observation is given in section 3.4.

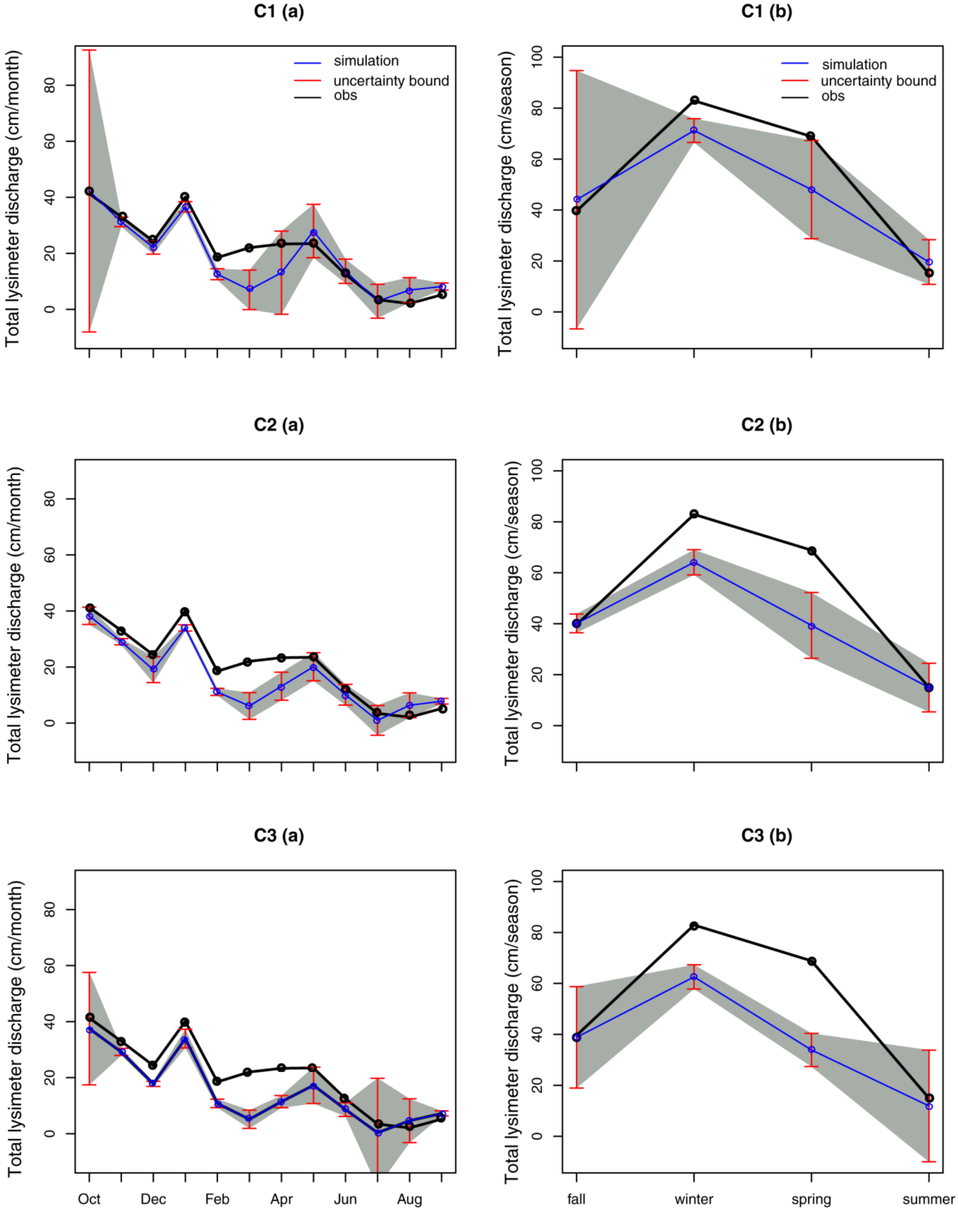

Figure 3-5. Total lysimeter discharge for monthly (a) and seasonal (b) time scales versus observed values from October 2002 to June 2005 based on conceptual models C1, C2 and C3. Error bars indicate the 95% confidence intervals.

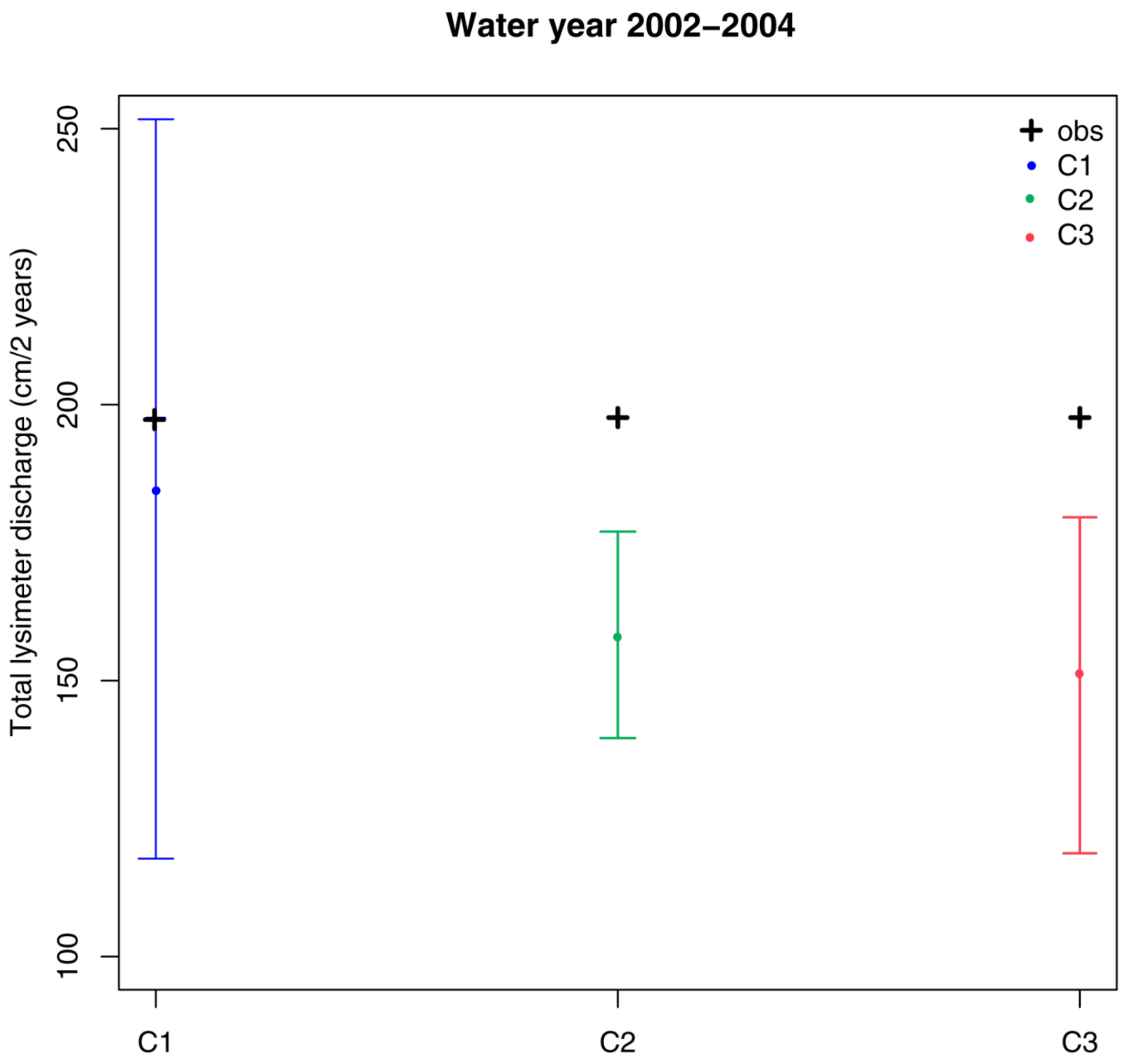

Figure 3-6. Total lysimeter discharge versus observed values from October 2002 to October 2004 based on conceptual models of C1, C2 and C3. Error bars indicate the 95% confidence intervals.

Figure 3-6 displays the estimated total lysimeter discharge for models C1 to C3 from October 2002 to October 2004 (two consecutive water years). In contrast to shorter time scales, C1 outperforms C2 and C3, even though the uncertainty bound in C1 is much larger than in the other two conceptual models. In fact, while C1 underestimates discharge values in winter and spring and overestimates discharge in summer and fall (Figure S 3-1 in supplementary materials), models C2 and C3 underestimate discharge values only in winter and spring (Figure S3-2 and S3-3 in supplementary materials). The reason for under-estimation of discharge in winter and spring based on C2 and C3 was discussed earlier in Section 3.3.1.

3.4 Discussion

Daily comparison of water balance components indicated that the concepts that comprised the preferential flow component, i.e., C2 and C3 outperformed C1 which was formed based just on the existence of matrix flow in the lysimeter. This result is in line with the findings of Lafond et al. (2014) and (Wegehenkel and Gerke, 2015) who showed that ignoring preferential flow for the simulation of daily drainage water flux from a lysimeter led to the biased estimation of the simulated values.

Figure 3-2c and Figure 3-3c clearly indicate the underestimation of profile-averaged water content. This is more pronounced in the calibration period (Figure 3-2c), compared to the validation period. One of the reasons that may explain this underestimation are our measurement devices (TDRs). These devices are sensitive primarily to the typically large fraction of water that has limited mobility (Nimmo, 2012) and therefore they cannot provide reliable information regarding preferential flow. The second reason that may explain the underestimation of water contents is the different weights assigned to the components of the water balance. The weights are inversely proportional to the measurement error of the measuring devices. The weights for WC are one sixth of the discharge weights and one third of the ET weights. Therefore, a better fit for discharge, compared to WC can be expected.

It is worth noting that the uncertainty bounds of the lysimeter discharge become very large for all the models, particularly for C1 and C2, after the event which occurred on 28th of May 2013. This might be the effect of increased rainfall magnitude which has been shown by Munoz et al. (2014) to increase model output uncertainty. However, the bounds of uncertainty for C3 do not change as much as for the other two models. This could be the effect of layered heterogeneity (existence of different textures) which can be as important as structural heterogeneity (existence of macropores) in the model structure. We also note that the uncertainty bounds at daily (event) scale for C3 are generally larger than those of C1 and C2. This can be due to the non-identifiability of some of the model parameters leading to equifinality. Cibin et al. (2010) showed that equifinality becomes apparent when non-identifiable parameters are calibrated during the model calibration. This issue is addressed in details in the next chapter. The uncertainty bounds of the seasonal estimations are also noteworthy (Figure 3-5). The fact that all the models fail at reproducing the lysimeter discharge within the 95% confidence intervals may indicate that this issue is more relative to the uncertainty in input data rather than uncertainty in model structure. On the other hand, the bounds of C3 increase, compared to C2, in summer and fall, when it is believed that the input data is less biased. However, these bounds in C3 either decrease or remain the same as in C2 (winter and spring). We assume that this issue is related to the temporal sensitivity of the parameters. We believe that the lysimeter discharge in winter and spring is more sensitive to preferential flow parameters, while in summer and fall, it is the matrix flow parameters which highly affect the lysimeter discharge estimation. This belief is in line with similar studies which have shown that wetter soils generate more macropore flow (Graham and Lin, 2011; Jarvis, 2007; Kramers et al., 2012; Nimmo, 2012). A global sensitivity analysis can help to find out the sensitivity of our C3 model parameters in spring (wetter) and summer (drier) seasons. A global sensitivity analysis is carried out as the next step in our research to distinguish the temporal sensitivity of model parameters.

We also showed on Figure 3-6 that annual estimation of lysimeter discharge is better reproduced with C1 than C2 and C3. The reason that may explain this fact is that the over-estimations (summer and fall) and under-estimations (spring and summer) of the lysimeter discharge in C1 can cancel the effect of each other to a certain extent and therefore a better match is obtained at annual time scales (Figure S 3-4 in supplementary materials).

3.5 Summary and Conclusion

The aim of this research was to evaluate the effect of increasing model complexity on simulating daily water balance components in addition to exploring prediction uncertainty of recharge estimations, represented by lysimeter discharge, at different time scales. To accomplish the task, the physically-based model HydroGeoSphere was applied to simulate daily water content, evapotranspiration and discharge from a weighing lysimeter with three different conceptual models. The applied model performance metrics indicated that the model performance in reproducing daily water balance components improves as model complexity increases. To validate the effect of increasing complexity on model predictions at different time scales, all three conceptual models were employed at event-based, monthly, seasonal and yearly time scales to estimate the lysimeter discharge. Analysis of the results revealed that the simplest model (C1) is not reliable in reproducing the event-based discharge values, as none of the uncertainty bounds of the simulations included observed values. It was also concluded that moving from C1 to the more complex model (C2) lowers the uncertainty bounds over monthly, seasonal and yearly time scales. On the contrary, increasing model complexity from C2 to the most complex model (C3) leads to an increase in uncertainty bounds of the predictions over all time scales, particularly when input data are less biased. It was also shown that the simplest model (C1) has a better performance than the more complex models (C2 and

C3) over annual cycles. This behavior was due to seasonal errors in the simplest model in which discharge was underestimated in winter and spring while being overestimated in summer and fall. Indeed, the annual simulation results for C1 were balanced. Therefore, if the research aims at estimating only the annual recharge, the analysis can be done with simple models, without including preferential flow. On the other hand, where subsurface flow is the focus of the study, more complex models which include preferential flow are required. Due to the limitations of our study, such as neglecting horizontal flow in the porous medium, and also disregarding the spatial distribution of preferential flow patterns, our findings should be verified at the larger scale.

The four major points of our study are as follows. 1) Our study explored the effect of increasing model complexity for a wide range of time scales. It was shown in our study that different levels of complexities can lead to different model behavior at various time scales. This is an issue which has not been adequately addressed in the literature. 2) In comparison to many similar studies of lysimeter modeling, our study included preferential flow in addition to a very targeted calibration strategy based on high quantity and quality data such as evapotranspiration, discharge and soil moisture. 3) Our study addressed the issue of model complexity with a physically-based model based on real-world data rather than laboratory experiments or synthetic examples. 4) A smaller number of parameters were calibrated in comparison to similar applications of the HGS model. In fact, we tried to keep the sense of physically-based model applications consistent through using the measured values of the parameters, where possible, rather than calibrating them.

Acknowledgement

We would like to thank 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 for financing this study. We would also like to thank Sonia Senevirante, Martin Hirschi and Dominik Michel for providing the data and the discussions over the lysimeter measurements. We are also very grateful to the three anonymous reviewers whose constructive comments helped a consistent manuscript to be out.

Appendix 3-A

HydroGeoSphere (HGS) is a three-dimensional surface-subsurface physically-based numerical model (Therrien et al., 2010). It utilizes a rigorous mass conservative approach that fully couples all key components of the hydrologic cycle based on the following equation:

$$P - I - ET - R - O - \frac{\Delta S}{\Delta t} = 0 \tag{3-A.1}$$

in which P is the precipitation, ET is the evapotranspiration, R is the recharge or deep percolation, O is the overland flow, I is the interception and $\frac{\Delta S}{\Delta t}$ is the water storage over the time step. How components of equation (3-A.1) are calculated in HGS is briefly described in the following.

Interception is the retention of a certain amount of rainfall on the leaves, branches and stems of vegetation. Precipitation exceeding the interception storage and evaporation from the canopy reaches ground surface. The way interception and transpiration are calculated in HGS is based on Panday and Huyakorn (2004) as follows:

$$I_{max} = C_{\text{int}} LAI \tag{3-A.2}$$

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.

$$S_{\text{int}}^{*} = \min(S_{\text{int}}^{\max}, S_{\text{int}}^{0} + P_p \Delta t) \tag{3-A.3}$$

$$E_{can} \Delta t = \min(S_{\text{int}}^{*}, ET_p \Delta t) \tag{3-A.4}$$

$$S_{\text{int}} = S_{\text{int}}^{*} - E_{can} \Delta t \tag{3-A.5}$$

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 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:

$$T_P = f_1(LAI) * f_2(\theta) * RDF * \left[ET_P - E_{can}\right] \tag{3-A.6}$$

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

$$f_1(LAI) = max\{0, min[1, (c_2 + c_1 LAI)]\} \tag{3-A.7}$$

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:

$$RDF = \frac{\int_{z_1}^{z_2} rf(z)dz}{\int_0^{L_r} rf(z)dz} \tag{3-A.8}$$

In equation (3-A.8) $rf(z)$ 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 (3-A.6) relates transpiration to the soil moisture and is expressed as:

$$\begin{bmatrix} 0 \text{ if } 0 \le \theta \le \theta_{wp} \\ \text{f}_3 \text{ if } \theta_{wp} \le \theta \le \theta_{fc} \\ 1 \text{ if } \theta_{fc} \le \theta \le \theta_o \\ \text{f}_4 \text{ if } \theta_o \le \theta \le \theta_{an} \\ 0 \text{ if } \theta_{an} \le \theta \end{bmatrix} \tag{3-A.9}$$

$$f_3 = 1 - \left[\frac{\theta_{fc} - \theta}{\theta_{fc} - \theta_{wp}}\right]^{c_3} \tag{3-A.10}$$

$$f_4 = 1 - \left[\frac{\theta_{an} - \theta}{\theta_{an} - \theta_0}\right]^{c_3} \tag{3-A.11}$$

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_o is the moisture content at oxic limit and θ_{an} is the soil moisture content at anoxic limit. The logic supporting equation (3-A.9) 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 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:

$$E_s = \alpha(E_P - E_{can})[1 - f_1(LAI)]EDF \qquad (3\text{-A.12})$$

where EDF is the evaporation density function and is assumed to be effective from 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 (3-A.12) is a moisture dependent coefficient which varies between zero and one and is expressed as:

$$\begin{bmatrix} 0 \text{ if } \theta < \theta_{e2} \\ \dfrac{\theta - \theta_{e2}}{\theta_{e1} - \theta_{e2}} \text{ if } \theta_{e2} \le \theta \le \theta_{e1} \\ 1 \text{ if } \theta > \theta_{e1} \end{bmatrix} \qquad (3\text{-A.13})$$

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.

The a priori assumption in the calculation of ETa is that it cannot exceed the potential evapotranspiration, regardless of the type of vegetation considered. Therefore, an accurate estimation of ET_p is of high importance. In this research, the Penman method (Penman, 1948) was applied in order to compute ET_p.

In order to define LAI, we assumed that LAI grows linearly from its minimum to maximum during each cutting period. The grass on the top of the lysimeter was cut four times from June to October 2012 with an average time interval of 45 days. The range of variability for leaf area index of the lysimeter crop was extracted from *Scurlock et al. (2001)*.

3.6 Supplementary Data

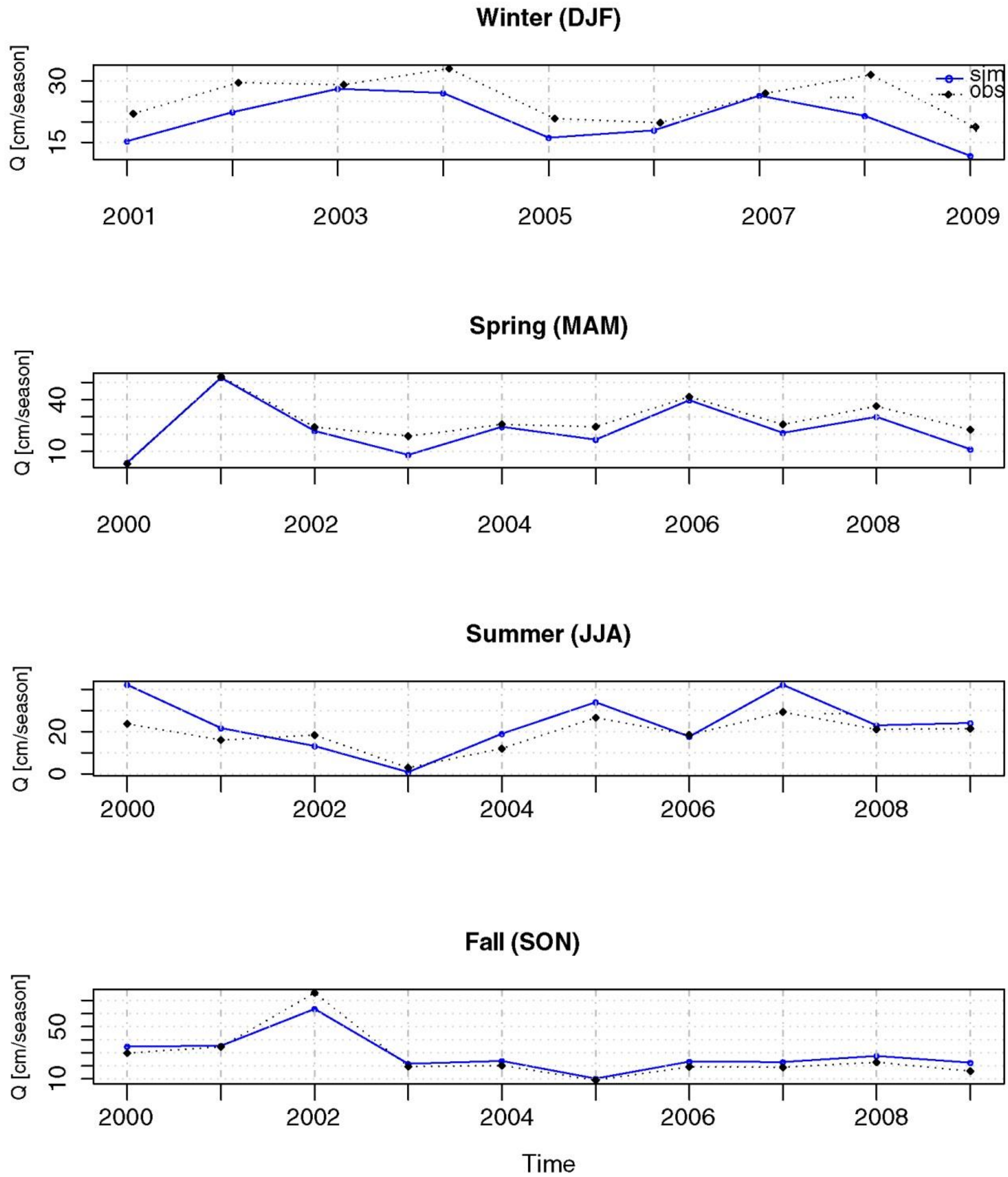

Figure S 3-1. Seasonal versus observed variation of lysimeter discharge based on C1 from January 2000 to December 2009 (DJF: December, January, February; MAM: March, April, May; JJA: June, July, August; SON: September, October, November).

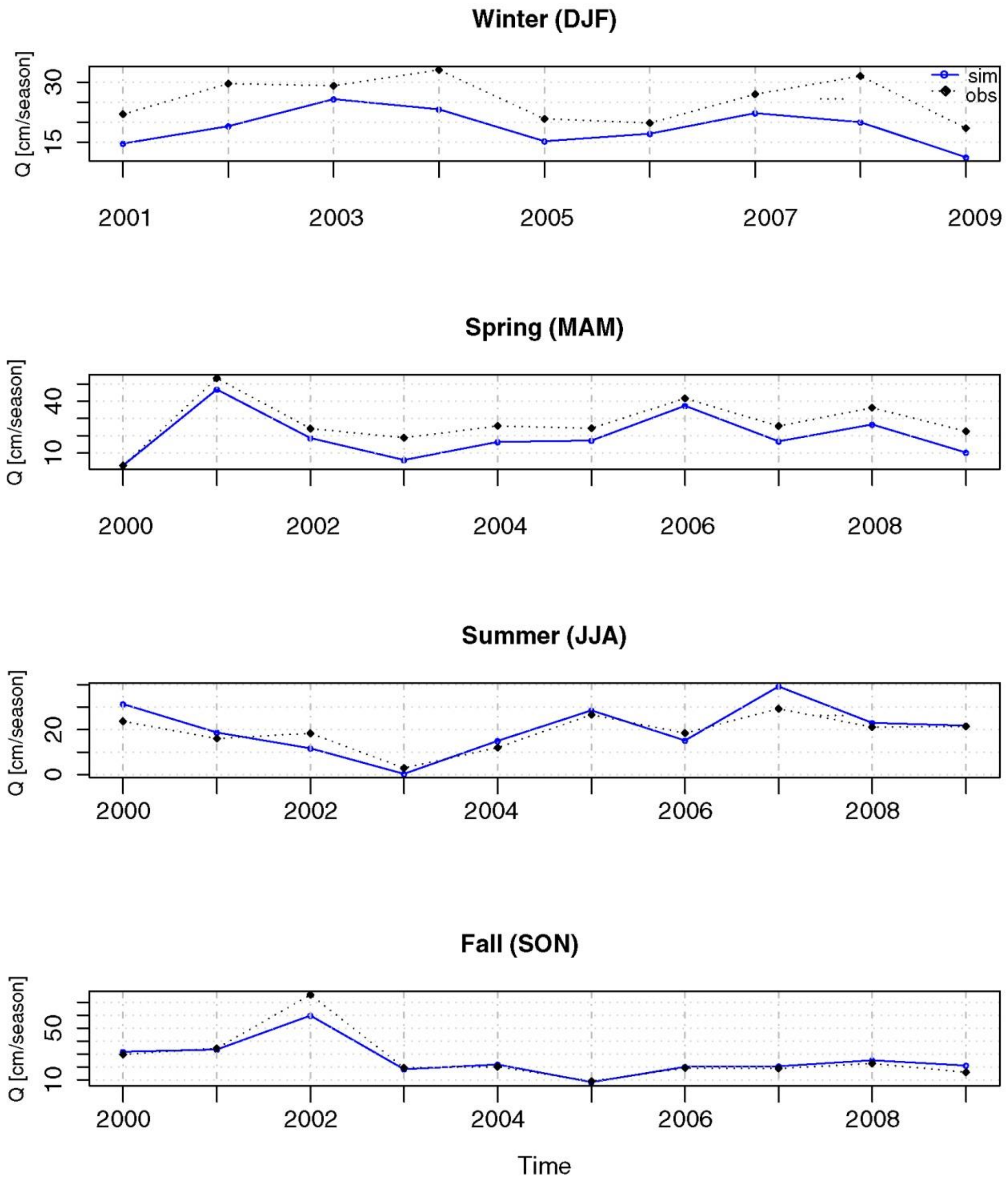

Figure S3-2. Seasonal versus observed variation of lysimeter discharge based on C2 from January 2000 to December 2009 (DJF: December, January, February; MAM: March, April, May; JJA: June, July, August; SON: September, October, November).

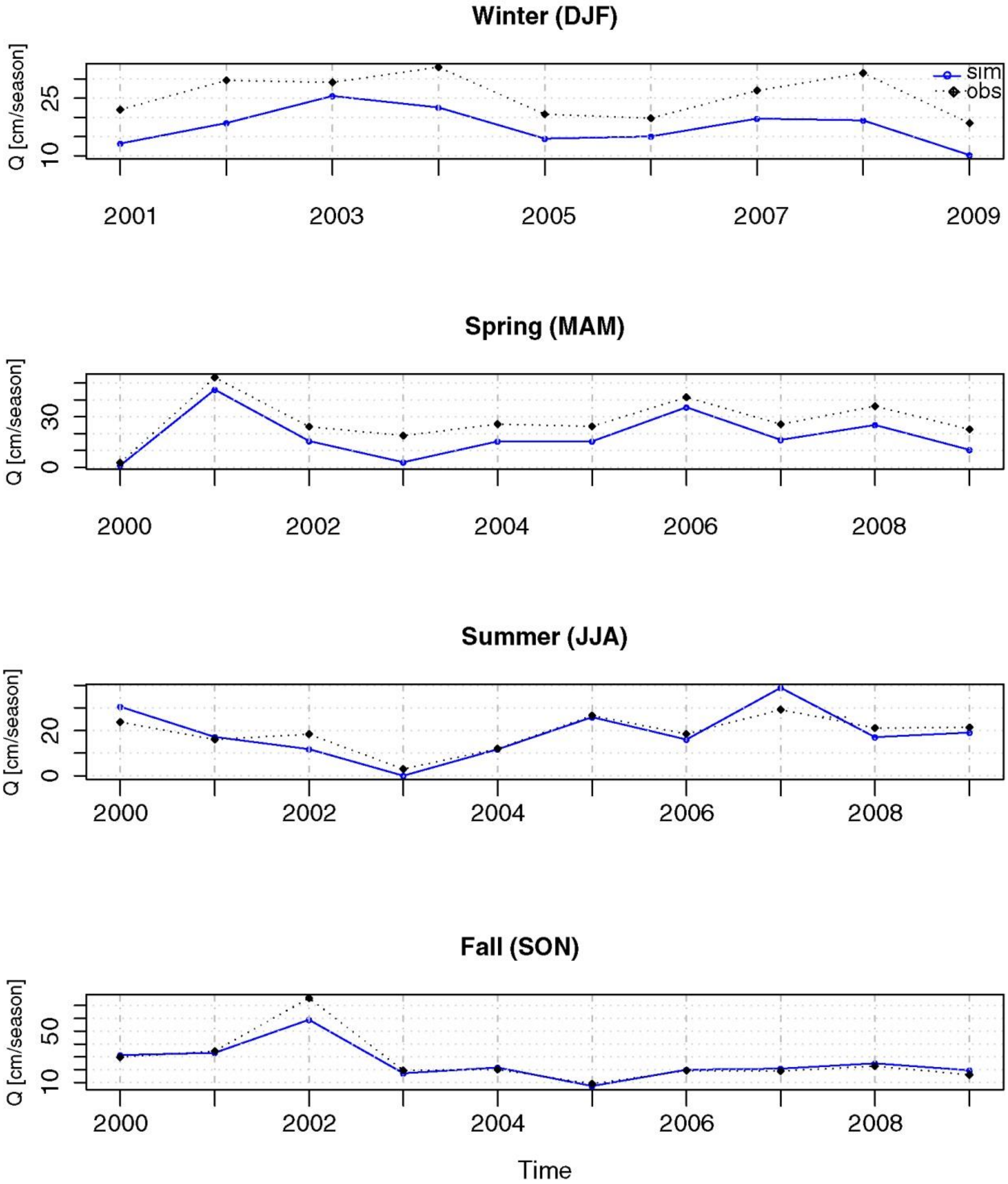

Figure S 3-3. Seasonal versus observed variation of lysimeter discharge based on C3 from January 2000 to December 2009 (DJF: December, January, February; MAM: March, April, May; JJA: June, July, August; SON: September, October, November).

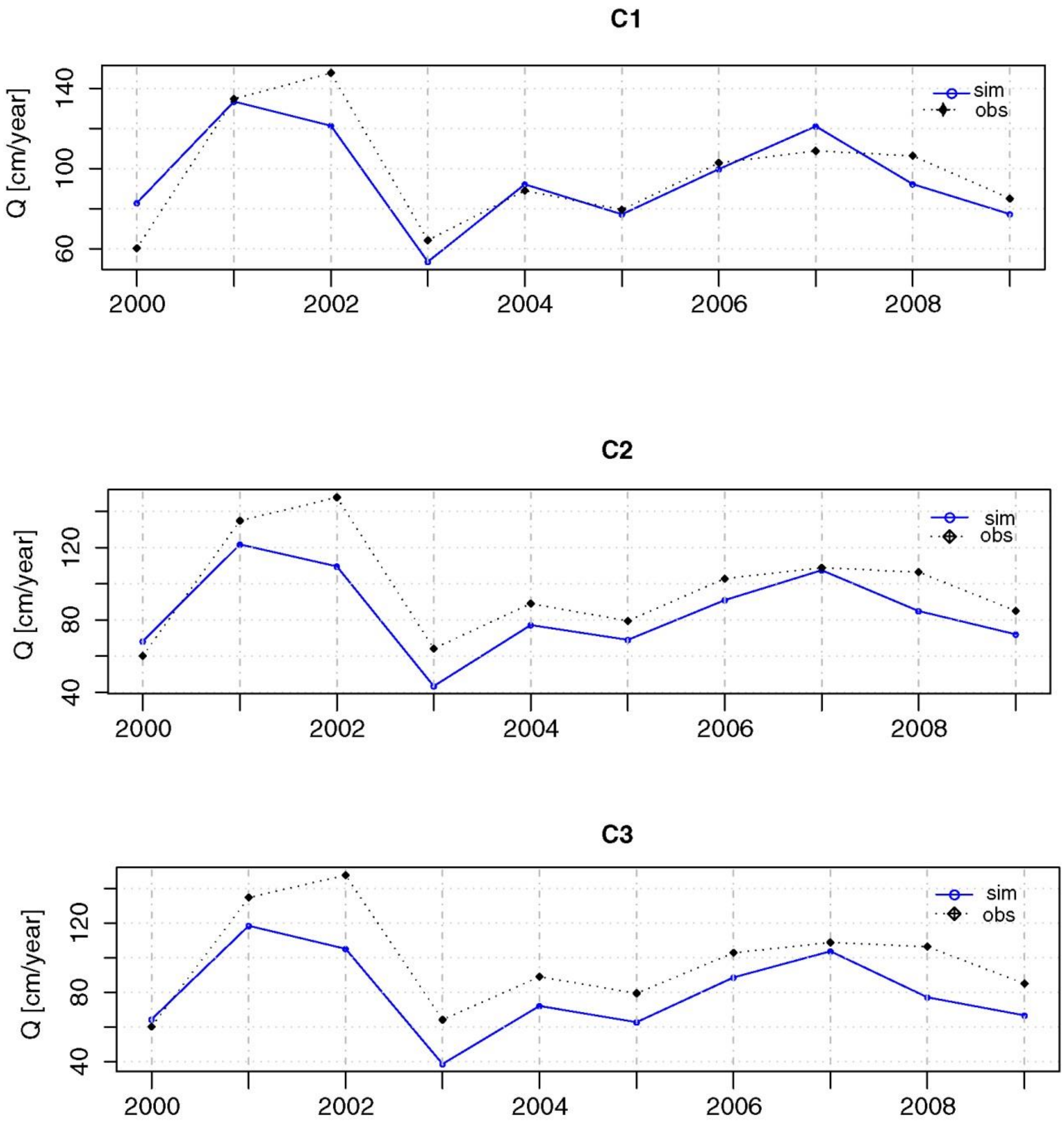

Figure S 3-4. Annual observation versus simulation of lysimeter discharge from January 2000 to December 2009 for C1, C2 and C3.

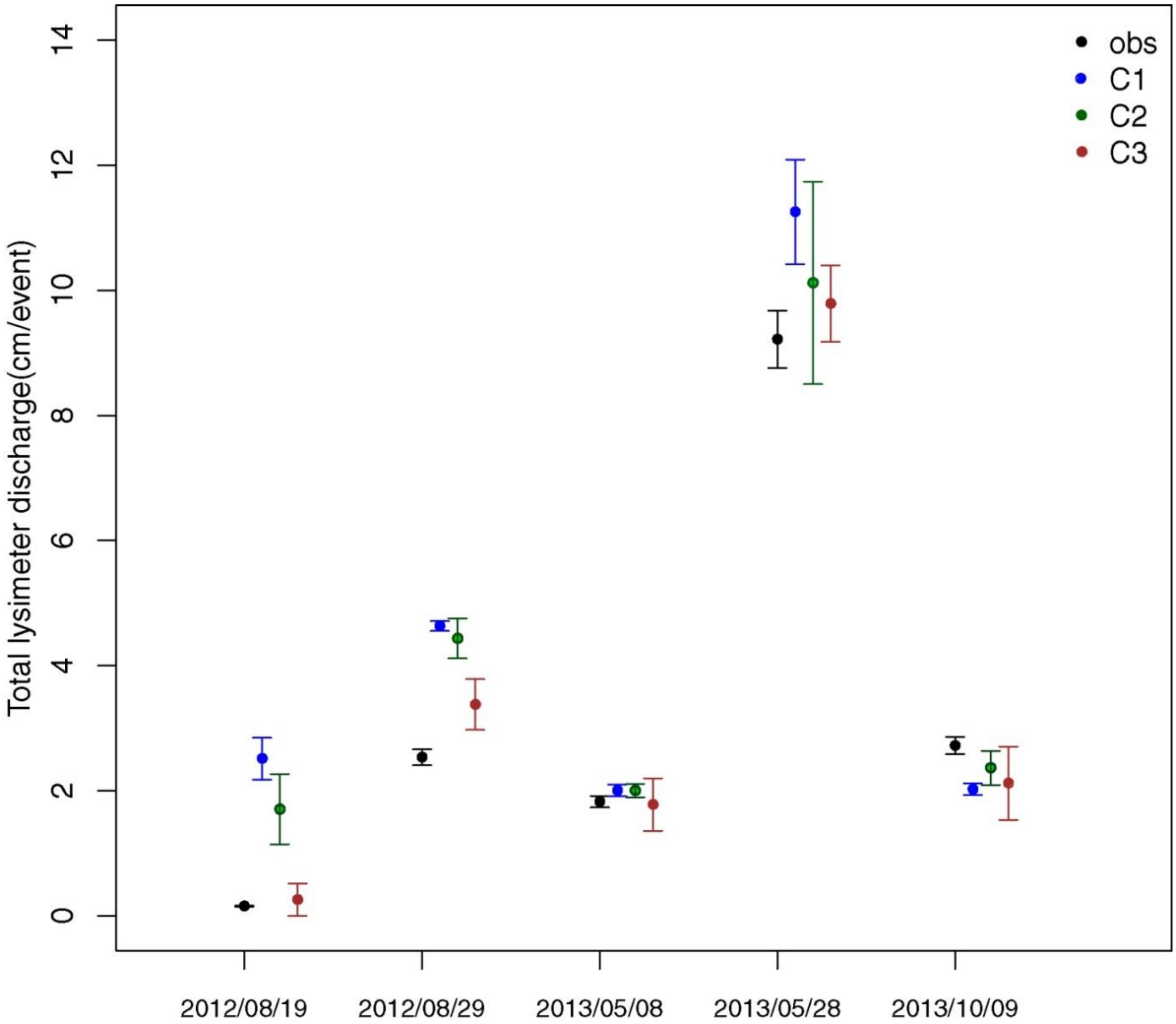

Figure S 3-5. Event-based observed (including measurement uncertainty) versus simulated lysimeter discharge for conceptual models C1, C2 and C3. Error bars indicate the 95% uncertainty bounds.

The relative accuracy of the discharge measurements varies with time and can rise up to ±5% for discharge measurements. However, due to not knowing the exact values, it was assumed in this research that the measurement uncertainty was just due to the measurement error of the measuring devices (precision) and therefore it was constant over time. To make sure that this assumption does not violate the conclusions, the maximum error (i.e., ±5%) of the measured lysimeter discharge was considered for the comparisons of event discharge simulations. The results, indicated above show that only the error bars of the observation for the third event from left overlap with the error bars of the C1 simulations. In fact, C1 still has the weakest performance in comparison to the other conceptual models which comprised the preferential flow component and therefore adding measurement uncertainty does not change the drawn conclusions.

Chapter 4

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

Submitted to the journal of Environmental Modelling & Software

Ghasemizade, M., Baroni, G., Abbaspour, K., Schirmer, M. Combined analysis of time-varying sensitivity and identifiability indices to diagnose the response of complex environmental models

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 temporal sensitivity analysis (TSA) using the variance-based method of Sobol, and the temporal identifiability (TIA) of model parameters, using the dynamic identifiability method (DYNIA) and discussed the relationship between the two analyses. HydroGeoSphere was used to simulate daily evapotranspiration, water content, and seepage at the lysimeter scale. We found that identifiability is a necessary but not sufficient condition for a parameter to be used for uncertainty reduction. The comparison of the two analyses revealed 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.

4.1 Introduction

The advances in computer science have supported 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 as well as incorporating all the relevant processes in space and time. 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 (see Beven, 2006). Among others, parametrization is recognized as a crucial step for a proper model application. On the one hand, a large number of parameters usually required by the 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, several studies have shown the excellent capability of complex models to describe the processes but with the cost of high uncertainties in prediction when the model is extrapolated beyond the calibration period (e.g., Laloy et al., 2010).

Diagnostic tools are often applied to overcome these limitations (Gupta et al., 2008; Matott et al., 2009; Uusitalo et al., 2015). They are used to explore the input-output response of the models and thus they give a better overview of the model behavior. One of the main popular methods currently applied is the variance-based global sensitivity analysis (GSA). This analysis separates the contribution of each individual parameter to the output variance as well as the interacting contributions of the parameters to the outputs. It incorporates the influence of input parameters over the whole range of variation, and is independent of model linearity and model monotonicity (Saltelli et al., 2010). In environmental modeling, this analysis has been recognized as an integrated step of the model application giving relevant information on the behavior of the model response (Shin et al., 2015; Song et al., 2015). GSA has been used for complex environmental models (Nossent et al., 2011), to investigate inadequacy in the model structure (Herman et al., 2013) and to incorporate all the possible sources of uncertainty (Baroni and Tarantola, 2014). GSA has been shown to be the only existing approach that provides a meaningful global measure of interacting parameters (Razavi and Gupta, 2015). Analysis at different temporal scales has also been recognized as a useful tool. This application 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 help to identify the dominant processes and conditions for deriving more accurate estimations (Guse et al., 2014; Reusser and Zehe, 2011).

In general, global sensitivity analysis has two main features. The first feature helps to identify the parameters that are not important (not sensitive) to model outputs. The aim of this feature is to reduce the number of free-varying and less influential parameters by removing or assigning constant values to them. These parameters have minimum potential to reduce the output uncertainty (factor fixing). It has to be noted that this feature helps to reduce the complexity of calibration but it does not guarantee the reduction of the output uncertainty. For example, van Werkhoven et al. (2009) performed GSA 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 these parameters from the calibration, their model results had essentially the same predictive performance as the non-fixed parameter case. For this reason, the second feature of GSA is to identify the important parameters (sensitive) in the model output response (factors prioritization) and for that the focus is on the parameters that have the potential to maximally reduce the output uncertainty and to improve the performance of the model output (Saltelli et al., 2006).

Another relevant diagnostic method, which is used to explore the input-output response of a model, is the so-called identifiability analysis (IA). This analysis helps to derive the maximum likelihood values of the parameters, given available observations. For this purpose, several methods have been developed (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 (Matott et al., 2009; Shin et al., 2015). The dynamic identifiability analysis (DYNIA) method developed by Wagener et al. (2003) is one of the methods specifically developed for the identifiability analysis. It has been widely used due to being relatively immune from the effects of model nonlinearity on parameter estimates and also the ease of use of its employed algorithm. The advantage of 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. The reason is that DYNIA calculates the objective function for running windows (time periods) that are shorter than the entire simulation time.

It is a necessary, but not sufficient condition that parameters must be sensitive to be identifiable (Guse et al., 2014; Reusser and Zehe, 2011). While sensitivity and identifiability analysis are relevant and commonly used for the study of model behaviors, the two analyses are rarely compared in literature. In most of the studies, identifiability of model parameters have been discussed in terms of their sensitivity, assuming that the chances of parameter identifiability increases as the parameter becomes more sensitive (Shin et al., 2013). Cibin et al. (2010) showed that sensitivity and identifiability analysis provide complementary information to model behavior, even if parameters identifiability was evaluated just by visual exploration.

This paper aims at a better evaluation of the relationship and dependency between the two analyses. We conduct a detailed study of a model response based on: 1) a temporal GSA (the study of how the uncertainty in the output of the model can be apportioned to different inputs is the primary focus here); 2) TIA to extract the maximum information content from available observations (quantifying the temporal identifiability of model parameters versus different datasets is the focus in this part); finally, 3) we discuss the relationship between TSA and TIA results. We applied the physically-based model HydroGeoSphere (HGS) (Therrien et al., 2010) to represent a complex and nonlinear system in which the links between formulation and behavior are difficult to discern a priori. The simulations were carried out in the framework of a weighing lysimeter, based on the analysis of actual evapotranspiration, lysimeter discharge and water content data.

4.2 Methods

4.2.1 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).

4.2.2 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 Richard's 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 (1993a), 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 4.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 (ETp), 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.

4.2.3 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:

$$v = \sum_i v_i + \sum_{i<j} v_{ij} + \sum_{i<j<m} v_{ijm} + \ldots v_{12..z} \tag{4.1}$$

Where

$$v_i = v\left(E\left(Y \middle| X_i = x_i^*\right)\right) \tag{4.2}$$

$$v_{ij} = v\left(E\left(Y \middle| X_i = x_i^*, X_j = x_j^*\right)\right) - v\left(E\left(Y \middle| X_i = x_i^*\right)\right) - v\left(E\left(Y \middle| X_j = x_j^*\right)\right) \tag{4.3}$$

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 $E\left(Y \middle| X_i = x_i^*\right)$ 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:

$$S_i = \frac{v_{Xi}\left(E_{X\sim i}\left(Y \middle| X_i\right)\right)}{v(Y)} \tag{4.4}$$

$$S_{Ti} = \frac{E_{X\sim i}\left(v_{Xi}\left(Y \middle| X_{\sim i}\right)\right)}{v(Y)} \tag{4.5}$$

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 0 to 1 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.

Table 4-1. 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

Sensitivity indices in equations (4.4) and (4.5) 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 (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 4-1. Total number of simulations required for Sobol computation is $(k+2)n$, 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).

It is worth noting that sample based methods of sensitivity analysis are all prone to uncertainty due to the limitation in the number of samples (Pappenberger et al., 2006). This issue may lower the reliability of SA results and therefore needs to be addressed carefully before making any interpretations. In order to analyze the adequacy of the sample size, we increased the number of samples for each parameter by increments of 50. Bootstrapping (Rindskopf, 1997) was applied at each increment to compute the 95% confidence intervals. Figure 4-1 indicates the convergence of the main and total effects for two parameters of the model with acceptable confidence intervals. Similar results apply to the other parameters of the model (not shown).

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).

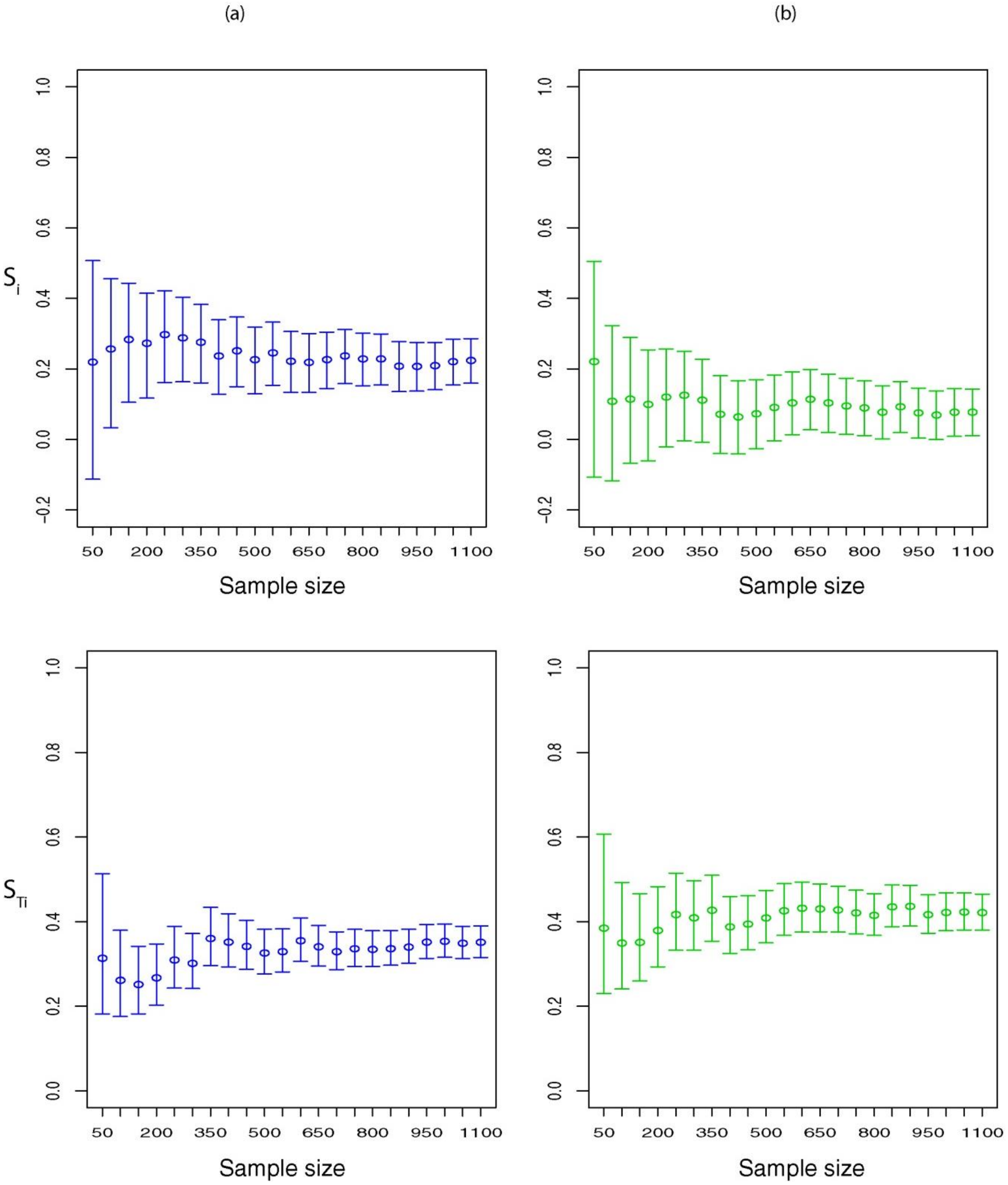

Figure 4-1. Convergence of the estimated sensitivity indices with increasing sample size and 95% confidence bounds. (a) Main (Si) and Total (STi) effects for the parameter 3 (α_w); (b) Main (Si) and Total (STi) effects for the parameter 10 (K_{s4}). Refer to the Table 4-2 to see the description of parameters.

4.2.4 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, which were employed for SA to do the IA. RMSE of daily observed versus simulated values were used. In particular, at each time step, the algorithm first selects *m* time steps before and after the current time step (i.e., a window size of *2m+1*) and then calculates the average model error (RMSE) for the specified time frame. 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. It varies in the range of zero to one. This facilitates the comparison of results obtained from SA. The same window sizes used for SA were also used for IA to keep consistency between the two analyses. IA was performed in the Matlab environment based on the SAFE toolbox (Pianosi et al., 2015).

4.3 Results

4.3.1 Global sensitivity analysis

Figure 4-2 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 4-2a). 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 4-2b). This finding is in line with Ghasemizade et al. (2015) who found, based on the comparison of three model concepts, the necessity of including preferential flow component for daily simulation of seepage (lysimeter discharge). The presence of a visual correlation between the interacting sensitivity of parameter 3 (α_w) and smoothed variation of water content suggests that soil moisture has a control on the contribution of preferential flow to discharge over different hydrologic conditions. The present results also suggest the difficulties to calibrate the preferential flow parameters (α_w, α) due to their interactions 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).

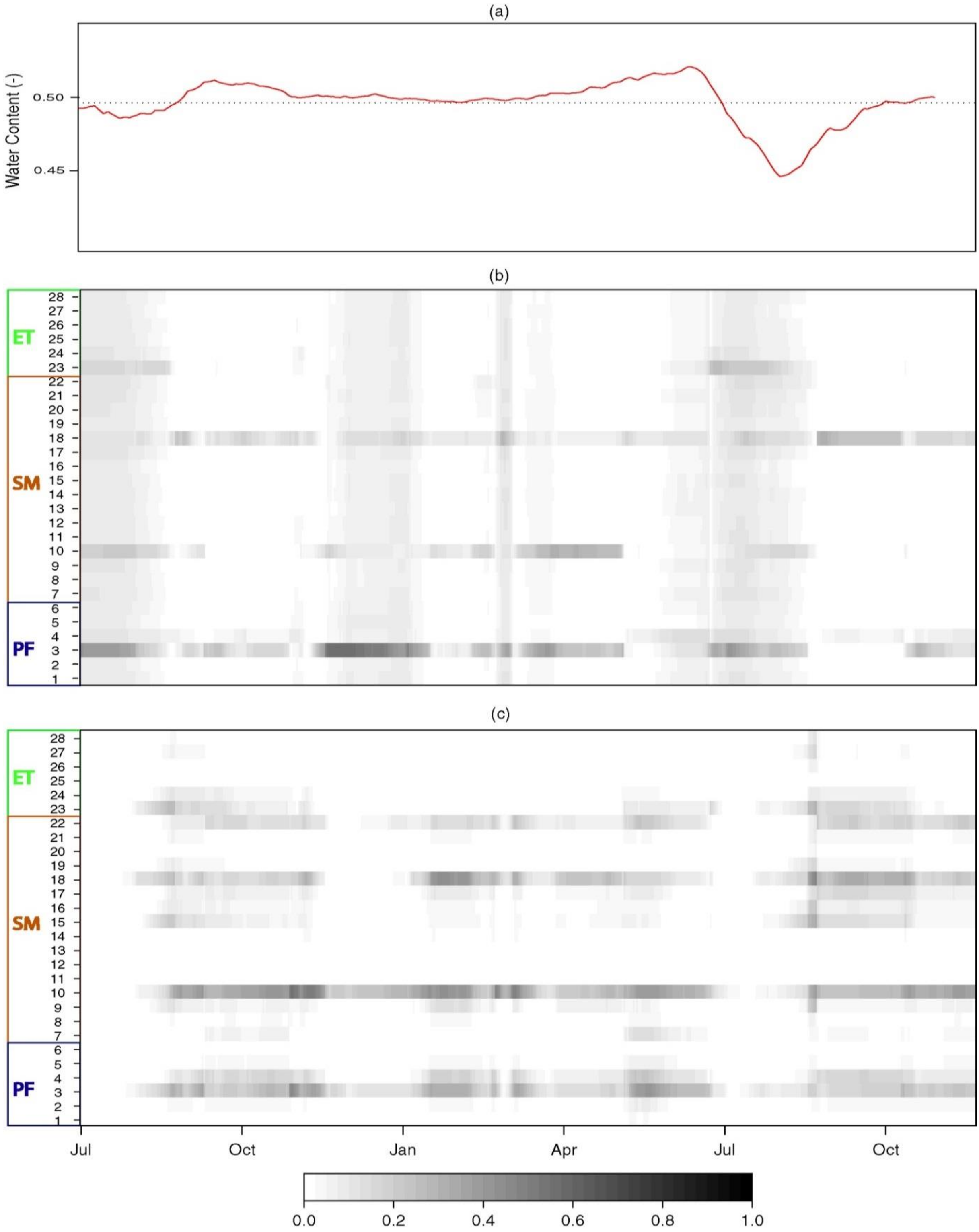

Figure 4-2. Time-varying daily sensitivity of the discharge RMSE metric to model parameters within the simulation period, i.e., July 2012 to November 2013; (a) smoothed average water content with a window of 47 days. The dotted line indicates the average of the water content for the hydrologic year of 2012-2013; (b) main effect of model parameters; (c) interaction effect of model parameters; PF, SM and ET show groups of preferential flow, soil matrix and evapotranspiration parameters, respectively.

Parameters 10 (k_{s4}) and 18 (α_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 parameters 10 and 18 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 on Figure 4-2 is that 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. This 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. However, 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 (Wagener et al., 2003).

Figure 4-3 indicates the results of the temporal sensitivity analysis of simulated daily ET versus observed values. It is not surprising that the parameters 23 (C1) and 24 (C2), 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 4-3b). However, they show a different temporal dynamic for different water content conditions. The main effect of parameter 24 (C2) 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 (C1) 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 (C3) 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 4-3c). 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. This could be due to the relatively small window size (5 days) for evapotranspiration, in comparison to other studied model outputs. The temporal sensitivity of the model parameters of water content versus observed values with a window size of 35 days are displayed on Figure 4-4.

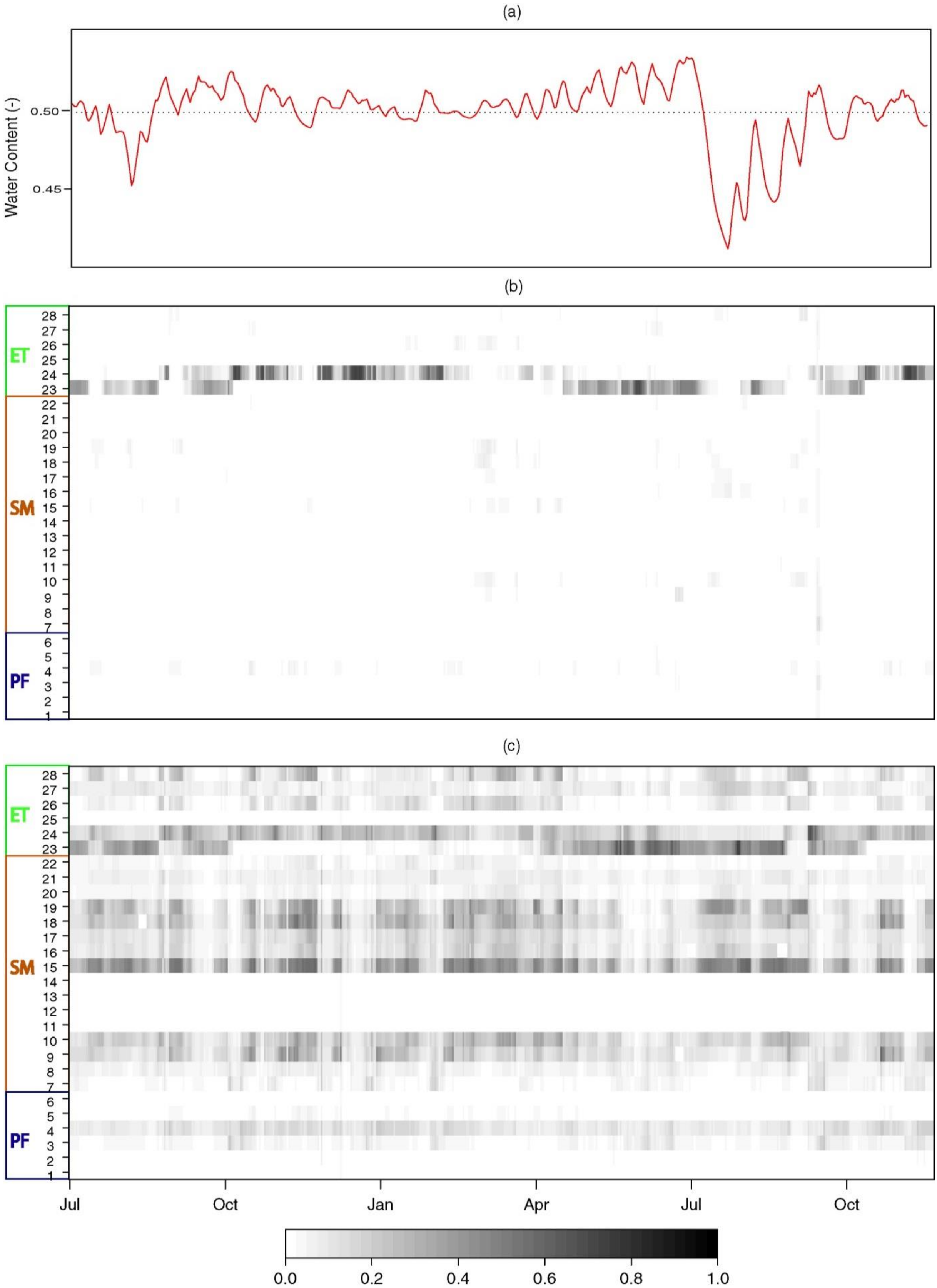

Figure 4-3. Time-varying daily sensitivity of the evapotranspiration RMSE metric to model parameters within the simulation period, i.e., July 2012 to November 2013; (a) smoothed average water content with a window of 5 days. The dotted line indicates the average of water content for the hydrologic year of 2012-2013; (b) main effect of model parameters; (c) interaction effect of model parameters; PF, SM and ET show groups of preferential flow, soil matrix and evapotranspiration parameters, respectively.

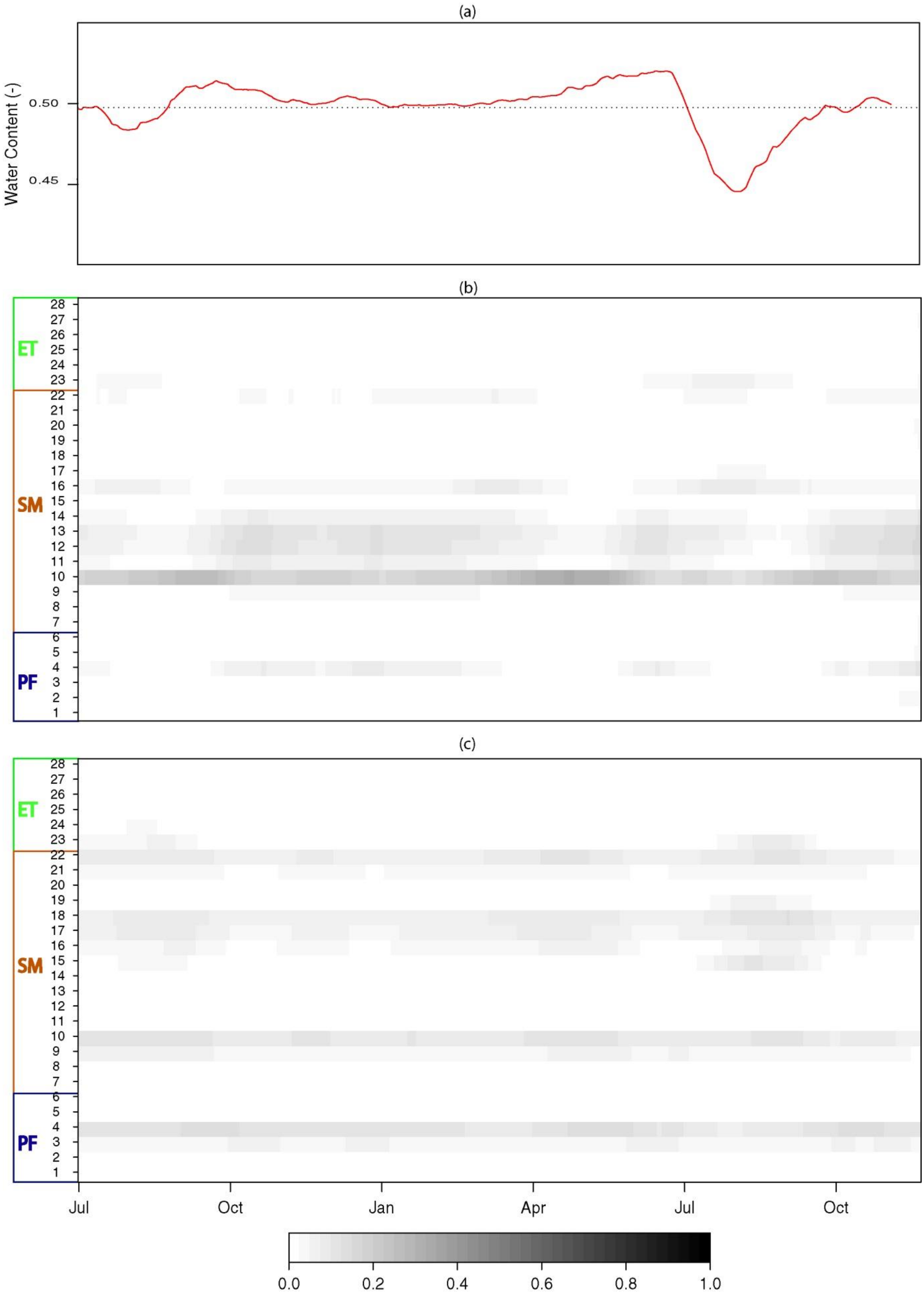

Figure 4-4. Time-varying daily sensitivity of the water content RMSE metric to model parameters within the simulation period, i.e., July 2012 to November 2013; (a) smoothed average water content with a window of 35 days. The dotted line indicates the average of water content for the hydrologic year of 2012-2013; (b) main effect of model parameters; (c) interaction effect of model parameters; PF, SM and ET show groups of preferential flow, soil matrix and evapotranspiration parameters, respectively.

It is clearly visible on Figure 4-4b that parameter 10 (k_{S4}) has the biggest 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. It is noteworthy that the interactions of these four parameters with other parameters are very close to zero (Figure 4-4c), implying their independent and crucial role in affecting the water content. It is also clear on Figure 4-4c 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). Overall, comparing Figure 4-2, 4-3 and 4-4 indicates that different parameters are sensitive to different outputs. This issue suggests the necessity of including multiple datasets for parameter calibration.

4.3.2 Temporal identifiability analysis

The information content of all the 28 parameters, based on the different studied outputs, is shown in Figure 4-5. 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 4-5a).

Figure 4-5b indicates the variable information content of the model parameters versus the lysimeter discharge. The results show a visual correlation between the information content of parameter 3 (α_w) and the variation of water content. 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.

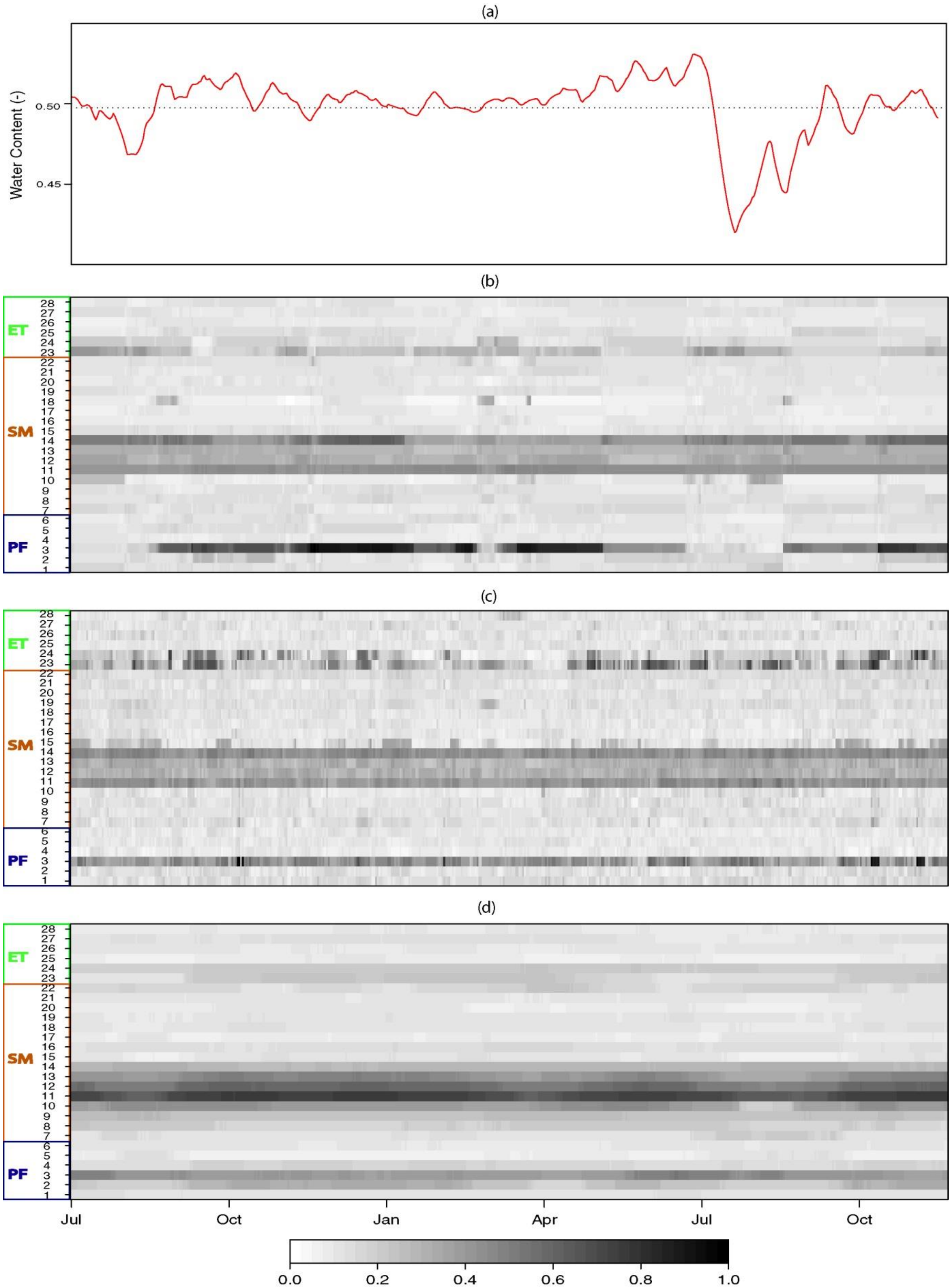

Figure 4-5. Smoothed average water content with a window of 10 days for the simulation period, i.e., July 2012 to November 2013 in addition to the annual average of the water content (the dotted line) for the hydrologic year of 2012-2013 (a). Time-varying daily information content of the discharge RMSE (b), evapotranspiration RMSE (c) and water content RMSE metric (d) for model parameters. PF, SM and ET show groups of preferential flow, soil matrix and evapotranspiration parameters, respectively.

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 4-5c. It is discernable that parameters 23 (C1) and 24 (C2) 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 4-5d) is slightly different than the other information content plots. While porosity values of matrix flow in layers 1 to 4 ($\theta_{S1}, \theta_{S2}, \theta_{S3}, \theta_{S4}$) 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 4-5b to Figure 4-5d reveals the fact that relatively few parameters are identifiable and that the different datasets (discharge, evapotranspiration, water content) are contributing most equally to identifying the model parameters.

4.3.3 Relation between TSA and TIA

The analysis presented so far shows a complex behavior of the model parameters in comparison to the different outputs investigated. The sensitivity analysis shows how the parameters with high main effects vary when different outputs are considered (e.g., discharge: parameter 3; evapotranspiration: parameters 23 and 24; water content: parameter 10). This seems to contradict the results obtained with the identifiability analysis where almost the same parameters (i.e., parameter 3, 10:14, 23 and 24) were found to have relatively high information content for each model output. To further explore this behavior, sensitivity indices and information content for all the considered outputs obtained at each time step are plotted in Figure 4-6. On the one hand, we observe (Figure 4-6a) that there is a tendency

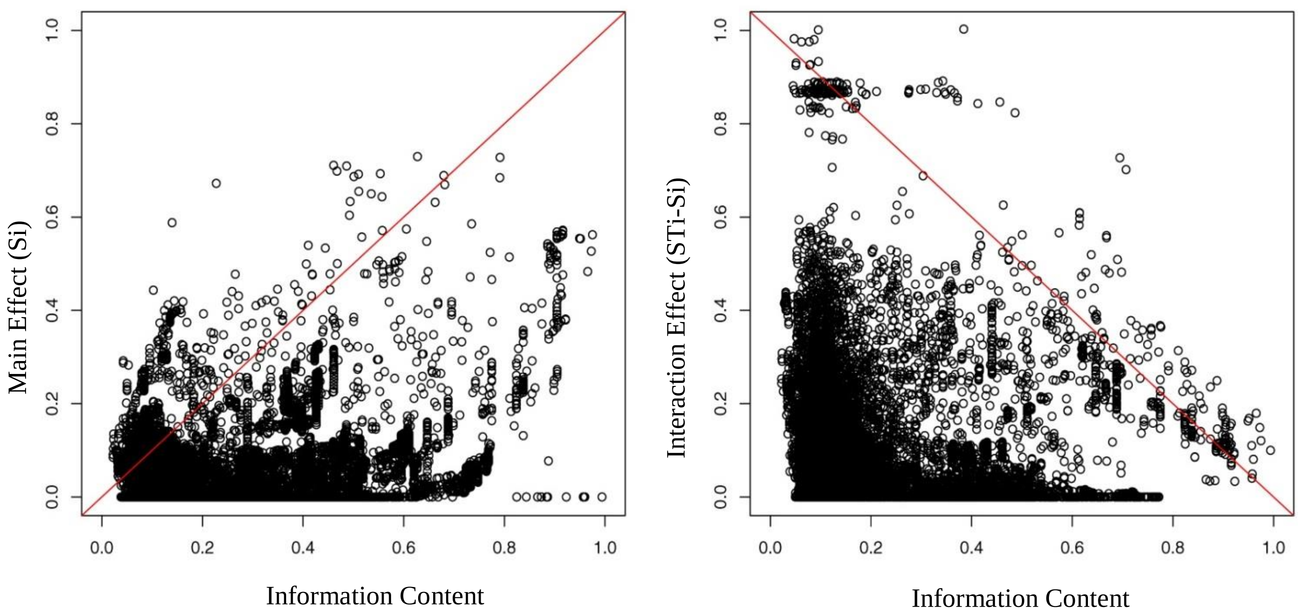

Figure 4-6. Information content of discharge, evapotranspiration and water content versus (a) main effects (b) interaction effects for all the model parameters for the entire simulation period, i.e., July 2012 to November 2013.

The comparison between information content and interaction (i.e., total effect – main effect) is shown in Figure 4-6b. It is noteworthy that an inverse type of relationship, in comparison to the one described before, is apparent. 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 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 this sense the results show how low interaction is a necessary but not sufficient condition for a parameter to be identifiable.

4.4 Discussion

Figure 4-3 clearly indicates that evapotranspiration is not sensitive to parameter *C3* as opposed to *C1* and *C2*. Three reasons can be discussed in that sense; time scale of parameter importance (window time), objective function, and hydrologic conditions. Massmann et al. (2014) analyzed the sensitivities of a rainfall-runoff model parameters at different time scales (window sizes) and found out that the window size 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. The other factor that can significantly impact the identifiability and sensitivity of a parameter is the choice of model performance metric (Gupta et al., 2008;

Wagener et al., 2009; Zhang et al., 2013). In this study, running-window RMSE was used as the only model metric since the focus of our study was to highlight the relationship between identifiability and sensitivity rather than focusing on each individually. Hydrologic conditions have also been shown to affect the sensitivity of model parameters (Linhoss et al., 2013; van Werkhoven et al., 2008). In a comparison study, Herman et al. (2013) applied three rainfall-runoff models different in concepts and formulations and found out that the parameters associated with evapotranspiration generally dominate under dry conditions. The time frame of our analyses (June 2012-Nov. 2013) is assumed to be normal to wet in terms of hydrologic conditions and therefore the simulated period does not include dry or extreme dry conditions (year 2003). This fact may have hindered the impact of *C3* on evapotranspiration.

Assuming that the model parameters should be sensitive and identifiable to decrease the uncertainty in the model outputs, the analysis and the comparison presented in this study show how both sensitivity indices should be considered for a proper assessment of the model response. This issue has not been addressed clearly in literature due to the fact that just the main effect or just the total effect has been used to identify the important parameters for calibration. It is worth noting that factors prioritization assumes that the parameter (factor), which will lead to the greatest reduction in output variance can be fixed to its true value. This assumption rarely holds because it is very unlikely to obtain the true value of any uncertain factor (Allaire and Willcox, 2012). We showed how identifiability of the parameters is a necessary but not sufficient condition for a parameter in order to reduce the uncertainty in the model output. In addition, we presented that the parameters should have high main effects as well as low interactions to increase the likelihood of identifiability and reduce the uncertainty of the model output.

Considering the specific model application (HGS), it is interesting to note that some model parameters such as 10 (k_{S4}) and 18 (α_4) show a relatively high level of interaction for all the outputs investigated. Thus, these parameters are not well determined and therefore are not identifiable. On the other hand, only parameters 3, 23, 24 (α_w, *C1*, *C2*) meet the two requirements explained above (i.e., high main effects and low interactions). Moreover, they meet these requirements only under some specific conditions (i.e., water content conditions). For example, parameter 3 (α_w) meets all the requirements when the soil moisture is close to saturation (December to May). This finding is in line with Ghasemizade et al. (2015) who applied the same modeling concept for the same case study and found that the output uncertainty is significantly higher in summer and fall than winter and spring. Finally, relatively few parameters ($k_{sm}, w_f, \varphi, \theta_{S1}, \theta_{S2}, \theta_{S3}, \theta_{S4}, C3$) have in general low interaction effects, suggesting that most of the parameters are important for the simulation of the processes considered. In this sense, 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 features of the environmental system but they have the severe difficulties of parametrization, leading to their reduced predictive capabilities.

4.5 Summary and conclusion

This paper investigates the temporal sensitivity and identifiability of model parameters in providing a framework for reducing the dimensionality of 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 their values 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 available information to adjust model parameters to more (high) certain values. Note that our analysis were based on one type of objective function (RMSE) and did not address the effect of different window sizes. Also, our modeling approach was based on 1D vertical flow in the variably saturated zone and did not include the effect of horizontal flow. Nonetheless, the major points of our study are:

- We explored the relationship between temporal sensitivity of model parameters and their corresponding identifiability values. This issue has been rarely addressed in environmental modeling.
- An attempt was made here to relate the temporal sensitivity and identifiability values with the change in the soil moisture as an index for describing the status of the system.
- Our study addressed the effect of preferential flow on all water balance components rather than focusing on just discharge. Identification of the parameters of the dual permeability concept with evapotranspiration data is very worthwhile since the seepage (discharge in our work) data is not easily available.
- 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").

Appendix 4-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:

$$\frac{\partial \theta_f}{\partial t} = \frac{\partial}{\partial z}\left(k_f \frac{\partial h_f}{\partial z} - k_f \right) - \frac{\Gamma_w}{w_f} \tag{4-A.1}$$

$$\frac{\partial \theta_m}{\partial t} = \frac{\partial}{\partial z}\left(k_m \frac{\partial h_m}{\partial z} - k_m \right) + \frac{\Gamma_w}{1-w_f} \tag{4-A.2}$$

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 :

$$\Gamma_w = \alpha_w (h_f - h_m) \tag{4-A.3}$$

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 (1993a) given by :

$$\alpha_w = \frac{\beta}{a^2} \gamma k_{sa} \tag{4-A.4}$$

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:

$$S_w = S_{wr} + (1 - S_{wr})\left[1 + |\alpha h|^n\right]^{-m} \tag{4-A.5}$$

$$K_r = S_e^{(\ell_p)}\left[1 - \left(1 - S_e^{\frac{1}{m}}\right)^m\right]^2 \tag{4-A.6}$$

$$m = 1 - \frac{1}{n} \tag{4-A.7}$$

$$S_e = \frac{(S_w - S_{wr})}{(1 - S_{wr})} \quad (4\text{-A.8})$$

$$S_w = \frac{\theta}{\theta_S} \quad (4\text{-A.9})$$

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: (4-A.10)

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

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.

$$S_{int}^{*} = min(S_{int}^{max}, S_{int}^{0} + P_p \Delta t) \quad (4\text{-A.11})$$

$$E_{can} \Delta t = min(S_{int}^{*}, ET_p \Delta t) \quad (4\text{-A.12})$$

$$S_{int} = S_{int}^{*} - E_{can} \Delta t \quad (4\text{-A.13})$$

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:

$$T_P = f_1(LAI) * f_2(\theta) * RDF * [ET_P - E_{can}] \quad (4\text{-A.14})$$

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

$$f_1(LAI) = max\{0, min[1, (c_2 + c_1 LAI)]\} \quad (4\text{-A.15})$$

C1 and *C2* are dimensionless fitting parameters. The features of vegetation root are summarized in RDF which is a root distribution function and is described as:

$$RDF = \frac{\int_{z_1}^{z_2} rf(z) dz}{\int_{0}^{L_r} rf(z) dz} \quad (4\text{-A.16})$$

In equation (4-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 (4-A.14) relates transpiration to the soil moisture and is expressed as:

$$\begin{bmatrix} 0 \text{ if } 0 \le \theta \le \theta_{wp} \\ f_3 \text{ if } \theta_{wp} \le \theta \le \theta_{fc} \\ 1 \text{ if } \theta_{fc} \le \theta \le \theta_o \\ f_4 \text{ if } \theta_o \le \theta \le \theta_{an} \\ 0 \text{ if } \theta_{an} \le \theta \end{bmatrix} \tag{4-A.17}$$

$$f_3 = 1 - \left[\frac{\theta_{fc} - \theta}{\theta_{fc} - \theta_{wp}} \right]^{c_3} \tag{4-A.18}$$

$$f_4 = 1 - \left[\frac{\theta_{an} - \theta}{\theta_{an} - \theta_0} \right]^{c_3} \tag{4-A.19}$$

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_o is the moisture content at the oxic limit and θ_{an} is the soil moisture content at the anoxic limit. The logic supporting equation (4-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:

$$E_s = \alpha^* (E_P - E_{can})[1 - f_1(LAI)]EDF \tag{4-A.20}$$

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 (4-A.20) is a moisture dependent coefficient which varies between zero and one and is expressed as:

$$\begin{bmatrix} 0 \text{ if } \theta < \theta_{e2} \\ \dfrac{\theta - \theta_{e2}}{\theta_{e1} - \theta_{e2}} \text{ if } \theta_{e2} \le \theta \le \theta_{e1} \\ 1 \text{ if } \theta > \theta_{e1} \end{bmatrix} \tag{4-A.21}$$

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.

Chapter 5

5.1 Summary and Conclusion

This PhD thesis focused on seeking the optimum level of model complexity as well as investigating the worth of different observation datasets for representing vadose zone processes and water balance components at the scale of a lysimeter. An attempt was made to compare different levels of complexities through increasing the number of parameters and processes in vadose zone. Moreover, parameter and output uncertainty as well as evaluating the worth of different datasets to constrain the values of model parameters were performed.

To address the objectives of the thesis, three conceptual models for simulating the water content (WC), evapotranspiration (ET) and discharge (D) in a weighing lysimeter were developed, using the physically-based model HydroGeoSphere. Conceptual Model 1 (C1) considered the lysimeter as a homogeneous zone with matrix flow, while conceptual Model 2 (C2) comprised an added preferential flow component to evaluate the effect of adding extra processes. Conceptual Model 3 (C3) included layered heterogeneity in addition to the matrix and preferential flow components and included four soil layers in order to evaluate the effect of increasing parameters. All the models were calibrated versus observed data to optimize the values of unknown parameters. In order to assess the effects of increasing complexity on output uncertainty, the 95% confidence bounds of discharge values, which were simulated based on the three concepts, were analyzed at the time scales of days, months, seasons and years.

In the second step, the global and variance-based sensitivity method of Sobol was applied on C3 to partition the uncertainty in WC, ET and D into the individual uncertainty of the input parameters. This analysis was performed temporally to see to what extent the parameters can impact the model performance in varying times. Also, in order to realize when the available datasets can highly constrain the range of variations of the parameter values , a temporal identifiability analysis was performed, using the dynamic identifiability analysis approach (DYNIA).

Using the methodology explained above, the main findings of this thesis are summarized as follows:

- Adding complexity increases the model performance measured by R^2 and RMSE metrics for all the outputs (WC, EC, D). Our analyses indicated that the performance metrics improved when the complexity increased from C1 to C3 in calibration and validation periods.
- Layered heterogeneity and preferential flow are required to be included in vadose zone modeling if it aims at simulating the flow processes at the time scale of an event. It was shown in this thesis that C3 had a better performance than C1 and C2 in simulating the events discharges due to the fact that more of the C3 uncertainty bounds could capture the observed values.
- Generally, adding layered heterogeneity exacerbates the reliability of discharge predictions at the time scales of months and seasons, due to wider uncertainty bounds. The investigation of the uncertainty bounds at the times where less biased precipitation data (summer and fall) were available revealed that the uncertainty bounds of C3 were significantly bigger than uncertainty bounds in C2.
- Simple models can compete with more complex models at reproducing discharge values on annual time scales. In this thesis, C1 underestimated discharge values in winter and spring and overestimated discharge in summer and fall for a time period of ten years (2000 to 2009). Therefore, the over- and under-estimations of lysimeter discharge in C1 balanced each other to an extent and consequently a better match was obtained at annual time scales.
- The sensitivity of the simulated outputs (water balance components) to the individual model parameters varies in time, independent of the considered model output. This implies that the dominant processes are time dependent. It was found in this work that preferential flow (parameter 3 of C3) was dominant in reproducing the lysimeter discharge when the soil moisture was close to saturation.
- Both sensitivity indices of the Sobol analysis (main effects and interaction effects) should be considered to identify the important parameters when the modeler aims at reducing the dimension of model parameters for calibration and output uncertainty reduction. Having high main effects and low interaction effects increase the likelihood of parameter identifiability and therefore allows reduction in the output uncertainty.
- High temporal variability in sensitivity and identifiability analyses indicates the importance of taking into account different conditions (i,e., wetness, precipitation)

in the model evaluations. It was shown in our analyses that soil moisture affects temporal variability of sensitivity and identifiability.

5.2 Outlook

Proper management of water resources requires good knowledge of water balance components at different scales. This Ph.D. thesis is one step towards better evaluation of physically-based and integrated models and provides methods and diagnostic tools that may facilitate the challenging application of such models within the context of pre-alpine environments. We indicated in this Ph.D. thesis that HydroGeoSphere as a physically-based and integrated model is a valuable tool to better understand the hydro(geo)logical processes and their interactions in a closed system. However, the model application and the outputs of this research motivated some other research questions. In the following, limitations of the work in this thesis as well as recommendations for future research are briefly discussed.

Calibration and uncertainty analysis: The algorithm of the parameter estimation code (PEST), which was employed in this thesis to optimize (calibrate) the unknown values of parameters and to perform the uncertainty analysis is based on Gauss-Marquardt-Levenberg. This approach is a local optimization method and does not guaranty that it searches the model space adequately. This issue becomes more important when one notes that the linear uncertainty method used in PEST is implicitly dependent on the optimized values. In fact, the sensitivity matrix (Jacob matrix), which is embedded in the formulation of linear uncertainty analysis in PEST, calculates the sensitivity of model outputs to the parameters at their optimized values. Therefore, it is highly recommended that a modeler repeats the analysis using a number of widely different realizations of parameter values for calculation of the sensitivities that underpin linear analysis. It was also assumed in this thesis that the assumptions of the linear uncertainty analysis i.e., randomness of the residuals as well as their zero autocorrelation holds. The modeler should use diagnostic tools to check for the validity of the assumptions and use non-linear uncertainty approaches such as GLUE (Beven and Binley, 2014), which can be used for informal likelihood measures, if the assumptions do not hold.

Performance metrics: The application of a performance metrics for complexity comparisons (chapter 2) were restricted to R^2 and RMSE while for sensitivity and identifiability analyses (chapter 3) only RMSE was employed. These metrics are statistical metrics and therefore in order to generalize the conclusions that were drawn based on the analyses, it is highly recommended to use other metrics such as water balance objective function (Wagener et al., 2009) as well.

Window size: It is also worth noting that the identifiability and sensitivity analyses were performed based on fixed window sizes, due to the lack of a clear guideline for the window sizes. However, it is worth noting that different window sizes may reveal different aspects of a model behavior (Massmann and Holzmann, 2012; Massmann et al., 2014) and therefore it is suggested to repeat the analyses with varying time windows.

Input uncertainty: In this thesis, output uncertainty (chapter 3) and parameter uncertainty (chapter 4) were investigated as the main sources of uncertainty in our modeling. In fact, it was assumed that the input uncertainty is zero and therefore the precipitation values were used without any corrections. However, it has been shown by many that precipitation values measured at standard gauges are biased due to wind speed as the main source of errors (Mekonnen et al., 2015; Seibert and Moren, 1999; Sevruk et al., 1991). These biases can be minimized by measuring liquid precipitation at ground level due to close-to-zero speed of wind at ground level (Mekonnen et al., 2015), although this approach is not efficient for solid precipitation due to the movement of snow drift to the gauge when wind blows. Therefore, correction of wind-induced precipitation measurements are recommended to include more events for comparisons of model complexities. It is noteworthy that the precipitations which were used for events modeling (chapter 2) were selected at periods where the maximum hourly wind speeds were lower than 4 m s^{-1}.

Bibliography

Abaspour, K.C., Sonnleitner, M.A., Schulin, R., 1999. Uncertainty in estimation of soil hydraulic parameters by inverse modeling: Example lysimeter experiments. Soil Sci Soc Am J, 63(3): 501-509. DOI:10.2136/sssaj1999.03615995006300030012x

Abdou, H.M., Flury, M., 2004. Simulation of water flow and solute transport in free-drainage lysimeters and field soils with heterogeneous structures. Eur J Soil Sci, 55(2): 229-241. DOI:10.1046/j.1365-2389.2004.00592.x

Abdul, A.S., Gillham, R.W., 1984. Laboratory Studies of the Effects of the Capillary-Fringe on Streamflow Generation. Water Resour Res, 20(6): 691-698. DOI:10.1016/0022-1694(89)90177-7

Abdul, A.S., Gillham, R.W., 1989. Field Studies of the Effects of the Capillary-Fringe on Streamflow Generation. J Hydrol, 112(1-2): 1-18.

Allaire, D.L., Willcox, K.E., 2012. A variance-based sensitivity index function for factor prioritization. Reliab Eng Syst Safe, 107: 107-114. DOI:10.1016/j.ress.2011.08.007

Ameli, A.A., Craig, J.R., McDonnell, J.J., 2015. Are all runoff processes the same? Numerical experiments comparing a Darcy-Richards solver to an overland flow-based approach for subsurface storm runoff simulation. Water Resour Res, 51(12): 10008-10028. DOI:10.1002/2015wr017199

Andersen, J., Dybkjaer, G., Jensen, K.H., Refsgaard, J.C., Rasmussen, K., 2002. Use of remotely sensed precipitation and leaf area index in a distributed hydrological model. J Hydrol, 264(1-4): 34-50. DOI:10.1016/S0022-1694(02)00046-X

Anderson, S.P. et al., 1997. Subsurface flow paths in a steep, unchanneled catchment. Water Resour Res, 33(12): 2637-2653. DOI:10.1029/97wr02595

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

Augenstein, M., Goeppert, N., Goldscheider, N., 2015. Characterizing soil water dynamics on steep hillslopes from long-term lysimeter data. J Hydrol, 529 795-804. DOI:10.1016/j.jhydrol.2015.08.053

Bachmair, S., Weiler, M., 2011. New dimensions of hillslope hydrology. In: Levia, D., Carlyle-Moses, D., Tanaka, T. (Eds.), Forest Hydrology and Biogeochemistry.

Bae, D.H., Jung, I.W., Lettenmaier, D.P., 2011. Hydrologic uncertainties in climate change from IPCC AR4 GCM simulations of the Chungju Basin, Korea. J Hydrol, 401(1-2): 90-105. DOI:10.1016/j.jhydrol.2011.02.012

Baillie, M.N., Hogan, J.F., Ekwurzel, B., Wahi, A.K., Eastoe, C.J., 2007. Quantifying water sources to a semiarid riparian ecosystem, San pedro river, Arizona. J Geophys Res-Biogeo, 112(G3). DOI:10.1029/2006jg000263

Bakker, M., Bartholomeus, R.P., Ferre, T.P.A., 2013. "Groundwater recharge: processes and quantification" Preface. Hydrol Earth Syst Sc, 17(7): 2653-2655. DOI:10.5194/hess-17-2653-2013

Baroni, G., Tarantola, S., 2014. A General Probabilistic Framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study. Environ Modell Softw, 51: 26-34. DOI:10.1016/j.envsoft.2013.09.022

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:Doi 10.1029/1999jd900155

Bazemore, D.E., Eshleman, K.N., Hollenbeck, K.J., 1994. The Role of Soil-Water in Stormflow Generation in a Forested Headwater Catchment - Synthesis of Natural Tracer and Hydrometric Evidence. J Hydrol, 162(1-2): 47-75. DOI:10.1016/0022-1694(94)90004-3

Becker, A., 2005. Runoff Processes in Mountain Headwater Catchments: Recent Understanding and Research Challenges In: Huber, U.M., Bugmann, H.K.M., Reasoner, M.A. (Eds.), Global Change and Mountain Regions. Advances in Global Change Research. Springer Netherlands, pp. 283-295. DOI:10.1007/1-4020-3508-x_29

Bennett, N.D. et al., 2013. Characterising performance of environmental models. Environ Modell Softw, 40: 1-20. DOI:10.1016/j.envsoft.2012.09.011

Betson, R.P., 1964. What Is Watershed Runoff. J Geophys Res, 69(8).

Beven, K., 1993. Prophecy, Reality and Uncertainty in Distributed Hydrological Modeling. Adv Water Resour, 16(1): 41-51. DOI:10.1016/0309-1708(93)90028-E

Beven, K., 2006. A manifesto for the equifinality thesis. J Hydrol, 320(1-2): 18-36. DOI:10.1016/j.jhydrol.2005.07.007

Beven, K., Binley, A., 2014. GLUE: 20 years on. Hydrological Processes, 28(24): 5897-5918. DOI:10.1002/hyp.10082

Beven, K., Germann, P., 2013. Macropores and water flow in soils revisited. Water Resour Res, 49(6): 3071-3092. DOI:10.1002/Wrcr.20156

Beven, K.J., 2000. Uniqueness of place and process representations in hydrological modelling. Hydrol Earth Syst Sc, 4(2): 203-213. DOI:10.5194/hess-4-203-2000, 2000.

Birkel, C., Tetzlaff, D., Dunn, S.M., Soulsby, C., 2010. Towards a simple dynamic process conceptualization in rainfall-runoff models using multi-criteria calibration and tracers in temperate, upland catchments. Hydrological Processes, 24(3): 260-275. DOI:10.1002/Hyp.7478

Bishop, K., Seibert, J., Nyberg, L., Rodhe, A., 2011. Water storage in a till catchment. II: Implications of transmissivity feedback for flow paths and turnover times. Hydrological Processes, 25(25): 3950-3959. DOI:10.1002/Hyp.8355

Bishop, K.H., 1991. Episodic increase in stream acidity, catchment flow pathways and hydrograph separation. PhD Thesis, University of Cambridge, 246 pp.

Blumstock, M., Tetzlaff, D., Malcolm, I.A., Nuetzmann, G., Soulsby, C., 2015. Baseflow dynamics: Multi-tracer surveys to assess variable groundwater contributions to montane streams under low flows. J Hydrol, 527: 1021-1033. DOI:10.1016/j.jhydrol.2015.05.019

Bolger, B.L., Park, Y.J., Unger, A.J.A., Sudicky, E.A., 2011. Simulating the pre-development hydrologic conditions in the San Joaquin Valley, California. J Hydrol, 411(3-4): 322-330. DOI:10.1016/j.jhydrol.2011.10.013

Bowling, L.C., Kane, D.L., Gieck, R.E., Hinzman, L.D., Lettenmaier, D.P., 2003. The role of surface storage in a low-gradient Arctic watershed. Water Resour Res, 39(4). DOI:10.1029/2002wr001466

Brauer, C.C., Teuling, A.J., Torfs, P.J.J.F., Uijlenhoet, R., 2013. Investigating storage-discharge relations in a lowland catchment using hydrograph fitting, recession analysis, and soil moisture data. Water Resour Res, 49(7): 4257-4264. DOI:10.1002/Wrcr.20320

Bredehoeft, J., 2010. Models and Model Analysis. Ground Water, 48(3): 328-328. DOI:10.1111/j.1745-6584.2009.00631.x

Brunner, P., Doherty, J., Simmons, C.T., 2012. Uncertainty assessment and implications for data acquisition in support of integrated hydrologic models. Water Resour Res, 48. DOI:10.1029/2011wr011342

Burhenne, S., Jacob, D., Henze, G., 2011. SAMPLING BASED ON SOBOL0 SEQUENCES FOR MONTE CARLO TECHNIQUES APPLIED TO BUILDING SIMULATIONS, 12th Conference of International Building Performance Simulation Association, Sydney.

Burt, T.P., Butcher, D.P., 1985. On the Generation of Delayed Peaks in Stream Discharge. J Hydrol, 78(3-4): 361-378. DOI:10.1016/0022-1694(85)90113-1

Buttle, J.M., 1994. Isotope Hydrograph Separations and Rapid Delivery of Pre-Event Water from Drainage Basins. Prog Phys Geog, 18(1): 16-41.

Carrera-Hernandez, J.J., Smerdon, B.D., Mendoza, C.A., 2012. Estimating groundwater recharge through unsaturated flow modelling: Sensitivity to boundary conditions and vertical discretization. J Hydrol, 452: 90-101. DOI:10.1016/j.jhydrol.2012.05.039

Carsel, R.F., Parrish, R.S., 1988. Developing Joint Probability-Distributions of Soil-Water Retention Characteristics. Water Resour Res, 24(5): 755-769. DOI:10.1029/Wr024i005p00755

Chanat, J.G., Hornberger, G.M., 2003. Modeling catchment-scale mixing in the near-stream zone - Implications for chemical and isotopic hydrograph separation. Geophys Res Lett, 30(2). DOI:10.1029/2002gl016265

Chen, J.S. et al., 2012. Isotopic constraints on the origin of groundwater in the Ordos Basin of northern China. Environ Earth Sci, 66(2): 505-517. DOI:DOI 10.1007/s12665-011-1259-6

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

Cloke, H.L., Anderson, M.G., McDonnell, J.J., Renaud, J.P., 2006. Using numerical modelling to evaluate the capillary fringe groundwater ridging hypothesis of streamflow generation. J Hydrol, 316(1-4): 141-162. DOI:10.1016/j.jhydrol.2005.04.017

Clow, D.W. et al., 2003. Ground Water Occurrence and Contributions to Streamflow in an Alpine Catchment, Colorado Front Range. Ground Water, 41(7): 937-950. DOI:10.1111/j.1745-6584.2003.tb02436.x

Combalicer, E., Lee, S., Ahn, S., Kim, D., Im, S., 2008. Comparing groundwater recharge and base flow in the Bukmoongol small-forested watershed, Korea. J Earth Syst Sci, 117(5): 553-566. DOI:10.1007/s12040-008-0052-8

Cornelissen, T., Diekkruger, B., Bogena, H.R., 2014. Significance of scale and lower boundary condition in the 3D simulation of hydrological processes and soil moisture variability in a forested headwater catchment. J Hydrol, 516: 140-153. DOI:10.1016/j.jhydrol.2014.01.060

Coron, L. et al., 2012. Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments. Water Resour Res, 48. DOI:10.1029/2011wr011721

Creed, I.F., Band, L.E., 1998. Exploring functional similarity in the export of nitrate-N from forested catchments: A mechanistic modeling approach. Water Resour Res, 34(11): 3079-3093. DOI:10.1029/98WR02102

Dahlke, H.E. et al., 2012. Dissecting the variable source area concept – Subsurface flow pathways and water mixing processes in a hillslope. J Hydrol, 420–421(0): 125-141. DOI:10.1016/j.jhydrol.2011.11.052

Detty, J.M., McGuire, K.J., 2010. Topographic controls on shallow groundwater dynamics: implications of hydrologic connectivity between hillslopes and riparian zones in a till mantled catchment. Hydrological Processes, 24(16): 2222-2236. DOI:10.1002/Hyp.7656

Dickinson, R.E., Henderson-Sellers, A., Rosenzweig, C., Sellers, P.J., 1991. Evapotranspiration Models with Canopy Resistance for Use in Climate Models - a Review. Agr Forest Meteorol, 54(2-4): 373-388. DOI:10.1016/0168-1923(91)90014-H

Doherty, J., Christensen, S., 2011. Use of paired simple and complex models to reduce predictive bias and quantify uncertainty. Water Resour Res, 47. DOI:10.1029/2011wr010763

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

Doherty, J., Hunt, R.J., Tonkin, M.J., 2010. Approaches to highly parameterized inversion: A guide to using PEST for model-parameter and predictive-uncertainty analysis, U.S. Geological Survey Scientific Investigations Report 2010-5211, 71 p.

Doherty, J., Welter, D., 2010. A short exploration of structural noise. Water Resour Res, 46. DOI:10.1029/2009wr008377

Dooge, J.C.I., 1986. Looking for Hydrologic Laws. Water Resour Res, 22(9): S46-S58. DOI:10.1029/Wr022i09sp0046s

Downer, C.W., Ogden, F.L., 2003. Prediction of runoff and soil moistures at the watershed scale: Effects of model complexity and parameter assignment. Water Resour Res, 39(3). DOI:10.1029/2002wr001439

Durner, W., Jansen, U., Iden, S.C., 2008. Effective hydraulic properties of layered soils at the lysimeter scale determined by inverse modelling. Eur J Soil Sci, 59(1): 114-124. DOI:10.1111/j.1365-2389.2007.00972.x

Ebel, B.A., Loague, K., 2006. Physics-based hydrologic-response simulation: Seeing through the fog of equifinality. Hydrological Processes, 20(13): 2887-2900. DOI:10.1002/Hyp.6388

Edlefsen, N.E., Bodman, G.B., 1941. Field measurements of water movement through a silt loam soil. Amer. Soc. Agron. Jour, 33: 713-731.

Elci, A., Molz, F.J., 2009. Identification of Lateral Macropore Flow in a Forested Riparian Wetland through Numerical Simulation of a Subsurface Tracer Experiment. Water Air Soil Poll, 197(1-4): 149-164. DOI:10.1007/s11270-008-9798-5

Ewen, T., Lehner, I., Seibert, J., Seneviratne, S.I., 2011. Climate patterns in the long-term hydrometeorological data series of the Rietholzbach catchmentt. Bodenkultur. 62: 53–58.

Fienen, M.N., Doherty, J.E., Hunt, R.J., Reeves, H.W., 2010. Using Prediction Uncertainty Analysis to Design Hydrologic Monitoring Networks: Example Applications from the Great Lakes Water Availability Pilot Project:, U.S. Geological Survey Scientific Investigations Report 2010-5159, 44 p. [http://pubs.usgs.gov/sir/2010/5159]

Fiori, A., Romanelli, M., Cavalli, D.J., Russo, D., 2007. Numerical experiments of streamflow generation in steep catchments. J Hydrol, 339(3-4): 183-192. DOI:10.1016/j.jhydrol.2007.03.014

Fiori, A., Russo, D., 2007. Numerical analyses of subsurface flow in a steep hillslope under rainfall: The role of the spatial heterogeneity of the formation hydraulic properties. Water Resour Res, 43(7). DOI:10.1029/2006wr005365

Flugel, W.A., 1995. Delineating Hydrological Response Units by Geographical Information-System Analyses for Regional Hydrological Modeling Using Prms/Mms in the Drainage-Basin of the River Brol, Germany. Hydrological Processes, 9(3-4): 423-436. DOI:10.1002/hyp.3360090313

Freeze, R.A., 1972. Role of Subsurface Flow in Generating Surface Runoff .1. Base Flow Contributions to Channel Flow. Water Resour Res, 8(3). DOI:10.1029/Wr008i003p00609

Freeze, R.A., 1974. Streamflow Generation. Rev Geophys, 12(4): 627-647.

Frei, S., Lischeid, G., Fleckenstein, J.H., 2010. Effects of micro-topography on surface-subsurface exchange and runoff generation in a virtual riparian wetland - A modeling study. Adv Water Resour, 33(11): 1388-1401. DOI:10.1016/j.advwatres.2010.07.006

Frohlich, K., Frohlich, W., Wittenberg, H., 1994. Determination of Groundwater Recharge by Baseflow Separation - Regional Analysis in Northeast China. Iahs-Aish P(221): 69-75.

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., Kohne, J.M., 2004. Dual-permeability modeling of preferential bromide leaching from a tile-drained glacial till agricultural field. J Hydrol, 289(1-4): 239-257. DOI:10.1016/j.jhydrol.2003.11.019

Gerke, H.H., van Genuchten, M.T., 1993a. 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

Gerke, H.H., van Genuchten, M.T., 1993b. Evaluation of a 1st-Order Water Transfer Term for Variably Saturated Dual-Porosity Flow Models. Water Resour Res, 29(4): 1225-1238. DOI:10.1029/92wr02467

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

Gillham, R.W., 1984. The Capillary-Fringe and Its Effect on Water-Table Response. J Hydrol, 67(1-4): 307-324. DOI:10.1016/0022-1694(84)90248-8

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

Granovsky, A.V., Mccoy, E.L., Dick, W.A., Shipitalo, M.J., Edwards, W.M., 1994. Impacts of Antecedent Moisture and Soil Surface Mulch Coverage on Water and Chemical-Transport through a No-Till Soil. Soil Till Res, 32(2-3): 223-236. DOI:10.1016/0167-1987(94)90022-1

Grayson, R., Blöschl, G.n., 2001. Spatial patterns in catchment hydrology : observations and modelling. Cambridge University Press, Cambridge, U.K. ; New York, xii, 404 p. pp.

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

Gurtz, J. et al., 2003. Long-term hydrometeorological measurements and model-based analyses in the hydrological research catchment Rietholzbach. J. Hydrol. Hydromech.(154): 162-174.

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

Harr, R.D., 1977. Water Flux in Soil and Subsoil on a Steep Forested Slope. J Hydrol, 33(1-2): 37-58. DOI:10.1016/0022-1694(77)90097-X

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

Hewlett, J.D., 1974. Letters Relating to Role of Subsurface Flow in Generating Surface Runoff .2. Upstream Source Areas by Freeze, Ra. Water Resour Res, 10(3): 605-607.

Hewlett, J.D., Hibbert, A.R., 1963. Moisture and Energy Conditions within a Sloping Soil Mass during Drainage. J Geophys Res, 68(4). DOI:10.1029/JZ068i004p01081

Hilberts, A.G.J., Troch, P.A., Paniconi, C., Boll, J., 2007. Low-dimensional modeling of hillslope subsurface flow: Relationship between rainfall, recharge, and unsaturated storage dynamics. Water Resour Res, 43(3). DOI:10.1029/2006wr004964

Hill, C., Tiedeman, C.R., 2007. Effective groundwater model calibration : with analysis of data, sensitivities, predictions, and uncertainty. Wiley-Interscience, Hoboken, N.J., xviii, 455 p. pp.

Hills, R.C., 1971. The influence of land management and soil characteristics on infiltration and the occurrence of overland flow. J Hydrol, 13(0): 163-181. DOI:10.1016/0022-1694(71)90213-7

Horton, R.E., 1933. The role of infiltration in the hydrologic cycle. Eos T Am Geophys Un, 14: 446-460.

Huth, A.K., Leydecker, A., Sickman, J.O., Bales, R.C., 2004. A two-component hydrograph separation for three high-elevation catchments in the Sierra Nevada, California. Hydrological Processes, 18(9): 1721-1733. DOI:10.1002/Hyp.1414

James, A.L., McDonnell, J.J., Tromp-van Meerveld, I., Peters, N.E., 2010. Gypsies in the palace: experimentalist's view on the use of 3-D physics-based simulation of hillslope hydrological response. Hydrological Processes, 24(26): 3878-3893. DOI:10.1002/Hyp.7819

James, S.C., Doherty, J.E., Eddebbarh, A.A., 2009. Practical Postcalibration Uncertainty Analysis: Yucca Mountain, Nevada. Ground Water, 47(6): 851-869. DOI:10.1111/j.1745-6584.2009.00626.x

Jarvis, N.J., 2007. A review of non-equilibrium water flow and solute transport in soil macropores: principles, controlling factors and consequences for water quality. Eur J Soil Sci, 58(3): 523-546. DOI:10.1111/j.1365-2389.2007.00915.x

Jayatilaka, C.J., Gillham, R.W., 1996. A deterministic-empirical model of the effect of the capillary fringe on near-stream area runoff .1. Description of the model. J Hydrol, 184(3-4): 299-315. DOI:10.1016/0022-1694(95)02985-0

Johansson, B., 1985. A Study of Soilwater and Groundwater-Flow of Hillslopes - Using a Mathematical-Model. Nord Hydrol, 16(2): 67-78.

Jones, J.P., Sudicky, E.A., Brookfield, A.E., Park, Y.J., 2006a. An assessment of the tracer-based approach to quantifying groundwater contributions to streamflow. Water Resour Res, 42(2). DOI:10.1029/2005wr004130,

Jones, J.P., Sudicky, E.A., McLaren, R.G., 2008. Application of a fully-integrated surface-subsurface flow model at the watershed-scale: A case study. Water Resour Res, 44(3). DOI:10.1029/2006wr005603

Jones, P.D., Lister, D.H., Wilby, R.L., Kostopoulou, E., 2006b. Extended riverflow reconstructions for England and Wales, 1865-2002. Int J Climatol, 26(2): 219-231. DOI:10.1002/Joc.1252

Kashyap, P.S., Panda, R.K., 2001. Evaluation of evapotranspiration estimation methods and development of crop-coefficients for potato crop in a sub-humid region. Agr Water Manage, 50(1): 9-25. DOI:10.1016/S0378-3774(01)00102-0

Kendall, K.A., Shanley, J.B., McDonnell, J.J., 1999. A hydrometric and geochemical approach to test the transmissivity feedback hypothesis during snowmelt. J Hydrol, 219(3-4): 188-205.

Kendy, E. et al., 2003. A soil-water-balance approach to quantify groundwater recharge from irrigated cropland in the North China Plain. Hydrological Processes, 17(10): 2011-2031. DOI:10.1002/Hyp.1240

Kirnbauer, R., Haas, P., 1998. Observations on runoff generation mechanisms in small Alpine catchments. Hydrology, Water Resources and Ecology in Headwaters(248): 239-247.

Kish, G.R., Stringer, C.E., Stewart, M.T., Rains, M.C., Torres, A.E., 2010. A Geochemical Mass-Balance Method for Base-Flow Separation, Upper Hillsborough River Watershed, West-Central Florida, 2003-2005 and 2009. 2010–5092,Survey, U.S.G.

Kohne, J.M., Kohne, S., Gerke, H.H., 2002. Estimating the hydraulic functions of dual-permeability models from bulk soil data. Water Resour Res, 38(7). DOI:10.1029/2001wr000492

Kohne, J.M., Kohne, S., Simunek, J., 2009. A review of model applications for structured soils: a) Water flow and tracer transport. J Contam Hydrol, 104(1-4): 4-35. DOI:10.1016/j.jconhyd.2008.10.002

Kohne, J.M., Mohanty, B.P., Simunek, J., 2006. Inverse dual-permeability modeling of preferential water flow in a soil column and implications for field-scale solute transport. Vadose Zone J, 5(1): 59-76. DOI:10.2136/Vzj2005.0008

Kordilla, J., Sauter, M., Reimann, T., Geyer, T., 2012. Simulation of saturated and unsaturated flow in karst systems at catchment scale using a double continuum approach. Hydrol Earth Syst Sc, 16(10): 3909-3923. DOI:10.5194/hess-16-3909-2012

Kouwen, N., Soulis, E.D., Pietroniro, A., Donald, J., Harrington, R.A., 1993. Grouped Response Units for Distributed Hydrologic Modeling. J Water Res Pl-Asce, 119(3): 289-305. DOI:10.1061/(Asce)0733-9496(1993)119:3(289)

Kramers, G., Holden, N.M., Brennan, F., Green, S., Richards, K.G., 2012. Water Content and Soil Type Effects on Accelerated Leaching after Slurry Application. Vadose Zone J, 11(1). DOI:10.2136/vzj2011.0059

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

Lafond, J.A., Piette, E.B., Caron, J., Rancourt, G.T., 2014. Evaluating fluxes in Histosols for water management in lettuce: A comparison of mass balance, evapotranspiration and lysimeter methods. Agr Water Manage, 135: 73-83. DOI:10.1016/j.agwat.2013.12.016

Laloy, E., Weynants, M., Bielders, 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

Laudon, H., Seibert, J., Kohler, S., Bishop, K., 2004. Hydrological flow paths during snowmelt: Congruence between hydrometric measurements and oxygen 18 in meltwater, soil water, and runoff. Water Resour Res, 40(3). DOI:10.1029/2003wr002455

Lee, C.H., Chen, W.P., Lee, R.H., 2007. Estimation of groundwater recharge using water balance coupled with base-flow-record estimation and stable-base-flow analysis (vol 51, pg 73, 2006). Environ Geol, 51(5): 869-869. DOI:10.1007/s00254-006-0561-1

Lehner, B., Doll, P., Alcamo, J., Henrichs, T., Kaspar, F., 2006. Estimating the impact of global change on flood and drought risks in europe: A continental, integrated analysis. Climatic Change, 75(3): 273-299. DOI:10.1007/s10584-006-6338-4

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

Liang, D.F., Falconer, R.A., Lin, B.L., 2007. Coupling surface and subsurface flows in a depth averaged flood wave model. J Hydrol, 337(1-2): 147-158. DOI:10.1016/j.jhydrol.2007.01.045

Linhoss, A., Munoz-Carpena, R., Kiker, G., Hughes, D., 2013. Hydrologic Modeling, Uncertainty, and Sensitivity in the Okavango Basin: Insights for Scenario Assessment. J Hydrol Eng, 18(12): 1767-1778. DOI:10.1061/(Asce)He.1943-5584.0000755

Liu, L., Cheng, S.Y., Guo, H.C., 2004. A simulation-assessment modeling approach for analyzing environmental risks of groundwater contamination at waste landfill sites. Hum Ecol Risk Assess, 10(2): 373-388. DOI:10.1080/10807030490438436

Liu, Y.J., Luo, Y., 2010. A consolidated evaluation of the FAO-56 dual crop coefficient approach using the lysimeter data in the North China Plain. Agr Water Manage, 97(1): 31-40. DOI:10.1016/j.agwat.2009.07.003

Loague, K. et al., 2005. Further testing of the Integrated Hydrology Model (InHM): event-based simulations for a small rangeland catchment located near Chickasha, Oklahoma. Hydrological Processes, 19(7): 1373-1398. DOI:10.1002/Hyp.5566

Loague, K. et al., 2006. Physics-based hydrologic-response simulation: foundation for hydroecology and hydrogeomorphology. Hydrological Processes, 20(5): 1231-1237. DOI:10.1002/Hyp.6179

Marechal, J.C., Dewandel, B., Ahmed, S., Galeazzi, L., Zaidi, F.K., 2006. Combined estimation of specific yield and natural recharge in a semi-arid groundwater basin with irrigated agriculture. J Hydrol, 329(1-2): 281-293. DOI:10.1016/j.jhydrol.2006.02.022

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

Matonse, A.H., Kroll, C.N., 2013. Applying hillslope-storage models to improve low flow estimates with limited streamflow data at a watershed scale. J Hydrol, 494: 20-31. DOI:10.1016/j.jhydrol.2013.04.032

Mcdonnell, J.J., 1990. A Rationale for Old Water Discharge through Macropores in a Steep, Humid Catchment. Water Resour Res, 26(11): 2821-2832. DOI:10.1029/WR026i011p02821

McDonnell, J.J., Buttle, J.M., 1998. Comment on "A deterministic-empirical model of the effect of the capillary-fringe on near-stream area runoff. 1. Description of the model" by Jayatilaka, C. J. and Gillham, R. W. (Journal of Hydrology vol. 184 (1996) 299-315). J Hydrol, 207(3-4): 280-285.

McDonnell, J.J. et al., 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. Water Resour Res, 43(7). DOI:10.1029/2006wr005467

McGuire, K.J., McDonnell, J.J., 2006. A review and evaluation of catchment transit time modeling. J Hydrol, 330(3-4): 543-563. DOI:10.1016/j.jhydrol.2006.04.020

McGuire, K.J., Weiler, M., McDonnell, J.J., 2007. Integrating tracer experiments with modeling to assess runoff processes and water transit times. Adv Water Resour, 30(4): 824-837. DOI:10.1016/j.advwatres.2006.07.004

McKerchar, A.I., Schmidt, J., 2007. Decreases in low flows in the lower Selwyn River? Journal of Hydrology: New Zealand, 46(2): 63-72.

Mekonnen, G.B., Matula, S., Dolezal, F., Fisak, J., 2015. Adjustment to rainfall measurement undercatch with a tipping-bucket rain gauge using ground-level manual gauges. Meteorol Atmos Phys, 127(3): 241-256. DOI:10.1007/s00703-014-0355-z

Menzel, L., Demuth, N., 1993. Tracerhydrologische Untersuchungen am Lysimeter Rietholzbach [in German], Berichte und Skripten, Rep. 52,, ETH Zürich, 27 pp.

Merdun, H., Meral, R., Demirkiran, A.R., 2008. Effect of the Initial Soil Moisture Content on the Spatial Distribution of the Water Retention. Eurasian Soil Sci+, 41(10): 1098-1106. DOI:10.1134/S1064229308100128

Mertens, J., Stenger, R., Barkle, G.F., 2006. Multiobjective inverse modeling for soil parameter estimation and model verification. Vadose Zone J, 5(3): 917-933. DOI:10.2136/Vzj2005.0117

Merz, R., Parajka, J., Bloschl, G., 2011. Time stability of catchment model parameters: Implications for climate impact analyses. Water Resour Res, 47.

Mirus, B.B., Ebel, B.A., Heppner, C.S., Loague, K., 2011. Assessing the detail needed to capture rainfall-runoff dynamics with physics-based hydrologic response simulation. Water Resour Res, 47. DOI:10.1029/2010wr009906

Mittelbach, H., Lehner, I., Seneviratne, S.I., 2012. Comparison of four soil moisture sensor types under field conditions in Switzerland. J Hydrol, 430: 39-49. DOI:10.1016/j.jhydrol.2012.01.041

Moeck, C., Hunkeler, D., Brunner, P., 2015. Tutorials as a flexible alternative to GUIs: An example for advanced model calibration using Pilot Points. Environ Modell Softw, 66: 78-86. DOI:10.1016/j.envsoft.2014.12.018

Mohanty, S.K., Saiers, J.E., Ryan, J.N., 2014. Colloid-Facilitated Mobilization of Metals by Freeze-Thaw Cycles. Environ Sci Technol, 48(2): 977-984. DOI:10.1021/Es403698u

Morbidelli, R., Corradini, C., Saltalippi, C., Flammini, A., Rossi, E., 2011. Infiltration-soil moisture redistribution under natural conditions: experimental evidence as a guideline for realizing simulation models. Hydrol Earth Syst Sc, 15(9): 2937-2945. DOI:10.5194/hess-15-2937-2011

Morbidelli, R., Saltalippi, C., Flammini, A., Rossi, E., Corradini, C., 2014. Soil water content vertical profiles under natural conditions: matching of experiments and simulations by a conceptual model. Hydrological Processes, 28(17): 4732-4742. DOI:10.1002/hyp.9973

Mualem, Y., 1976. New Model for Predicting Hydraulic Conductivity of Unsaturated Porous-Media. Water Resour Res, 12(3): 513-522. DOI:10.1029/Wr012i003p00513

Munoz, E., Tume, P., Ortiz, G., 2014. Uncertainty in rainfall input data in a conceptual water balance model: effects on outputs and implications for predictability. Earth Sci Res J, 18(1). DOI:http://dx.doi.org/10.15446/esrj.v18n1.38760

Nimmo, J.R., 2012. Preferential flow occurs in unsaturated conditions. Hydrological Processes, 26(5): 786-789. DOI:10.1002/hyp.8380

Nixon, P.R., Lawless, G.P., 1960. Translocation of Moisture with Time in Unsaturated Soil Profiles. J Geophys Res, 65(2): 655-661.

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

Orr, H.G., Carling, P.A., 2006. Hydro-climatic and land use changes in the river lune catchment, North West England, implications for catchment management. River Res Appl, 22(2): 239-255. DOI:10.1002/Rra.908

Panday, S., Huyakorn, P.S., 2004. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. Adv Water Resour, 27(4): 361-382. DOI:10.1016/j.advwatres.2004.02.016

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

Park, Y.J., Sudicky, E.A., Brookfield, A.E., Jones, J.P., 2011. Hydrologic response of catchments to precipitation: Quantification of mechanical carriers and origins of water. Water Resour Res, 47. DOI:10.1029/2011wr010075

Payn, R.A., Gooseff, M.N., McGlynn, B.L., Bencala, K.E., Wondzell, S.M., 2012. Exploring changes in the spatial distribution of stream baseflow generation during a seasonal recession. Water Resour Res, 48. DOI:10.1029/2011wr011552

Penman, H.L., 1948. Natural Evaporation from Open Water, Bare Soil and Grass. Proc R Soc Lon Ser-A, 193(1032). DOI:10.1098/rspa.1948.0037

Penna, D., Mantese, N., Hopp, L., Dalla Fontana, G., Borga, M., 2015. Spatio-temporal variability of piezometric response on two steep alpine hillslopes. Hydrological Processes, 29(2): 198-211. DOI:10.1002/hyp.10140

Perrin, C., Michel, C., Andreassian, V., 2001. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. J Hydrol, 242(3-4): 275-301. DOI:10.1016/S0022-1694(00)00393-0

Peters, E., Bier, G., van Lanen, H.A.J., Torfs, P.J.J.F., 2006. Propagation and spatial distribution of drought in a groundwater catchment. J Hydrol, 321(1-4): 257-275. DOI:10.1016/j.jhydrol.2005.08.004

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

Pierson, T.C., 1980. Piezometric response to rainstorms in forested hillslope drainage depressions. J. Hydrol. : New Zealand, 19: 1 - 10.

Pinder, G.F., Jones, J.F., 1969. Determination of Ground-Water Component of Peak Discharge from Chemistry of Total Runoff. Water Resour Res, 5(2). DOI:10.1029/WR005i002p00438

Pujol, G., Iooss, B., Janon, A., 2015. sensitivity package.

Rahim, B.E.E.A., Yusoff, I., Jafri, A.M., Othman, Z., Ghani, A.A., 2012. Application of MIKE SHE modelling system to set up a detailed water balance computation. Water Environ J, 26(4): 490-503. DOI:10.1111/j.1747-6593.2012.00309.x

Rawitz, E., Engman, E.T., Cline, G.D., 1970. Use of Mass Balance Method for Examining Role of Soils in Controlling Watershed Performance. Water Resour Res, 6(4). DOI:10.1029/WR006i004p01115

Razavi, S., Gupta, H.V., 2015. What do we mean by sensitivity analysis? The need for comprehensive characterization of "global'' sensitivity in Earth and Environmental systems models. Water Resour Res, 51(5): 3070-3092. DOI:10.1002/2014wr016527

Reggiani, P., Sivapalan, M., Hassanizadeh, S.M., 1998. A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics. Adv Water Resour, 22(4): 367-398. DOI:10.1016/S0309-1708(98)00012-8

Reggiani, P., Sivapalan, M., Hassanizadeh, S.M., 2000. Conservation equations governing hillslope responses: Exploring the physical basis of water balance. Water Resour Res, 36(7): 1845-1863. DOI:10.1029/2000wr900066

Remson, I., Randolph, J.R., Barksdale, H.C., 1960. The zone of aeration and ground-water recharge in sandy sediments at Seabrook, New Jersey. Soil Sci., 89: 145-156.

Renaud, J.P., Cloke, H.L., Weiler, M., 2007. An assessment of the tracer-based approach to quantifying groundwater contributions to streamflow - Comment. Water Resour Res, 43(9). DOI:10.1029/2006wr005157

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

Rinderer, M., van Meerveld, I., Stähli, M., Seibert, J., 2015. Is groundwater response timing in a pre-alpine catchment controlled more by topography or by rainfall? Hydrological Processes. DOI:10.1002/hyp.10634

Rindskopf, D., 1997. An introduction to the bootstrap - Efron,B, Tibshirani,RJ. J Educ Behav Stat, 22(2): 245-245.

Rodriguez-Iturbe, I., Porporato, A., Ridolfi, L., Isham, V., Cox, D.R., 1999. Probabilistic modelling of water balance at a point: the role of climate, soil and vegetation. P Roy Soc a-Math Phy, 455(1990): 3789-3805. DOI:10.1098/rspa.1999.0477

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

Rubin, J., 1966. Theory of Rainfall Uptake by Soils Initially Drier Than Their Field Capacity and Its Applications. Water Resour Res, 2(4). DOI:0.1029/WR002i004p00739

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

Schelle, H., Durner, W., Iden, S.C., Fank, J., 2013. Simultaneous estimation of soil hydraulic and root distribution parameters from lysimeter data by inverse modeling. Procedia Environ Sci, 19: 564-573. DOI:10.1016/j.proenv.2013.06.064

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

Schwarzel, K., Simunek, J., van Genuchten, M.T., Wessolek, G., 2006. Measurement and modeling of soil-water dynamics and evapotranspiration of drained peatland soils. J Plant Nutr Soil Sc, 169(6): 762-774. DOI:10.1002/jpln.200621992

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.

Seibert, J., Moren, A.S., 1999. Reducing systematic errors in rainfall measurements using a new type of gauge. Agr Forest Meteorol, 98-9: 341-348. DOI:10.1016/S0168-1923(99)00107-0

Seiller, G., Anctil, F., Perrin, C., 2012. Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions. Hydrol. Earth Syst. Sci., 16(4): 1171-1189. DOI:10.5194/hess-16-1171-2012

Selle, B., Minasny, B., Bethune, M., Thayalakumaran, T., Chandra, S., 2011. Applicability of Richards' equation models to predict deep percolation under surface irrigation. Geoderma, 160(3-4): 569-578. DOI:10.1016/j.geoderma.2010.11.005

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

Sevruk, B., Hertig, J.A., Spiess, R., 1991. The Effect of a Precipitation Gauge Orifice Rim on the Wind-Field Deformation as Investigated in a Wind-Tunnel. Atmos Environ a-Gen, 25(7): 1173-1179. DOI:10.1016/0960-1686(91)90228-Y

Shaman, J., Stieglitz, M., Burns, D., 2004. Are big basins just the sum of small catchments? Hydrological Processes, 18(16): 3195-3206. DOI:10.1002/Hyp.5739

Shen, C.P., Niu, J., Phanikumar, M.S., 2013. Evaluating controls on coupled hydrologic and vegetation dynamics in a humid continental climate watershed using a subsurface- land surface processes model. Water Resour Res, 49(5): 2552-2572. DOI:10.1002/Wrcr.20189

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

Shipitalo, M.J., Edwards, W.M., 1996. Effects of initial water content on macropore/matrix flow and transport of surface-applied chemicals. J Environ Qual, 25(4): 662-670. DOI:10.2134/jeq1996.00472425002500040005x

Sidle, R.C., Hornbeck, J.W., 1991. Cumulative Effects - a Broader Approach to Water-Quality Research. J Soil Water Conserv, 46(4): 268-271.

Simunek, J., Jarvis, N.J., van Genuchten, M.T., Gardenas, A., 2003. Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone. J Hydrol, 272(1-4): 14-35. DOI:10.1016/S0022-1694(02)00252-4

Sklash, M.G., Farvolden, R.N., 1979. Role of Groundwater in Storm Runoff. J Hydrol, 43(1-4): 45-65. DOI:10.1016/0022-1694(79)90164-1

Smakhtin, V.U., 2001. Low flow hydrology: a review. J Hydrol, 240(3-4): 147-186. DOI:10.1016/S0022-1694(00)00340-1

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

Soldevilla-Martinez, M., Quemada, M., Lopez-Urrea, R., Munoz-Carpena, R., Lizaso, J.I., 2014. Soil water balance: Comparing two simulation models of different levels of complexity with lysimeter observations. Agr Water Manage, 139: 53-63. DOI:10.1016/j.agwat.2014.03.011

Squillace, P.J., 1996. Observed and simulated movement of bank-storage water. Ground Water, 34(1): 121-134. DOI:10.1111/j.1745-6584.1996.tb01872.x

Stoddard, J.L. et al., 1999. Regional trends in aquatic recovery from acidification in North America and Europe. Nature, 401(6753): 575-578. DOI:10.1038/44114

Stumpp, C., Maloszewski, P., 2010. Quantification of preferential flow and flow heterogeneities in an unsaturated soil planted with different crops using the environmental isotope delta O-18. J Hydrol, 394(3-4): 407-415. DOI:10.1016/j.jhydrol.2010.09.014

Sudicky, E.A., Jones, J.P., Brookfield, A.E., Park, Y.J., 2007. An assessment of the tracer-based approach to quantifying groundwater contributions to streamflow - Comment - Reply. Water Resour Res, 43(9). DOI:10.1029/2006wr005416

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. J Geophys Res-Atmos, 106(D7): 7183-7192. DOI:10.1029/2000jd900719

Tetzlaff, D., Soulsby, C., 2008. Sources of baseflow in larger catchments - Using tracers to develop a holistic understanding of runoff generation. J Hydrol, 359(3-4): 287-302. DOI:10.1016/j.jhydrol.2008.07.008

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.

Thompson, J.R., Green, A.J., Kingston, D.G., 2014. Potential evapotranspiration-related uncertainty in climate change impacts on river flow: An assessment for the Mekong River basin. J Hydrol, 510: 259-279. DOI:10.1016/j.jhydrol.2013.12.010

Torres, R., Dietrich, W.E., Montgomery, D.R., Anderson, S.P., Loague, K., 1998. Unsaturated zone processes and the hydrologic response of a steep, unchanneled catchment. Water Resour Res, 34(8): 1865-1879. DOI:10.1029/98WR01140

Touhami, I. et al., 2013. Recharge estimation of a small karstic aquifer in a semiarid Mediterranean region (southeastern Spain) using a hydrological model. Hydrological Processes, 27(2): 165-174. DOI:10.1002/Hyp.9200

Twarakavi, N.K.C., Sakai, M., Simunek, J., 2009. An objective analysis of the dynamic nature of field capacity. Water Resour Res, 45. DOI:10.1029/2009wr007944

Uhlenbrook, S., Hoeg, S., 2003. Quantifying uncertainties in tracer-based hydrograph separations: a case study for two-, three- and five-component hydrograph separations in a mountainous catchment. Hydrological Processes, 17(2): 431-453. DOI:10.1002/Hyp.1134

Uhlenbrook, S., Seibert, J., Leibundgut, C., Rodhe, A., 1999. Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. Hydrolog Sci J, 44(5): 779-797. DOI:10.1080/02626669909492273

van Dam, J.C., Feddes, R.A., 2000. Numerical simulation of infiltration, evaporation and shallow groundwater levels with the Richards equation. J Hydrol, 233(1-4): 72-85. DOI:10.1016/S0022-1694(00)00227-4

van der Perk, M., 1997. Effect of model structure on the accuracy and uncertainty of results from water quality models. Hydrological Processes, 11(3): 227-239. DOI:10.1029/2007JD008662

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., 2008. Characterization of watershed model behavior across a hydroclimatic gradient. Water Resour Res, 44(1). DOI:10.1029/2007wr006271

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

VanderKwaak, J.E., 1999. Numerical simulation of flow and chemical transport in integrated surface-subsurface systems. Ph.D Thesis, University of Waterloo,Ontario, Canada, 242 pp.

Vaze, J. et al., 2010. Climate non-stationarity - Validity of calibrated rainfall-runoff models for use in climate change studies. J Hydrol, 394(3-4): 447-457. DOI:10.1016/j.jhydrol.2010.09.018

Vazquez, R.F., Feyen, J., 2003. Effect of potential evapotranspiration estimates on effective parameters and performance of the MIKE SHE-code applied to a medium-size catchment. J Hydrol, 270(3-4): 309-327. DOI:10.1016/S0022-1694(02)00308-6

Vidon, P., 2012. Towards a better understanding of riparian zone water table response to precipitation: surface water infiltration, hillslope contribution or pressure wave processes? Hydrological Processes: n/a-n/a. DOI:10.1002/hyp.8258

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

Vitvar, T., Gurtz, J., Lang, H., 1999. Application of GIS-based distributed hydrological modelling for estimation of water residence times in the small Swiss pre-alpine catchment Rietholzbach. Integrated Methods in Catchment Hydrology: Tracer, Remote Sensing and New Hydrometric Techniques(258): 241-248.

Viviroli, D., Zappa, M., Gurtz, J., Weingartner, R., 2009. An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools. Environ Modell Softw, 24(10): 1209-1222. DOI:10.1016/j.envsoft.2009.04.001

Vogel, H.J., Ippisch, O., 2008. Estimation of a critical spatial discretization limit for solving Richards' equation at large scales. Vadose Zone J, 7(1): 112-114. DOI:10.2136/Vzj2006.0182 von Freyberg, J., Moeck, C., Schirmer, M., 2015. Estimation of groundwater recharge and drought severity with varying model complexity. J Hydrol, 527: 844-857. DOI:10.1016/j.jhydrol.2015.05.025

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

Wang, Y.S., Bradford, S.A., Simunek, J., 2014. Estimation and upscaling of dual-permeability model parameters for the transport of E. coli D21g in soils with preferential flow. J Contam Hydrol, 159: 57-66. DOI:10.1016/j.jconhyd.2014.01.009

Watson, T.A., Doherty, J.E., Christensen, S., 2013. Parameter and predictive outcomes of model simplification. Water Resour Res, 49(7): 3952-3977. DOI:10.1002/Wrcr.20145

Wegehenkel, M., Gerke, H.H., 2015. Water table effects on measured and simulated fluxes in weighing lysimeters for differently-textured soils. J Hydrol Hydromech, 63(1): 82-92. DOI:10.1515/johh-2015-0004

Weiler, M., 2001. Mechanisms controlling macropore flow during infiltration: Dye tracer experiments and simulations, ETH Zürich, 1-172 pp.

Weiler, M., Fluhler, H., 2004. Inferring flow types from dye patterns in macroporous soils. Geoderma, 120(1-2): 137-153. DOI:10.1016/j.geoderma.2003.08.014

Weiler, M., McDonnell, J., 2004. Virtual experiments: a new approach for improving process conceptualization in hillslope hydrology. J Hydrol, 285(1-4): 3-18. DOI:10.1016/S0022-1694(03)00271-3

Weiler, M., Naef, F., 2003. Simulating surface and subsurface initiation of macropore flow. J Hydrol, 273(1-4): 139-154.

Weill, S., Mazzia, A., Putti, M., Paniconi, C., 2011. Coupling water flow and solute transport into a physically-based surface-subsurface hydrological model. Adv Water Resour, 34(1): 128-136. DOI:10.1016/j.advwatres.2010.10.001

Wenninger, J., Uhlenbrook, S., Tilch, N., Leibundgut, C., 2004. Experimental evidence of fast groundwater responses in a hillslope/floodplain area in the Black Forest Mountains, Germany. Hydrological Processes, 18(17): 3305-3322. DOI:10.1002/Hyp.5686

Whipkey, R.Z., 1965. Subsurface stormflow from forested slopes. Bull. Int. Ass. Sci. Hydrol, 10(2): 74-85.

Whitaker, M.P.L., 2000. Estimating bank storage and evapotranspiration using soil physical and hydrological techniques in a gaining reach of the San Pedro River, Arizona. Ph.D. Thesis, University of Arizona.

Winter, T.C., 2007. Role of ground water in generating streamflow in headwater areas and in maintaining base flow. J Am Water Resour As, 43(1): 15-25. DOI:10.1111/j.1752-1688.2007.00003.x

Winter, T.C., Harvey, J.W., Lehn Franke, O., Alley, W.M., 1998. Ground water and surface water - a single resource,US Geological Survey Circular 1139, U.D.o.t.I.

WMO, 1974. World Meteorological Organization, Geneva.

Wondzell, S.M., Gooseff, M.N., McGlynn, B.L., 2007. Flow velocity and the hydrologic behavior of streams during baseflow. Geophys Res Lett, 34(24). DOI:10.1029/2007gl031256

Zaltsberg, E., 1986. Laboratory Studies of the Effects of the Capillary-Fringe on Streamflow Generation by Abdul,A.S. And Gillham,R.W. - Comment. Water Resour Res, 22(5): 837-838. DOI:10.1029/WR022i005p00837

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

Zhang, G.P., Savenije, H.H.G., 2005. Rainfall-runoff modelling in a catchment with a complex groundwater flow system: application of the Representative Elementary Watershed (REW) approach. Hydrol Earth Syst Sc, 9(3): 243-261. DOI:10.5194/hess-9-243-2005

Zheng, C.M., Gorelick, S.M., 2003. Analysis of solute transport in flow fields influenced by preferential flowpaths at the decimeter scale. Ground Water, 41(2): 142-155. DOI:10.1111/j.1745-6584.2003.tb02578.x

Zhu, J.T., Sun, D.M., 2009. Effective Soil Hydraulic Parameters for Transient Flows in Heterogeneous Soils. Vadose Zone J, 8(2): 301-309. DOI:10.2136/Vzj2008.0004

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