ALBERT-LUDWIGS
UNIVERSITÄT FREIBURG

HYDROLOGIE FREIBURG

Inverse Hydrological Modelling of Headwater Basins with Sensor Network Data

Till H.M. Volkmann

Diplomarbeit unter Leitung von Prof. Dr. Markus Weiler
Freiburg im Breisgau, März 2011

Albert-Ludwigs-Universität Freiburg im Breisgau
Institut für Hydrologie

Inverse Hydrological Modelling of Headwater Basins with Sensor Network Data

Till H.M. Volkmann

Referent: Prof. Dr. Markus Weiler
Koreferent: Prof. Dr. Harry Vereecken

Diplomarbeit unter Leitung von Prof. Dr. Markus Weiler
Freiburg im Breisgau, März 2011

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass die Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt wurde.

Freiburg i.Br., 30.3.2011 Till H.M. Volkmann

Acknowledgements

I would like to thank my supervisor Prof. Dr. M. Weiler to encourage me to this thesis and Prof. Dr. H. Vereecken for volunteering to co-review this thesis.

Many people helped to get the work done. It is a pleasure to thank those who made this thesis possible, to Jasper Vrugt for advice with the MOSCEM setup, Hoshin V. Gupta who shared his MATLAB MOSCEM computer code, Thorsten Wagener who provided the MCAT toolbox, Bodo Friedrich who provided solar radiation data, Andreas Hartmann who helped to setup programs on the server, and Julian Haas who helped getting me started with IDL. Special thanks go to Heye Bogena for his continued support and effort. I owe my deepest gratitude to my girlfriend Milena and my parents who were always there when I needed them.

Table of content

List of figures

List of tables

Abbreviations

ASCE	American Society of Civil Engineers
CORR	Coefficient of Correlation
CV	Coefficient of Variation
DEM	Digital Elevation Model
GLUE	Generalized Likelihood Uncertainty Estimation
IDL	Interactive Data Language
MAE	Mean Absolute Error
MAHE	Mean Absolute Histogram Error
MCAT	Monte Carlo Analysis Toolbox
MOSCEM	Multiobjective Shuffled Complex Evolution Metropolis
MCMC	Markov Chain Monte Carlo
OF	Objective Function
PTF	Pedotransfer Functions
SCEM	Shuffled Complex Evolution Metropolis
TERENO	Terrestrial Enviromental Observatoria

Variables

b	empirical drainage coefficient	m
C	canopy storage	m
CFR	Refreezing Coefficient	
D	canopy drainage	h^{-1}
d	saturation vapor pressure deficit	
E_c	canopy evaporation rate	$m\ h^{-1}$
E_p	potential evaporation rate	$m\ h^{-1}$
ET	evapotranspiration rate	$m\ h^{-1}$
f	hydraulic conductivity shape factor	m
G	ground heat flux	$W\ m^{-2}$
h	overland flow water level	m
k_s	saturated hydraulic conductivity	$m\ h^{-1}$
k_0	surface saturated hydraulic conductivity	$m\ h^{-1}$
k_b,	saturated hydraulic conductivity of the bedrock	$m\ h^{-1}$
LAI	Leaf Area Index	
$MELT$	melt rate	$m\ h^{-1}$
n_0	drainable porosity at the soil surface	
Q	runoff yield	$m\ h^{-1}$
Q_{of}	overland flux	$m^3\ h^{-1}$
Q_{ssf}	Lateral subsurface flow	$m^3\ h^{-1}$
P	gross rainfall rate	$m\ h^{-1}$
$REFR$	Refrozen water	

R	recharge to the saturated zone	m h^{-1}
R_n	net radiation	W m^{-2}
RMC	residual moisture content	%
RS	relative saturation	%
r_a	aerodynamic resistance	s m^{-1}
r_c	canopy resistance	s m^{-1}
r_{sc}	minimum canopy resistance	s m^{-1}
r_s	resistance to vapor transport	s m^{-1}
r_{ss}	surface resistance	s m^{-1}
δ	local surface slope	
s	minimum canopy storage capacity	m
S	Seepage flux	m h^{-1}
SMC	soil moisture content	%
T	transmissivity	m^2 h^{-1}
T_t	threshold temperature	°C
T_a	actual air temperature	°C
u	wind speed	m h^{-1}
h_w	height of the water table	m
w	width of flow area	m
z	soil depth below the surface	m
α	soil moisture stress factor	
β	water table gradient	
p	free throughfall coefficient	
ρ_a	density of air	kg m-3
φ	observed porosity	

Abstract

This thesis developed an inverse modeling approach to directly incorporate high resolution spatiotemporal soil moisture content data produced by the wireless sensor network SoilNet for the temperate humid forested headwater basin Wüstebach to parameterize the distributed hydrological model Hill-Vi. A one year period of six-hourly soil moisture measurements from 150 locations in the 0.27 km^2 basin was available for this study. The model was coupled to the Multi-Objective Shuffled Complex Evolution Metropolis (MOSCEM-UA) algorithm to facilitate efficient sampling of the parameter space with respect to a suite of objective functions constraining the model to runoff at the catchment outlet and internal soil moisture dynamics. State-of-the-art hydrological inverse modeling methods were applied to assess the identifiability of model parameters and model structural uncertainty. The approach was then used to preliminarily investigate the value of soil moisture data in constraining the parameter space and to study the influence of increasing model complexity by including bedrock seepage, variable soil depth, and variable throughfall into the model. As for now, the model was found to be incapable of appropriately simulating the basin dynamics. This may be primarily attributed to biases in the meteorological data and the limited amount of parameter samples generated. More research is needed to derive reliable conlcusions on both the utility of soil moisture data and the value of added model complexity for the simulation of internal soil moisture dynamics and runoff response.

Keywords: Inverse modelling, model calibration, distributed hydrological models, soil moisture, wireless sensor networks, multi-criteria optimization, model complexity

Zusammenfassung

In dieser Arbeit wurden Methoden der inversen Modellierung zur Parametrisierung verteilter hydrologischer Modelle unter Verwendung räumlich und zeitlich hochaufgelöster Bodenfeuchtedaten. Die Daten stammen von dem Wireless Sensor Network SoilNet das in dem bewaldeten Einzugsgebiet Wüstenbach (0,27 km^2) installiert wurde. Für diese Studie wurden Bodenfeuchtemessungen in einem 6-stündigen Zeitschritt von 150 Sensoren verwendet. Das Modell wurde mit dem Multi-Objective Shuffled Complex Evolution Metropolis (Moscem-UA) Algorithmus gekoppelt um eine effiziente Beprobung des Parameterraums hinsichtlicher multipler Gütemaße zu ermöglichen. Die Gütemaße messen die Übereinstimmung der Simulationen von Abfluss am Gebietsauslass und Bodenfeuchte mit den entsprechenden Messungen. Um die Identifizierbarkeit der Parameter und die Unsicherheit in der Modellstruktur zu erfassen wurden inverse Modellierungsverfahren nach dem Stand der Forschung angewendet. Die Methode wurde angewendet um den Wert der Bodenfeuchtedaten zur Begrenzung des Parameterraums sowie den Einfluss erhöhter Modellkomplexität bezüglich der Einbeziehung von Versickerung in den Untergrund, variabler Bodentiefen und variablen Kronendurchlass vorläufig zu untersuchen. Bei erster Betrachtung der Ergebnisse wurde keine adäquate Übereinstimmung zwischen den Simulationen und Messungen gefunden. Dies ist vermutlich hauptsächlich systematischen Fehlern in den meteorologischen Daten und der geringen Anzahl an möglichen Parameterstichproben zuzuordnen. Weitere Untersuchungen sind erforderlich um zuverlässige Schlüsse über den Nutzen der Bodenfeuchtedaten und der erhöhten Modellkomplexität für die Simulation der Bodenfeuchtedynamik und der Abflusses ziehen zu können.

1. Introduction

1.1. Motivation

According to Bogena et al. (2010) a remaining challenge in hydrology is to explain the observed patterns of hydrologic behavior across multiple space–time scales as a result of interacting environmental factors. Furthermore, there is an increasing demand for spatially explicit predictions to address complex environmental problems concerning surface water acidification, soil erosion, pollutant leaching, and possible consequences of land use or climatic changes (Grayson et al., 2002). The complexity of problems hydrologists are asked to investigate has grown over the years (Wagener et al., 2007).

Observations and measurements are vital to improving our understanding of hydrological response. However, to understand the dynamics of hydrological processes, a framework to facilitate hypothesis testing is needed. Computer-based modelling is used throughout hydrology for this purpose (e.g., Wealands, 2006). There are many different types of models, ranging from those that estimate bulk quantities to those that produce spatially explicit estimates across an area. There are many comprehensive reviews of hydrological modelling available, which provide examples and classifications of models (e.g. Singh, 1995; Abbott and Refsgaard, 1996; Grayson and Blöschl, 2000a,b,c). In this thesis, the focus is on spatial models, which are used for testing hypotheses about the behaviour of hydrological systems. Models provide the platform on which conceptualisations of hydrological processes are combined to simulate hydrological response. If models prove to adequately simulate a certain response, they can also be used for predicting the effects of changed conditions on hydrological response (e.g. land use change).

The past decades have seen the development and application of numerous physically based distributed models (i.e., models that explicitly represent spatially varying fields) of diverse levels of complexity over a range of scales, from hillslopes (Faeh et al., 1997; Weiler et al., 1998; Calver and Cammeraat, 1993; Sloan and Moore, 1984) to mesoscale and largescale basins (e.g., Abbott et al., 1986; Beven et al., 1987; Grayson et al., 1992; Julien and Saghafian, 1991; Wigmosta et al., 1994; Garrote and Bras, 1995; Ivanov et al., 2004). Yet, the anticipated utility of such models (e.g., Beven, 1989; Goodrich et al., 1995) to significantly advance the skill to simulate and forecast hydrologic response, to serve as tools for scientific hypothesis testing and to elucidate the complexity of distributed and interacting hydrologic processes in time and space has not yet fully emerged (Finnerty et al., 1997). In attempting to describe the complex spatial behavior of hydrologic systems, distributed models tend to be complex (refering to the detail of process representation) in structure and contain numerous parameters to be estimated. Among the major critiques expressed in the literature are the lack of parameter identifiability, model structural uncertainty and possible overparameterization of the process description (e.g., Refsgaard, 1997; Beven and Freer, 2001; Seibert, 2001; Grayson et al., 2002; Gupta et al., 2005). Nevertheless, spatially distributed hydrological models can provide insights into questions that can not be addressed based on point field observations, laboratory experiments, or lumped models (Beven, 2000).

At present, one of the most severe constraints for the further development of distributed hydrological modelling and its utility for prediction and analysis can be found in the general inability to thoroughly evaluate and constrain the distributed model dynamics with available data (Refsgaard, 1997; Grayson et al., 2002). Usually, model evaluation and calibration has been mainly based on a comparison of observed versus simulated runoff at the basin outlet. It has been pointed out numerous times that this is a very weak constraint on the adequacy of a model and its parameters. Many modeling studies have shown that only matching the simulated and observed integrated catchment response (i.e. streamflow) is no guarantee that the internal, spatially distributed hydrologic response is correct (Grayson and Blöschl, 2000a,b,c; Seibert et al., 1997). Many different parameter combinations and even model structures describing different processes can exist that may produce a wide array of internal states yet very similar runoff outputs.

Spatial observations provide the ability to evaluate the internal behavior of the models, in terms of simulated patterns of state variables and model output. This can not only improve the identification of model parameters and increase the reliability and precision of predictions (e.g., Franks et al., 1998). Seeing how well a measured internal response is simulated provides a much more rigorous test of model structures, process conceptualizations and assumptions. This can provide a more reliable basis to answer questions about what complexity (i.e., level of detail in process representations) is actually needed (Beven, 1989; Grayson et al., 1992; Jakeman and Hornberger, 1993) and how can experimental findings be incorporated to arrive at models that are known to „work", to some more testable degree, for the right reasons, i.e. are consistent with the current process understanding (e.g., Seibert and McDonnell, 2002; Weiler and McDonnell, 2004; Tromp-van Meerveld and Weiler, 2008). By including spatial pattern comparisons in model assessments, we will "improve the confidence with which we can claim our models do indeed represent the right processes and get the right answers for the right reasons" (Grayson and Blöschl, 2000a,b,c).

Spatial data is becoming increasingly available in recent years (Grayson et al., 2002; Vereecken et al., 2007). One promising new technology is emerging with wireless sensor networks (e.g., Cardell-Oliver et al., 2005; Trubilowicz et al., 2009; Bogena et al., 2009, 2010). These networks provide data on important environmental variables such as, for instance, soil moisture content, with unprecedented spatiotemporal resolution and can bridge the scale gap between local hydrogeophysical measurements and remotely based sensor systems (Bogena et al., 2009). Although the technology is still in its infancy, it provides the potential to revolutionize data collection at least at the experimental catchment scale (Soulsby et al., 2008).

Now it is necessary to find ways to make optimal use of this data to constrain the models to be consistent with key signatures this data may contain. While it is a common believe that detailed spatial observations are vital to improving our understanding of catchment hydrological behavior, still relatively little experience seems to exist on what constitutes appropriate data in a given situation and how to make optimal use of the data (Wealands, 2006). This will form an interesting and important field of research in both hydrological process understanding and modeling (Grayson et al., 2002).

1.2. Scope

The primary goal of this thesis is the development of an inverse modeling approach to directly incorporate the high resolution spatiotemporal soil moisture content data produced by a wireless sensor network into distributed hydrological models.

This study uses the distributed model Hill-Vi (Weiler and McDonnell, 2004) and soil moisture data produced by the wireless sensor network SoilNet (Bogena et al., 2010) for a temperate humid forested headwater basin. The inverse modeling application involved

- setting up the model to the study site,
- preparation of the meteorological and sensor network data,
- selection of appropriate objective functions to constrain the model to major aspects of the observed soil moisture data,
- coupling of the model to an effective and efficient parameter sampling scheme and implementations of state-of-the-art inverse modeling methods to allow assessment of model structural and parameter uncertainty.

Finally, this study intends to investigate the importance of representing several processes that have recently received attention in experimental studies, for instance, bedrock seepage, variable soil depth (or bedrock topography), and variable throughfall, for the simulation of internal soil moisture dynamics as well as the runoff response at the catchment outlet to be investigated.

1.3. Background

1.3.1. Hydrological inverse modeling

The hydrological behavior of any hillslope or catchment involves a number of spatially distributed and interacting water, energy and vegetation processes. Therefore every computer-based hydrological model, regardless of how detailed or spatially explicit, how physically-founded or conceptual in nature, is necessarily a simplified and to some degree spatiotemporally aggregated representation of the highly complex and heterogeneous reality (Gupta et al., 2005; Wagener and Gupta, 2005). As a consequence, at least some of the model parameters are – while often still physically interpretable and related to properties of the system – not directly measurable. Instead, they have to be identified via an indirect process of parameter estimation, during which the model parameters are iteratively adjusted such that the model simulations match, as closely and consistently as possible, the observed behavior of the system under study. This process is variously referred to as inverse modeling or model calibration (e.g., Hornberger and Spear, 1981; Young, 1983; Beven, 2005; Wagener and Gupta, 2005; Vereecken et al., 2007). A further important aspect is that while the model structure is most commonly fixed a priori to any modeling attempts (Wheather, 1993), a variety of model structures, representing different degrees of complexity and varying process conceptualizations and assumptions, may appear equally possible for a given situation. The selection process usually amounts to a subjective decision by the modeller (Wagener, 1998), since objective decision criteria are often lacking (Mroczkowski et al., 1997).
Calibration and testing of hydrological models has been an active area of research in recent years. The greater use of complex models has increased the problems of balancing data availability, predictive performance and model complexity (Grayson et al., 2002), which has led to questioning the classical calibration paradigm (Gupta et al., 1998). Sophisticated automatic global optimization algorithms (see Gupta et al., 2005; Vrugt et al., 2008; and

references therein) are now available to reliably locate the global optimum for some predefined mathematical measure of „goodness-of-fit“ to the data (henceforth referred to as objective function; OF). However, it has become clear that usually a large number of models or parameter sets exist that result in very similar values for the selected OF. This is referred to in the optimization literature as the problem of nonuniqueness, indeterminacy, nonidentifiability, or more recently “equifinality” (Beven, 2001). Consequently, the inherent uncertainty in any model application must be explicitly considered during both calibration and prediction. While the entire uncertainty is often projected into the parameter space, it has to be clear that all components of the modeling process, including the measurements of system input and output and the model structure, are uncertain and errorprone (e.g., Wagener and Gupta, 2005).

Several approches exist in the literature to respond to the problem of perceived equifinality.

First, the finding can be interpreted as the need for set theoretic approaches, which assume that all plausible models should be retained unless and until evidence to the contrary becomes apparent. Many of these set theoretic approaches are related to the Regional Sensitivity Analysis (RSA; also called the Hornberger- Spear-Young approach) concept advanced by Spear and Hornberger (1980) that evaluates the sensitivity of the model output to changes in parameters without referring to a specific point in the parameter space. These techniques commonly apply Monte Carlo sampling procedures to explore the feasible parameter space in search for plausible behavioral (the terms „behavioural“ and „non-behavioural“ are often used to describe models that „match“ or „do not match“ the observations (Hornberger and Spear, 1981) models. Examples of the set theoretic approach applied to hydrological modeling include the Generalized Likelihood Uncertainty Estimation (GLUE) technique (Beven and Binley,1992), the Dynamic Identifiability Analysis (DYNIA) approach (Wagener et al., 2003), the Parameter Identification Method based on the Localization of Information (PIMLI) approach (Vrugt et al., 2002), the Monte Carlo Set Membership (MCSM) approach (van Straten and Keesman, 1991), the Explicit Bayesian Approach (Kuczera and Mroczkowski, 1998), the Bayesian Recursive Estimation (BARE) technique (Thiemann et al., 2001), and the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm (Vrugt et al., 2003a).

Second, it has been argued that more powerful methods are needed to properly exploit the information contained in the data. Various research efforts have shown that the amount of information retrieved using a single OF is insufficient to identify more then three to five parameters (e.g. Beven, 1989; Jakeman and Hornberger, 1993; Gupta, 2000). More information can become available through the use of multiple objective functions to increase the discriminative power of the calibration procedure (e.g. Gupta et al., 1998; Gupta, 2000). These measures can either retrieve different types of information from a single time-series, e.g. streamflow (e.g. Gupta et al., 1998; Dunne, 1999; Boyle et al., 2000; Wagener et al., 2001), or describe the performance of individual models with respect to different measured variables, including internal state variables, e.g. groundwater levels (e.g. Kuczera and Mroczkowski, 1998; Seibert, 2000) or saturated areas (Franks et al., 1998). The multi-objective approach proposed by Gupta (1998) based on the concepts of Pareto optimality further allows to gain insights (that may also be used for improvements in the model structure) into consequences of model structural uncertainty by revealing trade-offs between the models capabilities in reproducing several data types or aspects and portions of the data equally well with a single parameter set.

Third, the finding that parameter non-identifiability can be attributed to overly complex model structures with too many tunable parameters given the information content in the data led Wheater et al. (1993), Jakeman and Hornberger (1993), Young et al. (1996); Wagener et al. (2003) to apply more parsimonious model structures with only as many parameters as can confidently be identified. This is in contrast to an argument sometimes made in model development that processes that are perceived to have an effect in the real system should be represented in the model as well (Beven, 2001). However, the increase in identifiability is at the price of a decrease in the number of separate processes described by the model. There is therefore a danger of building a model structure that is too simplistic for the anticipated purpose (Kuczera and Mroczkowski, 1998). An important question is therefore how much complexity is really needed or warranted in hydrological models (Beven, 1989; Grayson et al., 1992; Jakeman and Hornberger, 1993; Tromp-van Meerveld and Weiler, 2008).

1.3.2. Use of spatial patterns in distributed hydrological modeling

The recent advances in calibration and testing methodology described above have highlighted the importance of additional information to augment standard runoff data. Hence several studies have used spatial data to acknowledge the limited amount of information contained in stream flow or any other integrated flux data to identify model parameters (e.g. Wheater et al., 1996; Beven, 1989; Jakeman and Hornberger, 1993; Ye et al., 1997) and to assess issues of model complexity and realism (Beven and Freer, 2001; Seibert and McDonnell, 2002).

Extensive reviews on this topic are found in Grayson et al. (2002) and Wealands (2006). The use of other auxiliary data types has been reviewed e.g. by Seibert and McDonnell (2002).

Important differences in these studies include (1) the type of data or pattern information used, (2) the strategy to compare observed and simulated data, and (3) the strategy to combine multiple data sources. Some key aspects are summarized below.

1.The type of data commonly used to describe spatial patterns can either consist of point measurements, categorical and binary data or surrogate data (i.e. data that shows a degree of correlation to the spatial pattern of interest). In the context of this study, use of point measurements is of particular interest. Point measurements used as basis for spatial model evaluation include soil moisture measurements (e.g., Chirico et al., 2003), snow depth measurements (e.g., Davis et al., 1998), groundwater levels (Lamb et al., 1998; Blazkova et al., 2002), and internal stream stage measurements (Hunter et al., 2005). They may either be used directly or interpolated to the model grid using methods of varying complexity (Grayson et al., 2002) to produce a predicted pattern. Following Grayson and Blöschl (2000), a spatial pattern refers to any image or surface showing the spatial distribution of an attribute, especially where there is a degree of organisation, as opposed to the spatial pattern being random. Interpolated maps of continuously valued data may further be converted to categorical or binary maps. While this results in some loss of information, it allows for application of a different set of evaluation metrics (Wealands, 2006). Important considerations in the use of point measurements are how representative the point measurement is of a larger area and whether there are sufficient measurements to characterize the field and justify interpolation to form a spatial field or pattern (i.e. how the support for the measurement relates to the support of the model, and how the measurement error compares to any underlying pattern in the field; Grayson and Blöschl, 2000; Grayson et al., 2002). For example, Anderton et al. (2002b) found difficulties in using limited soil moisture and phreatic surface information in the validation of the SHE-TRAN model due to both the sparseness of the data and the 'mismatch' of the measurement scale to the model gridscale. Therefore, a simple direct comparison of simulated model variables to observed data for specific points representing intermediate locations on the model grid will be of limited value (Rosso, 1994; Gupta et al., 2005).

Binary patterns include snow cover derived from aerial photographs or from satellite remote sensing (e.g. Owe et al. 2008), uncertain estimates of saturated areas derived from high-resolution synthetic aperture radar (SAR) imagery (Franks et al. (1998) from photographs or high-resolution optical instruments (e.g. Land-Sat TM, AVHRR; see also Jensen and Calabresi (1997), for examples from a range of platforms). Hunter et al. (2005) used SAR and aerial photographs of inundation extent or stream networks (Stoll and Weiler, 2010), while Peschke et al. (1999) mapped the type of runoff generation mechanism that occurs for a given catchment state.

Surrogate patterns are useful when the attribute of interest is difficult to collect, as they provide a means of assessing spatial predictions albeit with greater uncertainty. For example, terrain has been used as a surrogate for solar radiation exposure, soil properties, vegetation distributions, (e.g. Wilson and Gallant, 2000), soil texture to infer hydraulic properties and remote sensing data, for example, for surface soil moisture derived from SAR (e.g., Satalino et al., 2002; Montanari et al., 2009).

2. Comparing observed versus simulated fields has often been limited to visual comparison – arguably a very powerful method, particularly when combined with detailed process understanding (e.g., Tromp-van Meerveld and Weiler, 2008), yet qualitative, subjective, and limited to selected points in time (Grayson et al., 2002; Wealands, 2006). It is thus not possible to extend this method to automated optimization techniques. Quantitative comparison techniques may be categorized into global and local (cell-by-cell) comparisons. In global comparisons, each spatial field is either aggregated into a number or into a graph from which the characteristics are derived. The disagreement in these global characteristics is then used to produce a measure of global similarity or error. This includes basic methods such as simple least squares type errors or bias in comparing the mean of observed and simulated fields (Wealands, 2006). Western et al. (2001) used variograms to describe soil moisture patterns and compare pattern characteristics over time. They further investigated the characterization of spatial connectivity within patterns using connectivity functions. Local or cell-by-cell methods are based on comparing simulated and observed values at each grid cell. Thus Güntner et al. (2004) compared landscape metrics characterizing the general size, shape and arrangement for simulated and observed saturated area patches. The differences (residuals) can be aggregated in terms of measures of error variance (e.g., Lamb et al., 1998) or bias, somewhat similar to statistics used in traditional model evaluation using time series. In map comparison of categorical or binary data, the Kappa measure (Cohen, 1960) is frequently used (e.g., Pontius, 2000; Sciuto, 2009; Stoll and Weiler, 2010). Some studies have employed extensions of the strict cell-by-cell comparison using fuzzy measures (see Ross, 1995), e.g., accounting for shifts in the location of patterns (Grayson and Blöschl, 2002) or accounting for

the cell neighborhood (Constanza, 1989) due to uncertainty of location (Güntner et al., 2004). This is particularly useful to avoid unduly biases in the assessment of model performance when measurements are sparse in time and/or space and/or particularly uncertain. However, no standard does currently exist on which approaches should be used for spatial data comparison in hydrological applications for a given question of interest, data type, and scale or resolution. This is a very new area for hydrology, and techniques are still being trialed and developed. An extensive suite of potentially useful measures for quantitative map comparison has been recently compiled by Wealands (2006). He reviewed methods used in other disciplines such as image processing and pattern recognition with the intention to develop comparison strategies that emulate the powerful yet subjective and non-reproducable approach of visual comparison. He recommended the simultaneous use of multiple comparison measures (see also Legates and McCabe, 1999; Boyle et al., 2000), to focus on functionally important parts of the information contained in the data (see also Gupta et al., 1998), the use of image segmentation and clustering methods to delineate coherent regions within organized data fields, tolerance for unimportant disagreements (i.e., use of fuzzy measures) and the comparison on multiple scales. However, many of the more advanced methods have not yet been applied in model calibration. This may in part be due to inexperience in the interpretation of the results, computational extensiveness and increased effort including user interaction.

Both set theoretic methods and the concept of the Pareto optimality provide useful frameworks to incorporate spatial data in inverse modeling and model evaluation. A set of studies (e.g., Franks et al.,1998; Lamb et al., 1998; Blazkova et al., 2002; Freer et al., 2003; Hunter et al., 2005) have used spatial data sources in conjunction with the GLUE methodology (Beven and Binley, 1992; Freer et al., 1996). They have demonstrated that updating of generalised likelihoods based on such data can substantially reduce the uncertainty in parameter estimates and response predictions as compared to use of only dicharge to constrain the parameter space. Madsen (2003) used the Pareto optimality approach to reveal trade-offs between the model performance of groundwater level simulations and the catchment runoff and to find a balanced optimal solution.

Finally it is noteworthy that integration of spatial data often reveals insufficiencies in the models used concerning individual processes, that could be used for informed model improvement (e.g., Tromp-van Meerveld and Weiler, 2008). Particular data sets may only constrain particular parameters (e.g., Lamb et al, 1998). Also, model calibration using spatial or other auxiliary information can result in reduced runoff simulation performance and may not necessarily reduce runoff prediction uncertainty, while it is often concluded that the realism and definition of internal processes of the model can be increased (e.g., Blazkova et al., 2002; Seibert and McDonnell, 2002).

1.3.3. Soil moisture variability

Soil moisture has a major influence on a range of hydrological processes such as flooding, erosion, solute transport and land–atmosphere interactions, as well as a range of pedogenic processes (Western et al., 2004). The soil moisture content (SMC) of a basin exhibits large spatial and temporal variability. According to Vereecken et al. (2007) understanding, characterizing and predicting this spatial variability is one of the major challenges in hydrologic science. One important step is the improved ability to measure soil moisture at various scales with new techniques, such as remote sensing and geophysical methods (see Vereecken et al., 2007 and Famiglietti et al., 2008; and references therein) or, for instance, wireless sensor networks (e.g., Cardell-Oliver et al., 2005; Trubilowicz et al., 2009; Bogena et al., 2009, 2010). A comprehensive review on the utility and applications of soil moisture data can be found in Vereecken et al. (2007). Only a few relevant aspects can be adressed here.

Several modeling and field studies have been conducted to address the properties of soil moisture spatiotemporal variability across a range of spatial scales. At the small catchment and hillslope scales, soil moisture variability or pattern is determined by water-routing processes (e.g., Dunne et al., 1975; Beven and Kirkby, 1979; Moore et al., 1988) radiative (aspect) effects (Moore et al., 1993), heterogeneity in vegetation (e.g., Tromp-van Meerveld and McDonnell, 2006; Ivanov et al., 2010), and soil characteristics (e.g., Vereecken et al., 2007). At this scale, spatial patterns of soil moisture can excert a major control on the rainfall–runoff response, especially where saturation excess runoff processes dominate (e.g., Merz and Plate, 1997; Grayson et al., 1997; Western and Grayson, 1998). On larger scales, the value of soil moisture data for the prediction of runoff is still under debate (e.g., Parajka et al., 2006).

An important research topic has been the identification and characterization of spatial organization of soil moisture (Grayson et al., 1997; Western et al., 1999; Western and Blöschl, 1999; Rodriguez-Iturbe et al., 1995; Oldak et al., 2002; Thierfelder et al., 2003). Western et al. (2001) found that spatial organization had a significant effect on the rainfall-runoff behaviour with event-based model simulations. Grayson et al. (1997) argued that soil water patterns in temperate regions switch relatively rapid between two different preferred states, controlled by different processes. During the wet state, moisture patterns are highly organized or connected and controled by nonlocal factors (e.g., topography and landscape position or upslope contributing area) given the dominance of lateral soil water fluxes. In the dry state moisture patterns are disorganized (or random) because of the influence of local catchment attributes (e.g., soil and vegetation characteristics and terrain slope), and the predominance of vertical soil water fluxes.

Spatial organization can also be analyzed within a geostatistical framework. However, spatial connectivity is a spatial organization feature that is not captured by standard geostatistics (variograms) or indicator geostatistics (indicator variograms). In contrast, connectivity statistics (Allard, 1993; Allard et al., 1994) provide an appropriate tool for characterizing spatial connectivity (Western et al., 2001). Several studies have advocated that connectivity in shallow soil moisture patterns induces threshold-like changes in runoff in temperate rangeland catchments (e.g., Western et al., 2001). However, James and Roulet (2007) found that this was not the case in a temperate humid forested catchment and attributed this to the differences in climate settings and to the fact that forested catchments exhibit larger variability in soil hydrologic properties than rangelands. Ali et al. (2010) argue that climate may rather act as an indirect control while differences in the dominant runoff processes, i.e. saturation excess overland flow versus perched water tables and shallow subsurface stormflow above low-permability layers, explain the differences.

A variety of studies analyzed soil moisture variability in terms of the spatial variance, spatial standard deviation, and/or coefficient of spatial variation (CV), in relation to the mean moisture content. Conclusions considering these relationships generally varied. For example, several studies reported soil moisture variability to increase with decreasing mean moisture content (e.g. Famiglietti et al., 1999; Hupet and Vanclooster, 2002), while others found opposite trends (e.g. Western and Grayson, 1998; Famiglietti et al., 1998). Owe et al. (1982) observed maximum soil moisture variance in the mid-range of mean soil moisture, resulting in the change of soil moisture variability along a convex-upward curve with increasing mean soil moisture. Crow and Wood (1999) suggest that different relationships may exist for different scales. Although many authors have speculated about the origin of soil moisture variability, there have been only few quantitative studies looking at how different processes act to either increase or decrease the spatial variability of soil moisture. By using the similar media concept, Salvucci (1998) showed how variability in soil texture leads to different soil moisture variability states in different limiting cases. Peters-Lidard et al. (2001) attributed the convex-upward relationship to the heterogeneity of soil texture, suggesting that soil moisture variance increases with drying if the mean soil moisture content is between saturation (i.e., volumetric soil moisture content is equivalent to porosity of the soil) and field capacity of the soil, but that it decreases with drying if mean soil moisture is lower than field capacity of the soil. Hydraulic conductivity of a soil medium is greatly affected by its texture, and the difference in the drainage rate among different soil textures is largest when the soil moisture content is between saturation and field capacity. Albertson and Montaldo (2003) showed how covariances between soil moisture and fluxes, originating from variability in soil moisture, forcing and/or land surface properties, can lead to either an increase or decrease in soil moisture variability. Albertson and Montaldo (2003) showed that heterogeneous atmospheric forcing over the land surface can also result in a variance-mean moisture content relationship that peaks in the mid-range. Teuling and Troch (2005) used model simulations to explain trends for different data sets and show how vegetation, soil and topography controls interact to either create or destroy spatial variance. Vereecken et al. (2007a,b) predicted the relationship between soil moisture variance and mean by stochastic analysis of the unsaturated Brooks-Corey flow in heterogeneous soils and showed that parameters of the moisture retention characteristic and their spatial variability largely determine the shape of this relation.

2. Study site and data sets

2.1. Site description

The study site is the small forested headwater basin Wüstebach (Figure 1). This experimental test site is a subcatchment of the River Rur basin and part of the TERENO Eifel/Lower Rhine Valley Observatory. The basin covers an area of 0.27 km2 (27 ha) and is located in the low mountain ranges approximately 600 m above sea level within the National Park Eifel in [central] Germany. The climate of the area is temperate and maritime with a mean temperature of above 7 °C and a mean annual precipitation typically ranging from 1100 to 1200 mm. Substantial snow coverage of part of the basin can be present for several weeks of the year (Sciuto and Diekkrüger, 2009). The bedrock is composed of Devonian shales with occasional sandstone inclusions. The geomorphology is plateau like with a mean and maximum slopes of 3.6% and 10.4%, respectively. The soils developed on periglacial solifluction layers, with an average thickness of 1 to 2 m. While cambisols predominate on the hillslopes, gleyic soils and half-bogs have developed near the river. The soil texture is loamy silt. The basin is densely vegetated, predominantly by Norway spruce [*Picea abies* (L.) H. Karst.], a species characterized by a shallow root system. This location is also rich in low vegetation like plants that grow in water-saturated areas, e.g., *Sphagnum spp*. and *Cirsium palustre* (L.) Scop. (Sciuto and Diekkrüger, 2009). The plant coverage is about 90% (Bogena et al.,2010). Many irrigation ditches and drainage channels from pre-war period, as well as bunkers and bomb craters from the war exist.

2.2. The sensor network SoilNet and soil moisture data preparation

Bogena et al. (2009; 2010) presented the development of SoilNet. SoilNet is a hybrid wireless soil moisture underground network which consists of soil moisture sensors that are embedded in a new low-cost Zigbee radio network and enables near real-time monitoring of soil moisture variations at high spatial and temporal resolution. SoilNet uses a mixture of underground devices, each wired to several soil sensors, and aboveground router devices. Bogena et al. (2009) developed and validated a semi-empirical model to demonstrate that in the case of a 5-cm soil layer, data communication over longer distances (e.g., 100 m) is possible for most soil conditions. The SoilNet instrumentation of the Wüstebach test site is part of the TERENO activity (TERENO, 2010), and was accomplished in close cooperation with the DFG/TR32 (Transregional Centre 32, 2009). Itcomprises a total of 600 EC-5 sensors and 300 5TE sensors (Decagon Devices) at 150 locations (a combination of 50 sensor units in a 60- by 60-m raster and 100 randomly distributed sensor units) and three depths (5, 20, and 50 cm). Two sensors were installed at each depth with a small separation to increase the measurement volume and to enable the examination of inconsistencies. At the 5- and 50-cm depths, one EC5 and one 5TE sensor were installed, whereas two EC-5 sensors were installed at the 20-cm depth. The network is producing soil moisture content measurements since August 2009 which are stored in a central database with a measurement frequency of 15 min (Bogena et al., 2010).

For the inverse modeling, the simple arithmetic mean of measurements available for the 5- and 20-cm depths at each of the 150 sensors for the time period from August 16, 2009 to August 16, 2010 was used. Each sensor record was carefully examined for obviously “unrealistic” behavior, failures and outliers. Time steps during which more than 1/3 of the 150 sensors failed were excluded from the analysis.

The individual measurements were interpolated to the model grid using ordinary kriging (e.g., Goovaerts, 1997). Spherical variogram functions (with a nugget variance) were fitted to the experimental variogram using least squares non-linear optimization. The kriging was performed for visualization purposes only.

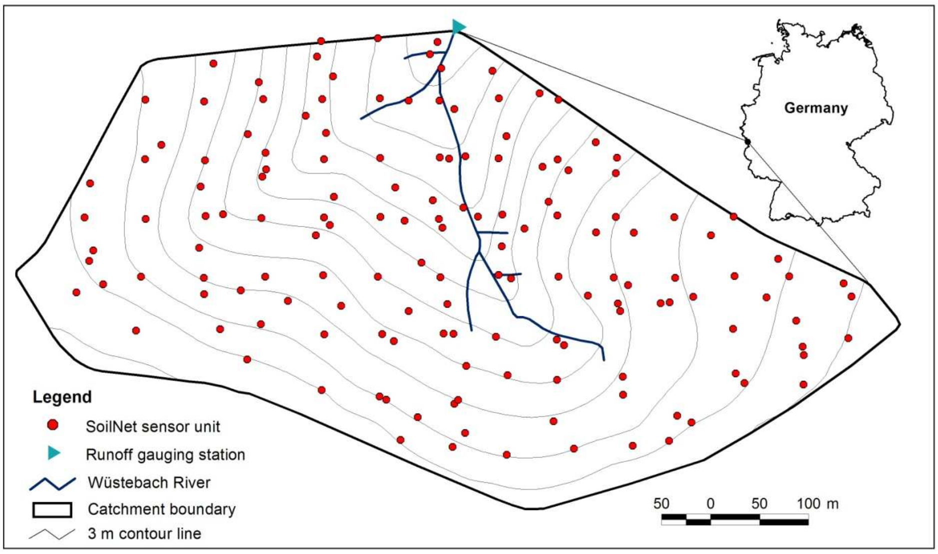


Figure 1: Location and map of the experimental Wüstebach catchment and the SoilNet instrumentation (sensors only).

2.3. Meteorological and hydrometric data

Discharge is monitored at the catchment outlet (Fig. 1). It is noteworthy that the measured discharge time series contains several data gaps, which were disregarded for the analysis. No meteorological data is collected in the basin. Unfortunately, continuous quality-controled meteorological data from nearby stations was also lacking for this study. Instead, the required data was gathered from several other sources. It has to be noted therefore that the meteorological data is subject to unquantified uncertainty in terms of both measurement accuracy and locational representativeness.

Continuous 1-hourly precipitation data was retrieved from the weather station Schleiden-Schöneseiffen of the private weather network Meteomedia (www.meteomedia.de). The station is located approximately 10 km from the basin (latitude: 50.52°N; longitude: 6.37°E) and at a similar elevation (620 m above sea level).

Further meteorological data was required to run the snow routine of the model and to compute the potential evaporation rate according to the Penman combination equation. Data on air temperature, relative humidity, and wind speed at a 1-hour time step was also retrieved for the Meteomedia station Schleiden-Schöneseiffen. Incoming solar or shortwave radiation data was taken from the station Schleiden operated by the Forschungszentrum Jülich (www.fz-juelich.de ODER Heye Bogena, personal communication). The measurements were made at a 15-min resolution and were aggregated to hourly values by arithmetic averaging. Small data gaps (i.e., shorther than 3 hours) in all meteorological data time series (except for precipitation) were filled using linear interpolation between the immediately preceding and following data points. However, the solar radiation data did contain large data gaps during the month of December 2009 as well as during the months of March and April 2010, which are part of the model calibration period. These gaps were filled using incoming solar radiation data from a private weather station (type Davis Vantage Pro 2 aktive plus) located in Monschau-Mützenich (Bodo Friedrich, personal communication; http://www.ew-messnetz.de/stationsdetails/wetterstation-muetzenich.php), approximately 15 km from the Wüstebach basin and at a similar elevation (600 m above sea level). This data was available with a 1- to 2-min time step and again aggregated to hourly values by arithmetic averaging. A further large data gap in the solar radiation data remained for the months of February and March 2009, which are part of the spin up period of the model runs. These gaps were filled using computed clear-sky (i.e., under cloud-free conditions) solar radiation. Making use of a set of standard equations for hourly calculation periods provided in ASCE (1990) and Allen et al. (1999), the values of potentially incoming solar radiation at each time step were first computed as a function of the time of year, the time of day, and latitude. The clear-sky solar radiation was then computed from these values as a function of station elevation, serving as a surrogate for total air mass and atmospheric transmissivity above the measurement site. Given the lack of information on cloudiness, the computed clear sky solar radiation during the period without data was reduced by multiplication with the arithmetic average of the ratio of measured incoming solar radiation and computed clear-sky solar radiation obtained for the preceding and following 30-day period with available measurements to avoid strong bias in the data, and is henceforth assumed to equal the incoming solar radiation.The net shortwave radiation was then computed from measured and computed incoming solar radiaton (e.g., Dingman, 2000) using a fixed value of 0.12 for the albedo based on literature values provided for coniferous forests (Schulla, 1997).To close the radiation balance (e.g., Dingman, 2000) and to obtain the net radiation, the net longwave radiation has to be known in addition to the net shortwave radiation. Here, a standardized procedure (ASCE, 1990); Allen et al., 1999)) which is based on the Brunt (1932, 1952) approach for predicting net surface emissivity is used to compute net longwave radiation as function of the air temperature, actual vapor pressure (computed as a function of air temperature and relative humidity), and the relative short-wave radiation (i.e. the ratio of measured or calculated solar radiation to calculated clear-sky radiation) to indicate relative cloudiness. For nighttime hours (i.e., when the solar radiation equals zero), the value of relative short-wave radiation was computed by linearly interpolating between values occurring 2 hours before and 2 hours after sunset as recommended by Dong et al. (1992). Finally, the obtained continuous time series data of meteorological variables was aggregated to 6-hourly values as arithmetic averages (for air temperature, relative humidity, wind speed, and net radiation) and totals (for precipitation), respectively. All dates and times are given in Central European Time (Coordinated Universal Time plus one hour).

3. Methodology

In this section, the Hill-Vi model (Weiler and McDonnell, 2004) used for this study is first described. Four complexity levels of the model will be investigated. The specifications of these model complexities can be found in Tab. 1. Subsequently, a description of the model set up and the methods used to estimate feasible ranges and fixed values, respectively, for the model parameters is provided. Finally, the inverse modeling strategy used to assess the value of soil moisture data in addition to runoff and to test the different model complexities is described.

Table 1: Specifications of the different Model complexities evaluated.

Model Complexity	Bedrock Seepage	Variable Soildepth	Variable Throughfall
1			
2	**x**		
3	**x**	**x**	
4	**x**	**x**	**x**

3.1. Model description

The model used is the physically-based Hill-Vi model, introduced by Weiler and McDonnell (2004) to study process controls on subsurface flow generation via virtual experiments on hillslopes. The model has been subsequently modified in the context of its various applications (e.g., Weiler and McDonnell, 2006; McGuire et al., 2007; Weiler and McDonnell, 2007; Tromp-van Meerveld and Weiler, 2008). For this work, further modifications were applied, partly in order to increase the computational efficiency and stability during Monte Carlo experiments, and partly to implement or refine representations of processes deemed important in the context of the study and given experience gained during initial model testing. Thereby, attention was paid to stay in line with the original philosophy of the model, that is, to describe the major controls on flow processes while being simple in terms of its structure and number of tunable parameters (Weiler and McDonnell, 2004). Yet, application to a specific real-world domain in a specific climate and landscape and at larger scales compared to hillslopes requires additional processes to be parameterized.

Detailed descriptions of the fundamental concepts of the model may be found in Weiler and McDonnell (2004, 2006, 2007) and McGuire et al. (2007) and are only briefly reviewed below, followed by a description of the model structure as used for this study.

3.1.1. Basic concepts

Hill-Vi is a spatially explicit model, where the model domain is discretized into a uniform raster of grid cells and extends vertically from the soil surface to an impermeable or semi-permeable bedrock. The domain is laterally delimited by no-flow boundaries and may, such as in this case, include a network of channel cells treated as constant head boundaries. All water entering channel cells, including overland and subsurface fluxes as well as channel precipitation is instantaneously removed from the domain as runoff.

The core model (i.e., the soil routine) solves basic continuity equations for tightly coupled unsaturated and saturated zones within each grid cell. This unsaturated-saturated zone coupling was implemented to represent the frequently observed (Dunne, 1978; Bonell, 1998; McGlynn et al., 2002) unsaturated zone conversion to transient saturation during storm events. The unsaturated zone is defined by the depth from the soil surface to the water table and time-variable water content. The saturated zone is defined by the height of the water table above the bedrock surface and the porosity n.

In light of field observations (e.g., Weiler, 2001), an exponential depth function for the drainable porosity n_d (defined by the difference in volumetric water content between 0 and 100 cm of water potential) was included in Hill-Vi (WEILER & MCDONNELL 2006), representing changes in soil structure, macropore development and presence, or increased skeleton content with depth. This function can be written as:

$$n_d(z) = n_0 \exp\left(-\frac{z}{b}\right)$$

where n_0 is the drainable porosity at the soil surface, z is the soil depth below the surface and b is a decay coefficient. Similarly, an exponential decline of the saturated hydraulic conductivity k_S is represented by the following function:

$$k_s(z) = k_0 \exp\left(-\frac{z}{m}\right)$$

where k_0 is the saturated hydraulic conductivity at the soil surface and m is the hydraulic conductivity shape factor.
The core model was extended by simplified formulations to simulate snow storage and melt, interception, variable throughfall, and overland flow routing.

3.1.2. Snow melt

Water may enter the basin as either rain or snow. Snow melt and storage are described using a simple degree-day routine as implemented by Stoll and Weiler (2010). Below a threshold temperature T_t, all precipitation accumulates as snow. When the threshold temperature is exceeded, snowmelt occurs according to

$$MELT = ddf \cdot \max(0, T_a - T_t)$$

where $MELT$ is the melt rate, T_a is the actual air temperature, and ddf is the degree-day factor. Melt water is retained in the snow storage until a specified portion of the snow water equivalent is exceeded and may refreeze when the actual air temperature falls below the threshold temperature:

$$REFR = CFR \cdot ddf \cdot (T_t - T_a)$$

where *REFR* is the amount of refrozen water and *CFR* is the refreezing coefficient.

3.1.3. Interception and Evapotranspiration

An interception routine was implemented, since both smoothing of rainfall intensities and reduction of water input to the system due to temporal storage on and evaporation loss from vegetation surfaces were deemed important aspects for the response and water balance of the densely vegetated study site. This consideration was supported by initial model testing using high-quality meteorological data for the year 2007, since hydrographs were found to be too flashy and the catchment was constantly too wet. The interception model relates changes in the canopy storage C to the gross rainfall rate P, canopy drainage D, and canopy evaporation rate E_c in the form

$$\frac{dC}{dt} = (1 - p)P - D - E_c$$

where p is the free throughfall coefficient. A simple linear threshold model (Calder, 1977; Vrugt et al., 2003) is used to compute canopy drainage, accounting for water losses from leaf dripping and stemflow:

$$D = b(C - s),\ C > s$$

where b is an empirical drainage coefficient and s is the minimum canopy storage capacity. The latter parameter exerts a major control on the total evaporative loss from the vegetative surfaces (i.e., the interception loss). Stem storage and flow is not, such as in more complex models (e.g., Rutter et al., 1975), explicitly represented. The average throughfall T_{av} is then obtained as the sum of free throughfall and canopy drainage.

A stepwise approach, largely following Eltahir and Bras (1993), Wigmosta et al. (1994), and Ivanov et al. (2004), is used to compute the canopy evaporation rate and the actual evapotranspiration rate from the root zone. This approach allows the vegetation to change states from wet to dry during a time step. The evaporation of intercepted water from wet vegetation surfaces is assumed to occur at the potential rate E_p (equal to evaporation from an open water surface), which is appropriate since stomata are not involved. When the canopy storage falls below its minimum capacity, it is assumed that part of the canopy is dry and evaporation is reduced proportionally. The canopy evaporation rate is thus computed according to:

$$E_c = E_p,\ ET = 0,\ C \geq s$$

$$E_c = \frac{C}{s} E_p,\ ET > 0,\ C < s$$

where E_p is the potential evaporation rate. The potential evaporation rate is computed using the Penman combination equation (Penman, 1948):

$$E_p = \frac{1}{\lambda}\left[\frac{\Delta(R_n - G) + \rho_a c_p d / r_a}{\Delta + \gamma}\right]$$

where R_n is the net radiation, G is the ground heat flux, Δ is the slope of the Clausius-Clayperon relationship between saturation vapor pressure and temperature, γ is the psychometric constant, ρ_a is the density of air, c_p is the specific heat of air at a constant pressure, λ_v is the latent heat of vaporization, d is the saturation vapor pressure deficit, and r_a is the aerodynamic resistance. The state variables were calculated using standard meteorological relationships described in ASCE (1990) and Allen et al. (1998).

Subsequently, the actual evapotranspiration rate from the unsaturated zone $ET_{a,un}$ is computed according to:

$$ET_{a,un} = \alpha \cdot \left[(E_p - E_c) \cdot \frac{\Delta + \gamma}{\Delta + \gamma(1 + r_a / r_c)}\right]$$

where α is a soil moisture stress factor and r_c is the canopy resistance to vapor transport. Note that the term in square brackets is equivalent to the Penman-Monteith equation (Monteith, 1965) for time steps where the canopy remains completely dry.

The soil moisture stress factor accounts for the current soil moisture stress limiting the root water uptake and is computed using a simple quasi-linear function (e.g., Davies and Allen, 1973; Federer, 1979, 1982; Spittlehouse and Black, 1981) of relative saturation in the unsaturated zone:

$$\alpha = \min\left\{1, \left[\frac{S_{un}}{S_{max,un}}\right]\right\}$$

where S_{un}, is the actual soil moisture content of the unsaturated zone and $S_{max,un}$ is the maximum storage capacity of the unsaturated zone (obtained by subtracting the drainable porosity from the total porosity). The respective values are obtained by integrating the variables over the time variable vertical extend of the unsaturated zone (i.e., from the soil surface to the water table).

To avoid underestimation of ET in cases where the water table is very close to or at the soil surface, Eq. (11) is solved again, this time integrating over a fixed root zone depth (estimated as 50 cm based on Sciuto and Diekkrüger (2009)) and the result is used with Eq. (10) to compute the actual evapotranspiration rate from the root zone $ET_{a,rz}$. The difference between $ET_{a,un}$ and $ET_{a,rz}$ then yields the actual evapotranspiration from the saturated zone, $ET_{a,sat}$. Note that $ET_{a,sat}$ equals zero when the watertable falls below the root zone depth.

3.1.4. Throughfall Variability

A simple throughfall variability option was implemented in Hill-Vi for this study. The implementation is based on results of a recent study by Keim et al. (2005). This study found that temporally and spatially persistent throughfall patterns existed beneath forest stands. In a deciduous stand, the spatial correlation length was about one crown diameter and maximally 10 m. Given the grid size of the model of 10 m, it seems unnecessary to implement across-grid-cell autocorrelation in throughfall. Instead, a random field following a standard normal distribution was generated (Fig. 2) and used to transform the uniform gross rainfall input into a variable throughfall field at each time step. The (variable) throughfall T_i at a location i given the uniform throughfall input T_{av} is obtained according to:

$$T_i = T_{av} \cdot \left(1 + \tilde{T}_i \cdot CV_T\right)$$

where $\tilde{T}_i$ is the normalized throughfall drawn randomly from a standard normal distribution, and CV_T is the predefined spatial coefficient of variation of throughfall. The CV_T was set to 0.25 and the random field was initially generated and then fixed during the analysis.

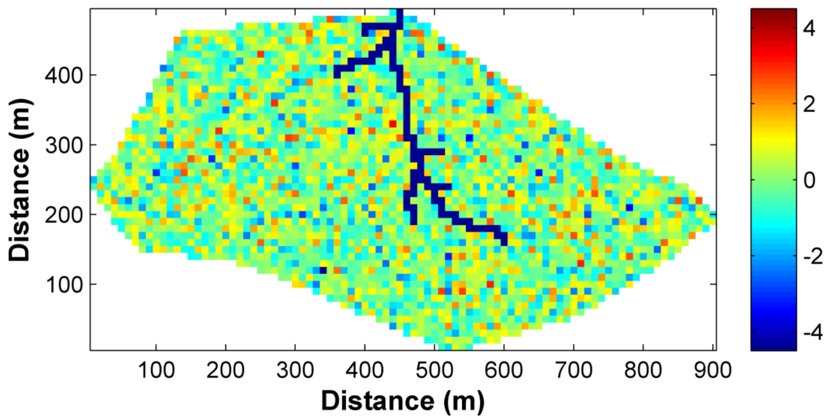


Figure 2: Map of normalized throughfall for the Wüstebach basin used to simulate throughfall variability with the Hill-Vi model.

3.1.5. Overland flow

A new and computationally efficient overland flow routing routine was implemented in Hill-Vi, which allows capturing runoff-runon-effects. In Hill-Vi, overland flow can only be generated by the mechanisms of saturation excess and return flow, i.e. when the soil becomes entirely saturated and no further water can infiltrate or the water table rises above the surface due to subsurface flow convergence. Infiltration excess overland flow is unlikely an important process in temperate forested headwaters with high infiltration capacities and was thus neglected.

The routine is based on the Manning-Strickler equation, which is simple and only requires one parameter, the Manning roughness coefficient, *a*, to be estimated. The overland flow velocity, *v*, is computed according to:

$$v = \frac{1}{a} \cdot h_o^{2/3} \cdot \delta^{1/2}$$

where ho is the overland flow water level (approx. equal to the hydraulic radius) and δ is the local surface slope based on a DEM. The overland flux Q_{of} at a given location is then computed as:

$$Q_{of} = v \cdot h \cdot w$$

where *w* is the flow width, which is set equal to the width of the model grid cells.

The flow direction and partitioning of overland flux leaving a grid cell, respectively, were derived from a DEM using the D∞ algorithm proposed by Tarboton (1997), which assigns a flow direction based on the steepest slope on a triangular facet. If the flow direction falls on cardinal or diagonal direction, then the flow from each cell drains to one neighbour. If the flow direction falls between the direct angles to two adjacent neighbours, the flow is apportioned between these two cells depending on how close the flow direction angle is to the direct angle for those cells.

3.1.6. Soil routine

For each grid cell, the water balance of the unsaturated zone is defined by the infiltration, vertical recharge into the saturated zone, actual evapotranspiration, and change in water content. No lateral flow can occur in the unsaturated zone. Recharge R, from the unsaturated zone to the saturated zone is described by a power law function according to:

$$R = \left(\frac{S_{un}}{S_{\max,un}} \right)^{c} \cdot k_s\left(z'\right)$$

where R is the recharge to the saturated zone, c is the power coefficient reflecting a nonlinear response to increased wetness, $k_s(z')$ is the saturated hydraulic conductivity at the depth of the water table, z'.

The water balance of the saturated zone is defined by the recharge input from the unsaturated zone, the lateral subsurface inflow and outflow, seepage into bedrock (when included) and the corresponding change of water table depth. Lateral subsurface flow, Q_{ssf}, is computed using the Dupuit–Forchheimer assumption (Freeze and Cherry, 1979):

$$Q_{ssf} = T \cdot \beta \cdot w$$

where T is the transmissivity, β is the water table gradient, and w is the flow width. The saturated subsurface flow is routed downslope using an explicit grid cell by grid cell approach (Wigmosta and Lettenmaier, 1999), with the flow direction and thus the partitioning of outflow from each grid cell being recalculated for each time step. This facilitates simulations in the presence of variable soil depth and bedrock topography, respectively (Weiler and McDonnell, 2004).

Seepage S to bedrock is computed on the basis of the hydraulic head above the bedrock surface and the saturated hydraulic conductivity, k_b, of the bedrock (Tromp-van Meerveld and Weiler, 2008):

$$S = k_s \cdot (1 + h_w)$$

where h_w is the height of the water table above the soil-bedrock interface. Under unsaturated conditions, seepage is limited by the recharge from the unsaturated zone. Water entering the bedrock is assumed to be lost from the system.

3.2. Model Set Up and Estimation of Parameters and Feasible Ranges

The model was run and evaluated on a 6-hour time step. The model evaluation is performed for the 1-year period from August 16, 2009 to August 15, 2010. A half-year spin-up period preceding the evaluation period is used to avoid any influences of the pre-specified initial conditions on the model performance assessment.

The 0.27 km^2 model domain was discretized into a uniform 10- by 10-m raster of grid cells based on a DEM (Fig. 3). The channel network (i.e., a set of cells with a constant water level) was derived by rasterization of an available polyline shapefile using a Geographic Information System. The bedrock topography was characterized based on approximately 100 measurements carried out using a Pürckhauer-drill which were interpolated to the model grid using ordinary kriging (Heye Bogena, personal communication). The soil depth for each grid cell (Fig. 3) is then defined by the difference in elevation between the land surface and the bedrock surface. Catchment mean soil depth (Tab. 1) was computed as the arithmetic average of soil depths across the catchment.

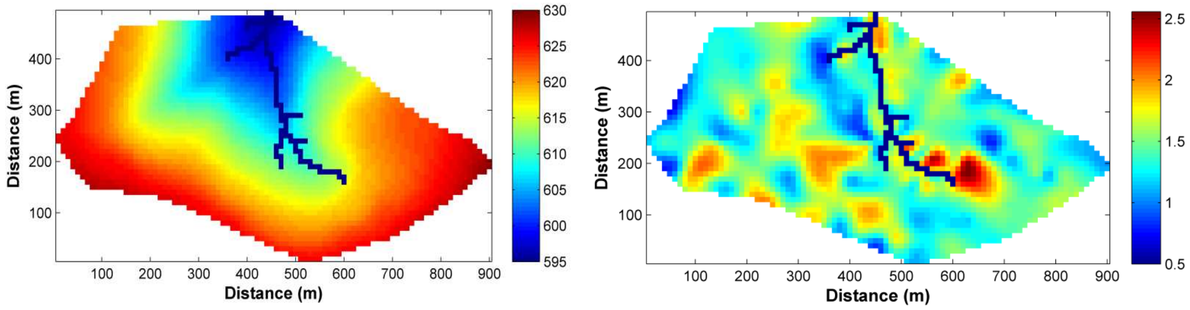


Figure 3: DEM (top) of the Wüstebach basin and map of soil depth (bottom).

The feasible ranges of model parameters related to soil hydraulic properties – i.e. the surface saturated hydraulic conductivity, the surface drainable porosity, and the shape factors of the respective depth functions – were constrained based on available soil data by applying pedotransfer functions (PTFs). Characterization of basin soils was performed by the Geologischer Dienst Nordrhein-Westfalen (Geological Survey of North-Rhine-Westphalia) at a scale of 1:2500 (Fig. 4).The basin area was sub-divided into 61 soil mapping units, each of which was assigned information for three soil horizons, amounting to a total of 183 characterized soil layers. However, most of the soil units have the same properties. There are 31 different soil textures present in the catchment, for which the texture (sand, clay, and silt fractions), humus content, and skeleton (Ø > 2 mm) content are available.

In a first step, the well-known PTFs of Vereecken et al. (1989, 1990) were implemented in the MATLAB environment (MathWorks Inc.) and applied to determine the soil hydraulic parameters of the van Genuchten model for the moisture retention characteristic and the saturated hydraulic conductivity for each soil horizon from soil texture, bulk density, and organic carbon content. The bulk density was computed from the porosity assuming standard densities for mineral (2.65 g m-3) and organic materials (1.2 g m-3). The organic carbon content was determined by dividing the observed humus content by the ratio of organic matter to organic carbon, which was set to 1.72 as suggested by the Bodenkundliche Kartieranleitung (2005). Soil textural fractions were converted to the international scale after log-linear transformation (see Wösten et al., 1998). Further, the resulting retention curve and saturated hydraulic conductivity were corrected for skeleton content, using equations provided in Brakensiek and Rawls (1994). A minor fraction of the topsoil horizons was characterized as 100% organic. Given the high degree of variation in literature values and class PTFs for hydraulic properties of organic or peat soils (e.g., Wösten et al., 2001), estimation of a particular value seemed rather arbitrary and these horizons were therefore disregarded. The drainable porosity was determined as the difference between the saturated water content and the water content at a soil water tension of 100 cm (approximately field capacity) as in Weiler and McDonnell (2004) and McGuire et al. (2006).

In a second step, the parameters of the exponential depth functions for the saturated hydraulic conductivity and the drainable porosity employed in the Hill-Vi model were estimated by fitting these functions to the PTF-based estimates of the respective soil property for the three horizons (estimated properties were assigned to the center of the respective horizon) in each soil unit. This nonlinear least-squares curve fitting problem was solved using the Levenberg-Marquardt non-linear optimization algorithm (e.g., Moré, 1977). Thus, for each soil unit, estimates of the surface drainable porosity and saturated hydraulic conductivity as well as the shape parameters are now obtained.

The feasible range for the total porosity prameter *n* in Hill-Vi was estimated from the observed total porosity ϕ according to:

$$n = (\varphi - RMC) \cdot (1 - Zvol)$$

where RMC is the volumetric residual moisture content estimated based on the PTFs of Vereecken et al. (1989) and *Zvol* is the observed volumetric skeleton content, both determined for a given soil horizon. To obtain values for each soil unit, the computed values for the three horizons were arithemtically averaged.

The recharge power coefficient c of the Hill-Vi model was estimated based on the Brooks and Corey (1964) pore-size distribution index which in turn was determined for each horizon using PTFs of Rawls and Brakensiek (1985) and then arithmetically averaged for each soil unit.

Finally, the model parameter ranges were estimated as the 95 % central range of estimated properties for all soil units. Due to the numerous uncertainties in the parameter estimation approach, including uncertainties in the data itself, in the PTFs, as well issues of scale, the thus obtained limits of the parameter ranges were extended by 25 %. A summary of the estimated feasible parameter ranges is provided in Tab. 2.

Note that the database of Vereecken et al. (1989, 1990) contains data mostly on agricultural soils in Belgium and application to forest soils is thus generally questionable, because these show distinctively different hydraulic properties. Among others, they are less compacted, show a greater aggregate stability and macro-porosity and therefore, a greater saturated hydraulic conductivity and air capacity (e.g., Fisher and Binkley, 2000). However, the PTFs of Vereecken et al. (1989, 1990) have been reported by several studies to be among the most accurate that were evaluated (e.g., Tietje and Tapkenhinrichs, 1993; Romano and Santini, 1997; Cornelis et al., 2001; Wagner et al., 2002; Mermoud and Xu, 2006) and were reported to provide relatively good results even for forest soils (Hammel and Kennel, 2001).

The parameters of the snow model, the saturated hydraulic conductivity of the bedrock and several other auxiliary parameters were taken from the literature and are provided in Tab. 2.

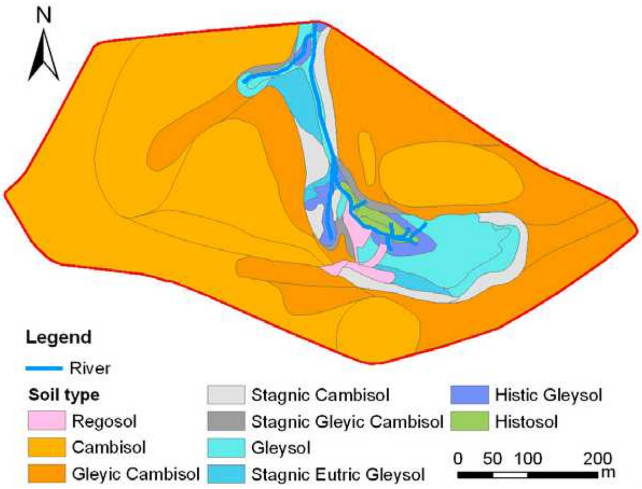


Figure 4: Map of soil types in the Wüstebach basin as provided by the Geological Survey of North-Rhine-Westphalia (modified from Sciuto and Diekkrüger, 2009).

Table 2: The Hill-Vi Model Parameters and Their Respective Ranges Used to Simulate the Hydrologic Response of the Wüstebach basin.

Parameter Symbol	Description	Lower limit	Upper limit	Source
Soil routine				
n	Total porosity	0.23	0.45	soil data and PTFs (Vereecken et al., 1989, 1990)
b (m)	Shape factor for drainable porosity function	0.7	2	soil data and PTFs (Vereecken et al. , 1989, 1990)
n_0	Drainable porosity at the soil surface	0.05	0.18	soil data and PTFs (Vereecken et al. , 1989, 1990)
m (m)	Shape factor for hydraulic conductivity function	0.2	1.9	soil data and PTFs (Vereecken et al. , 1989, 1990)
k_0 (m h^{-1})	Saturated hydraulic conductivity at the soil surface	0.02	0.85	soil data and PTFs (Vereecken et al. , 1989, 1990)
c	Recharge power law exponent	12	26	soil data and PTFs (Rawls and Brakensiek, 1985)

Overland flow routine/routing				
a ($m^{-1/3}$ h)	Manning's roughness coefficient	0.001	0.3	model testing
Snow routine				
ddf ($m\ K^{-1}\ d^{-1}$)	Degree-day factor	3		Seibert (2002) and model testing
T_T °C	Threshold temperature	0		Seibert (2002) and model testing
CWH	Water holding capacity	0.15		Seibert (2002) and model testing
CFR	Refreezing coefficient	0.05		Seibert (2002) and model testing
Sonstige				
k_b ($m\ h^{-1}$)	Hydraulic conductivity of bedrock (m/h)	$1\ 10^{-8}$	$3\ 10^{-4}$	Freeze and Cherry (1976) and hydrogeological map (1:200,000 sheet CC 5502 Cologne)
sd	Mean Soil Depth	1.46		
LAI	Leaf Area Index	6.7		

For the application of the Penman equation and the Penman-Monteith equation, the resistance terms have to be estimated. The aerodynamic resistance r_a was computed following an equation for forest sites provided by Schulla (1997):

$$r_a = 64/(1 + 0.54u)$$

where u is the wind speed measured at a height of 2 m above ground. The canopy resistance r_c was computed following the MORECS scheme (Thompson et al., 1981). The following equation is used during day time periods:

$$\frac{1}{r_c} = \frac{(1 - 0.65^{LAI})}{r_{sc}} + \frac{0.65^{LAI}}{r_{ss}}$$

where r_c is the canopy resistance, r_{sc} is the minimum canopy resistance in case of optimal water supply and dense plant coverage, r_{ss} is the surface resistance for bare soil (set to 150 s m^{-1}), and LAI is the leaf area index. The following equation is used during night time periods (i.e., when the stomata openings are largely closed):

$$\frac{1}{r_c} = \frac{LAI}{2500} + \frac{1}{r_{ss}}$$

Monthly values for the minimum canopy resistance r_{sc} (Tab. 3) were taken from Schulla et al. (1997) and further corrected as a function of air temperature and vapour pressure deficit using equations provided in Wigmotsa et al. (1994).

It should be noted that the decisions to modify or add particular aspects of the model were met on the basis of test runs using quality-controlled data available in the initial phase of this study. Test model runs using the (uncertain) forcing data ultimately used for model simulations and testing did reveal problems to generate sufficient runoff to closely match the hydrograph. The rainfall and runoff data indicate that significant negative bias may be present in the former. Also, the Wüstebach basin may actually extend beyond the currently defined divide, such that the actual runoff contributing area may be larger than considered for this study. Thus, while seemingly needed in light of the test runs, the interception routine was deactivated for the model evaluation runs.

Table 3: Monthly Values of Minimum Canopy Resistances r_{sc} for coniferous forests (modified from Schulla, 1997).

Month	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Okt**	**Nov**	**Dec**
r_{sc} (s m^{-1})	70	70	60	45	45	45	45	50	65	65	70	70

3.3. Inverse modeling strategy

The selection of an appropriate inverse modeling strategy to directly use internal state variables in addition to runoff data involves several important considerations, including (1) which quantities are actually to be compared in terms of the observed and simulated state variable, (2) what are appropriate numerical metrics (OFs) to compare simulated and observed variables given the data and model used, (3) how to assess the added value of the soil moisture data to constrain the parameter space, (4) how to assess improvements in the model structure and (5) how to efficiently and effectively sample the parameter space with respect to multiple criteria. The approaches to deal with these issues are described in the following subsections.

3.3.1. Which quantities are to be compared?

While the comparison of simulated and observed runoff at the outlet of a basin is relatively unambiguous (i.e., the measured discharges may be considered to be essentially the same variable as that predicted by the model), careful consideration has to be given when comparing simulated and observed state variables such as, for instance, soil moisture content. It is important to note that the state variable within a model element is an ‘‘effective’’ state, i.e. the distribution of moisture content within the model element is usually lumped into a single aggregate quantity, both vertically and laterally (e.g. Wagener and Gupta, 2005). Care must be taken to compare variables that are at least similar in their meaning.

The soil moisture sensor measures the volumetric soil moisture content (%-vol/vol; i.e., the ratio of water volume to soil volume) of the fine earth fraction at a point in a specific depth, while the Hill-Vi model produces for each grid element a laterally and vertically lumped value of water storage height in the unsaturated zone excluding a residual moisture content and the vertical extend of the unsaturated model compartment as defined by the depth to the water table. In order to make the two quantities comparable, it was decided to convert both to a value of relative saturation RS, where a value of 0 % corresponds to a moisture content equal to the residual moisture content and a value of 100 % corresponds to a moisture content equal to the total porosity. The observed relative saturation $RS_{obs,i,z}$ at a sensor location i and depth z is therefore computed as:

$$RS_{obs,i,z} = \frac{SMC_{obs,i,z} - RMC_{i,z}}{\varphi_{i,z} - RMC_{i,z}}$$

where $SMC_{obs,i,z}$ is the observed soil moisture content, $\phi_{,i,z}$ is the (uncorrected) total porosity of the fine earth fraction (or the maximum SMC observed) and $RMC_{,i,z}$ is the residual moisture content as obtained using the PTFs (or the minimum SMC observed). The specific values of $\phi_{i,z}$ and RMCi,z were obtained from linear regression equations fitted to observed values of the respective variable for bordering soil layers at each sensor location.

The simulated relative saturation $RS_{sim,i,z}$ at a specific sensor location i (i.e., at a specific grid cell containing the sensor) and depth z is computed as:

$$RS_{sim,i,z} = \begin{cases} 100 \cdot \dfrac{S_{un,i}}{z_i' \cdot n}, \ z_i'(t) < z_i \\ 100, \ z_i' \geq z_i \end{cases}$$

where $S_{un,i,z}$ is the water storage height in the unsaturated zone, z_i' is the depth to the watertable.

For any given sensor and grid cell location, *i*, in the basin the values of RS_{sim} and RS_{obs}, respectively, obtained for the depths of 5- and 20-cm are then arithmetically averaged. This avoids any assumptions about the soil moisture distribution between the sensors. The resulting values are then used for further analysis. In terms of the terminology, it should be noted that the definition of relative saturation as used in this study does not correspond to the common definition of relative saturation (i.e., ratio of water volume to total pore volume).

3.3.2. Objective Functions

A set of four OFs is used to compare the simulated and observed runoff response as well as soil moisture state. The most commonly used OFs to assess the agreement between simulated and observed runoff response are clearly those of the simple least squares type (e.g., Gupta et al., 2005), such as the Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970). However, as outlined above, the data on rainfall and runoff yield available for this study is subject to potentially large uncertainty and bias. Indeed, initial model testing revealed that the model was incapabale of getting even close to matching the observed runoff volume and dynamics at the same time. Also, particularly strong bias was observed during periods of potentially strong snow melt (i.e., during periods of winter high-flows with little rainfall but air temperature rising above zero), indicating insufficiencies in the respective model routine which is, however, not subject to particular interest in this study. Since it is undesirable to use an error metric dominated by irresolvable (in this study) limitations in the data and the model, it was decided to use a measure that is independent of biases between simulated and observed runoff and more appropriate to constrain the runoff dynamics in this case. For this purpose, the coefficient of correlation (CORR) is used:

$$\mathrm{CORR} = \frac{\sum_{t=1}^{n}\left(Q_{sim}(t)-\overline{Q_{sim}}\right)\left(Q_{obs}(t)-\overline{Q_{obs}}\right)}{\sqrt{\sum_{t=1}^{n}\left(Q_{sim}(t)-\overline{Q_{sim}}\right)^2 \sum_{t=1}^{n}\left(Q_{obs}(t)-\overline{Q_{obs}}\right)^2}} \tag{1}$$

where $Q_{sim}(t)$ and *Qobs(t)* denote the simulated and observed runoff yield, respectively, at time step *t*, and the over score operator (as in $\overline{Q_{sim}}$) indicates the temporal average of the variable (here Q_{sim}) over all *n* time intervals considered.

Several OFs are used to constrain the model to adequate representation of various aspects of the soil moisture observations deemed important. Unless indicated otherwise, the following statistical characterizations are based on direct usage of $RS_{obs,i}$ for the sensor location *i* (as distinguished from interpolated values) and of $RS_{sim,i}$ for model grid cell that contain a sensor *i*. As a first OF, the global mean absolute error (MAE) is used as a basic and well-known metric that ensures tracking of observed spatial mean RS over time without particularly emphasizing agreement during phases with high or low relative saturation:

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|RS_{sim}(t)-RS_{obs}(t)\right|$$

where RSsim is the simulated spatial average relative saturation and RSobs is the observed spatial average relative saturation.

Given the common usage and importance (e.g., Vereecken et al., 2007a,b) of the relationship between the spatial variability and the average RS (or SMC), a signature performance criterion (Gupta et al., 2008) was constructed that ensures proper representation of this relationship by the model. This measure, CSMR, is defined as the correlation coefficient between the observed and simulated relationship between the spatial standard deviation and the spatial average relative saturation, discretized into 15 equally spaced bins according to their numeric range, and can be expressed as:

$$\mathrm{CSMR} = \frac{\sum_{k=1}^{K}\left(\sigma_{sim}(k)-\overline{\sigma_{sim}}\right)\left(\sigma_{obs}(k)-\overline{\sigma_{obs}}\right)}{\sqrt{\sum_{k=1}^{K}\left(\sigma_{sim}(k)-\overline{\sigma_{sim}}\right)^2 \sum_{k=1}^{K}\left(\sigma_{obs}(k)-\overline{\sigma_{obs}}\right)^2}}$$

where $\sigma_{sim}(k)$ denotes the simulated average standard deviation during time steps grouped into the bin k (according to the spatial mean RS_{sim}), $\sigma_{obs}(k)$ denotes the observed average standard deviation during time steps grouped into the bin k, K denotes the number of bins and the over score operator denotes temporal averaging.
In order to characterize the error in the distribution of simulated RS at a given time step, a further OF is introduced in form of the mean squared histogram error MSHE. This OF is obtained as the average of the mean squared error between the histograms of observed and simulated RS at any given time step:

$$\mathrm{MSHE} = \frac{1}{n}\sum_{t=1}^{n}\left[\frac{1}{H}\sum_{h=1}^{H}\left| f_{sim}(h,t) - f_{obs}(h,t)\right|\right]$$

where $f_{sim}(h, t)$ denotes the relative frequency associated with the bin h of the histogram of simulated relative saturation at time step t, $f_{obs}(h, t)$ denotes the relative frequency associated with the bin h of the histogram of observed relative saturation at time step t, and H denotes the number of (equally spaced) bins (here, H = 15). While this OF may be rather unconventional, it is expected to enable a more holistic assessment of the error in the RS distribution at a given time step compared to more conventional statistical tests to compare distributions, e.g. the Kolmogorov-Smirnoff two-sample-test (which uses the maximum error between two distributions as the test statistic). Note also that the value of this OF is influenced by inadequacies of the shape of the RS distribution at a given time step, but it is independent of errors in the mean RS and therefore complementary to the MAE. This is because the histogram ranges for simulated and observed fields are determined according to their respective numeric range at a given time step and independent of each other.

3.3.3. Model calibration and identifiability analysis

To assess the value of the soil moisture data and related OFs in addition to runoff data in constraining the parameters of a given model structure and to evaluate the value of incoporating more complexity in the model for the simultaneous simulation of soil moisture and runoff behavior and with regard to the identifiability of the parameters, several state-of-the-art inverse modling methods are applied as described below. Instead of simply evaluating the performance of the models in terms of a single „best“ parameter set, these methods enable to determine the uncertainty in the parameters (resulting from a lack of identifiability) and, to a certain degree, the uncertainty in the model structure by retaining multiple parameter sets according to differing concepts. Further, the methods can be used to propagate these uncertainties into the model predictions or output.

3.3.3.1. Asessment of model identifiability

Following Wagener and Kollat (2007), model identification here refers to „the identification of all models (parameter sets) within a given model structure, that can be considered feasible (behavioral) representations of the natural system under study". The narrower the space that these chosen models cover within the feasible parameter space, the more identifiable is the model. Lack of identifiability is present when different combinations of parameters (e.g., Johnston and Pilgrim, 1976; Beven and Binley, 1992) yield similar results in terms of the defined OFs.

The procedure applied here largely follows methodologies proposed by Wagener et al. (1999, 2001, 2003) and makes extensive use of computer code provided with the Monte Carlo analysis toolbox (MCAT; Wagener et al., 1999, 2001; Wagener and Kollat, 2007). The core of this methodology is based on the concept of RSA (Spear and Hornberger, 1980; Hornberger and Spear, 1981), and its extension to the GLUE technique (Beven and Binley, 1992; Freer et al., 1996) which has been extensively applied to estimate the uncertainty associated with model outputs and parameter estimates and also to assess the value of spatial data (e.g., Lamb et al., 1998). A major difference to the methodology of Wagener et al. (2001, 2003) or most GLUE applications in general is that parameter samples generated using the Multiobjective Shuffled Complex Evolution Metropolis (MOSCEM-UA) algorithm by Vrugt et al. (2003a) as described below form the basis of the analysis instead of Monte Carlo sampling based on a uniform prior distribution. The basic steps in the procedure applied here are described in the following.

As in GLUE, no single optimum parameter set is identified. Instead, recognizing that it is impossible to identify a single "best" model given errors in both the model and the data, a set of models is selected where each model has a certain likelihood (pseudo probability) of being the correct representation of the system. Likelihoods or likelihood functions are any performance metrics that can be used to differentiate how likely it is that the model (i.e. a specific parameter set and model structure combination) is representative for the system at hand. The likelihood function value is then used to distinguish between behavioral (i.e. acceptable) and non-behavioral (i.e. non-acceptable) solutions. Likelihood measures must have the characteristic that they sum to unity, are greater than zero and higher values indicate better performing parameter values. The objective function values $F(\varphi)$ for each given realization of a model and a parameter set φ were transformed into likelihood values, $L(\varphi)$, following the methodology used by Blasone et al. (2008):

$$L\left(\phi|Y\right)=\frac{1}{F\left(\phi\right)}\cdot\frac{1}{G}$$

where Y indicates a set of observations and G is a scaling constant, which ensures that the cumulative sum of $L(\varphi|Y)$ over all the behavioral parameter sets equals unity.

An important property of the GLUE methodology for this study is that it allows including multiple sources of information in the likelihood function and thus in the uncertainty estimation procedure. Multiple criteria can be accounted for in different ways. The most common aggregation method used in GLUE applications (Freer et al., 1996; Lamb et al., 1998) is to perform Bayesian updating, i.e. by further conditioning the likelihood function, L, when data of different types are available:

$$L\left(\phi|Y_{1,2}\right)=L\left(\phi|Y_2\right)\cdot L\left(\phi|Y_1\right)/G$$

where $L(\varphi|Y_{1,2})$ is the posterior likelihood function of the parameter set φ obtained after conditioning on the observed variables Y_1 and Y_2, $L(\varphi|Y_{1,2})$ is the prior likelihood of the parameter set φ calculated using the observation set $Y1$ and $L(\varphi|Y_{1,2})$ is the likelihood measure calculated with the observations Y_2 and G is a scaling constant, which ensures that the cumulative sum of $L(\varphi|Y_{1,2})$ over all the behavioral parameter sets equals unity.

A user-defined threshold criterion is then required to select the set of behavioral solutions. This may be either a percentage of best performing models or a subjectively selected OF value. Here, subjective thresholds were set first specifically for each of the OFs. Relatively weak thresholds were chosen to eliminate inaceptably performing models (CORR = 0.3; MAE = 15; CSMR = 0; MSHE = 0.02) primarily with regard to the Likelihood function updating using multiple objectives. Subsequently, the best performing 10 % of the evaluated parameter sets in terms of a given likelihood function and combined Likelihood functions, respectively, are retained as behavioral. The likelihood functions of the accepted solutions are then rescaled again.

To propagate the parameter uncertainty into the model output, the 95 % central range of outputs based on the behavioral parameter sets is computed at each point in time. In addition, an output estimate is computed as the median of this distribution and a further output estimate is computed based on the best or most likely parameter set.

To assess the uncertainty (which is assumed to be the opposite of parameter identifiability) in posterior parameter distributions, derived from conditioning an initial distribution on a selected OF, methods developed by Wagener et al. (2001) are used. The best performing 10 % of the parameter population are selected and their cumulative distribution is computed. The gradient of the cumulative distribution is the marginal probability distribution of the parameter, and therefore an indicator of the strength of the conditioning by the data, and of the identifiability of the parameter. To obtain a measure of identifiability, the range of each parameter is segmented into 10 containers and the gradient in each container is computed. The highest value marks the location (or segment) of greatest identifiability of the parameter. Further, the cumulative distributions can be used to derive confidence limits for the different parameters (here 90 %). Wide confidence limits suggest that parameter values associated with equally good performance are distributed widely over the parameter space, while narrow limits suggest that the best performing parameters are focused in a small area of the feasible range. By computing the identifiability after conditioning on the various OFs, the information provided by the OF to constrain the parameter space can be assessed. By repeating the computations after combining the likelihood functions, the value of using runoff and soil moisture data simultaneously can be investigated.

3.3.3.2. Assesment of model structural uncertainty

The methods described above analyze and propagate parameter uncertainty. Several researchers (e.g., Yapo et al., 1996; Gupta et al., 1998) emphasized that the factor currently limiting model performance is model (structural) error arising from the imperfect and aggregated representation of the real system. It is therefore generally advisable to explicitly address the uncertainty originating from model structural inadequacies and errors and particularly of interest to assess the usefulness of increased model complexity. Yet, the nature of model structural error does not allow the estimation of a probabilistic structure (e.g., in the construction of an appropriate OF) to describe it, since the errors are not random in a probabilistic sense (Gupta et al., 1998). However, some of the consequences of this uncertainty can be detected and even used for improvements in the model structure.

A major consequence of model structural imperfection is that the model is incapable of reproducing all aspects and portions of the catchment behavior equally well with a single parameter set (Gupta et al., 1998). Thus, structural insufficiency and uncertainty does become visible in the finding that strongly differing parameter sets are required to enable the model to reproduce, for instance, the hydrograph and different aspects soil moisture state simultaneously, with strong trade-offs between the conflicting objectives. This can be analyzed in terms of the multiple-criteria strategy for watershed model parameter estimation proposed by Gupta et al. (1998) based on methods from the field of economic analysis (Pareto, 1906). Following Gupta et al. (1998), the multi-criteria model calibration problem can then be formally stated as the optimization problem:

$$\min_{\phi \in \Phi} F(\phi) = \left[F_1(\phi), F_2(\phi),\ \dots\ , F_m(\phi) \right]$$

where the goal is to find the parameter set φ within the feasible set Φ that simultaneously minimizes all of the m criteria (here, m equals 4). This problem does not, in general, have a unique solution that simultaneously optimizes each criterion due to errors in the model structure (and other possible sources). Instead, it is generally necessary to adopt a Pareto set of solutions (often times referred to as the trade-off set, non-inferior set, non-dominated set, or the efficient set) which have the property that moving from one solution to another will result in the improvement of at least one criterion while causing deterioration in at least one other. It is thus impossible to distinguish any of the Pareto solutions as being objectively better than any of the other Pareto solutions, such that the Pareto set defines the minimum uncertainty in the parameter selection that can be achieved without stating a subjective relative preference for minimizing one specific component of $F(\varphi)$ at the expense of another (Gupta et al., 1998; Vrugt et al., 2003a). However, the identified Pareto set can be used by the analysis of multiple objectives which allows to evaluate the correlation and trade-offs between different objective functions. Further, the Pareto-solution set can be used to generate a Pareto-ensemble of simulated responses and can be displayed as a trade-off-uncertainty region on the runoff or soil mositure timeseries plots. This shows the uncertainty in the model simulations due to different possible ways of trading-off the model errors (and other errors) (Gupta et al., 1998). The use of Pareto parameter sets to represent model structural uncertainty and Pareto-ensemble simulations to represent model output uncertainty can provide useful ways for evaluating models and their performance.

3.3.3.3. Parameter sampling

Both approaches used here to assess model related uncertainties require sampling of the parameter space with respect to multiple criteria. The selection of an appropriate sampling strategy is particularly important when using distributed models, which are more computationally demanding than lumped models and require a larger number of model runs for calibration and uncertainty asessment owing to their complexity (Blasone et al., 2008). While lumped model runs can be performed in seconds, even a simple distributed model like Hill-Vi requires several minutes for a single run (depending on the size of the domain and number of time steps evaluated). This clearly limits the applicability of the commonly used sampling schemes such as uniform random or Latin hypercube sampling with distributed models (McMichael et al., 2006). Instead, a more efficient sampling scheme is preferable that further has the capability to handle multiple OFs.

Blasone et al. (2008) have recently demonstrated that using a Markov chain Monte Carlo (MCMC) sampling scheme in combination with GLUE significantly improves the efficiency and effectiveness of the methodology. In their revised version of the GLUE procedure, the shuffled complex evolution metropolis (SCEM-UA) algorithm by Vrugt et al. (2003b) is used as sampler of the prior parameter distributions. They further used a flexible objective (likelihood) function, balancing different calibration criteria to include multiple information in their uncertainty assessment. However, since an estimate of the Pareto solution set (see below) is also required in this study, the Multiobjective Shuffled Complex Evolution Metropolis (MOSCEM-UA) algorithm by Vrugt et al. (2003a) is used here. The MOSCEM-UA algorithm is a MCMC sampler that merges the strengths of complex shuffling employed in the shuffled complex evolution (SCE-UA) algorithm (Duan et al., 1992) with the probabilistic covariance-based search methodology of the Metropolis algorithm and an improved fitness assignment concept of Zitzler and Thiele (1999) to construct an efficient and uniform estimate of the Pareto solution set. It uses an innovative concept of Pareto dominance rather than direct-objective function evaluations (such as the SCEM-UA algorithm) and is capable of generating a fairly uniform approximation of the ‘‘true’’ Pareto frontier (which should include the single-criteria end points of the Pareto solution set) within a single optimization run (Vrugt et al., 2003a).

A MATLAB implementation of the MOSCEM-UA algorithm provided by Hoshin V. Gupta (personal communication) was used for this study. Since the Hill-Vi model was programmed in IDL, the computer code of both the algorithm and the model had to be modified to facilitate their simultaneous usage. Since IDL programs are not compilable, a communication (i.e. transfer of generated parameter samples and model outputs) between the programming environments was set up via text files.

For each model complexity investigated, the MOSCEM-UA algorithm was run for all possible combinations (six in total) of the four objective functions CORR, MAE, MSHE, and CSMR. The algorithm (as used here) has two algorithmic parameters that must be specified by the user. The number of complexes was set to a value of two and the population size was set to a value of 16 (i.e., eight members per complex) as recommended by Jasper Vrugt (personal communication). For each pair of OFs, the algorithm was run for 22 loops (resulting in a total of 368 model evaluations). The entire population of parameter sets visited during the individual MOSCEM-UA runs for a given model complexity were then merged (resulting in a total of 2208 parameter sets and model evaluations per complexity) and used as bais for the further analysis. This parameter sampling strategy should generally be well-suited to obtain (1) a sufficiently dense sample of parameter sets in the high probability density region of the feasible parameter space, (2) a reasonable estimate of the Pareto set, and (3) maintain computational feasibility at the same time. It provides the advantage of allowing for both the GLUE and Pareto concepts to be applied from a single sample. However, the rather small number of model runs feasible given the temporal constraints imposed on this work is likely insufficient to obtain a reliable estimate of the high probability density region in the parameter space and the Pareto set such that the subsequent analysis is subject to uncertainty.

4. Results and Discussion

A set of state-of-the-art inverse modeling methods was coupled to the Hill-Vi to provide the ability for a detailed assessment of the model structural and parameter uncertainty. However, while a large amount of effort was spent to compile an appropriate data set and set up the model such as to properly suite the catchment under study, no sufficient agreement between the model and the data could be established that would allow for reliable conclusions about either the value of the soil moisture data to constrain the model parameterization or for the rejection/justification of an increase in model complexity. Also, it is evident that the MOSCEM-UA algorithm did not properly converge during several trials. This should not be attributed to the incapability of the algorithm, but rather to the limited number of model evaluations that could be conducted.

Nevertheless, a suite of figures is provided that should, however, be considered as indicative of the possibilities provided by the framework set up in this work to evaluate the model with respect to the soil moisture data. Fig. 5 shows identifiability plots of the hill-Vi model parameters by conditioning on the various OFs and for various model complexities. High gradients in the cumulative distribution indicate high identifiability in the top performing model parameters whereas shallower gradients indicate low identifiability. The maximum gradient can be considered as a metric of parameter identifiability, and likewise the percentile ranges associated with the cumulative parameter distribution shown in Fig. 6. Given these preliminary results, the runoff data clearly imposes a much stronger constraint on the model parameterization than the relative saturation and associated OFs, repectively. When both data sources are combined, the resulting parameter uncertainty is clearly larger compared to the case when only the runoff-based OF CORR is used. In fact, this is not an unlikely result, given that adequate representation of runoff and relative saturation may provide conflicting targets to track for a model. The existence of a strong trade-off between the ability of the model to simultaneously match the various OFs used in this study is indicated by the multi-objective plots shown in Fig. 7. Note that these two-dimensional plots show the four-criteria rather than the two-criteria Pareto optimal sets, i.e. the rank one solutions with respect to all four criteria. Fig. 7 further indicates problems of algorithm onvergence to the true extend of the Pareto front.

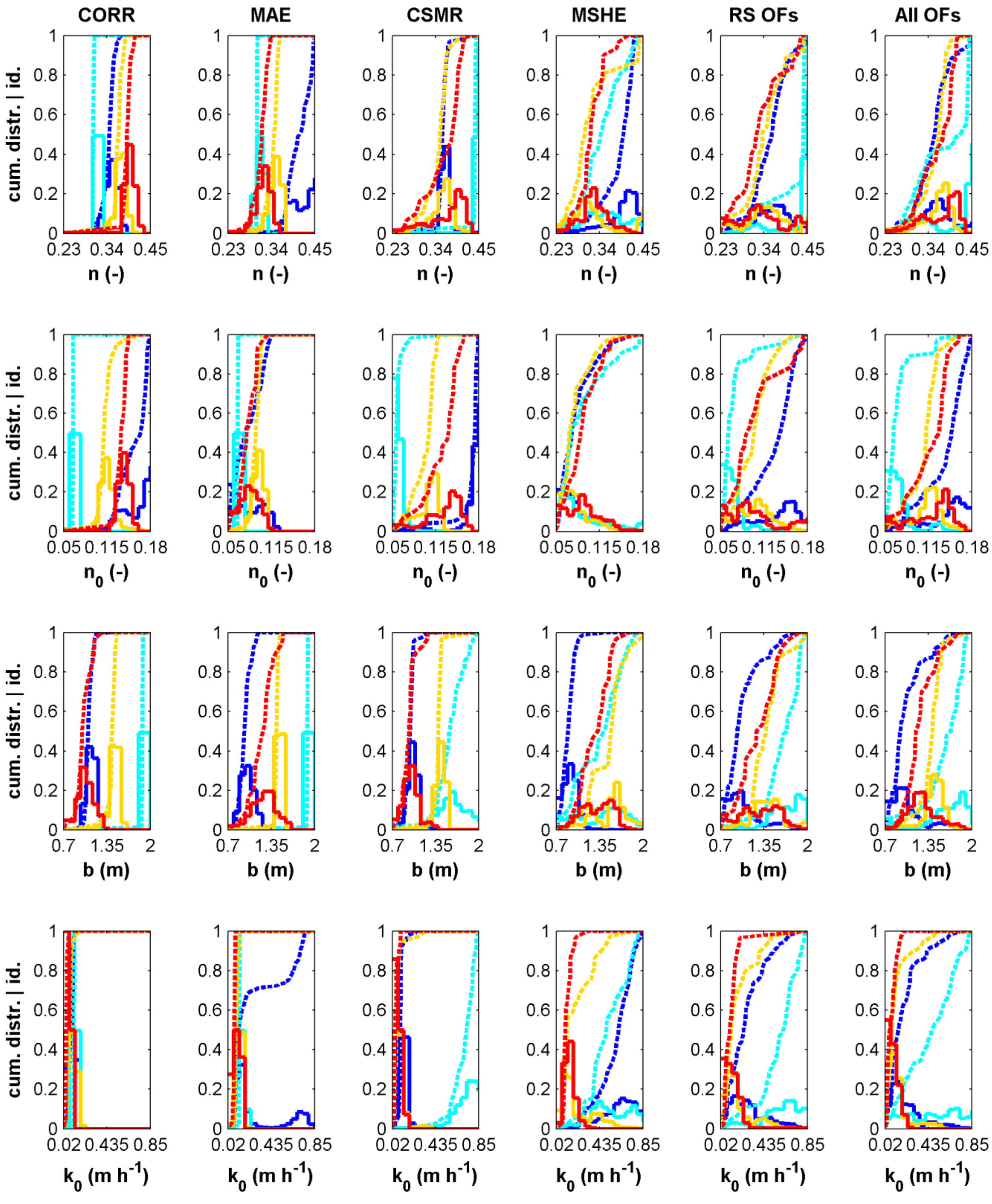
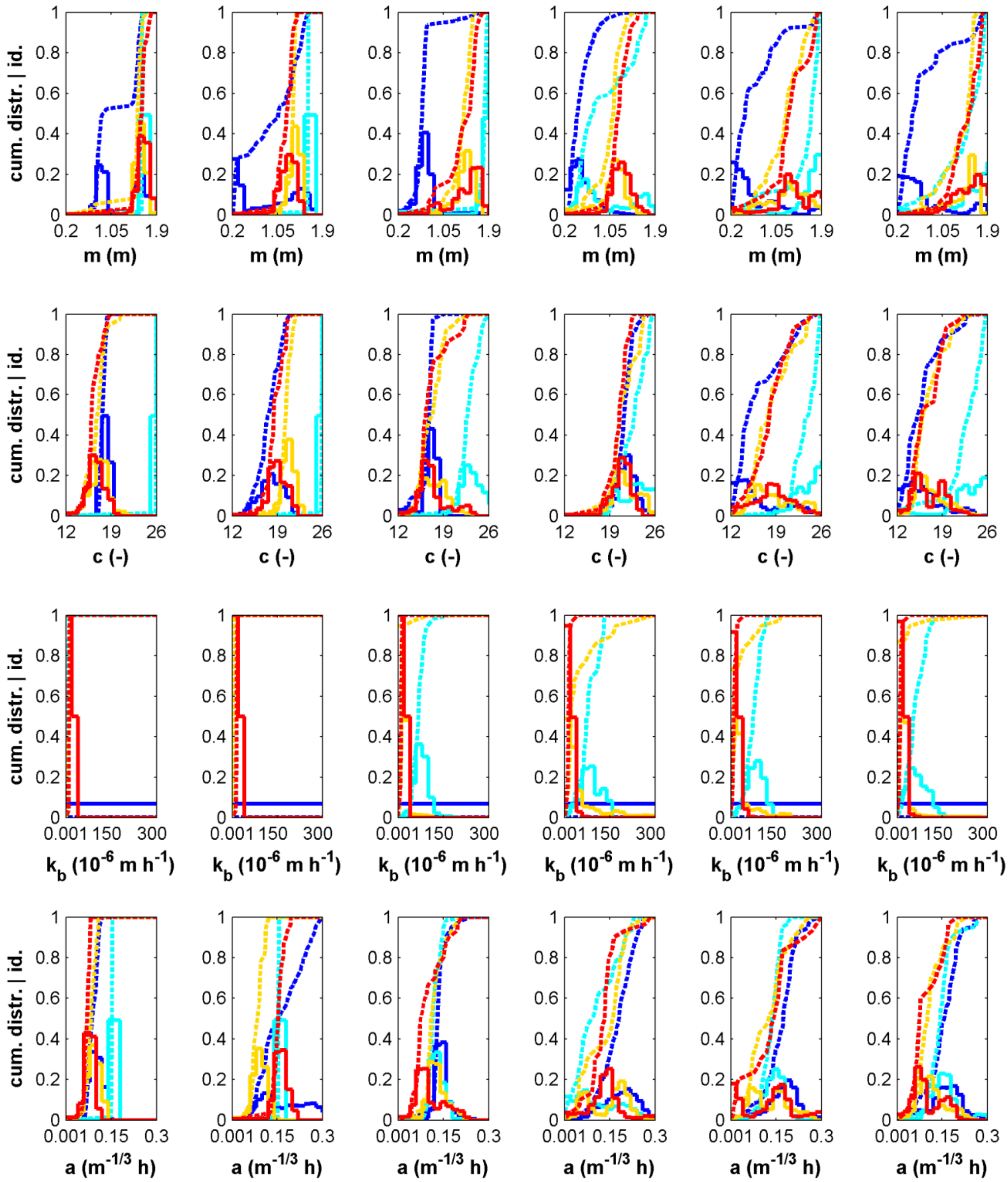


Figure 5: Identifiability plots showing the cumulative distribution (dashed lines) of the top 10-percent of the parameter populations for model complexities 1 (blue), 2 (turquoise), 3 (orange), and 4 (red) and in terms of the objective function values for (from left to right) CORR, MAE, CSMR, MSHE, all OFs with respect to relative saturation data and all OFs. Stair plots indicate the (rescaled) distribution of gradients of the cumulative distribution across the parameter range splitted into ten bins.

(Figure 5 continued)



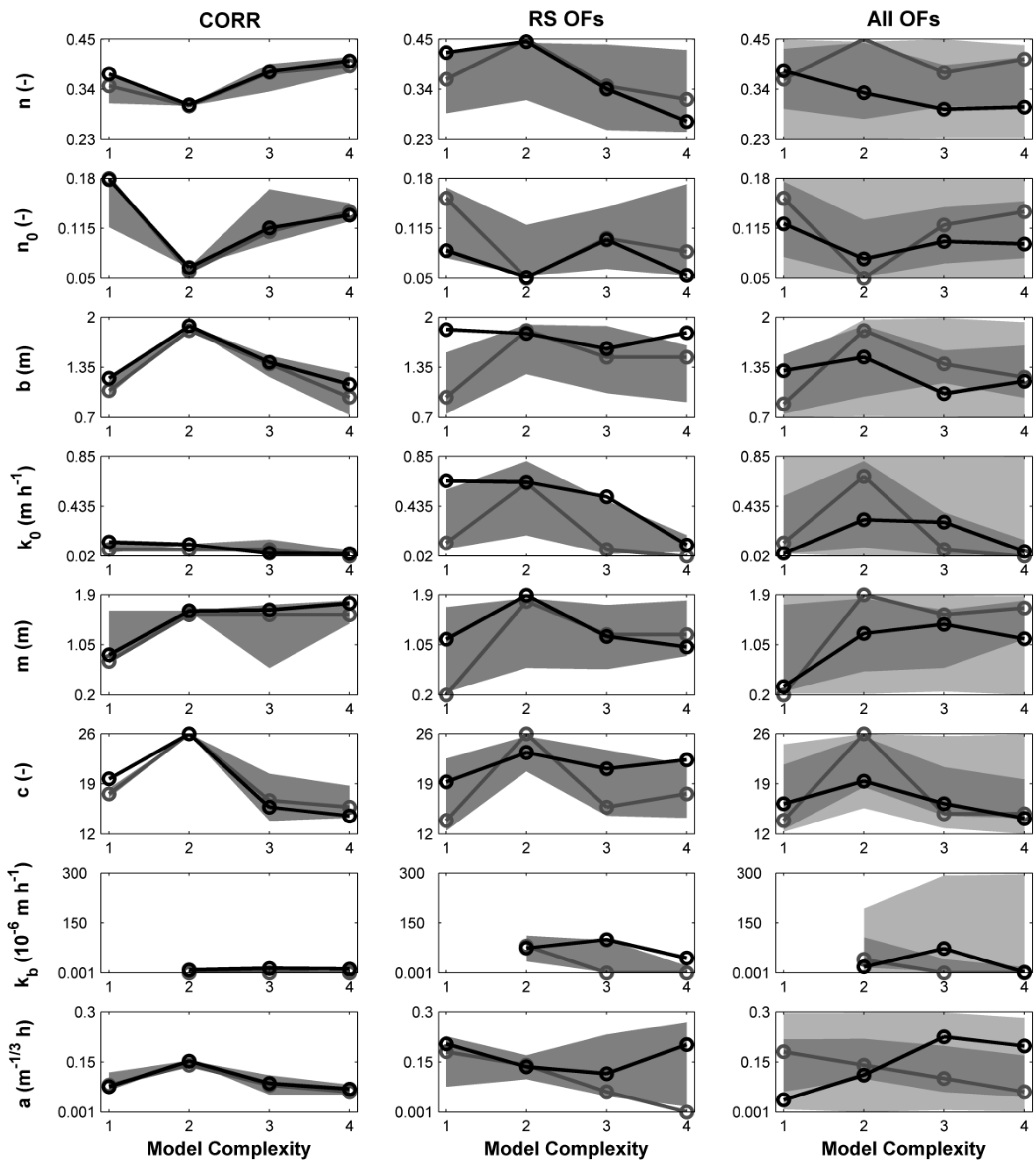


Figure 6: Parameter uncertainly of the Hill-Vi model vs. the model complexity after conditioning on runoff data only (OF: CORR; left column), relative saturation data only (OFs: MAE, CMSR, MSHE; central column), and all data and objective functions, respectively (right column). Light gray shading indicates parameter uncertainty associated with the Pareto set of solutions. Medium grey shading and dark grey line indicate the 95 % confidence limits and median parameter values based on the cumulative parameter distributions (top 10 %). The best performing parameter set is shown as a black line.

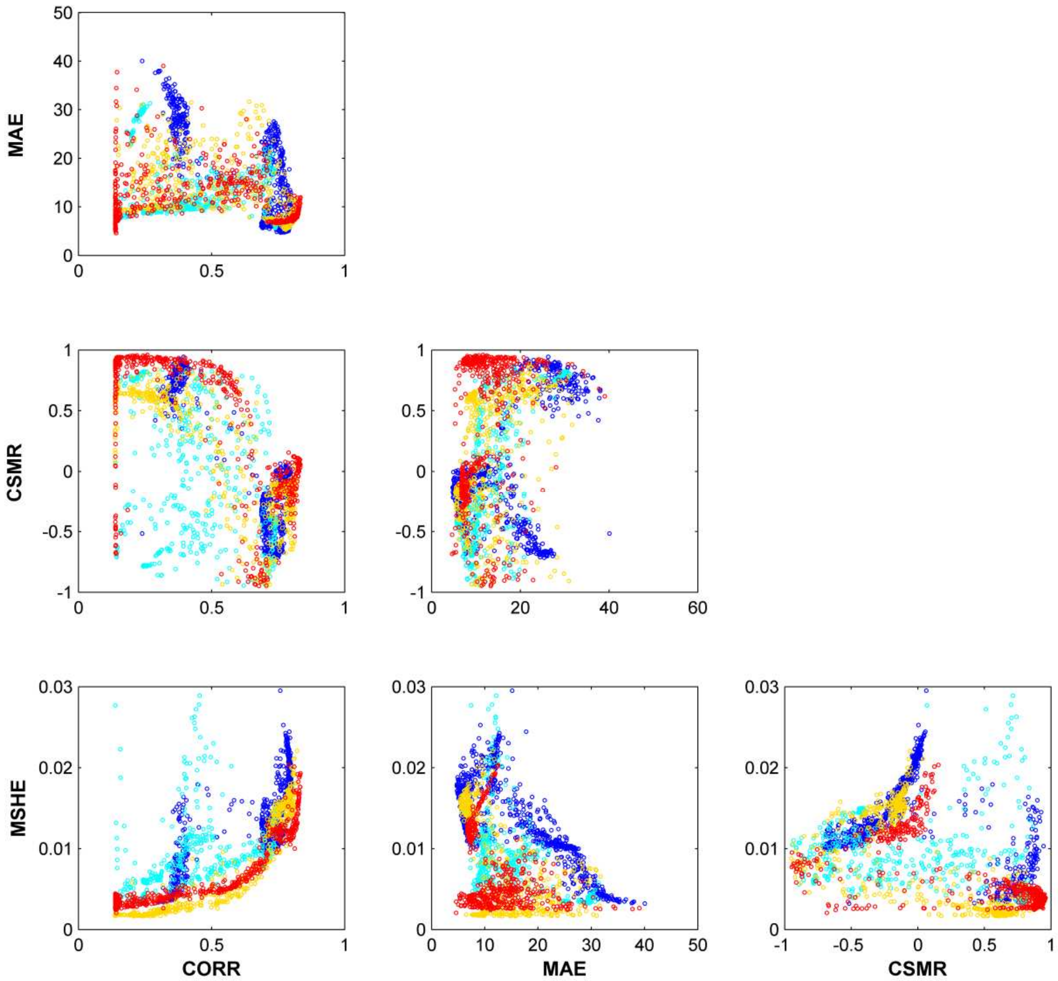


Figure 7: Four-criteria trade-off surfaces in two-dimensional objective spaces for model complexities 1 (blue), 2 (turquoise), 3 (orange), and 4 (red).

The parameter uncertainty is propagated into the output uncertainty as shown in Figs 8 to 11. Constraining the data to runoff does result in extremely narrow output confidence intervals for both runoff and basin mean relative saturation simulations. The precision of the simulated response is high, yet the accuracy of the runoff simulations (Figs 8 and 9) is low. While “acceptably“ high correlation coefficients were archieved for all of the four model complexities, none of the models is capable of achieving a bias smaller than 35 % in simulated flow volumes. Given a runoff coefficient of 0.77 over the evaluation period, this is not surprising. Clearly, the model fails to closely match the runoff particularly during high flow periods in the winter, which can be related to the insufficiency of the snow module. This is true for all simulations, including the single-objective optimal solution. Including the soil moisture data to condition the model parameterization leads to strongly widened confidence intervals, reflecting the increased parameter uncertainty (Figs 5 and 6). Note also the wide uncertainty intervals associated with the Pareto set of solutions, indicating the strong trade-off between the ability of the model to represent the behaviors associated with the multiple-objectives simultaneously.

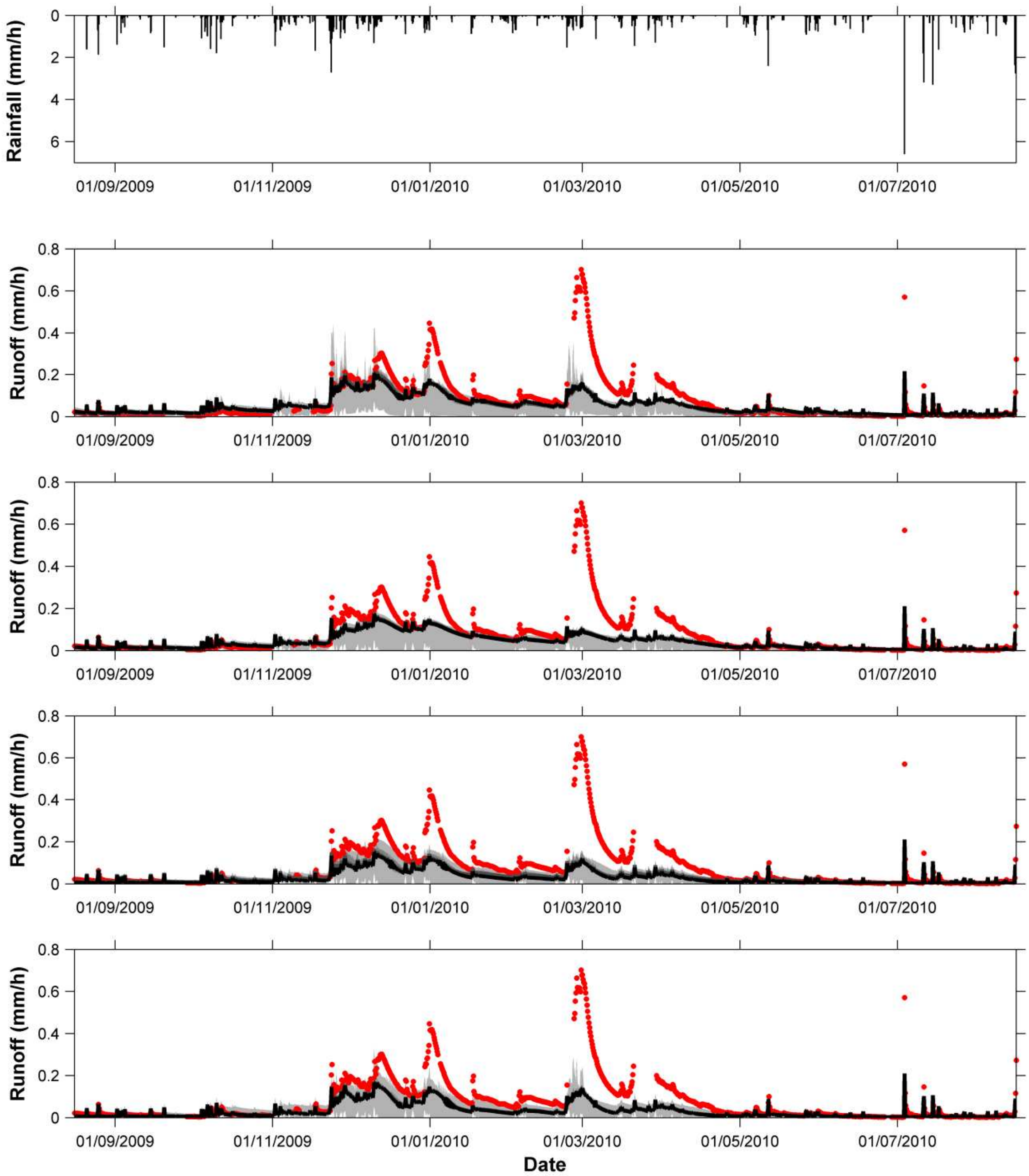


Figure 8: Timeseries of rainfall and observed (red) and simulated runoff for parameter populations conditioned on runoff only and for model complexities 1 to 4 (top to bottom). Light grey shading indicates output uncertainty ranges associated with the Pareto set of solutions. Medium grey shading and dark grey line indicate the 95 % central range based on the behavioral parameter sets. The output associated with the best parameter set is shown as a black line.

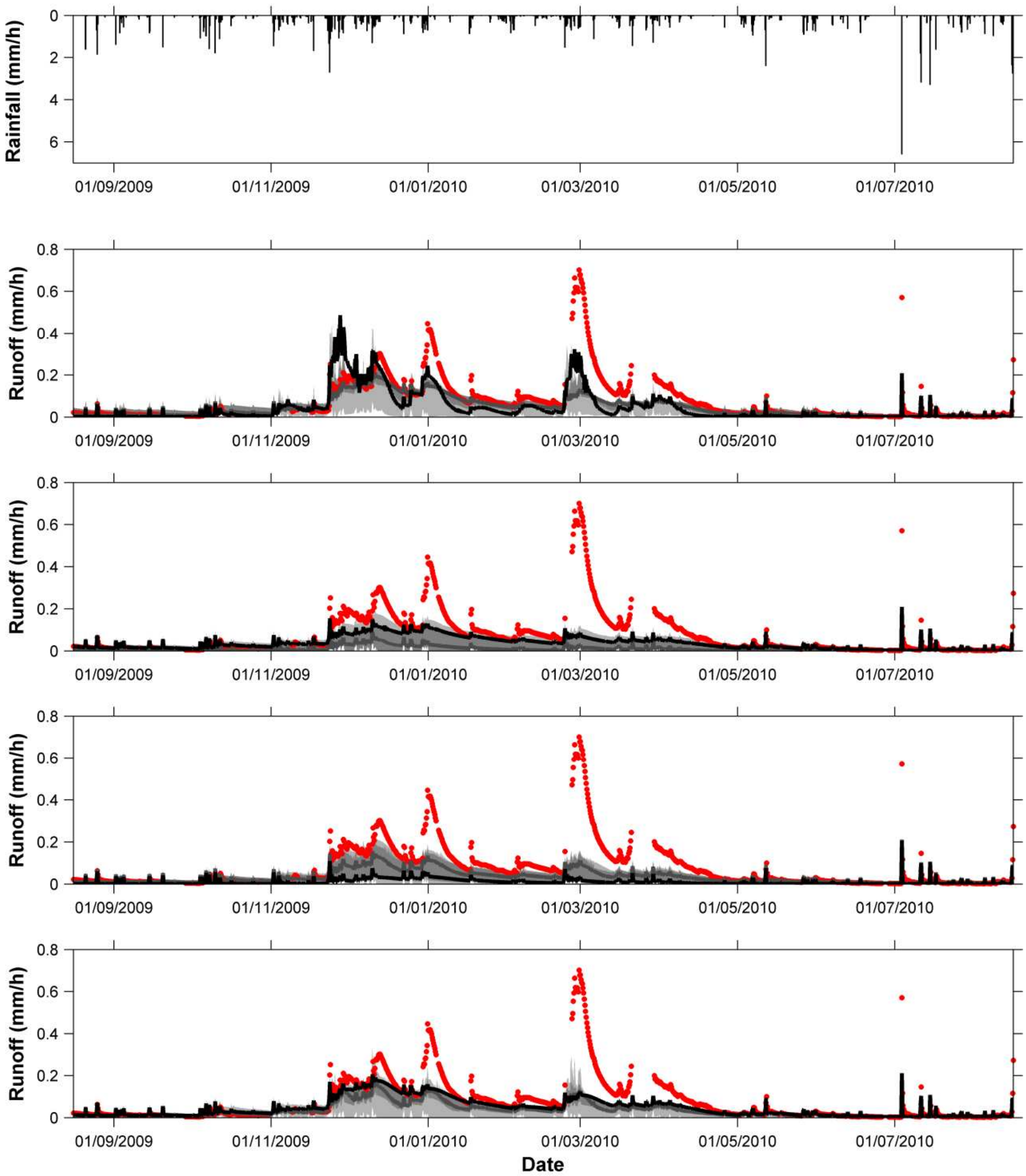


Figure 9: Timeseries of rainfall and observed (red) and simulated basin mean relative saturation for parameter populations conditioned on runoff and relative saturation observations and for model complexities 1 to 4 (top to bottom). Light grey shading indicates output uncertainty ranges associated with the Pareto set of solutions. Medium grey shading and dark grey line indicate the 95 % central range based on the behavioral parameter sets. The output associated with the best parameter set is shown as a black line.

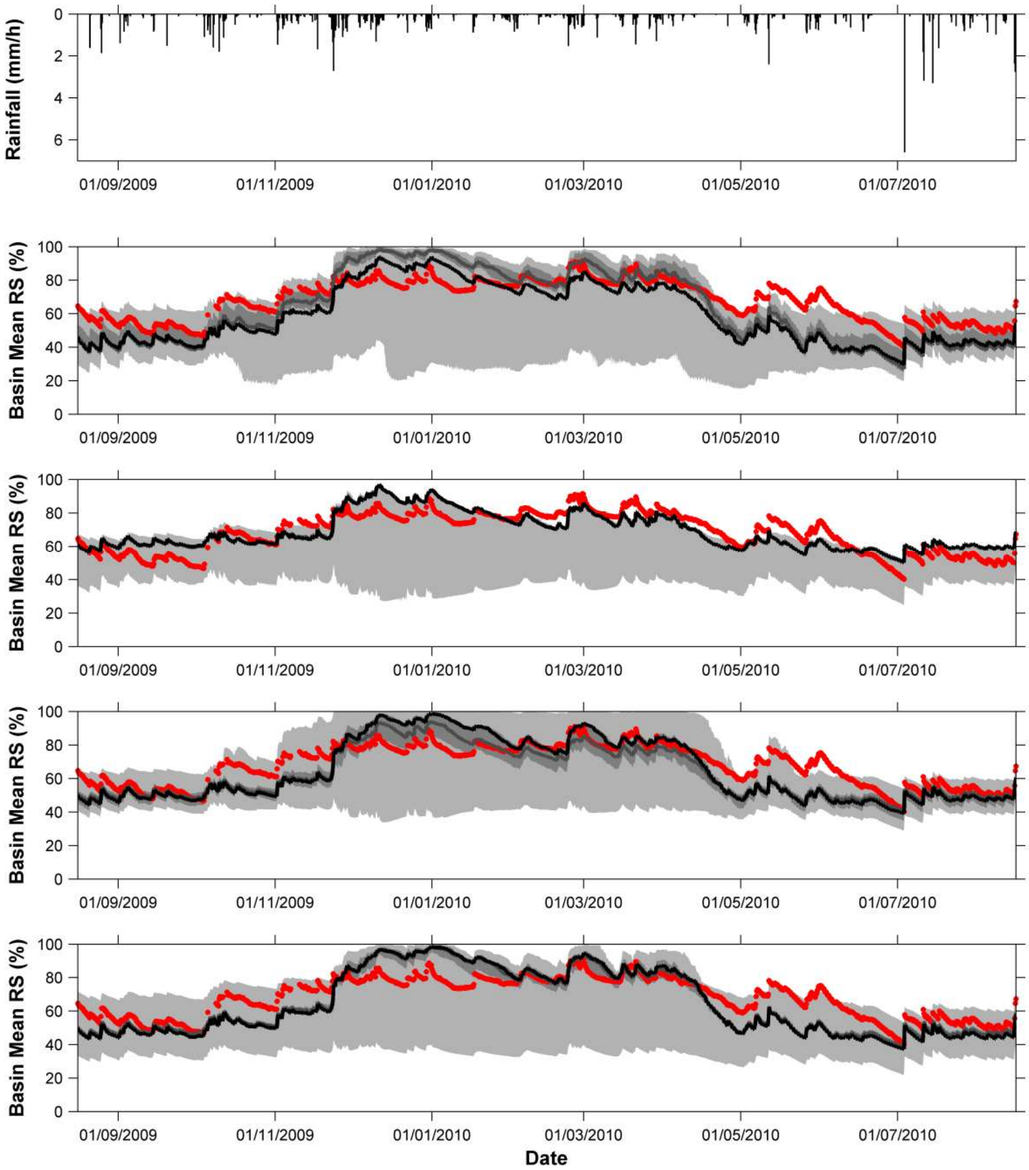


Figure 10: Timeseries of rainfall and observed (red) and simulated basin mean relative saturation for parameter populations conditioned on runoff only and for model complexities 1 to 4 (top to bottom). Light grey shading indicates output uncertainty ranges associated with the Pareto set of solutions. Medium grey shading and dark grey line indicate the 95 % central range based on the behavioral parameter sets. The output associated with the best parameter set is shown as a black line.

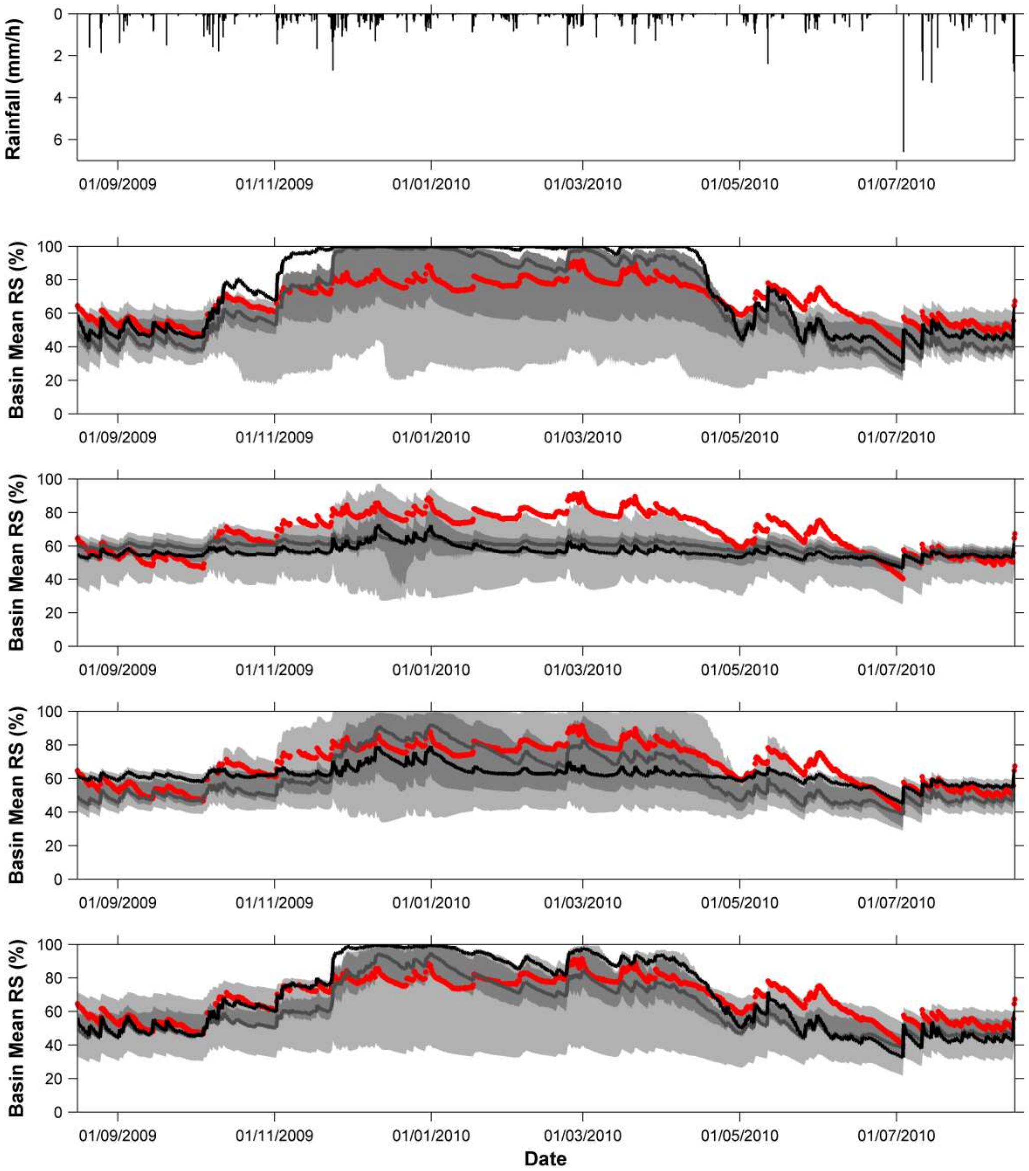


Figure 11: Timeseries of rainfall and observed (red) and simulated basin mean relative saturation for parameter populations conditioned on runoff and relative saturation observations and for model complexities 1 to 4 (top to bottom). Light grey shading indicates output uncertainty ranges associated with the Pareto set of solutions. Medium grey shading and dark grey line indicate the 95 % central range based on the behavioral parameter sets. The output associated with the best parameter set is shown as a black line.

As an example, Fig. 12 shows a map of observed relative saturation on January 15, 2009, which is representative of a wet catchment state, when lateral water movement can be expected to be the main process controlling the soil moisture variability across the entire catchment. It will be interesting to further investigate such maps for different diagnostic periods to assess deficiencies and benefits, respectively, of vaying model complexities.

Table 4: Model performance with respect to the four objective functions for the best parameter set and the best parameter set after conditioning (cond) the model to all objective functions.

	CORR		**MAE**		**CSMR**		**MSHE**	
	best	cond	best	cond	best	cond	best	cond
1	0.20	0.21	7.2	10.4	0.94	0.57	0.0032	0.0100
2	0.25	0.26	6.5	8.5	0.86	0.42	0.0032	0.0088
3	0.18	0.18	7.1	8.2	0.90	0.46	0.0016	0.0113
4	0.175	0.19	7.5	9.3	0.93	0.5	0.0024	0.0101

The assessment of model complexity using auxiliary spatial data has proven a valuable approach (e.g., Tromp-van Meerveld and Weiler, 2008). As for now, increasing the model complexity did result in neither improved identifiability of model parameters (Figs 6 and 7) nor in improved performance (Tab. 4) in simulating the observed runoff and soil moisture dynamics. However, given the uncertainties involved in this study, it would be assumptive to draw any conclusions on the importance of the investigated model refinements. Clearly, the impact of bedrock seepage (e.g., Tromp-van Meerveld and Weiler, 2008), throughfall variability (e.g., Keim et al., 2005), and variable soil depth (e.g., Woods and Rowe, 1996; Freer et al., 1997) on the simulated runoff response and soil moisture variability and connectivity deserve further investigation. It remains to be investigated whether or not soil moisture data is appropriate to constrain the Hill-Vi model and judge the value of the applied modifications. Further modifications may as well be tested, such as the impact of pipeflow and spatially variable soil properties. A representation of pipeflow is implemented in Hill-Vi (Weiler and McDonnell, 2007) and a parameter-free representation of variable soil properties between cambisols and near stream stagnic and gleyic soils was already implemented for the study site and is ready to be tested.

In this study, only OFs have been used that may be classified as global. This avoids the otherwise necessary locationally explicit comparison of point measurments and model grid simulations or interpolation to the grid assuming a fitted statistical model to be true. A further practical advantage of using global metrics is the ability to apply them in cases where a model has a stochastic component, such as the throughfall variability representation implemented for this study. It is, however, still unclear which OF should be used for a given data. Further metrics were implemented in MATLAB are ready to be used when more reliable data becomes available and with more model evaluations. Among others, connectifity functions

(Western et al., 2001) were implemented based on pseudo code provided by Andrew Western (personal communication). These functions provide promising means for characterizing organized features that exist in observed spatial fields and that can have an important influence on hydrologic behavior. Based on these functions, the integral connectivity scale can be computed and used to asess as a measure of the presence of hydrologic connectivity. The correlation of the relation between mean relative saturation and standard deviation also provides an interesting function to investigate. Note that the standard deviation versus mean relation may be better suited for the OF type used here as compared to the coefficient of variation, given that the former was often found to have a convex rather than linear form (Owe et al., 1982), thus imposing a stronger constraint on the value of a correlation coefficient. An interesting aspect to consider is that this function may even be applied to constrain the parameter space and evaluate models even when no soil moisture data is available. This could be implemented, for example, by assuming an appropriate functional form (e.g., Famiglietti et al. 2008; Vereecken et al., 2008) for a standardized soil moisture variability-mean relation for a given location and scale the model is to be applied to and then evaluating the standardized model simulations against this relation.

This study has unwantedly demonstrated that it is highly important to be able to appropriatetely characterize not only one state variable such as, for instance, soil moisture, but to have reliable data available with respect to all major hydrological fluxes and ideally additional important state variables such as groundwater level as well as soil properties. As stated by Vereecken et al. (2008), the collection of such data sets at the catchment scale is an important challenge that should be addressed in the terrestrial observatories that are currently being established. It is hoped that more appropriate forcing data will be available for the Wüstebach site to make optimal use of the high-quality soil moisture data in the framework of hydrological inverse modeling.

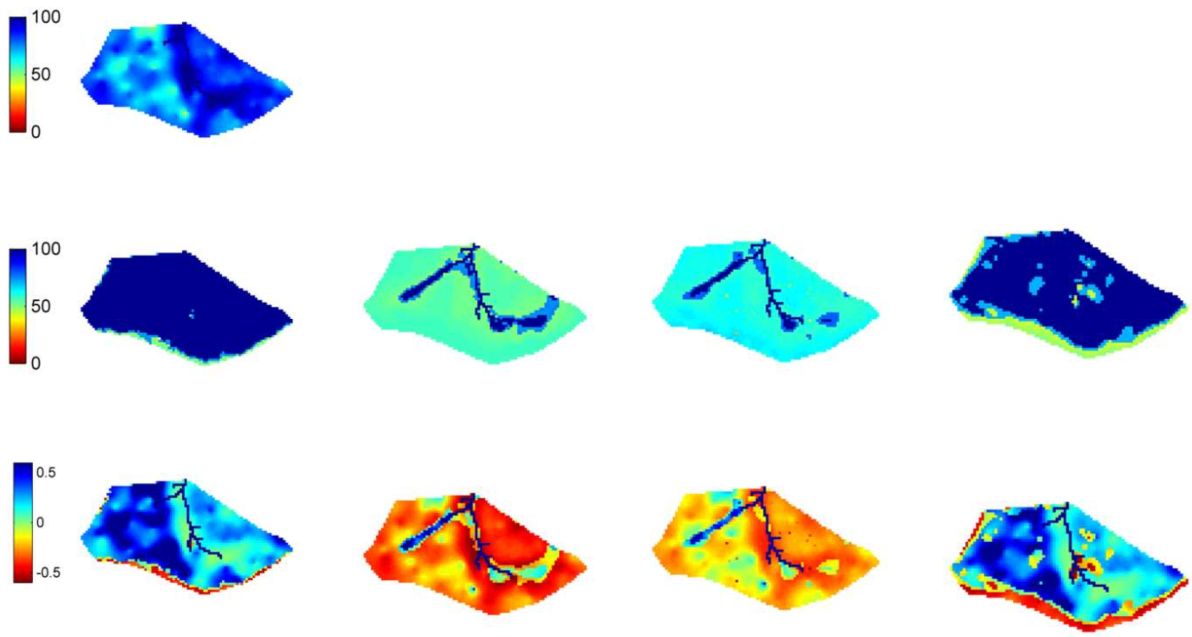


Figure 12: Top: Map of observed relative saturation (%) during a wet period on January 15, 2009 (interpolated using Ordinary Kriging). Middle row: Maps of simulated relative saturation using increasingly complex models (from left to right) conditioned on relative saturation and runoff. Bottom row: Maps of relative residuals for the increasingly complex models (from left to right).

References

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J. (1986) An introduction to the European Hydrological System – Système Hydrologique Européen, "SHE," 1: History and philosophy of a physically-based, distributed modelling system, Journal of Hydrology, 87, 45-59.

Abbott, M.B. and Refsgaard J.C. (eds) (1996) Distributed hydrological modelling, Water Science and Technology Library, 22, 321 pp.

Albertson, J. D. and Montaldo, N. (2003) Temporal dynamics of soil moisture variability: 1. Theoretical basis, Water Resources Research, 39(10), 1274, doi:10.1029/2002WR001616.

Ali, G. A., Roy, A.G. and Legendre, P. (2010) Spatial relationships between soil moisture patterns and topographic variables at multiple scales in a humid temperate forested catchment, Water Resources Research, 46, W10526, doi:10.1029/2009WR008804.

Allard, J.F. (1993) Sound proppagation in porous media: Modeling sound absorbing materials. Elsevier, London.

Allard, J.F., Castagnede, B., Henry, M. and Lauriks, W. (1994) Evaluation of tortuosity in acoustic porous materilas saturated by air. Revue of Scientific Instruments, 65, 754-755.

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M. (1998) Crop evapotranspiration – Guidelines for computing crop water requiresments - FAO Irrigation and drainage paper 56, FAO Rome, M-56, ISBN: 92-5-10421.

Anderton, S.P., Latron, J., White, S. M., Llorens, P., Gallart, F., Salvany, C., O'Connell, P. E. (2002) Internal evaluation of a physically-based distributed model using data from a Mediterranean mountain catchment, Hydrology and Earth System Sciences, Volume 6, Issue 1, 2002, pp. 67-84.

Anderson, M. G. and Bates, P. D. (eds) (2001) Model Validation: Perspectives in Hydrological Science, Wiley: Chichester.

ASCE - Task Committee on definition of criteria for evaluation of watershed models (1993) Criteria for evaluation of watershed models, Journal of irrigation and drainage engineering, American Society of Civil Engineers, 119, 429-442.

Beier, C., Hansen, K. and Gundersen, P. (1993) Spatial variability of throughfall fluxes in a spruce forest. Environmental Pollutution, 81, 257- 267.

Beven, K. J. (1987) Towards the use of catchment geomorphology in flood frequency predictions, Earth Surface Processes and Landforms, 12, 1, 69-82.

Beven, K.J. (1989) Changing ideas in hydrology - The case of physically based models, Journal of Hydrology, 105, 1-2, 157-172.

Beven, K.J. (2000) Uniqueness of place and process representations in hydrological modelling, Hydrology and Earth System Sciences, 4, 203–213.

Beven, K. J. (2005) On the concept of model structural error, Water Science and Technology 52,167-175.

Beven, K.J. and Kirkby, M. J. (1979) A physically based variable contributing area model of basin hydrology, Hydrological Sciences Bulletin, 24, 43-69.

Beven, K.J., Calven, A. and Morris, E.M. (1987) The Institute of Hydrology Distributed Model, Institute of Hydrology Report 98, Wallingford, UK.

Beven, K.J. and Binley, A.M. (1992) The future of distributed models: Model calibration and uncertainty prediction, Hydrological Processes, 6, 279-298.

Beven, K.J. and Freer, J. (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, Journal of Hydrology, 249, 11-29.

Blasone, R., Vrugt, J.A., Madesen, H., Rosbjerg, D., Robinson, B.A. and Zyvoloski, G.A. (2008) Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling, Advances in Water Resources, 31, 630-648.

Blazkova, S., Beven, K.J., Kulasova, A. (2002) On constraining TOPMODEL hydrograph simulations using partial saturated area information, Hydrological Processes, 16, 441-458.

Blazkova, S. and Beven, K. J. (2002) Flood Frequency Estimation by Continuous Simulation for a Catchment treated as Ungauged (with Uncertainty), Water Resources Research, 38, doi: 10.1029/2001/WR000500.

Bodenkundliche Kartieranleitung (2005) Ad-hoc-AG Boden. In Kommission: E. Schweizerbart´sche Verlagsbuchhandlung, Stuttgart, 438 pp.

Bogena, H.R., Huisman, J.A., Meier, H., Rosenbaum U. and Weuthen, A. (2009) Hybrid wireless underground sensor networks: Quantification of signal attenuation due to soil adsorption. Vadose Zone Journal, 8, 755-761.

Bogena, H.R., Herbst, M., Huismann, J.A., Rosenbaum, U., Weuthen, A. and Vereecken, H. (2010) Potential of Wireless Sensor Networks for Measuring Soil Water Content Variability, Vadose Zone Journal, 9, 1002-1013.

Bonell, M. (1993) Progress in the understanding of runoff generation dynamics in forests, Journal of Hydrology, 150, 2-4, 217-275.

Bonell, M. (1998) Selected challenges in outflow generation research in forests from the hillslope to headwater drainage basin scale, Journal of the American Water Resources Association, 34, 765-785.

Bouten, W., Heimovaara, T.J. and Tiktak, A. (1992) Spatial patterns of throughfall and soil-water dynamics in a Douglas-fir stand, Water Resources Research, 28, 3227-3233.

Boyle, D.P., Gupta, H.V. and Sorooshian, S. (2000) Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, Water Resources Research, 36, 12, 3663-3674.

Brakensiek, D.L. and Rawls, W.J. (1994) Soil containing rock fragments: effects on infiltration, Catena, 23, 99-110.

Brocca, L., Morbidelli, R., Melone, F. and Moramarco, T. (2007a) Soil moisture spatial variability in experimental areas of central Italy, Journal of Hydrology, 333, 2-4, 356-373.

Brocca, L., Melone, F. and Moramarco, T. (2007) Storm runoff estimation based on the soil conservation service – curve number method with soil moisture data assimilation., Uncertainties in the ‘monitoring-conceptualisation-modelling’ sequence of catchment research, L. Pfilster, L. Hoffmann Eds., IHP-VI, Technical Documents in Hydrology, 81 (UNESCO), 141-148.

Brooks, R.H. and Corey, A.T. (1964) Hydraulic properties of porous media, Hydrology Paper No.3, Civil Engineering Department. Colorado State University, 24 pp.

Buttafuoco, G., Castrignanò, A., Busoni, E. and Dimase, A. C. (2005) Studying the spatial structure evolution of soil water content using multivariate geostatistics, Journal of Hydrology, 311, 202-218.

Calder, I.R. (1977) A model of transpiration and interception loss from a spruce forest in Plynlimon, central Wales, Journal of Hydrology, 33, 247-265.

Calver, A. and Cammeraat, L.H. (1993) Testing a physically-based runoff model against field observations on a Luxembourg hillslope, Catena, 20, 3, 273-288.

Cardell-Oliver, R., Smettem, K., Kranz, M. and Mayer, K. (2005) A reactive soil moisture sensor network: Design and field evaluation, International Journal of Distributed Sensor Networks, 12, 149-162.

Chirico, G.B., Grayson, R.B. and Western, A.W. (2003) A downward approach to identifying the structure and parameters of a process-based model for a small experimental catchment, Hydrological Processes, 17, 2239-2258.

Cohen, J. (1960) A coefficient of agreement for nominal scales, Educational and Psychological Measurement, 20, 37-46.

Costanza, R. (1989) Model goodness of fit: a multiple resolution procedure, Ecological Modelling, 47, 199-215.Ecological Modelling, 47 (1989) 199-215 199

Cornelis, W.M., Ronsyn, J., Meirvenne, M.V. and Hartmann R. (2001) Evaluation of Pedotransfer Functions for Predicting the Soil Moisture Retention Curve, Soil Science Society of America Journal, 65, 638-648.

Coronato, R. and Bertiller, M. (1996) Precipitation, landscape and grazing related effects on soil moisture in semiarid Patagonia, Journal of Arid environments, 34, 1-9.

Cosh, M.H., Jackson, T.J., Starks, P. and Heathman G. (2006) Temporal stability of surface soil moisture in the Little Washita River watershed and its applications in satellite soil moisture product validation, Journal of Hydrology, 323, 168-177.

Cosh, M.H., Jackson, T.J., Bindlish, R. and Prueger, J.H. (2004) Watershed scale temporal and spatial stability of soil moisture and its role in validating satellite estimates, Remote Sensing of Environment, 92, 4, 427-435.

Crow, W.T. and Wood, E.F. (1999) Multi-scale dynamics of soil moisture variability observed during SGP'97, Geophysical Research Letters, 26, 23, 3485-3488.

Davies, J.A. and Allen, C.D. (1973) Equlibrium, potential and actual evaporation from cropped surfaces in southern Ontario, Journal of Applied Meteorology, 12, 649-657.

Davis, R.E., Metcalfe, J.R., Hardy, J.P. and Goodison, B. (1998) Variations of snow accumulation in the southern boreal forest: Analyses of snow measurements 1993±1994 and 1994±1995, Proceedings of the Eastern Snow Conference 55, 153-164.

De Lannoy, G. J. M., Houser, P.R., Pauwels, V.R.N. and Verhoest, N.E.C. (2006a) Assessment of model uncertainty for soil moisture through ensemble verification, Journal of Geophysical Research, 111, D10101, doi:10.1029/2005JD006367.

De Lannoy, G. J. M., Verhoest, N.E.C., Houser, P.R. and Gish, T. (2006b) Spatial and temporal characteristics of soil moisture in an intensively monitored agricultural field (OPE), Journal of Hydrology, 331, 719–730.

Duan, Q., Sorooshian, S. and Gupta, H. V. (1992) Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models, Water Resources Research, 28, 1015-1031.

Dunn, S.M. (1999) Imposing constraints on parameter values of a conceptual hydrological model using baseflow response, Hydrology and Earth Science, 3, 271-284.

Dunne T., Moore, T.R. and Taylor, C.H. (1975) Recognition and prediction of runoff-producing zones in humid regions, Hydrological Science Bulletin, 20, 305-327.

Dunne T. (1978) Field studies of hillsope flow processes, In: Kirkby, M.J. (ed). Hillslope hydrology. Chichester: Wiley;. pp. 227–93.

Eltahir, E. A. B. and Bras, R.L. (1993) A Description of Rainfall Interception Over Large-areas, Journal of Climate, 6, 1002-1008.

Eschner, A.R. (1967) Interception and soil moisture distribution. In: Sopper, W.E., Lull, H.W. (eds), Forest Hydrology. Pergamon, Oxford, pp. 191–200.

Faeh, A.O., Scherrer, S. and Naef, F. (1997) A combined field and numerical approach to Investigate flow processes in natural macroporous soils under extreme precipitation, Hydrology and Earth System Sciences, 1, 4, 787-800.

Famiglietti, J. S., Rudnicki, J.W. and Rodell, M. (1998) Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas, Journal of Hydrology, 210, 259-281.

Famiglietti, J. S., Devereaux, J. A., Laymon, C. A., Segaye, T., Houser, P. R., Jackson, T. J., Graham, S.T., Rodell, M. and van Oevelen, P. (1999) Ground- based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains (1997) Hydrology Experiment, Water Resources Research, 35, 1839-1851,

Famiglietti, J. S., D. Ryu, A. A. Berg, M. Rodell, and T. J. Jackson (2008) Field observations of soil moisture variability across scales, Water Resources Research, 44, W01423, doi:10.1029/2006WR005804.

Federer, C.A. and Lash, D. (1978) BROOK: A hydrologic simulation model for eastern forests. University of New Hampshire Water Resources Research Center, Research Report 19, 94 pp., Durham, New Hampshire.

Federer, C.A. (1982) Transpirational supply and demand: Plant, soil, and atmospheric effects evaluated by simulation, Water Resources Research, 18, 355-362.

Finnerty, B. D., Smith, M. B., Seo, D.-J., Koren, V. and Moglen, G. E. (1997) Space–time scale sensitivity of the Sacramento model to radar-gage precipitation inputs, Journal of Hydrology, 203, 21-38.

Fisher, R.F. and Binkley, D. (2000) Ecology and Management of Forest Soils. New York: John Wiley & Sons. 489 pp.

Franks, S.W., Gineste, P., Beven, K.J., and Merot, P. (1998) On constraining the predictions of a distributed model: the incorporation of fuzzy estimates of saturated areas in the calibration process, Water Resources Research, 34, 787-797.

Freer, J., B. Ambroise, and K.J. and Beven, K.J. (1996) Bayesian Estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach, Water Resources Research, 32, 7, 2161-2173.

Freer, J.E., McMillan, H., McDonnell, J.J. and Beven, K.J. (2003) Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures, Journal of Hydrology, 291, 254-277.

Freeze, R.A. and Cherry, J.A. (1979) Groundwater. Prentice-Hall, Inc., Englewood Cliffs, NJ, 604 pp.

Garrote, L. and Bras, R.L. (1995) A distributed model for real-time flood forecasting using digital elevation models, Journal of Hydrology, 167, 279-306.

Goodrich, D.C., Faures, J.M., Woolhiser, D.A., Lane, L.J. and Sorooshian, S. (1995) Measurement and analysis of small-scale convective storm rainfall variability, Journal of Hydrology, 173, 283-308.

Goovaerts P. (1997) Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Grayson, R. B., Moore, I. D. and McMahon, T. A. (1992) Physically based hydrologic modelling. 2: Is the concept realistic? Water Resources Research, 28, 2659-2666.

Grayson, R.B., Western, A.W., Chiew, F.H.S. and Blöschl, G. (1997) Preferred states in spatial soil moisture patterns: Local and nonlocal controls, Water Resources Research, 33, 12, 2897-2908.

Grayson, R.B. and Western, A.W. (1998) Towards areal estimation of soil water content from point measurements: time and space stability of mean response, Journal of Hydrology, 207, 68-82.

Grayson, R.B. and Blöschl, G. (2000a) Spatial processes, organisation and patterns. Chapter 1 in: Spatial Patterns in Catchment Hydrology: Observations and Modelling. Cambridge University Press. pp. 3-16.

Grayson, R.B. and Blöschl, G. (2000b) Spatial modelling of catchment dynamics. Chapter 3 in Spatial Patterns in Catchment Hydrology: Observations and Modelling, Cambridge University Press. pp. 51-81.

Grayson, R.B. and Blöschl, G. (2000c) Summary of pattern comparison and concluding remarks. Chapter 14 in: Spatial Patterns in Catchment Hydrology: Observations and Modelling. Cambridge University Press. pp. 355-367.

Grayson, R.B., Blöschl, G., Western, A.W. and McMahon, T.A. (2002) Advances in the use of observed spatial patterns of catchment hydrological response, Advances in Water Resources, 25, 1313-1334.

Güntner, A., Seibert, J. and Uhlenbrook, S. (2004) Modeling spatial pat- terns of saturated areas: An evaluation of different terrain indices, Water Resources Research, 40, W05114, doi:10.1029/2003WR002864.

Gupta, H.V., Beven, K.J. and Wagener, T. (2005) Model calibration and uncertainty estimation. In: Anderson, M.G. et al. (eds) Encyclopedia of hydrological sciences. John Wiley & Sons Ltd., Chichester, UK, pp. 1-17.

Gupta H.V., Sorooshian S., Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data, Journal of Hydrology, 181, 23-48

Gupta, V.K., Sorooshian, S., and Yapo, P.O. (1998) Toward improved calibration of hydrologic models: multiple and non-commensurable measures of information, Water Resources Research 34, 751-763.

Hammel, K. and Kennel, M. (2001) Charakterisierung und Analyse der Wasserverfügbarkeit und des Wasserhaushalts von Waldstandorten in Bayern mit dem Simulationsmodell Brook90. Forstwirtschaftliche Forschungsberichte, München 185 pp.

Hawley, M.E., Jackson, T.J. and McCuen, R.H. (1983) Surface soil moisture variation on small agricultural watersheds, Journal of Hydrology, 62, 179-200.

Hornberger G.M., and Spear, R.C. (1981) An approach to the preliminary analysis of environmental systems, Journal of Environmental Management, 12, 7-18.

Hunter, N.M., Bates, P.D., Horritt, M.S., De Roo, P.J. and Werner, M.G.F. (2005) Utility of different data types for calibrating flood inundation models within a GLUE framework, Hydrology and Earth System Sciences, 9, 4, 412-430.

Hupet, F. and Vanclooster, M. (2002) Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped field, Journal of Hydrology, 261, 86-101.

Ivanov, V.Y., Vivoni, E.R., Bras, R.L. and Entekhabi, D. (2004) Catchment hydrologic response with a fully distributed triangulated irregular network model, Water Resources Research, 40, W11102, doi:10.1029/2004WR003218.

Ivanov, V. Y., Fatichi, S., Jenerette, G.D., Espeleta, J.F., Troch, P.A. and Huxman T.E. (2010) Hysteresis of soil moisture spatial heterogeneity and the “homogenizing” effect of vegetation, Water Resources Research, 46, W09521, doi:10.1029/2009WR008611.

Jacques, D., Šimçnek, J., Timmerman, A. and J. Feyen, J. (2001) Calibration of the Richards’ and convection–dispersion equations to field-scale water flow and solute transport under rainfall conditions, Journal of Hydrology, 259, 15-31.

Jakeman, A. J. and Hornberger, G. M. (1993) How much complexity is warranted in a rainfall-runoff model?, Water Resources Research, 29, 2637-2649.

James, A.L. and Roulet, N.T. (2007) Investigating hydrologic connectivity and its association with threshold change in runoff response in a temperate forested watershed, Hydrological Processes, 21, 3391-3408.

Jensen, W. and Calabresi, G. (1997) Flood monitoring workshop, ESRIN, Frascati, Italy 26 27 June 1995. Hydrological Processes 11(10), 1357.

Johnston, P.R. and Pilgrim, D.H. (1976) Parameter Optimization for Watershed Models, Water Resources Research, 12, 3, 477-486.

Julien, P.Y. and Saghafian, B. (1991) A two-dimensional watershed rainfall-runoff model-CASC2D user´s manual, Civil Engineering Report R90-91PYJ-BS-12, Colorado State University.

Kachanoski, R.G., and de Jong, E. (1988) Scale dependence and the temporal persistence of spatial patterns of soil water storage. Water Resources Research, 24,85-91.

Kamgar, A., Hopmans, J.W., Wallender, W.W. and Wendworth, O. (1993) Plot size and sample number for neutron probe measurements, Soil Science, 156, 213-224.

Hornberger, G.M., and Spear , R.C. (1981) An approach to the preliminary analysis of environmental systems, Journal of Environmental Management, 12, 7-18.

Keim, R.F., Skaugset, A.E. and Weiler, M. (2005) Temporal persistence of spatial patterns in throughfall, Journal of Hydrology, 314, 1-4, 263-274.

Kuczera, G. and Mroczkowski, M. (1998) Assessment of hydrologic parameter uncertainty and the worth of multi-response data, Water Resources Research, 34, 1481-1490.

Lamb, R., Beven, K. and Myrabø, S. (1998) Use of spatially distributed water table observations to constrain uncertainty in a rainfall–runoff model, Advances in Water Resources, 22, 4, 305–317.

Legates, D.R.and McCabe, G.J. (1999) Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation, Water Resources Resarch, 35, 233-241.

Lin, H. (2006) Temporal Stability of Soil Moisture Spatial Pattern and Subsurface Preferential Flow Pathways in the Shale Hills Catchment, Vadose Zone Journal, 5, 317-340.

Madsen, H. (2003) Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, Advances in Water Resources, 26, 205-216.

McGlynn, B.L., McDonnell, J.J., Seibert, J. and Kendall, C. (2004) Scale effects on headwater catchment runoff timing, flow sources, and groundwater – streamflow relations, Water Resources Research, 40, W07504, doi:10.1029/2003WR002494.

McGuire, K., Weiler, M. and McDonnell, J.J. (2006) Integrating field experiments with modeling to infer water residence times, Advances in Water Resources. doi:10.1016/j.advwatres 2006.07.004

McGuire, K.J., Weiler, M. and McDonnel, J.J. (2007) Integrating tracer experiments with modeling to assess runoff processes and water transit times, Advances in Water Resources, 30, 4, 824-837.

McMichael, A. J., Woodruff, R. E. and Hales, S. (2006) Climate change and human health: present and future risks, Lancet, 367, 859-869.

Mermoud, A. and Xu, D. (2006) Comparative Analysis of Three Methods to Generate Soil Hydraulic Functions, Oil and Tillage Research, 87, 89-100.

Merz, B. and Plate, E. (1997) An Analysis of the Effects of Spatial Variability of Soil and Soil Moisture on Runoff. Water Resources Research, 33, 12, 2909-2922.

Mohanty, B. P. and Skaggs, T.H. (2001) Spatiotemporal evolution and timestable characteristics of soil moisture within remote sensing footprints with varying soil, slope, and vegetation, Advances in Water Resources, 24, 1051–1067.

Montanari, M., Hostache, R., Matgen, P., Schumann, G., Pfister, L. and Hoffmann, L. (2009) Calibration and sequential updating of a coupled hydrologic-hydraulic model using remote sensing-derived water stages, Hydrology and Earth System Sciences, 13, 367-380.

Monteith, J.L. (1965) Evaporation and environment. In: Fogg, G.E. (ed.) Symposium of the Society for Experimental Biology, The State and Movement of Water in Living Organisms, 19, 205-234. Academic Press, Inc., NY.

Moore, I.D., O'Loughlin, E.M. and Burch, G.J. (1988) A contour-based topographic model for hydrological and ecological applications, Earth Surface Processes and Landforms, 13, 305-320.

Moore, I.D., Gallant, J.C., Guerra, L. and Kalma, J.D. (1993) Modelling the spatial variability of hydrological processes using Geographical Information Systems, in: Kovar, K. and Nachtnebel, H.P. (eds). HydroGIS 93: Application of Geographical Information Systems in Hydrology and Water Resources, IAHS Publication. No. 211, 161-169.
Moré, J.J. (1977) the Levenberg-Marquardt Algorithm: Implementation and theory. In: Watson, G.A. (ed). Lecture Notes in Mathematics 630, Springer Verlag, pp. 105-116. Numerical Analysis.

Mroczkowski, M., Raper, G.P. and Kuczera, G. (1997) The quest for more powerful validation of conceptual catchment models, Water Resources Research, 33, 2325-2335.

Nash J.E. and Sutcliffe J.V. (1970) River forecasting through conceptual models. Part 1-A. Discussion of Principles, Journal of Hydrology, 10, 282-290.

Niemann, K.O. and Edgell, M.C.R. (1993) Preliminary analysis of spatial and temporal distribution of soil moisture on a deforested slope, Physical Geography, 14, 449-464.

Oldak, A., Pachepsky, Y., Jackson, T.J. and Rawls, W.J. (2002) Statistical properties of soil moisture images revisited, Journal of Hydrology, 255, 12-24.

Owe, M., Jones, E. and Schmugge, T. (1982), Soil moisture variation patterns observed in Hand County, South Dakota, Water Resources Bulletin, 18, 949–954.

Owe, M., de Jeu, R. and Holmes, T. (2008) Multisensor historical climatology of satellite-derived global land surface moisture, Journal of Geophyical Research, 113, doi:10.1029/2007JF000769

Parajka J., Merz, R. and Blöschl, G. (2005) Regionale Wasserbilanzkomponenten für

Österreich auf Tagesbasis (Regional water balance components in Austria on a daily basis), Österreichische Wasser- und Abfallwirtschaft, 57 (3/4), 43-56.

Parajka, J., Merz R. and Blöschl, G. (2007) Uncertainty and multiple objective calibration in regional water balance modeling - Case study in 320 Austrian catchments, Hydrological Processes, 21, 435-446.

Pareto, V. 1906 Manuale di Economia Politica, Edizione Critica, Aldo. Montesano, Alberto Zanni and Luigino Bruni (eds), (Milan: EGEA—Universita`Bocconi Editore).

Pariente, S. (2002) Spatial Patterns of Soil Moisture as Affected by Shrubs, in Different Climatic Conditions, Environmental and Assessment, 73, 237-251.

Penman, H.L. (1948) Natural evaporation from open water, bare soil and grass, Proceedings of the Royal Meteorological Society (seriesA), 193, 120-146.

Peschke, G., Etzenberg, C., Müller, G., Töpfer, J. and Zimmermann, S. (1999) Das wissensbasierte System FLAB – ein Instrument zur rechnergestützten Bestimmung von Landschaftseinheiten mit gleicher Abflussbildung. – IHI-Schriften 10, Zittau, 122 pp.

Peters-Lidard, C.D., Pan, F. and Wood, E.F. (2001) A reexamination of modeled and measured soil moisture spatial variability and its implications for land surface modeling, Advances in Water Resources, 24, 1069-1083.

Plaut-Berger, K. and Entekhabi D. (2001) Basin hydrologic response relations to distributed physiographic descriptors and climate, Journal of Hydrology, 247, 169-182.

Pontius, R.G. (2000) Quantification error versus locational error in comparison of categorial maps, Photogrammetric Engineering and Remote Sensing, 66,1011-1016.

Raat, K.J., Draaijers, G.P.J., Schaap, M.G., Tietema, A. and Verstraten, J.M. (2002) Spatial variability of throughfall water and chemistry and forest floor water content in a Douglas fir forest stand, Hydrology and Earth System Sciences, 6,363–374.

Rawls, W.J. and Brakensiek, D.L. (1985) Prediction of soil water properties for hydrologic modelling. In: Jones E. and Ward T.J. (eds) Watershed Management in the Eighties. Proc. Symp. Irrig. Drain. Div. ASCE. Denver, CO. ASCE, New York.

Rawls, W.J. and Brakensieks, D.L. (1985) Estimation of soil water retention and hydraulic properties. In: Unsaturated flow in hydrological modeling. in: Morel- Seytoux H.J. (ed.) Kluwer, Dordrecht.

Refsgaard, J.C. (1997) Validation and intercomparison of different updating procedures for real-time forecasting, Nordic Hydrology, 28, 65-84.

Rodriguez-Iturbe, I., Vogel, C. K., Rigon, R., Entekhabi, D., Castelli, F. and Rinaldo, A. (1995) On the spatial organization of soil moisture fields, Geophysical Research Letters, 22, 2757–2760.

Romano, N. and Santini, A. (1997) Effectiveness of using pedo-transfer functions to quantify the spatial variability of soil water retention characteristics, Journal of Hydrology, 202,137-157.

Ross, T.J. (1995) Fuzzy Logic, with engineering applications. John Wiley and Sons, Chichester

Rutter, A.J., Morton, A.J. and Robins, P.C. (1975) A predictive model of inter-ception loss in forests: II. Generalization of the model and comparison with observations in some coniferous and hardwood stands, Journal of Applied Ecology, 12, 367–380.

Salvucci, G. (1998) Limiting relations between soil moisture and soil texture with implications for measured, modeled and remotely sensed estimates, Geophysical Research Letters, 25, 10, 1757– 1760.

Satalino, G., Mattia, F., Davidson, M.W.J., Thuy Le Toan, Pasquariello, G. and Borgeaud, M. (2002) On current limits of soil moisture retrieval from ERS-SAR data, Geoscience and Remote Sensing, 40, 2438-2447.

Schulla J. (1997) Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaänderungen. Züricher Geographische Schriften, Heft 69, Geographisches Institut ETH, Zürich.

Schume, H., Jost, G. and Katzensteiner, K. (2003) Spatiotemporal analysis of the soil water content in a mixed Norway spruce (Picea abies (L.) Karst.)-European beech (Fagus sylvatica L.) stand. Geoderma, 112, 273–287.

Sciuto, G. and Diekkrüger, B. (2010) Influence of soil heterogeneity and spatial discretization on catchment water balance modeling. Vadose Zone J. 9:955–969

Seibert, J. (2000) Multi-criteria calibration of a conceptual runoff model using a genetic algorithm, Hydrology and Earth System Sciences, 4, 215 - 224.

Seibert, J. (2001) On the need for benchmarks in hydrological modelling, Hydrological Processes, 15, 1063-1064.

Seibert, J. (2002) HBV light 2, User´s Manual. Uppsala University, Institute of Earth Sciences, Department of Hydrology, Uppsala, Sweden, 16 pp.

Seibert, J., K. Bishop, and Nyberg, L. (1997), A test of TOPMODEL's ability to predict spatially distributed groundwater levels, Hydrological Processes, 11, 1131–1144.

Seibert, J., and J. J. McDonnell, (2002) On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resources Research, 38, 11, 1241, doi:1210.1029/2001WR000978

Seiler, J. and Matzner, E. (1995) Spatial variability of throughfall chemistry and selected soil properties as influenced by stem distance in a mature Norway spruce (Picea-abies, Karst) stand, Plant Soil, 176, 139–147.

Si, B.C. (2002) Spatial and statistical similarities of local soil water fluxes, Soil Science Society of America Journal 66, 753–759.

Singh, V. P. (1995) Computer models of watershed hydrology. Water Resources Publications (Highlands Ranch, Colo.),1995 pp.

Sloan, P. G. and Moore, I.D. (1984) Modeling subsurface stormflow on steeply sloping forested watersheds, Water Resources Research, 20, 1815-1822.

Soulsby, C., Neal, C., Laudon, H., Burns, D.A., Merot, P., Bonell, M., Dunn, S.M. and Tetzlaff, D. (2008) Catchment data for process conceptualization: simply not enough?, Hydrological Processes, doi:10.1002/hyp.7068.

Spear, R. and Hornberger, G. (1980) Eutrophication in peel inlet, ii, identification of critical uncertainties via general sensitivity analysis, Water Research, 14, 43- 49.

Spittlehouse, D.L. and Black, T.A. (1981) A growing season water balance model applied to two Douglas fir stands. Water Resources Research 17, 1651-1656

Starks, P. J., Heathman, G. C., Jackson, T. J., and Cosh, M. H. (2006) Temporal stability of profile soil moisture, Journal of Hydrology, 324, 400-411.

Stoll, S. and Weiler, M. (2010) Explicit simulations of stream networks to guide hydrological modelling in ungauged basins, Hydrology and Earth System Sciences,14, 8,1435-1448.

Tarboton, D.G. (1997) A new method for determination of flow directions and upslope areas in grid digital elevation models, Water Resources Research, 33, 309-319.

Teuling, A.J. and Troch, P.A. (2005) Improved understanding of soil moisture variability dynamics, Geophysical Research Letters, 32, L05404, DOI:10.1029/2004GL021935.

Teuling, A. J., Seneviratne, S.I., Williams, C. and Troch, P.A. (2006) Observed timescales of evapotranspiration response to soil moisture, Geophysical Research Letters, 33, L23403, doi:10.1029/2006GL028178.

Thiemann, M., Trosset, M., Gupta, H.V. and Sorooshian, S. (2001) Bayesian recursive parameter estimation for hydrologic models, Water Resources Research, 37, 10, 2521-2535.

Thierfelder, T.K., Grayson, R.B., von Rosena, D. and A.W. Western, A.W. (2003) Inferring the location of catchment characteristic soil moisture monitoring sites: Covariance structures in the temporal domain, Journal of Hydrology, 280,13-32.

Thompson, M.D., Preat, N.A. and Sommerfeld, A.Y. (1981) LANDSAT for determination and mapping of saline soils in dryland areas in southern Alberta. In: Proceedings of 4th Canadian Symposium on Remote Sensing, pp. 1213–1215. Winnipeg.

Tietje, O. and Tapkenhinrichs, M. (1993) Evaluation of pedotransfer functions, Soil Science Society Am. J., 57, 1088–1095.

Tromp-van Meerveld and McDonnell, J.J. (2006) Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis, Water Resources Research, 42, W02411, doi:10.1029/2004WR003800.

Tromp-van Meerveld, H. J. and J. J. McDonnell (2006a) Threshold relations in subsurface stormflow: 1. A 147-storm analysis of the Panola hillslope, Water Resources Research, 42, W02415, doi:10.1029/2004WR003778.

Tromp-van Meerveld, H. J., and McDonnell, J.J. (2006b) On the interactions between the spatial patterns of topography, soil moisture, transpiration and species distribution at the hillslope scale, Advances in Water Resources, 29, 293 – 310.

Tromp-van Meerveld, H.J. and M. Weiler (2007) How does soil depth variability influence subsurface flow variability? Advances in Water Resources, 30.

Tromp-van Meerveld, H.J. and Weiler, M. (2008), Hillslope dynamics modeled with increasing complexity, Journal of Hydrology 361, 24-40.

Trubilowicz, J., Cai, K. and Weiler, M. (2009) Viability of motes for hydrological measurement, Water Resources Research, 45:W00D22, doi:10.1029/2008WR007046.

Uhlenbrook, S., Seibert, J., Leibundgut, C. and Rodhe, A. (1999) Prediction uncertainty of conceptual rainfall-runoff models caused by problems to identify model parameters and structure, Hydrologycal Science Journal, 44, 5, 779-797.

Van Straten, G. and Keesman, K.J. (1991) Uncertainty propagation and speculation in projective forecasts of environmental change: a lake- eutrophication example. Journal of Forecasting, 10, 163-190.

Vereecken H., Maes J., Feyen J. and Darius P. (1989) Estimating the soil moisture retention characteristic from texture, bulk density, and carbon content, Soil Science, 148, 389-403.

Vereecken, H., Maes, J. and Feyen, J. (1990) Estimating unsaturated hydraulic conductivity from easily measured soil properties, Soil Science, 149, 1, 1-12.

Vereecken, H., Kasteel, R., Vanderborght, J. and Harter, T. (2007a) Upscaling Hydraulic Properties and Soil Water Flow Processes in Heterogeneous Soils: A Review, Vadose Zone Journal 6, 1-28.

Vereecken, H., Kamai, T., Harter, R., Kasteel, J.W., Hopmans, R. and Vanderborght, J. (2007b) Explaining soil moisture variability as a function of mean soil moisture: A stochastic unsaturated flow perspective, Geophysical Research Letters, 34, Art. No. L22402.

Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W. and Sorooshian, S. (2002) Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resources Research, 39, 1214, doi:10.1029/2002WR001746, 2003.

Vrugt, J.A., Gupta, H.V., Bouten, W. and Sorooshian, S. (2003a) A shuffled complex evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, Water Resources Research, 39(8), doi: 10.1029/2002WR001642.

Vrugt, J.A., Bouten, W., Gupta, H.V. and Hopmans, J.W. (2003b) Toward improved identifiability of soil hydraulic parameters: on the selection of a suitable parametric model, Vadose Zone Journal, 2, 98–113, doi:10.2136/vzj2003.0098. 2035, 2037

Vrugt, J. A., Stauffer, P. H., Wöhling, T., Robinson, B.A., and Vesselinov, V.V. (2008a) Inverse modeling of subsurface flow and transport properties: a review with new developments, Vadose Zone J., 7, 843–864, doi:10.2136/vzj2007.0078. 2021

Vrugt, J.A., Braak, C. J. F., Clark, M. P., Hyman, J. M. and Robinson, B. A. (2008b) Treatment of input uncertainty in hydrologic modeling: doing hydrology backward with Markov chain Monte Carlo simulation, Water Resources Research, 44, W00B09, doi:10.1029/2007WR0067. 2023, 2032.

Wagener, T. (1998) Developing a knowledgebased system to support rainfall-runoff modelling, MSc Thesis, Department of Civil Engineering, Delft University of Technology, Netherlands.

Wagener, T., Lees, M.J., Wheater, H.S. (1999) A generic rainfall-runoff modelling toolbox, EOS, Transactions of the American Geophysical Union, 80, F203.

Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V. and Sorooshian, S. (2001a) A framework for development and application of hydrological models, Hydrology and Earth System Sciences, 5, 1, 13-26.

Wagener T., Camacho, L.A., Lees, M.J. and Wheater, H.S. (2001b) Dynamic parameter identifiability of a solute transport model. In Beheshti R (ed.). Advances in Design Sciences and Technology-Proceedings of EuropIA'8, Delft, April; 251-264.

Wagner, B., Tarnawski, V.R., Hennings, V., Müller, U., Wessolek, G. and Plagge, R. (2002) Reply to comments on ''Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using an independent data set'', Geoderma, 108, 301- 304.

Wagener, T., McIntyre, N., Lees, M.J., Wheater, H.S. and Gupta, H.V. (2003) Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, Hydrological Processes, 17, 455-476.

Wagener, T. and Gupta, H.V. (2005) Model identification for hydrological forecasting under uncertainty, Stochastic Environmental Research and Risk Assessment, 19, 378-387.

Wagener, T. and Kollat, J. (2007a) Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox, Environmental Modelling & Software, 22, 1021-1033.

Wagener, T., Gupta, H., Yatheendradas, S., Goodrich, D., Unkirch, C. and Schaffner, M. (2007b) Understanding sources of uncertainty in flashflood forecasting for semi-arid regions, IAHS, 313, 204-212.

Whelan, K.R.T., Smith,T.J., Cahoon, D.R., Lynch, J.C. and Anderson, G.H. (1998) Groundwater control of mangrove surface elevation: Shrink and swell varies with soil, Estuaries and Coasts, 28, 833-843.

Wealands, S. (2006) Hydrological spatial field comparison, PhD thesis, University of Melbourne, Melbourne, Australia, 232 pp.

Weiler, M. (2001) Mechanisms controlling macropore flow during infiltration-dye tracer experiments and simulations, ETHZ, Zürich, Switzerland, 151 pp.

Weiler, M., Naef, F. and Leibundgut, C. (1998) Study of runoff generation on hillslopes using tracer experiments and a physically-based numerical hillslope model, IAHS Publication no. 248, 353-360.

Weiler, M., Uchida T. and McDonnell J.J. (2003) Connectivity due to preferential flow controls water flow and solute transport at the hillslope scale. In: Proceedings of MODSIM 2003, Townsville, Australia.

Weiler, M. and McDonnell, J.J. (2004a) Water storage and soil movement. In: Burley, J., Evans, J. and Youngquist , J. (eds), Encyclopedia of Forest Sciences. Elsevier Science Publishers, pp. 1253-1260.

Weiler, M. and McDonnell, J.J. (2004b) Soils and Site: Water Storage and Movement Encyclopedia of Forest Sciences, Academic Press.

Weiler, M. and McDonnell, J.J. (2006) Testing nutrient flushing hypotheses at the hillslope scale: A virtual experiment approach, Journal of Hydrology, 319, 339-356.

Weiler, M. and McDonnell, J.J. (2007) Conceptualizing lateral preferential flow and flow networks and simulating the effects on gauged and ungauged hillslopes, Water Resources Research, 43, W03403, doi:10.1029/2006WR004867.

Wendroth, O., Pohl, W., Koszinski, S., Rogasik, H., Ritsema, C.J. and Nelson, D.R. (1999) Spatiotemporal patterns and covariance structures of soil water status in two Northeast-German field sites, Journal of Hydrology, 215, 38-58.

Western, A.W. and Grayson, R.B. (1998) The Tarrawarra data set: soil moisture patterns, soil characteristics and hydrological flux measurements, Water Resources Research, 34, 2765- 68.
Western, A.W. and Blöschl, G. (1999a) On the spatial scaling of soil moisture, Journal of Hydrology, 217, 203-224.

Western, A.W., Grayson, R.B., Blöschl, G., Willgoose, G.R. and McMahon, T.A. (1999b) Observed spatial organization of soil moisture and its relation to terrain indices, Water Resources Research, 35, 3, 797-810.

Western, A.W., Blöschl, G. and Grayson, R.B. (2001) Toward capturing hydrologically significant connectivity in spatial patterns, Water Resources Research, 37, 1, 83-97.

Western, A.W., Zhou, S.-L., Grayson, R.B., McMahon, T.A., Blöschl, G. and Wilson, D.J. (2004) Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes, Journal of Hydrology, 286, 113-134.

Wheater, H.S., Bishop, K.H. and Beck, M.B. (1986) The identification of conceptual hydrological models for surface water acidification, Hydrological Processes, 1, 89-109.

Wheater, H.S., Jakeman, A.J. and Beven, K.J. (1993) Progress and directions in rainfall-runoff modeling. In Modeling Change in Environmental Systems, Jakeman, A.J., Beck, M.B. and McAleer, M.J. (eds), John Wiley & Sons: UK; 101-132.

Whelan, M.J., Anderson, J.M. (1996) Modelling spatial patterns of throughfall and interception loss in a Norway spruce (Picea abies) plantation at the plot scale, Journal of Hydrology, 186, 335-354.

Wigmosta, M.S., Lance, W.V. and Lettenmaier, D.P. (1994) A distributed hydrology-vegetation model for complex terrain, Water Resources Research, 30, 6, 1665-1679.

Wigmosta, M.S. and Lettenmaier, D.P. (1999) A comparison of simplified methods for routing topographically driven subsurface flow, Water Resources Research, 35, 255-64.

Wilson, J. P. and Gallant, J.C. (2000) Secondary Topographic Attributes. Terrain Analysis: Principles and Applications. J. P. Wilson, Gallant, J. C. New York, John Wiley and Sons: 87-131.

Wösten J.H.M., Lilly, A., Nemes, A., Le Bas, C. (1998): Using existing soil data to derive hydraulic param- eters for simulation models in environmental stud- ies and in land use planning, Final Rep. Eur. Un., Wageningen.Yapo P.O.,

Wösten, J.H.M., Pachepsky, Ya.A. and Rawls, W.J. (2001) Pedotransfer functions: bridging the gap between available basic soil data and missing hydraulic characterisitics, Journal of Hydrology, 251, 123-150.

Ye, Z.H., Baker, A.J.M., Wong, M.H. and Willis, A.J. (1997) Zinc, lead and cadmium tolerance, uptake and accumulation in populations of Phragmites australis (Cav.) Trin. ex Steudel, Annals of Botany, 80, 363–370.

Young, P.C. (1983) The validity and credibility of models for badly-defined systems. In: Beck M.B. and van Straten, G. (eds) Uncertainty and Forecasting of Water Quality, Springer-Verlag, Berlin, 69-98 pp.

Young, P.C., Parkinson S. and Lees, M.J. (1996) Simplicity out of complexity in environmental modelling: Occham's razor revisited, Journal of Applied Statistics, 23, 165-210.

Zhou, J.Z., Xia, B.C., Treves, D.S., Wu, L.Y., Marsh, T.L., O'Neill, R.V., Palumbo, A.V. and Tiedje, J.M. (2002) Spatial and resource factors influencing high microbial diversity in soil, Applied Environmental Microbiology, 68, 326-334.

Zitzler, E. and Thiele, L. (1999) Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach, IEEE Transactions on Evolutionary Computation, 3, 4, 257-271.