Hydrol. Earth Syst. Sci. Discuss., 10, 14801–14855, 2013
www.hydrol-earth-syst-sci-discuss.net/10/14801/2013/
doi:10.5194/hessd-10-14801-2013

Hydrology and
Earth System
Sciences
Discussions
Open Access

Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration

S. Gharari[1,2], **M. Hrachowitz**[1], **F. Fenicia**[1,2], **H. Gao**[1], **and H. H. G. Savenije**[1]

[1]Delft University of Technology, Faculty of Civil Engineering and Geosciences,
Water Resources Section, Delft, the Netherlands
[2]Public Research Center–Gabriel Lippmann, Belvaux, Luxembourg

Received: 10 November 2013 – Accepted: 18 November 2013 – Published: 5 December 2013

Correspondence to: S. Gharari (s.gharari@tudelft.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Conceptual environmental systems models, such as rainfall runoff models, generally rely on calibration for parameter identification. Increasing complexity of this type of model for better representation of hydrological process heterogeneity typically makes parameter identification more difficult. Although various, potentially valuable, strategies for better parameter identification were developed in the past, strategies to impose general conceptual understanding regarding how a catchment works into the process of parameterizing a conceptual model has still not been fully explored. In this study we assess the effect of imposing semi-quantitative, relational expert knowledge into the model development and parameter selection, efficiently exploiting the complexity of a semi-distributed model formulation. Making use of a topography driven rainfall-runoff modeling (FLEX-TOPO) approach, a catchment was delineated into three functional units, i.e. wetland, hillslope and plateau. Ranging from simplicity to complexity, three model set-ups, FLEX^A, FLEX^B and FLEX^C have been developed based on these functional units. While FLEX^A is a lumped representation of the study catchment, the semi-distributed formulations FLEX^B and FLEX^C introduce increasingly more complexity by distinguishing 2 and 3 functional units, respectively. In spite of increased complexity, FLEX^B and FLEX^C allow modelers to compare parameters as well as states and fluxes of their different functional units to each other. Based on these comparisons, expert knowledge based, semi-quantitative relational constraints have been imposed on three models structures. More complexity of models allows more imposed constraints. It was shown that a constrained but uncalibrated semi-distributed model, FLEX^C, can predict runoff with similar performance than a calibrated lumped model, FLEX^A. In addition, when constrained and calibrated, the semi-distributed model FLEX^C exhibits not only higher performance but also reduced uncertainty for prediction, compared to the calibrated, lumped FLEX^A model.

1 Introduction

Lumped conceptual and distributed physically based models are the two endpoints of the modeling spectrum in many environmental systems models, ranging from simplicity to complexity. These two approaches are characterized by their very own advantages and limitations. In hydrology, physically based models are typically applied under the assumptions that (a) the spatial resolution and the complexity of the model is warranted by the available data, and (b) the catchment response is a mere aggregation of small scale processes. However, these two fundamental assumptions are violated in many cases. As a result, not only the predictive power but also the hydrological insights that these models provide is limited (e.g. Beven, 1989, 2001; Grayson et al., 1992; Blöschl, 2001; Pomeroy et al., 2007; Sivapalan, 2006; McDonnell et al., 2007; Hrachowitz et al., 2013a).

In contrast, lumped conceptual models require less data for model parameterization. This advantage comes at the expense of considerable limitations. Representing system integrated processes, model structures and parameters are not directly linked to observable quantities. Their estimation therefore strongly relies on calibration. To limit parameter identifiability issues arising from calibration, these models are often oversimplified abstractions of the system. If inadequately tested they may act as mere mathematical marionettes (Kirchner, 2006), frequently resulting in models with good calibration performance, frequently outperforming more complex distributed models (e.g. Refsgaard and Knudsen, 1996; Ajami et al., 2004; Reed et al., 2004), but failing to provide realistic representations of the underlying processes, leading to limited predictive power (e.g. Freer et al. , 2003; Seibert, 2003; Kirchner, 2006; Beven, 2006; Kling and Gupta, 2009; Andréassian et al., 2012; Euser et al., 2013; Gharari et al., 2013a).

Various strategies have been suggested in the past to allow for increased model complexity and to thereby improve the physical realism of conceptual models. These strategies included on the one hand multi-criteria calibration, incorporating multiple response variables, such as ground- and soil water dynamics (e.g. Seibert et al., 2003; Freer et al., 2004; Fenicia et al., 2008a; Matgen et al., 2012; Sutanudjaja et al., 2013), remotely sensed evaporation (e.g. Winsemius et al., 2008), snow dynamics (e.g. Parajka and Blöschl, 2008) or tracer data (e.g. Vaché and McDonnell, 2006; Dunn et al., 2008; Son and Sivapalan, 2007; Birkel et al., 2011; Hrachowitz et al., 2013a). On the other hand, a complementary approach has been to simultaneously reproduce a set of hydrological signatures of one response variable, i.e. multi-objective calibration (e.g. Gupta et al., 1998, 2008; Boyle et al., 2000, 2001; Khu et al., 2008; Madsen, 2000; Fenicia et al., 2006; Rouhani et al., 2007; Bulygina and Gupta, 2010; Winsemius et al., 2009; McMillan et al., 2011; Clark et al., 2011; Euser et al., 2013; Hrachowitz et al., 2013a).

Traditionally, parameter estimation of conceptual models relied on the availability of calibration data, which, however, are frequently not available for the time period or the resolution of interest. A wide range of regionalization techniques for model parameters and hydrological signatures were thus developed to avoid calibration in such data scarce environments (e.g. Bárdossy, 2007; Yadav et al., 2007; Perrin et al., 2008; Zhang et al., 2008; Kling and Gupta, 2009; Samaniego et al., 2010; Kumar et al., 2010; Wagener and Montanari, 2011; Kapangaziwiri et al., 2012; Viglione et al., 2013). However, it was for a long time considered to be challenging to identify suitable functional relationships between catchment characteristics and model parameters (e.g. Merz and Blöschl, 2004; Kling and Gupta, 2009). Only recently, Kumar et al. (2010, 2013) showed that making use of multi-scale parameter regionalization (MPR) can yield model parameterizations which perform consistently over different scales catchments. In a further study they successfully transferred parameterizations obtained by the MPR technique to ungauged catchments in Germany and the USA (Samaniego et al., 2013). Without any further calibration the transferred parameterizations were capable to adequately reproduce runoff as well as other hydrological responses of the catchments.

Related to the above discussed difficulties with parameterization, the frequent lack of sufficient processes heterogeneity, i.e. complexity, in conceptual models introduces further limitations on the degree of realism in these models The concept of hydrological response unit (HRUs) can be exploited as a strategy for an efficient tradeoff between model simplicity, required for adequate parameter identifiability, and a realistic representation of hydrological processes. HRUs are units within a catchment, characterized by a different hydrological function. Individual HRUs can be represented by different model structures to account for hydrologically heterogeneous behavior based on data availability and desired resolution of process representation. This helps to enhance model realism while keeping the necessary complexity and related identifiability issues comparatively low. In most cases HRUs are defined based on soil type, land cover and similar physical catchment characteristics (e.g. Knudsen et al., 1986; Flügel, 1995; Grayson and Blöschl, 2000; Winter, 2001; Scherrer and Naef, 2003; Uhlenbrook et al., 2004; Wolock et al., 2004; Pomeroy et al., 2007; Scherrer et al., 2007; Schmocker-Fackel et al., 2007; Efstratiadis et al., 2008; Lindström et al., 2010; Nalbantis et al., 2011; Krcho, 2001; Kumar et al., 2010).

A wide range of studies also points towards the potential value of using topographical indices, which are readily available from digital elevation models (DEM) to account for process heterogeneity (e.g. McGlynn and McDonnell, 2003; Seibert et al., 2003; McGuire et al., 2005; Hrachowitz et al., 2009; Jencso et al., 2009; Detty and McGuire, 2010; Gascuel-Odoux et al., 2010). As standard metrics of landscape organization, such as absolute elevation, slope or curvature as used in the catena concept (Milne, 1935; Park and van de Giesen, 2004), are often not strong enough descriptors to infer hydrological function, alternative concepts were sought. The development of derived metrics such as the Topographic Wetness Index (Beven and Kirkby, 1979) facilitated an important step forward as it is at the core of TOPMODEL (Beven and Kirkby, 1979; Beven and Freer, 2001b), which proved to be a valuable approach in specific environmental settings meeting the assumptions of the model. A different descriptor allowing a potentially more generally applicable and hydrologically meaningful landscape classification has recently been suggested by Rennó et al. (2008): the Height Above the Nearest Drainage (HAND). Nobre et al. (2011) showed the hydrological relevance of

HAND by investigating at the long term groundwater behavior and land use. In a further study, this metric facilitated the identification of hydrologically similar landscape units, such as wetlands, hillslopes and plateaus in a Luxembourgish catchment (Gharari et al., 2011).

Explicitly invoking the co-evolution of topography, vegetation and hydrology, Savenije (2010) argued that catchments, as self-organizing systems, need to fulfill the contrasting hydrological functions of efficient drainage and sufficient water storage in order to allow, in a feedback process, topography and vegetation to develop the way they did. These distinct hydrological functions could then be associated with different landscape elements or HRUs as defined by HAND and slope, so that each HRU is represented by a model structure best representing its function in the ecosystem (cf. Savenije, 2010).

While HAND-based landscape classification can potentially show a way forward, it does not solve the problem arising when moving from lumped to HRU-guided semi-distributed model formulations: multiple parallel model structures result in an increased number of parameters, which, when not adequately constrained may cause equifinality and thereby increase predictive uncertainty (e.g. Gupta and Sorooshian, 1983; Beven, 2006; Gupta et al., 2008). In order to better satisfy the contrasting priorities of model complexity and predictive power, new strategies are sought to more efficiently utilize the modelers' understanding of the system and the frequently scarce available data for constraining the feasible model- and parameter space (e.g. Gupta et al., 2008; Wagener and Montanari, 2011; Andréassian et al., 2012; Gharari et al., 2013a; Hrachowitz et al., 2013b). In contrast to earlier attempts to constrain models using multiple evaluation criteria or a priori information on catchment properties such as land use or soil type as discussed above, the utility of a different and so far underexploited type of constraints, based on a priori understanding of the system, has been tested in this study. The concept of topography-driven conceptual modeling introduced by Savenije (2010) involves the identification of HRUs that operate in parallel. This opens the possibility to impose semi-quantitative, knowledge based, relational constraints of catchment behavior on model parameters, similar to what was suggested by Pokhrel et al. (2008) and Yilmaz et al. (2008). To restrict the resulting posterior parameter distributions, hydrologically meaningful relations between parallel HRUs are introduced. These relative relationship functions are based on expert knowledge as constraints to ensure that similar processes between parallel model structures are represented in an internally consistent way, thereby reducing the parameters' potential for compensating for errors. The advantage of this method is that there is only limited need to precisely quantify the constraints or the prior parameter distributions as the constraints are essentially relational (e.g. Koren et al., 2000, 2003; Kuzmin et al., 2008; Duan et al., 2006). This could allow for a meaningful and potentially more realistic representation of the system in which each model component is, within certain limits, forced to do what it is designed to do, rather than allowing it to compensate for data and model structural errors.

The objectives of this paper are thus to test the hypothesis if the use of semi-distributed, conceptual models, representing HRUs defined by hydrologically meaningful, topography-based landscape classification combined with model constraints can (1) increase model internal consistency and thus the level of process realism as compared to lumped model set-ups, (2) increase the predictive power of models compared to lumped model set-ups and (3) reduce the need for model calibration by the use of expert knowledge based on relations between parameters, fluxes and states.

2 Study area and data

The outlined methodology will be illustrated with a case study using data of the Wark catchment in the Grand Duchy of Luxembourg. The catchment has an area of 82 km^2 with the catchment outlet located downstream of the town of Ettelbrück at the confluence with the Alzette River (49.85° N, 6.10° E, Fig. 1). With an annual mean precipitation of 850 $\mathrm{mm\,yr^{-1}}$ and an annual mean potential evaporation of 650 $\mathrm{mm\,yr^{-1}}$ the annual mean runoff is approximately 250 $\mathrm{mm\,yr^{-1}}$. The geology in the northern part is dominated by schist while the southern part of the catchment is mostly underlain by sandstone and conglomerate. Hillslopes are generally characterized by forest, while plateaus and valley bottoms are mostly used as crop land and pastures, respectively. Drogue et al. (2002) quantified land use in the catchment as 4.3 % urban areas, 52.7 % agricultural land and 42.9 % forest. In addition they reported that 61 % of catchment is covered by permeable lithology while the remainder is characterized by lower permeability substrate. The elevation varies between 195 to 532 m. with a mean value of 380 m. The slope of the catchment varies between 0–200 %, with a mean value of 17 % (Gharari et al., 2011).

The hydrological data used in this study include discharge measured at the outlet of the Wark catchment, potential evaporation estimated by the Hamon equation (Hamon, 1961) with temperature data measured at Luxembourg airport (Fenicia et al., 2008a); and precipitation measured by three tipping bucket rain gauges located at Reichlange. The temporal resolution used in this study is 3 h.

3 FLEX-TOPO framework

Realizing the potential of "reading the landscape" in a systems approach (cf. Sivapalan et al., 2003), Savenije (2010) argued that due to the co-evolution of topography, soils and vegetation, all of which define the hydrological function of a given location, an efficient, hydrologically meaningful descriptor of topography together with land use could be used to distinguish different HRUs. HAND, which can be loosely interpreted as the hydraulic head at a given location in a catchment, may be such a descriptor as it potentially allows for meaningful landscape classification (e.g. Rennó et al., 2008; Gharari et al., 2011). It was argued previously (Gharari et al., 2011) that, in Central European landscapes, HAND can efficiently distinguish between wetlands, hillslopes and plateaus. These are landscape elements that may also be assumed to fulfill distinct hydrological functions (HRUs) in the study catchment (Savenije, 2010). Wetlands, located at low elevations above streams, are characterized by shallow ground water tables with limited fluctuations. Due to reduced storage capacity between ground water table and soil surface, potentially exacerbated by the relative importance of the capillary fringe, wetlands tend to be saturated earlier during a rainfall event than the two other landscape elements with arguably higher storage capacity, thus frequently becoming the dominant source of storm flow during comparably dry periods (e.g. Seibert et al., 2003; McGlynn et al., 2004; Molénat et al., 2005; Blume et al., 2008; Anderson et al., 2010; Kavetski et al., 2011). The dominant runoff process in wetlands can therefore be assumed to be saturation overland flow. Contrastingly, forested hillslopes, landscape elements with steeper slopes than the wetlands or plateaus, require a balance between sufficient storage capacity and efficient drainage to develop and maintain the ecosystem (Savenije, 2010). A dual system combining sufficient water storage in the root zone and efficient lateral drainage through preferential flow networks, controlled by a suite of activation thresholds as frequently observed on hillslopes (e.g. Hewlett, 1961; Beven and Germann, 1982; Sidle et al., 2001; Freer et al. , 2002; Weiler et al., 2003; McNamara et al., 2005; Tromp-van Meerveld and McDonnell, 2006a, b; Zehe and Sivapalan, 2009; Spence, 2010) can be seen as the dominant mechanism. Finally, plateaus are undulating landforms with low to moderate slopes and comparably deep ground water tables. In absence of significant topographic gradients and due to the potentially increased unsaturated storage capacity, it can be hypothesized that the primary functions of plateaus are sub-surface storage and groundwater recharge (Savenije, 2010). Although plateau may experience infiltration excess overland flow in specific locations, the topographical gradients may not be sufficient to generate surface runoff connected to the toward stream network. In the FLEX-TOPO approach the proportions of the hydrologically distinct landscape units, i.e. HRUs, in a given catchment need to be determined on the basis of topographical and land cover information. Subsequently suitable model structures and parameterizations will be assigned to the different HRUs (Fenicia et al., 2011; Kavetski et al., 2011; Clark et al., 2009). The integrated catchment output, i.e. runoff and evaporative fluxes, can then be obtained by combining the computed proportional output from the individual HRUs. Note that the three landscape classes tested for suitability in this study, i.e. wetland, hillslope and plateau together with their assumed dominant runoff process are designed for the Wark catchment and are likely to be different for other environmental settings.

3.1 Landscape classification

As the objective of FLEX-TOPO is to efficiently extract and use hydrologically relevant information from worldwide readily available topographic data, i.e. DEMs, the Height Above the Nearest Drainage (HAND; Rennó et al., 2008; Nobre et al., 2011) is a potentially powerful metric to classify landscapes into HRUs with distinct hydrological function, as discussed above. Testing a suite of HAND-based classification methods (Gharari et al., 2011) found that results best matching observed landscape types could be obtained by using HAND together with the local slope. Based on a probabilistic framework to map the desired HRUs which were then compared with in-situ observations they obtained a threshold for HAND and slope of approximately 5 m and 11 % for the Wark catchment. Following from that, wetlands were defined to be areas with HAND index lower than 5 m. Areas with HAND index more than 5 m and local slopes more than 11 % were classified as hillslopes, while areas with HAND index more than 5 m and local slope less than 11 % were defined as plateaus. The HAND and slope map of the study catchment together with the classified landscape entities (wetland, hillslope and plateau) are presented in Fig. 1. The proportion of the individual HRUs wetland, hillslope and plateau are 15 %, 45 % and 40 % respectively.

3.2 Model set up

In this study a lumped conceptual model of the Wark catchment, hereafter referred to as FLEX^A, is used as benchmark as lumped conceptual models are frequently used in catchment hydrology, particularly in small- to mesoscale catchments (e.g. Merz and Blöschl, 2004; Clark et al., 2008; Perrin et al., 2008; Seibert and Beven, 2009; Fenicia et al., 2013). The above discussed concept of FLEX-TOPO (Savenije, 2010) is thereafter tested with a stepwise increased number of landscape units (FLEX^B, FLEX^C), thereby increasing the conceptualized process heterogeneity and thus the model complexity. The core of the three model set-ups is loosely based on the FLEX model (Fenicia et al., 2006).

3.2.1 FLEX^A

This model set-up represents the catchment in a lumped way. The FLEX^A model structure consists of four storage elements representing interception, unsaturated, slow (i.e. groundwater) and fast responding reservoirs (i.e. preferential flow and saturation overland flow). A schematic illustration of FLEX^A is shown in Fig. 2a. The water balance and constitutive equations used are given in Table 2.

Interception reservoir (S_I)

The interception reservoir is characterized by its maximum storage capacity (I_{max} [L]). After precipitation (P [L T^{-1}]) enters this reservoir the excess precipitation, hereafter referred to as effective precipitation (P_e [L T^{-1}]), is distributed between the unsaturated (S_U), slow (S_S) and fast reservoir (S_F).

Unsaturated reservoir (S_U)

The unsaturated reservoir is characterized by a parameter that loosely reflects the maximum soil moisture capacity in the root zone ($S_{U,max}$ [L]). Part of the effective precipitation (P_e) enters the unsaturated zone according to the coefficient C_r, which here is defined by a power function with exponent β [–], reflecting the spatial heterogeneity of thresholds for activating fast lateral flows from S_F. This coefficient C_r will be 1 when soil moisture (S_U) is lower than a specific percentage of maximum soil moisture capacity ($S_{U,max}$) defined by field capacity (F_C[–]), meaning that the entire incoming effective precipitation (P_e) at a given time step is stored in the unsaturated reservoir (S_U). The soil moisture reservoir feeds the slow reservoir through matrix percolation (R_p [LT-1]), expressed as a linear relation of the available moisture in the unsaturated zone (S_U) and the maximum percolation capacity (P_{Per} [L T^{-1}]). The reverse process, capillary rise (R_C), feeds the unsaturated reservoir from the saturated zone. Capillary rise (R_C [L T^{-1}]) has an inverse linear relation with the moisture content in the unsaturated zone and is characterized by the maximum capillary rise capacity (C [L T^{-1}]). Soil moisture is depleted by plant transpiration. Transpiration is assumed to be moisture constrained when the soil moisture content is lower than a fraction L_p [–] of the maximum unsaturated capacity ($S_{U,max}$). When the soil moisture content in the unsaturated reservoir is higher than this fraction (L_p) transpiration is assumed to be equal to the potential evaporation (E_{pot} [L T^{-1}]).

Splitter and transfer functions

The proportion of effective rainfall which is not stored in the unsaturated zone, i.e. $1-C_r$, is further regulated by the partitioning coefficient (D [–]), distributing flows between preferential groundwater recharge (R_S [L T^{-1}]) to S_S and water that is routed to the stream by fast lateral processes from S_F (e.g. preferential flow or saturation overland flow, R_F). Both fluxes are lagged by rising linear lag functions with parameters N_{lagf} and N_{lags}, respectively (e.g. Fenicia et al., 2008b).

Fast reservoir (S_F)

The fast reservoir is a linear reservoir characterized by fast reservoir coefficient S_F.

Slow reservoir (S_S)

The slow reservoir is a linear reservoir characterized by slow reservoir coefficient S_S.

3.2.2 FLEX^B

As discussed above, a range of process studies suggested that wetlands can frequently exhibit storage-discharge dynamics that are decoupled from other parts of a catchment, in particular due to their typically reduced storage capacity and closeness to the stream. FLEX^B explicitly distinguishes wetlands from the rest of the catchment, the "remainder" (i.e. hillslopes and plateaus), which is represented in a lumped way, to account for this difference. The FLEX^B model set-up therefore consists of two parallel model structures which are connected with a common groundwater reservoir (Fig. 2b), similar to what has been suggested by Knudsen et al. (1986). One major difference between the two parallel structures is that capillary rise is assumed to be a relevant process only in the wetland, while it is considered negligible in the remainder of the catchment due to the deeper groundwater. Further, since the wetlands are predominantly ex-filtration zones of potentially low permeability, preferential recharge is considered negligible in wetlands. The two landscape units used in this model set-up, i.e. wetland and the remainder of the catchment, share a common slow reservoir. The areal proportions of wetland and the remainder (i.e. hillslope and plateau) of the catchment are 15 and 85 %, respectively (Gharari et al., 2011).

3.2.3 FLEX^C

This model set-up offers a complete representation of the three HRUs in the study catchment: wetland, hillslope and plateau (Fig. 2c). The formulation of the wetland module in FLEX^C is identical to the one suggested above for FLEX^B. The hillslope HRU is represented by a model structure resembling the FLEX^A set-up. Plateaus are assumed to be dominated by vertical fluxes, while direct lateral flows are considered negligible compared to those generated from hillslope and wetland HRUs. Therefore the plateau model structure does not account for these fast fluxes. Analogous to FLEX^B, the FLEX^C set-up is characterized by one single groundwater reservoir linking the three dominant HRUs in this catchment. The individual proportions of wetland, hillslope and plateau are 15, 45 and 40 %, respectively (Gharari et al., 2011). The proportions of these HRUs are used to compute the total discharge based on the contribution of each landscape unit.

The connection between the parallel structures of FLEX^B and FLEX^C is through the surface drainage network (the stream network) and through the slow (groundwater) reservoir.

3.3 Introducing realism constraints in selecting behavioral parameter sets

With increasing process heterogeneity from FLEX^A over FLEX^B to FLEX^C, the respective model complexities and therefore the number of calibration parameters also increase. This, in the frequent absence of sufficient suitable data to efficiently constrain a model, typically leads to a situation where parameters have increased freedom to compensate for errors in data and model structures, as recently reiterated by Gupta et al. (2008). As a consequence, the resulting higher risk of equifinality can substantially reduce a model's predictive power. As discussed earlier, to avoid the problem of equifinality, in this study, two fundamentally different types of constraints have been applied in the models to test their value for reducing equifinality in complex model set-ups. Firstly, conditions between parameters of the different parallel model units, hereafter referred to as parameter constraints, were imposed before each model evaluation run. These a priori constraints ensure that the individual parameter values for the same process in the parallel units, reflect the modeler's perception of the system. For example, it can be argued that the maximum interception capacity (I_{max}) of a forested HRU needs to be higher than the one in a not forested HRU. The second type of constraints is process constraints, which can only be applied after each evaluation run during the calibration phase. These a posteriori constraints compare the modeled output of the individual HRUs and ensure that these outputs follow the modeler's perception of the system's internal dynamics. For example, it can be argued that the modeled evaporation in forested HRUs needs to be higher than in not forested HRUs. The parameter and process constraints imposed on the models in this study are described in detail below. Note that which constraints to impose is the modeler's choice and that with increasing number of different HRUs more and more constraints can be applied. In this study, FLEX^A only allows one prior constraint related to the overall runoff coefficient, while all of the constraints suggested below can be applied to FLEX^C.

3.3.1 Parameter constraints

A number of a priori constraints is imposed on the relative value of different model parameters in order to exclude unrealistic parameter combinations. The constraints are guided by considerations on what the model components are designed to reproduce. The number of constraints that can be imposed increases with increasing model complexity. The full set of parameter constraints detailed below were applied to FLEX^C and when applicable also to FLEX^B. No parameter constraint could be used for FLEX^A, as for this model no obvious relationship between parameters could be identified. In the following, the subscripts w, h and p indicate parameters for wetland, hillslope and plateau, respectively.

Interception

The different land cover proportions of each landscape unit, here wetlands, hillslopes and plateaus, can be used to define the relation between interception thresholds (I_{max}) of these individual units. The land uses are defined as two general classes for this case study, forested areas and grass or pasture-land areas. The maximum interception capacity (I_{max}) for each landscape entity can be explained by proportion of every land-use class and their maximum interception capacity which is selected form defined ranges in Table 1. Therefore the maximum interception capacity for each landscape units can be derived as below:

$$I_{\text{max, w}} = a_\text{w} I_{\text{max, forest}} + b_\text{w} I_{\text{max, cropland}} \tag{1}$$

$$I_{\text{max, h}} = a_\text{h} I_{\text{max, forest}} + b_\text{h} I_{\text{max, cropland}} \tag{2}$$

$$I_{\text{max, p}} = a_\text{p} I_{\text{max, forest}} + b_\text{p} I_{\text{max, cropland}} \tag{3}$$

The proportions of forested area are indicated with a_w, a_h and a_p for wetland, hillslope and plateau and are fixed as 42, 60 and 29 % respectively. The proportions of cropland and grass land areas are indicated by b_w, b_h and b_p for wetland, hillslope and plateau and are fixed as 58, 40 and 71 % respectively. Moreover the parameter sets which are selected for maximum interception capacity of forest are expected to be higher than crop- or grassland:

$$I_{\text{max, cropland}} < I_{\text{max, forest}} \tag{4}$$

Lag functions

Preferential recharge (R_S) is routed to the slow reservoir by a lag function. Due to a deeper groundwater table on plateaus it can be assumed that the lag time for (R_S) is longer for plateaus than for hillslopes. It can also be assumed that the lag function used for fast reservoir for hillslopes is longer than for wetlands due to the on average higher distance of and therefore longer travel times from hillslopes to the stream.

$$N_{\text{lags, w}} \leq N_{\text{lags, h}} \leq N_{\text{lags, p}} \tag{5}$$

Soil moisture capacity

Many experimental studies support the assumption that wetlands have shallower groundwater tables than the other two landscape entities in this study. Therefore the unsaturated zone of wetland should be shallower, i.e. the maximum soil moisture capacity ($S_\text{U,max}$) compared to hillslopes and plateaus can be assumed to be higher. Moreover, as hillslopes in the study catchment are predominantly covered with forest, it can, due to the deeper root zone of forests, be expected that the maximum unsaturated soil moisture capacity ($S_{\mathrm{U,max}}$) in the root zone of hillslopes be deeper than the other two landscape entities.

$$S_{\mathrm{U,max,\ w}} < S_{\mathrm{U,max,\ p}} < S_{\mathrm{U,max,\ h}} \tag{6}$$

Reservoir coefficients

The reservoir coefficient of the wetland fast reservoir (K_F) is assumed to be lower than reservoir coefficient of the hillslope fast reservoir as, once connectivity is established, the flow velocities of saturation overland flow in wetlands are assumed to exceed the integrated flow velocities of preferential flow networks (cf. Anderson et al., 2009). Consequently, the retention time of the slow reservoir should be higher than both wetland and hillslope fast reservoirs.

$$K_{\mathrm{F,w}} < K_{\mathrm{F,h}} < K_\mathrm{S} \tag{7}$$

3.3.2 Process constraints

In contrast to the parameter constraints discussed above, which are set a priori, process constraints are applied a posteriori. Only parameterizations which generate model internal flux dynamics in agreement with the modeler's perception of these dynamics are retained as feasible. Hence, while with the use of parameter constraints there is no need to run the model, here it is necessary to run the model to evaluate the individual fluxes.

The process constraints are defined for dry and wet periods as well as for peak-, high- and low flows. Here wet periods were defined to be the months from December to March, while the dry periods in the study catchment occur between April and November. The thresholds for distinguishing between high and low flow were chosen to be 0.05 and 0.2 mm$\,(3\,\mathrm{h})^{-1}$ respectively for dry and wet periods. Furthermore events in which the discharge increases with a rate of more than $0.2\,\mathrm{mm}\,(3\,\mathrm{h})^{-2}$ are defined as peak flows. Note that in the following the subscripts peak, high and low indicate peak-, high- and low flows.

Transpiration

Transpiration typically exhibits a clear relationship with the normalized difference vegetation index (NDVI). Therefore the ratios between NDVI values of different landscape units can serve as constraints on modeled transpiration obtained from the individual parallel model components. A rough estimation of the ratio between transpiration from plateau and hillslope can be derived from LANDSAT 7 images. For this ratio seven cloud free images have been selected (acquisition dates of 20 April 2000, 06 May 2000, 11 September 2000, 18 February 2001, 06 March 2001, 26 July 2001 and 29 August 2001). The ratio of transpiration between hillslope and plateau (R_{trans}) can be estimated by assuming a linear relation (Szilagyi et al. , 1998) with slope of α and intercept zero between transpiration and mean NDVI for each landscape unit (μ_{NDVI}).

$$R_{\mathrm{trans}} = \frac{\alpha\mu_{\mathrm{NDVI,h}}}{\alpha\mu_{\mathrm{NDVI,p}}} = \frac{\mu_{\mathrm{NDVI,h}}}{\mu_{\mathrm{NDVI,p}}} \tag{8}$$

Mean ($\mu_{R_{\mathrm{trans}}}$) and standard deviation ($\sigma_{R_{\mathrm{trans}}}$) of R_{trans} can be used to estimate acceptable limits of the transpiration ratios for hillslope and plateau. Therefore the annual transpiration can be confined between two values as follows:

$$\mu_{R_{\mathrm{trans}}} - \sigma_{R_{\mathrm{trans}}} < \frac{\int T_{\mathrm{h}}\mathrm{d}t}{\int T_{\mathrm{p}}\mathrm{d}t} < \mu_{R_{\mathrm{trans}}} + \sigma_{R_{\mathrm{trans}}} \tag{9}$$

Based on the mean ($\mu_{R_{\mathrm{trans}}} = 1.2$) and standard deviation ($\sigma_{R_{\mathrm{trans}}} = 0.2$) of the seven LANDSAT 7 images used the following process constraint was imposed:

$$1.0 < \frac{\int T_{\mathrm{h}}\mathrm{d}t}{\int T_{\mathrm{p}}\mathrm{d}t} < 1.4 \tag{10}$$

Similar constraints can be imposed between transpiration fluxes from wetland, hillslope or plateau; in this study catchment defined wetland extent can only be capture by DEM resolution of 20 m and higher which is smaller compare to mentioned LANDSAT 7 images therefore the constraints comparing the transpiration fluxes from wetland to other landscape entities were not imposed.

Runoff coefficient

The runoff coefficient is a frequently used catchment signature (e.g. Sawicz et al., 2011; Euser et al., 2013) and can be used as a behavioral constraint (e.g. Duan et al., 2006; Winsemius et al., 2009). In this study the runoff coefficients of dry and wet periods as well as the annual runoff coefficient were used. Parameterizations that result in modeled runoff coefficients that substantially deviate from the observed ones are therefore discarded. In case of absence of suitable runoff data to estimate the runoff coefficient it can be derived from the regional Budyko curve using for example the Turc-Pike relationship (Turc, 1954; Pike, 1964; Arora, 2002). However in this study the runoff coefficients of each individual year, and of their respective dry and wet periods was used and determined the mean and standard deviation of the runoff coefficients for these periods. Here, as a conservative assumption, the limits are set to three times the standard deviation around the mean runoff coefficient. Note that the runoff coefficient is the only constraint that is not related to model structure in this study and can therefore also be applied to the lumped FLEX^A set-up.

$$\frac{\int Q_\mathrm{m} \mathrm{d}t}{\int P \mathrm{d}t} < 0.43 \tag{11}$$

$$\frac{\int Q_\mathrm{m} \mathrm{d}t}{\int P \mathrm{d}t} > 0.16 \tag{12}$$

$$\frac{\int Q_\mathrm{m,dry} \mathrm{d}t}{\int P_\mathrm{dry} \mathrm{d}t} < 0.36 \tag{13}$$

$$\frac{\int Q_\mathrm{m,dry} \mathrm{d}t}{\int P_\mathrm{dry} \mathrm{d}t} > 0 \tag{14}$$

$$\frac{\int Q_\mathrm{m,wet} \mathrm{d}t}{\int P_\mathrm{wet} \mathrm{d}t} < 0.71 \tag{15}$$

$$\frac{\int Q_\mathrm{m,wet} \mathrm{d}t}{\int P_\mathrm{wet} \mathrm{d}t} > 0.40 \tag{16}$$

Preferential recharge

The slow reservoir can be recharged by both preferential and matrix percolation from the unsaturated reservoirs. Both hillslopes and plateaus contribute to slow reservoir by preferential recharge. It can be assumed that in a realistic model setup the long term contribution volume of preferential recharge ratio between hillslope and plateau should not be unrealistically high or low. For example it was assumed unrealistic that the ratio is zero or infinity, meaning that one landscape unit is constantly feeding the slow reservoir while another one is not contributing at all. To avoid such a problem, a loose and very conservative constraint is imposed on the ratio of contribution of the two fluxes.

$$0.2 < \frac{\int R_{\mathrm{s,h}} \mathrm{d}t}{\int R_{\mathrm{s,p}} \mathrm{d}t} < 5 \tag{17}$$

Fast component discharge

During dry periods, hillslopes and plateaus can exhibit significant soil moisture deficits, limiting the amount of fast runoff generated from these landscape elements. In contrast, due to their reduced storage capacity, wetlands are likely to generate fast flows at lower moisture levels, thus dominating event response during dry periods (cf. Beven and Freer, 2001a; Seibert et al., 2003; Molénat et al., 2005; Anderson et al., 2010; Birkel et al., 2010). It can thus be assumed that during both, the entire dry periods as well as peak flows in dry periods the fast component of wetlands ($Q_{\mathrm{f,\ w,dry}}$; $Q_{\mathrm{f,\ w,dry,\ peaks}}$) contributes to runoff more than the fast component of hillslopes ($Q_{\mathrm{f,\ h,dry}}$; $Q_{\mathrm{f,\ h,dry,\ peaks}}$). In contrast, high flows during wet periods are predominantly generated by hillslopes ($Q_{\mathrm{f,\ h,wet}}$; $Q_{\mathrm{f,\ h,wet,\ high}}$).

$$\frac{\int Q_{\mathrm{f,h,dry,peaks}} \mathrm{d}t}{\int Q_{\mathrm{f,w,dry,peaks}} \mathrm{d}t} < 1 \tag{18}$$

$$\frac{\int Q_{\mathrm{f,h,wet,high}} \mathrm{d}t}{\int Q_{\mathrm{f,w,wet,high}} \mathrm{d}t} > 1 \tag{19}$$

$$\frac{\int Q_{\mathrm{f,h,wet}} \mathrm{d}t}{\int Q_{\mathrm{f,w,wet}} \mathrm{d}t} > 1 \tag{20}$$

3.3.3 Calibration algorithm and objective functions

Based on uniform prior parameter distributions, as well as on the parameter- and process constraints the model was calibrated using MOSCEM-UA (Vrugt et al., 2003). As a brief description, MOSCEM-UA uses a Latin Hypercube sampling strategy for the random sampling of the entire parameter space. Introducing constraints, however, may lead to non-smooth objective functions which potentially may cause instabilities in the search algorithm and/or create invalid results. A recently developed stepwise search algorithm was therefore used for finding parameter sets which satisfy both parameter- and process constraints (Gharari et al., 2013b). These parameter sets were then used as initial sampling parameter sets for MOSCEM-UA instead of the traditionally used Latin Hypercube sampling strategy.

The models were evaluated on the basis of three different objective functions to emphasize different characteristics of the system response: (i) the Nash-Sutcliffe efficiency of the flows (Nash and Sutcliffe, 1970, ; E_{NS}), (ii) the Nash-Sutcliffe efficiency of the logarithm of the flows ($E_{\mathrm{NS,log}}$) and (iii) the Nash-Sutcliffe efficiency of the flow duration curve ($E_{\mathrm{NS,FDC}}$). These criteria evaluate the models' ability to simultaneously reproduce high flows, low flows and flow duration curves respectively. The model set ups have been constrained and calibrated for the year 2006–2007 and validated for year 2008–2009. The year 2005 was used as warming up period.

3.4 Model validation and parameter evaluation

To assess the value of incorporating parameter and process constraints in increasingly complex models a four-step procedure as outlined below was followed. Note that for each step the respective model (parameterization) was evaluated against the constrained and calibrated lumped FLEX$^{\mathrm{A}}$ benchmark model.

3.4.1 Evaluating models with "constrained but uncalibrated" parameter sets

Firstly, the parameter sets which satisfy all the applied constraints were evaluated based on their ability to reproduce the observed hydrograph. Hereafter these parameter sets are referred to as *constrained but uncalibrated* parameter sets because they were obtained *without any calibration* on the observed hydrographs. Based on the retained, feasible parameter sets, the mean performance of the three *constrained but uncalibrated* models FLEX^A, FLEX^B and FLEX^C, for the three objective functions (E_{NS}, $E_{NS,log}$, $E_{NS,FDC}$) together with their uncertainty ranges for both the calibration and the validation periods are compared. FLEX^A, FLEX^B and FLEX^C have an increasing number of constraints. It was thus tested whether the higher complexity models also result in better model performance and how the predictive uncertainty is affected by increased complexity and model realism. To investigate how well the hydrographs generated with parameters satisfying all constraints match the observed hydrograph, the 95 % uncertainty intervals of simulated hydrographs based on these parameter sets were generated for the three models. The uncertainty was estimated on the basis of the area indicated by 95 % uncertainty intervals based on simulated hydrographs.

3.4.2 Evaluating models with "constrained and calibrated" parameter sets

In a second step the three models FLEX^A, FLEX^B and FLEX^C have been calibrated within the parameter space which satisfied all the imposed parameter and process constraints. The models were calibrated using a multi-objective strategy (E_{NS}, $E_{NS,log}$, $E_{NS,FDC}$). The obtained Pareto optimal model parameterizations are in the following referred to as *constrained and calibrated*. Analogous to the previous step uncertainty intervals based on the constrained and calibrated Pareto optimal parameterizations, were generated. The uncertainty was estimated on the basis of the area of the uncertainty bands.

3.4.3 Comparison of model performance and uncertainty for *constrained but un-calibrated* and *constrained and calibrated* parameterizations sets

To assess the added value of incorporating constraints in higher complexity models, the performance and uncertainties of the three models FLEX^A, FLEX^B and FLEX^C were compared for both the *constrained but un-calibrated* and the *constrained and calibrated* case during calibration (2006–2007) and validation (2008–2009) periods.

3.4.4 Comparison of modeled hydrograph components for different model structures

One of the main reasons for imposing constraints on model parameterization is to ensure the realistic internal dynamic of a model. Comparing different fluxes contributing to the modeled hydrograph can give an insight into the performance of imposed constrained on the model. The effect of imposing behavioral constraints on fast and slow components of the three models structures, FLEX$^{\mathrm{A}}$, FLEX$^{\mathrm{B}}$ and FLEX$^{\mathrm{C}}$ is compared visually. The fast component of lumped model, FLEX$^{\mathrm{A}}$, is compared with fast components of FLEX$^{\mathrm{B}}$ which are wetland and remainder of catchment and fast components of FLEX$^{\mathrm{C}}$ which are wetland and hillslope. This visual comparison is based on normalized average contribution of each component for Pareto optimal parameter sets in every time step

4 Results and discussion

4.1 Evaluating the performance of *constrained but uncalibrated* parameter sets

The median and the 95 % uncertainty intervals of the performance of modeled hydrographs for *constrained but uncalibrated* parameter sets is presented in Table 3 for the calibration period. The lumped FLEX$^{\mathrm{A}}$ model has only one process constraint, i.e. the runoff coefficient. Hence, this model is free within the limits of this apparently relatively weak condition, resulting in a wide range of possible parameterizations, many of which cannot adequately reproduce the system response. As a consequence, the overall performance is poor ($E_{\mathrm{NS,median}} = 0.31$, $E_{\mathrm{NS,log,\ median}} = -0.48$, $E_{\mathrm{NS,FDC,median}} = 0.66$) with considerable uncertainty in the modeled hydrograph (Table 3, Fig. 3).

FLEX$^{\mathrm{B}}$, run with the set of *constrained but uncalibrated* parameters shows a substantial improvement in overall performance ($E_{\mathrm{NS,median}} = 0.51$, $E_{\mathrm{NS,log,\ median}} = 0.30$, $E_{\mathrm{NS,FDC,median}} = 0.87$) compared to FLEX$^{\mathrm{A}}$, as it not only allows for more process heterogeneity but, more importantly, as it is conditioned with an increased number of constraints.

The additional process heterogeneity and constraints allowed by FLEX^C, results in the highest overall performance for all three objective functions ($E_{\text{NS,median}} = 0.65$, $E_{\text{NS,log, median}} = 0.55$, $E_{\text{NS,FDC,median}} = 0.97$) and narrowest uncertainty intervals for this comparatively complex model set-up (Table 3, Fig. 3).

Table 3, moreover, presents the performance of *constrained but uncalibrated* parameter sets from the calibration period in the validation period. The comparison of the respective 95 % uncertainty intervals of the three different models shows the capability of FLEX^C to reproduce the features of the hydrograph with lower uncertainty than FLEX^A or FLEX^B (Fig. 3). This is demonstrated by the reduction of the total uncertainty area with the gradual introduction of more constraints from FLEX^A over FLEX^B to FLEX^C (Table 3).

These results clearly illustrate that the imposed relational constraints force the model and its parameterization towards a more realistic behavior, which significantly improves model performance and considerably reduces predictive uncertainty even in the absence of actual calibration.

4.2 Evaluating the performance of *constrained and calibrated* parameter sets

The comparison of the constrained and calibrated model set-ups shows that all three models set-ups can reproduce the hydrograph similarly well (Table 4, Fig. 4). FLEX^A exhibits a slightly better calibration performance compared to the other two model set-ups. This can partly be attributed to the lower number of parameters which leads, with the same number of samples, to a more exhaustive sampling of the parameter space and a smoother identification of Pareto optimal solutions. In addition, FLEX^A has the lowest number of imposed constraints, i.e. only the runoff coefficient, compared to FLEX^B and FLEX^C. This model set-up therefore allows more freedom in exploiting the parameter space to produce mathematically good fits between observed and modeled system response in the calibration period.

For the validation period, arguably the more important evaluation period because, in contrast to the calibration period, it gives information on model consistency (cf. Klemeš, 1986; Andréassian et al., 2009; Euser et al., 2013) and predictive uncertainty, the performances of the three model set-ups exhibit quite different patterns (Table 4). The simplest model, the lumped FLEX$^{\mathrm{A}}$, is characterized by the highest performance deterioration from calibration to validation. FLEX$^{\mathrm{B}}$ shows a better validation/calibration performance ratio than FLEX$^{\mathrm{A}}$. Despite the expectation that increasingly complex models will have increasingly poor validation/calibration performance ratios, due to higher degrees of freedom, FLEX$^{\mathrm{C}}$ exhibited a more stable performance between calibration and validation. In addition, the absolute performance of FLEX$^{\mathrm{C}}$ in the validation period is in general higher than the performances of FLEX$^{\mathrm{A}}$ and FLEX$^{\mathrm{B}}$ (Table 4). Although, strictly speaking, no meaningful comparison between Nash-Sutcliffe efficiencies from different periods can be made, these results nevertheless indicate that the most complex model set-up, i.e. FLEX$^{\mathrm{C}}$, is the most consistent model-set-up with the lowest predictive uncertainty, which has important implications that will be discussed below. The explanation is that in spite of the high degree of process heterogeneity, the high number of constraints in FLEX$^{\mathrm{C}}$ prevents the calibration algorithm to over-fit this complex model set-up, thus reducing the probability of seriously misrepresenting reality.

4.3 Comparison of *constrained but uncalibrated* and *constrained and calibrated* models

The following comparison of the performances of FLEX$^{\mathrm{A}}$, FLEX$^{\mathrm{B}}$ and FLEX$^{\mathrm{C}}$ for *constrained but uncalibrated* and *constrained and calibrated* parameter sets focused on E_{NS} only, for the reason of brevity (Fig. 5). In Fig. 5a and b the model performances based on the *constrained but uncalibrated* parameter sets, that satisfy the full set of constraints, are shown for the calibration and validation periods. As discussed in detail above, even uncalibrated, increasing the number of constraints from FLEX^A to FLEX^C increases the overall performance of the models while reducing uncertainty (Fig. 5c and d; note that these are zoom-ins).

Figure 5e compares model performance based on *constrained and calibrated* parameter sets for the calibration period. As discussed earlier, it can be clearly seen that the simple lumped model, FLEX^A, shows the best calibration performance with lowest uncertainty. However, when comparing the individual model performances of the constrained and calibrated models during the validation period (Fig. 5f), it can be seen that FLEX^A not only shows the strongest performance deterioration compared to the calibration period but also that FLEX^A is also the model with the poorest performance in the validation period. This implies that although FLEX^C is the most complex model, the realism constraints imposed on this model generate the most reliable outputs when used for prediction, i.e. in the validation period. This strongly underlines that the widely accepted notion of complex models necessarily being subject to higher predictive uncertainty is not generally valid when the model parameters can be well constrained based on assumptions of realistic functionality of a catchment.

In addition a second crucial aspect was revealed by comparing *constrained but uncalibrated* and *constrained and calibrated* models. It can be seen that, for the study catchment, a *calibrated* lumped model, FLEX^A (Fig. 5f) can on average not clearly outperform a more complex *constrained but uncalibrated* model, i.e. FLEX^C (Fig. 5d). This has potentially important implications for the parameterization of models in ungauged basins as it highlights the value of semi- and non-quantitative hydrological expert knowledge, even in the absence of reliable model regionalization tools and detailed soil or geological information, as discussed in detail below. Note that when interpreting the results based on the low flow performance ($E_{\text{NS,log}}$), the conclusion is somewhat weaker, yet still clearly illustrating the value of relational constraints for applications in ungauged basins.

4.4 Comparison of flow contributions from different model components

The comparison of the fluxes generated from the individual model components in the three model set-ups helps to assess to which degree the model internal dynamics reflect the modeler's perception of the system and thus to a certain degree the realism of the models.

Fast and slow responses of each tested model set-up have been visually illustrated in Fig. 6. Predominance of slow responses of all the three models are indicated by green color; predominance of fast responses of FLEX^A, fast responses of the remainder of the catchment of FLEX^B and fast responses of hillslope of FLEX^C is indicated by red color; wetland fast responses of FLEX^B and FLEX^C are indicated by predominance of blue color.

The colors in Fig. 6 are an illustration using RGB (red, green and blue) color code for the models' responses based on their weight of contribution to the modeled runoff. As it can be seen in Fig. 6a the fast component of FLEX^A is dominant just during peak flows and even the recession shortly after peak flows are accounted for mainly by ground water. Analysis of the individual model components computed by Pareto optimal parameter sets (not shown here for brevity), indicates that some Pareto optimal parameterizations can generate peak flows by predominant contributions from slow responses while fast reaction is tend to be inactive during these events.

In contrast, Fig. 6b and c show that the early recessions after peak flows are mostly accounted for by fast components of hillslope in FLEX^B and FLEX^C while later the slower groundwater component becomes dominant. The fast response of hillslopes in FLEX^C is less dominant compare to fast reaction of remainder of catchment during dry period. This appears to be linked to the inclusion of the plateau, in FLEX^C. In accordance with the perception of the system that wetlands are predominantly responsible for peak flows during dry conditions, Fig. 6b and c show that wetland fast responses in FLEX^B and FLEX^C control the peaks during dry period as well as during wetting up periods (dry to wet transition), before hillslope fast processes become more important at higher moisture levels.

4.5 General discussion

The results of this study quite clearly indicate that discretizing the catchment into hydrological response units (HRUs) and incorporating expert knowledge in model development and testing is a potentially powerful strategy for runoff prediction, even where insufficient data for model calibration (e.g. Koren et al., 2003; Duan et al., 2006; Winsemius et al., 2009) or only comparatively unreliable regionalization tools are available (e.g. Wagener and Wheater, 2006; Bárdossy, 2007; Parajka et al., 2007; Oudin et al., 2008; Laaha et al., 2012). It was found that the performance and the predictive power of a comparatively complex uncalibrated conceptual model, based on posterior parameter distributions obtained merely from relational, semi- and non-quantitative realism constraints inferred from expert knowledge, can be as efficient as the calibration of a lumped conceptual model (Fig. 5).

Typically it is expected that, if not warranted by data, models with higher complexity suffer from higher predictive uncertainty. As stated by Beven (2001): “More complexity means more parameters, more parameters mean more calibration problems, more calibration problems will often mean more uncertainty in the predictions, particularly outside the range of the calibration data”. Thus, more parameters would allow better fits of the hydrograph but would not necessarily imply a better and more robust understanding of catchment behavior or more reliable predictions.

A complex model may include many processes, i.e. hypotheses, which can usually not be rigorously tested with the available data. However, a wide range of previous studies has demonstrated that hydrologically meaningful constraints can help to limit the increased uncertainty caused by incorporating additional processes, i.e. parameters (e.g. Yadav et al., 2007; Zhang et al., 2008; Kapangaziwiri et al., 2012). These studies generally include a large set of catchments and try to relate model parameters to catchment characteristics. Although regional constraints are important, the importance of expert knowledge on the catchment scale, which leads to better understanding of hydrological behavior is highlighted in this study.

In a similar attempt, Pokhrel et al. (2008, 2012) demonstrated use of regularization for model parameterization and reduction of model parameter space dimensionality by linking model parameters using super-parameters to catchment characteristics. However, no explicit hydrological reasoning is typically applied for such “regularization rules” (e.g. Pokhrel et al., 2012). On the other hand, Kumar et al. (2010, 2013) parameterize and successfully regionalize their models using empirical transfer functions with global parameters, developed from extensive literature study and iterative testing in a large sample of catchments In contrast, the use of relational parameter- and process constraints, as presented in this study, is based on semi-quantitative, hydrologically explicit and meaningful reasoning avoiding the need for empirical transfer functions to link catchments characteristics and model parameterizations.

Including prior knowledge for parameterization of physically-based models for estimating runoff in ungauged basins was quite successfully investigated in the past (e.g. Ott and Uhlenbrook, 2004; Vinogradov et al., 2011; Fang et al., 2013; Semenova et al., 2013). These studies specifically indicate that calibration can be replaced by prior information which is a significant contribution to Predictions in Ungauged Basins (PUB). While physically-based models need detailed information of catchment behavior for model parameterization, the here proposed semi-distributed conceptual modeling framework, exploiting relational constraints, can be more efficiently set up using the least prior information necessary. In this study, the performances and uncertainties of the three tested model set-ups for constrained but uncalibrated parameters indicate the potential of the presented TOPO-FLEX framework for Predictions in Ungauged Basins. Hence, this framework can efficiently use expert knowledge for improving model parameterization in complex conceptual hydrological models, not only to increase model performance but also to reduce model predictive uncertainty even in the absence of calibration.

It should be noted that the model set-ups suggested within the FLEX-TOPO framework are hypotheses that still need to undergo further tests, ideally confronting them with additional, system internal information, such as groundwater dynamics (e.g. Seibert, 2003; Fenicia et al., 2008b) or tracer data (e.g. Birkel et al., 2011; Capell et al., 2012; Hrachowitz et al., 2013a). To make more efficient use of relational constraints, model sensitivities to these constraints need to be evaluated in the future. It is also emphasized that the constraints introduced in this study are based on the authors' subjective understanding of catchment behavior and can and should be discussed further. However, we would like to stress the notion that reaching an agreement on the relation between parameters and fluxes in different landscape units is potentially much easier than finding the actually most adequate parameter values for a conceptual model based on field observations or available data on geology or soil types.

5 Conclusions

In this study it was tested if a topography-driven semi-distributed formulation of a catchment-scale conceptual model, conditioned by expert knowledge based relational parameter- and process constraints, can increase the level of process realism and predictive power while reducing the need for calibration compared to a lumped model set-up.

It was found that:

1. A constrained but uncalibrated semi-distributed model performed equally well as a constrained and calibrated lumped model when used for prediction. This illustrates the potential value of the combined use of higher complexity models and relational constraints for predictions in ungauged basins, where no calibration data are available.

2. The use of relational parameter- and process constraints in model calibration ensured a high degree of process realism. Thus, in spite of the comparatively high

complexity, the overall model performance and uncertainty showed better prediction results than for a lumped model. It was shown that higher model complexity therefore does not necessarily entail reduced predictive power.

3. Semi-distributing a model on the basis of HRUs derived from topographic data can increase model internal consistency as it better account for fundamentally different runoff generating processes active at different wetness conditions.

4. In contrast to constraints based on more detailed and frequently unavailable regionalization relationships or catchment data, such as geology and soils, hydrologically meaningful relational constraints can be applied with a minimum of information.

Acknowledgements. Anke Luijben's assistance is appropriated for data analysis of the study area. Shervan Gharari is funded during his PhD program by Fonds National de la Recherche (FNR) of Luxembourg with Aides à la Formation-Recherche (AFR) project number of 1383201.

References

Ajami, N. K., Gupta, H. V., Wagener, T., and Sorooshian, S.: Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system, J. Hydrol., 298, 112–135, doi:10.1016/j.jhydrol.2004.03.033, 2004. 14803

Anderson, A. E., Weiler, M., Alila, Y., and Hudson, R. O.: Subsurface flow velocities in a hillslope with lateral preferential flow, Water Resour. Res., 45, W11407, doi:10.1029/2008WR007121, 2009. 14817

Anderson, A. E., Weiler, M., Alila, Y., and Hudson, R. O.: Piezometric response in zones of a watershed with lateral preferential flow as a first-order control on subsurface flow, Hydrol. Process., 24, 2237–2247, doi:10.1002/hyp.7662, 2010. 14809, 14821

Andréassian, V., Perrin, C., Berthet, L., Le Moine, N., Lerat, J., Loumagne, C., Oudin, L., Mathevet, T., Ramos, M.-H., and Valéry, A.: HESS Opinions "Crash tests for a standardized evaluation of hydrological models", Hydrol. Earth Syst. Sci., 13, 1757–1764, doi:10.5194/hess-13-1757-2009, 2009. 14826

Andréassian, V., Le Moine, N., Perrin, C., Ramos, M.-H., Oudin, L., Mathevet, T., Lerat, J., and Berthet, L.: All that glitters is not gold: the case of calibrating hydrological models, Hydrol. Process., 26, 2206–2210, doi:10.1002/hyp.9264, 2012. 14803, 14806

Arora, V. K.: The use of the aridity index to assess climate change effect on annual runoff, J. Hydrol., 265, 164–177, doi:10.1016/S0022-1694(02)00101-4, 2002. 14819

Bárdossy, A.: Calibration of hydrological model parameters for ungauged catchments, Hydrol. Earth Syst. Sci., 11, 703–710, doi:10.5194/hess-11-703-2007, 2007. 14804, 14829

Beven, K.: INTERFLOW, Ursaturated Flow in Hydrologic Modeling Theory and Practice, 191–219, Springer, 1989. 14803

Beven, K.: Rainfall-runoff modelling: the primer, Wiley Chichester, Vol. 15, 2001. 14803, 14829

Beven, K.: A manifesto for the equifinality thesis, J. Hydrol., 320, 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006. 14803, 14806

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, doi:10.1016/S0022-1694(01)00421-8, 2001a. 14821

Beven, K. and Freer, J.: A dynamic TOPMODEL, Hydrol. Process., 15, 1993–2011, doi:10.1002/hyp.252, 2001b. 14805

Beven, K. and Germann, P.: Macropores and water flow in soils, Water Resour. Res., 18, 1311–1325, doi:10.1029/WR018i005p01311, 1982. 14809

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrol. Sci. J., 24, 43–69, doi:10.1080/02626667909491834, 1979. 14805

Birkel, C., Dunn, S. M., Tetzlaff, D., and Soulsby, C.: Assessing the value of high-resolution isotope tracer data in the stepwise development of a lumped conceptual rainfall-runoff model, Hydrol. Process., 24, 2335–2348, doi:10.1002/hyp.7763, 2010. 14821

Birkel, C., Tetzlaff, D., Dunn, S., and Soulsby, C.: Using time domain and geographic source tracers to conceptualize streamflow generation processes in lumped rainfall-runoff models, Water Resour. Res., 47, W02515, doi:10.1029/2010WR009547, 2011. 14804, 14831

Blöschl, G.: Scaling in hydrology, Hydrol. Process., 15, 709–711, doi:10.1002/hyp.432, 2001. 14803

Blume, T., Zehe, E., and Bronstert, A.: Investigation of runoff generation in a pristine, poorly gauged catchment in the Chilean Andes II: qualitative and quantitative use of tracers at three spatial scales, Hydrol. Process., 22, 3676–3688, doi:10.1002/hyp.6970, 2008. 14809

Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, Water Resour. Res., 36, 3663–3674, doi:10.1029/2000WR900207, 2000. 14804

Boyle, D. P., Gupta, H. V., Sorooshian, S., Koren, V., Zhang, Z., and Smith, M.: Toward improved streamflow forecasts: value of semidistributed modeling, Water Resour. Res., 37, 2749–2759, doi:10.1029/2000WR000207, 2001. 14804

Breuer, L., Eckhardt, K., and Frede, H.-G.: Plant parameter values for models in temperate climates, Ecol. Model., 169, 237–293, doi:10.1016/S0304-3800(03)00274-6, 2003. 14846

Bulygina, N. and Gupta, H.: How Bayesian data assimilation can be used to estimate the mathematical structure of a model, Stoch. Environ. Res. Risk Assess., 24, 925–937, doi:10.1007/s00477-010-0387-y, 2010. 14804

Capell, R., Tetzlaff, D., and Soulsby, C.: Can time domain and source area tracers reduce uncertainty in rainfall-runoff models in larger heterogeneous catchments?, Water Resour. Res., 48, W09544, doi:10.1029/2011WR011543, 2012. 14831

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): a modular framework to diagnose differences between hydrological models, Water Resour. Res., 44, W00B02, doi:10.1029/2007WR006735, 2008. 14810

Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-van Meerveld, H. J., Peters, N. E., and Freer, J. E.: Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale, Hydrol. Process., 23, 311–319, doi:10.1002/hyp.7154, 2009. 14809

Clark, M. P., Kavetski, D., and Fenicia, F.: Pursuing the method of multiple working hypotheses for hydrological modeling, Water Resour. Res., 47, W09301, doi:10.1029/2010WR009827, 2011. 14804

Detty, J. M. and McGuire, K. J.: Topographic controls on shallow groundwater dynamics: implications of hydrologic connectivity between hillslopes and riparian zones in a till mantled catchment, Hydrol. Process., 24, 2222–2236, doi:10.1002/hyp.7656, 2010. 14805

Drogue, G., Pfister, L., Leviandier, T., Humbert, J., Hoffmann, L., Idrissi, A. E., and Iffly, J.-F.: Using 3-D dynamic cartography and hydrological modelling for linear streamflow mapping, Comput. Geosci., 28, 981–994, doi:10.1016/S0098-3004(02)00028-6, 2002. 14808

Duan, Q., Schaake, J., Andréassian, V., Franks, S., Goteti, G., Gupta, H., Gusev, Y., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., and Wood, E.: Model Parameter Estimation Experiment (MOPEX): an overview of science strategy and major results from the second and third workshops, J. Hydrol., 320, 3–17, doi:10.1016/j.jhydrol.2005.07.031, 2006. 14807, 14819, 14829

Dunn, S. M., Bacon, J. R., Soulsby, C., Tetzlaff, D., Stutter, M. I., Waldron, S., and Malcolm, I. A.: Interpretation of homogeneity in ^{18}O signatures of stream water in a nested sub-catchment system in north-east Scotland, Hydrol. Process., 22, 4767–4782, doi:10.1002/hyp.7088, 2008. 14804

Efstratiadis, A., Nalbantis, I., Koukouvinos, A., Rozos, E., and Koutsoyiannis, D.: HYDROGEIOS: a semi-distributed GIS-based hydrological model for modified river basins, Hydrol. Earth Syst. Sci., 12, 989–1006, doi:10.5194/hess-12-989-2008, 2008. 14805

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., and Savenije, H. H. G.: A framework to assess the realism of model structures using hydrological signatures, Hydrol. Earth Syst. Sci., 17, 1893–1912, doi:10.5194/hess-17-1893-2013, 2013. 14803, 14804, 14819, 14826

Fang, X., Pomeroy, J. W., Ellis, C. R., MacDonald, M. K., DeBeer, C. M., and Brown, T.: Multivariable evaluation of hydrological model predictions for a headwater basin in the Canadian Rocky Mountains, Hydrol. Earth Syst. Sci., 17, 1635–1659, doi:10.5194/hess-17-1635-2013, 2013. 14830

Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Is the groundwater reservoir linear? Learning from data in hydrological modelling, Hydrol. Earth Syst. Sci., 10, 139–150, doi:10.5194/hess-10-139-2006, 2006. 14804, 14811

Fenicia, F., McDonnell, J., and Savenije, H. H. G.: Learning from model improvement: on the contribution of complementary data to process understanding, Water Resour. Res., 44, W06419, doi:10.1029/2007WR006386, 2008a. 14804, 14808

Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Understanding catchment behavior through stepwise model concept improvement, Water Resour. Res., 44, W01402, doi:10.1029/2006WR005563, 2008b. 14812, 14831

Fenicia, F., Kavetski, D., and Savenije, H. H. G.: Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, Water Resour. Res., 47, W11510, doi:10.1029/2010WR010174, 2011. 14809

Fenicia, F., Kavetski, D., Savenije, H. H. G., Clark, M. P., Schoups, G., Pfister, L., and Freer, J.: Catchment properties, function, and conceptual model representation: is there a correspondence?, Hydrol. Process., doi:10.1002/hyp.9726, in press, 2013. 14810

Flügel, W.-A.: Delineating hydrological response units by geographical information system analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the River Bröl, Germany, Hydrol. Process., 9, 423–436, doi:10.1002/hyp.3360090313, 1995. 14805

Freer, J., McDonnell, J. J., Beven, K. J., Peters, N. E., Burns, D. A., Hooper, R. P., Aulenbach, B., and Kendall, C.: The role of bedrock topography on subsurface storm flow, Water Resour. Res., 38, 1269, doi:10.1029/2001WR000872, 2002. 14809

Freer, J., Beven, K., and Peters, N.: Multivariate seasonal period model rejection within the generalised likelihood uncertainty estimation procedure, Water Sci. Appl., 6, 69–87, doi:10.1029/WS006p0069, 2003. 14803

Freer, J., McMillan, H., McDonnell, J., and Beven, K.: Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures, J. Hydrol., 291, 254–277, doi:10.1016/j.jhydrol.2003.12.037, 2004. 14804

Gascuel-Odoux, C., Aurousseau, P., Durand, P., Ruiz, L., and Molenat, J.: The role of climate on inter-annual variation in stream nitrate fluxes and concentrations, Sci. Total. Environ., 408, 5657–5666, doi:10.1016/j.scitotenv.2009.05.003, 2010. 14805

Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H. G.: Hydrological landscape classification: investigating the performance of HAND based landscape classifications in a central European meso-scale catchment, Hydrol. Earth Syst. Sci., 15, 3275–3291, doi:10.5194/hess-15-3275-2011, 2011. 14806, 14808, 14810, 14813, 14814, 14850

Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H. G.: An approach to identify time consistent model parameters: sub-period calibration, Hydrol. Earth Syst. Sci., 17, 149–161, doi:10.5194/hess-17-149-2013, 2013a. 14803, 14806

Gharari, S., Shafiei, M., Hrachowitz, M., Fenicia, F., Gupta, H. V., and Savenije, H. H. G.: A strategy for “constraint-based” parameter specification for environmental models, Hydrol. Earth Syst. Sci. Discuss., 10, 14857–14871, doi:10.5194/hessd-10-14857-2013, 2013b. 14822

Grayson, R. and Blöschl, G.: Spatial Patterns in Catchment Hydrology: Observations and Modelling., Chap. 14 Summary of Pattern Comparison and Concluding Remarks, Cambridge University Press, Cambridge, 355–367, 2000. 14805

Grayson, R. B., Moore, I. D., and McMahon, T. A.: Physically based hydrologic modeling: 1. a terrain-based model for investigative purposes, Water Resour. Res., 28, 2639–2658, doi:10.1029/92WR01258, 1992. 14803

Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information, Water Resour. Res., 34, 751–763, doi:10.1029/97WR03495, 1998. 14804

Gupta, H. V., Wagener, T., and Liu, Y.: Reconciling theory with observations: elements of a diagnostic approach to model evaluation, Hydrol. Process., 22, 3802–3813, doi:10.1002/hyp.6989, 2008. 14804, 14806, 14814

Gupta, V. K. and Sorooshian, S.: Uniqueness and observability of conceptual rainfall-runoff model parameters: the percolation process examined, Water Resour. Res., 19, 269–276, doi:10.1029/WR019i001p00269, 1983. 14806

Hamon, W. R.: Estimating potential evapotranspiration, J. Hydraul. Div., 87, 107–120, 1961. 14808

Hewlett, J. D.: Soil Moisture as a Source of Base Flow from Steep Mountain Watersheds, US Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, 1961. 14809

Hrachowitz, M., Soulsby, C., D., T., Dawson, J. J. C., and Malcolm, I. A.: Regionalization of transit time estimates in montane catchments by integrating landscape controls, Water Resour. Res., 45, W05421, doi:10.1029/2008WR007496, 2009. 14805

Hrachowitz, M., Savenije, H., Bogaard, T. A., Tetzlaff, D., and Soulsby, C.: What can flux tracking teach us about water age distribution patterns and their temporal dynamics?, Hydrol. Earth Syst. Sci., 17, 533–564, doi:10.5194/hess-17-533-2013, 2013a. 14803, 14804, 14831

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB) – a review, Hydrolog. Sci. J., 58, 1–58, doi:10.1080/02626667.2013.803183, 2013b. 14806

Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., and Marshall, L. A.: Hydrologic connectivity between landscapes and streams: transferring reach- and plot-scale understanding to the catchment scale, Water Resour. Res., 45, W04428, doi:10.1029/2008WR007225, 2009. 14805

Kapangaziwiri, E., Hughes, D., and Wagener, T.: Incorporating uncertainty in hydrological predictions for gauged and ungauged basins in southern Africa, Hydrolog. Sci. J., 57, 1000–1019, doi:10.1080/02626667.2012.690881, 2012. 14804, 14829

Kavetski, D., Fenicia, F., and Clark, M. P.: Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: Insights from an experimental catchment, Water Resour. Res., 47, W05501, doi:10.1029/2010WR009525, 2011. 14809

Khu, S. T., Madsen, H., and de Pierro, F.: Incorporating multiple observations for distributed hydrologic model calibration: an approach using a multi-objective evolutionary algorithm and clustering, Adv. Water Resour., 31, 1387–1398, doi:10.1016/j.advwatres.2008.07.011, 2008. 14804

Kirchner, J. W.: Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology, Water Resour. Res., 42, W03S04, doi:10.1029/2005WR004362, 2006. 14803

Klemeš, V.: Operational testing of hydrological simulation models, Hydrolog. Sci. J., 31, 13–24, doi:10.1080/02626668609491024, 1986. 14826

Kling, H. and Gupta, H.: On the development of regionalization relationships for lumped watershed models: the impact of ignoring sub-basin scale variability, J. Hydrol., 373, 337–351, doi:10.1016/j.jhydrol.2009.04.031, 2009. 14803, 14804

Knudsen, J., Thomsen, A., and Refsgaard, J. C.: WATBAL A semi-distributed, physically based hydrological modelling system, Nord. Hydrol., 17, 347–362, doi:10.2166/nh.1986.026, 1986. 14805, 14813

Koren, V., Smith, M., Wang, D., and Zhang, Z.: Use of soil property data in the derivation of conceptual rainfall-runoff model parameters, in: 15th Conference on Hydrology, Long Beach, American Meteorological Society, Paper, vol. 2, 2000. 14807

Koren, V., Smith, M., and Duan, Q.: Use of a Priori Parameter Estimates in the Derivation of Spatially, 2003. 14807, 14829

Krcho, J.: Modelling of georelief and its geometrical structure using DTM: positional and numerical accuracy, Q111 Vydavatel Stvo, 2001. 14805

Kumar, R., Samaniego, L., and Attinger, S.: The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics, J. Hydrol., 392, 54–69, doi:10.1016/j.jhydrol.2010.07.047, 2010. 14804, 14805, 14830

Kumar, R., Samaniego, L., and Attinger, S.: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resour. Res., 491, 360–379, doi:10.1029/2012WR012195, 2013. 14804, 14830

Kuzmin, V., Seo, D.-J., and Koren, V.: Fast and efficient optimization of hydrologic model parameters using a priori estimates and stepwise line search, J. Hydrol., 353, 109–128, doi:10.1016/j.jhydrol.2008.02.001, 2008. 14807

Laaha, G., Skøien, J., and Blöschl, G.: Spatial prediction on river networks: comparison of top-kriging with regional regression, Hydrol. Process., doi:10.1002/hyp.9578, 2012. 14829

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., and Arheimer, B.: Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales, Hydrol. Res, 41, 295–319, doi:10.2166/nh.2010.007, 2010. 14805

Madsen, H.: Automatic calibration of a conceptual rainfall–runoff model using multiple objectives, J. Hydrol., 235, 276–288, doi:10.1016/S0022-1694(00)00279-1, 2000. 14804

Matgen, P., Fenicia, F., Heitz, S., Plaza, D., de Keyser, R., Pauwels, V. R., Wagner, W., and Savenije, H.: Can ASCAT-derived soil wetness indices reduce predictive uncertainty in well-gauged areas? A comparison with in situ observed soil moisture in an assimilation application, Adv. Water Resour., 44, 49–65, doi:10.1016/j.advwatres.2012.03.022, 2012. 14804

McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M. L., Selker, J., and Weiler, M.: Moving beyond heterogeneity and process complexity: a new vision for watershed hydrology, Water Resour. Res., 43, W07301, doi:10.1029/2006WR005467, 2007. 14803

McGlynn, B. L. and McDonnell, J. J.: Quantifying the relative contributions of riparian and hillslopezones to catchment runoff, Water Resour. Res., 39, 1310, doi:10.1029/2003WR002091, 2003. 14805

McGlynn, B. L., McDonnell, J. J., Seibert, J., and Kendall, C.: Scale effects on headwater catchment runoff timing, flow sources, and groundwater-streamflow relations, Water Resour. Res., 40, W07504, doi:10.1029/2003WR002494, 2004. 14809

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M., and Seibert, J.: The role of topography on catchment-scale water residence time, Water Resour. Res., 41, W05002, doi:10.1029/2004WR003657, 2005. 14805

McMillan, H. K., Clark, M. P., Bowden, W. B., Duncan, M., and Woods, R. A.: Hydrological field data from a modeller's perspective: Part 1. Diagnostic tests for model structure, Hydrol. Process., 25, 511–522, doi:10.1002/hyp.7841, 2011. 14804

McNamara, J. P., Chandler, D., Seyfried, M., and Achet, S.: Soil moisture states, lateral flow, and streamflow generation in a semi-arid, snowmelt-driven catchment, Hydrol. Process., 19, 4023–4038, doi:10.1002/hyp.5869, 2005. 14809

Merz, R., and Blöschl, G.: Regionalisation of catchment model parameters, J. Hydrol., 287, 95–123, 2004. 14804, 14810

Milne, G.: Some suggested units of classification and mapping particularly for East African soils, Soil Research, 4, 183–198, 1935. 14805

Molénat, J., Gascuel-Odoux, C., Davy, P., and Durand, P.: How to model shallow water-table depth variations: the case of the Kervidy-Naizin catchment, France, Hydrol. Process., 19, 901–920, doi:10.1002/hyp.5546, 2005. 14809, 14821

Nalbantis, I., Efstratiadis, A., Rozos, E., Kopsiafti, M., and Koutsoyiannis, D.: Holistic versus monomeric strategies for hydrological modelling of human-modified hydrosystems, Hydrol. Earth Syst. Sci., 15, 743–758, doi:10.5194/hess-15-743-2011, 2011. 14805

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models Part I: A discussion of principles, J. Hydrol., 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970. 14822

Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., Waterloo, M., and Saleska, S.: Height above the nearest drainage, a hydrologically relevant new terrain model, J. Hydrol., 404, 13–29, 2011. 14805, 14810

Ott, B. and Uhlenbrook, S.: Quantifying the impact of land-use changes at the event and seasonal time scale using a process-oriented catchment model, Hydrol. Earth Syst. Sci., 8, 62–78, doi:10.5194/hess-8-62-2004, 2004. 14830

Oudin, L., Andréassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity, physical similarity, regression and ungaged catchments: a comparison of regionalization approaches based on 913 French catchments, Water Resour. Res., 44, W03413, doi:10.1029/2007WR006240, 2008. 14829

Parajka, J. and Blöschl, G.: Spatio-temporal combination of MODIS images – potential for snow cover mapping, Water Resour. Res., 44, W03406, doi:10.1029/2007WR006204, 2008. 14804

Parajka, J., Merz, R., and Blöschl, G.: Uncertainty and multiple objective calibration in regional water balance modelling: case study in 320 Austrian catchments, Hydrol. Process., 21, 435–446, doi:10.1002/hyp.6253, 2007. 14829

Park, S. and van de Giesen, N.: Soil-landscape delineation to define spatial sampling domains for hillslope hydrology, J. Hydrol., 295, 28–46, doi:10.1016/j.jhydrol.2004.02.022, 2004. 14805

Perrin, C., Andréassian, V., Serna, C. R., Mathevet, T., and Moine, N. L.: Discrete parameterization of hydrological models: evaluating the use of parameter sets libraries over 900 catchments, Water Resour. Res., 44, W08447, doi:10.1029/2007WR006579, 2008. 14804, 14810

Pike, J.: The estimation of annual run-off from meteorological data in a tropical climate, J. Hydrol., 2, 116–123, doi:10.1016/0022-1694(64)90022-8, 1964. 14819

Pokhrel, P., Gupta, H. V., and Wagener, T.: A spatial regularization approach to parameter estimation for a distributed watershed model, Water Resour. Res., 44, W12419, doi:10.1029/2007WR006615, 2008. 14806, 14830

Pokhrel, P., Yilmaz, K. K., and Gupta, H. V.: Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures, J. Hydrol., 418–419, 49–60, doi:10.1016/j.jhydrol.2008.12.004, 2012. 14830

Pomeroy, J. W., Gray, D. M., Brown, T., Hedstrom, N. R., Quinton, W. L., Granger, R. J., and Carey, S. K.: The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence, Hydrol. Process., 21, 2650–2667, doi:10.1002/hyp.6787, 2007. 14803, 14805

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., and Participants, D.: Overall distributed model intercomparison project results, J. Hydrol., 298, 27–60, doi:10.1016/j.jhydrol.2004.03.031, 2004. 14803

Refsgaard, J. C. and Knudsen, J.: Operational validation and intercomparison of different types of hydrological models, Water Resour. Res., 32, 2189–2202, doi:10.1029/96WR00896, 1996. 14803

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., and Waterloo, M. J.: HAND, a new terrain descriptor using SRTM-DEM: mapping terra-firme rainforest environments in Amazonia, Remote Sens. Environ., 112, 3469–3481, doi:10.1016/j.rse.2008.03.018, 2008. 14805, 14808, 14810

Rouhani, H., Willems, P., Wyseure, G., and Feyen, J.: Parameter estimation in semi-distributed hydrological catchment modelling using a multi-criteria objective function, Hydrol. Process., 21, 2998–3008, doi:10.1002/hyp.6527, 2007. 14804

Samaniego, L., Bárdossy, A., and Kumar, R.: Streamflow prediction in ungauged catchments using copula-based dissimilarity measures, Water Resour. Res., 46, W02506, doi:10.1029/2008WR007695, 2010. 14804

Samaniego, L., Rohini, K., and Matthias, Z.: Implications of parameter uncertainty on soil moisture drought analysis in Germany, J. Hydrometeorol., 14, 47–68, doi:10.1175/JHM-D-12-075.1, 2013. 14804

Savenije, H. H. G.: HESS Opinions "Topography driven conceptual modelling (FLEX-Topo)", Hydrol. Earth Syst. Sci., 14, 2681–2692, doi:10.5194/hess-14-2681-2010, 2010. 14806, 14808, 14809, 14810

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A., and Carrillo, G.: Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrol. Earth Syst. Sci., 15, 2895–2911, doi:10.5194/hess-15-2895-2011, 2011. 14819

Scherrer, S. and Naef, F.: A decision scheme to indicate dominant hydrological flow processes on temperate grassland, Hydrol. Process., 17, 391–401, doi:10.1002/hyp.1131, 2003. 14805

Scherrer, S., Naef, F., Faeh, A. O., and Cordery, I.: Formation of runoff at the hillslope scale during intense precipitation, Hydrol. Earth Syst. Sci., 11, 907–922, doi:10.5194/hess-11-907-2007, 2007. 14805

Schmocker-Fackel, P., Naef, F., and Scherrer, S.: Identifying runoff processes on the plot and catchment scale, Hydrol. Earth Syst. Sci., 11, 891–906, doi:10.5194/hess-11-891-2007, 2007. 14805

Seibert, J.: Reliability of model predictions outside calibration conditions, Nord. Hydrol., 34, 477–492, 2003. 14803, 14831

Seibert, J. and Beven, K. J.: Gauging the ungauged basin: how many discharge measurements are needed?, Hydrol. Earth Syst. Sci., 13, 883–892, doi:10.5194/hess-13-883-2009, 2009. 14810

Seibert, J., Bishop, K., Rodhe, A., and McDonnell, J. J.: Groundwater dynamics along a hillslope: a test of the steady state hypothesis, Water Resour. Res., 39, 1014, doi:10.1029/2002WR001404, 2003. 14804, 14805, 14809, 14821

Semenova, O., Lebedeva, L., and Vinogradov, Y.: Simulation of subsurface heat and water dynamics, and runoff generation in mountainous permafrost conditions, in the Upper Kolyma River basin, Russia, Hydrogeol. J., 21, 107–119, doi:10.1007/s10040-012-0936-1, 2013. 14830

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀▌ ▌▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Sidle, R. C., Noguchi, S., Tsuboyama, Y., and Laursen, K.: A conceptual model of preferential flow systems in forested hillslopes: evidence of self-organization, Hydrol. Process., 15, 1675–1692, doi:10.1002/hyp.233, 2001. 14809

Sivapalan, M.: Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the Catchment Scale, John Wiley & Sons, Ltd, doi:10.1002/0470848944.hsa012, 2006. 14803

Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R.: Downward approach to hydrologic prediction, Hydrol. Process., 17, 2101–2111, doi:10.1002/hyp.1425, 2003. 14808

Son, K. and Sivapalan, M.: Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data, Water Resour. Res., 43, W01415, doi:10.1029/2006WR005032, 2007. 14804

Spence, C.: A paradigm shift in hydrology: storage thresholds across scales influence catchment runoff generation, Geogr. Compass, 4, 819–833, doi:10.1111/j.1749-8198.2010.00341.x, 2010. 14809

Sutanudjaja, E., de Jong, S., van Geer, F., and Bierkens, M.: Using {ERS} spaceborne microwave soil moisture observations to predict groundwater head in space and time, Remote Sens. Environ., 138, 172–188, doi:10.1016/j.rse.2013.07.022, 2013. 14804

Szilagyi, J., Rundquist, D. C., Gosselin, D. C., Parlange, M. B.: NDVI relationship to monthly evaporation, Geophys. Res. Lett., 25, 1753–1756, doi:10.1029/98GL01176, 1998. 14818

Tromp-van Meerveld, H. J., and McDonnell, J. J.: Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis, Water Resour. Res., 42, W02411, doi:10.1029/2004WR003800, 2006a. 14809

Tromp-van Meerveld, H. J., and McDonnell, J. J.: Threshold relations in subsurface stormflow: 1. A 147-storm analysis of the Panola hillslope, Water Resour. Res., 42, W02410, doi:10.1029/2004WR003778, 2006b. 14809

Turc, L.: Le bilan d'eau des sols. Relation entre la precipitation, l'evaporation et l'ecoulement, Ann. Agron. 5, 1954. 14819

Uhlenbrook, S., Roser, S., and Tilch, N.: Hydrological process representation at the mesoscale: the potential of a distributed, conceptual catchment model, J. Hydrol., 291, 278–296, doi:10.1016/j.jhydrol.2003.12.038, 2004. 14805

Vaché, K., and McDonnell, J.: A process-based rejectionist framework for evaluating catchment runoff model structure, Water Resour. Res., 42, W02409, doi:10.1029/2005WR004247, 2006. 14804

Viglione, A., Parajka, J., Rogger, M., Salinas, J. L., Laaha, G., Sivapalan, M., and Blöschl, G.: Comparative assessment of predictions in ungauged basins – Part 3: Runoff signatures in Austria, Hydrol. Earth Syst. Sci., 17, 2263–2279, doi:10.5194/hess-17-2263-2013, 2013. 14804

Vinogradov, Y. B., Semenova, O. M., and Vinogradova, T. A.: An approach to the scaling problem in hydrological modelling: the deterministic modelling hydrological system, Hydrol. Process., 25, 1055–1073, doi:10.1002/hyp.7901, 2011. 14830

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resour. Res., 39, 1214, doi:10.1029/2002WR001746, 2003. 14821

Wagener, T. and Montanari, A.: Convergence of approaches toward reducing uncertainty in predictions in ungauged basins, Water Resour. Res., 47, W06301, doi:10.1029/2010WR009469, 2011. 14804, 14806

Wagener, T. and Wheater, H. S.: Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty, J. Hydrol., 320, 132–154, doi:10.1016/j.jhydrol.2005.07.015, 2006. 14829

Weiler, M., McGlynn, B., McGuire, K., and McDonnell, J.: How does rainfall become runoff? A combined tracer and runoff transfer function approach, Water Resour. Res., 39, 1315, doi:10.1029/2003WR002331, 2003. 14809

Winsemius, H. C., Savenije, H. H. G., and Bastiaanssen, W. G. M.: Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins?, Hydrol. Earth Syst. Sci., 12, 1403–1413, doi:10.5194/hess-12-1403-2008, 2008. 14804

Winsemius, H. C., Schaefli, B., Montanari, A., and Savenije, H. H. G.: On the calibration of hydrological models in ungauged basins: a framework for integrating hard and soft hydrological information, Water Resour. Res., 45, W12422, doi:10.1029/2009WR007706, 2009. 14804, 14819, 14829

Winter, T. C.: The concept of hydrologic landscapes, J. Am. Water Resour. As., 37, 335–349, doi:10.1111/j.1752-1688.2001.tb00973.x, 2001. 14805

Wolock, D. M., Winter, T. C., and McMahon, G.: Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses, Environ. Manage., 34, S71–S88, doi:10.1007/s00267-003-5077-9, 2004. 14805

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀| |▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, Adv. Water Resour., 30, 1756–1774, doi:10.1016/j.advwatres.2007.01.005, 2007. 14804, 14829

Yilmaz, K. K., Gupta, H. V., and Wagener, T.: A process-based diagnostic approach to model evaluation: application to the NWS distributed hydrologic model, Water Resour. Res., 44, W09417, doi:10.1029/2007WR006716, 2008. 14806

Zehe, E. and Sivapalan, M.: Threshold behaviour in hydrological systems as (human) geo-ecosystems: manifestations, controls, implications, Hydrol. Earth Syst. Sci., 13, 1273–1297, doi:10.5194/hess-13-1273-2009, 2009. 14809

Zhang, Z., Wagener, T., Reed, P., and Bhushan, R.: Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization, Water Resour. Res., 44, W00B04, doi:10.1029/2008WR006833, 2008. 14804, 14829

Table 1. Prior parameter distributions for the three model set-ups.

			FLEX^{A}	FLEX^{B}		FLEX^{C}		
	Unit			wetland	remainder	wetland	hillslope	plateau
$I_{\text{max}}{}^{*}$	mm	Interception storage for forest			2–5			
		Interception storage for grassland and pasture			1–3			
$S_{\text{U,max}}$	mm	Maximum unsaturated storage	0–500	0–100	0–500	0–100	0–500	0–500
β	–	Soil moisture distribution power	0–5	0–5	0–5	0–5	0–5	0–5
L_{P}	–	Transpiration coefficient	0.5	0.5	0.5	0.5	0.5	0.5
F_{C}	–	Field capacity	0–0.3	0	0–0.3	0	0–0.3	0–0.3
D	–	Partitioning fast and slow reservoir	0–1	0	0–1	0	0–1	1
C	$\text{mm}(3\,\text{h})^{-1}$	Maximum capillary rise rate	0	0–0.3	–	0–0.3	–	–
P_{per}	$\text{mm}(3\,\text{h})^{-1}$	Maximum percolation rate	0–0.5	0–0.5	0–0.5	0–0.5	0	0–0.5
N_{lagf}	3 h	Lag time for flux to fast reservoir	1–7	1–3	1–5	1–3	1–5	–
N_{lags}	3 h	Lag time for preferential recharge	1–7	–	1–7	–	1–7	1–7
K_{F}	$(3\,\text{h})^{-1}$	Fast reservoir coefficient	0–1	0–1	0–1	0–1	0–1	–
K_{S}	$(3\,\text{h})^{-1}$	Slow reservoir coefficient	0.005–0.05	0.005–0.05		0.005–0.05		

*Inferred from Breuer et al. (2003).

Table 2. Water balance and constitutive equations used in FLEX$^{\mathrm{A}}$.

Reservoir	Water balance equations	Constitutive relations
Interception reservoir	$\frac{\mathrm{d}S_\mathrm{I}}{\mathrm{d}t} = P - I - P_\mathrm{e}$	$I = \begin{cases} E_\mathrm{pot} & E_\mathrm{p}\mathrm{d}t < S_\mathrm{I} \\ S_\mathrm{I}/\mathrm{d}t & E_\mathrm{p}\mathrm{d}t \geq S_\mathrm{I} \end{cases}$ $P_\mathrm{e} = \begin{cases} 0 & S_\mathrm{I} < I_\mathrm{max} \\ (S_\mathrm{I} - I_\mathrm{max})/\mathrm{d}t & S_\mathrm{I} \geq I_\mathrm{max} \end{cases}$
Unsaturated reservoir	$\frac{\mathrm{d}S_\mathrm{U}}{\mathrm{d}t} = R_\mathrm{u} - T - R_\mathrm{P} + R_\mathrm{C}$	$R_\mathrm{u} = C_\mathrm{r}P_\mathrm{e}$ $C_\mathrm{r} = \begin{cases} 1 - \left[\frac{(S_\mathrm{U} - S_\mathrm{U,max}F_\mathrm{C})}{(S_\mathrm{U,max} - S_\mathrm{U,max}F_\mathrm{C})}\right]^\beta & S_\mathrm{U} \geq S_\mathrm{U,max} \\ 1 & S_\mathrm{U} < S_\mathrm{U,max} \end{cases}$ $T = K_\mathrm{T}(E_\mathrm{pot})$ $K_\mathrm{T} = \begin{cases} \left[\frac{S_\mathrm{U}}{S_\mathrm{U,max}L_\mathrm{p}}\right] & S_\mathrm{U} < S_\mathrm{U,max}L_\mathrm{p} \\ 1 & S_\mathrm{U} \geq S_\mathrm{U,max}L_\mathrm{p} \end{cases}$ $R_\mathrm{p} = [S_\mathrm{U}/S_\mathrm{U,max}]P_\mathrm{per}$ $R_\mathrm{C} = [1 - (S_\mathrm{U}/S_\mathrm{U,max})]C$
Fast reservoir	$\frac{\mathrm{d}S_\mathrm{F}}{\mathrm{d}t} = R_\mathrm{F,lag} - Q_\mathrm{F}$	$R_\mathrm{F} = (1 - D)(1 - C_\mathrm{r})P_\mathrm{e}$ $R_\mathrm{F,lag} = R_\mathrm{F} * N_\mathrm{lagf}$ $Q_\mathrm{F} = S_\mathrm{F}/K_\mathrm{F}$
Slow reservoir	$\frac{\mathrm{d}S_\mathrm{S}}{\mathrm{d}t} = R_\mathrm{S,lag} - Q_\mathrm{S}$	$R_\mathrm{S} = D(1 - C_\mathrm{r})P_\mathrm{e}$ $R_\mathrm{S,lag} = R_\mathrm{S} * N_\mathrm{lags}$ $Q_\mathrm{S} = S_\mathrm{S}/K_\mathrm{S}$

$*$ is the convolution operator.

Table 3. The median model performances (in brackets their corresponding 95 % uncertainty intervals) and the area spanned by the 95 % uncertainty interval of hydrograph derived from uncalibrated parameter sets which satisfy the complete set of constraints for the three model set-ups FLEX^A, FLEX^B and FLEX^C, for the three modeling objectives (E_{NS}, $E_{NS,log}$, $E_{NS,FDC}$) in the calibration (2006–2007) and validation (2008–2009) periods.

		E_{NS}	$E_{NS,log}$	$E_{NS,FDC}$	95 % uncertainty area [mm]
FLEX^A	Calibration	0.31 [0.15 0.51]	−0.48[−8.29 0.44]	0.66 [0.45 0.88]	970
FLEX^A	Validation	0.32 [−0.13 0.48]	−1.01 [−9.38 0.57]	0.64 [0.34 0.91]	954
FLEX^B	Calibration	0.51 [0.14 0.74]	0.30 [−0.22 0.68]	0.87 [0.68 0.96]	918
FLEX^B	Validation	0.51 [0.12 0.70]	0.50 [−0.77 0.75]	0.88 [0.66 0.97]	974
FLEX^C	Calibration	0.65 [0.25 0.80]	0.55 [−1.60 0.73]	0.97 [0.85 0.99]	625
FLEX^C	Validation	0.64 [0.20 0.78]	0.45 [−2.24 0.77]	0.97 [0.89 0.99]	667

Table 4. The median model performances (in brackets their corresponding Pareto uncertainty intervals) and the area spanned by uncertainty interval of hydrograph derived from the Pareto optimal solutions of the constrained and calibrated model set-ups FLEX^A, FLEX^B and FLEX^C for the three modeling objectives (E_{NS}, $E_{NS,log}$, $E_{NS,FDC}$) in the calibration and validation periods.

		E_{NS}	$E_{NS,log}$	$E_{NS,FDC}$	95 % uncertainty area [mm]
FLEX^A	Calibration	0.89 [0.84 0.91]	0.85 [0.72 0.89]	0.99 [0.92 0.99]	196
FLEX^A	Validation	0.73 [0.70 0.76]	0.74 [0.63 0.83]	0.93 [0.85 0.96]	265
FLEX^B	Calibration	0.86 [0.82 0.89]	0.83 [0.64 0.87]	0.99 [0.98 0.99]	263
FLEX^B	Validation	0.80 [0.76 0.81]	0.79 [0.40 0.87]	0.95 [0.94 0.96]	255
FLEX^C	Calibration	0.87 [0.81 0.89]	0.74 [0.29 0.85]	0.99 [0.97 0.99]	238
FLEX^C	Validation	0.80 [0.77 0.82]	0.66 [0.07 0.86]	0.95 [0.94 0.96]	281

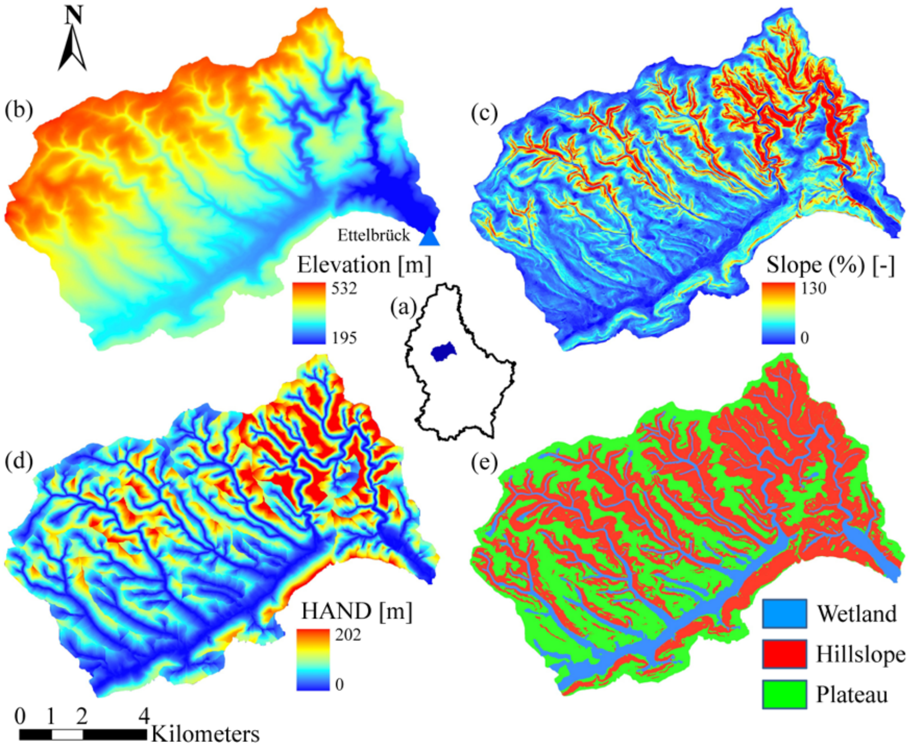

Fig. 1. **(a)** Location of the Wark catchment in the Grand Duchy of Luxembourg, **(b)** digital elevation model (DEM) of the Wark catchment with cell size of 5 m × 5 m [m], **(c)** local slopes (%) in the Wark catchment derived from a DEM with resolution of 5 m × 5 m [–], **(d)** HAND of the Wark Catchment derived from a DEM with resolution of 5 m × 5 m [m], **(e)** the classified landscape units, wetland, hillslope and plateau using the combined HAND and slope thresholds of 5 m and 11 %, respectively (from Gharari et al., 2011).

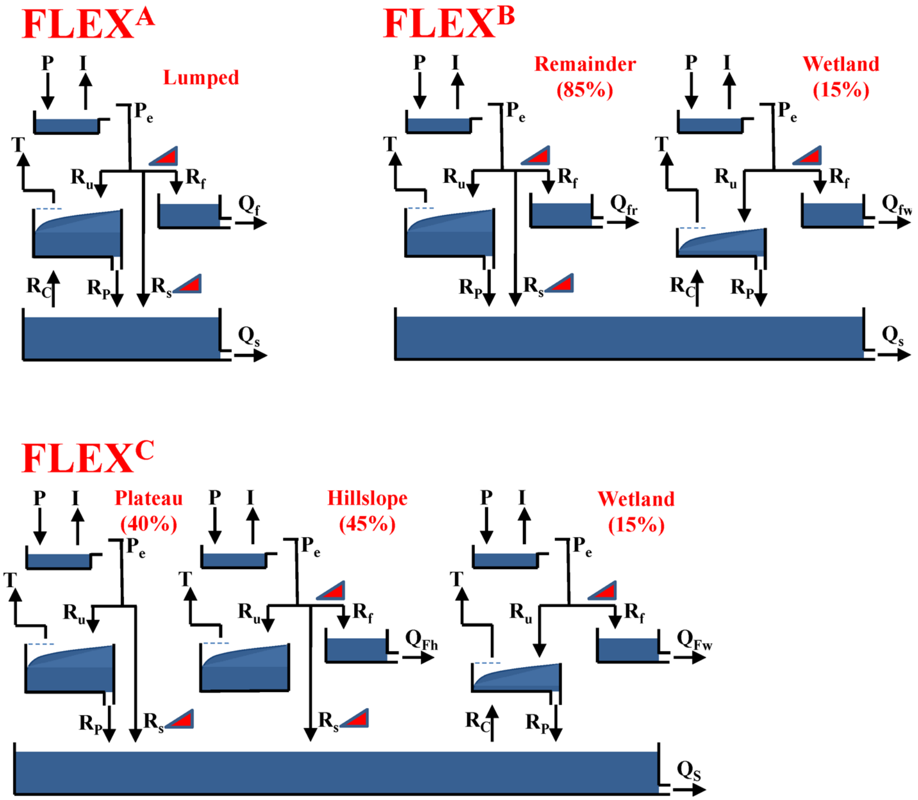

Fig. 2. The model structures for FLEX^A, FLEX^B and FLEX^C.

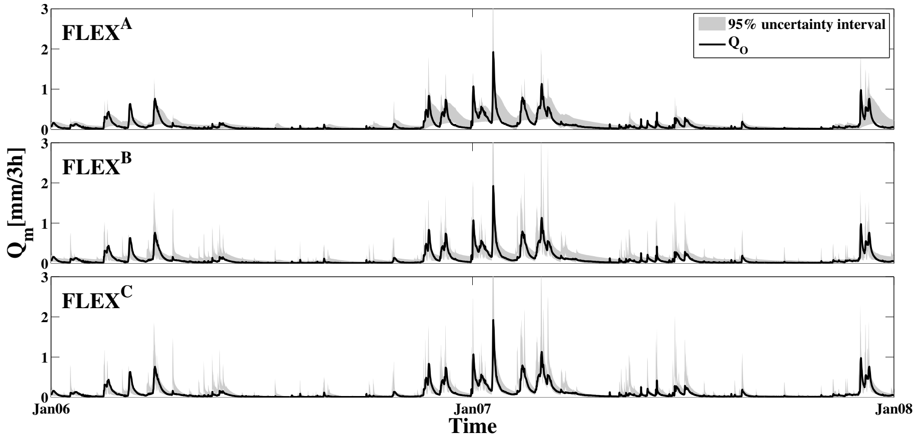

Fig. 3. The observed hydrograph and the 95 % uncertainty interval of the modeled hydrograph derived from the complete set of constrained but un-calibrated parameter sets for the three different model set-ups FLEX^A, FLEX^B and FLEX^C for the calibration (2006–2007).

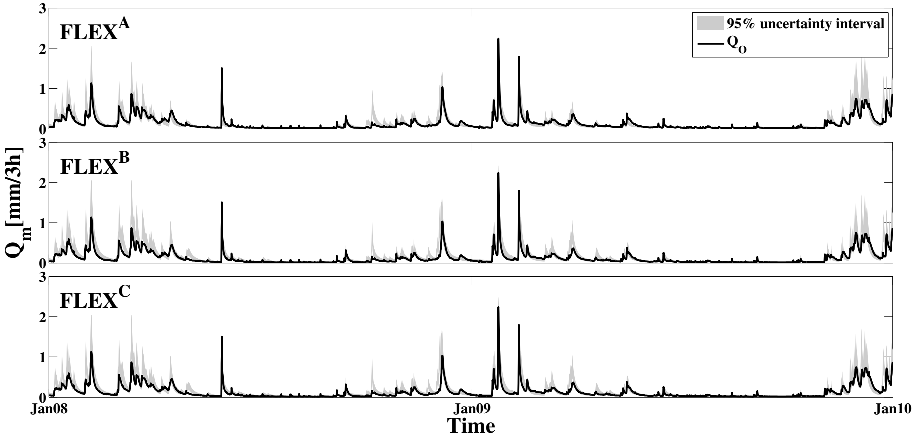

Fig. 4. The observed hydrograph and the 95 % Pareto uncertainty interval of the modeled hydrograph for constrained and calibrated parameter sets for the three different model set-ups FLEX^A, FLEX^B and FLEX^C for the validation period (2008–2009).

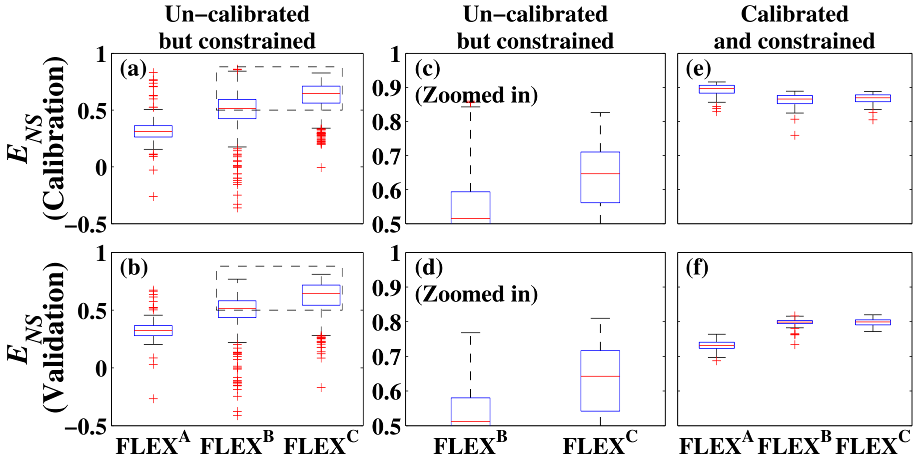

Fig. 5. Model performance (E_{NS}) based on constrained but uncalibrated **(a–d)** and constrained and calibrated **(e–f)** parameter sets for calibration (2006–2007) and validation (2008–2009) periods for the three different model set-ups FLEX$^{\mathrm{A}}$, FLEX$^{\mathrm{B}}$ and FLEX$^{\mathrm{C}}$. Note that **(c)** and **(d)** are zoom-ins of **(a)** and **(b)**.

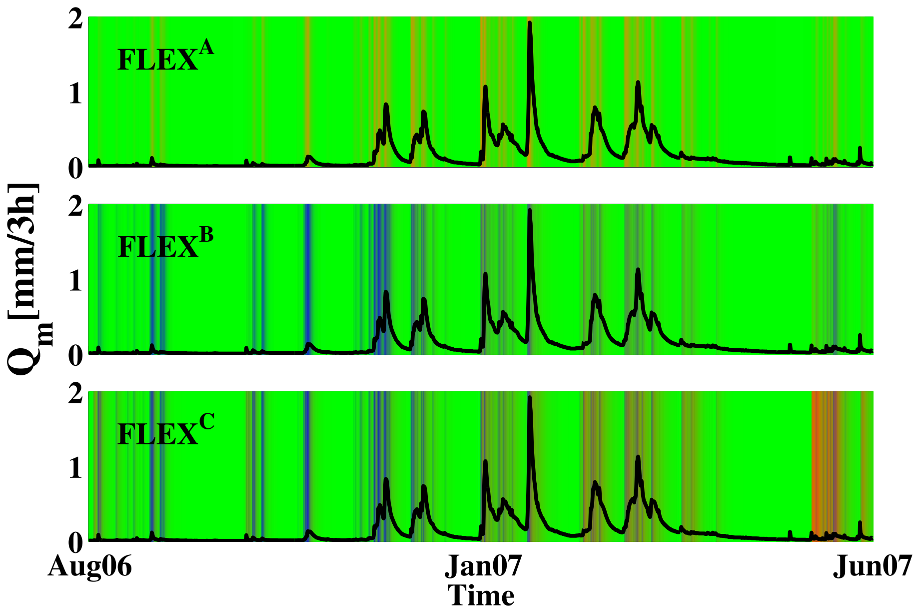

Fig. 6. The comparison between mean proportions of Pareto members for model components of the three model set-ups in part of the calibration periods (August 2006–June 2007) **(a)** FLEX^A, **(b)** FLEX^B, and **(c)** FLEX^C. The green color indicates the contribution of the slow reservoir for the three different models. The red indicates the fast component reaction from fast reservoir of FLEX^A, fasts reservoir of remainder of the catchment and fast reservoir of hillslope of FLEX^C. the blue color indicates the fast component of wetland of FLEX^B and FLEX^C. The colors are then made based on RGB color code based on the weight of the contribution of each flux to the model runoff.