

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

S. Gharari<sup>1,2</sup>, M. Hrachowitz<sup>1</sup>, F. Fenicia<sup>3,1</sup>, H. Gao<sup>1</sup>, H. H. G. Savenije<sup>1</sup>

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

3-Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

## Abstract

Conceptual environmental systems models, such as rainfall runoff models, generally rely on calibration for parameter identification. Increasing complexity of this type of models for better representation of hydrological process heterogeneity typically makes parameter identification more difficult. Although various, potentially valuable, approaches for better parameter identification were developed in the past, strategies to impose general conceptual understanding of how a catchment works into the process of parameter identification of a conceptual model has still not been fully explored. In this study we assess the effect of imposing semi-quantitative, relational expert knowledge for 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<sup>A</sup>, FLEX<sup>B</sup> and FLEX<sup>C</sup> have been developed based on these functional units. While FLEX<sup>A</sup> is a lumped representation of the study catchment, the semi-distributed formulations FLEX<sup>B</sup> and FLEX<sup>C</sup> introduce increasingly more complexity by distinguishing 2 and 3 functional units, respectively. In spite of increased complexity, FLEX<sup>B</sup> and FLEX<sup>C</sup> allow modelers to compare parameters as well as states and fluxes of their different functional units to each other. Parameter estimation was performed using semi-quantitative, relational constraints imposed onto three models structures. Increased model complexity allowed the identification of additional constraints. It was shown that a constrained but uncalibrated semi-distributed model, FLEX<sup>C</sup>, can predict runoff with similar performance to a calibrated lumped model, FLEX<sup>A</sup>. In addition, when constrained and calibrated, the semi-distributed model FLEX<sup>C</sup> exhibits not only higher performance but also lower predictive uncertainty than the calibrated, lumped FLEX<sup>A</sup> model.

## 1- Introduction:

Lumped conceptual and distributed physically based models are the two endpoints of the modeling spectrum, 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 are 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., 2013b).

1 In contrast, lumped conceptual models require less data for identifying model parameters.  
2 This advantage comes at the expense of considerable limitations. Representing system  
3 integrated processes, model structures and parameters are not directly linked to observable  
4 quantities. Their estimation therefore strongly relies on calibration. To limit parameter  
5 identifiability issues arising from calibration, these models are often oversimplified  
6 abstractions of the system. If inadequately tested they may act as “mathematical marionettes”  
7 (Kirchner, 2006), frequently resulting in good calibration performance. They may outperform  
8 more complex distributed models (e.g. Refsgaard and Knudsen 1996; Ajami et al., 2004;  
9 Reed et al., 2004), but they often fail to provide realistic representations of the underlying  
10 processes, leading to limited predictive power (e.g. Freer et al., 2003; Seibert, 2003; Kirchner,  
11 2006; Beven, 2006; Kling and Gupta, 2009; Andréassian et al., 2012; Euser et al., 2013;  
12 Gharari et al., 2013).

13 Various strategies have been suggested in the past to allow for increased model complexity  
14 and to thereby improve the physical realism of conceptual models. These strategies included  
15 the attempt to incorporate different data sources in the parameter estimation process, such as  
16 ground- and soil water dynamics (e.g. Seibert and McDonnell, 2002; Freer et al., 2004;  
17 Fenicia et al., 2008b; Matgen et al., 2012; Sutanudjaja et al., 2013), remotely sensed  
18 evaporation (e.g. Winsemius et al., 2008), snow dynamics (e.g. Parajka and Blöschl, 2008) or  
19 tracer data (e.g. Vache and McDonnell, 2006; Dunn et al., 2008; Son and Sivapalan, 2007;  
20 Birkel et al., 2011; Hrachowitz et al., 2013a). Alternatively, it was tried to extract more  
21 information from available data, for example through the development of signatures  
22 representing different aspects of the data (e.g. Gupta et al. 1998, 2008; Boyle et al., 2000,  
23 2001; Madsen 2000; Fenicia et al., 2006; Rouhani et al., 2007; Khu et al., 2008; Winsemius et  
24 al., 2009; Bulygina and Gupta, 2010; McMillan et al., 2011, Clark et al., 2011; Euser et al.,  
25 2013; Hrachowitz et al., 2013b).

26 Traditionally, parameter estimation of conceptual models relied on the availability of  
27 calibration data, which, however, are frequently not available for the time period or the spatio-  
28 temporal resolution of interest. A wide range of regionalization techniques for model  
29 parameters and hydrological signatures were thus developed to avoid calibration in such data  
30 scarce environments (e.g. Bardossy, 2007; Yadav et al., 2007; Perrin et al., 2008; Zhang et al.,  
31 2008; Kling and Gupta, 2009; Samaniego et al., 2010; Kumar et al., 2010; Wagener and  
32 Montanari, 2011; Kapangaziwiri et al., 2012, Viglione et al., 2013). However, it was for a  
33 long time considered to be challenging to identify suitable functional relationships between  
34 catchment characteristics and model parameters (e.g. Merz and Blöschl, 2004; Kling and  
35 Gupta, 2009). Only recently, Kumar et al. (2010, 2013a) showed that making use of multi-  
36 scale parameter regionalization (MPR) can yield global parameters which perform  
37 consistently over different catchment scales. In a further study they successfully transferred  
38 parameters obtained by the MPR technique to ungauged catchments in Germany and the USA  
39 (Kumar et al., 2013b). Without any further calibration the transferred global parameters were  
40 capable to adequately reproduce runoff as well as other hydrological responses of the  
41 catchments.

42 Related to the above discussed difficulties with parameterization, the frequent lack of  
43 sufficient processes heterogeneity, i.e. complexity, in conceptual models introduces further  
44 limitations to the degree of realism in these models. The concept of hydrological response  
45 unit (HRU) can be exploited as a strategy for an efficient tradeoff between model simplicity,  
46 required for adequate parameter identifiability, and a more realistic representation of  
47 hydrological processes. HRUs are units within a catchment, characterized by a different  
48 hydrological function. Individual HRUs can be represented by different model structures to  
49 account for hydrologically heterogeneous behavior based on data availability and desired  
50 resolution of process representation. This helps to enhance model realism while keeping the

1 necessary complexity and related identifiability issues comparatively low. In most cases  
2 HRUs are defined based on soil types, land cover and similar physical catchment  
3 characteristics (e.g. Knudsen et al., 1986; Flügel, 1995; Grayson and Blöschl, 2000; Krcho,  
4 2001; Winter, 2001; Scherrer and Naef, 2003; Uhlenbrook et al., 2004; Wolock et al., 2004;  
5 Pomeroy et al., 2007; Scherrer et al., 2007; Schmocker-Fackel et al., 2007; Efstratiadis et al.,  
6 2008; Lindström et al., 2010; Nalbantis et al., 2011; Kumar et al., 2010).

7 A wide range of studies also points towards the potential value of using topographical indices,  
8 which are readily available from digital elevation models (DEM) to account for process  
9 heterogeneity (e.g. McGlynn and McDonnell, 2003; Seibert et al., 2003; McGuire et al., 2005;  
10 Hrachowitz et al., 2009; Jensco et al., 2009; Detty and McGuire, 2010; Gascuel-Oudoux et al.,  
11 2010). As standard metrics of landscape organization, such as absolute elevation, slope or  
12 curvature, as used in the catena concept (Milne, 1935; Park and Van de Giesen, 2004), are  
13 often not strong enough descriptors to infer hydrological function, alternative concepts were  
14 sought. The development of derived metrics such as the *Topographic Wetness Index* (Beven  
15 and Kirkby, 1979) facilitated an important step forward, being at the core of TOPMODEL  
16 (e.g. Beven and Kirkby, 1979; Beven and Freer, 2001b), which has proven to be a valuable  
17 approach in specific environmental settings meeting the assumptions of the model. A different  
18 descriptor allowing a potentially more generally applicable and hydrologically meaningful  
19 landscape classification has recently been suggested by Rennó et al. (2008): the *Height Above*  
20 *the Nearest Drainage* (HAND). Nobre et al. (2011) showed the hydrological relevance of  
21 HAND by investigating long term groundwater behavior and land use. In a further study, this  
22 metric facilitated the identification of hydrologically similar landscape units, such as  
23 wetlands, hillslopes and plateaus in a Luxembourgish catchment (Gharari et al., 2011).

24 Explicitly invoking the co-evolution of topography, vegetation and hydrology, Savenije  
25 (2010) argued that catchments, as self-organizing systems, need to fulfill the contrasting  
26 hydrological functions of efficient drainage and sufficient water storage in order to allow, in a  
27 feedback process, topography and vegetation to develop the way they did. These distinct  
28 hydrological functions can then be associated with different landscape elements or HRUs as  
29 defined by HAND and slope, such that each HRU is represented by a model structure best  
30 representing its function in the ecosystem (cf. Savenije, 2010).

31 While HAND-based landscape classification can potentially show a way forward, it does not  
32 solve the problem arising when moving from lumped to HRU-guided, semi-distributed model  
33 formulations: multiple parallel model structures typically result in an increased number of  
34 parameters, which, when not adequately constrained, may increase equifinality and thereby  
35 predictive uncertainty (e.g. Gupta and Sorooshian, 1983; Beven, 2006; Gupta et al., 2008). In  
36 order to better satisfy the contrasting priorities of model complexity and predictive power,  
37 new strategies are sought to more efficiently utilize the modelers' understanding of the system  
38 and the frequently scarce available data for constraining the feasible model- and parameter  
39 space (e.g. Gupta et al., 2008; Wagener and Montanari, 2011; Singh and Bárdossy, 2012;  
40 Andréassian et al., 2012; Gharari et al., 2013; Hrachowitz et al., 2013b; Razavi and Tolson,  
41 2013). In contrast to earlier attempts to constrain models using multiple evaluation criteria or  
42 *a priori* information on catchment properties such as land use or soil type (e.g. Koren et al.,  
43 2008), the utility of a different and so far underexploited type of constraints, based on *a priori*  
44 understanding of the system, has been tested in this study. The concept of topography-driven  
45 conceptual modeling introduced by Savenije (2010) involves the identification of HRUs that  
46 operate in parallel. Linked to the technique of regularization (e.g. Tikhonov, 1963; Engl et al.,  
47 1996), this opens the possibility to impose semi-quantitative, expert knowledge based,  
48 relational constraints of catchment behavior on model parameters, similar to what was  
49 suggested by Pokhrel et al. (2008) and Yilmaz et al. (2008). To restrict the posterior  
50 parameter distributions, *hydrologically meaningful* relations between parallel HRUs are

1 introduced. Based on expert knowledge and expressed as relational constraints they ensure  
2 that similar processes between parallel model structures in the semi-distributed model are  
3 represented in an internally consistent way, thereby reducing the parameters' potential for  
4 compensating for errors. The advantage of this method is that there is only limited need to  
5 precisely quantify the constraints or the prior parameter distributions (e.g. Koren et al., 2000,  
6 2003; Kuzmin et al., 2008; Duan et al., 2006). This could allow for a meaningful and  
7 potentially more realistic representation of the system in which each model component is,  
8 within certain limits, forced to do what it is designed to do, rather than allowing it to  
9 compensate for data and model structural errors.

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

## 17 18 **2- Study area and data:**

19  
20 The outlined methodology will be illustrated and tested with a case study using data of the  
21 Wark catchment in the Grand Duchy of Luxembourg. The catchment has an area of 82 km<sup>2</sup>  
22 with the catchment outlet located downstream of the town of Ettelbrück at the confluence  
23 with the Alzette River (49.85° N, 6.10° E, Figure 1). With an annual mean precipitation of  
24 850 mm yr<sup>-1</sup> and an annual mean potential evaporation of 650 mm yr<sup>-1</sup> the annual mean runoff  
25 is approximately 250 mm yr<sup>-1</sup>. The geology in the northern part is dominated by schist while  
26 the southern part of the catchment is mostly underlain by sandstone and conglomerate.  
27 Hillslopes are generally characterized by forest, while plateaus and valley bottoms are mostly  
28 used as crop land and pastures, respectively. Drouge et al. (2002) quantified land use in the  
29 catchment as 4.3% urban areas, 52.7% agricultural land and 42.9% forest. In addition they  
30 reported that 61% of catchment is covered by permeable soils while the remainder is  
31 characterized by lower permeability substrate. The elevation varies between 195 to 532 m,  
32 with a mean value of 380 m. The slope of the catchment varies between 0-200%, with a mean  
33 value of 17 % (Gharari et al., 2011).

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

## 39 40 **3- FLEX-TOPO framework:**

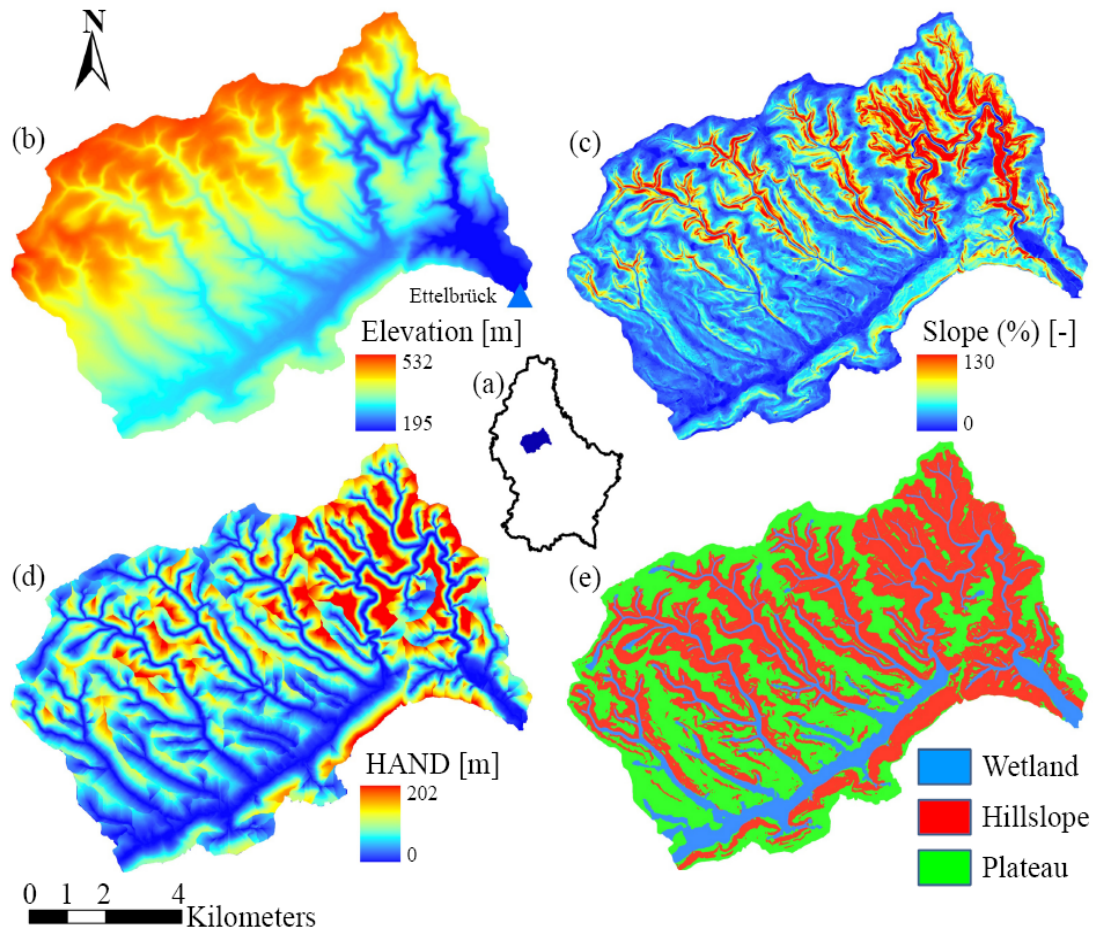
41  
42 Realizing the potential of “reading the landscape” in a systems approach (cf. Sivapalan,  
43 2003), Savenije (2010) argued that due to the co-evolution of topography, soil and vegetation,  
44 all of which define the hydrological function of a given location, an efficient, hydrologically  
45 meaningful descriptor of topography together with land use could be used to distinguish  
46 different HRUs. HAND, which can be loosely interpreted as the hydraulic head at a given  
47 location in a catchment, may be such a descriptor as it potentially allows for meaningful  
48 landscape classification (e.g. Rennó et al., 2008; Gharari et al., 2011). It was argued  
49 previously (Gharari et al., 2011) that, in Central European landscapes, HAND can efficiently  
50 distinguish between wetlands, hillslopes and plateaus. These are landscape elements that may

1 also be assumed to fulfill distinct hydrological functions (HRUs) in the study catchment  
2 (Savenije, 2010). Wetlands, located at low elevations above streams, are characterized by  
3 shallow ground water tables with limited fluctuations. Due to reduced storage capacity  
4 between ground water table and soil surface, potentially exacerbated by the relative  
5 importance of the capillary fringe, wetlands tend to be saturated, and thus connected, earlier  
6 during a rainfall event than the two other landscape elements with arguably higher storage  
7 capacity, thus frequently becoming the dominant source of storm flow during comparably dry  
8 periods (e.g. Seibert et al., 2003; McGlynn et al., 2004; Molenat et al., 2005; Blume et al.,  
9 2008; Anderson et al., 2010; Kavetski et al., 2011). The dominant runoff process in wetlands  
10 can therefore be assumed to be saturation overland flow. In contrast, forested hillslopes,  
11 landscape elements with steeper slopes than the wetlands or plateaus, require a balance  
12 between sufficient storage capacity and efficient drainage to develop and maintain the  
13 ecosystem (Savenije, 2010). A dual system combining sufficient water storage in the root  
14 zone and efficient lateral drainage through preferential flow networks, controlled by a suite of  
15 activation thresholds as frequently observed on hillslopes (e.g. Hewlett, 1961; Beven and  
16 Germann, 1982; Sidle et al., 2001; Freer et al., 2002; Weiler et al., 2003; McNamara et al.,  
17 2005; Tromp van Meerveld and McDonnell, 2006a, 2006b; Zehe and Sivapalan, 2009;  
18 Spence, 2010) can be seen as the dominant mechanism. Finally, plateaus are landforms with  
19 low to moderate slopes and comparably deep ground water tables. In absence of significant  
20 topographic gradients and due to the potentially increased unsaturated storage capacity, it can  
21 be hypothesized that the primary functions of plateaus are sub-surface storage and  
22 groundwater recharge (Savenije, 2010). Although plateaus may experience infiltration excess  
23 overland flow in specific locations, the topographical gradients may not be sufficient to  
24 generate surface runoff connected to the stream network. In the FLEX-TOPO approach the  
25 proportions of the hydrologically distinct landscape units, i.e. HRUs, in a given catchment  
26 need to be determined on the basis of topographical and land cover information. Subsequently  
27 suitable model structures and parameterizations (read constitutive functions) will be assigned  
28 to the different HRUs (Fenicia et al. 2011, Kavetski et al., 2011, Clark et al., 2009). The  
29 integrated catchment output, i.e. runoff and evaporative fluxes, can then be obtained by  
30 combining the computed proportional outputs from the individual HRUs. Note that the three  
31 landscape classes tested for suitability in this study, i.e. wetland, hillslope and plateau  
32 together with their assumed dominant runoff process are designed for the Wark catchment  
33 and are likely to be different for other environmental settings (e.g. Gao et al., 2014).

### 34 35 **3-1- Landscape classification:**

36  
37 As the objective of FLEX-TOPO is to efficiently extract and use hydrologically relevant  
38 information from worldwide readily available topographic data, i.e. DEMs, the *Height Above*  
39 *the Nearest Drainage* (HAND; Rennó et al., 2008; Nobre et al., 2011; Vannamettee et al.,  
40 2014) is a potentially powerful metric to classify landscapes into HRUs with distinct  
41 hydrological function, as discussed above. Testing a suite of HAND-based classification  
42 methods Gharari et al. (2011) found that results best matching observed landscape types could  
43 be obtained by using HAND together with the local slope. Based on a probabilistic  
44 framework to map the desired HRUs which were then compared with in-situ observations  
45 they obtained a threshold for HAND and slope of approximately 5 m and 11 % for the Wark  
46 catchment. Following that, wetlands were defined to be areas with  $HAND \leq 5$  m. Areas with  
47  $HAND > 5$  m and local slopes  $> 11$  % were classified as hillslopes, while areas with  $HAND >$   
48  $5$  m and slope  $\leq 11$  % were defined as plateaus. The HAND and slope map of the study  
49 catchment together with the classified landscape entities (wetland, hillslope and plateau) are

1 presented in Fig. 1. The proportion of the individual HRUs wetland, hillslope and plateau are  
 2 15%, 45% and 40% respectively.  
 3



4  
 5  
 6 Figure 1- (a) Location of the Wark catchment in the Grand Duchy of Luxembourg, (b) digital  
 7 elevation model (DEM) of the Wark catchment with cell size of 5 m × 5 m [m], (c) local  
 8 slopes (%) in the Wark catchment derived from a DEM with resolution of 5 m × 5 m [-], (d)  
 9 HAND of the Wark Catchment derived from a DEM with resolution of 5 m × 5 m [m], (e) the  
 10 classified landscape units wetland, hillslope and plateau using the combined HAND and slope  
 11 thresholds of 5 m and 11%, respectively (from Gharari et al., 2011).  
 12

### 13 3-2- Model setup:

14  
 15 In this study a lumped conceptual model of the Wark catchment, hereafter referred to as  
 16 FLEX<sup>A</sup>, is used as a benchmark since similar lumped conceptual models are frequently used  
 17 in catchment hydrology, particularly in small- to mesoscale catchments (e.g. Merz and  
 18 Blöschl, 2004; Clark et al., 2008; Perrin et al., 2008; Seibert and Beven, 2009; Fenicia et al.,  
 19 2013). The above discussed concept of FLEX-TOPO (Savenije, 2010) is thereafter tested with  
 20 a stepwise increased number of parallel landscape units (FLEX<sup>B</sup>, FLEX<sup>C</sup>), thereby increasing  
 21 the conceptualized process heterogeneity and thus the model complexity. The core of the  
 22 three model set-ups is loosely based on the FLEX model (Fenicia et al., 2006).  
 23

24 **3-2-1-FLEX<sup>A</sup>:** This model set-up represents the catchment in a lumped way. The FLEX<sup>A</sup>  
 25 model structure consists of four storage elements representing interception, unsaturated, slow

1 (i.e. groundwater) and fast responding reservoirs (i.e. preferential flow and saturation  
2 overland flow). A schematic illustration of FLEX<sup>A</sup> is shown in Fig. 2a. The water balance and  
3 constitutive equations used are given in Table 2.  
4

5 **3-2-1-1-Interception reservoir ( $S_I$ ):** The interception reservoir is characterized by its  
6 maximum storage capacity ( $I_{\max}$  [L]). After precipitation ( $P$  [L T<sup>-1</sup>]) enters this reservoir the  
7 excess precipitation, hereafter referred to as effective precipitation ( $P_e$  [L T<sup>-1</sup>]), is distributed  
8 between the unsaturated ( $S_U$ ), slow ( $S_S$ ) and fast reservoir ( $S_F$ ).  
9

10 **3-2-1-2-Unsaturated reservoir ( $S_U$ ):** The unsaturated reservoir is characterized by a  
11 parameter that loosely reflects the maximum soil moisture capacity in the root zone ( $S_{U,\max}$   
12 [L]). Part of the effective precipitation ( $P_e$ ) enters the unsaturated zone according to the  
13 coefficient  $C_r$ , which here is defined by a power function with exponent  $\beta$  [-], reflecting the  
14 spatial heterogeneity of thresholds for activating fast lateral flows from  $S_F$ . This coefficient  $C_r$   
15 will be 1 when soil moisture ( $S_U$ ) is lower than a specific percentage of maximum soil  
16 moisture capacity ( $S_{U,\max}$ ) defined by relative soil moisture at field capacity ( $F_C$ [-]), meaning  
17 that the entire incoming effective precipitation ( $P_e$ ) at a given time step is stored in the  
18 unsaturated reservoir ( $S_U$ ). The soil moisture reservoir feeds the slow reservoir through matrix  
19 percolation ( $R_p$  [LT<sup>-1</sup>]), expressed as a linear relation of the available moisture in the  
20 unsaturated zone ( $S_U$ ) and the maximum percolation capacity ( $P_{\text{Per}}$  [LT<sup>-1</sup>]). The reverse  
21 process, capillary rise ( $R_c$ ), feeds the unsaturated reservoir from the saturated zone. Capillary  
22 rise ( $R_c$  [LT<sup>-1</sup>]) has an inverse linear relation with the moisture content in the unsaturated zone  
23 and is characterized by the maximum capillary rise capacity ( $C$  [LT<sup>-1</sup>]). Soil moisture is  
24 depleted by plant transpiration. Transpiration is assumed to be moisture constrained when the  
25 soil moisture content is lower than a fraction  $L_p$  [-] of the maximum unsaturated capacity  
26 ( $S_{U,\max}$ ). When the soil moisture content in the unsaturated reservoir is higher than this  
27 fraction ( $L_p$ ) transpiration is assumed to be equal to the potential evaporation ( $E_{\text{pot}}$  [LT<sup>-1</sup>]).  
28

29 **3-2-1-3- Splitter and transfer functions:** The proportion of effective rainfall which is not  
30 stored in the unsaturated zone, i.e.  $1-C_r$ , is further regulated by the partitioning coefficient ( $D$   
31 [-]), distributing flows between preferential groundwater recharge ( $R_S$  [L T<sup>-1</sup>]) to  $S_S$  and water  
32 that is routed to the stream by fast lateral processes from  $S_F$  (e.g. preferential flow or  
33 saturation overland flow,  $R_F$ ). Both fluxes are lagged by rising linear lag functions with  
34 parameters  $N_{\text{lagf}}$  and  $N_{\text{lags}}$ , respectively (e.g. Fenicia et al., 2008b).  
35

36 **3-2-1-4-Fast reservoir ( $S_F$ ):** The fast reservoir is a linear reservoir characterized by reservoir  
37 coefficient  $K_F$ .  
38

39 **3-2-1-5-Slow reservoir ( $S_S$ ):** The slow reservoir is a linear reservoir characterized by a  
40 reservoir coefficient  $K_S$ .  
41

42 **3-2-2-FLEX<sup>B</sup>:** As discussed above, a range of process studies suggested that wetlands can  
43 frequently exhibit storage-discharge dynamics that are decoupled from other parts of a  
44 catchment, in particular due to their typically reduced storage capacity and closeness to the  
45 stream. FLEX<sup>B</sup> explicitly distinguishes wetlands from the rest of the catchment, the  
46 “remainder” (i.e. hillslopes and plateaus), which is represented in a lumped way, to account  
47 for this difference. The FLEX<sup>B</sup> model set-up therefore consists of two parallel model  
48 structures which are connected with a common groundwater reservoir (Figure 2b), similar to  
49 what has been suggested by Knudsen et al. (1986). One major difference between the two  
50 parallel structures is that capillary rise is assumed to be a relevant process only in the wetland,

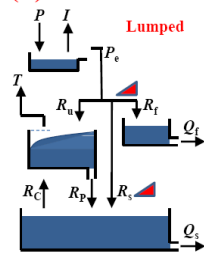


1 while it is considered negligible in the remainder of the catchment due to the deeper  
 2 groundwater. Further, since the wetlands are predominantly ex-filtration zones of potentially  
 3 low permeability, preferential recharge is considered negligible in wetlands. The areal  
 4 proportions of wetland and the remainder (i.e. hillslope and plateau) of the catchment are 15%  
 5 and 85%, respectively (Gharari et al., 2011).

6  
 7 **3-2-3-FLEX<sup>C</sup>**: This model set-up offers a complete representation of the three HRUs in the  
 8 study catchment: wetland, hillslope and plateau (Figure 2c). The formulation of the wetland  
 9 module in FLEX<sup>C</sup> is identical to the one suggested above for FLEX<sup>B</sup>. The hillslope HRU is  
 10 represented by a model structure resembling the FLEX<sup>A</sup> set-up. Plateaus are assumed to be  
 11 dominated by vertical fluxes, while direct lateral movement in the form of Hortonian overland  
 12 flow is considered negligible compared to those generated from hillslope and wetland HRUs.  
 13 Therefore the plateau model structure does not account for these fast fluxes. Analogous to  
 14 FLEX<sup>B</sup>, the FLEX<sup>C</sup> set-up is characterized by one single groundwater reservoir linking the  
 15 three dominant HRUs in this catchment. The individual proportions of wetland, hillslope and  
 16 plateau are 15%, 45% and 40%, respectively (Gharari et al., 2011). The proportions of these  
 17 HRUs are used to compute the total discharge based on the contribution of each landscape  
 18 unit.

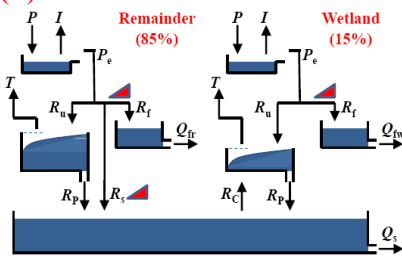
19 The connection between the parallel structures of FLEX<sup>B</sup> and FLEX<sup>C</sup> is through the surface  
 20 drainage network (the stream network) and through the slow (groundwater) reservoir.

(a) FLEX<sup>A</sup>



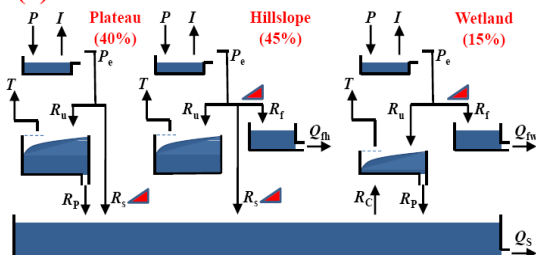
21  
 22

(b) FLEX<sup>B</sup>



23  
 24

(c) FLEX<sup>C</sup>



25  
 26

Figure 2- The model structures for (a) FLEX<sup>A</sup>, (b) FLEX<sup>B</sup> and (c) FLEX<sup>C</sup>.

27

28 Table 1- Uniform prior parameter distributions for the three model set-ups



	Unit		FLEX <sup>A</sup>		FLEX <sup>B</sup>		FLEX <sup>C</sup>		
					wetland	remainder	wetland	hillslope	plateau
$I_{max}^*$	mm	Interception storage for forest	2-5						
		Interception storage for grassland and pasture	1-3						
$S_{U,max}$	mm	Maximum unsaturated storage	0-500	0-100	0-500	0-100	0-500	0-500	
$\beta$	-	Soil moisture distribution exponent	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	
$F_C$	-	Relative soil moisture at 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	
$C$	mm(3h) <sup>-1</sup>	Maximum capillary rise rate	0	0-0.3	-	0-0.3	-	-	
$P_{per}$	mm(3h) <sup>-1</sup>	Maximum percolation rate	0-0.5	0-0.5	0-0.5	0-20.5	0	0-0.5	
$N_{lagf}$	3h	Lag time for flux to fast reservoir	1-7	1-3	1-5	1-3	1-5	-	
$N_{lags}$	3h	Lag time for preferential recharge	1-7	-	1-7	-	1-7	1-7	
$K_F$	(3h) <sup>-1</sup>	Fast reservoir coefficient	0-1	0-1	0-1	0-1	0-1	-	
$K_S$	(3h) <sup>-1</sup>	Slow reservoir coefficient	0.005-0.05	0.005-0.05	0.005-0.05	0.005-0.05	0.005-0.05	-	

1 \*Inferred from Breuer et al., 2003

2

3 Table 2- Water balance and constitutive equations used in FLEX<sup>A</sup>

Interception reservoir	$\frac{dS_I}{dt} = P - I - P_e$ (1)	$I = \begin{cases} E_{pot} & S_I > 0 \\ 0 & S_I = 0 \end{cases} \quad (2)$ $P_e = \begin{cases} 0 & S_I < I_{max} \\ P & S_I = I_{max} \end{cases} \quad (3)$
Unsaturated reservoir	$\frac{dS_U}{dt} = R_u - T - R_p + R_C$ (4)	$R_u = C_r P_e \quad (5)$ $C_r = \begin{cases} 1 - \left[ \frac{(S_u - S_{u,max} F_C)}{(S_{u,max} - S_{u,max} F_C)} \right]^\beta & S_u \geq S_{u,max} F_C \\ 1 & S_u < S_{u,max} F_C \end{cases} \quad (6)$ $T = K_T (E_{pot}) \quad (7)$ $K_T = \begin{cases} \left[ \frac{S_u}{S_{u,max} L_p} \right] & S_u < S_{u,max} L_p \\ 1 & S_u \geq S_{u,max} L_p \end{cases} \quad (8)$ $R_p = [S_u / S_{u,max}] P_{per} \quad (9)$ $R_C = [1 - (S_u / S_{u,max})] C \quad (10)$
Fast reservoir	$\frac{dS_F}{dt} = R_{F,lag} - Q_F$ (11)	$R_F = (1 - D)(1 - C_r) P_e \quad (12)$ $R_{F,lag} = R_F * N_{lagf} \quad (13)$ $Q_F = K_F S_F \quad (14)$
Slow reservoir	$\frac{dS_S}{dt} = R_{S,lag} - Q_S + R_p - R_C$ (16)	$R_S = D(1 - C_r) P_e \quad (17)$ $R_{S,lag} = R_S * N_{lags} \quad (18)$ $Q_S = K_S S_S \quad (19)$

4 \* is the convolution operator.

5

### 6 3-3- Introducing realism constraints in selecting behavioral parameter sets:

7

8 With increasing process heterogeneity from FLEX<sup>A</sup> over FLEX<sup>B</sup> to FLEX<sup>C</sup>, the respective  
9 model complexities and therefore the number of calibration parameters also increase. This, in

1 the frequent absence of sufficient suitable data to efficiently constrain a model, typically leads  
 2 to a situation where parameters have increased freedom to compensate for errors in data and  
 3 model structures, as recently reiterated by Gupta et al. (2008). As a consequence, the resulting  
 4 higher level of equifinality can substantially reduce a model's predictive power. In this study,  
 5 two fundamentally different types of model constraints were applied to test their value for  
 6 reducing equifinality in complex model set-ups. Firstly, conditions between parameters of  
 7 parallel model units, hereafter referred to as *parameter constraints*, were imposed *before* each  
 8 model evaluation run. These *a priori* constraints ensure that the individual parameter values  
 9 for the same process in the parallel units, reflect the modeler's perception of the system. For  
 10 example, it can be argued that the maximum interception capacity ( $I_{\max}$ ) of a forested HRU  
 11 needs to be higher than the one in a not forested HRU. In the absence of more detailed  
 12 information this does not only allow overlapping prior distributions but it also avoids the need  
 13 for quantification of the constraints themselves. The second type of constraints is *process*  
 14 *constraints*, which can only be applied *after* each evaluation run during the calibration phase.  
 15 These *a posteriori* constraints compare the modeled output of the individual HRUs and ensure  
 16 that these outputs follow the modeler's perception of the system's internal dynamics. For  
 17 example, it can be argued that the modeled evaporation in forested HRUs needs to be higher  
 18 than in not forested HRUs. The parameter and process constraints imposed on the models in  
 19 this study are described in detail below. Note that the choice of constraints to impose is the  
 20 modeler's choice and that with increasing number of different HRUs an increasing number of  
 21 constraints can be applied. While here FLEX<sup>A</sup> only allows for two constraints, i.e. one  
 22 parameter- and one process constraint, all constraints suggested below can be applied to  
 23 FLEX<sup>C</sup>.

### 24 3-3-1-Parameter constraints:

25  
 26  
 27 A number of *a priori* constraints is imposed on different model parameters in order to exclude  
 28 unrealistic parameter combinations. The constraints are guided by considerations on what the  
 29 model components are designed to reproduce. The number of constraints that can be imposed  
 30 increases with increasing model complexity. The full set of parameter constraints detailed  
 31 below were applied to FLEX<sup>C</sup> and when applicable also to FLEX<sup>B</sup>. In contrast, only one  
 32 parameter constraint could be used for FLEX<sup>A</sup>, as for this model no more obvious  
 33 relationships between parameters could be identified. In the following, the subscripts w, h and  
 34 p indicate parameters for wetland, hillslope and plateau, respectively.

#### 35 3-3-1-1-Interception:

36  
 37  
 38 The different land cover proportions of each landscape unit, here wetlands, hillslopes and  
 39 plateaus, can be used to define the relation between interception thresholds ( $I_{\max}$ ) of these  
 40 individual units. The land uses are defined as two general classes for this case study, forested  
 41 areas and grass or pasture-land areas. The maximum interception capacity ( $I_{\max}$ ) for each  
 42 landscape entity can be estimated from the proportion of land-use classes and their maximum  
 43 interception capacities, selected from their respective prior distributions as given in Table 1:

$$44 \quad I_{\max,w} = a_w I_{\max,forest} + b_w I_{\max,cropland} \quad (20)$$

$$45 \quad I_{\max,h} = a_h I_{\max,forest} + b_h I_{\max,cropland} \quad (21)$$

$$46 \quad I_{\max,p} = a_p I_{\max,forest} + b_p I_{\max,cropland} \quad (22)$$

47 The proportions of forested area are indicated with  $a_w$ ,  $a_h$  and  $a_p$  for wetland, hillslope and  
 48 plateau and are fixed at 42%, 60% and 29%, respectively. The proportions of cropland and  
 49 grass land areas are indicated by  $b_w$ ,  $b_h$  and  $b_p$  for wetland, hillslope and plateau and are fixed

1 at 58%, 40% and 71%, respectively. Moreover the parameter sets which are selected for  
2 maximum interception capacity of forest are expected to be higher than crop- or grassland:

$$3 \quad I_{\max, \text{cropland}} < I_{\max, \text{forest}} \quad (23)$$

### 7 **3-3-1-2- Lag functions:**

8  
9 Preferential recharge ( $R_s$ ) is routed to the slow reservoir by a lag function. Due to a deeper  
10 groundwater table on plateaus it can be assumed that the lag time for  $R_s$  is longer for plateaus  
11 than for hillslopes. It can also be assumed that the lag function used for fast reservoir for  
12 hillslopes is longer than for wetlands due to the on average higher distance of and therefore  
13 longer travel times from hillslopes to the stream.

$$14 \quad N_{\text{lags}, w} \leq N_{\text{lags}, h} \leq N_{\text{lags}, p} \quad (24)$$

### 17 **3-3-1-3-Soil moisture capacity:**

18  
19 Many experimental studies support the assumption that wetlands have shallower groundwater  
20 tables than the other two landscape entities in this study. Therefore the unsaturated zone of  
21 wetland should be shallower, i.e. the maximum soil moisture capacity ( $S_{U, \max}$ ) of hillslopes  
22 and plateaus can be assumed to be higher. Moreover, as hillslopes in the study catchment are  
23 predominantly covered with forest, it can, due to the deeper root zone of forests, be expected  
24 that the maximum unsaturated soil moisture capacity ( $S_{U, \max}$ ) in the root zone of hillslopes is  
25 deeper than the other two landscape entities.

$$26 \quad S_{U, \max, w} < S_{U, \max, p} < S_{U, \max, h} \quad (25)$$

### 29 **3-3-1-4-Reservoir coefficients:**

30  
31 The reservoir coefficient of the wetland fast reservoir ( $K_F$ ) is assumed to be lower than  
32 reservoir coefficient of the hillslope fast reservoir as, once connectivity is established, the  
33 flow velocities of saturation overland flow in wetlands are assumed to exceed the integrated  
34 flow velocities of preferential flow networks (cf. Anderson et al., 2009). Likewise, the  
35 retention time of the slow reservoir should be higher than both wetland and hillslope fast  
36 reservoirs.

$$37 \quad K_S < K_{F, h} < K_{F, w} \quad (26)$$

38 The reservoir constraints can be applied to all models while the other constraints can only be  
39 applied to FLEX<sup>B</sup> and FLEX<sup>C</sup>.

### 42 **3-3-2-Process constraints:**

43  
44 In contrast to the parameter constraints discussed above, which are set *a priori*, process  
45 constraints are applied *a posteriori*. Only parameters which generate model internal flux  
46 dynamics in agreement with the modeler's perception of these dynamics are retained as  
47 feasible. Hence, while with the use of parameter constraints there is no need to run the model

1 for rejected parameter sets, here it is necessary to run the model to evaluate it with respect to  
 2 the process constraints.

3 Process constraints are defined for dry and wet periods as well as for peak-, high- and low  
 4 flows. Here wet periods were defined to be the months from December to March, while the  
 5 dry periods in the study catchment occur between April and November. The thresholds for  
 6 distinguishing between high and low flow were chosen to be 0.05 and 0.2 mm(3h)<sup>-1</sup>  
 7 respectively for dry and wet periods. Furthermore, events during which discharge increases  
 8 with a rate of more than 0.2 mm(3h)<sup>-2</sup> are defined as peak flows. Note that in the following  
 9 the subscripts peak, high and low indicate peak-, high- and low flows.

### 11 3-3-2-1-Transpiration:

12  
 13 Transpiration typically exhibits a clear relationship with the normalized difference vegetation  
 14 index (NDVI, Szilagyi et al., 1998). Therefore the ratios between NDVI values of different  
 15 landscape units can serve as constraints on modeled transpiration obtained from the individual  
 16 parallel model components. A rough estimation of the ratio between transpiration from  
 17 plateau and hillslope can be derived from LANDSAT 7 images. For this ratio seven cloud free  
 18 images were selected (acquisition dates of 20/4/2000, 6/3/2000, 11/9/2000, 18/2/2001,  
 19 6/3/2001, 26/3/2001 and 29/8/2001). The ratio of transpiration between hillslope and plateau  
 20 ( $R_{trans}$ ) can be estimated by assuming a linear relation (Szilagyi et al., 1998) with slope of  $\alpha$   
 21 and intercept zero between transpiration and mean NDVI for each landscape unit ( $\mu_{NDVI}$ ).  
 22

$$23 \quad R_{trans} = \frac{\alpha \mu_{NDVI,h}}{\alpha \mu_{NDVI,p}} = \frac{\mu_{NDVI,h}}{\mu_{NDVI,p}} \quad (27)$$

24 Mean ( $\mu_{R_{trans}}$ ) and standard deviation ( $\sigma_{R_{trans}}$ ) of  $R_{trans}$  can be used to estimate acceptable limits  
 25 of the transpiration ratios for hillslope and plateau.

26 Therefore the annual transpiration can be confined between two values as follows:

$$27 \quad \mu_{R_{trans}} - \sigma_{R_{trans}} < \frac{\int T_h dt}{\int T_p dt} < \mu_{R_{trans}} + \sigma_{R_{trans}} \quad (28)$$

28 Based on the mean ( $\mu_{R_{trans}} = 1.2$ ) and standard deviation ( $\sigma_{R_{trans}} = 0.2$ ) of the seven LANDSAT 7  
 29 images used the following process constraint on transpiration from hillslope ( $T_h$ ) and plateau  
 30 ( $T_p$ ) was imposed:

$$31 \quad 1.0 < \frac{\int T_h dt}{\int T_p dt} < 1.4 \quad (29)$$

32 Similar constraints can be imposed between transpiration fluxes from wetland, hillslope or  
 33 plateau; however, the spatial resolution of LANDSAT 7 data with resolution of 30 meters is  
 34 coarser than the required 20-meter DEM resolution for distinguishing wetlands from other  
 35 landscape entities (Gharari et al., 2011).  
 36

### 37 3-3-3-2-Runoff coefficient:

38  
 39 The runoff coefficient is a frequently used catchment signature (e.g. Sawicz et al., 2011;  
 40 Euser et al., 2013) and can be used as a behavioral constraint (e.g. Duan et al., 2006;  
 41 Winsemius et al., 2009). In this study the runoff coefficients of dry and wet periods as well as  
 42 the annual runoff coefficient were used. Parameters that result in modeled runoff coefficients  
 43 that substantially deviate from the observed ones are therefore discarded. In case of absence  
 44 of suitable runoff data, the mean annual runoff coefficient can be estimated from the regional

1 Budyko curve using for example the Turc-Pike relationship (Turc, 1954; Pike, 1964; Arora,  
 2 2002). However in this study the runoff coefficients of each individual year, and their  
 3 respective dry and wet periods was used and determined the mean and standard deviation of  
 4 the runoff coefficients for these periods. Here, as a conservative assumption, the limits are set  
 5 to three times the standard deviation around the mean runoff coefficient. Note that the runoff  
 6 coefficient is the only process constraint that is not related to model structure in this study and  
 7 can therefore also be applied to the lumped FLEX<sup>A</sup> set-up.

$$8 \quad \frac{\int Q_m dt}{\int P dt} < 0.43 \quad (30)$$

$$9 \quad \frac{\int Q_m dt}{\int P dt} > 0.16 \quad (31)$$

$$10 \quad \frac{\int Q_{m,dry} dt}{\int P_{dry} dt} < 0.36 \quad (32)$$

$$11 \quad \frac{\int Q_{m,dry} dt}{\int P_{dry} dt} > 0 \quad (33)$$

$$12 \quad \frac{\int Q_{m,wet} dt}{\int P_{wet} dt} < 0.71 \quad (34)$$

$$13 \quad \frac{\int Q_{m,wet} dt}{\int P_{wet} dt} > 0.40 \quad (35)$$

### 14 3-3-3-3-Preferential recharge:

15  
 16 The slow reservoir can be recharged by both preferential and matrix percolation from the  
 17 unsaturated reservoirs. Here, hillslopes and plateaus contribute to the slow reservoir by  
 18 preferential recharge. It can be assumed that in a realistic model setup the long term  
 19 contribution volume of preferential recharge ratio between hillslope and plateau should not be  
 20 unrealistically high or low. For example, it can be assumed unrealistic that the ratio is zero or  
 21 infinity, meaning that one landscape unit is constantly feeding the slow reservoir while  
 22 another one is not contributing at all. To avoid such a problem, a loose and very conservative  
 23 constraint was imposed on the ratio of contribution of the two fluxes.  
 24  
 25

$$26 \quad 0.2 < \frac{\int R_{s,h} dt}{\int R_{s,p} dt} < 5 \quad (36)$$

### 27 3-3-3-4-Fast component discharge:

28  
 29 During dry periods, hillslopes and plateaus can exhibit significant soil moisture deficits,  
 30 limiting the amount of fast runoff generated from these landscape elements. In contrast, due to  
 31 their reduced storage capacity, wetlands are likely to generate fast flows at lower moisture  
 32 levels, thus dominating event response during dry periods (cf. Beven and Freer, 2001a;  
 33 Seibert et al., 2003; Molenat et al., 2005; Anderson et al., 2010; Birkel et al., 2010). It can  
 34 thus be assumed that during both, the entire dry periods as well as peak flows in dry periods  
 35

1 the fast component of wetlands ( $Q_{f,w,dry}$ ;  $Q_{f,w,dry,peaks}$ ) contributes to runoff more than the fast  
 2 component of hillslopes ( $Q_{f,h,dry}$ ;  $Q_{f,h,dry,peaks}$ ). In contrast, high flows during wet periods are  
 3 predominantly generated by hillslopes ( $Q_{f,h,wet}$ ;  $Q_{f,h,wet,high}$ ). This process constraint is also  
 4 applied to FLEX<sup>B</sup>.

$$5 \quad \frac{\int Q_{f,h,dry,peaks} dt}{\int Q_{f,w,dry,peaks} dt} < 1 \quad (37)$$

$$6 \quad \frac{\int Q_{f,h,wet,high} dt}{\int Q_{f,w,wet,high} dt} > 1 \quad (38)$$

$$7 \quad \frac{\int Q_{f,h,wet} dt}{\int Q_{f,w,wet} dt} > 1 \quad (39)$$

### 8 9 10 11 **3-3-4-Calibration algorithm and objective functions:**

12  
13 Based on uniform prior parameter distributions as well as on the parameter- and process  
 14 constraints the model was calibrated using MOSCEM-UA (Vrugt et al., 2003). As a brief  
 15 description, MOSCEM-UA uses a Latin Hypercube sampling strategy for the random  
 16 sampling of the entire parameter space. Penalizing the objective function(s) based on the  
 17 number of unsatisfied constraints, however, may lead to non-smooth objective functions  
 18 which potentially may cause instabilities in the search algorithm or create invalid results. A  
 19 recently developed stepwise search algorithm was therefore used for finding parameter sets  
 20 which satisfy both parameter- and process constraints (Gharari et al., this issue). These  
 21 parameter sets were then used as initial sampling parameter sets for MOSCEM-UA instead of  
 22 the traditionally used Latin Hypercube sampling strategy.

23 The models were evaluated on the basis of three different objective functions to emphasize  
 24 different characteristics of the system response: (i) the Nash-Sutcliffe efficiency of the flows  
 25 (Nash and Sutcliffe, 1970;  $E_{NS}$ ), (ii) the Nash-Sutcliffe efficiency of the logarithm of the flows  
 26 ( $E_{NS,log}$ ) and (iii) the Nash-Sutcliffe efficiency of the flow duration curve ( $E_{NS,FDC}$ ). These  
 27 criteria evaluate the models' ability to simultaneously reproduce high flows, low flows and  
 28 flow duration curves respectively. The model set-ups have been constrained and calibrated for  
 29 the year 2002-2005 and validated for year 2006-2009. The year 2001 was used as warm up  
 30 period.

### 31 32 **3-4-Model validation and parameter evaluation:**

33  
34 To assess the value of incorporating parameter and process constraints in increasingly  
 35 complex models a four-step procedure as outlined below was followed. Note that for each  
 36 step the respective model (parameters) was evaluated against the constrained and calibrated  
 37 lumped FLEX<sup>A</sup> benchmark model.

#### 38 39 **3-4-1- Evaluating models with “constrained but uncalibrated” parameter sets:**

40  
41 Firstly, all parameter sets which satisfy all the applied constraints were evaluated based on  
 42 their ability to reproduce the observed hydrograph. Hereafter these parameter sets are referred  
 43 to as *constrained but uncalibrated* parameter sets because they were obtained *without any*  
 44 *calibration* on the observed hydrographs. Based on the retained, feasible parameter sets, the

1 mean performance of the three *constrained but uncalibrated* models FLEX<sup>A</sup>, FLEX<sup>B</sup> and  
2 FLEX<sup>C</sup>, for the three objective functions ( $E_{NS}$ ,  $E_{NS,\log}$ ,  $E_{NS,FDC}$ ) together with their uncertainty  
3 ranges for both the calibration and the validation periods are compared. FLEX<sup>A</sup>, FLEX<sup>B</sup> and  
4 FLEX<sup>C</sup> have an increasing number of constraints. It was thus tested whether the higher  
5 complexity models also result in better model performance and how the predictive uncertainty  
6 is affected by increased complexity and model realism.

7 To investigate how well the hydrographs generated with parameters satisfying all constraints  
8 match the observed hydrograph, the 95% uncertainty intervals of simulated hydrographs  
9 based on these parameter sets were generated for the three models. The uncertainty was  
10 estimated on the basis of the area indicated by 95% uncertainty intervals based on simulated  
11 hydrographs.

### 12 **3-4-2- Evaluating models with “constrained and calibrated” parameter sets:**

13  
14  
15 In the second step, the three models FLEX<sup>A</sup>, FLEX<sup>B</sup> and FLEX<sup>C</sup> have been calibrated within  
16 the parameter space which satisfied all the imposed parameter and process constraints. The  
17 models were calibrated using a multi-objective strategy ( $E_{NS}$ ,  $E_{NS,\log}$ ,  $E_{NS,FDC}$ ). The obtained  
18 Pareto optimal model parameters are in the following referred to as *constrained and*  
19 *calibrated*.

20 Analogous to the previous step uncertainty intervals based on the constrained and calibrated  
21 Pareto optimal parameters, were generated. The uncertainty was estimated on the basis of the  
22 area of the uncertainty bands.

### 23 **3-4-3- Comparison of model performance and uncertainty for “constrained but un- 24 calibrated” and “constrained and calibrated” parameter sets:**

25  
26  
27 To assess the added value of incorporating constraints in higher complexity models, the  
28 performance and uncertainties of the three models FLEX<sup>A</sup>, FLEX<sup>B</sup> and FLEX<sup>C</sup> were  
29 compared for both the “constrained but un-calibrated” and the “constrained and calibrated”  
30 case during calibration (2002-2005) and validation (2006-2009) periods.

### 31 **3-4-4- Comparison of modeled hydrograph components for different model structures:**

32  
33  
34 One of the main reasons for imposing constraints on model parameters is to ensure the  
35 realistic internal dynamic of a model. Comparing different fluxes contributing to the modeled  
36 hydrograph can give an insight into the performance of imposed constrained on the model.  
37 The effect of imposing behavioral constraints on fast and slow components of the three  
38 models structures, FLEX<sup>A</sup>, FLEX<sup>B</sup> and FLEX<sup>C</sup> is compared visually. The fast component of  
39 lumped model, FLEX<sup>A</sup>, is compared with fast components of FLEX<sup>B</sup> which are wetland and  
40 remainder of catchment and fast components of FLEX<sup>C</sup> which are wetland and hillslope. This  
41 visual comparison is based on normalized average contribution of each component for Pareto  
42 optimal parameter sets in every time step.

## 43 **4- Results and discussion**

### 44 **4-1- Evaluating the performance of constrained but uncalibrated parameter sets:**

45  
46  
47 The median and the 95% uncertainty intervals of the performance of modeled hydrographs for  
48 *constrained but uncalibrated* parameter sets is presented in Table 3 for the calibration period.  
49  
50



1 The lumped FLEX<sup>A</sup> model has only one parameter and one process constraint, i.e. the runoff  
 2 coefficient. Hence, this model is free within the limits of this apparently relatively weak  
 3 condition, resulting in a wide range of possible parameters, many of which cannot adequately  
 4 reproduce the system response. As a consequence, the overall performance is poor  
 5 ( $E_{NS,median}=0.19$ ,  $E_{NS,log,median}=0.13$ ,  $E_{NS,FDC,median}=0.38$ ) (Table 3, Figure 3).

6 FLEX<sup>B</sup>, run with the set of *constrained but uncalibrated* parameters shows a substantial  
 7 improvement in overall performance ( $E_{NS,median}=0.59$ ,  $E_{NS,log,median}=0.44$ ,  $E_{NS,FDC,median}=0.92$ )  
 8 compared to FLEX<sup>A</sup>, as FLEX<sup>B</sup> not only allows for more process heterogeneity but, more  
 9 importantly, it is conditioned with an increased number of constraints.

10 The additional process heterogeneity and constraints allowed by FLEX<sup>C</sup>, results in the highest  
 11 overall performance for all three objective functions ( $E_{NS,median}=0.68$ ,  $E_{NS,log,median}=0.54$ ,  $E_{NS,$   
 12  $FDC,median}=0.95$ ) (Table 3, Figure 3).

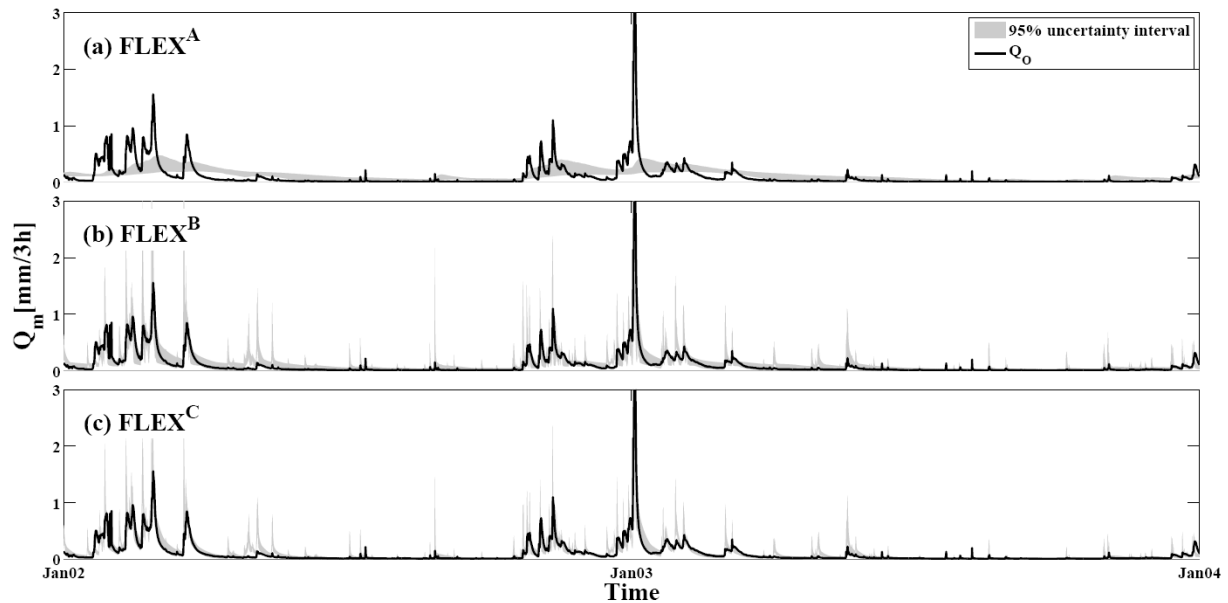
13 These results clearly illustrate that the imposed relational constraints force the model and its  
 14 parameters towards a more realistic behavior, which significantly improves model  
 15 performance.

16 The 95% uncertainty intervals indicate that FLEX<sup>C</sup>, which might be expected to produce the  
 17 highest uncertainty interval due to its complexity, is providing a lower uncertainty compared  
 18 to FLEX<sup>B</sup>. Although FLEX<sup>C</sup> cannot outperform FLEX<sup>A</sup> in terms of a lower uncertainty  
 19 interval, overall performance of this model is better than FLEX<sup>A</sup> as discussed earlier.

20 Table 3- The median model performances (in brackets their corresponding 95% uncertainty  
 21 intervals) and the area spanned by the 95% uncertainty interval of hydrograph derived from  
 22 uncalibrated parameter sets which satisfy the complete set of constraints for the three model  
 23 set-ups FLEX<sup>A</sup>, FLEX<sup>B</sup> and FLEX<sup>C</sup>, for the three modeling objectives ( $E_{NS}$ ,  $E_{NS,log}$ ,  $E_{NS,FDC}$ )  
 24 in the calibration (2002-2005) and validation (2006-2009) periods.

		$E_{NS}$	$E_{NS,log}$	$E_{NS,FDC}$	95% uncertainty area [mm]
FLEX <sup>A</sup>	Calibration	0.19 [0.12 0.28]	0.13 [-0.13 0.41]	0.38 [0.29 0.56]	801
	Validation	0.20 [0.10 0.33]	0.37 [0.18 0.56]	0.40 [0.27 0.61]	806
FLEX <sup>B</sup>	Calibration	0.59 [0.25 0.75]	0.44 [0.16 0.61]	0.92 [0.81 0.95]	1396
	Validation	0.54 [0.23 0.76]	0.59 [0.31 0.73]	0.92 [0.76 0.98]	1550
FLEX <sup>C</sup>	Calibration	0.68 [0.46 0.77]	0.54 [0.07 0.65]	0.95 [0.91 0.96]	878
	Validation	0.66 [0.42 0.77]	0.63 [0.05 0.77]	0.97 [0.94 0.99]	1025

25  
 26  
 27



1  
2 Figure 3- The observed hydrograph and the 95% uncertainty interval of the modeled  
3 hydrograph derived from the complete set of *constrained but un-calibrated* parameter sets for  
4 the three different model set-ups (a) FLEX<sup>A</sup>, (b) FLEX<sup>B</sup> and (c) FLEX<sup>C</sup> for two years (2002-  
5 2003) of calibration.

6  
7 **4-2- Evaluating the performance of *constrained and calibrated* parameter sets:**

8  
9 The comparison of the *constrained and calibrated* model set-ups shows that all three models  
10 set-ups can reproduce the hydrograph similarly well (Table 4, Figure 4). FLEX<sup>A</sup> exhibits a  
11 slightly better calibration performance compared to the other two model set-ups. This can  
12 partly be attributed to the lower number of parameters which leads, with the same number of  
13 samples, to a more exhaustive sampling of the parameter space and a smoother identification  
14 of Pareto optimal solutions. In addition, FLEX<sup>A</sup> has the lowest number of imposed  
15 constraints, i.e. only the runoff coefficient and one parameter constraints, compared to FLEX<sup>B</sup>  
16 and FLEX<sup>C</sup>. This model set-up therefore allows more freedom in exploiting the parameter  
17 space to produce mathematically good fits between observed and modeled system response in  
18 the calibration period.

19 For the validation period, arguably more important for model evaluation, as in contrast to the  
20 calibration period, it gives information on model consistency (cf. Klemes, 1986; Andréassian  
21 et al., 2009; Euser et al., 2013) and predictive uncertainty, the performances of the three  
22 model set-ups exhibit quite different patterns (Table 4). The simplest model, the lumped  
23 FLEX<sup>A</sup>, is characterized by the highest performance deterioration from calibration to  
24 validation. FLEX<sup>B</sup> shows a better validation/calibration performance ratio than FLEX<sup>A</sup>.  
25 Despite the expectation that increasingly complex models will have increasingly poor  
26 validation/calibration performance ratios, due to higher degrees of freedom, FLEX<sup>C</sup> exhibited  
27 a more stable performance between calibration and validation.

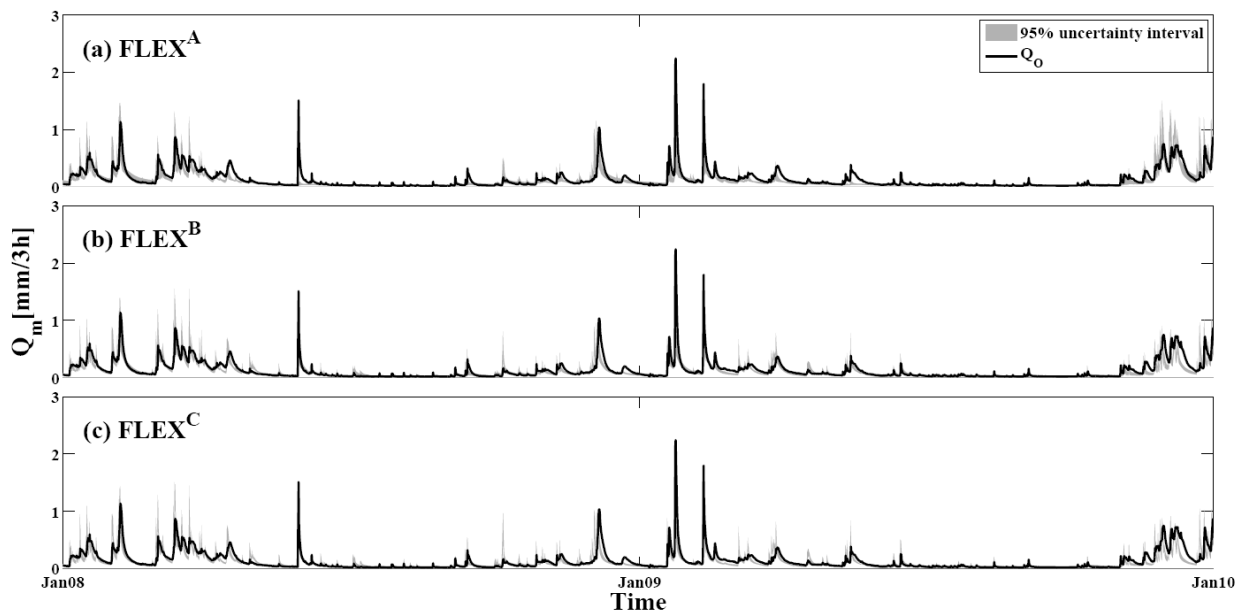
28 In addition, the absolute performance of FLEX<sup>C</sup> in the validation period is in general higher  
29 than the performances of FLEX<sup>A</sup> and FLEX<sup>B</sup> (Table 4). Although, strictly speaking, no  
30 meaningful comparison between Nash-Sutcliffe efficiencies from different periods can be  
31 made, these results nevertheless indicate that the most complex model set-up, i.e. FLEX<sup>C</sup>, is  
32 the most consistent model-set-up with the lowest predictive uncertainty, which has important  
33 implications that will be discussed below. The explanation is that in spite of the high degree  
34 of process heterogeneity, the high number of constraints in FLEX<sup>C</sup> prevents the calibration

1 algorithm to over-fit this complex model set-up, thus reducing the probability of seriously  
 2 misrepresenting reality.

3  
 4 Table 4- The median model performances (in brackets their corresponding Pareto uncertainty  
 5 intervals) and the area spanned by the uncertainty interval of the hydrograph derived from the  
 6 Pareto optimal solutions of the *constrained and calibrated* model set-ups FLEX<sup>A</sup>, FLEX<sup>B</sup> and  
 7 FLEX<sup>C</sup> for the three modeling objectives ( $E_{NS}$ ,  $E_{NS,\log}$ ,  $E_{NS,FDC}$ ) in the calibration and  
 8 validation periods.

		$E_{NS}$	$E_{NS,\log}$	$E_{NS,FDC}$	95% uncertainty area [mm]
FLEX <sup>A</sup>	Calibration	0.74 [0.51 0.82]	0.74 [0.66 0.80]	0.96 [0.95 0.98]	696
	Validation	0.66 [0.48 0.78]	0.75 [0.70 0.81]	0.97 [0.94 0.98]	826
FLEX <sup>B</sup>	Calibration	0.74 [0.60 0.80]	0.73 [0.59 0.79]	0.96 [0.94 0.98]	627
	Validation	0.72 [0.57 0.79]	0.78 [0.64 0.82]	0.96 [0.94 0.98]	719
FLEX <sup>C</sup>	Calibration	0.69 [0.60 0.79]	0.66 [0.60 0.67]	0.96 [0.94 0.96]	508
	Validation	0.68 [0.57 0.78]	0.74 [0.66 0.79]	0.98 [0.95 0.98]	593

9  
 10



11  
 12 Figure 4- The observed hydrograph and the 95% Pareto uncertainty interval of the modeled  
 13 hydrograph for constrained and calibrated parameter sets for the three different model set-ups  
 14 (a) FLEX<sup>A</sup>, (b) FLEX<sup>B</sup> and (c) FLEX<sup>C</sup> for the two years (2008-2009) of validation period.

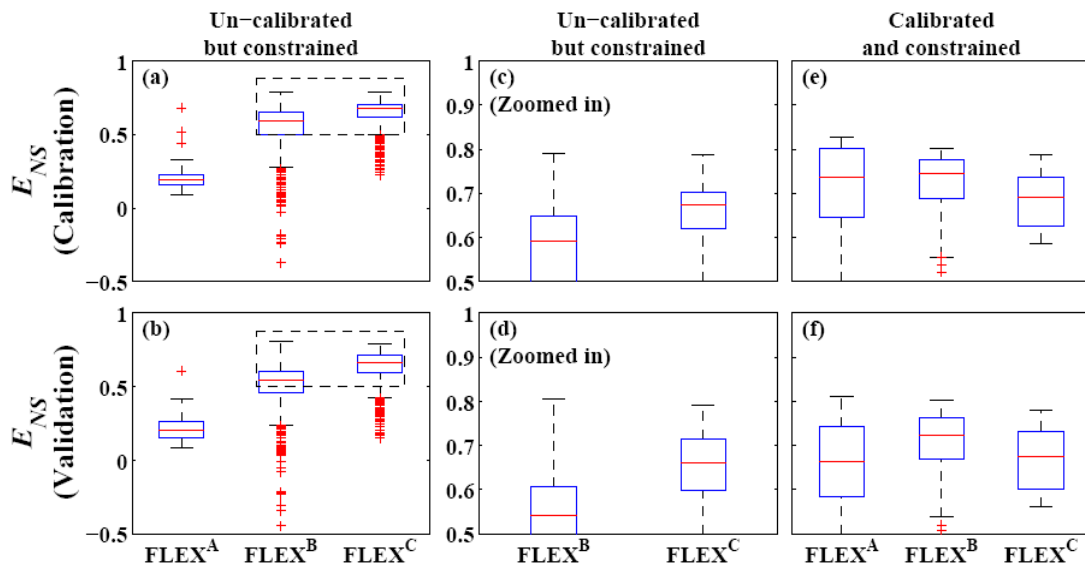
15  
 16 **4-3- Comparison of “constrained but uncalibrated” and “constrained and calibrated”  
 17 models:**

18  
 19 The following comparison of the performances of FLEX<sup>A</sup>, FLEX<sup>B</sup> and FLEX<sup>C</sup> for  
 20 “constrained but uncalibrated” and “constrained and calibrated” parameter sets focused on

1  $E_{NS}$  only, for the reason of brevity (Figure 5). In Figures 5a and 5b the model performances  
 2 based on the “constrained but uncalibrated” parameter sets, that satisfy the full set of  
 3 constraints, are shown for the calibration and validation periods. As discussed in detail above,  
 4 although uncalibrated, increasing the number of constraints from FLEX<sup>A</sup> to FLEX<sup>C</sup> increases  
 5 the overall performance of the models while reducing uncertainty (Figures 5c and 5d; note  
 6 that these are zoom-ins).

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

19 In addition a second crucial aspect was revealed by comparing “constrained but un-  
 20 calibrated” and “constrained and calibrated” models. It can be seen that, for the study  
 21 catchment, a calibrated lumped model, FLEX<sup>A</sup> (Figure 5f, left plot) can on average not  
 22 outperform a more complex *constrained but uncalibrated* model, i.e. FLEX<sup>C</sup> (Figure 5d, right  
 23 plot). This has potentially important implications for selecting suitable parameter values for  
 24 models applied in ungauged basins as it highlights the value of semi- and non-quantitative  
 25 hydrological expert knowledge, even in the absence of reliable model regionalization tools  
 26 and detailed soil or geological information, as discussed in detail below.



27  
 28  
 29 Figure 5- Model performance ( $E_{NS}$ ) based on *constrained but uncalibrated* (a-d) and  
 30 *constrained and calibrated* (e-f) parameter sets for calibration (2002-2005) and validation  
 31 (2006-2009) periods for the three different model set-ups FLEX<sup>A</sup>, FLEX<sup>B</sup> and FLEX<sup>C</sup>. Note  
 32 that (c) and (d) are zoom-ins of (a) and (b).  
 33

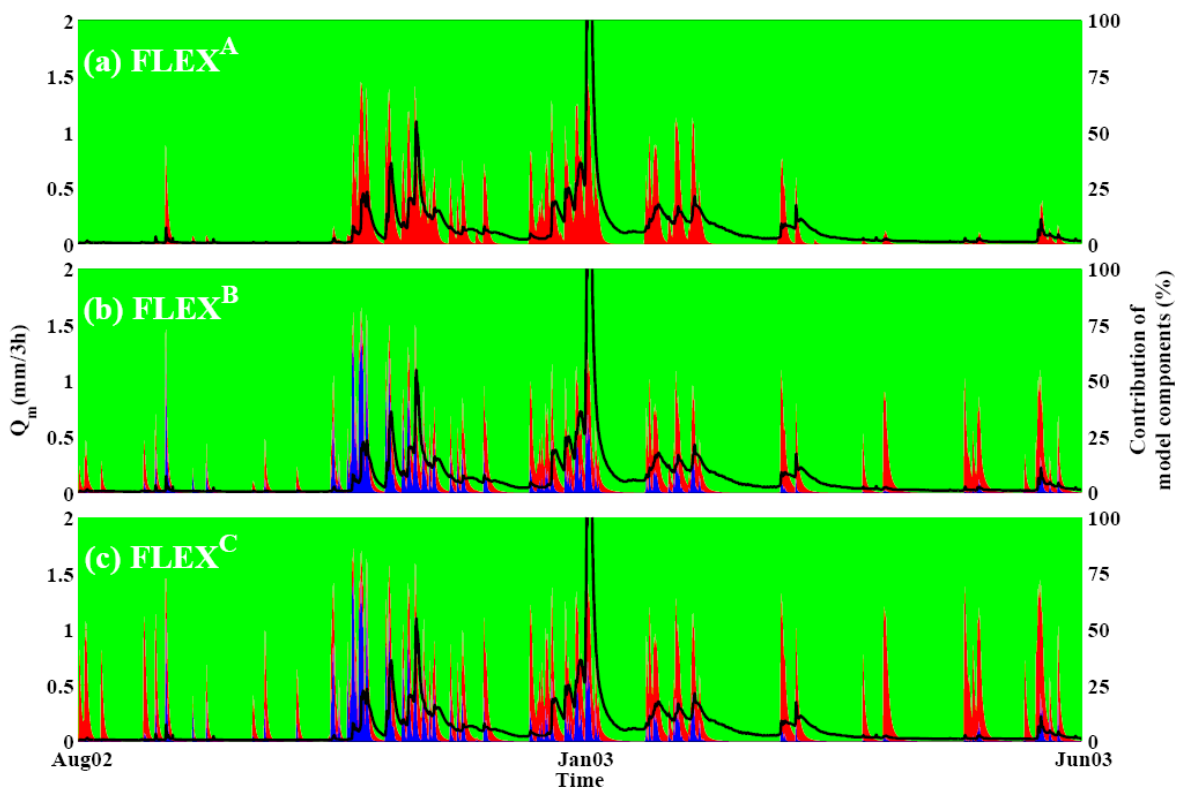
#### 4-4- Comparison of flow contributions from different model components:

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<sup>A</sup>, fast responses of the remainder of the catchment of FLEX<sup>B</sup> and fast responses of hillslope of FLEX<sup>C</sup> is indicated by red color; wetland fast responses of FLEX<sup>B</sup> and FLEX<sup>C</sup> are indicated by predominance of blue color.

The colors in Fig. 6 are an illustration using three colors (red, green and blue) 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<sup>A</sup> 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 parameters can generate peak flows by predominant contributions from slow responses while fast reaction is tend to be inactive during these events.

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<sup>B</sup> and FLEX<sup>C</sup> control the rapid response during wetting up periods (dry to wet transition), before hillslope fast processes become more important at higher moisture levels. When the system is saturated the hillslope contribution to modeled runoff becomes significantly higher compared to the wetland response. Note that the response of the wetland may not correspond well to individual events, as a consequence of the fact that the corresponding constraint was set for an aggregated period.



1 Figure 6- Comparison between mean proportions of Pareto members for model components  
2 of the three model set-ups in part of the calibration periods (August 2002- June 2003) (a)  
3 FLEX<sup>A</sup>, (b) FLEX<sup>B</sup>, and (c) FLEX<sup>C</sup>. The green color indicates the relative contribution of the  
4 slow reservoir for the three different models. Red indicates relative contribution from the fast  
5 components, i.e. fast reservoir in FLEX<sup>A</sup>, fast reservoir of the remainder of the catchment in  
6 FLEX<sup>B</sup> and fast reservoir of hillslope of FLEX<sup>C</sup>. The blue color indicates the relative  
7 contribution of fast wetland component of FLEX<sup>B</sup> and FLEX<sup>C</sup>.

#### 9 **4-5- General discussion:**

10  
11 The results of this study quite clearly indicate that discretizing the catchment into  
12 hydrological response units (HRUs) and incorporating expert knowledge in model  
13 development and testing is a potentially powerful strategy for runoff prediction, even where  
14 insufficient data for model calibration (e.g. Koren et al., 2003; Duan et al., 2006; Winsemius  
15 et al., 2009) or only comparatively unreliable regionalization tools are available (e.g. Wagener  
16 and Wheater, 2006; Bárdossy, 2007; Parajka et al., 2007; Oudin et al., 2008; Laaha et al.,  
17 2013). It was found that the performance and the predictive power of a comparatively  
18 complex uncalibrated conceptual model, based on posterior parameter distributions obtained  
19 merely from relational, semi- and non-quantitative realism constraints inferred from expert  
20 knowledge, can be as efficient as the calibration of a lumped conceptual model (Fig. 5).

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

28 A complex model may include many processes, i.e. hypotheses, which can usually not be  
29 rigorously tested with the available data. However, a wide range of previous studies has  
30 demonstrated that hydrologically meaningful constraints can help to limit the increased  
31 uncertainty caused by incorporating additional processes, i.e. parameters (e.g. Yadav et al.,  
32 2007; Zhang et al., 2008; Kapangaziwiri et al., 2012). These studies generally include a large  
33 set of catchments and try to relate model parameters to catchment characteristics. Although  
34 regional constraints are important, the importance of expert knowledge on the catchment  
35 scale, which leads to better understanding of hydrological behavior is highlighted in this  
36 study.

37 In a similar attempt, Pokhrel et al. (2008, 2012) demonstrated use of regularization for model  
38 parameters and reduction of model parameter space dimensionality by linking model  
39 parameters using super-parameters to catchment characteristics. However, no explicit  
40 hydrological reasoning is typically applied for such “regularization rules” (e.g. Pokhrel et al.,  
41 2012). On the other hand, Kumar et al. (2010, 2013) parameterize and successfully  
42 regionalize their models using empirical transfer functions with global parameters, developed  
43 from extensive literature study and iterative testing in a large sample of catchments In  
44 contrast, the use of relational parameter- and process constraints, as presented in this study, is  
45 based on semi-quantitative, hydrologically explicit and meaningful reasoning avoiding the  
46 need for empirical transfer functions to link catchments characteristics and model parameters.

47 Including prior knowledge for parameters of physically-based models for estimating runoff in  
48 ungauged basins was quite successfully investigated in the past (e.g. Otte and Uhlenbrook

1 2004, Vinegradov et al., 2011, Fang et al., 2013, Semenova et al., 2013). These studies  
2 specifically indicate that calibration can be replaced by prior information which is a  
3 significant contribution to Predictions in Ungauged Basins (PUB). While physically-based  
4 models need detailed information of catchment behavior for model parameters, the here  
5 proposed semi-distributed conceptual modeling framework, exploiting relational constraints,  
6 can be more efficiently set up using the least prior information necessary. In this study, the  
7 performances and uncertainties of the three tested model set-ups for *constrained but*  
8 *uncalibrated* parameters indicate the potential of the presented FLEX-TOPO framework for  
9 Predictions in Ungauged Basins (PUB). Hence, this framework can efficiently use expert  
10 knowledge for improving model parameter value selection in complex conceptual  
11 hydrological models, not only to increase model performance but also to reduce model  
12 predictive uncertainty even in the absence of calibration.

13 It should be noted that the model set-ups suggested within the FLEX-TOPO framework are  
14 hypotheses that still need to undergo further tests, ideally confronting them with additional,  
15 system internal information, such as groundwater dynamics (e.g. Seibert and McDonnell  
16 2003, 2013; Fenicia et al., 2008) or tracer data (e.g. Madsen 2006, Campbell et al., 2012;  
17 Birkel et al., 2011; Hrachowitz et al., 2013a). To make more efficient use of relational  
18 constraints, model sensitivities to these constraints need to be evaluated in the future. It is also  
19 emphasized that the constraints introduced in this study are based on the authors' subjective  
20 understanding of catchment behavior and can and should be discussed further. However, we  
21 would like to stress the notion that reaching an agreement on the relations between parameters  
22 and fluxes in different landscape units is potentially much easier than finding the most  
23 adequate parameter values together with associated uncertainties for a conceptual model  
24 based on field observations or available data on geology or soil types.

25

## 26 **5- Conclusion:**

27

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

32 It was found that:

33 (1) A constrained but uncalibrated semi-distributed model exhibited an equivalent  
34 performance compared to a constrained and calibrated lumped model when used for  
35 prediction. This illustrates the potential value of the combined use of higher complexity  
36 models and relational constraints for predictions in ungauged basins, where no calibration  
37 data are available.

38 (2) The use of relational parameter- and process constraints in model calibration ensured a  
39 high degree of process realism. Thus, in spite of the comparatively high complexity, the  
40 overall model performance and uncertainty showed better prediction results than for a lumped  
41 model. It was shown that higher model complexity therefore does not necessarily entail  
42 reduced predictive power.

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

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

49



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

## 6- References:

- 9 Ajami, N. K., V., G. H., Wagener, T., and Sorooshian, S.: Calibration of a semi-distributed  
10 hydrologic model for streamflow estimation along a river system, *Journal of Hydrology*,  
11 298(1-4), 112–135, doi:10.1016/j.jhydrol.2004.03.033, 2004.
- 12 Anderson, A. E., Weiler, M., Alila, Y., and Hudson, R. O.: Subsurface flow velocities in a  
13 hillslope with lateral preferential flow, *Water Resources Research*, 45, W11 407,  
14 doi:10.1029/2008WR007121, 2009.
- 15 Anderson, A. E., Weiler, M., Alila, Y., and Hudson, R. O.: Piezometric response in zones of a  
16 watershed with lateral preferential flow as a first-order control on subsurface flow,  
17 *Hydrological Processes*, 24(16), 2237–2247, doi:10.1002/hyp.7662,  
18 <http://dx.doi.org/10.1002/hyp.7662>, 2010.
- 19 Andréassian, V., Perrin, C., Berthet, L., Le Moine, N., Lerat, J., Loumagne, C., Oudin, L.,  
20 Mathevet, T., Ramos, M., and Valry, A.: HESS Opinions "Crash tests for a standardized  
21 evaluation of hydrological models", *Hydrology and Earth System Sciences*, 13, 1757–  
22 1764, 2009.
- 23 Andréassian, V., Le Moine, N., Perrin, C., Ramos, M.-H., Oudin, L., Mathevet, T., Lerat, J.,  
24 and Berthet, L.: All that glitters is not gold: the case of calibrating hydrological models,  
25 *Hydrological Processes*, 26(14), 2206–2210, doi:10.1002/hyp.9264,  
26 <http://dx.doi.org/10.1002/hyp.9264>, 2012.
- 27 Arora, V. K.: The use of the aridity index to assess climate change effect on annual runoff,  
28 *Journal of Hydrology*, 265(1-4), 164–177, doi:[http://dx.doi.org/10.1016/S0022-](http://dx.doi.org/10.1016/S0022-1694(02)00101-4)  
29 [1694\(02\)00101-4](http://www.sciencedirect.com/science/article/pii/S0022169402001014), <http://www.sciencedirect.com/science/article/pii/S0022169402001014>,  
30 2002.
- 31 Bárdossy, A.: Calibration of hydrological model parameters for ungauged catchments,  
32 *Hydrology and Earth System Sciences*, 11(2), 703–710, doi:10.5194/hess-11-703-2007,  
33 <http://www.hydrol-earth-syst-sci.net/11/703/2007/>, 2007.
- 34 Beven, K.: INTERFLOW, Unsaturated Flow in Hydrologic Modeling Theory and Practice,  
35 pp. 191–219, 1989.
- 36 Beven, K.: *Rainfall-Runoff Modelling, The Primer*, John Wiley and Sons, 2001.
- 37 Beven, K.: A manifesto for the equifinality thesis, *Journal of Hydrology*, 320(1-2), 18–36,  
38 doi:10.1016/j.jhydrol.2005.07.007, 2006.
- 39 Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in  
40 mechanistic modelling of complex environmental systems using the GLUE methodology,  
41 *Journal of Hydrology*, 249(1-4), 11–29, doi:10.1016/S0022-1694(01)00421-8,  
42 <http://www.sciencedirect.com/science/article/pii/S0022169401004218>, 2001a.
- 43 Beven, K. and Freer, J.: A dynamic TOPMODEL, *Hydrological Processes*, 15(10), 1993–  
44 2011, doi:10.1002/hyp.252, <http://dx.doi.org/10.1002/hyp.252>, 2001b.
- 45 Beven, K. and Germann, P.: Macropores and water flow in soils, *Water Resources Research*,  
46 18(5), 1311–1325, doi:10.1029/WR018i005p01311,  
47 <http://dx.doi.org/10.1029/WR018i005p01311>, 1982.
- 48 Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin  
49 hydrology, *Hydrology Sciences Journal*, 24(1), 43–69, doi:10.1080/02626667909491834,  
50 1979.

- 1 Birkel, C., Dunn, S. M., Tetzlaff, D., and Soulsby, C.: Assessing the value of high-resolution  
2 isotope tracer data in the stepwise development of a lumped conceptual rainfall-runoff  
3 model, *Hydrological Processes*, 24, 2335–2348, doi:10.1002/hyp.7763, 2010.
- 4 Birkel, C., D., T., Dunn, S., and Soulsby, C.: Using time domain and geographic source  
5 tracers to conceptualize streamflow generation processes in lumped rainfall-runoff models,  
6 *Water Resource Research*, 47, W02 515, doi:10.1029/2010WR009547, 2011.
- 7 Blöschl, G.: Scaling in hydrology, *Hydrological Processes*, 15(4), 709–711,  
8 doi:10.1002/hyp.432, 2001.
- 9 Blume, T., Zehe, E., and Bronstert, A.: Investigation of runoff generation in a pristine, poorly  
10 gauged catchment in the Chilean Andes II: Qualitative and quantitative use of tracers at  
11 three spatial scales, *Hydrological Processes*, 22(18), 3676–3688, doi: 10.1002/hyp.6970,  
12 <http://dx.doi.org/10.1002/hyp.6970>, 2008.
- 13 Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic  
14 models: Combining the strengths of manual and automatic methods, *Water Resource  
15 Research*, 36(12), 3663–3674, doi:10.1029/2000WR900207, 2000.
- 16 Boyle, D. P., Gupta, H. V., Sorooshian, S., Koren, V., Zhang, Z., and Smith, M.: Toward  
17 improved streamflow forecasts: value of semidistributed modeling, *Water Resource  
18 Research*, 37, 2749–2759, doi:10.1029/2000WR000207, 2001.
- 19 Breuer, L., Eckhardt, K., and Frede, H.-G.: Plant parameter values for models in temperate  
20 climates, *Ecological Modelling*, 169, 237293, doi:10.1016/S0304-3800(03)00274-6, 2003.
- 21 Bulygina, N. and Gupta, H.: How Bayesian data assimilation can be used to estimate the  
22 mathematical structure of a model, *Stochastic Environmental Research and Risk  
23 Assessment*, 24(6), 925–937, doi:10.1007/s00477-010-0387-y,  
24 <http://dx.doi.org/10.1007/s00477-010-0387-y>, 2010.
- 25 Capell, R., Tetzlaff, D., and Soulsby, C.: Can time domain and source area tracers reduce  
26 uncertainty in rainfall-runoff models in larger heterogeneous catchments?, *Water Resources  
27 Research*, 48, W09 544, doi:10.1029/2011WR011543, 2012.
- 28 Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener,  
29 T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular  
30 framework to diagnose differences between hydrological models, *Water Resource  
31 Research*, 44, W00B02, doi:10.1029/2007WR006735, 2008.
- 32 Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-van Meerveld, H. J., Peters, N. E., and Freer,  
33 J. E.: Consistency between hydrological models and field observations: linking processes  
34 at the hillslope scale to hydrological responses at the watershed scale, *Hydrological  
35 Processes*, 23(2), 311–319, doi:10.1002/hyp.7154, <http://dx.doi.org/10.1002/hyp.7154>,  
36 2009.
- 37 Clark, M. P., Kavetski, D., and Fenicia, F.: Pursuing the method of multiple working  
38 hypotheses for hydrological modeling, *Water Resource Research*, 47, W09 301,  
39 doi:10.1029/2010WR009827, 2011.
- 40 Detty, J. M. and McGuire, K. J.: Topographic controls on shallow groundwater dynamics:  
41 implications of hydrologic connectivity between hillslopes and riparian zones in a till  
42 mantled catchment, *Hydrological Processes*, 24, 2222–2236, doi:10.1002/hyp.7656, 2010.
- 43 Drogue, G., Pfister, L., Leviandier, T., Humbert, J., Hoffmann, L., Idrissi, A. E., and Iffly, J.-  
44 F.: Using 3D dynamic cartography and hydrological modelling for linear streamflow  
45 mapping, *Computers & Geosciences*, 28(8), 981–994, doi:10.1016/S0098-  
46 3004(02)00028-6, <http://www.sciencedirect.com/science/article/pii/S0098300402000286>,  
47 2002.
- 48 Duan, Q., Schaake, J., Andrassian, V., Franks, S., Goteti, G., Gupta, H., Gusev, Y., Habets,  
49 F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O.,  
50 Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., and Wood, E.: Model Parameter

1 Estimation Experiment (MOPEX): An overview of science strategy and major results from  
2 the second and third workshops, *Journal of Hydrology*, 320(12), 3–17,  
3 doi:<http://dx.doi.org/10.1016/j.jhydrol>, 2005.07.031,  
4 <http://www.sciencedirect.com/science/article/pii/S002216940500329X>,  
5 model parameter estimation experiment; MOPEX workshop; 2006.

7 Dunn, S. M., Bacon, J. R., Soulsby, C., Tetzlaff, D., Stutter, M. I., Waldron, S., and Malcolm,  
8 I. A.: Interpretation of homogeneity in  $^{18}\text{O}$  signatures of stream water in a nested  
9 subcatchment system in north-east Scotland, *Hydrological Processes*, 22(24), 4767–4782,  
10 doi:10.1002/hyp.7088, <http://dx.doi.org/10.1002/hyp.7088>, 2008.

11 Efstratiadis, A., Nalbantis, I., Koukouvinos, A., Rozos, E., and Koutsoyiannis, D.:  
12 HYDROGEIOS: a semi-distributed GISbased hydrological model for modified river  
13 basins, *Hydrology and Earth System Sciences*, 12(4), 989–1006, doi:10.5194/hess-12-989-  
14 2008, <http://www.hydrol-earth-syst-sci.net/12/989/2008/>, 2008.

15 Engl, Heinz Werner, Martin Hanke, and Andreas Neubauer. *Regularization of inverse*  
16 *problems*. Vol. 375. Springer, 1996.

17 Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., and Savenije, H. H.  
18 G.: A framework to assess the realism of model structures using hydrological signatures,  
19 *Hydrology and Earth System Sciences*, 17(5), 1893–1912, doi:10.5194/hess-17-1893-2013,  
20 <http://www.hydrol-earth-syst-sci.net/17/1893/2013/>, 2013.

21 Fang, X., Pomeroy, J. W., Ellis, C. R., MacDonald, M. K., De-Beer, C. M., and Brown, T.:  
22 Multi-variable evaluation of hydrological model predictions for a headwater basin in the  
23 Canadian Rocky Mountains, *Hydrology and Earth System Sciences*, 17(4), 1635–1659,  
24 doi:10.5194/hess-17-1635-2013, <http://www.hydrol-earth-syst-sci.net/17/1635/2013/>,  
25 2013.

26 Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Is the groundwater reservoir  
27 linear? Learning from data in hydrological modelling, *Hydrology and Earth System*  
28 *Sciences*, 10, 139–150, 2006.

29 Fenicia, F., McDonnell, J., and Savenije, H. H. G.: Learning from model improvement: On  
30 the contribution of complementary data to process understanding, *Water Resource*  
31 *Research*, 44, W06 419, doi:10.1029/2007WR006386, 2008a.

32 Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Understanding catchment behavior  
33 through stepwise model concept improvement, *Water Resource Research*, 44, W01 402,  
34 doi:10.1029/2006WR005563, 2008b.

35 Fenicia, F., Kavetski, D., and Savenije, H. H. G.: Elements of a flexible approach for  
36 conceptual hydrological modeling: 1. Motivation and theoretical development, *Water*  
37 *Resource Research*, 47, W11 510, doi:10.1029/2010WR010174, 2011.

38 Fenicia, F., Kavetski, D., Savenije, H. H. G., Clark, M. P., Schoups, G., Pfister, L., and Freer,  
39 J.: Catchment properties, function, and conceptual model representation: is there a  
40 correspondence?, *Hydrological Processes*, pp. n/a–n/a, doi:10.1002/hyp.9726,  
41 <http://dx.doi.org/10.1002/hyp.9726>, 2013.

42 Flügel, W.-A.: Delineating hydrological response units by geographical information system  
43 analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of  
44 the River Bröl, Germany, *Hydrological Processes*, 9(3-4), 423–436,  
45 doi:10.1002/hyp.3360090313, 1995.

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

49 Freer, J., McMillan, H., McDonnell, J., and Beven, K.: Constraining dynamic TOPMODEL  
50 responses for imprecise water table information using fuzzy rule based performance

1 measures, *Journal of Hydrology*, 291, 254 – 277, doi:10.1016/j.jhydrol.2003.12.037,  
2 <http://www.sciencedirect.com/science/article/pii/S0022169404000356>, catchment  
3 modelling: Towards an improved representation of the hydrological processes in real-  
4 world model applications, 2004.

5 Freer, J.; Beven, K. & Peters, N. Multivariate seasonal period model rejection within the  
6 generalised likelihood uncertainty estimation procedure *Water Science and Application*, 6,  
7 69-87, 2003.

8 Gao, H., Hrachowitz, M., Fenicia, F., Gharari, S., & Savenije, H. H. G., Testing the realism of  
9 a topography-driven model (FLEX-Topo) in the nested catchments of the Upper Heihe,  
10 China. *Hydrology and Earth System Sciences*, 18(5), 1895-1915, 2014.

11 Gascuel-Oudou, C., Arousseau, P., Durand, P., Ruiz, L., and Molenat, J.: The role of climate  
12 on inter-annual variation in stream nitrate fluxes and concentrations, *Science of The Total  
13 Environment*, 408(23), 5657–5666, doi:<http://dx.doi.org/10.1016/j.scitotenv.2009.05.003>,  
14 <http://www.sciencedirect.com/science/article/pii/S0048969709004574>,  
15 Section: Integrating Water and Agricultural Management Under Climate Change/ce:title,  
16 2010.

17 Gharari, S., Fenicia, F., Hrachowitz, M., and Savenije, H. H. G.: Land classification based on  
18 hydrological landscape units, *Hydrology and Earth System Sciences Discussions*, 8(3),  
19 4381–4425, doi:10.5194/hessd-8-4381-2011, [http://www.hydrol-earth-syst-sci-](http://www.hydrol-earth-syst-sci-discuss.net/8/4381/2011/)  
20 [discuss.net/8/4381/2011/](http://www.hydrol-earth-syst-sci-discuss.net/8/4381/2011/), 2011.

21 Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H. G.: An approach to identify time  
22 consistent model parameters: sub-period calibration, *Hydrology and Earth System  
23 Sciences*, 17(1), 149–161, doi:10.5194/hess-17-149-2013, [http://www.hydrol-earth-syst-](http://www.hydrol-earth-syst-sci-net/17/149/2013/)  
24 [sci-net/17/149/2013/](http://www.hydrol-earth-syst-sci-net/17/149/2013/), 2013.

25 Gharari, S., Shafiei, M., Hrachowitz, M., Fenicia, F., Gupta, H., and Savenije, H. H. G.: A  
26 Strategy for Constraint-based parameter specification for Environmental Models,  
27 *Hydrology and Earth System Sciences Discussion*, XXX, XXX, doi:XXX, 2014.

28 Grayson, R. and Blöschl, G.: Spatial patterns in catchment hydrology: observations and  
29 modelling., chap. Chapter 14 Summary of pattern comparison and concluding remarks.,  
30 pp. 355–367, Cambridge: Cambridge University Press, 2000.

31 Grayson, R. B., Moore, I. D., and McMahon, T. A.: Physically based hydrologic modeling: 1.  
32 A terrain-based model for investigative purposes, *Water Resource Research*, 28, 2639-2658,  
33 doi:10.1029/92WR01258, 1992.

34 Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic  
35 models: Multiple and noncommensurable measures of information, *Water Resource  
36 Research*, 34(4), 751–763, doi:10.1029/97WR03495, 1998.

37 Gupta, H. V., Wagener, T., and Liu, Y.: Reconciling theory with observations: elements of a  
38 diagnostic approach to model evaluation, *Hydrological Processes*, 22(18), 3802–3813,  
39 doi:10.1002/hyp.6989, 2008.

40 Gupta, V. K. and Sorooshian, S.: Uniqueness and observability of conceptual rainfall-runoff  
41 model parameters: The percolation process examined, *Water Resources Research*, 19(1),  
42 269–276, doi:10.1029/WR019i001p00269, <http://dx.doi.org/10.1029/WR019i001p00269>,  
43 1983.

44 Hamon, W. R.: Estimating potential evapotranspiration, *Journal Hydraulic Division*, 87, 107–  
45 120, 1961.

46 Hewlett, J. D.: Soil moisture as a source of base flow from steep mountain watersheds, US  
47 Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, 1961.

48 Hrachowitz, M., Soulsby, C., D., T., Dawson, J. J. C., and Malcolm, I. A.: Regionalization of  
49 Transit Time Estimates in montane catchments by integrating landscape controls, *Water  
50 Resources Research*, 45, W05 421, doi:10.1029/2008WR007496, 2009.

- 1 Hrachowitz, M., Savenije, H., Bogaard, T. A., Tetzlaff, D., and Soulsby, C.: What can flux  
2 tracking teach us about water age distribution patterns and their temporal dynamics?,  
3 Hydrology and Earth System Sciences, 17(2), 533–564, doi:10.5194/hess-17-533-2013,  
4 <http://www.hydrol-earth-syst-sci.net/17/533/2013/>, 2013a.
- 5 Hrachowitz, M., Savenije, H. H. G., Blschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J.  
6 W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A.,  
7 Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.  
8 A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and  
9 Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB)a review, Hydrological  
10 Sciences Journal, 58(6), 1–58, doi:10.1080/02626667.2013.803183,  
11 <http://www.tandfonline.com/doi/abs/10.1080/02626667.2013.803183>, 2013b.
- 12 Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., and  
13 Marshall, L. A.: Hydrologic connectivity between landscapes and streams: transferring  
14 reach- and plot-scale understanding to the catchment scale, Water Resources Research, 45,  
15 W04 428, doi:10.1029/2008WR007225, 2009.
- 16 Kapangaziwiri, E., Hughes, D., and Wagener, T.: Incorporating uncertainty in hydrological  
17 predictions for gauged and ungauged basins in southern Africa, Hydrological Sciences  
18 Journal, 57(5), 1000–1019, doi:10.1080/02626667.2012.690881, <http://www.tandfonline.com/doi/abs/10.1080/02626667.2012.690881>, 2012.
- 20 Kavetski, D., Fenicia, F., and Clark, M. P.: Impact of temporal data resolution on parameter  
21 inference and model identification in conceptual hydrological modeling: Insights from an  
22 experimental catchment, Water Resource Research, 47, W05 501, doi:  
23 10.1029/2010WR009525, 2011.
- 24 Khu, S. T., Madsen, H., and de Pierro, F.: Incorporating multiple observations for distributed  
25 hydrologic model calibration: An approach using a multi-objective evolutionary algorithm  
26 and clustering, Advances in Water Resources, 31(10), 1387–1398, doi:  
27 10.1016/j.advwatres.2008.07.011, 2008.
- 28 Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements,  
29 analyses, and models to advance the science of hydrology, Water Resource Research, 42,  
30 W03S04, doi: 10.1029/2005WR004362, 2006.
- 31 Klemeš, V.: Operational testing of hydrological simulation models, Hydrological Sciences  
32 Journal, 31(1), 13–24, doi:10.1080/02626668609491024, 1986.
- 33 Kling, H. and Gupta, H.: On the development of regionalization relationships for lumped  
34 watershed models: The impact of ignoring sub-basin scale variability, Journal of  
35 Hydrology, 373(3-4), 337–351, doi:10.1016/j.jhydrol.2009.04.031,  
36 <http://www.sciencedirect.com/science/article/pii/S0022169409002893>, 2009.
- 37 Knudsen, J., Thomsen, A., and Refsgaard, J. C.: WATBAL A Semi-Distributed, Physically  
38 Based Hydrological Modelling System, Nordic Hydrology, 17(4-5), 347–362,  
39 doi:10.2166/nh.1986.026, 1986.
- 40 Koren, V., Smith, M., Wang, D., and Zhang, Z.: Use of soil property data in the derivation of  
41 conceptual rainfall-runoff model parameters, in: 15th Conference on Hydrology, Long  
42 Beach, American Meteorological Society, Paper, vol. 2, 2000.
- 43 Koren, V., Smith, M., and Duan, Q.: Use of a Prior Parameter Estimates in the Derivation of  
44 Spatially, 2003.
- 45 Krcho, J.: Modelling of georelief and its geometrical structure using DTM: positional and  
46 numerical accuracy, Q111 Vydavatel Stvo, 2001.
- 47 Kumar, R., Samaniego, L., and Attinger, S.: The effects of spatial discretization and model  
48 parameterization on the prediction of extreme runoff characteristics, Journal of Hydrology,  
49 392(1–2), 54–69, doi:10.1016/j.jhydrol. 2010.07.047,  
50 <http://www.sciencedirect.com/science/article/pii/S0022169410004865>, 2010.

- 1 Kumar, R., Samaniego, L., and Attinger, S.: Implications of distributed hydrologic model  
2 parameterization on water fluxes at multiple scales and locations, *Water Resources*  
3 *Research*, 491(1), 360–379, doi:10.1029/2012WR012195,  
4 <http://dx.doi.org/10.1029/2012WR012195>, 2013.
- 5 Kumar, R.; Livneh, B. & Samaniego, L. Toward computationally efficient large-scale  
6 hydrologic predictions with a multiscale regionalization scheme *Water Resources*  
7 *Research*, 49(9), 2013b.
- 8 Kuzmin, V., Seo, D.-J., and Koren, V.: Fast and efficient optimization of hydrologic model  
9 parameters using a priori estimates and stepwise line search, *Journal of Hydrology*, 353(1–  
10 2), 109–128, doi:<http://dx.doi.org/10.1016/j.jhydrol.2008.02.001>,  
11 <http://www.sciencedirect.com/science/article/pii/S0022169408000668>, 2008.
- 12 Laaha, G., Skien, J., and Blöschl, G.: Spatial prediction on river networks: comparison of top-  
13 kriging with regional regression, *Hydrological Processes*, pp. n/a–n/a,  
14 doi:10.1002/hyp.9578, <http://dx.doi.org/10.1002/hyp.9578>, 2012.
- 15 Lindström, G., Pers, C., Rosberg, J., Strmqvist, J., and Arheimer, B.: Development and testing  
16 of the HYPE (Hydrological Predictions for the Environment) water quality model for  
17 different spatial scales, *Hydrology Research*, 41(3-4), 295–319, doi:10.2166/nh.2010.007,  
18 2010.
- 19 Madsen, H.: Automatic calibration of a conceptual rainfallrunoff model using multiple  
20 objectives, *Journal of Hydrology*, 235(3-4), 276–288, doi:10.1016/S0022-1694(00)00279-  
21 1, <http://www.sciencedirect.com/science/article/pii/S0022169400002791>, 2000.
- 22 Matgen, P., Fenicia, F., Heitz, S., Plaza, D., de Keyser, R., Pauwels, V. R., Wagner, W., and  
23 Savenije, H.: Can ASCAT derived soil wetness indices reduce predictive uncertainty in  
24 well-gauged areas? A comparison with in situ observed soil moisture in an assimilation  
25 application, *Advances in Water Resources*, 44(0), 49–65,  
26 doi:<http://dx.doi.org/10.1016/j.advwatres.2012.03.022>,  
27 <http://www.sciencedirect.com/science/article/pii/S0309170812000772>, 2012.
- 28 McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C.,  
29 Hooper, R., Kirchner, J., Roderick, M. L., Selker, J., and Weiler, M.: Moving beyond  
30 heterogeneity and process complexity: A new vision for watershed hydrology, *Water*  
31 *Resource Research*, 43, W07 301, doi:10.1029/2006WR005467, 2007.
- 32 McGlynn, B. L. and McDonnell, J. J.: Quantifying the relative contributions of riparian and  
33 hillslopezones to catchment runoff, *Water Resource Research*, 39(11), 1310,  
34 doi:10.1029/2003WR002091, 2003.
- 35 McGlynn, B. L., McDonnell, J. J., Seibert, J., and Kendall, C.: Scale effects on headwater  
36 catchment runoff timing, flow sources, and groundwater-streamflow relations, *Water*  
37 *Resource Research*, 40, W07 504, doi:10.1029/2003WR002494, 2004.
- 38 McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M., and  
39 Seibert, J.: The role of topography on catchment-scale water residence time, *Water*  
40 *Resources Research*, 41, W05 002, doi:10.1029/2004WR003657, 2005.
- 41 McMillan, H. K., Clark, M. P., Bowden, W. B., Duncan, M., and Woods, R. A.: Hydrological  
42 field data from a modeller's perspective: Part 1. Diagnostic tests for model structure,  
43 *Hydrological Processes*, 25(4), 511–522, doi:10.1002/hyp.7841,  
44 <http://dx.doi.org/10.1002/hyp.7841>, 2011.
- 45 McNamara, J. P., Chandler, D., Seyfried, M., and Achet, S.: Soil moisture states, lateral flow,  
46 and streamflow generation in a semi-arid, snowmelt-driven catchment, *Hydrological*  
47 *Processes*, 19(20), 4023–4038, doi:10.1002/hyp.5869, <http://dx.doi.org/10.1002/hyp.5869>,  
48 2005.
- 49 Merz, R. and Blöschl, G.: Regionalisation of catchment model parameters, *Journal of*  
50 *Hydrology*, 287, 95123, 2004.

- 1 Milne, G.: Some suggested units of classification and mapping particularly for East African  
2 soils, *Soil Research*, 4(3), 183–198, 1935.
- 3 Molnat, J., Gascuel-Oudou, C., Davy, P., and Durand, P.: How to model shallow water-table  
4 depth variations: the case of the Kervidy-Naizin catchment, France, *Hydrological  
5 Processes*, 19(4), 901–920, doi:10.1002/hyp.5546, <http://dx.doi.org/10.1002/hyp.5546>,  
6 2005.
- 7 Nalbantis, I., Efstratiadis, A., Rozos, E., Kopsiafti, M., and Koutsoyiannis, D.: Holistic versus  
8 monomeric strategies for hydrological modelling of human-modified hydrosystems,  
9 *Hydrology and Earth System Sciences*, 15, 743758, doi:10.5194/hess-15-743-2011, 2011.
- 10 Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A  
11 discussion of principles, *Journal of Hydrology*, 10(3), 282 – 290, doi:DOI:10.1016/0022-  
12 1694(70)90255-6, <http://www.sciencedirect.com/science/article/pii/0022169470902556>,  
13 1970.
- 14 Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A.,  
15 Waterloo, M., and Saleska, S.: Height above the Nearest Drainage, a hydrologically  
16 relevant new terrain model, *Journal of Hydrology*, 404(1-2), 13–29, 2011.
- 17 Ott, B. and Uhlenbrook, S.: Quantifying the impact of land-use changes at the event and  
18 seasonal time scale using a processoriented catchment model, *Hydrology and Earth System  
19 Sciences*, 8(1), 62–78, 2004.
- 20 Oudin, L., Andrassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity,  
21 physical similarity, regression and ungaged catchments: A comparison of regionalization  
22 approaches based on 913 French catchments, *Water Resource Research*, 44, W03 413,  
23 doi:10.1029/2007WR006240, 2008.
- 24 Parajka, J. and Blöschl, G.: Spatio-temporal combination of MODIS images potential for  
25 snow cover mapping, *Water Resources Research*, 44(3), n/a–n/a,  
26 doi:10.1029/2007WR006204, <http://dx.doi.org/10.1029/2007WR006204>, 2008.
- 27 Parajka, J., Merz, R., and Blöschl, G.: Uncertainty and multiple objective calibration in  
28 regional water balance modelling: case study in 320 Austrian catchments, *Hydrological  
29 Processes*, 21(4), 435–446, doi:10.1002/hyp.6253, 2007.
- 30 Park, S. and van de Giesen, N.: Soil-landscape delineation to define spatial sampling domains  
31 for hillslope hydrology, *Journal of Hydrology*, 295(1-4), 28–46,  
32 doi:10.1016/j.jhydrol.2004.02.022, 2004.
- 33 Perrin, C., Andréassian, V., Serna, C. R., Mathevet, T., and Moine, N. L.: Discrete  
34 parameterization of hydrological models: Evaluating the use of parameter sets libraries  
35 over 900 catchments, *Water Resources Research*, 44, W08 447, doi:10.1029/  
36 2007WR006579, 2008.
- 37 Pike, J.: The estimation of annual run-off from meteorological data in a tropical climate,  
38 *Journal of Hydrology*, 2(2), 116–123, doi:[http://dx.doi.org/10.1016/0022-1694\(64\)90022-](http://dx.doi.org/10.1016/0022-1694(64)90022-8)  
39 8, <http://www.sciencedirect.com/science/article/pii/0022169464900228>, 1964.
- 40 Pokhrel, P., Gupta, H. V., and Wagener, T.: A spatial regularization approach to parameter  
41 estimation for a distributed watershed model, *Water Resources Research*, 44(12), n/a–n/a,  
42 doi:10.1029/2007WR006615, <http://dx.doi.org/10.1029/2007WR006615>, 2008.
- 43 Pokhrel, P., Yilmaz, K. K., and Gupta, H. V.: Multiplecriteria calibration of a distributed  
44 watershed model using spatial regularization and response signatures, *Journal of  
45 Hydrology*, 418419(0), 49–60, doi:<http://dx.doi.org/10.1016/j.jhydrol.2008.12.004>,  
46 <http://www.sciencedirect.com/science/article/pii/S0022169408005970>, <http://www.sciencedirect.com/science/article/pii/S0022169408005970>,  
47 *The Distributed Model Intercomparison Project (DMIP) - Phase 2 Experiments in the  
48 Oklahoma Region, USA*, 2012.
- 49 Pomeroy, J. W., Gray, D. M., Brown, T., Hedstrom, N. R., Quinton, W. L., Granger, R. J.,  
50 and Carey, S. K.: The cold regions hydrological model: a platform for basing process



1 representation and model structure on physical evidence, *Hydrological Processes*, 21(19),  
2 2650–2667, doi:10.1002/hyp.6787, <http://dx.doi.org/10.1002/hyp.6787>, 2007.

3 Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., and Participants, D.:  
4 Overall distributed model intercomparison project results, *Journal of Hydrology*, 298(1-4),  
5 27–60, doi:<http://dx.doi.org/10.1016/j.jhydrol.2004.03.031>,  
6 <http://www.sciencedirect.com/science/article/pii/S0022169404002380>, *ce:title*The  
7 Distributed Model Intercomparison Project (DMIP)*ce:title*, 2004.

8 Refsgaard, J. C. and Knudsen, J.: Operational Validation and Intercomparison of Different  
9 Types of Hydrological Models, *Water Resource Research*, 32(7), 2189–2202,  
10 doi:10.1029/96WR00896, 1996.

11 Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., and  
12 Waterloo, M. J.: HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme  
13 rainforest environments in Amazonia, *Remote Sensing of Environment*, 112(9), 3469–  
14 3481, doi:10.1016/j.rse.2008.03.018, 2008.

15 Rouhani, H., Willems, P., Wyseure, G., and Feyen, J.: Parameter estimation in semi-  
16 distributed hydrological catchment modeling using a multi-criteria objective function,  
17 *Hydrological Processes*, 21(22), 2998–3008, doi:10.1002/hyp.6527,  
18 <http://dx.doi.org/10.1002/hyp.6527>, 2007.

19 Samaniego, L., Bárdossy, A., and Kumar, R.: Streamflow prediction in ungauged catchments  
20 using copula-based dissimilarity measures, *Water Resource Research*, 46, W02 506,  
21 doi:10.1029/2008WR007695, 2010.

22 Savenije, H. H. G.: HESS Opinions ”Topography driven conceptual modelling (FLEX-  
23 Topo)”, *Hydrology and Earth System Sciences*, 14, 2681–2692, doi:10.5194/hess-14-2681-  
24 2010, <http://www.hydrol-earth-syst-sci.net/14/2681/2010/>, 2010.

25 Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A., and Carrillo, G.: Catchment  
26 classification: empirical analysis of hydrologic similarity based on catchment function in  
27 the eastern USA, *Hydrology and Earth System Sciences*, 15(9), 2895–2911,  
28 doi:10.5194/hess-15-2895-2011, <http://www.hydrol-earth-syst-sci.net/15/2895/2011/>,  
29 2011.

30 Scherrer, S. and Naef, F.: A decision scheme to indicate dominant hydrological flow  
31 processes on temperate grassland, *Hydrological Processes*, 17(2), 391–401,  
32 doi:10.1002/hyp.1131, <http://dx.doi.org/10.1002/hyp.1131>, 2003.

33 Scherrer, S., Naef, F., Faeh, A. O., and Cordery, I.: Formation of runoff at the hillslope scale  
34 during intense precipitation, *Hydrology and Earth System Sciences*, 11(2), 907–922, 2007.

35 Schmockler-Fackel, P., Naef, F., and Scherrer, S.: Identifying runoff processes on the plot and  
36 catchment scale, *Hydrology and Earth System Sciences*, 11, 891–906, 2007.

37 Seibert, J.: Reliability of model predictions outside calibration conditions, *Nordic Hydrology*,  
38 34(5), 477–492, 2003.

39 Seibert, J. and Beven, K. J.: Gauging the ungauged basin: how many discharge measurements  
40 are needed?, *Hydrology and Earth System Sciences*, 13(6), 883–892, doi:10.5194/hess-13-  
41 883-2009, <http://www.hydrol-earth-syst-sci.net/13/883/2009/>, 2009.

42 Seibert, J., Bishop, K., Rodhe, A., and McDonnell, J. J.: Groundwater dynamics along a  
43 hillslope: A test of the steady state hypothesis, *Water Resources Research*, 39, 1014,  
44 doi:10.1029/2002WR001404, 2003.

45 Seibert, J. & McDonnell, J. J. On the dialog between experimentalist and modeler in  
46 catchment hydrology: Use of soft data for multicriteria model calibration *Water Resource*  
47 *Research*, 38(11), 1241, 2002.

48 Seibert, J. & McDonnell, J. Gauging the Ungauged Basin: Relative Value of Soft and Hard  
49 Data *Journal of Hydrologic Engineering*, 2013, 0, A4014004

- 1 Semenova, O., Lebedeva, L., and Vinogradov, Y.: Simulation of subsurface heat and water  
2 dynamics, and runoff generation in mountainous permafrost conditions, in the Upper  
3 Kolyma River basin, Russia, *Hydrogeology Journal*, 21(1),107–119, doi:10.1007/s10040-  
4 012-0936-1, <http://dx.doi.org/10.1007/s10040-012-0936-1>, 2013.
- 5 Sidle, R. C., Noguchi, S., Tsuboyama, Y., and Laursen, K.: A conceptual model of  
6 preferential flow systems in forested hillslopes: evidence of self-organization,  
7 *Hydrological Processes*, 15(10), 1675–1692, doi:10.1002/hyp.233,  
8 <http://dx.doi.org/10.1002/hyp.233>, 2001.
- 9 Singh, S. K. & Bárdossy, A.: Calibration of hydrological models on hydrologically unusual  
10 events, *Advances in Water Resources*, 38(0), 81-91, doi: 10.1016/j.advwatres.2011.12.006,  
11 2012.
- 12 Sivapalan, M.: Pattern, Process and Function: Elements of a Unified Theory of Hydrology at  
13 the Catchment Scale, John Wiley & Sons, Ltd, doi:10.1002/0470848944.hsa012,  
14 <http://dx.doi.org/10.1002/0470848944.hsa012>, 2006.
- 15 Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R.: Downward approach to hydrologic  
16 prediction, *Hydrology Processes*, 17(11), 2101–2111, doi:10.1002/hyp.1425, 2003.
- 17 Son, K. and Sivapalan, M.: Improving model structure and reducing parameter uncertainty in  
18 conceptual water balance models through the use of auxiliary data, *Water Resource*  
19 *Research*, 43, W01 415, doi:10.1029/2006WR005032, 2007.
- 20 Spence, C.: A Paradigm Shift in Hydrology: Storage Thresholds Across Scales Influence  
21 Catchment Runoff Generation, *Geography Compass*, 4(7), 819–833, doi:10.1111/j.1749-  
22 8198.2010.00341.x, <http://dx.doi.org/10.1111/j.1749-8198.2010.00341.x>, 2010.
- 23 Sutanudjaja, E., de Jong, S., van Geer, F., and Bierkens, M.: Using fERSg spaceborne  
24 microwave soil moisture observations to predict groundwater head in space and time,  
25 *Remote Sensing of Environment*, 138(0), 172 – 188,  
26 doi:<http://dx.doi.org/10.1016/j.rse.2013.07.022>,  
27 <http://www.sciencedirect.com/science/article/pii/S0034425713002344>, 2013.
- 28 Tikhonov, Andrey. "Solution of incorrectly formulated problems and the regularization  
29 method." *Soviet Math. Dokl.* Vol. 5. 1963.
- 30 Razavi, S. and Tolson, B. A.: An efficient framework for hydrologic model calibration on  
31 long data periods *Water Resources Research*, 49(12), 8418-8431, doi:  
32 10.1002/2012WR013442, 2013.
- 33 Tromp-van Meerveld, H. J. and McDonnell, J. J.: Threshold relations in subsurface  
34 stormflow: 2. The fill and spill hypothesis, *Water Resources Research*, 42, W02 411,  
35 doi:10.1029/2004WR003800, 2006a.
- 36 Tromp-van Meerveld, H. J. and McDonnell, J. J.: Threshold relations in subsurface  
37 stormflow: 1. A 147-storm analysis of the Panola hillslope, *Water Resources Research*, 42,  
38 W02 410, doi:10.1029/2004WR003778, 2006b.
- 39 Turc, L.: Le bilan deau des sols. Relation entre la precipitation, levaporation et lecoulement,  
40 *Ann. Agron.* 5, 1954.
- 41 Uhlenbrook, S., Roser, S., and Tilch, N.: Hydrological process representation at the meso-  
42 scale: the potential of a distributed, conceptual catchment model, *Journal of Hydrology*,  
43 291(3-4), 278–296, doi:10.1016/j.jhydrol.2003.12.038, 2004.
- 44 Vaché, K. and McDonnell, J.: A process-based rejectionist framework for evaluating  
45 catchment runoff model structure, *Water Resource Research*, 42, W02 409,  
46 doi:10.1029/2005WR004247, 2006.
- 47 Vannamettee, E.; Babel, L.; Hendriks, M.; Schuur, J.; de Jong, S.; Bierkens, M. and  
48 Karssenberg, D.: Semi-automated mapping of landforms using multiple point geostatistics,  
49 *Geomorphology*, 2014.

- 1 Viglione, A., Parajka, J., Rogger, M., Salinas, J. L., Laaha, G., Sivapalan, M., and Blöschl,  
2 G.: Comparative assessment of predictions in ungauged basins &ndash; Part 3: Runoff  
3 signatures in Austria, *Hydrology and Earth System Sciences*, 17(6), 2263–2279,  
4 doi:10.5194/hess-17-2263-2013, <http://www.hydrol-earth-syst-sci.net/17/2263/2013/>,  
5 2013.
- 6 Vinogradov, Y. B., Semanova, O. M., and Vinogradova, T. A.: An approach to the scaling  
7 problem in hydrological modelling: the deterministic modelling hydrological system,  
8 *Hydrological Processes*, 25(7), 1055–1073, doi:10.1002/hyp.7901,  
9 <http://dx.doi.org/10.1002/hyp.7901>, 2011.
- 10 Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and  
11 efficient algorithm for multiobjective optimization of hydrologic models, *Water Resource*  
12 *Research*, 39(8), 1214, doi:10.1029/2002WR001746, 2003.
- 13 Wagener, T. and Montanari, A.: Convergence of approaches toward reducing uncertainty in  
14 predictions in ungauged basins, *Water Resources Research*, 47(6), n/a–n/a,  
15 doi:10.1029/2010WR009469, <http://dx.doi.org/10.1029/2010WR009469>, 2011.
- 16 Wagener, T. and Wheater, H. S.: Parameter estimation and regionalization for continuous  
17 rainfall-runoff models including uncertainty, *Journal of Hydrology*, 320(1–2), 132–154,  
18 doi:<http://dx.doi.org/10.1016/j.jhydrol.2005.07.015>,  
19 <http://www.sciencedirect.com/science/article/pii/S0022169405003410>, *ice:title*ꞤThe model  
20 parameter estimation experiment*/ice:title*Ꞥ*ice:subtitle*ꞤMOPEX*/ice:subtitle*Ꞥ *xocs:full-*  
21 *name*ꞤMOPEX workshop*/xocs:full-name*Ꞥ, 2006.
- 22 Weiler, M., McGlynn, B., McGuire, K., and McDonnell, J.: How does rainfall become  
23 runoff? A combined tracer and runoff transfer function approach, *Water Resource*  
24 *Research*, 39, 1315, doi:10.1029/2003WR002331, 2003.
- 25 Winsemius, H. C., Savenije, H. H. G., and Bastiaanssen, W. G. M.: Constraining model  
26 parameters on remotely sensed evaporation: justification for distribution in ungauged  
27 basins?, *Hydrology and Earth System Sciences*, 12(6), 1403–1413, doi:10.5194/hess-12-  
28 1403-2008, <http://www.hydrol-earth-syst-sci.net/12/1403/2008/>, 2008.
- 29 Winsemius, H. C., Schaeffli, B., Montanari, A., and Savenije, H. H. G.: On the calibration of  
30 hydrological models in ungauged basins: A framework for integrating hard and soft  
31 hydrological information, *Water Resources Research*, 45(12), n/a–n/a,  
32 doi:10.1029/2009WR007706, <http://dx.doi.org/10.1029/2009WR007706>, 2009.
- 33 Winter, T. C.: The Concept OF Hydrologic Landscapes, *Journal of the American Water*  
34 *Resources Association*, 37(2), 335–349, doi:10.1111/j.1752-1688.2001.tb00973.x, 2001.
- 35 Wolock, D. M., Winter, T. C., and McMahon, G.: Delineation and evaluation of hydrologic-  
36 landscape regions in the United States using geographic information system tools and  
37 multivariate statistical analyses, *Environmental Management*, 34(1), S71–S88,  
38 doi:10.1007/s00267-003-5077-9, 2004.
- 39 Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed  
40 response behavior for improved predictions in ungauged basins, *Advances in Water*  
41 *Resources*, 30(8), 1756–1774, doi:10.1016/j.advwatres.2007.01.005,  
42 <http://www.sciencedirect.com/science/article/pii/S0309170807000140>, 2007.
- 43 Yilmaz, K. K., Gupta, H. V., and Wagener, T.: A process-based diagnostic approach to model  
44 evaluation: Application to the NWS distributed hydrologic model, *Water Resources*  
45 *Research*, 44(9), n/a–n/a, doi:10.1029/2007WR006716,  
46 <http://dx.doi.org/10.1029/2007WR006716>, 2008.
- 47 Zehe, E. and Sivapalan, M.: Threshold behaviour in hydrological systems as (human) geo-  
48 ecosystems: manifestations, controls, implications, *Hydrology and Earth System Sciences*,  
49 13(7), 1273–1297, doi:10.5194/hess-13-1273-2009, [http://www.hydrol-earth-syst-](http://www.hydrol-earth-syst-sci.net/13/1273/2009/)  
50 [sci.net/13/1273/2009/](http://www.hydrol-earth-syst-sci.net/13/1273/2009/), 2009.

1 Zhang, Z., Wagener, T., Reed, P., and Bhushan, R.: Reducing uncertainty in predictions in  
2 ungauged basins by combining hydrologic indices regionalization and multiobjective  
3 optimization, *Water Resources Research*, 44(12), n/a– n/a, doi:10.1029/2008WR006833,  
4 <http://dx.doi.org/10.1029/2008WR006833>, 2008.