1 **A "constraint-based" strategy for parameter specification of environmental models**
2
3 S. Gharari[1,2], M. Shafiei[3,1], M. Hrachowitz[1], R. Kumar[4], F. Fenicia[5,1], H.V. Gupta[6], H. H. G.
4 Savenije[1]
5 1- Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University
6 of Technology, Delft, the Netherlands
7 2-Public Research Center-Gabriel Lippmann, Belvaux, Luxembourg
8 3-Department of Water Engineering, Faculty of Agriculture, Ferdowsi University of
9 Mashhad, Mashhad, Iran
10 4- UFZ - Helmholtz Centre for Environmental Research, Leipzig, Germany
11 5-Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf,
12 Switzerland
13 6-Department of Hydrology and Water Resources, the University of Arizona, Tucson,
14 Arizona, USA
15
16 **Abstract:**
17
18 Many environmental systems models, such as conceptual rainfall-runoff models, rely on
19 model calibration for parameter identification. For this, an observed output time series (such
20 as runoff) is needed, but frequently not available (e.g. when making predictions in ungauged
21 basins). In this study, we provide an alternative approach for parameter identification using
22 constraints based on two types of restrictions derived from prior (or expert) knowledge. The
23 first, called "*parameter constraints*", restricts the solution space based on realistic
24 relationships that must hold between the different model parameters while the second, called
25 "*process constraints*" requires that additional realism relationships between the fluxes and
26 state variables must be satisfied. Specifically, we propose a strategy for finding parameter sets
27 that simultaneously satisfy such constraints, based on stepwise sampling of the parameter
28 space. Such parameter sets have the desirable property of being consistent with the modeler's
29 intuition of how the catchment functions, and can (if necessary) serve as prior information for
30 further investigations by reducing the prior uncertainties associated with both calibration and
31 prediction.
32
33 **1 Introduction:**
34
35 Environmental systems models, such as conceptual rainfall-runoff (CRR) models, are abstract
36 simplifications of real system behavior. Often, the parameters in such models cannot be
37 specified through direct measurements of physical properties of the system. Further, even
38 when a parameter is related to measurable quantities, its value in the model typically
39 represents an integrated value over a much larger scale than the measurement scale. For this
40 reason, such models typically rely upon calibration (tuning to match system input-output
41 behavior for some historical data period) to ensure satisfactory performance when applied to
42 specific hydrological systems of interest (*Wheater et al., 1993; Beven 2001).*
43 In the case of CRR, parameter values are typically specified through a process of calibration
44 that seeks to match the model runoff simulations to observed hydrographs by the use of an
45 objective function (e.g. root mean square error, RMSE). Expert knowledge is brought to bear
46 implicitly, by the prior selection of the model and the specification of parameter ranges that
47 define the prior parameter space. Recently, several studies have tested strategies that relate the
48 parameters of CRR models to catchment physical characteristics (*Koren et al., 2000, 2003;*
49 *Anderson et al., 2006; Yadav et al., 2007; Pokhrel et al., 2008, 2012; Kling and Gupta, 2009;*
50 *Hrachowitz et al., 2013b).* The general picture that emerges from these studies is that

exploiting expert knowledge (by somehow imposing more rigorous constraints on the parameters) has the potential to result in more realistic models (*Martinez and Gupta 2011*) that in several cases has practical benefits.

For example, *Pokhrel et al. (2008, 2012)* linked the parameters of a spatially-distributed model to catchment physical characteristics via a set of regularization (or regionalization) relationships, and thereby converting the original high-dimensional parameter space to a much reduced set of "super-parameters", thereby leading to a dramatically simplification for a calibration problem. Similarly, *Merz and Blöschl (2004) and Kling and Gupta (2009)* and *Yadav et al., (2007)*, amongst others, investigated explicit links between catchment characteristics and the parameters of a simple lumped conceptual model; they concluded, however, that such relationships are difficult to establish and may not be often possible given the available data. Other studies have used a comparison of catchment characteristics based on similarities between catchment responses to constrain parameter values. For example, *Zhang et al. (2008),* imposed a set of three constraints to infer the runoff characteristics of catchments*,* following the concept of regionalization of hydrological signatures (initially developed by *Yadav et al. 2007)*. Recently, *Kapangaziwiri et al. (2012)* constrained the Pitman monthly rainfall runoff model (*Hughes et al. 2006*) based on a regionalization of runoff signatures. *Perrin et al. (2008)* proposed a method called *discrete parameterization* based on the use of parameter sets compiled *a priori* via calibration to other catchments. Their approach "*abandon(s) the idea of searching for an optimum parameter set in the continuous, n-dimensional parameter space*" and instead "*limit(s) the calibration process to a search within a finite collection (a library) of predefined parameter sets*". More recently, *Samaniego et al. (2010)* and *Kumar et al. (2010, 2013)* demonstrated that a multi-scale approach to parameter regionalization can provide consistent model performance for both gauged and ungauged catchments.

In a complementary direction, the use of multiple objective functions or multiple system responses for calibration (*Gupta et al 1998*) has been shown to result in more realistic parameter sets that achieve improved simulations of system dynamics. The multi-objective approach seeks to identify parameter sets that simultaneously provide "*optimal*" performance for different aspects of system response (*Gupta et al. 1998; Boyle et al. 2000, 2001*). This can include constraining the model to reproduce multiple system fluxes and state variables such as runoff, evaporation, groundwater levels or tracer concentrations (e.g. *Gupta et al., 1999; Bastidas et al 1999; Freer et al., 2002; Seibert and McDonnell 2002; Khu ans Madsen., 2005; Fenicia et al., 2008; Winsemius et al., 2008; Birkel et al., 2011; Hrachowitz et al., 2013a; Seibert and McDonnell 2013*).

While the aforementioned studies have demonstrated that incorporation of expert- and *a priori* knowledge can help improve the realism of models, to our knowledge, no systematic strategy has been presented in the literature for constraining the model parameters to be consistent with the patchy understanding of a modeler regarding how the real system might work. Part of the difficulty in doing this is that expert knowledge may not directly translate to quantifiable relationships (e.g. between catchment physical characteristics and model parameters) rather, it may consist of conceptual understanding about consistency relationships that must exist among modeled state variables and/or fluxes, as well as, among various model parameters. For example, the geology of a given catchment may suggest that the catchment response during intense rainfall events is characterized by a slow responding groundwater component accompanied by fast responding Hortonian overland flow. In this case, any model result that imply peak flows are composed of a strong groundwater response should be discarded or should be given low importance.

An example of such an approach toward modeling was mentioned by *Götzinger and Bárdossy (2007),* where they impose the Lipschitz and monotonic conditions to avoid the abrupt jump

in soil moisture values for the neighboring cells of a distributed model based on the physical premises that such jumps are numerical artifacts and hydrologically unrealistic. Such kinds of information, which are not explicitly provided during model calibration, act as constraints to limit the feasible extent of the model parameter space, thus resulting in physically more meaningful model simulations. As pointed out by *Efesteridiatis et al. (2010)* "*It also offers a means to partially handle the huge uncertainty resulting from the complexity of model parameterizations in contrast to data scarcity, which is a global engineering problem that is getting increasingly severe. Actual research should provide more guidance on the effective combination of statistical and expert-based evaluation procedures.*"

In this study, we present a "*constraint-based*" strategy to limit the feasible parameter space of a conceptual hydrologic model, based on relational constraints inferred from expert knowledge regarding plausible catchment behavior. The approach is applicable to both lumped/semi-distributed and spatially distributed catchment models.

2 Constraints in environmental models:

In the companion paper, *Gharari et al. (2014a)*, we test the importance of constraining the feasible parameter space. The case study focuses on the meso-scale catchment of the Wark in Luxembourg covering an area of approximately 82 km^2. Three different models with varying (spatial) complexity are developed to capture hydrological processes across different delineated landscapes, wetland, hillslope and plateau, following the work by *Gharari et al. (2011)*. The first model, FLEX^A, is a lumped model which considers the whole catchment as a single unit. A more complex model, FLEX^B, takes into account wetlands (i.e., riparian zones), and considers the rest of the catchment as another unit. The most complex model of all, FLEX^C, distinguishes between the aforementioned landscape classes (wetland, hillslope and plateau).

A set of parameter and process constraints is defined for each of the three landscape entities, based on the expert knowledge. Parameter constraints are considered to be *a priori* because they can be imposed without actually running the model, while process constraints can only be imposed after a model is run with selected parameter sets. The number of constraints may vary from model to model, and here depends on the model complexity. For example all of the pre-defined constraints (see the following section for more detail) can be applied to the most complex model, FLEX^C, while only a subset of constraints can be applied to FLEX^B and just few constraints to FLEX^A. Using the proposed search method, explained later in this study, we define the behavioral set of parameters that satisfy all the parameter and process constraints. These resulting parameter sets are termed as "*constrained but un-calibrated*". The performance of the "*constrained but un-calibrated*" parameter sets for the three models could then be evaluated based on efficiency measures of Nash-Sutcliffe efficiency, E_{NS} (Nash and Sutcliffe, 1970). Figure 1 illustrates the 95% uncertainty interval of the model simulations generated using the "*constrained but un-calibrated*" parameter sets, and reports the median Nash-Sutcliffe efficiency ($E_{NS,median}$) for three model structures.
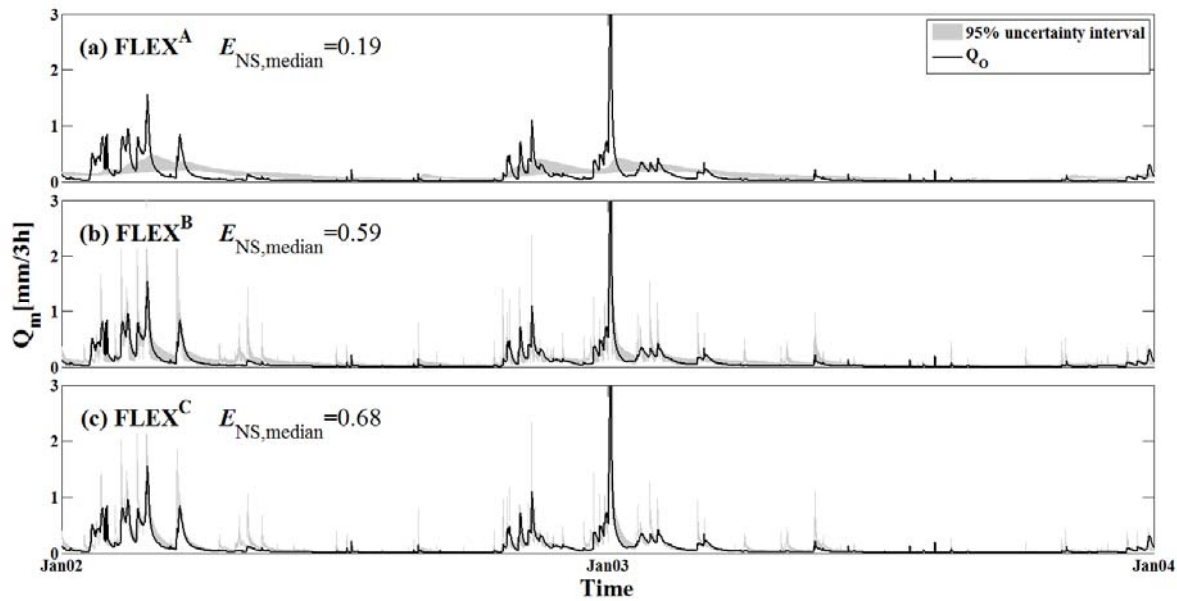
Figure 1- The observed hydrograph and the 95% uncertainty interval of the modeled hydrograph derived from the complete set of *constrained but un-calibrated* parameter sets for the three different model set-ups (**a**) FLEX[A], (**b**) FLEX[B] and (**c**) FLEX[C] for the two years (2002-2003) of calibration period. The median performance of the un-calibrated but constrained models based on Nash-Sutcliffe efficiency ($E_{NS,median}$) are reported.

The results show that the constrained and calibrated FLEX[C], the most complex model, has lower predictive uncertainty compared to a (loosely) constrained and calibrated lumped model, FLEX[A]. This result contradicts to what appears to be a general belief about the higher predictive uncertainty of complex models, and shows that if a complex model is properly constrained then it can perform well during the evaluation period while providing even lower predictive uncertainty than a simpler model. Result of the recent study by *Gao et al. (2014)* had shown that a more complex, but simultaneously constrained and calibrated model, can provide better predictions of extreme events outside of the calibration period as well as capturing the flow duration curve in nested catchments for the large scale Heihe basin in China. Meanwhile, *Hrachowitz et al. (in review)* reported similar results for a small-scale catchment in France, using 11 different model structures and 20 different runoff signatures. In these studies, the introduction of constraints was found to be crucial for ensuring better model performance, particularly outside of a calibration period.

As mentioned earlier, constraints on a model consists of two main types, *a priori* constraints applicable to model parameters (i.e. *parameter constraints)* and *a posteriori* constraints on model states and fluxes (i.e. *process constraints;* e.g., see *Bulygina and Gupta 2009, 2010, 2011*). These two types of constrains are elaborated in the following section.

2.1 Parameter constraints

Parameter constraints provide information regarding the relationships between parameters of the same process that correspond to different spatial components (or units or grid cells) of a (semi-) distributed model. Such constraints can be expressed by equality or inequality constraints; for example:

$$\frac{A_1}{A_2} < G, \text{ Where } G \text{ is a dimensionless constant} \quad (1)$$

$$A_1 < A_2 + G, \text{ Where } G \text{ has the same units as } A_1 \text{ and } A_2 \quad (2)$$

$$A_1 B_1 < A_2 B_2 \quad (3)$$

As a simple illustration of this concept, the maximum interception capacity of a forested area ($I_{max,forest}$) can typically be assumed to be larger than the maximum interception capacity of a grassland area ($I_{max,grass}$).

2.2. Process constraints

Process constraints provide comparative information regarding the fluxes and/or states of a model at each time step, or integrated over some specific time period. Examples of such constraints include:

$$\frac{\int_{t_1}^{t_2} F_1 \mathrm{d}t}{\int_{t_1}^{t_2} F_2 \mathrm{d}t} < G \quad (4)$$

$$\frac{F_1(t)}{F_2(t)} < G \quad (5)$$

Where G is a dimensionless constant.

As an illustration, one can compare the transpiration fluxes from different spatial entities of a (semi-) distributed model. For example for two regions having similar soil type and aspect, the region with smaller normalized difference vegetation index (NDVI) is expected to transpire at a lower rate.

It is worth noting that in either case, parameter sets that satisfy the constraints are not conditional on the information provided by *observations* (or measurements) of the output response of the system (e.g. the runoff hydrograph), and these can therefore be determined without resorting to model calibration. Moreover, parameter sets that satisfy all of the constraints (within some acceptable range) can provide insights into how the real system can be expected to behave, assuming that it corresponds to the expert's perception of realistic (behavioral) system properties and dynamics.

Unfortunately, the use of available evolutionary algorithms to search for parameter sets that satisfy such constraints is complicated by the non-convex and potentially non-continuous parameter search space that results. In the following section, we propose a stepwise search strategy that can be used to identify behavioral parameter sets fulfilling the constraints imposed by expert knowledge.

3 Methodology and algorithm - Constraint Based Search (CBS):

The Constraint Based Search (CBS) method is based on a simple parameter sampling approach to identify the parts of the feasible parameter space that satisfy the set of constraints as discussed in the previous section. At each step, the algorithm tries to generate new parameter sets that satisfy the parameter constraints, while only violating the process constraints to an '*acceptable*' level. This level could be set-up based on the desired model performance. The process of search continues until all of the parameter sets generated properly satisfy the set of imposed process constraints.

In the following description M refers to the total number of <u>*process*</u> constraints and *m,* $m \in \{1,\dots, M\}$, is an index indicating how many of the process constraints are satisfied by a given parameter set; for example, if a parameter set satisfies two process constraints then *m* will be equal to 2. The algorithm ultimately generates a set P of *n* number of *behavioral*

parameter sets that satisfy all of the parameter and process constraints (i.e. $M=m$ for all of members of P).

The search method is as follows; let C be a counter which identifies the minimum number of process constraints to be satisfied (at a given stage of the search) by the parameter sets considered to be behavioral, N is the initial sample size, and S is the sample size at each stage. The user-specified parameter f represents the memory in CBS corresponding to which of the existing parameter sets should be used to generate the new parameter sets at each stage; specifically, it indicates the extent to which the algorithm is allowed to relax the acceptable minimum number of process constraints (C) when considering parameter sets for use in generating the new samples.[1]

Step 0: Begin with $C = 2$.

Step 1: Generate N random samples (parameter sets) across the entire feasible parameter space using uniform prior distributions.

Step 2: Evaluate the parameter constraints and identify the K_1 samples that satisfy them.

Step 3: Run the model K_1 times, one time for each of the samples identified in step 2, evaluate the M process constraints for each sample and assign a value of m to each parameter set corresponding to the number of process constraints satisfied.

Step 4: Place the K_2 samples that satisfy C or more process constraints in set P, and the K_3 samples that satisfy exactly C-f-1 to C-1 process constraints in set P'. Discard samples that satisfy C-f-2 or fewer process constraints.

Step 5: Use the members of sets P and P' to generate S new samples by applying each of the three Monte Carlo based rules below to generate $S/3$ of the samples, where θ_{new} is the newly generated sample. θ_P and $\theta_{P'}$ are samples selected randomly from sets P and P' respectively and α is a random value between 0 and 1 which differs for each newly generated sample. Figure 2 shows a graphical illustration of these rules.

$$\theta_{new} = \alpha\theta_p + (1-\alpha)\theta_p \quad (6)$$

$$\theta_{new} = \alpha\theta_p + (1-\alpha)\theta_{p'} \quad (7)$$

$$\theta_{new} = \alpha\theta_{p'} + (1-\alpha)\theta_{p'} \quad (8)$$

Step 6: Discard all existing members of set P' which satisfy exactly C-f-1 process constraints[2].

Step 7: Increase C by one and return to step 2. Repeat this process until C becomes equal to the total number of process constraints (i.e. $C=M$), or the maximum number of model evaluation is exceeded, or the size of P is satisfactory.

In step 5, the selection of parameter sets from P and P' can also be based on the number of satisfied process constraints or any other distribution different from a uniform distribution.

Note that any member of set P is within the space marked by members of set P'. Using members of P' to generate new parameter sets (step 5) helps to identify the boundary between the parameter space that satisfies exactly C-f-1 to C-1 (set P') and C or more (set P) constraints. The intention is to obtain a diverse parameter representation for set P by including the set P' (Figure 2).

---

[1] C, S, N and f can be changed by the user in accordance with the model, nature and number of constraints. We propose C to be set at 2 initially.

[2] In the case that f is set to 0 there won't be any existing parameter sets in P' (i.e. $P'=\Phi$).
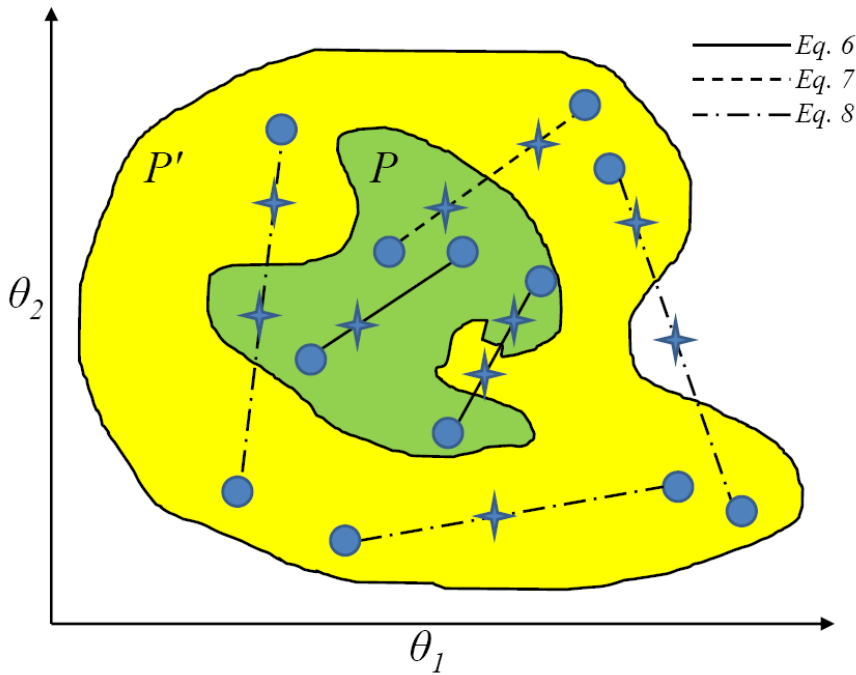
1
2 Figure 2- A conceptual illustration of possible positions of newly generated parameter sets
3 based on parameter sets randomly drawn from P and P' for a two dimensional parameter
4 space. The area indicated in yellow represents the set P' that satisfies exactly C-f-1 to C-1
5 process based constraints. The area indicated in green represents the set of P that satisfies C
6 or more process constraints. The circles indicate randomly selected parameter sets drawn from
7 the sets P or P'. Different line styles indicate different parameter generation rules
8 (corresponding to the different equations). Solid lines represent the first rule where the
9 parameter sets are randomly selected from set P *(Eq.6)*, dashed lines illustrate the second rule
10 where one parameter set is randomly selected from P and one from P' *(Eq.7)*, the dash-dot
11 line represents the third rule where both randomly selected parameter sets are selected from
12 set P' *(Eq.8)*. Note that due to possible non-convexity of sets of P and P' the newly generated
13 parameter sets based on the three rules can be outside of sets P and P'.
14
15 The final set P contains parameters that simultaneously satisfy all of the parameter and
16 process constraints. These parameter sets can be referred to as *constrained but un-calibrated,*
17 as they are not calibrated to match observed data sets or target variables (e.g. discharge time-
18 series). Figure 3 presents a graphical illustration of the steps of CBS to find set *constrained*
19 *but un-calibrated* parameter sets.
20 Note that the set P can also be used to constrain a search for 'optimal' parameter sets within
21 this space of *constrained but un-calibrated* parameter sets. This can be easily achieved by
22 evaluating them based on model performance in regard to a target variable (e.g. observed
23 runoff). The set P can be used as an initial sample for any evolutionary algorithm. In this case,
24 any new parameter set generated by an optimization algorithm would need to be checked for
25 both parameter and process constraints and would be retained only if they satisfy the entire set
26 of constraints.
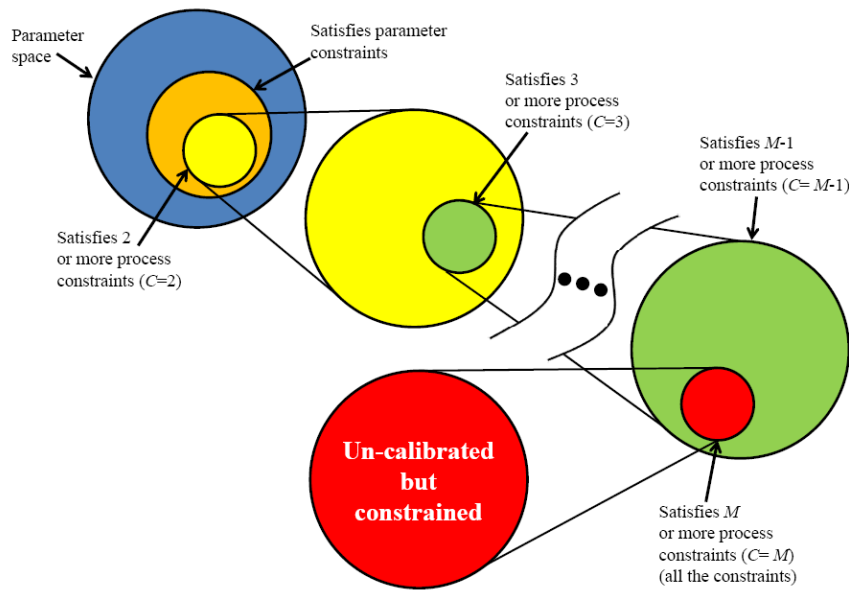
1
2   Figure 3- A conceptual illustration of stepwise search for the parameter space satisfying all of
3   the parameter and process constraints.
4
5
6   **4 Case study**
7
8   A synthetic case study was designed to illustrate the efficiency of the proposed constraint-
9   based search method. First, the lumped model FLEX[A] (figure 4; see companion paper Gharari
10 et al., 2014a) was calibrated to runoff data from the Wark catchment for the year 2002-2005,
11 using year 2001 as warm up period. A multi-criteria calibration was conducted *without*
12 imposing any parameter or process constraints, based on three objective functions: (i) the
13 Nash-Sutcliffe efficiency of the flows (E_{NS}), (ii) the Nash-Sutcliffe efficiency of the logarithm
14 of the flows ($E_{NS,log}$) and (iii) the Nash-Sutcliffe efficiency of the flow duration curve
15 ($E_{NS,FDC}$). From the parameter sets belonging to the Pareto front, the parameter set having
16 minimum distance to the origin was taken as the "*best*" behavioral parameter set (θ_{best}).
17 A set of (parameter and process) constraints was then designed based on the model
18 simulations (flux and state trajectories) provided by this best parameter set (θ_{best}). Since the
19 constraints were constructed so that the best performing parameter set (θ_{best}) is "*behavioral*",
20 we are guaranteed the existence of feasible parameter sets that satisfy all of those constraints.
21 Table 1 summarizes the parameter and process constraints designed in this manner; note that
22 the value for each constraint corresponding to θ_{best} is mentioned, along with the acceptable
23 range limits for each constraint to be used during the search.
24

FLEX^A
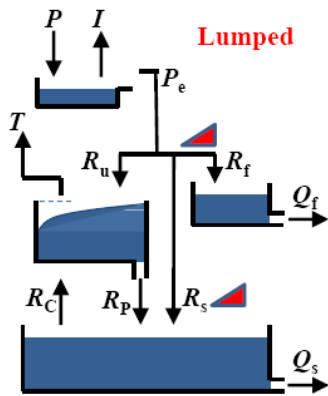
1
2 Figure 4- The models structure of FLEX^A. Precipitation is indicated by P; I and T represent
3 components of evaporation which are interception and transpiration respectively. Q_f and Q_s
4 indicate the fast and slow components that make the total model output.
5
6 Table 1- The "true" values for each of the design constraints (corresponding to θ_{best}) along
7 with the acceptable limits assigned to each constraint for use during the search. Dry and wet
8 periods are defined as April to November and December to March respectively. Events in
9 which the discharge increases with a rate of more than 0.2 mm(3h)$^{-2}$ are defined as peak
10 flows.

No	Quantification of designed constraints based on θ_{best}	Acceptable range for each designed constraints
	Parameter constraint	
1	$\dfrac{K_F}{K_S} = 4.18$	$3 < \dfrac{K_F}{K_S} < 5$
	Process constraints	
1	$\dfrac{\int I \mathrm{d}t}{\int I \mathrm{d}t + \int T \mathrm{d}t} = 0.30$	$0.25 < \dfrac{\int I \mathrm{d}t}{\int I \mathrm{d}t + \int T \mathrm{d}t} < 0.35$
2	$\dfrac{\int Q_f \mathrm{d}t}{\int Q_f \mathrm{d}t + \int Q_s \mathrm{d}t} = 0.33$	$0.25 < \dfrac{\int Q_f \mathrm{d}t}{\int Q_f \mathrm{d}t + \int Q_s \mathrm{d}t} < 0.35$
3	$\dfrac{\int Q_{f,wet,peaks} \mathrm{d}t}{\int Q_{f,wet,peaks} \mathrm{d}t + \int Q_{s,wet,peaks} \mathrm{d}t} = 0.57$	$0.50 < \dfrac{\int Q_{f,wet,peaks} \mathrm{d}t}{\int Q_{f,wet,peaks} \mathrm{d}t + \int Q_{s,wet,peaks} \mathrm{d}t} < 0.60$
4	$\dfrac{\int Q_{f,dry,peaks} \mathrm{d}t}{\int Q_{f,dry,peaks} \mathrm{d}t + \int Q_{s,dry,peaks} \mathrm{d}t} = 0.60$	$0.55 < \dfrac{\int Q_{f,dry,peaks} \mathrm{d}t}{\int Q_{f,dry,peaks} \mathrm{d}t + \int Q_{s,dry,peaks} \mathrm{d}t} < 0.65$

|		
5	$\dfrac{\int Q_{m,dry,2002}dt}{\int P_{dry,2002}dt}=0.10$, $\dfrac{\int Q_{m,wet,2002}dt}{\int P_{wet,2002}dt}=0.59$, $\dfrac{\int Q_{m,2002}dt}{\int P_{2002}dt}=0.39$	$0.15<\dfrac{\int Q_{m,dry,2002}dt}{\int P_{dry,2002}dt}<0.05$, $0.55<\dfrac{\int Q_{m,wet,2002}dt}{\int P_{wet,2002}dt}<0.65$, $0.35<\dfrac{\int Q_{m,2002}dt}{\int P_{2002}dt}<0.45$
6	$\dfrac{\int Q_{m,dry,2003}dt}{\int P_{dry,2003}dt}=0.06$, $\dfrac{\int Q_{m,wet,2003}dt}{\int P_{wet,2003}dt}=0.44$, $\dfrac{\int Q_{m,2003}dt}{\int P_{2003}dt}=0.27$	$0.00<\dfrac{\int Q_{m,dry,2003}dt}{\int P_{dry,2003}dt}<0.10$, $0.40<\dfrac{\int Q_{m,wet,2003}dt}{\int P_{wet,2003}dt}<0.50$, $0.20<\dfrac{\int Q_{m,2003}dt}{\int P_{2003}dt}<0.30$
7	$\dfrac{\int Q_{m,dry,2004}dt}{\int P_{dry,2004}dt}=0.10$, $\dfrac{\int Q_{m,wet,2004}dt}{\int P_{wet,2004}dt}=0.35$, $\dfrac{\int Q_{m,2004}dt}{\int P_{2004}dt}=0.20$	$0.05<\dfrac{\int Q_{m,dry,2004}dt}{\int P_{dry,2004}dt}<0.15$, $0.30<\dfrac{\int Q_{m,wet,2004}dt}{\int P_{wet,2004}dt}<0.40$, $0.15<\dfrac{\int Q_{m,2004}dt}{\int P_{2004}dt}<0.25$
8	$\dfrac{\int Q_{m,dry,2005}dt}{\int P_{dry,2005}dt}=0.05$, $\dfrac{\int Q_{m,wet,2005}dt}{\int P_{wet,2005}dt}=0.26$, $\dfrac{\int Q_{m,2005}dt}{\int P_{2005}dt}=0.17$	$0.00<\dfrac{\int Q_{m,dry,2005}dt}{\int P_{dry,2005}dt}<0.05$, $0.20<\dfrac{\int Q_{m,wet,2005}dt}{\int P_{wet,2005}dt}<0.30$, $0.10<\dfrac{\int Q_{m,2005}dt}{\int P_{2005}dt}<0.20$
9	$\dfrac{\sum_{i=1}^{t}\left\lfloor \dfrac{S_{u,i}}{0.5S_{u\,max}}\right\rfloor}{t}=0.53$	$0.4<\dfrac{\sum_{i=1}^{t}\left\lfloor \dfrac{S_{u,i}}{0.5S_{u\,max}}\right\rfloor}{t}<0.6$

1
2
3   The proposed algorithm was then applied to search for parameter sets that satisfy the
4   constraints mentioned in Table 1. The initial sample size (N) was taken as 50000 parameter
5   sets. S, f and C were set to be 5000, 1 and 2 respectively. The search was terminated when the
6   number of generated parameter sets satisfying all of the parameter and process constraints
7   exceeded 8000.
8   The parameter sets identified by the proposed search methodology were then compared with
9   the "best" value (θ_{best}). Figure 5 shows the results using a normalized parameter plot (Gupta et
10   al 1998), where the red trajectory indicates θ_{best} and the hue of the grey-to-black lines
11   indicates the number of process constraints satisfied by a given parameter set (the darker the
12   line, the higher number of process constraints satisfied). The result clearly shows that with the
13   increasing number of process constraints satisfied, the feasible solution space converges
14   towards the "best" parameter set (θ_{best}).
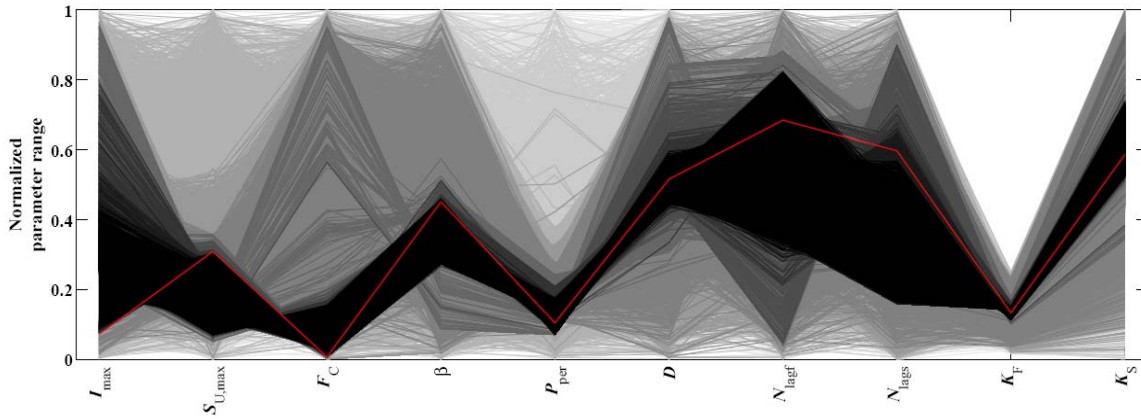
1
2   Figure 5- Normalized parameter plot for FLEX[A]. Darker colors indicate parameter sets that
3   satisfy larger numbers of process constraints. The red trajectory indicates θ_{best}.
4
5   A similar comparison (Figure 6) was conducted for the modeled hydrographs associated with
6   these parameter sets. The hydrograph for the best performing parameter set (θ_{best}) is indicated
7   in red, and the simulated hydrographs for different parameter sets generated by the proposed
8   search method are depicted in different shades ranging from grey-to-black (darker colors
9   indicate that larger numbers of process constraints are satisfied). Clearly, with increasing
10 number of satisfied process constraints, the hydrographs progressively approach the one
11 simulated using the θ_{best}.
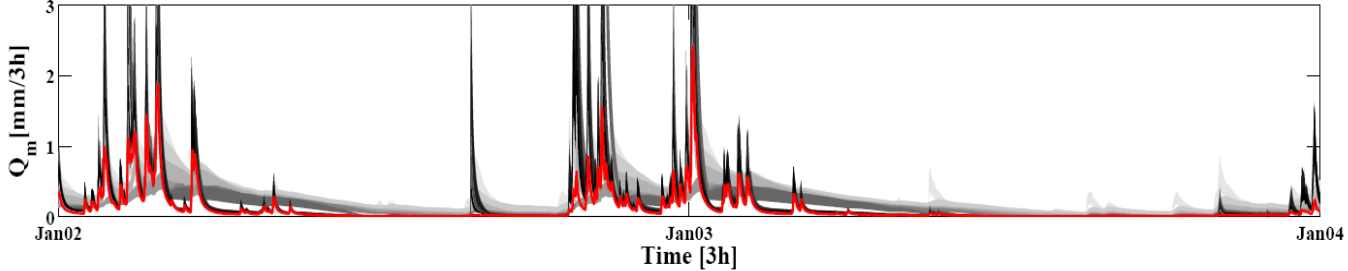
12
13 Figure 6- 95% uncertainty intervals for parameter sets satisfying different numbers of process
14 constraints. Darker colors indicate the uncertainty intervals which satisfy more of the process
15 constraints. The hydrograph generated by θ_{best} is indicated in red.
16
17 Overall the search method generated and evaluated 102106 parameter sets to find 8000
18 feasible solutions that satisfy all of the nine constraints imposed, which corresponds to
19 approximately 8% efficiency. In comparison, when a conventional Monte-Carlo sampling
20 approach was applied using 102106 samples, _none_ of the samples were found to be able to
21 satisfy all of the constraints, while only two of the samples were able to satisfy at least 7 of
22 the 9 process constraints imposed. Clearly, the proposed search method is capable of
23 relatively rapid convergence towards the region of the parameter space where all of the
24 constraints are satisfied.
25 Figure-7 illustrates how quickly the search algorithm is able to locate the behavioral
26 parameter sets. It depicts the percentages of generated samples satisfying a given number of
27 process constraints at each step of the search. Darker colors are used to indicate the
28 proportions of parameter sets that satisfy progressively more of the process constraints (white
29 indicates none of the parameter constraints being satisfied). The initial sample of 50,000
30 parameter sets, see region (**a**) in Figure-7, consists of samples drawn uniformly from the
31 entire parameter space, of which less than 10% satisfy any of the imposed constraints, and
32 only a very few satisfy 1, 2 and 3 constraints (progressively darker shades of grey). Each

progressive step (see region (**b**) in Figure-7) then consists of S=5000 samples generated according to the strategy discussed above, in which the value of C is increased by one. Clearly, at each step, the proportion of *newly generated samples* satisfying larger numbers of constraints increases rapidly. Note that when C was set equal to 9 (i.e., the maximum number of design constraints), the search was continued until the pre-specified 8000 "behavioral" parameter sets had been found, and then terminated. At this point the fraction of number of newly generated samples (behavioral parameter sets) that satisfy all of the constraints has reached approximately 40% (region (**c**) in Figure-7).
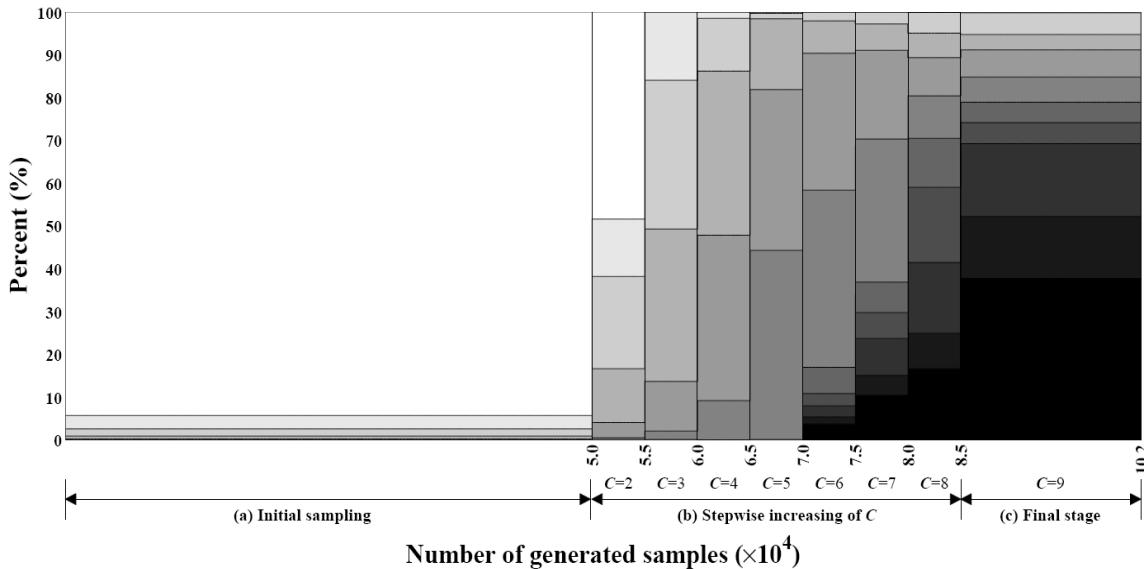
Figure 7- Percentages of samples generated at each stage that satisfy a given number of process constraints. White indicates no satisfied constraints, and progressively darker colors indicate increasing numbers of satisfied constraints. (**a**) The initial sampling (N=50000). (**b**) Stepwise increase of C, thus requiring an increasing number of process constraints to be satisfied at each stage, and (**c**) the final stage of the search when C is set to the maximum number of design constraints (here 9), at which point the search continues until 8000 behavioral parameter sets have been located.

Of course, in this illustrative case study, the constraints were specifically designed to guarantee that the observed hydrograph corresponding to the "best" performing parameter set lie within a predetermined feasible space. In principle the constraints can be specified without recourse to the information contained in the discharge time series (as discussed earlier). The main purpose of this synthetic case study was to illustrate the capability of the proposed CBS search method to efficiently locate behavioral parameter sets that satisfy user-specified *a priori* parameter and *a posteriori* process constraints.

5 Conclusions

One of the most challenging tasks in the development of complex conceptual hydrological models for simulation of catchment responses to inputs is the realistic specification of parameter values. We have presented a constraint-based search strategy that facilitates the incorporation of expert knowledge (i.e. the modeler's perception of catchment behavior and characteristics) into the parameter specification process. Because the CBS algorithm does not require observational data regarding the target system output (e.g. runoff) it can provide a way forward for better prediction in ungauged basins in absence of streamflow (or other system

output) data for model calibration. As constraints are much easier for understand, rather than parameters, when discussing system behavior, they can potentially be used as an efficient tools to bridge the gap in the dialogue between modelers and experimentalists. Further, the approach can help to provide behaviorally more conceptually realistic parameter sets when used in conjunction with model calibration. Future study may apply the proposed CBS algorithm to identify behavioral parameter sets for different kind of hydrologic models in different regions. A Matlab^{TM} code of the CBS algorithm can be obtained through personal communication with the lead author.

6 References

Anderson, R. M.; Koren, V. I. & Reed, S. M. Using SSURGO data to improve Sacramento Model a priori parameter estimates Journal of Hydrology, 320(1-2), 103-116, 2006.

Bastidas, L. A.; Gupta, H. V.; Sorooshian, S.; Shuttleworth, W. J. & Yang, Z. L. Sensitivity analysis of a land surface scheme using multicriteria methods Journal of Geophysical Research: Atmospheres, 104(D16), 19481-19490, 1999.

Beven, K. Rainfall-Runoff Modelling, The Primer John Wiley and Sons, 372, 2001.

Birkel, C., D., T., Dunn, S., and Soulsby, C.: Using time domain and geographic source tracers to conceptualize streamflow generation processes in lumped rainfall-runoff models, Water Resource Research, 47, W02 515, doi:10.1029/2010WR009547, 2011.

Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, Water Resource Research, 36(12), 3663–3674, doi:10.1029/2000WR900207, 2000.

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

Bulygina, N. and Gupta, H.: Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation, Water Resources Research, 45(12), n/a–n/a, doi:10.1029/2007WR006749, http://dx.doi.org/10.1029/2007WR006749, 2009.

Bulygina, N. and Gupta, H.: How Bayesian data assimilation can be used to estimate the mathematical structure of a model, Stochastic Environmental Research and Risk Assessment, 24(6), 925–937, doi:10.1007/s00477-010-0387-y, http://dx.doi.org/10.1007/s00477-010-0387-y, 2010.

Bulygina, N. and Gupta, H.: Correcting the mathematical structure of a hydrological model via Bayesian data assimilation, Water Resources Research, 47(5), n/a–n/a, doi:10.1029/2010WR009614, http://dx.doi.org/10.1029/2010WR009614, 2011.

Efstratiadis, A. and Koutsoyiannis, D.: One decade of multiobjective calibration approaches in hydrological modelling: a review, Hydrological Sciences Journal, 55(1), 58–78, doi:10.1080/02626660903526292, 2010.

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

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

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

Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H., and Savenije, H. H. G.: Using Expert Knowledge to Increase Realism in Environmental System Models can dramatically Reduce the Need for Calibration, Hydrology and Earth System Sciences Discussions, XXX, XXX, doi:XXX, 2014a.

Gharari, S.; Hrachowitz, M.; Fenicia, F. & Savenije, H. H. G. Hydrological landscape classification: investigating the performance of HAND based landscape classifications in a central European meso-scale catchment Hydrology and Earth System Sciences, 15(11), 3275-3291 ,2011.

Gőtzinger, J. & Bárdossy, A. Comparison of four regionalisation methods for a distributed hydrological model Journal of Hydrology, 333, 374-384, 2007.

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

Gupta, H. V., Bastidas, L. A., Sorooshian, S., Shuttleworth, W. J., and Yang, Z. L.: Parameter estimation of a land surface scheme using multicriteria methods, Journal of Geophysical Research: Atmospheres, 104(D16), 19 491–19 503, doi:10.1029/1999JD900154, http://dx.doi.org/10.1029/1999JD900154, 1999.

Hrachowitz, M.; Savenije, H.; Bogaard, T. A.; Tetzlaff, D. & Soulsby, C. What can flux tracking teach us about water age distribution patterns and their temporal dynamics? Hydrology and Earth System Sciences, 533-564, 17(2), 2013a.

Hrachowitz, M., Savenije, H. H. G., Blschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB)a review, Hydrological Sciences Journal, 58(6), 1–58, doi:10.1080/02626667.2013.803183, http://www.tandfonline.com/doi/abs/10.1080/02626667.2013.803183, 2013b.

Hrachowitz, M.; Fovet, O.; Ruiz, L.; Euser, T.; Gharari, S.; Freer, J.; Savenije, H. H. G.; Gascuel-Odoux, C.: Process Consistency in Models: the Importance of System Signatures, Expert Knowledge and Process Complexity, Water Resource Research, accepted.

Hughes, D. A., Andersson, L., Wilk, J., and Savenije, H. H.: Regional calibration of the Pitman model for the Okavango River, Journal of Hydrology, 331(1-2), 30–42, doi: http://dx.doi.org/10.1016/j.jhydrol.2006.04.047, http://www.sciencedirect.com/science/article/pii/S0022169406002514, ¡ce:title¿Water Resources in Regional Development: The Okavango River¡/ce:title¿, 2006.

Kapangaziwiri, E., Hughes, D., and Wagener, T.: Incorporating uncertainty in hydrological predictions for gauged and ungauged basins in southern Africa, Hydrological Sciences Journal, 57(5), 1000–1019, doi:10.1080/02626667.2012.690881, http://www.tandfonline.com/doi/abs/10.1080/02626667.2012.690881, 2012.

Khu, S. T. and Madsen, H.: Multiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration, Water Resource Research, 41, W03 004, doi:10.1029/2004WR003041, 2005.

Kling, H. and Gupta, H.: On the development of regionalization relationships for lumped watershed models: The impact of ignoring sub-basin scale variability, Journal of Hydrology, 373(3-4), 337–351, doi:10.1016/j.jhydrol.2009.04.031, http://www.sciencedirect.com/science/article/pii/S0022169409002893, 2009.

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

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

Kumar, R., Samaniego, L., and Attinger, S.: The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics, Journal of Hydrology, 392(1–2), 54–69, doi:10.1016/j.jhydrol.2010.07.047, http://www.sciencedirect.com/science/article/pii/S0022169410004865, 2010.

Kumar, R., Samaniego, L., and Attinger, S.: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resources Research, 491(1), 360–379, doi:10.1029/2012WR012195, http://dx.doi.org/10.1029/2012WR012195, 2013.

Martinez, G. F. and Gupta, H. V.: Hydrologic consistency as a basis for assessing complexity of monthly water balance models for the continental United States, Water Resources Research, 47(12), n/a–n/a, doi:10.1029/2011WR011229, http://dx.doi.org/10.1029/2011WR011229, 2011.

Merz, R. and Blőschl, G.: Regionalisation of catchment model parameters, Journal of Hydrology, 287, 95123, 2004.

Perrin, C., Andréassian, V., Serna, C. R., Mathevet, T., and Moine, N. L.: Discrete parameterization of hydrological models: Evaluating the use of parameter sets libraries over 900 catchments, Water Resources Research, 44, W08 447, doi:10.1029/2007WR006579, 2008.

Pokhrel, P., Gupta, H. V., and Wagener, T.: A spatial regularization approach to parameter estimation for a distributed watershed model, Water Resources Research, 44(12), n/a–n/a, doi:10.1029/2007WR006615, http://dx.doi.org/10.1029/2007WR006615, 2008.

Pokhrel, P., Yilmaz, K. K., and Gupta, H. V.: Multiplecriteria calibration of a distributed watershed model using spatial regularization and response signatures, Journal of Hydrology, 418419(0), 49–60, doi:http://dx.doi.org/10.1016/j.jhydrol.2008.12.004, http://www.sciencedirect.com/science/article/pii/S0022169408005970, ¡ce:title¿The Distributed Model Intercomparison Project (DMIP) - Phase 2 Experiments in the Oklahoma Region, USA¡/ce:title¿, 2012.

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

Seibert, J. and McDonnell, J.: Gauging the Ungauged Basin: The Relative Value of Soft and Hard Data, Journal of Hydrologic Engineering, doi:10.1061/(ASCE)HE.1943-5584.0000861, http://ascelibrary.org/doi/abs/10.1061/%28ASCE%29HE.1943-5584.0000861, 2013.

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resource Research, 38(11), 1241, doi:10.1029/2001WR000978, 2002.

Wheater, H. S., Jakeman, A. J., and Beven, K. J.: Progress and directions in rainfall-runoff modeling, JohnWiley & Sons, 1993.

Winsemius, H. C., Savenije, H. H. G., and Bastiaanssen, W. G. M.: Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins?, Hydrology and Earth System Sciences, 12(6), 1403–1413, doi:10.5194/hess-12-1403-2008, http://www.hydrol-earth-syst-sci.net/12/1403/2008/, 2008.

Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, Advances in Water

Resources, 30(8), 1756–1774, doi:10.1016/j.advwatres.2007.01.005, http://www.sciencedirect.com/science/article/pii/S0309170807000140, 2007.

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