A comparison of alternative multiobjective calibration strategies for hydrological modeling

Fabrizio Fenicia,[1,2] Hubert H. G. Savenije,[2] Patrick Matgen,[1] and Laurent Pfister[1]

[1]   A conceptual hydrological model structure contains several parameters that have to be estimated through matching observed and modeled watershed behavior in a calibration process. The requirement that a model simulation matches different aspects of system response at the same time has led the calibration problem toward a multiobjective approach. In this work we compare two multiobjective calibration approaches, each of which represents a different calibration philosophy. The first calibration approach is based on the concept of Pareto optimality and consists of calibrating all parameters with respect to a common set of objectives in one calibration stage. This approach results in a set of Pareto-optimal solutions representing the trade-offs between the selected calibration objectives. The second is a stepped calibration approach (SCA), which implies a stepwise calibration of sets of parameters that are associated with specific aspects of the system response. This approach replicates the steps followed by a hydrologist in manual calibration and develops a single solution. The comparison is performed considering the same set of objectives for the two approaches and two model structures of a different level of complexity. The difference in the two approaches, their reciprocal utility, and the practical implications involved in their application are analyzed and discussed using the Hesperange catchment case, an experimental basin in the Alzette River basin in Luxembourg. We show that the two approaches are not necessarily conflicting but can be complementary. The first approach provides useful information about the deficiencies of a model structure and therefore helps the model development, while the second attempts at determining a solution that is consistent with the data available. We also show that with increasing model complexity it becomes possible to reproduce the observations more accurately. As a result, the solutions for the different calibration objectives become less distinguishable from each other, indicating that calibration results become less dependent on the objective functions used when the model is a better representation of reality and has a higher potential to reproduce the observations.

1. Introduction

[2] Conceptual hydrological models commonly operate with several connected stocks representing physical elements in a catchment. Model parameters define the behavior of the various conceptual elements and the way they relate to each other. As conceptual elements represent averages of various subcatchment processes that contribute to the overall catchment response, model parameters are conceptual representations of abstract watershed characteristics and cannot be assessed from direct measurements. Instead they have to be determined by calibration, which is a process of changing parameter values until a satisfactory agreement between simulated and observed catchment behavior is obtained [*Sorooshian and Gupta*, 1995].

[3] In manual calibration a process of trial and error parameter adjustment is made, and the simulated and observed watershed behavior is compared using visual inspection and different measures of performance. While manual calibration can produce good results, it can be time consuming and it involves a great deal of subjective judgment.

[4] The shortcomings of manual calibration have motivated the automation of the calibration process. This has transformed the calibration problem into an optimization problem, consisting in determining the set of model parameters that optimizes (maximizing or minimizing) a number of objective functions. Objective functions are single valued equations that depend on model parameters and express the agreement between observed and simulated catchment behavior in numerical form.

[5] Single objective calibration consists of determining the set of model parameters that optimizes a single objective function. Such an approach to model calibration, however, is subject to limitations that restrict its applicability. Calibration based on a single objective function, in fact, often

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

results in hydrograph representations that are considered unrealistic from the operational hydrologist's point of view. This can be due to the following reasons. First, a single objective function may enhance the error with respect to the simulation of some aspects of the observed signal at the expense of other aspects, therefore constraining the calibration to fit certain characteristics of the system response while neglecting others. Second, the integration of the residuals into one value may hide or underestimate the information content of the data available, therefore not capturing and not exploiting all the information that is present in the data. These limitations suggest the need of constraining the calibration processes by a larger number of objective functions, leading to a multiobjective view of the calibration problem.

[6] In this paper we compare two multiobjective approaches, representative of different ways of interpreting the calibration process. The first approach refers to the concept of Pareto optimality [*Gupta et al.*, 1998] and consists in calibrating all model parameters simultaneously with respect to a common set of objective functions. The approach results in the determination of a set of Pareto-optimal solutions, reflecting various trade-offs between parameters and calibration objectives. The second is a "stepped" calibration approach [*Hogue et al.*, 2000], and consists in associating model parameters with calibration objectives based on the processes that each parameter is designed to represent and on the role of each process on the overall system response. The parameter sets associated with the different objectives are calibrated in separate stages, reflecting the procedure that is followed by operational hydrologists in manual calibration. The approach provides a single solution that represents a balance between the selected calibration objectives. The purpose is to demonstrate the principles and implications of each approach in a comparative evaluation. The two approaches are examined in a case study that considers the calibration of two models of different levels of complexity.

[7] The set of objective functions, the same for the two approaches, is chosen to evaluate model performances with respect to three aspects of the stream hydrograph simulation, namely, low flows, high flows and lag time of the system. A comparison is made between two model structures with different levels of complexity. Initially a simple model structure is used and calibration results are evaluated. According to the calibration results and to the hydrological insight of the catchment, the initial model structure is improved by introducing additional processes and components. The calibration procedure is repeated for the improved model structure. This comparison gives the opportunity not only of evaluating the performance of multiobjective calibration at different levels of model complexity, but also allows a discussion of the results of calibration strategies as a means of understanding model deficiencies and helping model development.

2. Description of the "All at Once" Pareto-Based Calibration Approach

[8] In the field of hydrology, the principles of multiobjective calibration based on the concept of Pareto optimality were first presented by *Gupta et al.* [1998]. Subsequently, the concept has been applied in several case studies with different kinds of model structures and calibration objectives [*Bastidas et al.*, 1999; *Boyle et al.*, 2000; *Madsen*, 2000, 2003; *Wagener et al.*, 2001; *Meixner et al.*, 2002; *Crow et al.*, 2003; *Vrugt et al.*, 2003]. In such a framework, the multiobjective calibration problem can be stated as follows [*Madsen*, 2000]:

$$\text{Min}\{F_1(\vartheta), F_2(\vartheta), \ldots, F_m(\vartheta)\} \qquad \vartheta \in \Theta$$

where the solution ϑ is a vector of model parameters, which is constrained to vary within the feasible parameter space Θ. The objective functions $F_i(\vartheta)$, $i = 1 \ldots m$; are scalars that reflect the model performance with respect to the selected calibration objectives. Lower values in $F_i(\vartheta)$ indicate better model performances.

[9] The concept of Pareto optimality is based on the notion of Domination and is defined as follows: (1) A solution ϑ_1 is said to dominate another solution ϑ_2 when ϑ_1 is better than ϑ_2 in at least one objective (meaning $F_i(\vartheta_1) < F_i(\vartheta_2)$ for at least one value of i), and not worse than ϑ_2 in any of the others (meaning $F_i(\vartheta_1) \leq F_i(\vartheta_2)$ for all values of i). (2) The Pareto-optimal set of solutions is composed of those solutions that are not dominated by any solution of the feasible search space. The mapping of the Pareto-optimal solutions in the objective function space is defined as the Pareto-optimal front.

[10] The Pareto-optimal set of solutions will in general consist of more than one solution. When this is the case, the objective functions are said to be conflicting with each other, in a sense that moving from one optimal solution to the other determines an improvement in one or more objectives, and a deterioration in the others.

[11] When applied to hydrological modeling, the existence of multiple optimal solutions can be related to a systematic component of the modeling error. This component can be determined by errors in the model structure, in the boundary conditions, and in the process of data collection and preparation. When the modeling task is seen as an effort to represent as closely as possible the observed behavior of the catchment, which can appear reasonable when no information is available to correct eventual distortions in the measured data, the existence of multiple Pareto-optimal solutions can be regarded as a failure of the model to perform this task. In this view, the condition of multiple optimal solutions can be regarded as an inability of the model to reproduce simultaneously different aspects of the system behavior, and therefore is related to model structural limitations.

[12] The Pareto-based approach, apart from helping the identification of model limitations, can also be useful to compare relative merits of different models and to track changes in model performance. The Pareto-optimal front, in fact, marks the best performance that a model can reach given a record of calibration data. A shifting of the front toward the origin of the axes with the same calibration record would therefore indicate a better model, or a successful modification of the model structure. As an example, *Xia et al.* [2002] use this approach to evaluate differences between alternative model structures and parameterizations, concluding that such an approach provides a useful guidance for model improvement.

[13] In the light of the Pareto-based approach, all Pareto-optimal solutions are equally important, as it is difficult to prefer one solution over another without any further information about the problem. This does not mean, however, that an operational hydrologist who is interested in model simulations that fulfill the selected calibration objectives would regard all Pareto-optimal solutions as equally good. In the presence of conflicting objectives, for instance, the solutions that optimize each individual objective function and that belong to the Pareto-optimal set of solutions would probably suffer from the same problems that affect single objective calibration and that lead to the determination of often unsatisfactory outcomes. Those solutions could in fact result in a biased performance that fits one aspect of the observed system behavior (the one related to the specific objective function) but neglects other aspects. *Madsen* [2003], for instance, performed multiobjective calibration by considering the simulation of groundwater levels and catchment runoff as calibration objectives. They observed that the Pareto-optimal solution that optimizes the performance with respect to groundwater levels performed "very badly" in simulating catchment runoff. A significant improvement in the catchment runoff simulation could be obtained by slightly relaxing the performance of the groundwater level simulation. *Xia et al.* [2002, 2005], analyzing several land surface models, determined that some parameter sets, while being optimal, corresponded to an unrealistic description of processes.

[14] This is to say that the solutions that fulfill all the selected calibration objectives may have to be found in a proper balance of the corresponding objective functions. Balanced solutions may be contained in the Pareto-optimal set of solutions, but not necessarily represented by all solutions of the set. Hence Pareto-based calibration can be considered as a valuable tool to evaluate a model with respect to its ability to reproduce different aspects of the observations and to compare performances of different models, but it has to be considered that the solutions obtained by this approach, while being equally good from an "optimization" point of view, may not all be acceptable from a "calibration" perspective.

3. Description of the Stepped Calibration Approach

[15] The stepped calibration approach (SCA) is a calibration strategy that mimics the steps that would be followed by operational hydrologists who manually calibrate a model to fit certain aspects of the observed watershed behavior. The approach determines a single parameter set, which corresponds to an acceptable simulation according to the selected criteria. The approach consists of the following main points: (1) selecting some specific characteristic of the recorded time series that should be well simulated according to our needs (e.g., high flows or low flows), (2) defining objective functions that represent performance measures for the simulation of the selected characteristics, (3) associating model parameters with the selected objective functions based on the process description associated to each parameter and on the influence that each process has on the simulated system response, and (4) calibrating the parameters associated with each objective function in separate stages.

[16] The calibration in separate stages can be done by adopting different schemes and assigning different priorities. An approach could be to first calibrate all parameters with respect to the first objective function, then recalibrating all parameters except those related to the first objective function (which are fixed at the calibrated values) with respect to the second objective function, then recalibrate all parameters except those related to the first and second objective function against the third objective function, and so forth. Alternatively, after a first calibration of all parameters with respect to the first objective function, each group of parameters associated with each objective function can be calibrated individually, while fixing the remaining parameters at the calibrated values [e.g., *Hogue et al.*, 2000]. The solution that is determined through the SCA is clearly dependent on the succession of steps that is followed, in the sense that altering the order of the different steps would lead to different solutions. However, as this approach replicates in an automated fashion the steps that are normally undertaken by operational hydrologists during manual calibration, the methodology that is followed is to start by fitting first those characteristics of the hydrograph that are more regular and better identifiable (e.g., low-flow recessions), while gradually proceeding to the calibration of the others (e.g., timing, bias).

[17] The application of the SCA can be trivial in those cases where model components perform operations that are not easily recognizable as specific features of the system response, or where their effect applies to aspects of the simulation that affect different calibration objectives. However, it is not infrequent that at least some of the model parameters express processes that have a specific influence on the simulation. When this is the case, a possibility could be to use in the first stage of the SCA a general objective function representing an overall measure of performance, and then proceed with more specific calibration objectives in the following stages, adjusting only those parameters that have a direct impact on the selected objectives.

[18] The SCA is based on the assumption that model calibration should assure that all components perform the operations for which they are intended. In this context, it appears logical to adjust model parameters to reproduce the hydrograph characteristics that they are designed to influence. The purpose of the SCA is to obtain a single parameter set that corresponds to a model simulation that is consistent with the expert's understanding of reality represented by the model structure. Hence it tries to avoid fitting some aspects of the simulations at the expense of others and tries to prevent compensation of internal model structural errors by adjusting model parameters to unrealistic values.

[19] With respect to the Pareto-based calibration, the SCA looks at the calibration problem more in a "calibration" optic than in an "optimization" optic. As it proceeds in successive single objective calibration stages, in fact, the final solution does not necessarily result in a Pareto-optimal solution. This issue will be illustrated further in the application. Moreover, unlike the Pareto-based calibration, the SCA results in a single solution, and does not provide alternative possible combinations of parameters that can produce equally good results.

[20] The SCA has been applied by *Hogue et al.* [2000] for the calibration of the Sacramento soil moisture accounting (SAC-SMA) and snow accumulation and
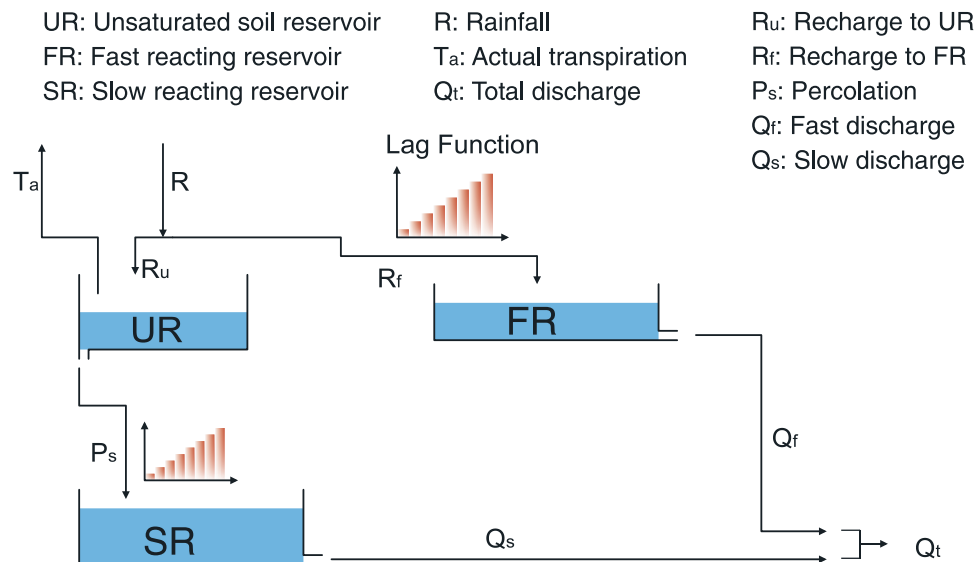
Figure 1. Structure schematization of the FLEX^A model.

ablation (SNOW-17) models. Other applications refer to *Harlin* [1991], who developed a process oriented calibration scheme for the Hydrologiska Byråns Vattenbalansavdelning (HBV) model, further improved by *Zhang and Lindström* [1997]. In those applications, the SCA is mostly regarded as an automated alternative to manual calibration. In the case of this paper, we provide more insights with respect to the principles and implications of this calibration strategy.

4. Case Study 1: Lower-Complexity Model Structure

4.1. FLEX^A Model Description

[21] In this study we use two versions with different complexities of the flux exchange (FLEX) hydrological model, introduced by *Fenicia et al.* [2006]. The first version, named FLEX^A, is composed of three reservoirs: an unsaturated soil reservoir (UR), which represents the storage capacity of the soil, a fast reacting reservoir (FR) accounting for the formation of fast runoff components and a slow reacting reservoir (SR), representing the slow runoff components (Figure 1).

4.1.1. Unsaturated Soil Module

[22] Rainfall R is partitioned in a component that produces runoff R_f and a component that infiltrates into the soil R_u through a rainfall excess model that assumes a distribution of storage capacity into the catchment (equations (1), (2), and (3))

$$C_r = \frac{1}{1 + \exp\left(\dfrac{-S_u/S_{fc} + 1/2}{\beta}\right)} \qquad (1)$$

$$R_f = C_r R \qquad (2)$$

$$R_u = (1 - C_r)R \qquad (3)$$

where C_r is the runoff coefficient, expressed as an S-shaped function dependent on the ratio between the storage S_u in

UR and the maximum storage S_{fc}, β is a shaping parameter, R_f is the contribution to FR, and R_u is the flux that is added to UR.

[23] Percolation P_s from UR to SR is calculated as a linear function of S_u through the coefficient P_{max}.

$$P_s = P_{max}\left(S_u/S_{fc}\right) \qquad (4)$$

[24] The potential transpiration is converted into actual transpiration according to the following formula:

$$T_a = T_p \cdot \min\left(1, \frac{S_u}{S_{fc}} \frac{1}{L_p}\right) \qquad (5)$$

where L_p is the fraction of S_{fc} below which T_p is constrained by S_u.

4.1.2. Transfer Routine

[25] As shown in Figure 1, the transfer routine of the model consists of two lag functions and two reservoirs. The two lag functions are characterized by a triangular distribution of linearly increasing weights and are defined by the parameters N_{lagf} and N_{lags} that determine the number of time steps in the transformation routine. Those functions are used to offset the flux P_s that enters SR and the flux R_f that enters FR, and mainly control the lag time of the system and the simulation of the rising limbs of the hydrograph.

[26] The FR and SR reservoirs are linear reservoirs defined by the timescales K_f and K_s respectively. The drainage equations for the two recession components are expressed as follows:

$$Q_f = S_f/K_f \qquad (6)$$

$$Q_s = S_s/K_s \qquad (7)$$

where Q_f and Q_s are the fast and slow discharges and S_f and S_s are the storages of FR and SR respectively. These reservoirs mainly control the simulation of the recession

Table 1. FLEX^A Model Parameters and Corresponding Units

Parameter	Units	Definition
S_{fc}	mm	maximum UR storage
L_p	-	limit for potential transpiration
β	-	shape parameter of runoff generation
P_{max}	mm/h	maximum percolation rate
N_{lagf}	h	lag time of FR transfer function
N_{lags}	h	lag time of SR transfer function
K_f	h	FR timescale
K_s	h	SR timescale

limbs of the hydrograph. Some more details on the characteristics of the transfer module composed lag function and reservoir are given in Appendix A. The model has a total of 8 parameters that are summarized in Table 1 together with their corresponding units.

4.2. Objective Functions Definition

[27] For this study we chose three main hydrograph characteristics that the model should correctly simulate: high flows, low flows, and timing. Model performance regarding those characteristics is evaluated by the following objective functions, respectively:

$$F_{LF} = \frac{1}{n} \left(\sum_{i=1}^{n} \left(\ln Q_{s,i} - \ln Q_{o,i} \right)^2 \right) \quad (8)$$

$$F_{HF} = \frac{1}{n} \left(\sum_{i=1}^{n} \left(Q_{s,i} - Q_{o,i} \right)^2 \right) \quad (9)$$

$$F_{LT} = 1 - R = 1 - \frac{\sum_{i=1}^{n} \left(Q_{s,i} - \overline{Q}_s \right)\left(Q_{o,i} - \overline{Q}_o \right)}{\sqrt{\sum_{i=1}^{n} \left(Q_{s,i} - \overline{Q}_s \right)^2 \sum_{i=1}^{n} \left(Q_{o,i} - \overline{Q}_o \right)^2}} \quad (10)$$

where Q represents discharge, n is the total number of time steps on the calibration period, the subscripts i, s, and o stand for time of observation, simulated, and observed respectively, and the overbar indicates the average during the observation period. Because of the use of the logarithmic function, F_{LF} weighs the error (absolute difference between observed and simulated values) on the low flows more than the error on the high flows. Therefore F_{LF} places a strong constraint on the simulation of the lower portion of the hydrograph. F_{HF} gives the same weight to the error on different portions of the hydrograph. However, considering that the error on the high flows is normally larger than the error on the low flows, and as the error is squared, F_{HF} gives a strong weight to the error in the peaks of the hydrograph. R is the correlation coefficient, which is maximized when the time shift between the fluctuations of the observed and simulated discharge is minimal. The reason of using 1-R (F_{LT}) is because the problem formulation uses a "minimization" objective.

4.3. Study Area and Data Description

[28] The study area is part of the experimental Alzette river basin located in the Grand Duchy of Luxembourg (Figure 2). For this study three years of streamflow hourly data from 1 August 2000 until 31 July 2003 are used. The data are recorded at the Hesperange gauging station, which is located along the main course of the Alzette River and drains a catchment of 288 km^2. The land use consists of agriculture (27%), grass (26%), forest (29%) and urban area (18%). The lithology is mainly characterized by marl and marly sandstone on the left bank tributaries, and limestone on the right bank tributaries of the Alzette River. Limestone areas represent zones of infiltration of rainfall water, and constitute a permeable aquifer, which represents the main reservoir that sustains the streamflow during dry weather periods. Marl areas are relatively impermeable to rainfall, therefore determining high runoff volumes during or shortly after rainfall events, and little or no discharge during dry periods. The density of rain gauges in the study site is of about one instrument per 30 km^2. Instruments consist of tipping buckets and automatic samplers that measure at different time intervals of 20 min or shorter. Hourly rainfall is calculated by averaging the time series at the various stations through the Thiessen polygon method. Daily potential evaporation series are calculated through the Penman-Monteith formula using temperature, wind speed, humidity and net radiation. The necessary data were measured at the meteorological station located at Luxembourg airport. Hourly estimates are calculated distributing the daily amounts using a sine curve distribution between the hours of sunlight.

4.4. Application of the Pareto-Based Calibration Approach

[29] The goal of multiobjective optimization is to sample the search space in such a way that the sampling converges toward the globally Pareto-optimal set of solutions. As an efficient multiobjective optimization algorithm, the multiobjective shuffled complex evolution Metropolis algorithm (MOSCEM-UA) [*Vrugt et al.*, 2003] has been used. The MOSCEM-UA algorithm starts by randomly generating a number s of samples within the feasible parameter space, which are subsequently sorted based on their objective function values. The points are partitioned into q complexes, and in each complex a parallel sequence is launched. The sequences proceed by generating a new candidate point from a multivariate normal distribution with the mean at the current value and covariance matrix derived from the history of each sequence. The new candidate point is added to the sequence based on the outcome of a Metropolis-type of acceptance rule. The algorithm proceeds until a predefined number of iterations n is reached. On the basis of the research of *Vrugt et al.* [2003] the iterative application of the various algorithmic steps causes the population to converge toward the globally Pareto-optimal set of solutions. The algorithm has three algorithmic parameters that have to be selected: s, q and n. In this study, for the FLEX^A model, which has 8 parameters to be optimized, the initial number of random samples s is set at 1000, the number of complexes q is set at 10 and the maximum number of iterations n at 10,000.
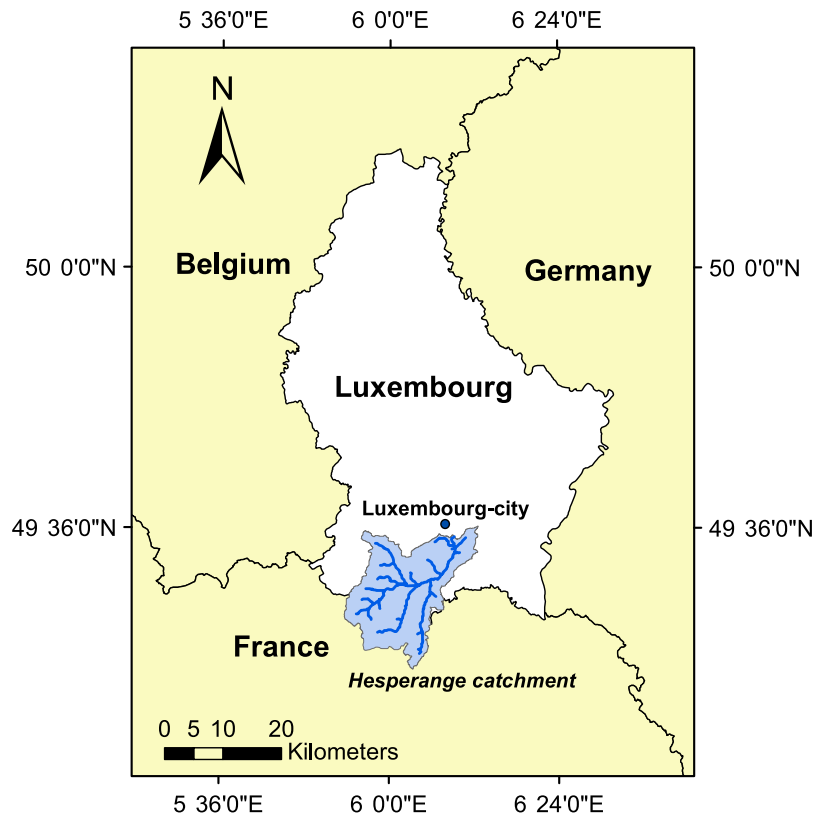
Figure 2. Location of the Hesperange catchment in Luxembourg.

[30] The outcomes of the study are represented in Figure 3. It can be seen that significant tradeoffs exist among the various objective functions, which indicates that no single parameter set is able to simultaneously optimize all individual objectives.

[31] The various Pareto-optimal solutions correspond to representations of the hydrograph that can be significantly different. As an example, the hydrographs corresponding to the individual solutions that optimize F_{HF} and F_{LF} are represented in Figure 4. It is possible to observe that the individual solutions are visibly different. Figure 5 shows a scatterplot of model residuals with respect to observed flow. For lower flows the scatter shown by the best low-flow

model tends to be lower than that corresponding to the high-flow model. For higher flows the best low-flow model tends to underestimate hydrograph peaks, while the high-flow model shows a more centered scatter. The initial parameter ranges and the ranges corresponding to the Pareto-optimal solutions are represented in Table 2.

4.5. Application of the Stepped Calibration Approach

[32] The application of the SCA requires the association of model parameters with a specific objective function. The association is performed based on the specific role that model parameters have on the modeled processes and on the main role that each process has on the composition of the
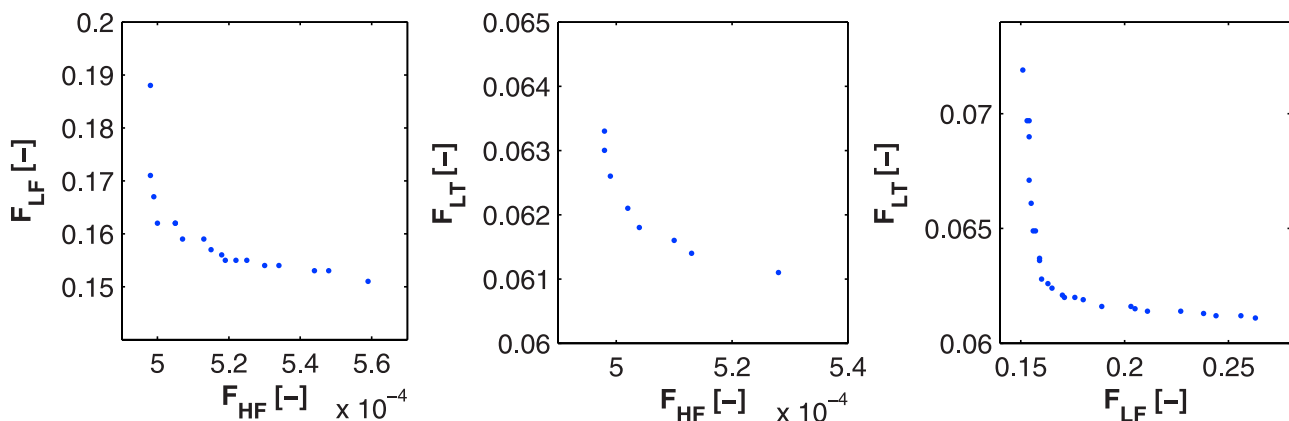

Figure 3. Estimated Pareto-optimal fronts according to the selected objective functions for the FLEX[A] model.
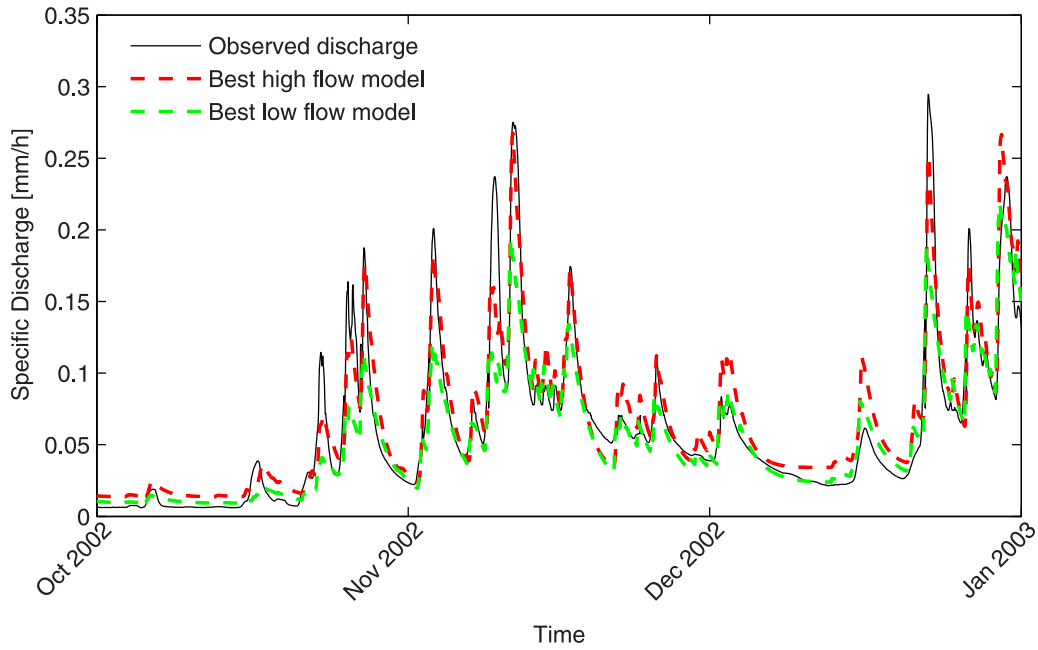
Figure 4. Comparison of FLEX^A model simulations corresponding to the solutions that optimize each of the individual objectives F_{HF} and F_{LF}.

integrated signal that represents the global response of the catchment. This means that each parameter is calibrated on the process that it is designed to influence, even if it also influences other processes (sometimes even more than the process it is supposed to influence). If, for instance, a groundwater related parameter is found to have a strong influence on high-flow reproduction, which is possible in case of structural deficiencies of a model, in the SCA this parameter will be adjusted to fit the base flow dynamics rather than the peaks, as the base flow is the process that such parameter is designed to represent. Hence the association of model parameters with a specific objective function is not determined on the basis of prior sensitivity analysis,

but on the judgment of the modeler and his subjectivity to introduce certain processes and components into the model structure.

[33] In the present case, the following parameter associations are used:

$$F_{LF} \leftarrow P_{\max}, K_s$$
$$F_{HF} \leftarrow L_p, S_{fc}, \beta, K_f$$
$$F_{LT} \leftarrow N_{lagf}, N_{lags}$$

[34] This association is based on the observations that $P_{\max}$ and K_s influence the parameterization of SR, which is designed to represent the groundwater processes. K_f
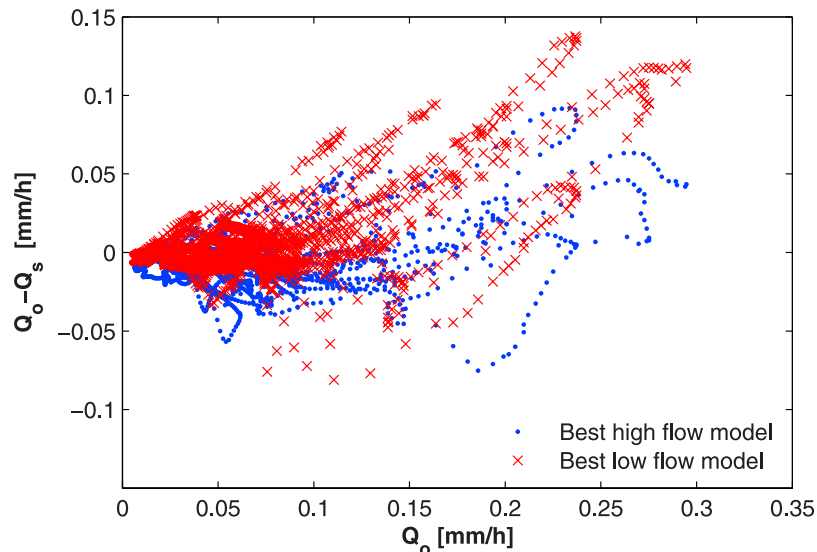

Figure 5. Residuals plots of the FLEX^A model simulations corresponding to the solutions that optimize F_{HF} and F_{LF}.

Table 2. Initial and Final Parameter Ranges

Parameter	Units	Initial Range	Final Range
S_{fc}	mm	$200-600$	$310-422$
L_p	-	$0.1-1.0$	$0.18-0.47$
β	-	$0.05-3.00$	$0.097-0.144$
P_{max}	mm/h	$0.04-0.20$	$0.060-0.094$
N_{lagf}	h	$1-10$	$2.88-7.42$
N_{lags}	h	$2-600$	$4.72-554$
K_f	h	$20-100$	$29-44$
K_s	h	$150-2000$	$420-1197$

influences the recession after peak flow. S_{fc}, and β play a role in the nonlinearity of the rainfall-infiltration-direct runoff relation, which to our judgment plays a major role in the simulation of peaks, since the peak depends on the antecedent moisture condition. L_p is a parameter that affects transpiration and is connected to the total volume of water discharged. In this case, as this parameter is related to the parameterization of UR, we associate it to F_{HF} rather than to the other calibration objectives. N_{lagf} and N_{lags} have been introduced to correctly simulate the lag time of the system.

[35] The calibration phase proceeds in three steps: in the first step all parameters are calibrated against F_{LF}, then the parameters related to high flows are recalibrated with respect to F_{HF}, and finally the parameters related to the lag time of the system are recalibrated toward F_{LT}. In the second and third calibration steps, parameters that are not recalibrated are kept constant at the values determined during the previous steps.

[36] As a search method to identify the global optimum in the parameter space we have selected the adaptive cluster covering (ACCO) strategy with local search developed by *Solomatine* [1995, 1999], which proves to be effective and efficient in global optimization problems. This algorithm is implemented in the global optimization tool GLOBE [*Solomatine*, 1999], which has been configured to calibrate the model parameters.

[37] The hydrograph reproduction after the first two steps is shown in Figure 6. Compared to Figure 4, showing the two best models with respect to low flow and high flow, respectively, we observe that the recalibration of high-flow-related parameters only has an impact on the peaks, while leaving the reproduction of low flows almost unchanged. The effect of the third calibration step is shown in Figure 7. The "loops" in the scatterplots, which indicate a time lag between observed and modeled watershed behavior, become smaller after the recalibration of lag time related parameters.

[38] A comparison of the outcomes of the two approaches is shown in Figure 8. Each point in the objective function space can be represented by the multiobjective vector **F** (F_{LF}, F_{HF}, F_{LT}). The three subplots represent a projection in two dimensions of the three-dimensional criterion space showing the Pareto-optimal front corresponding to the three calibration objectives as well as the successive steps $\mathbf{F}^{S1}$, $\mathbf{F}^{S2}$, and $\mathbf{F}^{SCA}$ obtained through the application of the SCA. The following steps of the SCA determine a progressive deterioration of the first calibration objective, and a progressive improvement in the other two objectives.

[39] The multiobjective vector $\mathbf{F}^{SCA}$ does not necessarily represent a Pareto-optimal point. The progression of steps in fact determines orthogonal movements in the parameter space and therefore disregards the eventual influence that each parameter might have on each objective function. However, as this solution represents a balance of calibration objectives, it can be exploited to develop optimal solutions that reflect such a balance. This way, the two approaches can be used in a combined, synergistic manner that exploits the strengths of each.

[40] A possibility to identify optimal balanced solutions could be to exclude from the Pareto-optimal set of solutions those parameter sets that correspond to objective function values that exceed in at least one of the objectives the solution developed through the SCA. If $\mathbf{F}^{SCA}$ is regarded as a balanced solution, all the remaining solutions, which are
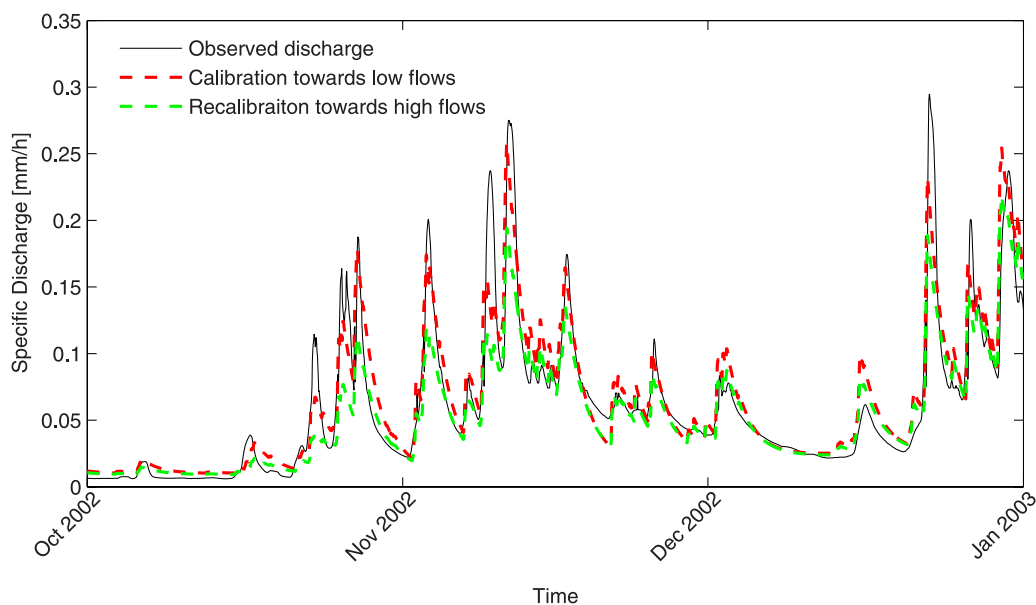

Figure 6. Comparison of hydrograph simulations corresponding to the first two calibration steps of the SCA generated by the FLEX^A model.
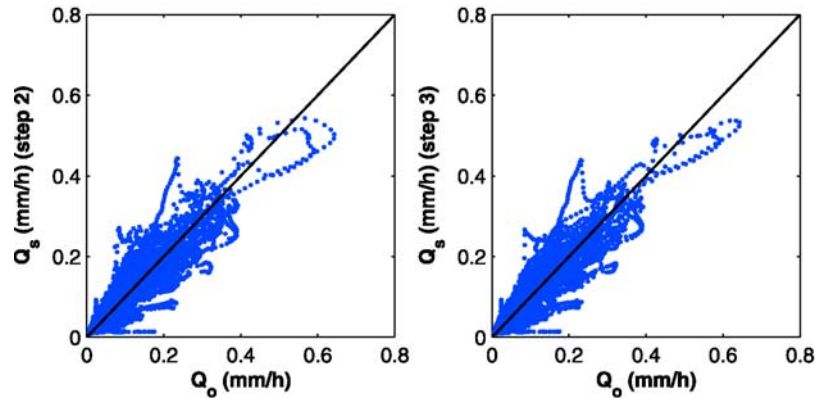
Figure 7. Scatterplot showing the improvement following the application of the third calibration step applied to the FLEX^A model.

characterized by F values that are all lower than those of $\mathbf{F}^{SCA}$, would demonstrate an improvement with respect to $\mathbf{F}^{SCA}$ and would therefore be preferable. As an example, Figure 8 shows the solution that minimizes the Euclidean distance to the line that connects $\mathbf{F}^{SCA}$ to the origin of the axes. In order to calculate such a distance, the objective space has been previously normalized by dividing each variable by the correspondent value of $\mathbf{F}^{SCA}$. The normalization operation is necessary to avoid giving a larger weight to objectives with larger F values. It is possible to observe that the so calculated balanced optimum $\mathbf{F}^{BAL}$ demonstrates an improvement with respect to all F values of $\mathbf{F}^{SCA}$. As an indication of such an improvement and of the progress of the successive calibration steps, the Euclidean distance of each point from $\mathbf{F}^{SCA}$ in the normalized objective space is reported in Table 3.

4.6. Posterior Parameter Sensitivity Analysis

[41] In order to illustrate the implications of the SCA in the parameter space, a posterior parameter sensitivity analysis is performed. The objective of the analysis is twofold. First it serves to evaluate the parameter sensitivity around the optimal parameter values. Second, it shows the effects on the optimal parameter values subsequent to the recalibration of model parameters at successive calibration steps.

[42] The sensitivity analysis is performed stepwise. Each group of parameters is evaluated with respect to each objective function in successive stages. During each stage, the parameters that are not concerned are kept constant at the values determined through the SCA. The parameter

sensitivity is expressed by the parameter ranges that correspond to values of the objective function that differ less than 10% from the optimum value. As the optimum value of each group of parameters might change during the recalibration of other parameters at successive calibration stages, the 10% threshold is evaluated with respect to the new value of the optimum. Since the ranges where to perform the sampling are not known in advance, and in order to perform an efficient sampling in the region of interest, a Markov chain sampling strategy has been adopted. The sampling strategy uses a Metropolis-type of acceptance rule and an adaptive algorithm proposed by *Haario et al.* [2001], which updates the covariance matrix at predefined steps. In this analysis, for each group of model parameters a total of 10,000 parameter samples were generated, with a burn-in period of 1000 iterations discarded for subsequent analyses.

[43] Parameter samples obtained are summarized by the density histograms shown in Figure 9. The conclusions that can be drawn from the analysis are summarized as follows.

[44] 1. The ranges of variation that are calculated indicate that model parameters are in general well identified within relatively small intervals. A noteworthy exception is the parameter N_{lags}, which simulates the lag time of the system related to groundwater flow. The fact that this parameter shows such a high degree of variability indicates that the process cannot be identified with the data available or with the objective functions used.

[45] 2. The optimal values of the calibrated model parameters change when other parameters are recalibrated, indicating that the optimum of a group of parameters associated
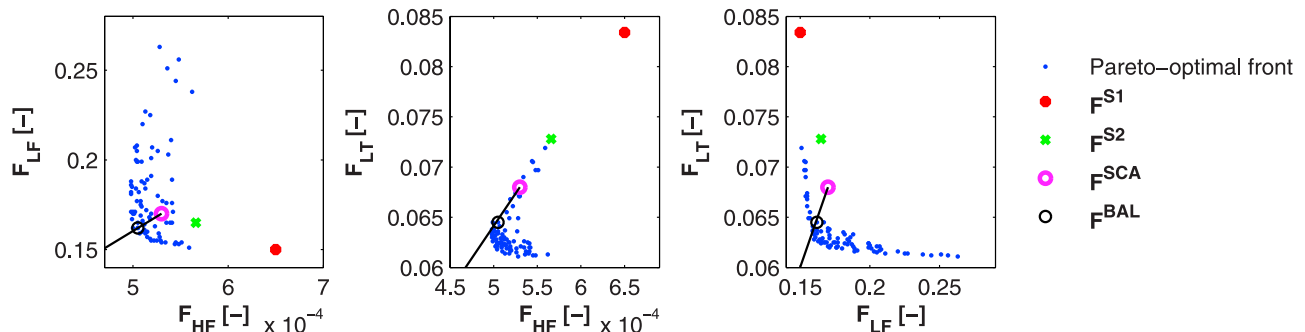

Figure 8. Comparison of the Pareto-based approach and the SCA applied to the FLEX^A model.

Table 3. Euclidean Distance in the Normalized Objective Space From the Solution Developed Through the SCA[a]

Parameter	Value
$\overline{F^{SCA}F^{S1}}$	3.41E-01
$\overline{F^{SCA}F^{S2}}$	1.01E-01
$\overline{F^{SCA}F^{BAL}}$	8.61E-02
$\overline{F^{SCA}F^{R1}}$	1.81E-02
$\overline{F^{SCA}F^{R2}}$	6.48E-02

[a]Read 3.41E-01 as 3.41×10^{-1}.

with one calibration objective also depends on the value of the parameters that are related to the other calibration objectives. When the parameters related to the other objectives change, because of recalibration, the optimal value of the already calibrated parameters changes as well. The most biased parameters are the ones that are calibrated first, while the parameters that are calibrated last, in this case the lag time related parameters, are not biased. Figure 9 shows the optimum corresponding to the parameter values determined through the SCA, and the new optimum that corresponds to the recalibration of each group of parameters while the other parameters are kept constant at the values determined through the SCA. It is possible to observe that the two

values differ. The difference however is not large, as the new optimum remains close to the initial one in most cases, and in all cases within the calculated confidence intervals. The distance between the parameter values due to recalibration with respect to their range of variation is indicated in Table 4. In order to show the effect of recalibration in the objective space, the distance between the solution $\mathbf{F}^{SCA}$ and the solutions corresponding to recalibration of the first and second groups of parameters $\mathbf{F}^{R1}$ and $\mathbf{F}^{R2}$ are indicated in Table 3. It can be observed that those distances are smaller than those corresponding to the solutions of the first two calibration steps $\mathbf{F}^{S1}$ and $\mathbf{F}^{S2}$.

5. Case Study 2: Higher-Complexity Model Structure

[46] The comparison between observations and model results allows an analysis of the hypotheses made, as it indicates whether the conceptualization proposed is acceptable, or, to the contrary, it is not adequate and requires further improvements. In the latter case, it is possible to go back to the conceptual model and modify it by introducing different processes and components. In this sense, the modeling becomes a learning process, as the attempt to model the hydrological processes allows a testing of our
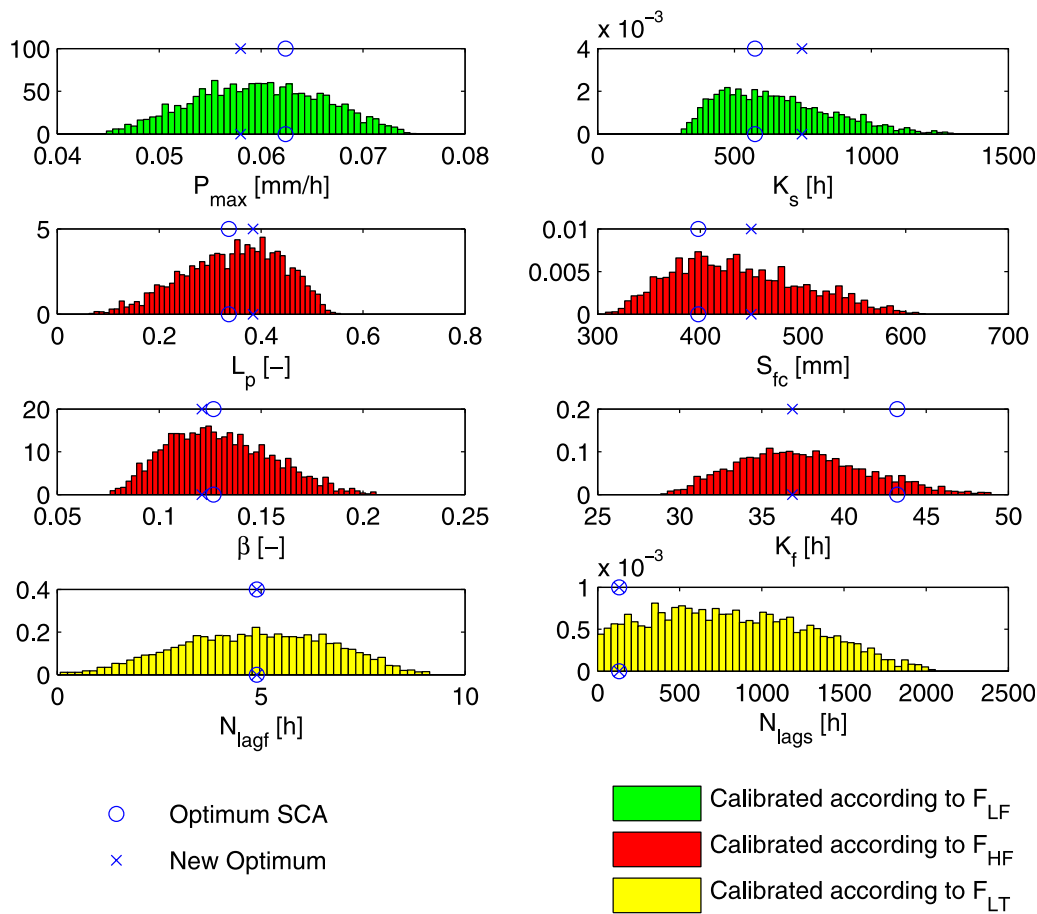

Figure 9. Sensitivity of FLEX[A] model parameters.

Table 4. Distance of Parameter Values due to Recalibration With Respect to Their Range of Variation

Parameter	Value
P_{max}	6.2E + 00%
K_s	4.0E-05%
L_p	9.6E + 00%
S_{fc}	1.7E + 01%
β	4.4E + 00%
K_f	3.2E + 01%

perceptions of the hydrological behavior of a catchment [Beven, 2001].

[47] In the present case, the application of the two calibration approaches provides some useful information about the structural limitations of the model used. The existence of significant tradeoffs between model parameters and calibration objectives represents a symptom of model structural limitations. The application of the SCA develops a compromise solution that at a careful analysis of the observed and simulated hydrographs shows poor overall performances.

[48] The analysis of calibration results as well as some insights on the dominant hydrological processes in the study area have led to a more complex version of the model, which is for convenience named FLEX^B. With respect to the previous version, the new model structure adds an interception reservoir (IR), described by two calibration parameters, and introduces a process of preferential recharge, which requires the definition of an additional calibration parameter (Figure 10). In total, the FLEX^B model structure presents 11 parameters, three more than the FLEX^A model.

[49] The interception process has been included because, apart from being an essential process in the hydrologic cycle, it is also indicated as one of those processes that could improve the accuracy of the simulation. In this case,

the careful examination of the observed and simulated hydrographs has pointed out that some systematic overestimations of the low flows took place in correspondence of rainfall events that occurred after prolonged dry conditions. Those overestimations could be attributed to interception, which the model does not simulate. The interception process is often neglected in conceptual models, however it is an important process that might have significant repercussions on the resulting hydrograph [Savenije, 2004]. This was also concluded by Zhang and Savenije [2005] who improved overall model performance by including an interception component in the model conceptualization.

[50] Preferential recharge has been introduced in order to better characterize the process of recharging the groundwater reservoir during and shortly after rainfall events. In limestone areas, part of the rainfall quickly reaches the groundwater through preferential pathways, without being stored in the soil matrix. As shown in the structure description in the following paragraph, the amount of water that reaches the aquifer through preferential recharge is made implicitly dependent on the storage of UR. This caters for the fact that for wet conditions the preferential recharge is a higher fraction of the effective rainfall than for dry conditions.

5.1. FLEX^B Model Description

[51] The structure of the model is shown in Figure 10 and the differences with the FLEX^A version are here briefly described.

5.1.1. Interception Module

[52] Rainfall reaches the IR, which can be filled up to a specified threshold, represented by I_{max}. Evaporation from intercepted water E_i can occur as long as water is available in the reservoir, and it is assumed to be linearly related to the potential evaporation E_p through the coefficient I_c:
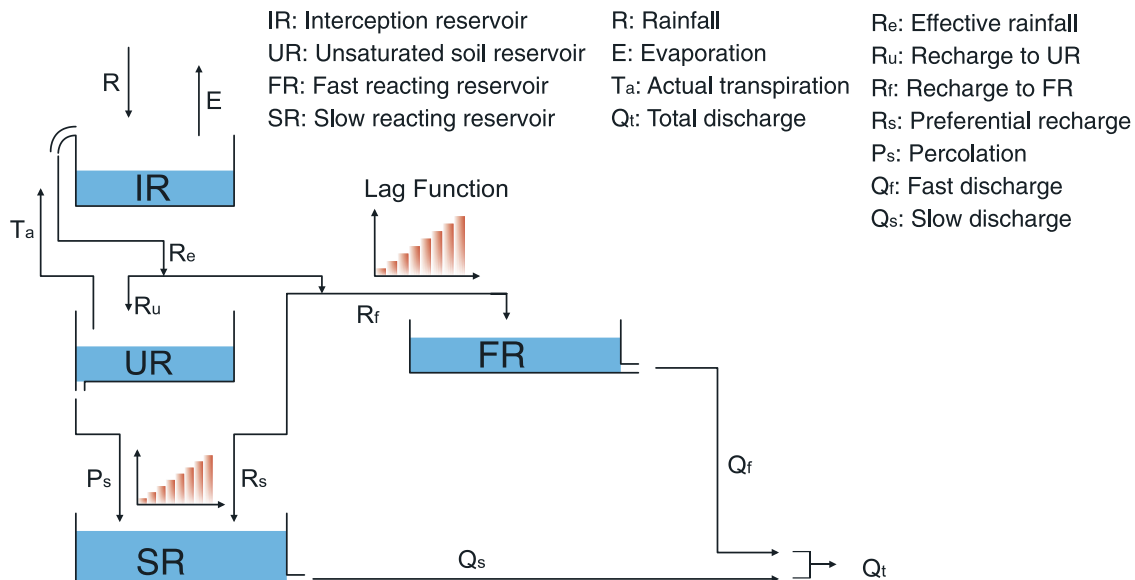
$$E_i = I_c E_p \tag{11}$$

IR: Interception reservoir   R: Rainfall   Re: Effective rainfall
UR: Unsaturated soil reservoir   E: Evaporation   Ru: Recharge to UR
FR: Fast reacting reservoir   Ta: Actual transpiration   Rf: Recharge to FR
SR: Slow reacting reservoir   Qt: Total discharge   Rs: Preferential recharge
Ps: Percolation
Qf: Fast discharge
Qs: Slow discharge

Figure 10. Structure schematization of the FLEX^B model.

Table 5. FLEX[B] Model Parameters and Corresponding Units

Parameter	Units	Definition
I_c	-	evaporation coefficient
I_{max}	mm	interception threshold
S_{fc}	mm	maximum UR storage
L_p	-	limit for potential transpiration
β	-	shape parameter of runoff generation
D	-	runoff partition coefficient
P_{max}	mm/h	maximum percolation rate
N_{lagf}	h	lag time of FR transfer function
N_{lags}	h	lag time of SR transfer function
K_f	h	FR timescale
K_s	h	SR timescale

5.1.2. Unsaturated Soil Module

[53] Effective rainfall R_e leaves the IR when the threshold I_{max} is exceeded. This amount is then partitioned into various components based on the value of an effective (i.e., after subtraction of interception [see *Savenije*, 2004]) runoff coefficient C_r, which is expressed by the same formula used in the previous structure (equation (1)). Part of R_e infiltrates into UR (R_u), excess water from UR is partitioned through the coefficient D into preferential recharge R_s, which flows to SR, and surface runoff R_f, which enters FR (equations (12), (13), and (14)).

$$R_u = (1 - C_r)R_e \qquad (12)$$

$$R_s = (R_e - R_u)D \qquad (13)$$

$$R_f = R_e - R_u - R_s \qquad (14)$$

[54] Percolation P_s from UR to SR and the potential transpiration are calculated as in the FLEX[A] model through equations (4) and (5), respectively.

5.1.3. Transfer Routine

[55] The transfer routines remain the same as in the FLEX[A] model, with the difference that the fluxes P_s and R_s are first added and their sum is routed through the triangular transfer function. The transformed flux is then added to SR. The 11 model parameters are summarized in Table 5.

5.2. Application of the Pareto-Based Calibration Approach and SCA to the FLEX[B] Model Structure

[56] The two calibration approaches are applied to the FLEX[B] model structure. The application uses the same "set up" of the calibration problem as in the FLEX[A] model version. For the application of the Pareto-based approach,

the MOSCEM-UA algorithm has been applied with the same algorithmic parameters as for the FLEX[A] version except from the number of complexes q, which has been increased to 15 as a result of the increased number of model parameters.

[57] The SCA has been applied with the following parameter associations:

$$\begin{aligned} F_{LF} &\leftarrow I_{max}, I_c, P_{max}, D, K_s, \\ F_{HF} &\leftarrow L_p, S_{fc}, \beta, K_f \\ F_{LT} &\leftarrow N_{lagf}; N_{lags} \end{aligned}$$

[58] The newly introduced parameters are all associated to the F_{LF} objective function. This is motivated by considering that the parameter D is introduced to better characterize the groundwater recharge process, and the interception processes should manifest its impact on the final hydrograph by improving the simulation of the catchment response to rainfall after prolonged dry conditions.

[59] The comparison of the two calibration approaches for the two model structures is shown in Figure 11. We observe that the Pareto-optimal fronts corresponding to the two model structures are clearly different, indicating that the increase of complexity has an explanatory power on the observed data. The more complex model is also a better model, as it produces a visible improvement of the accuracy of the simulation.

[60] The best models corresponding to F_{LF} and F_{HF} are represented in Figure 12. If compared to Figure 4, which represents the same situation for the FLEX[A] structure, we can see that the two models corresponding to the FLEX[B] structure represent the hydrograph better than any of the two models calibrated on the FLEX[A] structure. Moreover, the two models developed for the FLEX[B] structure are visually closer to each other than the corresponding ones for the FLEX[A] structure. This also appears from the scatterplot shown in Figure 13. The residuals shown by the two models vary in a similar interval in the range of observed discharges. On the objective function space, the Pareto-optimal front and the points that correspond to the stages followed in the application of the SCA are also more narrowly spaced.

[61] The hydrograph improvement obtained at following stages during the application of the SCA on the FLEX[B] model is less dramatic than for the FLEX[A] structure. This indicates that when a model has the potential of reproducing the data well, a calibration based on a single objective
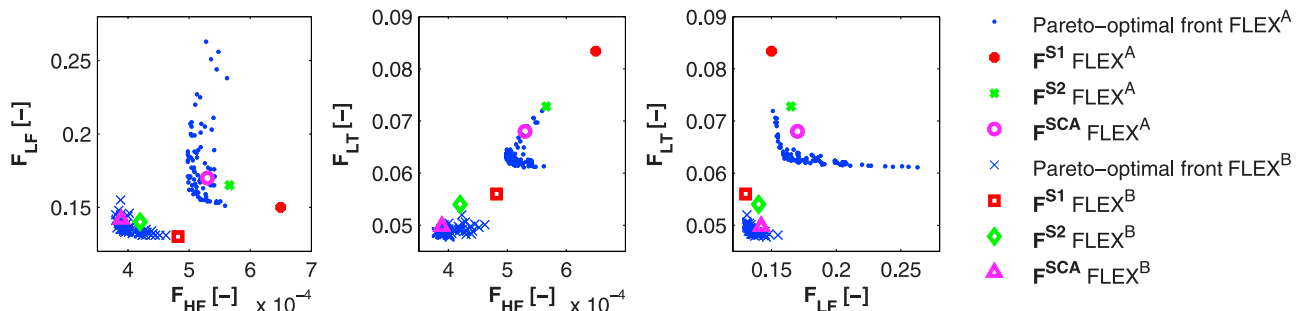

Figure 11. Comparison of the Pareto-based calibration approach and SCA for the FLEX[A] and FLEX[B] model structures in the objective space.
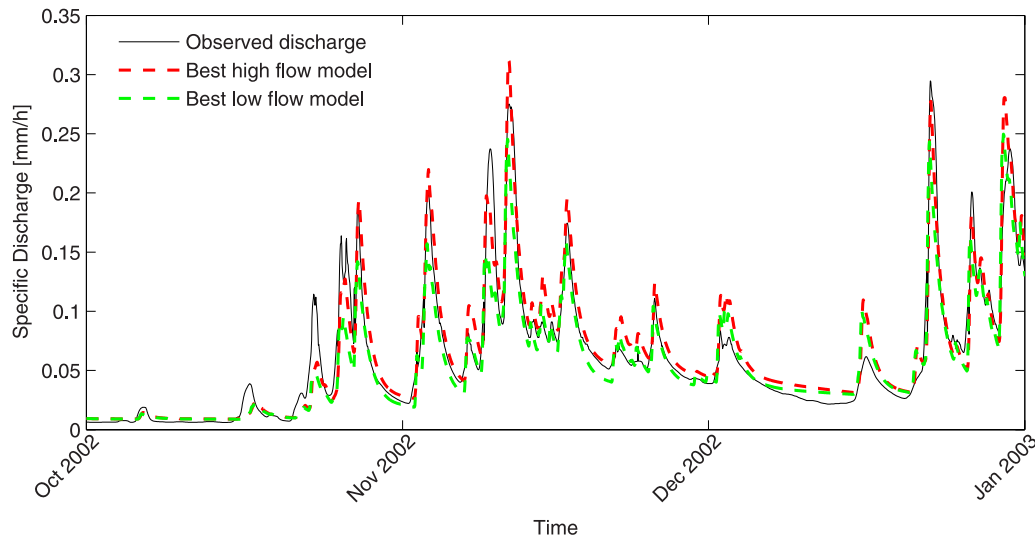
Figure 12. Comparison of FLEX^B model simulations corresponding to the solutions that optimize each of the individual objectives F_{HF} and F_{LF}.

taking into account the overall model performance can lead to good results, and the difference between the optimal simulations associated to different objective functions becomes less evident.

6. Conclusions

[62] This paper presents a comparison between two different calibration approaches and discusses their principles and implications in hydrological modeling. The first calibration approach is based on the concept of Pareto-optimality, and develops a set of optimal solutions according to the tradeoffs between different objectives. The second approach associates different groups of parameters with specific calibration objectives based on the processes that those parameters are designed to influence. This approach replicates, in an automated fashion, the steps that are

undertaken by operational hydrologists during manual calibration, and develops one optimal parameter set that is considered acceptable according to the selected calibration objectives.

[63] The analysis is performed on two model structures of different levels of complexity. Initially the two approaches are evaluated on a model structure of lower complexity. Subsequently, based on the outcomes of the analysis and the hydrological understanding of the system in question, the model structure is improved by adding new processes and components. The calibration analysis is repeated for the model structure of higher complexity, and calibration results are again analyzed. The main conclusions that can be drawn from this work are summarized as follows.

[64] 1. The Pareto-based calibration approach can be useful to visualize the structural limitations of a model, in terms of the inability of the model to reproduce the
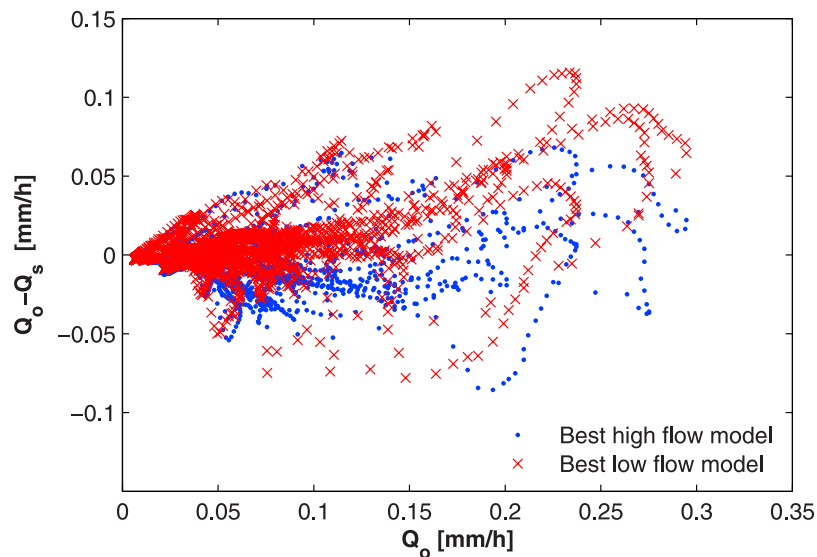

Figure 13. Residuals plots of the FLEX^B model simulations corresponding to the solutions that optimize F_{HF} and F_{LF}.
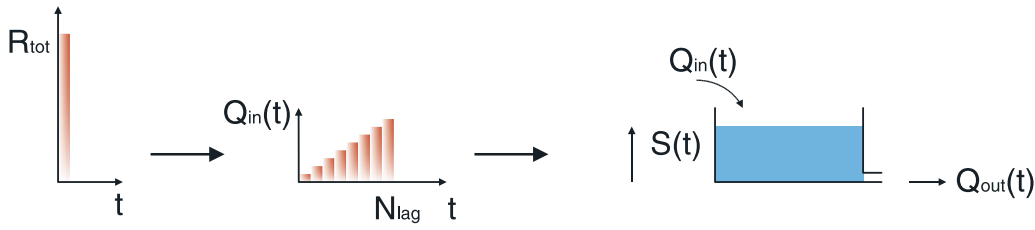
Figure A1. Schematization of an instantaneous rainfall through the transfer module.

observations. The Pareto-optimal set of solutions provides model simulations that are all equally important from a multiobjective optimization point of view, but can contain solutions that can be considered unacceptable with respect to the accomplishment of the various criteria that are demanded for calibration. Acceptable solutions, in fact, are likely to be contained in the Pareto-optimal set of solutions, but there may be many unacceptable solutions within the set.

[65] 2. The Pareto-based approach is also useful to compare merits of different models and to track changes of model performance as model structures are modified. Model improvements can in fact be identified as the Pareto-optimal front progressively moves toward the origin of the objective function space. This property has been used to evaluate the effect of a structural modification of the initial model.

[66] 3. The stepped calibration approach works more in a "calibration" optic than in an "optimization" optic. The aim is to determine a single parameter set that causes model components to reproduce the processes that they are designed to represent. The parameter set that is developed is not necessarily a Pareto-optimal solution; however, as it represents a balance of calibration objectives, it can be used to determine optimal solutions that reflect such a balance. This way, the two calibration approaches can be used in a combined, synergistic manner that exploits the strengths of each.

[67] 4. The more complex model structure has been developed from the simpler version by adding the interception process and a better representation of groundwater recharge. The multiobjective analysis leads to the conclusion that the new model structure allows a better accuracy of the simulation of the observed data. The additional complexity clearly contributes to a better understanding of the observed system behavior.

[68] 5. In subsequent steps, the model of higher complexity displays a Pareto-optimal front which is more narrowly spaced than for the lower-complexity model, moreover the hydrograph simulations corresponding to the single best solutions associated to different objectives and to the various steps of the SCA are closer to each other. This indicates that when a model structure has a high capability of simulating the observations, the use of different objective functions, even if they stress the simulation of different aspects of the simulations, will lead to similar results. This suggests that sometimes, instead of putting much effort on trying to improve the fit of a poor model by sophisticated calibration processes, it is more efficient to try to understand the model's limitations and to correct these by an improved schematization.

Appendix A: Transfer Module Details

[69] In this section some more details on the properties of the transfer module composed by the triangular transfer function and the linear reservoir are given. This module could be interesting for hydrological applications as it is described by parameters that represent specific outflow characteristics. Figure A1 shows the processing of an instantaneous rainfall R_{tot} through the transfer module. The discharge in time $Q_{out}(t)$ from the reservoir can be calculated by solving the system of closure and balance equations for the reservoir. N_{lag} represents the length of the distribution function, and the discharge can be calculated separately for $t \leq N_{lag}$ and for $t > N_{lag}$.

A1.   Calculation for $t \leq N_{lag}$

[70] The input to the reservoir $Q_{in}(t)$ can be written as follows:

$$Q_{in}(t) = Bt \qquad (A1)$$

where

$$B = \frac{2R_{tot}}{N_{lag}^2} \qquad (A2)$$

[71] The output from the reservoir $Q_{out}(t)$ is linearly proportional to the level in the reservoir $S(t)$ through the coefficient K. The storage function can be represented as:

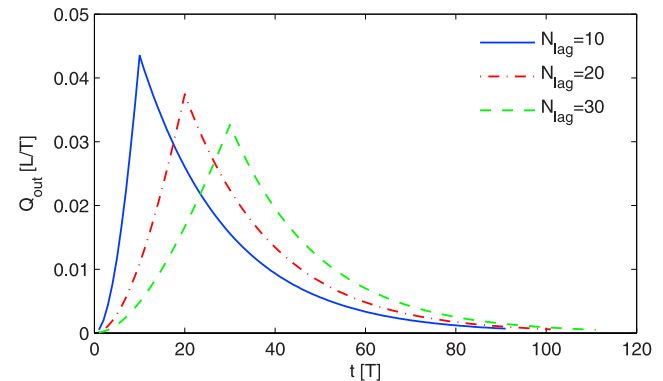$$Q_{out}(t) = \frac{S(t)}{K} \qquad (A3)$$

Figure A2. Q_{out} for a unitary rainfall, $K = 20$ time units, and different values of N_{lag}.

[72] The water balance in the reservoir can be expressed by the following continuity equation:

$$\frac{dS}{dt} = Q_{in} - Q_{out} \qquad (A4)$$

[73] Therefore

$$\frac{dS}{dt} + \frac{S}{K} = Bt \qquad (A5)$$

which is a first-order linear differential equation, whose solution is expressed by

$$S(t) = BKt - BK^2 + Ce^{-\frac{t}{k}} \qquad (A6)$$

where C is an arbitrary constant of integration.

[74] C can be determined imposing the constraint that for $t = 0$, $S = 0$ (empty reservoir for $t = 0$), and equation (A6) becomes

$$S(t) = BKt - BK^2\left(1 - e^{-\frac{t}{K}}\right) \qquad (A7)$$

which allows calculating Q_{out} through equation (A3).

A2. Calculation for $t > N_{lag}$

[75] For $t > N_{lag}$ the reservoir will empty as a linear reservoir. The maximum storage reached is

$$S_{\max} = S(N_{lag}) = BKN_{lag} - BK^2\left(1 - e^{-\frac{N_{lag}}{K}}\right) \qquad (A8)$$

[76] From this level the reservoir will empty in time according to

$$S(t) = S_{\max}e^{-\frac{t-N_{lag}}{K}} \qquad (A9)$$

and Q_{out} can be calculated through equation (A3).

[77] Figures A2 and A3 show the outflow from the reservoir for different values of N_{lag} and K. The maximum storage in the reservoir is reached for $t = N_{lag}$, the parameter N_{lag} therefore represents the time to peak of the system. K affects the slope of the falling and rising limbs of the
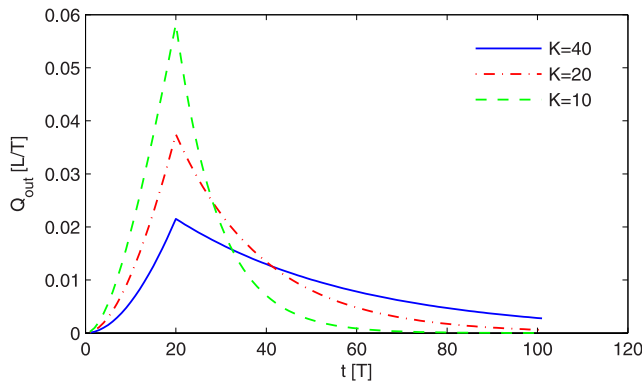

Figure A3. Q_{out} for a unitary rainfall, $N_{lag} = 20$ time units, and different values of K.
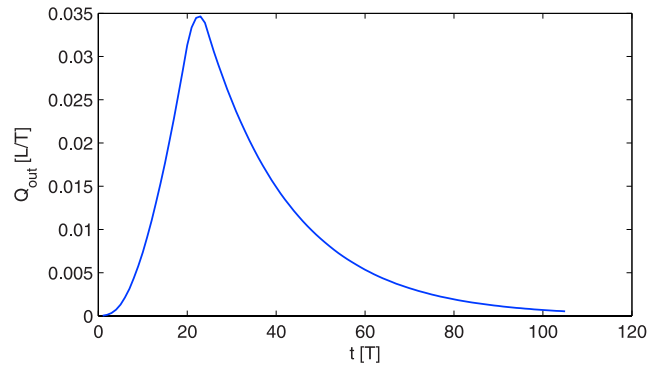

Figure A4. Transfer module in operational conditions: effect of unitary rainfall distributed in five time steps ($N_{lag} = 20$, $K = 20$).

hydrograph. This separation of the role that different parameters have on the outflow can be useful as these parameters can be calibrated for the aspects they influence. Figure A4 shows the outflow for a constant rainfall distributed over five time units. The shape resembles that of a natural hydrograph.

References

Bastidas, L. A., H. V. Gupta, S. Sorooshian, W. J. Shuttleworth, and Z. L. Yang (1999), Sensitivity analysis of a land surface scheme using multi-criteria methods, *J. Geophys. Res.*, *104*, 19,481–19,490.

Beven, K. J. (2001), *Rainfall-Runoff Modelling: The Primer*, John Wiley, Hoboken, N. J.

Boyle, D. P., H. V. Gupta, and S. Sorooshian (2000), Towards improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resour. Res.*, *36*, 3663–3674.

Crow, W. T., E. F. Wood, and M. Pan (2003), Multiobjective calibration of land surface model evapotranspiration predictions using streamflow observations and spaceborne surface radiometric temperature retrievals, *J. Geophys. Res.*, *108*(D23), 4725, doi:10.1029/2002JD003292.

Fenicia, F., H. H. G. Savenije, P. Matgen, and L. Pfister (2006), Is the groundwater reservoir linear? Learning from data in hydrological modelling, *Hydrol. Earth Syst. Sci.*, *10*, 139–151.

Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, *34*, 751–763.

Haario, H., E. Saksman, and J. Tamminen (2001), An adaptive Metropolis algorithm, *Bernoulli*, *7*, 223–242.

Harlin, J. (1991), Development of a process oriented calibration scheme for the HBV Hydrologic Model, *Nord. Hydrol.*, *22*, 15–36.

Hogue, T. S., S. Sorooshian, H. Gupta, A. Holz, and D. Braatz (2000), A multi-step automatic calibration scheme for river forecasting models, *J. Hydrometeorol.*, *1*, 524–542.

Madsen, H. (2000), Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, *J. Hydrol.*, *235*, 276–288.

Madsen, H. (2003), Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Adv. Water Res.*, *26*, 205–216.

Meixner, T., L. A. Bastidas, H. V. Gupta, and R. C. Bales (2002), Multi-criteria parameter estimation for models of stream chemical composition, *Water Resour. Res.*, *38*(3), 1027, doi:10.1029/2000WR000112.

Savenije, H. H. G. (2004), The importance of interception and why we should delete the term evapotranspiration from our vocabulary, *Hydrol. Processes*, *18*, 1507–1511.

Solomatine, D. P. (1995), The use of global random search methods for model calibration, in *Proceedings of the XXVIth Congress of IAHR (London, September 1995)*, vol. 1, pp. 224–229, Thomas Telford, London.

Solomatine, D. P. (1999), Two strategies of adaptive cluster covering with descent and their comparison to other algorithms, *J. Global Optim.*, *14*, 55–78.

Sorooshian, S., and V. K. Gupta (1995), Model calibration, in *Computer Models of Watershed Hydrology*, edited by V. P. Singh, pp. 23–68, Water Resour. Publ., Highlands Ranch, Colo.

Vrugt, J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian (2003), Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, *39*(8), 1214, doi:10.1029/2002WR001746.

Wagener, T., D. P. Boyle, M. J. Lees, H. S. Wheater, H. V. Gupta, and S. Sorooshian (2001), A framework for the development and application of hydrological models, *Hydrol. Earth Syst. Sci.*, *5*, 13–26.

Xia, Y., A. J. Pitman, H. V. Gupta, M. Leplastrier, A. Henderson-Sellers, and L. A. Bastidas (2002), Calibrating a land surface model of varying complexity using multi-criteria methods and the Cabauw data set, *J. Hydrometeorol.*, *3*, 181–194.

Xia, Y., Z. L. Yang, P. L. Stoffa, and M. K. Sen (2005), Optimal parameter and uncertainty estimation of a land surface model: Sensitivity to parameter ranges and model complexities, *Adv. Atmos. Sci.*, *22*, 142–157.

Zhang, G. P., and H. H. G. Savenije (2005), Rainfall-runoff modelling in a catchment with a complex groundwater flow system: Application of the representative elementary watershed (REW) approach, *Hydrol. Earth Syst. Sci.*, *9*, 243–259.

Zhang, X., and G. Lindström (1997), Development of an automatic calibration scheme for the HBV hydrological model, *Hydrol. Processes*, *11*, 1671–1682.

_____

F. Fenicia, P. Matgen, and L. Pfister, Public Research Center–Gabriel Lippmann, 41 rue du Brill, L-4422,Belvaux, Luxembourg. (fenicia@lippmann.lu)

H. H. G. Savenije, Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, NL-2600 GA Delft, Netherlands.