

This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

# Moving beyond traditional model calibration or how to better identify realistic model parameters: sub-period calibration

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

<sup>1</sup>Delft University of Technology, Faculty of Civil Engineering and Geosciences, Water Resources Section, Delft, The Netherlands

<sup>2</sup>Public Research Center–Gabriel Lippmann, Belvaux, Luxembourg

Received: 2 February 2012 – Accepted: 3 February 2012 – Published: 13 February 2012

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

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

**HESSD**

9, 1885–1918, 2012

**Sub-period  
calibration**

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Abstract

Conceptual hydrological models often rely on calibration for the identification of their parameters. As these models are typically designed to reflect real catchment processes, a key objective of an appropriate calibration strategy is the determination of parameter sets that reflect a “realistic” model behavior. Previous studies have shown that parameter estimates for different calibration periods can be significantly different. This questions model transposability in time, which is one of the key conditions for the set-up of a “realistic” model. This paper presents a new approach that selects parameter sets that provide a consistent model performance in time. The approach consists of confronting model performance in different periods, and selecting parameter sets that are as close as possible to the optimum of each individual sub-period. While aiding model calibration, the approach is also useful as a diagnostic tool, illustrating tradeoffs in the identification of time consistent parameter sets. The approach is demonstrated in a case study where we illustrate the multi-objective calibration of the HyMod hydrological model to a Luxembourgish catchment.

## 1 Introduction

Conceptual hydrological models represent an abstraction of real world processes, and are typically constituted of a number of interconnected reservoirs which are supposed to represent the main catchment compartments and dominant processes (Wagener et al., 2003). It is typically the case that several model parameters are not measurable, even when they are supposed to represent physical catchment characteristics, and therefore they have to be determined by calibration (Wheater et al., 1993). Different approaches to infer parameter values and their likelihood distribution have been developed, for example single or multi-objective calibration (Gupta et al., 1998), generalized likelihood uncertainty estimation (GLUE, Beven and Binley, 1992), dynamic identifiability analysis

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



(DYNIA, Wagener et al., 2003) and Bayesian inference (Wood and Rodríguez-Iturbe, 1975).

A key objective for hydrological modeling is the development of “realistic” models, that is, models which are able to reflect real catchment processes (Wagener, 2003).

The set-up of a realistic model requires the determination of a realistic structure and a suitable parameter set. While the determination of a suitable structure is a theoretical development in its own right (e.g. Clark et al., 2008; Fenicia et al., 2011), we here focus on the determination of a realistic parameter set, and in particular, on parameter sets that reflect a consistent model behavior in time.

Model transposability in time is in fact recognized as one of the main requirement to a successful “validation” of model performance (Klemeš, 1986). Hartmann and Bárdossy (2005) advocate that “if a model is to be used under non-stationary conditions, its parameters and process descriptions should be transferable.”

The calibration-validation framework (or the split-sample test proposed by Klemeš, 1986) has become standard in hydrological practice (Andréassian et al., 2009). A model is calibrated for a period of time and the parameter sets which are selected as behavioral in calibration period are evaluated for a different validation period. Different combination of calibration and validation were suggested by Mroczkowski et al. (1997) however the proposed combinations were constrained in calibration-validation framework for different time periods. This fact also is repeated in comprehensive model developing scheme by Refsgaard et al. (2005).

Seibert (2003) mentioned that the attempt to identify the best parameter sets (or model structure) is constrained in split sample test for time period with mostly similar characteristics. He argued that the reason of the scarce literature on models which perform well in time periods with completely different hydrological characteristics is due to the fact that they most probably fail this test (or the differential split sample test proposed by Klemeš, 1986). Kirchner (2006) criticized commonly used model evaluation following Seibert (2003), he argued “Such models are often good mathematical marionettes; they often can dance to the tune of the calibration data. However, their

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



predictive validity is often in doubt”. This shortcoming was repeatedly addressed in the literature (Anderson and Woessner, 1992; Hassan, 2004; Gupta et al., 2008; Refsgaard and Hansen, 2010). Different methods and strategies were suggested to overcome this shortcoming (Bárdossy and Singh, 2008; Schaeffli et al., 2011; Nalbantis et al., 2011).

5 In addition to model performance, it is also important to see how model parameters are affected by the calibration period. In this respect, previous research has shown that optimal parameter sets for different periods can change substantially.

Wagener et al. (2003) (DYNIA) have developed a method to screen across the time series of model prediction in order to investigate the identifiability of model parameters. 10 They show that uncertainties associated to model parameters can vary substantially in different time periods.

Previously, Freer et al. (2003) assessed Dynamic TOPMODEL using GLUE based on different objective functions and rising or falling limbs of the hydrograph. They showed that it may be difficult to propose a consistently parameterized model structure due to 15 the significant variability of the observed responses. They concluded that the model fails to meet even relaxed acceptable thresholds. Hartmann and Bárdossy (2005) investigated parameter transferability in different climatic conditions (“warm”, “cold”, “wet” and “dry”) and for different time scales (days up to years). They designed a calibration method that allows a good performance on different time scales simultaneously. 20 Li et al. (2011) investigate the transferability of model parameters for dry and wet conditions. They showed dry period contain more information for model calibration than the wet one. Bárdossy and Singh (2008) using depth function (Tukey, 1975) concluded “that equally performing parameters are not necessarily equally transferable or equally sensitive”.

25 Boyle et al. (2000, 2001) used the multi-objective calibration framework proposed by Gupta et al. (1998) to calibrate a model for different flow segments of hydrograph. The multi-objective framework makes it possible to identify optimal parameter sets for a set of objective function. This approach was extensively used in several applications (see Efstratiadis and Koutsoyiannis, 2010, for a review). Incorporating multi-criteria, as

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



an example, tracer data or remotely sensed evaporation into model calibration helps identification of more realistic model structure and parameter sets (Dunn and Colohan, 1999; Seibert and McDonnell, 2002; Weiler et al., 2003; Freer et al., 2004; Uhlenbrook and Sieber, 2005; Vaché and McDonnell, 2006; Son and Sivapalan, 2007; Winsemius et al., 2008; Dunn et al., 2008; Birkel et al., 2010).

Both, the multi-objective and the multi-criteria calibration strategies, constrain the feasible parameter space and facilitate parameter selection on basis of performance trade-offs, i.e. Pareto fronts. However, as argued by Beven (2006), the mere mappings of optimum parameter sets after calibration are “too simplistic, since they arbitrarily exclude many models that are very nearly as good as the ‘optima’”. This simply means the parameter realization should include “sub-optimal” parameter sets as well.

This paper introduces a new framework for parameter identification including optimal and sub-optimal parameter sets which are more time consistent. The method is based on the calibration on different periods, and determines the parameter sets which perform best for all sub-periods. As the selected parameter sets are evaluated in different periods, only the time consistent parameter sets are selected. The new method is applied on a case study and compared with a calibration-validation framework with respect to parameter identifiability and performance for the Wark catchment located in the Grand Duchy of Luxembourg, using the lumped conceptual model HyMod.

## 2 Sub-period calibration framework

The sub-period calibration framework involves two crucial steps in extracting the most realistic parameterizations for a given model structure. Firstly the available input and output data sets are split into (ideally equal-length)  $k$  sub-periods. These sub-periods and their lengths can be arbitrarily chosen (e.g. month, season, etc). It can, however, be convenient to base the analysis on full years. Alternatively, the full observation period could, for instance, be split up according to wetness conditions (e.g. Hartmann and Bárdossy, 2005). Each sub-period is then calibrated individually in a  $n$ -dimensional

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



multi-objective calibration framework (it can also be a single objective), which result in a  $n$ -dimensional Pareto front for each sub-period. Therefore  $k$   $n$ -dimensional calibration Pareto fronts (CPF) are obtained separately. Subsequently, for each parameter set its distance to the  $k$  Pareto fronts are calculated. Therefore  $k$  distances are obtained for the  $k$  sub-periods. The goal is to find parameter sets that minimize the distances to all Pareto fronts. In order to achieve this, the  $k$ -dimensional Pareto front of distances is determined. The sub-period calibration concept is illustrated in Fig. 1. The Parameters are acceptable which have the most consistent performance regarding the optimum performance of each sub-period.

The concept is further illustrated with an abstract 2-objective, 2-sub-period example in Fig. 2. The conventional CPFs for the two sub-periods, are shown in Fig. 2a. Symbol 1 (circle) represents a parameter set that is a Pareto-member of the first sub-period; however, it does not perform well compared to the best possible outcome, i.e. CPF<sub>2</sub>, when applied in the second sub-period. The parameter set represented by symbol 2 (star), on the other hand, although not a member of the CPF<sub>1</sub> in the first sub-period, performs rather well in the second sub-period as can be seen by the short distance to CPF<sub>2</sub>. In other words, parameter sets which are slightly sub-optimal in one sub-period may perform significantly better than “optimal” parameter sets, i.e. CPF members, in other sub-periods.

Figure 2b and c illustrates the set-up of the sub-period calibration framework. For each parameter set in each sub-period the distance to the two CPFs was calculated. The distances to the Pareto fronts of each parameter set represent a bi-dimensional space (Fig. 2c). Typically, model parameters on the CPF<sub>1</sub> will not be part of the CPF<sub>2</sub>. Hence, there will be a tradeoff when the objective is to minimize the distance between both Pareto fronts. Hereafter this tradeoff will be referred as the Minimum Distance Pareto Front (MDPF, Fig. 2c). At each edge of the MDPF are the points on the two CPFs of Fig. 2a and b. In between, there are points that have an overall good performance in both sub-periods. We consider all points on the MDPF as “acceptable” parameter sets.

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The sub-period calibration framework can be expressed in formal notation as follows:

$$Y(\theta, \xi) = \gamma(\theta|\xi) \quad (1)$$

where  $Y$ ,  $\gamma$ ,  $\xi$  and  $\theta$  are the model output, the hydrological model, forcing and parameter set, respectively. The objective function ( $O$ ) can be described as an error function ( $E$ ) which returns the difference between the observed and model values:

$$O_j(\theta, \xi_j) = E_j(Y(\theta, \xi_j), Y_{oj}) = \{o_{1j}, o_{2j}, \dots, o_{nj}\}, \quad j = 1, \dots, k \quad (2)$$

where  $n$  is the number of objective functions which is used for evaluation of the model performance and have to be optimized and  $k$  is the number of sub-periods and  $Y_{oj}$  indicates the observed time series for  $j$ -th subperiod. The calibration Pareto front (CPF) for each of the  $k$  sub-periods can be described by optimizing the objective function related to the same sub-period:

$$CPF_j = \min(O_j), \quad j = 1, \dots, k. \quad (3)$$

This results in  $k$  CPFs, each of which has the dimension  $n$  of the original objective space. The optimization space is then transformed into another multi-objective space with  $k$  (number of sub-periods) objective functions in which the difference in model performance for each sub-period with its related Pareto front is evaluated:

$$L(\theta) = G(CPF_j, O_j(\theta, \xi_j)), = \{l_1, l_2, \dots, l_k\} \quad (4)$$

where  $G(\cdot)$  quantifies the error between model performance for  $j$ -th sub-period and calibration Pareto front ( $CPF_j$ ) for the same sub-period. The final solution can be obtained by minimizing  $L$ . The method will in the following be referred to as SuPer (sub-period) calibration.

### 3 Case study

#### 3.1 Study area and data

The outlined methodology will in the following be illustrated with a case study using data from the Wark catchment in the Grand Duchy of Luxembourg. The catchment has an area of 82 km<sup>2</sup> with the catchment outlet located downstream of the town of Ettelbrück at the confluence with the Alzette River (49.85° N, 6.10° E). With an average annual precipitation of 850 mm yr<sup>-1</sup> and an average annual potential evaporation of 650 mm yr<sup>-1</sup> the annual runoff is approximately 250 mm yr<sup>-1</sup>. The geology in the northern part is dominated by schist while the southern part of the catchment is mostly underlain by sandstone and conglomerate. The dominant land uses are forest on hillslopes, agricultural land on plateaus and pastures in the valley bottoms. The elevation varies between 195 to 532 m with an average of 380 m a.s.l. The slope of the catchment varies between 0–200%, with an average value of 17% (Gharari et al., 2011). The hydrological data including discharge at the outlet of the Wark catchment, evaporation estimated by the Hamon equation (Hamon, 1961) with data measured at Findel (Luxembourg airport; Fenicia et al., 2008) and rainfall via three rain gauges with a 12-h resolution for the 2001–2004 period were used. While 2001 was used as model training period, the years 2002–2004 exhibited rather distinct meteorological conditions as summarized in Table 1, with 2003 clearly being the driest and 2002 the wettest year.

#### 3.2 Hydrological model

The rainfall-runoff model applied in the Wark catchment to illustrate the effects of the sub-period calibration framework was the lumped conceptual HyMod (Wagener et al., 2001). HyMod was chosen for its limited number of parameters while still maintaining adequate process representation including slow and fast responses together

## HESSD

9, 1885–1918, 2012

### Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





with a non-linear soil moisture component. To simulate runoff a forward explicit Euler method was used.

HyMod is characterized by five states, including the soil moisture reservoir ( $S_M$  (mm)), three linear reservoirs in series ( $S_{F_1}$  (mm),  $S_{F_2}$  (mm),  $S_{F_3}$  (mm)), mimicking the fast runoff component and one slow reservoir ( $S_{S_1}$  (mm)). The five parameters represent the maximum soil moisture storage ( $S_{M,max}$  (mm)), the spatial variability of soil moisture ( $\beta$  (-)), the partitioning between fast reservoirs and slow reservoir ( $\alpha$  (-)), as well as the timescales of the fast and slow reservoirs ( $R_Q$  ((12 h)<sup>-1</sup>),  $R_S$  ((12 h)<sup>-1</sup>)).

$P$  (mm (12 h)<sup>-1</sup>),  $E$  (mm (12 h)<sup>-1</sup>),  $E_p$  (mm (12 h)<sup>-1</sup>) and  $Q_m$  (mm (12 h)<sup>-1</sup>) represent precipitation, actual evaporation, potential evaporation and modeled runoff, respectively. The simulated runoff by the model ( $Q_m$ ) is the summation of slow and fast components ( $Q_m = Q_{S_1} + Q_{F_3}$ ). The water balance equations and constitutive relations are listed in Table 2 and HyMod schematic illustration is depicted in Fig. 3.

### 3.3 Implementation of sub-period calibration

Using 2001 as model warm-up period, the remaining 2002–2004 observation period was decomposed into three 1-yr sub-periods (2002, 2003 and 2004). The three sub-periods were calibrated individually to obtain the independent calibration Pareto fronts for each sub-period (CPF<sub>2002</sub>, CPF<sub>2003</sub> and CPF<sub>2004</sub>) as well as for the entire 2002–2003 (CPF<sub>2002–2003</sub>) and 2002–2004 periods (CPF<sub>2002–2004</sub>). Based on these premises, two example implementations of SuPer calibration are given below. In the first implementation, the parameter sets minimizing the Euclidean distance of performance to the CPFs and generating the minimum distance Pareto front (MDPF, Fig. 2c) were identified based on CPF<sub>2002</sub> and CPF<sub>2003</sub> only, making it a two-dimensional (i.e. two sub-periods) multi-objective problem. In this implementation CPF<sub>2004</sub> was explicitly not considered for constructing the MDPF for the purpose of *independently* demonstrating the effect of SuPer calibration. The year 2004 is here rather used as a validation or target year to compare the results of SuPer calibration with traditional calibration strategies. In operational applications of SuPer calibration CPF<sub>2004</sub> would

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



thus not be excluded in order to ensure efficient exploitation of the information content in the available data. A full operational application of SuPer calibration, including CPF<sub>2004</sub> in a three-dimensional (i.e. three sub-periods) multi-objective practice is thus shown in the second implementation.

In this case study, HyMod was evaluated for high and low flows in a multi-objective optimization approach. The respective objective functions used are the Root Mean Square Error of the flows (RMSE) and the Root Mean Square error of the logarithm of flows (LRMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Q_m - Q_o)^2} \quad (5)$$

$$LRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\log(Q_m) - \log(Q_o))^2} \quad (6)$$

where  $Q_m$  is the modeled flow,  $Q_o$  is observed flow, respectively and  $N$  is the number of time steps. RMSE was used rather than Nash Sutcliffe efficiency as RMSE does not need a base, which may be different in different year (or sub periods), for evaluating the performance (Schaefli and Gupta, 2007).

Here, the calibration to find the best parameter sets and the related CPF was based on the MOSCEM-UA algorithm (Vrugt et al., 2003). This was chosen as SuPer calibration identifies parameter sets with the best performance relative to CPFs, and as MOSCEM-UA uses Zitzler strength Pareto ranking (Zitzler and Thiele, 1999), which gives a better and more uniform estimation of CPF. Note, however, that the choice of sub-periods, calibration objectives and criteria as well as of the calibration algorithm used for SuPer calibration can in principle be arbitrarily adapted to available data and modeling requirements.

**Sub-period calibration**

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## 4 Result

### 4.1 Implementation 1: calibration based on years 2002 and 2003

As a first step HyMod was calibrated individually for the chosen sub-periods 2002, 2003, 2004 as well as for the entire 2002–2003 period. The resulting  $CPF_{2002}$ ,  $CPF_{2003}$ ,  $CPF_{2004}$  and  $CPF_{2002-2003}$  are shown as lines in Fig. 4. The dots with the same colors indicate the performance of the CPF member parameter sets in sub-periods the respective CPF has not been calibrated for, e.g. the performance of  $CPF_{2002}$  members in 2004, which is effectively a model validation in traditional terms. The best available calibrated performance of HyMod for validation period or target year 2004, i.e.  $CPF_{2004}$ , is represented by the black CPF in Fig. 4. As it is visible in Fig. 4, the optimal performance in 2003, as represented by  $CPF_{2003}$ , is better than the performance in 2002, as represented by  $CPF_{2002}$ . Yet, when using the  $CPF_{2002}$  members to run the model in the validation or target period 2004 they perform better, i.e. they plot closer to  $CPF_{2004}$ , than when using  $CPF_{2003}$  members. This clearly indicates that a better performance for one specific time period does not necessarily imply a better performance for other periods as well. On the other hand  $CPF_{2002}$  shows skewed performance where different parameter sets can equally well fit high flows (RMSE) while they result in varying performance for low flows (LRMSE).

The problem of skewed performance of the model also remains obvious for  $CPF_{2002-2003}$ . The best performing realistic parameter sets as identified by SuPer calibration are shown by light blue in Fig. 4. These parameter sets were identified based on the Euclidean distances of their performance to  $CPF_{2002}$  and  $CPF_{2003}$ , resulting in the minimum distance Pareto front (MDPF, Fig. 5). According to the trade-off in the MDPF they thus perform as closely as possible to both calibration Pareto fronts,  $CPF_{2002}$  and  $CPF_{2003}$ . The light blue dots in Fig. 4 represent the model performance, when it is run for target year 2004 with the parameter sets identified by SuPer calibration with the MDPF based on  $CPF_{2002}$  and  $CPF_{2003}$ . Further, performance of the parameter sets obtained from the MDPF by SuPer calibration in 2002 and 2003 are illustrated by light

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



blue crosses and stars, respectively. They exhibit significant skewed behavior towards similar part of the CPFs for the two sub-periods. For both sub-periods SuPer calibration chooses the parameter sets which perform close to the low flow (LRMSE) end of CPF<sub>2002</sub> and CPF<sub>2003</sub> which shows the chosen model structure can simultaneously identify low flow better than high flow in both sub-periods.

## 4.2 Implementation 2: calibration based on years of 2002, 2003 and 2004

In this section the entire time series, 2002–2004, was used to construct the MDPF for SuPer calibration in order to show the effects of SuPer calibration under operational conditions. The target sub-period 2004 was thus part of the calibration period. As three sub-periods (2002, 2003 and 2004) were used for constructing the MDPF, which is the basis of SuPer calibration, the calibration space was transformed into a three-dimensional multi-objective practice defined by Euclidean distances  $D_1$  to CPF<sub>2002</sub>,  $D_2$  to CPF<sub>2003</sub> and distance  $D_3$  to CPF<sub>2004</sub> (Fig. 6). The performances of CPF<sub>2002</sub>, CPF<sub>2003</sub> and CPF<sub>2004</sub> parameter sets in target year 2004 are illustrated in Fig. 7a for different methods. The dark blue, yellow, red, purple dots illustrate the performance of CPF<sub>2002</sub>, CPF<sub>2003</sub>, CPF<sub>2002–2003</sub> and CPF<sub>2002–2004</sub> members for target year 2004, respectively (validation of CPFs' members in 2004). The green dots represent the performance of SuPer calibration parameter sets, based on MDPF<sub>2002–2004</sub> members in target year 2004. Furthermore the performance of MDPF<sub>2002–2004</sub> members for sub-period 2002 and 2003 are presented by green stars and crosses in Fig. 7b, respectively.

Comparing the performance Parameter sets identified by SuPer calibration based on MDPF<sub>2002–2004</sub> in the three sub-periods reveals the goodness of model regarding evaluation objective functions (RMSE and LRMSE) in each sub-period. For MDPF<sub>2002–2004</sub>, SuPer calibration picks the parameters which focuses on low flow for sub-period 2002, and high flow for sub-period 2004 while covering the entire Pareto front of 2003 (Fig. 7b).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



### 4.3 Parameters identifiability

Parameter behavior of HyMod's fast reservoir coefficient and slow reservoir coefficient was evaluated. The reason for this selection is that the slow reservoir coefficient has overlap values in the three sub-periods and fast reservoir coefficient does not have feasible overlap values for optimum parameter value.

The optimum parameter behavior is depicted for slow reservoir coefficient ( $R_S$ ) of HyMod in Fig. 8. As it is clear from Fig. 8b–e and Fig. 8i–l the parameter ranges for  $CPF_{2002}$ ,  $CPF_{2002-2003}$  and  $CPF_{2002-2004}$  are between 0.01 and 0.07 ( $(12h)^{-1}$ ) for both objective functions (RMSE and LRMSE). The only year with a well identified slow reservoir coefficient  $R_S$  is 2003. Figure 8f,g and Fig. 8m,n show the parameter range of  $R_S$  as obtained by SuPer calibration, based on  $MDPF_{2002-2003}$  as well as on  $MDPF_{2002-2004}$ , respectively; SuPer calibration reduces the parameter range to values between to 0.01–0.03 ( $(12h)^{-1}$ ). This is further illustrated by comparing the distributions of CPF and SuPer calibration MDPF member parameter sets as shown in Fig. 8a, based on the normalized cumulative frequency of the respective Pareto member parameters.

The steeper cumulative frequency distribution for SuPer calibration parameter sets as shown in Fig. 8a indicates that parameter identifiability of SuPer calibration based on  $MDPF_{2002-2003}$  is higher than when calibrating the model to the 2002–2003 period following traditional calibration strategies, i.e.  $CPF_{2002-2003}$ . Similarly, SuPer calibration identifies behavioral parameter sets sharper than the traditional strategies for the 2002–2004 period, as well. Figure 8a also reveals that identifiability of  $R_S$  based on SuPer calibration is as good as that of the best calibrated performance of the hydrological model for target year 2004. This can be seen by comparing the  $CPF_{2004}$ , the black line, and the SuPer calibration results based on  $MDPF_{2002-2004}$  which is the green line. It also becomes evident in Fig. 8a that the actual parameter range as well as their distributions obtained from SuPer calibration based on  $MDPF_{2002-2004}$  is more consistent

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



with  $CPF_{2004}$ , the best available parameter set of target period 2004, than parameter sets obtained from  $CPF_{2002}$ ,  $CPF_{2003}$ ,  $CPF_{2002-2003}$  and  $CPF_{2002-2004}$ .

The optimum parameter behavior is depicted for fast reservoir coefficient ( $R_Q$ ) of HyMod in Fig. 9. Due to the contrasting characteristics of the sub-periods 2002 and 2003 the feasible parameter ranges of  $CPF_{2002}$  and  $CPF_{2003}$  members vary considerably. However, when the model is calibrated based on 2002–2003, i.e.  $CPF_{2002-2003}$ , the obtained parameter range is between the parameter sets obtained from  $CPF_{2002}$  and  $CPF_{2003}$  (Fig. 9b–d and Fig. 9i–k). This is also the case when the entire time series 2002–2004 is used for calibration and parameters are obtained according to  $CPF_{2002-2004}$  (Fig. 9e,l). However, SuPer calibration can detect the inconsistencies between the best parameter ranges of the individual sub-periods and thus widens the feasible ranges for these parameters (Fig. 9f,g,m,n).

## 5 Discussion

The fact that SuPer calibration focuses on different parts of sub-period calibration Pareto fronts, CPFs, helps to indicate how Pareto members should be retained as “realistic” (Figs. 4 and 7b). Pareto fronts of a calibrated model (CPF) may show a skewed behavior with respect to one or more objective functions ( $CPF_{2002}$  and  $CPF_{2002-2003}$  in Fig. 4). For traditional calibration strategies this introduces the requirement for a subjective decision on the parameter acceptance threshold (Fig. 10) as highlighted by Efstratiadis and Koutsoyiannis (2010). Khu and Madsen (2005) suggested a methodology to choose appropriate CPF members based on investigating the performance of CDF in its different sub-dimensional spaces. Birkel et al. (2010) selected “realistic” parameter sets by confronting the “best fit” parameter sets with tracer data. In contrast with mentioned methodologies, SuPer calibration does not require a subjective threshold for identifying parameter sets as this threshold is implicitly given by the MDPF. This threshold is not subjective rather it is the best compromise between CPFs of sub-periods that can be achieved by a given model structure. Furthermore SuPer

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



calibration doesn't need additional data (although additional data can be incorporated with SuPer calibration); it uses no more information than the data which is needed to calibrate a rainfall-runoff model traditionally.

Behavior of optimal parameter sets by SuPer calibration can be used as a criterion for parameter time consistency in different sub-periods. With time consistent parameters it is expected that the parameter ranges obtained by SuPer calibration are lower or equal to those obtained from long-term calibration, e.g.  $CPF_{2002-2003}$  or  $CPF_{2002-2004}$ . By identifying non-time consistent parameters, SuPer calibration can be used as a diagnostic tool for identifying model structure deficiencies (cf. Clark et al., 2008). This design also allows the reduction of both, type I and type II errors on model selection (false positives and false negatives, Beven, 2010). Furthermore, SuPer calibration can provide information about the behavior of each parameter with respect to the hydrological condition of that period. As an example, the fast reservoir coefficient  $R_Q$  shows higher values for the sub-period 2003 than for 2002; 2003 is hydrologically distinct to the other two years 2002 and 2004 (Fig. 9). Analyses like this, similar to the DYNIA (Wagener et al., 2003) can help the modeler to evaluate which and how a parameter or a function in the model structure should be changed or amended.

The proposed SuPer calibration framework is thus a method that allows identifying realistic model parameterizations based on the premises that acceptable parameterizations have to perform consistently well when predicting the response variable in independent model validation, which is implicitly enforced in SuPer calibration. To some extent it also has the potential to reduce epistemic error in models, i.e. the error due to disinformation (Beven and Westerberg, 2011) or inaccurate input data (Kavetski et al., 2002, 2006). As a thought experiment, consider a catchment with an adequate long term average representation of precipitation. In the case of a significant storm event with small spatial extent, which is not picked up by the gauges, a peak in runoff will be observed. A model will, through traditional calibration, be *forced* to mimic this peak even if there was no observed precipitation. This implies that the model will have to reproduce the "correct" output with "incorrect" input, hence the best fit parameter set

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





will be one that does exactly that: reproduce the “real” output with the “incorrect” input. As a consequence, the chosen parameters will misrepresent reality and result in low predictive power of the model. As it is unlikely that identical storm configuration and timing will occur in any of the other sub-periods, SuPer calibration will most likely discard this parameterization if it performs far from the calibration of the other sub-periods (cf. Fig. 8). Furthermore SuPer calibration can be used for storm events with different magnitude and return period separately to retain their characteristic during calibration process, as an example, sub-periods can be defined as different part of flow duration curve (Westerberg et al., 2011).

Although SuPer calibration framework can in principle be implemented with different calibration methods, its dependency on Pareto fronts requires calibration methods which represent the Pareto front position in the objective space adequately well. The uncertainty in Pareto front identification may introduce uncertainty in the final selected parameter set chosen by SuPer calibration. In this study MOSCEM-UA (Vrugt et al., 2003) was used to generate Pareto fronts in both steps of the procedure (creating CPFs and MDPFs). However, future research should investigate the effectiveness of MOSCEM-UA for the generation of MDPF in the second step of SuPer calibration, as the distance to Pareto fronts (e.g. line or surface) needs to be minimized instead of the vector toward a point (origin of objective space), which MOSCEM-UA was originally designed for. To ensure that using MOSCEM-UA in the second step of SuPer calibration performs well in parameter identification, SuPer calibration was implemented in both steps with Monte-Carlo sampling using the same parameter range for a million random samples. The result were consistent with the result obtained by MOSCEM-UA; however this may be case specific and not valid for other case studies or models with higher complexity therefore investigation the performance of optimization algorithm specially in second step of SuPer calibration is highly recommended.

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





## 6 Conclusions

In this paper a calibration framework, based on splitting the available data sets into sub-periods was proposed. The SuPer calibration framework is based on the extension of traditional split sample tests which can also be seen as an additional layer of model testing, independent from modeling objectives and criteria as well as calibration algorithms. By extracting more information from the available data and by avoiding the “loss” of data otherwise used for validation, it allows the identification of more realistic model parameterizations. Although this comes at the cost of potentially reduced performance during calibration, model parameterizations as obtained by SuPer calibration give consistently better prediction performances, which is what modelers actually should look for. The design of SuPer calibration is such that acceptable parameterizations have to perform consistently well when predicting any of the defined sub-periods, which is implicitly enforced in SuPer calibration, thus avoiding the need for explicit model validation. Furthermore, by the transformation of the traditional objective-space into a minimum Euclidean distance space the need for subjective choices of parameter acceptance thresholds is avoided.

It should be again emphasized here that SuPer calibration is not a calibration algorithm, nor is it explicitly addressing parameter uncertainty. It is rather a more advanced method of model testing, building on traditional split sample tests and making more efficient use of available data. SuPer calibration can in principle be done with any number and type of objective functions (e.g. NSE or RMSE) but also with any number and type of calibration criteria (e.g. only using runoff or using runoff and tracer dynamics). A Matlab function of the SuPer calibration framework based on Monte-Carlo calibration strategy for the same case study presented in this paper is available at <http://supercalibration.weblog.tudelft.nl/> or can be obtained by personal communication with the lead author.

**HESSD**

9, 1885–1918, 2012

### Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



*Acknowledgements.* The authors would like to thank the people who facilitated this research at Gabriel Lippmann Research Institute. Shervan Gharari is funded for his PhD by Fonds National de la Recherche (AFR) of Luxembourg with project reference number of 1383201.

## References

- 5 Anderson, M. P. and Woessner, W. W.: The role of the postaudit in model validation, *Adv. Water Res.*, 15, 167–173, 1992. 1888
- Andréassian, V., Perrin, C., Berthet, L., Le Moine, N., Lerat, J., Loumagne, C., Oudin, L., Mathévet, T., Ramos, M.-H., and Valéry, A.: HESS Opinions “Crash tests for a standardized evaluation of hydrological models”, *Hydrol. Earth Syst. Sci.*, 13, 1757–1764, doi:10.5194/hess-13-1757-2009, 2009. 1887
- 10 Bárdossy, A. and Singh, S. K.: Robust estimation of hydrological model parameters, *Hydrol. Earth Syst. Sci.*, 12, 1273–1283, doi:10.5194/hess-12-1273-2008, 2008. 1888
- Beven, K.: A manifesto for the equifinality thesis, *J. Hydrol.*, 320, 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006. 1889
- 15 Beven, K. and Westerberg, I.: On red herrings and real herrings: disinformation and information in hydrological inference, *Hydrol. Process.*, 25, 1676–1680, doi:10.1002/hyp.7963, 2011. 1899
- Beven, K. J.: Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models, *Hydrol. Process.*, 24, 1537–1547, doi:10.1002/hyp.7718, 2010. 1899
- 20 Beven, K. J. and Binley, A. M.: The future of distributed models: model calibration and uncertainty prediction, *Hydrol. Process.*, 6, 279–298, 1992. 1886
- Birkel, C., Dunn, S. M., Tetzlaff, D., and Soulsby, C.: Assessing the value of high-resolution isotope tracer data in the stepwise development of a lumped conceptual rainfall-runoff model, *Hydrol. Process.*, 24, 2335–2348, doi:10.1002/hyp.7763, 2010. 1889, 1898
- 25 Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resour. Res.*, 36, 3663–3674, doi:10.1029/2000WR900207, 2000. 1888

---

### Sub-period calibration

S. Gharari et al.

---

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



---

**Sub-period  
calibration**

S. Gharari et al.

---

[Title Page](#)
[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

- Boyle, D. P., Gupta, H. V., Sorooshian, S., Koren, V., Zhang, Z., and Smith, M.: Toward improved streamflow forecasts: value of semidistributed modeling, *Water Resour. Res.*, 37, 2749–2759, doi:10.1029/2000WR000207, 2001. 1888
- Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, 44, W00B02, doi:10.1029/2007WR006735, 2008. 1887
- Dunn, S. M. and Colohan, R. J. E.: Developing the snow component of a distributed hydrological model: a step-wise approach based on multi-objective analysis, *J. Hydrol.*, 223, 1–16, doi:10.1016/S0022-1694(99)00095-5, 1999. 1889
- Dunn, S. M., Bacon, J. R., Soulsby, C., Tetzlaff, D., Stutter, M. I., Waldron, S., and Malcolm, I. A.: Interpretation of homogeneity in  $18^{\circ}$  signatures of stream water in a nested sub-catchment system in north-east Scotland, *Hydrol. Process.*, 22, 4767–4782, doi:10.1002/hyp.7088, 2008. 1889
- Efstratiadis, A. and Koutsoyiannis, D.: One decade of multi-objective calibration approaches in hydrological modelling: a review, *Hydrolog. Sci. J.*, 55, 58–78, doi:10.1080/02626660903526292, 2010. 1888, 1898, 1918
- Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Understanding catchment behavior through stepwise model concept improvement, *Water Resour. Res.*, 44, W01402, doi:10.1029/2006WR005563, 2008. 1892
- Fenicia, F., Kavetski, D., and Savenije, H. H. G.: Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development, *Water Resour. Res.*, 47, W11510, doi:10.1029/2010WR010174, 2011. 1887
- Freer, J., Beven, K., and Peters, N.: Multivariate seasonal period model rejection within the generalised likelihood uncertainty estimation procedure, *Water Sci. Appl.*, 6, 69–87, doi:10.1029/WS006p0069, 2003. 1888
- Freer, J., McMillan, H., McDonnell, J., and Beven, K.: Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures, *J. Hydrol.*, 291, 254–277, doi:10.1016/j.jhydrol.2003.12.037, 2004. 1889
- Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H. G.: Hydrological landscape classification: investigating the performance of HAND based landscape classifications in a central European meso-scale catchment, *Hydrol. Earth Syst. Sci.*, 15, 3275–3291, doi:10.5194/hess-15-3275-2011, 2011. 1892

**Sub-period  
calibration**

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, 34, 751–763, doi:10.1029/97WR03495, 1998. 1886, 1888
- Gupta, H. V., Wagener, T., and Liu, Y.: Reconciling theory with observations: elements of a diagnostic approach to model evaluation, *Hydrol. Process.*, 22, 3802–3813, doi:10.1002/hyp.6989, 2008. 1888
- Hamon, W. R.: Estimating potential evapotranspiration, *J. Hydrol. Div.*, 87, 107–120, 1961. 1892
- Hartmann, G. and Bárdossy, A.: Investigation of the transferability of hydrological models and a method to improve model calibration, *Adv. Geosci.*, 5, 83–87, doi:10.5194/adgeo-5-83-2005, 2005. 1887, 1888, 1889
- Hassan, A. E.: Validation of numerical ground water models used to guide decision making, *Ground Water*, 42, 277–290, doi:10.1111/j.1745-6584.2004.tb02674.x, 2004. 1888
- Kavetski, D., Franks, S. W., and Kuczera, G.: Confronting input uncertainty in environmental modelling, American Geophysical Union, Washington, DC, 2002. 1899
- Kavetski, D., Kuczera, G., and Franks, S. W.: Bayesian analysis of input uncertainty in hydrological modeling: 1. theory, *Water Resour. Res.*, 42, W03407, doi:10.1029/2005WR004368, 2006. 1899
- Khu, S. T. and Madsen, H.: Multiobjective calibration with Pareto preference ordering: an application to rainfall-runoff model calibration, *Water Resour. Res.*, 41, W03004, doi:10.1029/2004WR003041, 2005. 1898
- Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resour. Res.*, 42, W03S04, doi:10.1029/2005WR004362, 2006. 1887
- Klemeš, V.: Operational testing of hydrological simulation models, *Hydrolog. Sci. J.*, 31, 13–24, doi:10.1080/02626668609491024, 1986. 1887
- Li, C. Z., Zhang, L., Wang, H., Zhang, Y. Q., Yu, F. L., and Yan, D. H.: The transferability of hydrological models under nonstationary climatic conditions, *Hydrol. Earth Syst. Sci. Discuss.*, 8, 8701–8736, doi:10.5194/hessd-8-8701-2011, 2011. 1888
- Mroczkowski, M., Raper, P. G., and Kuczera, G.: The quest for more powerful validation of conceptual catchment models, *Water Resour. Res.*, 33, 2325–2335, doi:10.1029/97WR01922, 1997. 1887

**Sub-period  
calibration**

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Nalbantis, I., Efstratiadis, A., Rozos, E., Kopsiafti, M., and Koutsoyiannis, D.: Holistic versus monomeric strategies for hydrological modelling of human-modified hydrosystems, *Hydrol. Earth Syst. Sci.*, 15, 743–758, doi:10.5194/hess-15-743-2011, 2011. 1888
- 5 Refsgaard, J. C. and Hansen, J. R.: A good-looking catchment can turn into a modeller's nightmare, *Hydrolog. Sci. J.*, 55, 899–912, doi:10.1080/02626667.2010.505571, 2010. 1888
- Refsgaard, J. C., Henriksen, H. J., Harrar, W. G., Scholten, H., and Kassahun, A.: Quality assurance in model based water management – review of existing practice and outline of new approaches, *Environ. Modell. Softw.*, 20, 1201–1215, doi:10.1016/j.envsoft.2004.07.006, 2005. 1887
- 10 Schaefli, B. and Gupta, H. V.: Do Nash values have value?, *Hydrol. Process.*, 21, 2075–2080, doi:10.1002/hyp.6825, 2007. 1894
- Schaefli, B., Harman, C. J., Sivapalan, M., and Schymanski, S. J.: HESS Opinions: Hydrologic predictions in a changing environment: behavioral modeling, *Hydrol. Earth Syst. Sci.*, 15, 635–646, doi:10.5194/hess-15-635-2011, 2011. 1888
- 15 Seibert, J.: Reliability of model predictions outside calibration conditions, *Nord. Hydrol.*, 34, 477–492, 2003. 1887
- 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 Resour. Res.*, 38, 1241, doi:10.1029/2001WR000978, 2002. 1889
- 20 Son, K. and Sivapalan, M.: Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data, *Water Resour. Res.*, 43, W01415, doi:10.1029/2006WR005032, 2007. 1889
- Tukey, J. W.: Mathematics and the Picturing of Data, in: *Proceedings of the 1975 International 17 Congress of Mathematics*, 2, 523–531, 1975. 1888
- 25 Uhlenbrook, S. and Sieber, A.: On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model, *Environ. Modell. Softw.*, 20, 19–32, doi:10.1016/j.envsoft.2003.12.006, 2005. 1889
- Vaché, K. and McDonnell, J.: A process-based rejectionist framework for evaluating catchment runoff model structure, *Water Resour. Res.*, 42, W02409, doi:10.1029/2005WR004247, 2006. 1889
- 30 Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, 39, 1214, doi:10.1029/2002WR001746, 2003. 1894, 1900

**Sub-period  
calibration**

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Wagener, T.: Evaluation of catchment models, *Hydrol. Process.*, 17, 3375–3378, doi:10.1002/hyp.5158, 2003. 1887

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V., and Sorooshian, S.: A framework for development and application of hydrological models, *Hydrol. Earth Syst. Sci.*, 5, 13–26, doi:10.5194/hess-5-13-2001, 2001. 1892

Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., and Gupta, H. V.: Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis, *Hydrol. Process.*, 17, 455–476, doi:10.1002/hyp.1135, 2003. 1886, 1887, 1888, 1899

Weiler, M., McGlynn, B., McGuire, K., and McDonnell, J.: How does rainfall become runoff? A combined tracer and runoff transfer function approach, *Water Resour. Res.*, 39, 1315, doi:10.1029/2003WR002331, 2003. 1889

Westerberg, I. K., Guerrero, J.-L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., and Xu, C.-Y.: Calibration of hydrological models using flow-duration curves, *Hydrol. Earth Syst. Sci.*, 15, 2205–2227, doi:10.5194/hess-15-2205-2011, 2011. 1900

Wheater, H. S., Jakeman, A. J., and Beven, K. J.: *Progress and Directions in Rainfall-Runoff Modeling*, John Wiley & Sons, 1993. 1886

Winsemius, H. C., Savenije, H. H. G., and Bastiaanssen, W. G. M.: Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins?, *Hydrol. Earth Syst. Sci.*, 12, 1403–1413, doi:10.5194/hess-12-1403-2008, 2008. 1889

Wood, E. F. and Rodríguez-Iturbe, I.: Bayesian inference and decision making for extreme hydrologic events, *Water Resour. Res.*, 11, 533–542, doi:10.1029/WR011i004p00533, 1975. 1887

Zitzler, E. and Thiele, L.: Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach, *IEEE T. Evol. Comput.*, 3, 257–271, doi:10.1109/4235.797969, 1999. 1894

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Table 1.** Rainfall, runoff and potential evaporation for year 2002 to 2004 for the Wark catchment.

Year	Rainfall (mm yr <sup>-1</sup> )	Runoff (mm yr <sup>-1</sup> )	Potential evaporation (mm yr <sup>-1</sup> )
2002	980	410	692
2003	744	226	738
2004	882	249	679

## Sub-period calibration

S. Gharari et al.

**Table 2.** Equations used in HyMod.

Reservoir	Water balance equations	Constitutive relations	
Soil moisture ( $S_M$ )	$dS_M/dt = P - P_e - E_a$	$P_e = FP$ $E_a = WE_p$	$F = 1 - (1 - S_M/S_{M,max})^\beta$ $W = \left[ \frac{S_M}{S_{M,max}} \right]$
First fast reservoir ( $S_{F_1}$ )	$dS_{F_1}/dt = \alpha P_e - Q_{F_1}$	$Q_{F_1} = S_{F_1} R_Q$	
Second fast reservoir ( $S_{F_2}$ )	$dS_{F_2}/dt = Q_{F_1} - Q_{F_2}$	$Q_{F_2} = S_{F_2} R_Q$	
Third fast reservoir ( $S_{F_3}$ )	$dS_{F_3}/dt = Q_{F_2} - Q_{F_3}$	$Q_{F_3} = S_{F_3} R_Q$	
Slow reservoir ( $S_{S_1}$ )	$dS_{S_1}/dt = (1 - \alpha)P_e - Q_{S_1}$	$Q_{S_1} = S_{S_1} R_S$	

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

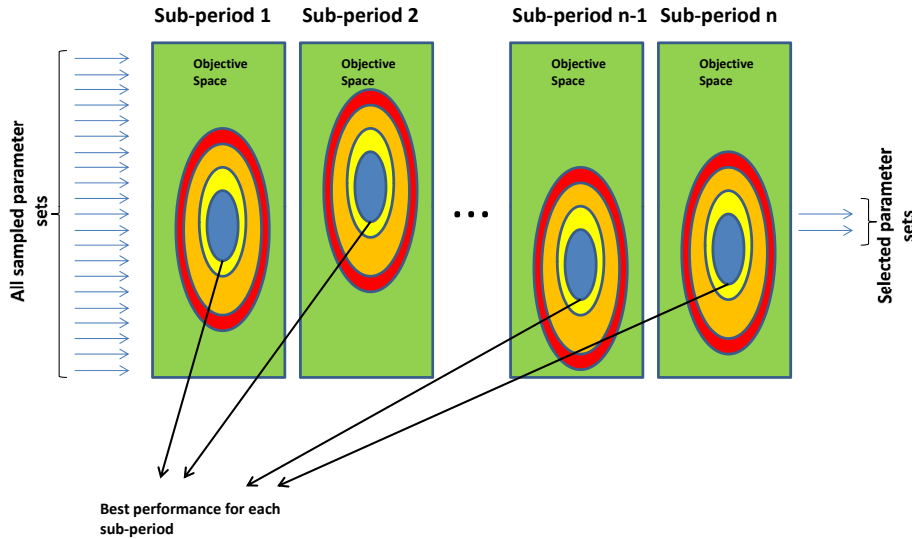
Full Screen / Esc

Printer-friendly Version

Interactive Discussion







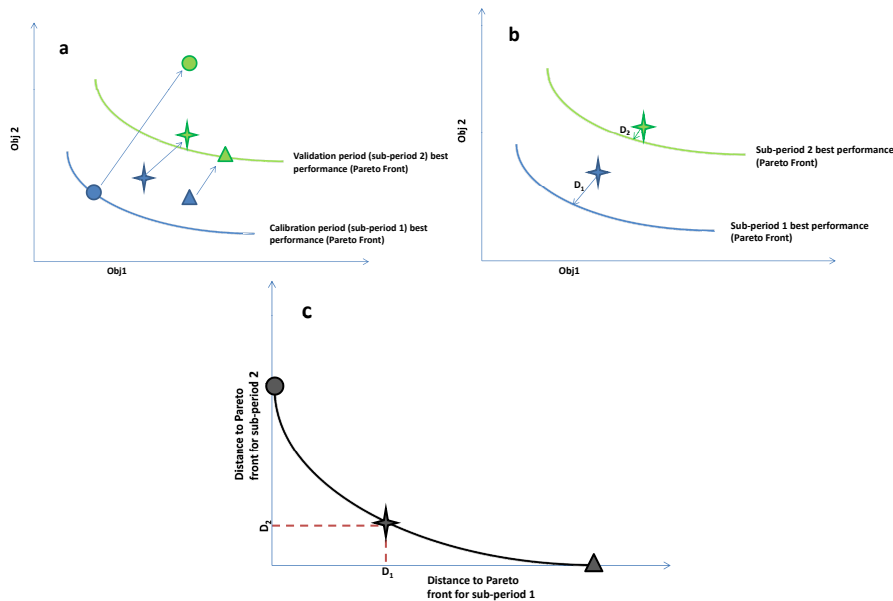
**Fig. 1.** Schematic illustration of sub-period calibration. The parameter sets which perform well for the entire sub-periods are retained.

## Sub-period calibration

S. Gharari et al.

<a href="#">Title Page</a>	
<a href="#">Abstract</a>	<a href="#">Introduction</a>
<a href="#">Conclusions</a>	<a href="#">References</a>
<a href="#">Tables</a>	<a href="#">Figures</a>
<a href="#">◀</a>	<a href="#">▶</a>
<a href="#">◀</a>	<a href="#">▶</a>
<a href="#">Back</a>	<a href="#">Close</a>
<a href="#">Full Screen / Esc</a>	
<a href="#">Printer-friendly Version</a>	
<a href="#">Interactive Discussion</a>	





**Fig. 2.** (a) Calibration-validation of a two dimensional abstract optimization problem; the lines represent the best available performance during calibration and validation periods Pareto fronts (CPF<sub>1</sub> and CPF<sub>2</sub> for sub-period calibration, respectively). The blue circle shows a CPF<sub>1</sub> member which performs poorly when validating it during the second sub-period, i.e. it plots far from the best available results as shown by CPF<sub>2</sub>, the reverse situation is illustrated by the green triangle which is a member of second sub-period (CPF<sub>2</sub>) but performing far from first sub-period Pareto front (CPF<sub>1</sub>), while the stars shows the performance of a non-CPF parameter set which performs relatively well in both sub-periods, i.e. for calibration and validation (CPF<sub>1</sub> and CPF<sub>2</sub>), (b) proposed method of calibration with reducing the distance to the optimal solution, i.e. to the Calibration Pareto Fronts (CPF<sub>1</sub> and CPF<sub>2</sub>), of each sub-period (c) Minimum Distance Pareto Front (MDPF) as generated by sub-period calibration; Star shows the trade of between performance related to each sub-period performance (CPF<sub>1</sub>, CPF<sub>2</sub>).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

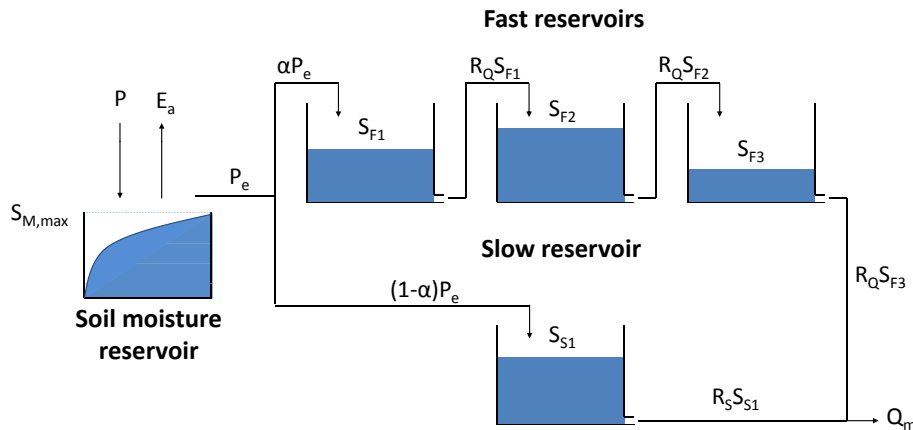
Printer-friendly Version

Interactive Discussion



## Sub-period calibration

S. Gharari et al.



**Fig. 3.** Schematic illustration of HyMod rainfall/runoff conceptual model.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

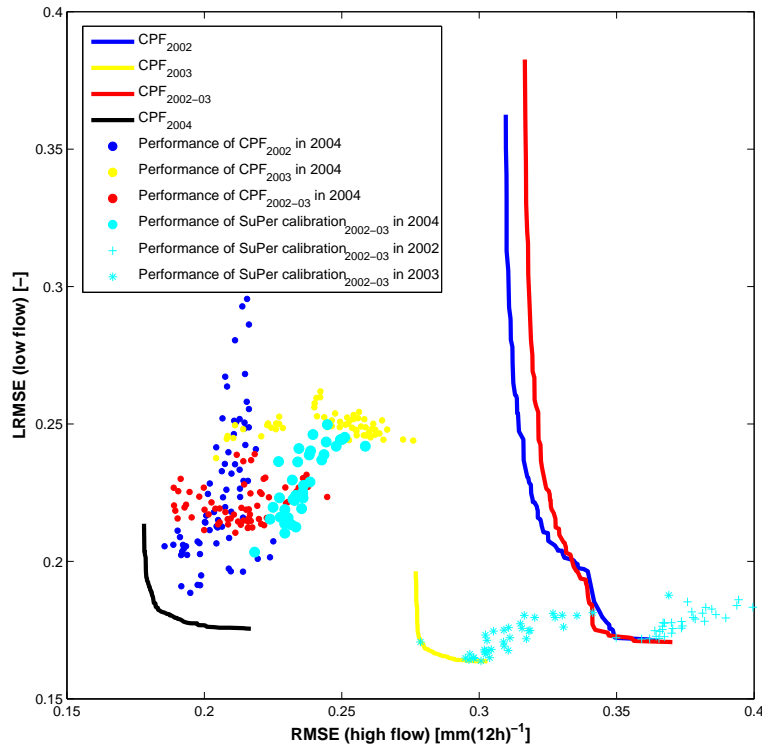
Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





**Fig. 4.** The calibration Pareto fronts based on 2002, 2003, 2004 and 2002–2003 ( $CPF_{2002}$ ,  $CPF_{2003}$ ,  $CPF_{2004}$ ,  $CPF_{2002-2003}$ ) are illustrated by dark blue, yellow, black and red lines, respectively. The dots of the same colors represent model performances using the  $CPF_{2002}$ ,  $CPF_{2003}$  and  $CPF_{2002-2003}$  members for target year 2004, i.e. the performance in traditional model validation. The light blue symbols show the performance of SuPer calibration parameter sets as identified by Minimum Distance Pareto Front members,  $MDPF_{2002-2003}$ , in 2002 (+), 2003 (\*) and 2004 (•).

**Sub-period calibration**

S. Gharari et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

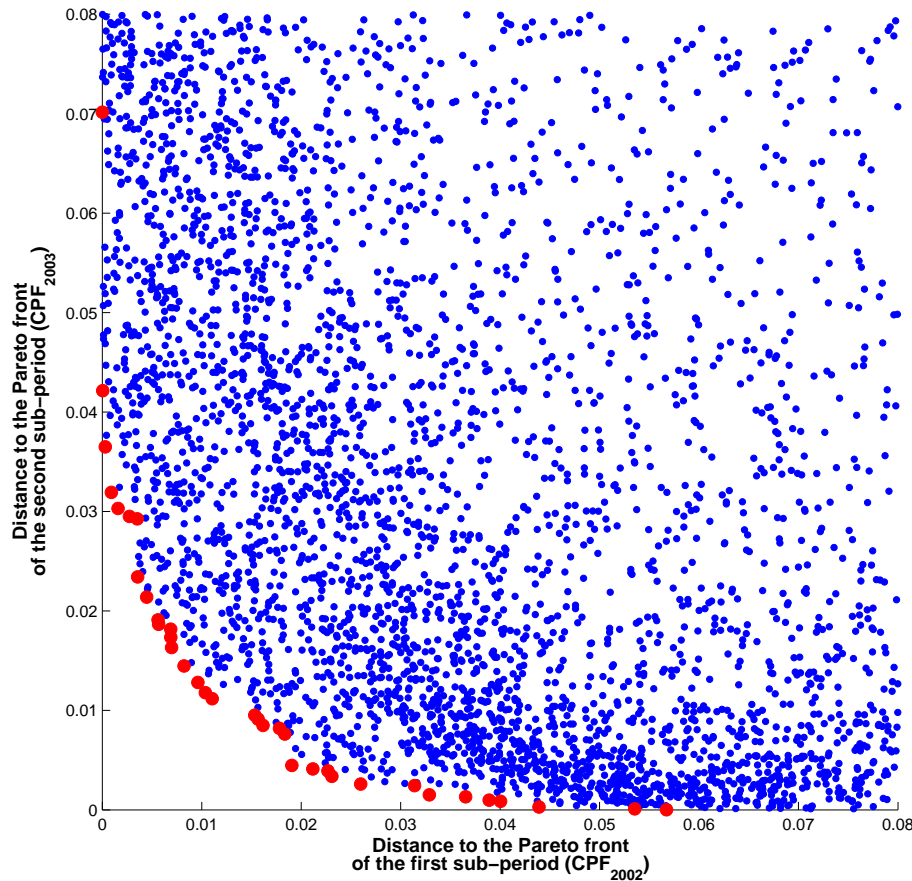
Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





**Fig. 5.** The two-dimensional Minimum Distance Pareto Front (MDPF, red dots) of SuPer calibration based on years 2002–2003 ( $\text{MDPF}_{2002-2003}$ ). MDPF indicates the trade of between performance of a parameter set regarding sub-period calibration Pareto fronts (CPF).

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

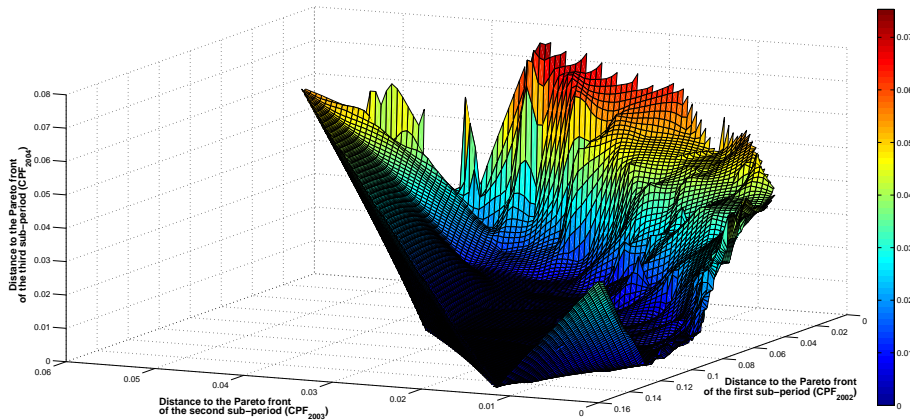
Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





**Fig. 6.** The three-dimensional Minimum Distance Pareto Front (MDPF, surface) of SuPer calibration based on years 2002–2004 ( $MDPF_{2002-2004}$ ). The color bar represents vertical values or the same distance to the Pareto front of the third sub-period ( $CPF_{2004}$ ). MDPF indicates the trade of between performance of a parameter set regarding sub-period calibration Pareto fronts (CPFs).

## Sub-period calibration

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

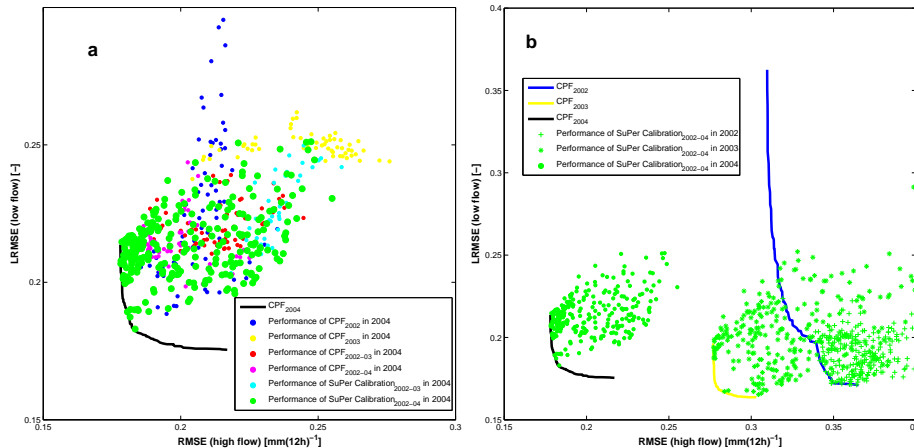
Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





**Fig. 7. (a)** The calibration Pareto fronts of 2004 is illustrated by black line. The dark blue, yellow, red, purple dots indicate the performance of Pareto members calibrated based 2002, 2003, 2002–2003, 2002–2004 ( $CPF_{2002}$ ,  $CPF_{2003}$ ,  $CPF_{2002-2003}$ ,  $CPF_{2002-2004}$ ) for target year (2004), respectively. Light blue and green dots indicate the performance of selected parameter sets by SuPer calibration based on 2002–2003 and 2002–2004 ( $MDPF_{2002-2003}$  and  $MDPF_{2002-2004}$ ) for target year (2004), respectively. **(b)** The calibration Pareto front of 2002, 2003 and 2004 ( $CPF_{2002}$ ,  $CPF_{2003}$ ,  $CPF_{2004}$ ) are illustrated by blue, yellow and black lines, respectively. The green symbols show behavior of SuPer calibration in 2002 (+), 2003 (\*) and 2004 (•). MDPF indicates the trade of between performance of a parameter set regarding sub-period calibration Pareto fronts (CPFs).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

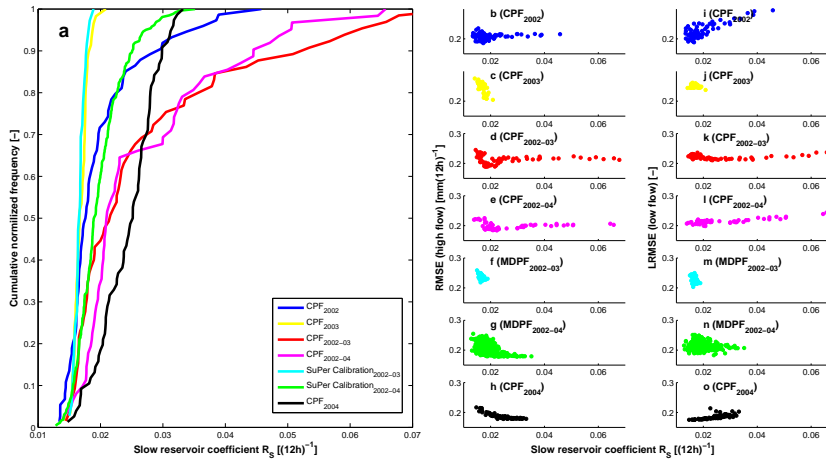
Close

Full Screen / Esc

Printer-friendly Version

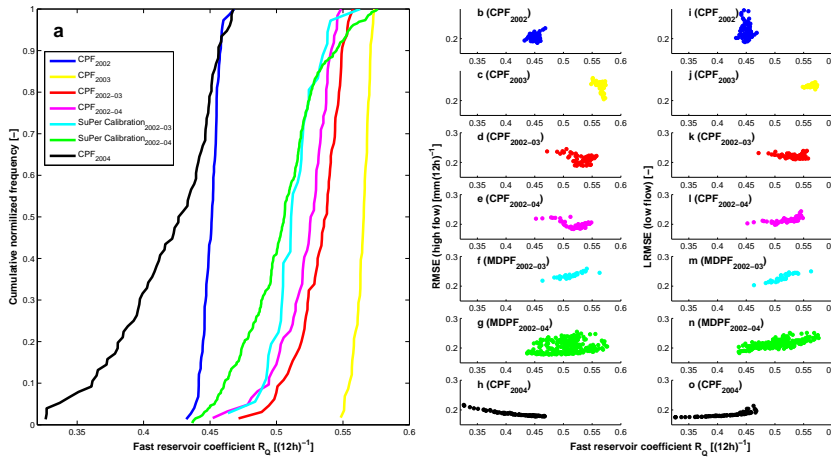
Interactive Discussion



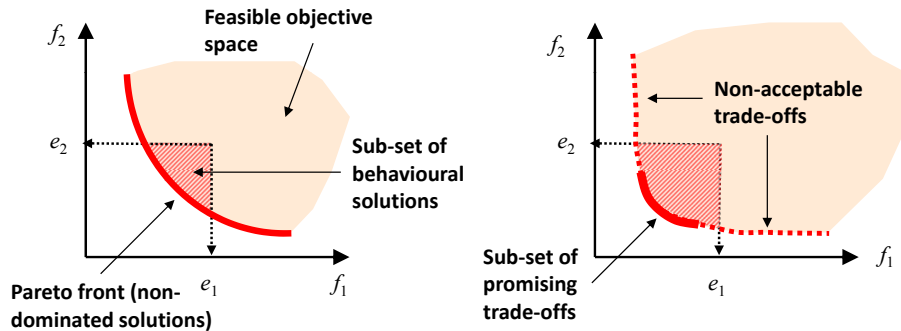


**Fig. 8.** Example results of the parameter  $R_S$ , describing the slow reservoir coefficient of Hy-Mod, (a) normalized cumulative frequency of Pareto members for different calibration strategies (CPFs and MDPFs). (b–e) Show the parameter ranges for the calibration Pareto fronts (CPFs): blue, yellow, red and purple dots are the  $R_S$  ranges which are the results of calibration based on 2002, 2003, 2002–2003 and 2002–2004, respectively (CPF<sub>2002</sub>, CPF<sub>2003</sub>, CPF<sub>2002–2003</sub>, CPF<sub>2002–2004</sub>) for RMSE (high flow). (f,g) Shows the parameter ranges obtained from SuPer calibration: light blue and green dots show the ranges of  $R_S$  regarding SuPer calibration based on 2002–2003 and 2002–2004 (MDPF<sub>2002–2003</sub> and MDPF<sub>2002–2004</sub>), respectively for RMSE (high flow). (h) Black dots show the range of  $R_S$  for CPF<sub>2004</sub> members with respect to RMSE (high flow). (i–o) Show the parameter ranges for the calibration Pareto fronts (CPFs): blue, yellow, red and purple dots are the  $R_S$  ranges which are the results of calibration based on 2002, 2003, 2002–2003 and 2002–2004, respectively (CPF<sub>2002</sub>, CPF<sub>2003</sub>, CPF<sub>2002–2003</sub>, CPF<sub>2002–2004</sub>) for LRMSE (low flow). (m,n) Shows the parameter ranges obtained from SuPer calibration: light blue and green dots show the ranges of  $R_S$  regarding SuPer calibration based on 2002–2003 and 2002–2004 (MDPF<sub>2002–2003</sub> and MDPF<sub>2002–2004</sub>), respectively for LRMSE (low flow). (o) Black dots show the range of  $R_S$  for CPF<sub>2004</sub> members with respect to LRMSE (low flow). MDPF indicates the trade of between performance of a parameter set regarding sub-period calibration Pareto fronts (CPFs).





**Fig. 9.** Example results of the parameter  $R_Q$ , describing the fast reservoir coefficient of Hy-Mod, **(a)** normalized cumulative frequency of Pareto members for different calibration strategies (CPFs and MDPFs). **(b–e)** Show the parameter ranges for the calibration Pareto fronts (CPFs): blue, yellow, red and purple dots are the  $R_Q$  ranges which are the results of calibration based on 2002, 2003, 2002–2003 and 2002–2004, respectively (CPF<sub>2002</sub>, CPF<sub>2003</sub>, CPF<sub>2002–2003</sub>, CPF<sub>2002–2004</sub>) for RMSE (high flow). **(f,g)** Shows the parameter ranges obtained from SuPer calibration: light blue and green dots show the ranges of  $R_Q$  regarding SuPer calibration based on 2002–2003 and 2002–2004 (MDPF<sub>2002–2003</sub> and MDPF<sub>2002–2004</sub>), respectively for RMSE (high flow). **(h)** Black dots show the range of  $R_Q$  for CPF<sub>2004</sub> members with respect to RMSE (high flow). **(i–o)** Show the parameter ranges for the calibration Pareto fronts (CPFs): blue, yellow, red and purple dots are the  $R_Q$  ranges which are the results of calibration based on 2002, 2003, 2002–2003 and 2002–2004, respectively (CPF<sub>2002</sub>, CPF<sub>2003</sub>, CPF<sub>2002–2003</sub>, CPF<sub>2002–2004</sub>) for LRMSE (low flow). **(m,n)** Shows the parameter ranges obtained from SuPer calibration: light blue and green dots show the ranges of  $R_Q$  regarding SuPer calibration based on 2002–2003 and 2002–2004 (MDPF<sub>2002–2003</sub> and MDPF<sub>2002–2004</sub>), respectively for LRMSE (low flow). **(o)** Black dots show the range of  $R_Q$  for CPF<sub>2004</sub> members with respect to LRMSE (low flow). MDPF indicates the trade of between performance of a parameter set regarding sub-period calibration Pareto fronts (CPFs).



**Fig. 10.** Graphical examples illustrating Pareto-optimal and behavioral solutions in the objective space, for two hypothetical problems of simultaneous minimization of two criteria  $[f_1, f_2]$  with smooth (left diagram) and steep (right diagram) trade-offs. Vector  $e = [e_1, e_2]$  indicates limits of acceptability, i.e. cut-off thresholds for distinguishing behavioral and non-behavioral solutions (Source: Efstratiadis and Koutsoyiannis, 2010, with permission of first author and publisher).

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

