

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.

Robust multi-objective calibration strategies – chances for improving flood forecasting

T. Krauß¹, J. Cullmann², P. Saile², and G. H. Schmitz¹

¹Institute of Hydrology and Meteorology, Department of Hydrology, Dresden University of Technology, Bergstr. 66, 01062 Dresden, Germany

²IHP/HWRP Secretariat, Federal Institute of Hydrology, Am Mainzer Tor 1, 56068 Koblenz, Germany

Received: 24 March 2011 – Accepted: 24 March 2011 – Published: 15 April 2011

Correspondence to: T. Krauß (thomas.krausse@tu-dresden.de)

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

HESSD

8, 3693–3741, 2011

**Robust
multi-objective
calibration strategies**

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

Process-oriented rainfall-runoff models are designed to approximate the complex hydrologic processes within a specific catchment and in particular to simulate the discharge at the catchment outlet. Most of these models exhibit a high degree of complexity and require the determination of various parameters by calibration. Recently automatic calibration methods became popular in order to identify parameter vectors with high corresponding model performance. The model performance is often assessed by a purpose-oriented objective function. Practical experience suggests that in many situations one single objective function cannot adequately describe the model's ability to represent any aspect of the catchment's behaviour. This is regardless whether the objective is aggregated of several criteria that measure different (possibly opposite) aspects of the system behaviour. One strategy to circumvent this problem is to define multiple objective functions and to apply a multi-objective optimisation algorithm to identify the set of Pareto optimal or non-dominated solutions. One possible approach to estimate the Pareto set effectively and efficiently is the particle swarm optimisation (PSO). It has already been successfully applied in various other fields and has been reported to show effective and efficient performance. Krauß and Cullmann (2011b) presented a method entitled $ROPE_{PSO}$ which merges the strengths of PSO and data depth measures in order to identify robust parameter vectors for hydrological models. In this paper we present a multi-objective parameter estimation algorithm, entitled the Multi-Objective Robust Particle Swarm Parameter Estimation (MO-ROPE). The algorithm is a further development of the previously mentioned single-objective $ROPE_{PSO}$ approach. It applies a newly developed multi-objective particle swarm optimisation algorithm in order to identify non-dominated robust model parameter vectors. Subsequently it samples robust parameter vectors by the application of data depth metrics. In a preliminary assessment MO-PSO-GA is compared with other multi-objective optimisation algorithms. In the frame of a real world case study MO-ROPE is applied identifying robust parameter vectors of a distributed hydrological model with focus on

HESSD

8, 3693–3741, 2011

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



flood events in a small, pre-alpine, and fast responding catchment in Switzerland. The method is compared with existing robust parameter estimation methods.

1 Introduction

Hydrological models are simplified, conceptual representations of a part of the hydro-
logic cycle. They relate rainfall to streamflow on a continuous basis. Many of those
models are driven by a vector of parameters that cannot be measured directly, but
must be determined through indirect methods. This is an important aspect of model
calibration. Efficient and effective parameter estimation techniques are a crucial factor
for the successful application of hydrological models. In the process of parameter esti-
mation the values of the model parameters are adjusted until the catchment behaviour
is closely matched.

Traditionally, this calibration is performed manually adjusting the parameters while vi-
sually inspecting the agreement between observations and model predictions. For the
application of automatic approaches the calibration is formulated as an optimisation
problem. A purpose-specific objective function f quantifies the agreement between
observations and simulation results. Many practical studies suggest that one single
objective function, no matter how carefully chosen, is often insufficient to represent all
characteristics of the system behaviour (Gupta et al., 1998; Cullmann, 2006; Gill et al.,
2006; Fenicia et al., 2007). For instance the mean absolute error of the discharge at
the catchment outlet might be a good indicator for the ability to represent the water
balance, however it is likely to be inadequate to measure the model performance for
flood forecasts where a correct simulation of the peak flow value and timing is crucial.
Consequently, single-objective calibration approaches that provide one unique global
best parameter vector are in many cases not considered acceptable by experienced
hydrologists. The most elementary solution to circumvent this problem is to aggregate
several objective functions. However, this approach involves a great deal of subjective
judgment and neglects the global bests for individual objective functions. Another

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



advanced option is a multi-objective view of the optimisation problem referring to the concept of Pareto optimality (Gupta et al., 1998). A multi-objective optimisation algorithm approximates the set of non-dominated, i.e. Pareto-optimal solutions with respect to a set of given objectives. The Pareto-optimal set or Pareto set reflects the trade-offs among all given objectives. Recently evolutionary algorithms, i.e. particle swarm optimisation or genetic algorithms have found favour in order to approximate this set.

A major disadvantage of automatic calibration procedures which understand the problem of model calibration just as an optimisation problem was studied by Bárdossy and Singh (2008): due to the complex-shaped response surface the solution of the optimisation problem can result in different near optimum parameter vectors that can lead to a very different performance on the validation data. The actual goal of a good model calibration should not be to find parameter vectors which perform best for the calibration period but to find parameter vectors which are robust. One possible approach to achieve this goal is the Robust Parameter Estimation approach that was first provided by Bárdossy and Singh (2008). It applies data depth metrics in order to sample robust parameter vectors with respect to a set of identified parameter vectors with good model performance. Recent studies (Bárdossy and Singh, 2008; Krauß and Cullmann, 2011a,b) presented the success of several development steps of this method. However, all ROPE approaches published so far just identify robust model parameter vectors with respect to one single objective. The consideration of multiple objectives is just possible by aggregation.

In this paper we present a new method, entitled Multi-Objective Robust Parameter Estimation (MO-ROPE) that synthesizes the advantages of both the multi-objective view and robust parameter estimation. In order to quantify the uncertainty of the parametrisation with respect to the given objectives, the method estimates a set of robust model parameter vectors applying a two-step approach. Within the first step, an evolutionary algorithm that combines ideas of swarm intelligence and genetic algorithms is used to approximate the Pareto-optimal set. In a second step parameter vectors with high data depth (with respect to the Pareto set) are sampled assuming that

**Robust
multi-objective
calibration strategies**

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



those parameter vectors are more robust than the complete Pareto set. The developed approach is tested on synthetical data. The results are compared with existing multi-objective calibration approaches. The real world case study shows the success of the MO-ROPE method in the calibration of a distributed hydrological model focussing on flood events in comparison with existing robust parameter estimation methods.

2 Overview over existing calibration strategies

The newly developed MO-ROPE synthesizes the concepts of multi-objective model calibration and robust parameter estimation. Therefore we will briefly introduce both concepts in this section.

2.1 Multi-objective model calibration

Multi-objective model calibration estimates the considered parameters of a given model by simultaneously optimising two or more conflicting objectives. The objectives may be subject to certain constraints. This is done referring to the concept of Pareto optimality or Pareto efficiency. A multi-objective optimisation problem can be stated as the minimisation of a set of objective functions which quantify the degree of match between the simulated and observed system behaviour. Table 3 gives an overview about different objective functions used in this paper. For a d -dimensional calibration task with p objectives the optimisation problem is defined as follows:

$$\min_{\theta \in \Theta} F(\theta) = \{f_1(\theta), f_2(\theta), \dots, f_p(\theta)\} \quad (1)$$

where $\theta = (\theta_1, \dots, \theta_d)$ is a d -dimensional vector in $\mathbb{R}^d$.

The optimisation task is to identify decision vectors¹ out of the feasible decision space Θ that will minimise all objective functions in F simultaneously. However, a

¹In the context of parameter estimation the decision vectors consist of the model parameters to be estimated and are thus often entitled parameter vectors.

unique global minimum for all objective functions in F only exists for exceptional cases². Thus, a set of non-dominated decision vectors is to be estimated according to the following definitions:

- a decision vector $\vec{u}$ *weakly dominates* another vector $\vec{v}$ if and only if $\forall i \in \{1, \dots, p\}$: $f_i(\vec{u}) \leq f_i(\vec{v})$. This is denoted by $\vec{u} \leq \vec{v}$.
- a decision vector $\vec{u}$ *dominates* another vector $\vec{v}$ if and only if $\vec{u} \leq \vec{v}$ and $\exists i \in \{1, \dots, p\}$: $f_i(\vec{u}) < f_i(\vec{v})$. This is denoted by $\vec{u} < \vec{v}$.
- a decision vector $\vec{u}$ is called *Pareto optimal* or *non-dominated* if and only if $\nexists \vec{w} \in \Theta : \vec{w} < \vec{u}$

The set of Pareto optimal solutions is denoted as the Pareto optimal set $\tilde{P}$ and the image of $\tilde{P}$ under the mapping F is called Pareto front. Often both terms are used synonymously. Consider that for objectives that have to be maximised either the negative of the objective has to be minimised or the operators $<$ and $\leq$ have to be substituted by $>$ and $\geq$ in the definition above. The principle is illustrated in Fig. 1 with the help of a multi-objective optimisation problem with two objectives. Hence, solving a multi-objective optimisation problem is to find the Pareto optimal solutions in the feasible decision space Θ .

The Pareto set characterises the uncertainty in the identification of a unique global best parameter vector with respect to the chosen objectives. Additionally it provides useful information about the limitations of the model. Unlike other single-objective uncertainty approaches, e.g. GLUE, BATEA, and SCEM-UA (Beven and Binley, 1992; Thyer et al., 2007; Vrugt et al., 2003b) a multi-objective view does not require arbitrary

²Consider that in such case multi-objective optimisation is not necessary, because the problem can reformulated as a single-objective optimisation problem with the same solution.

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



uncertainty thresholds, e.g. the 5th and 9th percentile. “Any fuzziness in the specification of the Pareto parameter spaces in the multi-objective approach arises from only two factors: (a) subjectivity in the selection of the measures in F and (b) the statistical uncertainty in the computation of each measure arising from sampling considerations” (s. Gupta et al., 1998, p. 756).

Recently evolutionary algorithms have been reported to successfully approximate the Pareto set for complicated non-linear problems (e.g. Zitzler et al., 2008). That is why they found favour in order to solve multi-objective calibration problems in many fields of study, including computational chemistry, bioinformatics, economics, computational science, and environmental sciences (e.g. Madsen, 2000; Gill et al., 2006; Tsou, 2008).

2.2 Robust parameter estimation

The application of automatic calibration procedures for model parameter estimation often completely neglects possible uncertainties in the observations used to quantify the matching of simulated values and measurements. Bárdossy and Singh (2008) studied this problem: due to the complex-shaped response surface and the erroneous observations the solution of the optimisation problem can lead to very different near optimum parameter vectors that correspond to a much different model performance on the validation data. The actual goal of a good model calibration should not be to find parameter vectors which perform best for the calibration period but to find parameter vectors which:

- lead to good model performance over the selected time period;
- lead to a hydrologically reasonable representation of the corresponding processes;
- are not sensitive: small changes of the parameters should not lead to very different results;

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



each point of the set with respect to the point set itself.

Using the example of the calibration of a hydrological model Bárdossy and Singh (2008) showed that parameter vectors with high data depth (with respect to a set of parameter vectors with good model performance) are on average more robust than such with low depth. In general ROPE consists of two steps:

1. In a first step a set of model parameter vectors with a reasonable good model performance X^* is identified. The estimated parameter vectors achieve the best possible performance with respect to the given objectives and calibration data. Thus, they are from now on called the good parameter vectors.
2. Afterwards a set of parameter vectors Y with high data depth with respect to the set of good parameter vectors is generated under the assumption that those parameter vectors are more robust than the complete set of good parameter vectors.

As mentioned above, the outcome of recent studies with further developed algorithms applying the ROPE principle showed that this concept can be very useful for the estimation of robust hydrological model parameters. For further details refer to the studies provided by Bárdossy and Singh (2008); Krauß and Cullmann (2011a,b). They apply ROPE to the calibration of hydrological models with daily and hourly time step.

3 Multi-objective robust parameter estimation

We propose to synthesize the strengths of multi-objective optimisation and robust parameter estimation with data depth functions in a new algorithm, entitled Multi-Objective Robust Parameter Estimation by Particle Swarm Optimisation (MO-ROPE). In this section we will introduce the newly developed MO-ROPE approach. It synthesizes the advantages of a multi-objective view and the robust parameter estimation. The approach applies a newly developed evolutionary algorithm in order to estimate the Pareto set.

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



In a second step a set of parameter vectors with high data depth with respect to the Pareto set is sampled by an algorithm called GenDeep. The introduction of the algorithm is rounded out by two synthetical benchmarks that have been widely used to assess the effectiveness and efficiency of multi-objective optimisation algorithms.

5

3.1 Multi-objective optimisation by particle swarm optimisation

Algorithm 3.1 MO-PSO-GA

1: initialise the non-dominated front: $\tilde{A} \leftarrow \emptyset$
2: **for all** particles i **do**
3: initialise position: $\mathbf{x}_i \leftarrow \mathcal{U}[\mathbf{x}_{lb}, \mathbf{x}_{ub}]$,
4: velocity $\mathbf{v}_i \leftarrow \mathbf{0}$,
5: local best $\mathbf{y}_i \leftarrow \mathbf{x}_i$
6: **end for**

// Iteration loop
7: **while** stopping criteria not met **do**
8: **for all** particles i **do**
9: compute $F(\mathbf{x}_i) = [f_1(\mathbf{x}_i), \dots, f_p(\mathbf{x}_i)]$
10: **end for**
// Estimate Pareto set
11: **for all** particles i **do**
12: **if** isempty($\tilde{A}$) **then**
13: $\tilde{A} \leftarrow \mathbf{x}_i$
14: **else if** $\exists u \in \tilde{A}$, where: $u \not\leq \mathbf{x}_i$ **then**
15: $\forall u \in \tilde{A}$, where: $\mathbf{x}_i < u$ **do**: $\tilde{A} \leftarrow \tilde{A} \setminus u$
16: $\tilde{A} \leftarrow \tilde{A} \cup \mathbf{x}_i$
17: **else**
18: do nothing to $\tilde{A}$
19: **end if**
20: **if** $\mathbf{x}_i \leq \mathbf{y}_i$ **then**
21: $\mathbf{y}_i \leftarrow \mathbf{x}_i$
22: **end if**
23: **end for**
// GA

Title Page

AbstractIntroduction

ConclusionsReferences

TablesFigures

◀▶

◀▶

BackClose

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



```

24:  $n_{\psi} \leftarrow \psi \cdot \#\{\text{particles}\}$ 
25: discard the worst  $n_{\psi}$  particles from the population
26: initialise the genetic offspring:  $o_{ga} \leftarrow \emptyset$ 
27: for  $i = 1$   $\frac{n_{\psi}}{2}$  do
28:   select a pair  $\{\mathbf{x}_1, \mathbf{x}_2\}$  from the population by tournament
   selection
29:   apply the VPAC operator to generate new offspring
    $\{\mathbf{x}_{1_o}, \mathbf{x}_{2_o}\} \leftarrow \text{VPAC}(\{\mathbf{x}_1, \mathbf{x}_2\})$ 
30:    $o_{ga} \leftarrow o_{ga} \cup \{\mathbf{x}_{1_o}, \mathbf{x}_{2_o}\}$ 
31: end for
   // PSO
32: for all particles  $i$  do
33:    $\mathbf{g}_i \leftarrow \text{random}(\tilde{A})$ 
   // update velocity and position
34:    $\mathbf{v}_i(t+1) \leftarrow \omega \cdot \mathbf{v}_i(t) + \phi_1 \cdot R_1 \cdot (\mathbf{g}_i(t) - \mathbf{x}_i(t)) + \phi_2 \cdot R_2 \cdot$ 
    $(\mathbf{y}_i(t) - \mathbf{x}_i(t))$ 
35:    $\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1)$ 
36: end for
   // merge the new population with genetic offspring
37: particles  $\leftarrow$  particles  $\cup o_{ga}$ 
38: end while

```

Recently several evolutionary search strategies, e.g. genetic algorithms (GA) and particle swarm optimisation (PSO) became popular to approximate the Pareto set. Both are population based algorithms that have proven to be successful in solving this task for an extensive variety of problems even in high dimensions effectively and efficiently.

We implemented a hybrid multi-objective search strategy based on both approaches with the goal to exploit their strengths, and to minimise their weaknesses. The developed approach is entitled Multi-Objective optimisation by Particle Swarm Optimisation and Genetic Algorithm (MO-PSO-GA).

A pseudocode listing of the algorithm is given in Algorithm 3.1. For the swarming of the particles we modified the approach provided by Tsou (2008). In difference to Gill et al. (2006) the (individual) global best for each particle is not the closest particle of the so far approximated Pareto optimal set $\tilde{A}$, but a random member of $\tilde{A}$. Furthermore

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



each particle i has a local best $\vec{y}_i$. The local best is updated by the current position $\vec{x}_i$, if it weakly dominates all so far found positions of the particle. Additionally we modified the algorithm according to the ideas provided by Settles and Soule (2005). They introduced a hybrid between a genetic algorithm (GA) and PSO. The algorithm behaves as a normal PSO algorithm besides an additional parameter, the breeding ratio ψ . This parameter determines the proportion of the population which are not moved according to PSO but will undergo breeding in the current generation. First of all the worst $\psi \cdot n$ particles are discarded, where n denotes the population size. Afterwards, candidates for breeding are nominated from the remaining population by tournament selection and recombined by a so called Velocity Propelled Averaged Crossover (VPAC) operator. This operator was introduced by Settles and Soule (2005). Its goal is to create two child particles whose position is between its parents' position, but accelerated away from their current direction (negative velocity) in order to increase diversity in the population. After a given number of iterations $\tilde{A}$ holds a set of parameter vectors that represent the Pareto set of the given multi-optimisation problem.

3.2 Synthesizing multi-objective optimisation and robust parameter estimation

Algorithm 3.2 MO-ROPE

- 1: Execute the multi-objective particle swarm based MO-PSO-GA procedure to estimate an approximation of the Pareto front of the problem ($\tilde{A}$).
- 2: Sample parameter vectors $\tilde{A}_d$ with high data depth w.r.t. $\tilde{A}$ by the *GenDeep* strategy provided in Krauß and Cullmann (2011a)
- 3: **return** Y

The set of non-dominated solutions $\tilde{A}$ estimated by MO-PSO-GA³ is the initial point of the second algorithmic step the sampling of robust parameter vectors. The *GenDeep*

³Consider that MO-PSO-GA can be substituted by any suitable multi-objective optimisation algorithm.

function presented in Krauße and Cullmann (2011a) samples a set of parameter vectors $\tilde{A}_d$ with high data depth with respect to $\tilde{A}$.

The used data depth function is the halfspace depth or Tukey depth, first introduced by Tukey (1975). According to Donoho and Gasko (1992) the halfspace depth of an arbitrary point $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$ with respect to a d -dimensional data set Z is defined as the smallest number of data points in any closed halfspace with boundary through θ . This is also called the Tukey or location depth, and it can be written as:

$$\text{hdepth}(\theta|Z) := \min_{\|\vec{u}\|=1} \#\{i, \vec{u}^\top \vec{x}_i \geq \vec{u}^\top \theta\} \quad (2)$$

where $\vec{u}$ ranges over all vectors in $\mathbb{R}_d$ with $\|\vec{u}\| = 1$.

Very often the halfspace depth is normalised by dividing the hdepth value by the number of points in the set Z :

$$\text{hdepth}^*(\theta|Z) := \frac{\text{hdepth}(\theta|Z)}{\#\{Z\}} \quad (3)$$

A pseudocode listing of the proposed MO-ROPE is provided in Algorithm 3.2. The developed solution addresses some of the drawbacks of existing multi-objective and robust single-objective calibration procedures. It provides a good possibility for the identification of robust model parameter vectors with respect to multiple calibration objectives. The algorithm was implemented in MATLAB and C++ and is embedded in a robust parameter estimation framework which comprises other published robust parameter estimation approaches. The framework is open source and available from the author.

3.3 Synthetic test problems

The published results of single-objective robust parameter estimation approaches have shown that the quality of the set of good parameter vectors plays an important role

for the robustness of the estimated model parameter vectors (Krauß and Cullmann, 2011a,b). This is why an effective estimation of the non-dominated solutions in the first algorithmic step is an important prerequisite for a successful application of a multi-objective robust parameter estimation approach. Thus, in a preliminary case study we applied the MO-PSO-GA algorithm in order to approximate the Pareto set for several synthetic problems with analytical solutions, in order to compare our MO-PSO-GA algorithm with already published approaches for multi-objective optimisation.

3.3.1 Test function I

For a first comparison of MO-PSO-GA with other approaches we selected a benchmark that was already chosen by previous publications dealing with multi-objective particle swarm optimisation algorithms (e.g. Coello et al., 2004; Gill et al., 2006). The benchmark, we refer to as *test function I*, was first used by Kita et al. (1996). The problem is to maximise the following function:

$$\max_{(x,y) \in \Theta} F = \{f_1(x,y), f_2(x,y)\} \quad (4)$$

where

$$f_1(x,y) = -x^2 + y, \quad f_2(x,y) = \frac{x}{2} + y + 1 \quad (5)$$

subject to

$$0 \geq \frac{x}{6} + y - 6.5, \quad 0 \geq \frac{x}{2} + y - 7.5,$$

$$0 \geq 5x + y - 30, \quad x, y \geq 0$$

in the range $0 \leq x, y \leq 7$.

Following the settings in Coello et al. (2004) and Gill et al. (2006). we ran the MO-PSO-GA algorithm with the same population size (30) and the same number of function

evaluations (5000). The true front of test function I and the non-dominated front estimated by MO-PSO-GA are shown in Fig. 3. It is obvious that the true front is approximated well and all sections of the front are uniformly covered. For an objective comparison of the approximated front against the true front we calculated several performance metrics and compared them with a published overview according to Gill et al. (2006). We computed the generational distance (GD) metric introduced by van Veldhuizen and Lamont (1998) and the spacing metric (SP) introduced by Schott (1995). The GD is a measure of the distance between the elements of the current non-dominated set and the true Pareto front, whereas the SP measures the mutual distance between those elements. Lower values of GD and SP denote a better approximation and a more uniform spread, respectively, with zero being the optimum. An overview of the performance metrics for the estimates of *test function I* computed by several multi-objective optimisation algorithms is given in Table 1. The nearly perfect GD value of 0.0006 indicates that the MO-PSO-GA estimates almost perfectly approximate the true Pareto front of *test function I*. Furthermore the SP value of 0.062 shows that the solutions are well spread all over the complete estimated front. The given results were computed by averaging the performance metrics for 25 individual algorithm runs. The particle swarm based MOPSO presented by Gill et al. (2006) and all other approaches compared in that paper are outperformed by the MO-PSO-GA.

3.3.2 Test function II

Additionally we applied the algorithm to another synthetical multi-objective test problem, which was used as a benchmark in Vrugt et al. (2003a) and Gill et al. (2006). We refer to this test problem as *test function II*. The problem is to minimise three two dimensional functions as follows:

$$\min_{\theta \in \Theta} F(\theta) = \{f_1, f_2, f_3\} \quad (6)$$

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



where

$$f_1(\theta_1, \theta_2) = \theta_1^2 + \theta_2^2$$

$$f_2(\theta_1, \theta_2) = (\theta_1 - 1)^2 + \theta_2^2$$

$$f_3(\theta_1, \theta_2) = \theta_1^2 + (\theta_2 - 1)^2$$

The Pareto solution set corresponding to Eq. (6) consists of a triangular-shaped area in the parameter space, having the corner points (0,0), (0,1) and (1,0) for θ_1 and θ_2 , respectively.

Referring to the original publication of Vrugt et al. (2003a) we applied MO-PSO-GA to estimate an approximation of the actual Pareto set of this problem. Just like Vrugt et al. (2003a) we set the maximum number of function evaluations to 5000. We set the population size to values of 10, 20, 50, and 100 respectively. Figure 5 provides scatter plots of the corresponding results. It is obvious that MO-PSO-GA performs well with uniform approximations of the true front irrespective of the specified population size. Consider that the higher population sizes do not forcefully yield better results, because the maximum number of function evaluations also limits the maximum number of iteration. Hence, for a maximum number of function evaluations of 5000 the run with a population size of 50 was iterated 100 times whereas the run with population size 10 was iterated 500 times. These results for *test function II* are very comparable to those of the MOSCEM provided by Vrugt et al. (2003a) and another multi-objective optimisation algorithm based on swarm intelligence presented by Gill et al. (2006). A detailed visual comparison confirms that the non-dominated front estimated by MO-PSO-GA is less clustered and provides a fairly better approximation of the true Pareto front. We repeated the runs once again but set the number of generations to 50 for all population sizes. The results are given in Fig. 5. Already the run with population size 20 and a corresponding number of function evaluations of just 1000 provides better results than the runs with MOSCEM and MO-PSO presented by Vrugt et al. (2003a)

Robust multi-objective calibration strategies

T. Krauße et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



and Gill et al. (2006) respectively. This underlines the efficiency of the MO-PSO-GA algorithm.

Even though these results are quite convincing, Wolpert (1997) showed that there is “no free lunch”, i.e. it is impossible to construct one single (parameter) search algorithm that will always outperform any other algorithm. In the last years some revolutionary approaches have been developed to benefit from the strengths of multiple search strategies, e.g. the multi-algorithm, genetically adaptive multi-objective method (AMAL-GAM) presented by Vrugt and Robinson (2007). Consider that it is easily possible to integrate the developed search algorithm into such a multi-algorithm framework.

4 Calibrating WaSiM with MO-ROPE focussing on flood forecasting

As the second application, the complete MO-ROPE approach is tested on the calibration of the distributed hydrological model WaSiM-ETH/6.4 model (in the further referred to as WaSiM). The model was introduced by Schulla (1997). WaSiM transforms rainfall into runoff according to the scheme shown in Fig. 6. It exemplarily shows that the soil water compartments receive infiltration which is computed by a modified approach according to Green and Ampt (1911). This module is also used to determine the direct runoff Q_d and the interflow Q_{ifl} in the model. Q_d is then routed via a flow-time grid and finally projected cell-wise to the catchment outlet by means of a simple bucket type function. The recession coefficient of this function is the model parameter k_d . The soil water movement through the different soil layers is modeled by means of the discrete form of the Richards-equation.

The WaSiM model was calibrated for the Rietholzbach catchment (3.18 km²) a small prealpine catchment located in the north-east of Switzerland. Figure 7 provides a 3-D-view of the catchment area. A significant number of studies have been conducted in this basin. For further information refer to Gurtz et al. (1999); Zappa (2002) and the website <http://www.iac.ethz.ch/research/rietholzbach>. In this case study WaSiM will be calibrated for the simulation of extreme discharges. Out of a time series of

**Robust
multi-objective
calibration strategies**

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



27 yr (1981–2008) of hourly measurements (precipitation, temperature, wind speed, global radiation, and streamflow), 24 significant flood events were identified. For further details and a more comprehensive overview refer to Krauß and Cullmann (2011a). Five flood events were used for calibration. The estimated parameter vectors were validated on 19 other flood events.

Preliminary sensitivity analyses (Cullmann, 2006; Seifert, 2010) showed that both the uncertainty in the identification of the soil hydraulic parameters and three further conceptual model parameters have a tremendous influence on the simulation results. These parameters were thus considered for calibration (Table 2). The simultaneous calibration of both soil hydraulic and conceptual model parameters was done following an approach provided in Grundmann (2010) that was already successfully applied in several other case studies (e.g. Krauß and Cullmann, 2011b). An overview over all possible objectives referred to in the following case studies is provided in Table 3.

4.1 Comparing single and multi-objective ROPE algorithms

In a first case study we calibrated WaSiM with the MO-ROPE algorithm in terms of two objective functions: rPD and NS. The number of function evaluations was set to 5000. Additionally we applied the single-objective robust parameter estimation algorithm $ROPE_{PSO}$ to this problem using rPD, NS and their aggregate FloodSkill as objective functions. The distributions of the estimated parameter vectors and some basic statistical properties, i.e. the mean value μ and the standard deviation σ , are provided in Table 5. In general the variance of the estimated distributions is higher for the multi-objective MO-ROPE approach than for the single-objective $ROPE_{PSO}$. Within the single-objective calibration runs the variance of the estimated distributions for the model parameters is in general the higher the more global the considered objective. For instance the variance of the $ROPE_{PSO}$ NS run estimates tends to be higher than those of the $ROPE_{PSO}$ rPD estimates. Furthermore there is an obviously strong dependence of the parameters k_d and k_{rec} on the used criterion. The more a correct representation of the observed peak flow value is measured by the used objective criterion, the lower

**Robust
multi-objective
calibration strategies**

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the value of k_d and the higher the value of k_{rec} . In general these results are reasonable and consistent with the model structure and the corresponding understanding of the hydrological system. Lower values of k_d increase the dynamics of the generated direct runoff, a higher value of k_{rec} decreases the effective saturated conductivity of deeper soil layers. This leads to a faster generation of direct runoff. However, within the basin direct runoff on the surface has hardly ever been observed, even not during large and intensive convective storm events. Thus these results are already a subtle hint to shortcomings in the model structure.

This assumption is supported by the achieved calibration performance for the multi-objective run. The trade-off surface for both rPD and NS on the basis of all parameter vectors evaluated by the MO-PSO-GA and the estimated approximation of the Pareto optimal set $\tilde{A}$ is given in Fig. 8. The evaluation of the model performances corresponding to the estimated parameter vectors show a clear tradeoff between both used objectives. That means that the best parameter vectors with respect to the peak flow value cannot represent the global behaviour of the catchment for flood events equally well. The advantage of the depth based sampling gets obvious from the plot given in Fig. 9. The sampled deep parameter vectors $\tilde{A}_d$ (with respect to the approximated Pareto optimal set $\tilde{A}$) show a better performance on the validation data. The deep parameter vectors are a good approximation of the (theoretical) Pareto front in the objective space based on the validation data. The tails of the Pareto front estimated in the calibration are not required. For example the best parameter vectors with respect to rPD on the calibration data do not have a better rPD in the validation than the sampled deep parameter vectors. However those vectors have a significantly worse NS. This shows that the deep parameter vectors are better transferable to other periods and events and thus more robust. The advantage of the depth based sampling gets even more obvious from the plots provided in Fig. 10. It visualises the dependency between the model performance of a sampled parameter vector on the validation data in terms of the FloodSkill criterion⁴ and its data depth with respect to the approximated

⁴The FloodSkill criterion is a compromise between rPD and NS (see Table 3 at the beginning

Robust multi-objective calibration strategies

T. Krauße et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Pareto optimal set. The parameter vectors with high data depth show a on average better model performance (lower FloodSkill values) with less variance than the hull of the Pareto optimal set $\tilde{A}_h$. This implies that parameter vectors with low data depth are more likely to be an outlier with bad model performance. The estimated results confirm the underlying assumption of the ROPE approach for multi-objective calibration tasks.

The model performances for multi-objective calibration were compared with the results of the single-objective robust parameter estimation runs. They are provided in Fig. 11. It is evident that due to the tradeoff the multi-objective calibration makes sense. It provides the best solution with respect to both objectives rPD and NS. The sets of robust parameter vectors Y_{rPD} and Y_{NS} estimated by ROPE_{PSO} considering just the rPD or NS show indeed a good validation performance on those criteria. However the model performance of the other criterion not considered for calibration is not sufficient. Also the single-objective solutions for the aggregated FloodSkill criterion $Y_{FloodSkill}$ overemphasizes the NS and provides just slightly better results in terms of the peak flow deviation in comparison to the single-objective calibration considering just the NS. The corresponding hydrographs of three validation events and the corresponding parameter and model uncertainties for both the solution $\tilde{A}_d$ estimated by MO-ROPE and $Y_{FloodSkill}$ estimated by ROPE_{PSO} are provided in Fig. 12. The complete model uncertainty was computed by two normal distributions fitted on the positive and negative discharge errors, transformed with the normal quantile transformation (NQT) (Krzysztofowicz, 1997) according to a method presented by Engeland et al. (2010). Contrary to Engeland et al. (2010) we considered observations with a discharge greater or equal $0.33 \text{ mm}^{-1} \text{ h}$ to account for the focus on the correct simulation of higher flow values. The hydrographs confirm the results of the previous analysis of the model performances. The peak flow values are better represented by the parameter vectors estimated by the multi-objective calibration. The confidence band of the parameter uncertainty is slightly smaller for the multi-objective calibration than for the single-objective calibration. However, the complete model uncertainty is approximately the same for both approaches. That is due to

of this section)

**Robust
multi-objective
calibration strategies**

T. Krauß et al.

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



the discussed shortcomings of the model structure and other neglected uncertainties in the observations.

4.1.1 Using soil moisture information and multi-objective parameter estimation to improve model calibration

In another case study we calibrated WaSiM again. However, this time we also accounted for the provided information of the soil moisture in the catchment for model calibration. The Rietholzbach catchment has been intensively monitored as a scientific research catchment since 1975. Lysimeter measurements deliver the necessary soil moisture information on hourly basis. On average the measurements are representative for the whole catchment for the long term. This is due to the small size of the basin, the location of the lysimeter station in the centre of the basin situated on grassland which is the major land use in the catchment. Unfortunately there are no significant studies for the catchment that examine the soil moisture dynamics for storm events in a more distributed way. We assume that a good approximation of the catchment behaviour (when focussing on flood events) requires a correct match of the dynamics of the soil moisture but no exact approximation of the measured water contents. Hence, we chose the correlation between observed and simulated as an additional objective. The underlying idea is that parameter vectors that do not only correspond to a good representation of catchment's behaviour with respect to the measured discharge but also the mean soil moisture dynamics are potentially better transferable to other flood events and can be used for extrapolation, i.e. they are potentially more robust.

Again we calibrated WaSiM with the MO-ROPE algorithm in terms of the two objective functions already used in the previous case study (rPD, NS) and the correlation between the measured and simulated soil moisture in the upper 50 cm of the soil layer. The assignment of calibration and validation events was changed due to some gaps in the provided soil moisture measurements. All other settings remained the same. In two runs robust parameter vectors were estimated in terms of all three objectives and just the rPD and NS. The distributions of the estimated parameter vectors are provided

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



in Table 5. For a better comparison the results of the previous case study with just 2 objectives considered are given as well. Considering three objectives the distribution of the most sensitive parameter k_d tends to smaller values than for the calibration run considering just the objectives rPD and NS. The distribution of the parameter k_i has approximately the same mean and variance but a significantly larger negative skewness. The mean value of the parameter dr is approximately the same with a significantly lower variance. The same holds for the soil parameters β_{SL} and β_{SiL} . An obvious difference can be observed in the distribution of the parameter k_{rec} . The mean value is significantly lower with a higher variance. That means that the effective saturated conductivity for deeper soil layers decreases faster, i.e. the soils in the model have on average a lower saturated conductivity.

The corresponding model performances on the validation data for both the 2-objective calibration and the 3-objective calibration considering the soil moisture dynamics are provided in Fig. 14. The scatter plots show the rPD and the NS for the approximated Pareto-optimal set $\tilde{A}$ and the subsequently sampled robust parameter vectors $\tilde{A}_d$. The boxplot below provides the correlation between observed and simulated soil moisture for the robust sets $\tilde{A}_d$. The improvements in terms of the correlation between the observed and simulated soil moisture are negligibly small for the validation period. An increase of the correlation coefficient of just one or two hundredths is no significant improvement. However, the validation performance in terms of the criteria assessing a good representation of the observed runoff is significantly worse, particularly for the rPD. The rPD on the validation events is approximately in the range of 0.25–0.40 instead of 0.19–0.34 for the 2-objective calibration, the NS is in the range of 0.41–0.64 instead of 0.34–0.67. Hence, here the soil moisture dynamics is no suitable additional criterion in order to improve the calibration results. This result is rather surprising.

One possible explanation might be the shortcomings in the model structure, already discussed in the previous case study. The soils of the Rietholzbach catchment are characterised by many macropores and some smaller drainages. This induces the

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



generation of preferential flow which is a very fast runoff component. However, the water movement in the unsaturated zone is described by the model in terms of the Richards equation that account just for the matrix flow in the soil but cannot represent the preferential flow component natively. Thus, the fast runoff components are supposed to be modeled by the direct runoff. The identified parameter vectors provide a good representation of the catchment's behaviour in terms of the discharge at the outlet. However, the additional criterion which assesses a reasonable representation of the observed soil moisture dynamics is rather counterproductive. This result is supported by some further evaluation of the previous case study. Figure 15 provides information about the dependance between soil moisture correlation r_Θ and data depth with respect to the Pareto set $\tilde{A}$ for both $\tilde{A}$ and the accordingly sampled deep parameter vectors $\tilde{A}_d$. The hull and the deep vectors are indicated by blue and red dots respectively. It is obvious that deep parameter vectors in $\tilde{A}_d$ which are robust in terms of rPD and NS provide no better correlation between observed and simulated soil moisture. Another possible explanation for the worse calibration results are the observations themselves. Due to the potentially very heterogeneous conditions in the unsaturated zone, one single lysimeter might notwithstanding the above-mentioned considerations not be significant for the short-term soil moisture dynamics in the whole catchment. Other case studies on a larger scale with a lot of measurements showed that the soil moisture can be an appropriate additional calibration objective, also for flood events (e.g Norbiato et al., 2008).

This underlines that robust parameter estimation can identify the most robust solutions within the given constraints. However, a good selection of appropriate calibration objectives and a suitable model structure are as important as a reliable and robust model parametrisation.

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



5 Discussion and conclusions

- This paper presents a multi-objective parameter estimation approach which is well suited for solving multi-criterial parameter estimation problems. The developed approach, entitled Multi-Objective Robust Parameter Estimation by Particle Swarm Optimisation and Genetic Algorithms (MO-ROPE), merges the strength of evolutionary multi-objective optimisation algorithms and depth based parameter sampling. MO-ROPE consists of two sub-components that are applied subsequently. A newly developed evolutionary optimisation algorithm (MO-PSO-GA) employs the concepts of particle swarm optimisation and genetic programming in order to effectively estimate the Pareto-optimal set. As a kind of post-processor we apply the concept of data depth to sample robust model parameter vectors with respect of the identified Pareto-optimal set. We study the efficiency and effectiveness of the developed solution by means of two synthetical benchmarks and the calibration of a process-oriented hydrological model.
- Two synthetical benchmarks assess the efficiency and effectiveness of the newly developed MO-PSO-GA algorithm in order to estimate Pareto-optimal set for a given (constrained) multi-objective optimisation problem. For the given problems the MO-PSO-GA proved its reliability and efficiency in comparison with other established approaches. The quality of the MO-PSO-GA estimates are at least as good as, in most cases even better than those estimated by approaches provided in Vrugt et al. (2003b); Coello et al. (2004); Gill et al. (2006).
- In a real world case study we compared the multi-objective MO-ROPE approach with the single-objective robust parameter estimation approach ROPE_{PSO} estimating three conceptual model parameters and three soil parameters of the hydrological model WaSiM focussing on flood events. Previous studies have already shown that the model has problems to represent the global catchment behaviour and the peak flow values equally well. That applies in particular to small and fast

HESSD

8, 3693–3741, 2011

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



responding catchments. Possible causes include problems with the model structure, uncertainty in the measurements, etc. The results of this study show that the multi-objective robust parameter estimation approach is a preferable option in such cases.

- In the scope of a second application we studied the calibration of WaSiM considering both observed discharge and soil moisture in order to obtain hydrologically reasonable parameter vectors representing the catchment behaviour with respect to both the discharge at the outlet and soil moisture dynamics. However, the calibration did not improve. That was due to shortcomings in the model structure and possibly insignificant soil moisture measurements at only one single spot. We strongly propose similar applications at larger scale with a sufficient set of gaging stations.
- Again we showed that parameter vectors with equal model performance on the calibration data can lead to very different results in validation. The proposed method of an evolutionary sampling of model parameter vectors by the help of data depth functions can help to identify sets of robust parameter vectors. In general parameters with low data depth are sensitive to small changes and cannot be tranfered to other time periods as well as those with high depths.
- In this paper, the Pareto optimal set was estimated by the MO-PSO-GA algorithm. The presented algorithm can be easily substituted by any other suitable multi-objective optimisation algorithm. Furthermore the developed MO-PSO-GA can be easily integrated into a multi-algorithm framework, e.g. AMALGAM (Vrugt and Robinson, 2007).

The application of the robust parameter estimation approach is a relatively new method which was applied to a limited number of case studies. We strongly propose its application to further models, catchments and also other fields of study where measurement errors with unknown distribution and model structures that cannot be easily identified are present.

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Acknowledgements. We would like to thank the German Research Foundation and the Cusanuswerk for the funding of this work. Furthermore we thank Irene Lehner from the land-climate interactions group at the ETH Zurich for the provision of all the measurement data in the Rietholz bach catchment and the Center for Information Services and High Performance Computing (ZIH) at the University of Technology, Dresden for the supply with the required computing power to carry out all the optimisation runs.

References

- Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Process.*, 6, 279–298, doi:10.1002/hyp.3360060305, 1992. 3698
- Brakensiek, D., Rawls, W., and Stephenson, G.: Modifying SCS hydrological soil groups and curve numbers for rangeland soils, *ASAE Paper*, 203, 1984. 3723
- Bremner, D., Chen, D., Iacono, J., Langerman, S., and Morin, P.: Output-sensitive algorithms for Tukey depth and related problems, *Stat. Comput.*, 18, 259–266, 2008. 3700
- 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. 3696, 3699, 3700, 3701
- Coello, C., Pulido, G., and Lechuga, M.: Handling multiple objectives with particle swarm optimization, *IEEE T. Evolut. Comput.*, 8, 256–279, 2004. 3706, 3716, 3722
- Cramer, K.: Multivariate Ausreißer und Datentiefe, Ph.D. thesis, Wirtschafts- und Sozialwissenschaftliche Fakultät, Universität zu Köln, 2003. 3700
- Cullmann, J.: Online flood forecasting in fast responding catchments on the basis of a synthesis of artificial neural networks and process models, Vol. 4 of *Dresdner Schriften zur Hydrologie*, Institut für Hydrologie und Meteorologie, Technische Universität Dresden, 2006. 3695, 3710
- Donoho, D. L. and Gasko, M.: Breakdown Properties of Location Estimates Based on Halfspace Depth and Projected Outlyingness, *The Annals of Statistics*, 20, 1803–1827, available at: <http://www.jstor.org/stable/2242368>, 1992. 3705
- Engeland, K., Renard, B., Steinsland, I., and Kolberg, S.: Evaluation of statistical models for forecast errors from the HBV model, *J. Hydrol.*, 384, 142–155, doi:10.1016/j.jhydrol.2010.01.018, 2010. 3712
- Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: A comparison of alternative multi-

HESSD

8, 3693–3741, 2011

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- objective calibration strategies for hydrological modeling, *Water Resour. Res.*, 43, W03434, doi:10.1029/2006WR005098, 2007. 3695
- Gill, M. K., Kaheil, Y. H., Khalil, A., McKee, M., and Bastidas, L.: Multiobjective particle swarm optimization for parameter estimation in hydrology, *Water Resour. Res.*, 42, W07417, doi:10.1029/2005WR004528, 2006. 3695, 3699, 3703, 3706, 3707, 3708, 3709, 3716, 3722
- Green, W. and Ampt, G.: Studies on Soil Physics: I. The flow of air and water through soils, *J. Agr. Sci.*, 4, 1–24, 1911. 3709
- Grundmann, J.: Analyse und Simulation von Unsicherheiten in der flächendifferenzierten Niederschlags-Abfluss-Modellierung, Vol. 8 of *Dresdner Schriften zur Hydrologie*, Institut für Hydrologie und Meteorologie, Technische Universität Dresden, 2010. 3710
- 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, 1998. 3695, 3696, 3699
- Gurtz, J., Baltensweiler, A., and Lang, H.: Spatially distributed hydrotope-based modelling of evapotranspiration and runoff in mountainous basins, *Hydrol. Process.*, 13, 2751–2768, 1999. 3709
- Kita, H., Yabumoto, Y., Mori, N., and Nishikawa, Y.: Parallel Problem Solving From Nature, chap. Multi-objective optimization by means of the thermodynamical genetic algorithm, *PPSN IV*, Springer, New York, 504–512, 1996. 3706
- Krauß, T. and Cullmann, J.: Identification of hydrological model parameters for flood forecasting using data depth measures, *Hydrol. Earth Syst. Sci. Discuss.*, 8, 2423–2476, doi:10.5194/hessd-8-2423-2011, 2011a. 3696, 3701, 3704, 3705, 3706, 3710
- Krauß, T. and Cullmann, J.: Towards a more representative parametrisation of hydrological models via synthesizing the strengths of particle swarm optimisation and robust parameter estimation, *Hydrol. Earth Syst. Sci. Discuss.*, 8, 2373–2422, doi:10.5194/hessd-8-2373-2011, 2011b. 3694, 3696, 3701, 3706, 3710
- Krzysztofowicz, R.: Transformation and normalization of variates with specified distributions, *J. Hydrol.*, 197, 286–292, doi:10.1016/S0022-1694(96)03276-3, 1997. 3712
- Liu, R. Y., Serfling, R., and Souvaine, D. L. (Eds.): *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications*, Vol. 72 of *Series in Discrete Mathematics and Theoretical Computer Science*, American Mathematical Society, 2006. 3700
- Madsen, H.: Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, *J. Hydrol.*, 235, 276–288, doi:10.1016/S0022-1694(00)00279-1, 2000. 3699

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Norbiato, D., Borga, M., Esposti, S. D., Gaume, E., and Anquetin, S.: Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins, *J. Hydrol.*, 362, 274–290, doi:10.1016/j.jhydrol.2008.08.023, 2008. 3715
- Schott, J.: Fault tolerant design using single and multicriteria genetic algorithm optimization, Master's thesis, Dep. of Aeronaut. and Astronaut., Mass. Inst. of Technol., Cambridge, Mass., 1995. 3707
- Schulla, J.: Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaveränderungen, Ph.D. thesis, Eidgenössische Technische Hochschule Zürich, diss. ETH Nr. 12018, 1997. 3709
- Seifert, M.: Untersuchungen zum Einfluss unsicherer Bodeninformationen im Schweizer Testeinzugsgebiet Rietholzbach, Master's thesis, Institut für Hydrologie und Meteorologie, Technische Universität Dresden, 2010. 3710
- Settles, M. and Soule, T.: Breeding swarms: a ga/pso hybrid, in: GECCO '05: Proceedings of the 2005 conference on Genetic and evolutionary computation, ACM Press, 161–168, 2005. 3704
- Thyer, M., Frost, A. J., and Kuczera, G.: Parameter estimation and model identification for stochastic models of annual hydrological data: Is the observed record long enough?, *J. Hydrol.*, 330, 313–328, available at: <http://www.sciencedirect.com/science/article/B6V6C-4K428WH-2/2/899aa9b003ba9b596314c0e934b79334>, 2006. 3700
- Thyer, M., Renard, B., Kavetski, D., Kuczera, G., and Srikanthan, S.: Bayesian Total Error Analysis For Hydrological Models: Preliminary Evaluation Using Multi-Site Catchment Rainfall Data, in: MODSIM 2007 International Congress on Modelling and Simulation, edited by: Oxley, L. and Kulasiri, D., Modelling and Simulation Society of Australia and New Zealand, ISBN: 978-0-9758400-4-7, 2007. 3698
- Tsou, C.-S.: Multi-objective inventory planning using MOPSO and TOPSIS, *Expert Systems with Applications*, 35, 136–142, doi:10.1016/j.eswa.2007.06.009, 2008. 3699, 3703
- Tukey, J. W.: Mathematics and the picturing of data, in: Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), *Canad. Math. Congress*, Montreal, Que., 2, 523–531, 1975. 3705
- van Veldhuizen, D. A. and Lamont, G.: Multiobjective evolutionary algorithm research: A history and analysis, Tech. Rep. TR-98-03, Dep. of Electr. Comput. Eng., Grad. Sch. of Eng., Air Force Inst. of Technol., Wright-Patterson Air Force Base, Ohio, 1998. 3707
- Vrugt, J. A. and Robinson, B. A.: Improved evolutionary optimization from genetically adaptive

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- multimethod search, in: Proceedings of the National Academy of Sciences of the United States of America, available at: www.pnas.org/cgi/content/full/06104711104/DC1 (last access: 15 March 2011), 2007. 3709, 3717
- 5 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, 2003a. 3707, 3708
- Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, 39, 1201, doi:10.1029/2002Wr001642, 2003b. 3698, 3716
- 10 Wagener, T. and Gupta, H.: Model identification for hydrological forecasting under uncertainty, *Stoch. Env. Res. Risk A.*, 19, 378–387, doi:10.1007/s00477-005-0006-5, 10.1007/s00477-005-0006-5, 2005. 3700
- Wagener, T. and Wheeler, H. S.: A Generic Framework for the Identification of Parsimonious Rainfall-Runoff Models, in: *Integrated Assessment and Decision Support (Proc. First Biennial Meeting of the International Environmental Modelling and Software Society)*, edited by: Rizzoli, A. E. and Jakeman, A. J., 1, 434–439, 2002. 3700
- 15 Wolpert, D. H.; Macready, W.: No free lunch theorems for optimization, *IEEE T. Evolut. Comput.*, 1, 67–82, doi:10.1109/4235.585893, 1997. 3709
- Wösten, J. H. M., Lilly, A., Nemes, A., and Le Bas, C.: Development and use of a database of hydraulic properties of European soils, *Geoderma*, 90, 169–185, available at: <http://www.sciencedirect.com/science/article/B6V67-3WWRF2R-C/2/50af110f999083a308c55eedc997ed4c>, 1999. 3723
- 20 Zappa, M.: Multiple-Response Verification of a Distributed Hydrological Model at Different Spatial Scales, Ph.D. thesis, Swiss Federal Institute of Technology (ETH) Zürich, 2002. 3709
- 25 Zitzler, E., Knowles, J., and Thiele, L.: Quality Assessment of Pareto Set Approximations, in: *Multiobjective Optimization: Interactive and Evolutionary Approaches*, edited by: Branke, J., Deb, K., Miettinen, K., and Slowinski, R., 373–404, Springer, 2008. 3699

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Robust
multi-objective
calibration strategies**

T. Krauß et al.

Table 1. Performance metrics for the non-dominated fronts estimated by several evolutionary multi-objective optimisation algorithms edited according to Gill et al. (2006).

Algorithm	GD	SP
MO-PSO-GA	0.0006	0.0621
MOPSO; Gill et al. (2006)	0.0122	0.1415
MOPSO; Coello et al. (2004)	0.0365	0.1095
NSGA-II	0.0842	0.0985
Micro-GA	0.1508	0.315
PAES	0.1932	0.1101

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 2. Overview of the used model parameters considered for calibration; the reference parameter vector θ_{wb} was estimated in order to use WaSiM for water-balance simulations in the Rietholzbach catchment; the parameterisation of the soil hydraulic parameters is done for each soil according to the pedotransfer functions provided in Wösten et al. (1999) and Brakensiek et al. (1984).

Parameter	Unit	Reference (θ_{wb})	Lower/Upper boundary		Description
k_d	[h]	7	0.01	25	storage coefficient of direct runoff
k_i	[h]	20	1	60	storage coefficient of interflow
dr	[–]	2.1	1	80	drainage density
k_{rec}	[–]	0.1	0.01	1	gradient of k_s with increasing depth
$\beta_{SL} = \left\{ \begin{array}{l} k_s \\ \Theta_s \\ \Theta_r \\ \alpha \\ n \end{array} \right.$	k_s [ms ^{−1}]	$2.22e-5$	$3.19e-6$	$1.32e-4$	saturated hydraulic conductivity
	Θ_s [–]	0.412	0.38	0.46	saturation water content
	Θ_r [–]	0.01	0.01	0.01	residual water content
	α [–]	4.60	1.62	8.97	empirical shape parameter (MVG)
	n [–]	1.29	1.18	1.45	empirical shape parameter (MVG)
$\beta_{SiL} = \left\{ \begin{array}{l} k_s \\ \Theta_s \\ \Theta_r \\ \alpha \\ n \end{array} \right.$	k_s [ms ^{−1}]	$7.12e-7$	$1.07e-7$	$2.95e-6$	saturated hydraulic conductivity
	Θ_s [–]	0.42	0.41	0.46	saturation water content
	Θ_r [–]	0.01	0.01	0.01	residual water content
	α [–]	1.36	0.67	2.17	empirical shape parameter (MVG)
	n [–]	1.26	1.14	1.46	empirical shape parameter (MVG)

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Robust multi-objective calibration strategies

T. Krauß et al.

Table 3. Objective functions used in this study, where q_{x_i} , $q_{y_i}(\theta)$, and Θ_{x_i} , $\Theta_{y_i}(\theta)$ are the observed and simulated discharge and mean soil moisture at time-step i , respectively. The simulated values are computed by the parameter vector θ . n denotes the number of observation points

Name	Description	Formula
NS	Nash-Sutcliffe efficiency	$1 - \frac{\frac{1}{n} \sum_{i=1}^n (q_{x_i} - q_{y_i}(\theta))^2}{\frac{1}{n} \sum_{i=1}^n (q_{x_i} - \bar{q}_x)^2}$
rPD	rel. peak flow deviation	$\frac{q_{x_{\max}} - q_{y_{\max}}(\theta)}{q_{x_{\max}}}$
FloodSkill	aggregate between NS and rPD	$0.5(-(\text{NS} - 1)) + 0.5(\text{rPD})$
r_{Θ}	moisture correlation coefficient	$\frac{\sum_{i=1}^n (\Theta_{x_i} - \bar{\Theta}_x) \cdot (\Theta_{y_i} - \bar{\Theta}_y)}{\sqrt{(\sum_{i=1}^n (\Theta_{x_i} - \bar{\Theta}_x)^2 \cdot \sum_{i=1}^n (\Theta_{y_i} - \bar{\Theta}_y)^2)}}$

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


**Robust
multi-objective
calibration strategies**

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

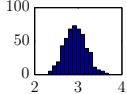
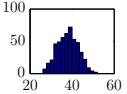
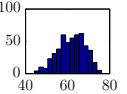
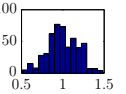
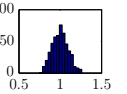
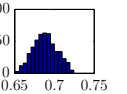
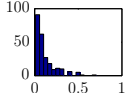
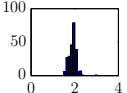
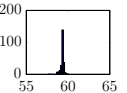
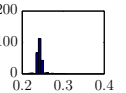
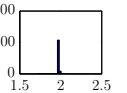
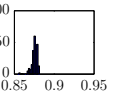
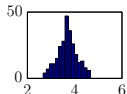
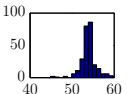
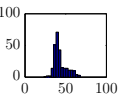
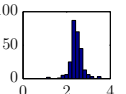
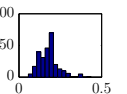
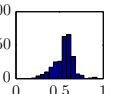
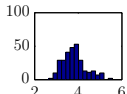
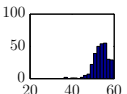
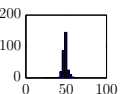
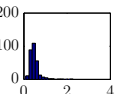
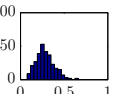
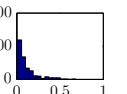
Full Screen / Esc

Printer-friendly Version

Interactive Discussion



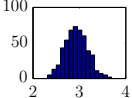
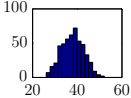
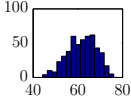
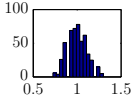
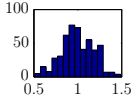
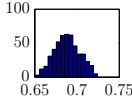
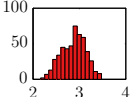
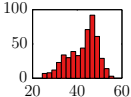
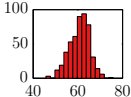
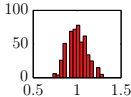
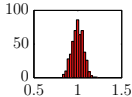
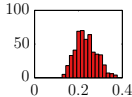
Table 4. Distribution of the robust parameter vectors of the WaSiM model estimated by MO-ROPE and the three different single-objective ROPE_{PSO} runs.

	k_d	k_i	dr	β_{SL}	β_{SiL}	k_{rec}
Initial range	0.01–25	1–60	1–80	0.06–16	0.04–9.5	0.01–1
MO-ROPE						
μ	2.92	38.64	62.74	1.00	1.00	0.69
σ	0.59	12.73	15.96	0.43	0.24	0.04
ROPE _{PSO} rPD						
μ	0.12	1.90	59.31	0.24	1.97	0.88
σ	0.12	0.17	0.22	0.01	0.02	0.00
ROPE _{PSO} FloodSkill						
μ	3.69	53.87	43.07	2.44	0.18	0.55
σ	0.41	1.85	7.88	0.29	0.06	0.11
ROPE _{PSO} rPD						
μ	3.82	53.62	48.51	0.49	0.28	0.11
σ	0.54	3.56	3.40	0.23	0.10	0.12

**Robust
multi-objective
calibration strategies**

T. Krauße et al.

Table 5. Distribution of the estimated robust parameter vectors of the WaSiM model w.r.t. 2 and 3 objectives, respectively.

	k_d	k_i	dr	β_{SL}	β_{SiL}	k_{rec}
2-objective						
μ	2.92	38.64	62.74	1.00	1.00	0.69
σ	0.59	12.73	15.96	0.43	0.24	0.04
3-objective						
μ	2.77	38.43	61.12	0.99	1.00	0.25
σ	0.56	12.69	11.02	0.25	0.18	0.13

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



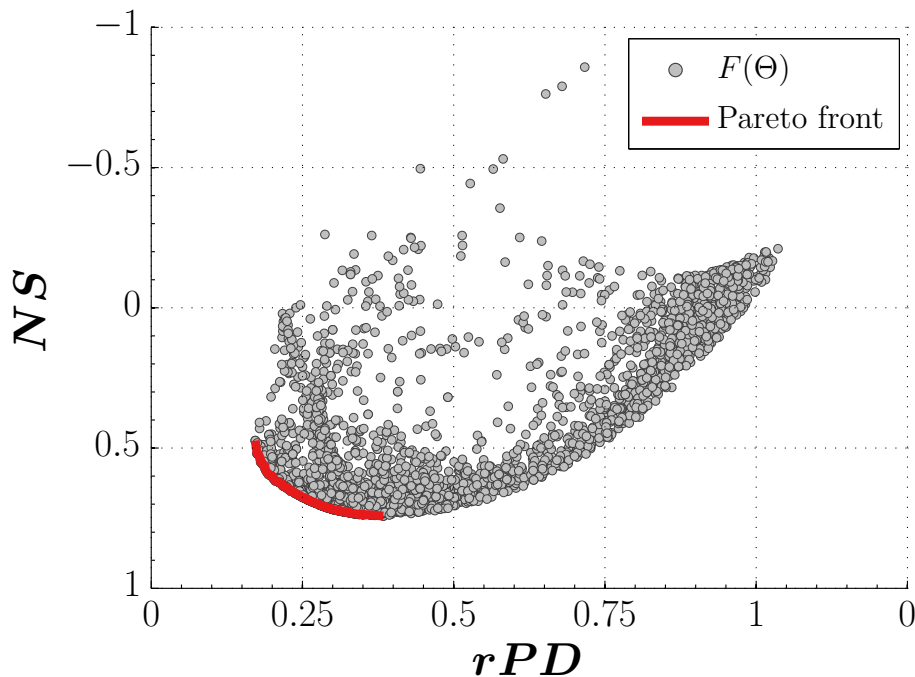


Fig. 1. Example of a bi-objective space (rPD, NS), spanned by the model performances computed for feasible parameter vectors of a hydrological model and the corresponding Pareto front. Consider that the objective rPD has to be minimised, whereas the NS has to be maximised.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



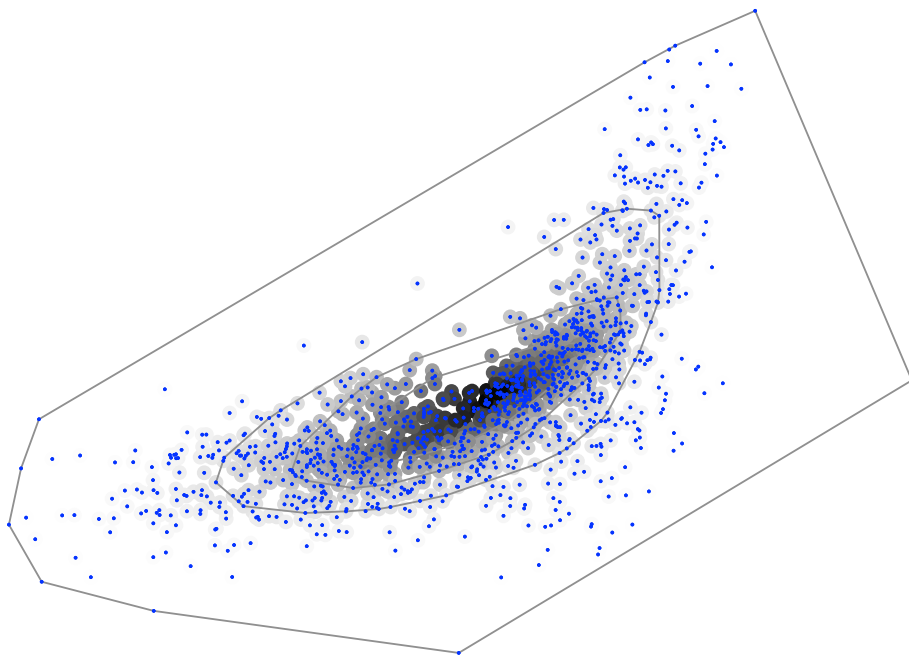


Fig. 2. 2-dimensional point set shaded according to assigned depth. A darker point represents higher depth. The lines indicate convex hulls enclosing the 25%, 50%, 75% and 100% deepest points. The used depth function was halfspace depth.

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



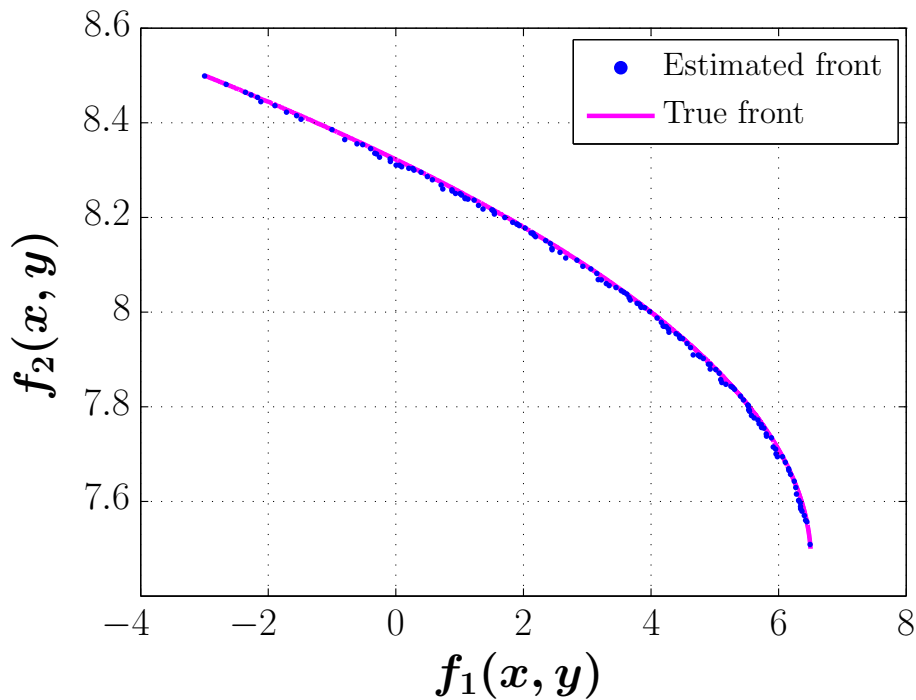


Fig. 3. True Pareto front and the non-dominated front estimated by MO-PSO-GA for *test function I*.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

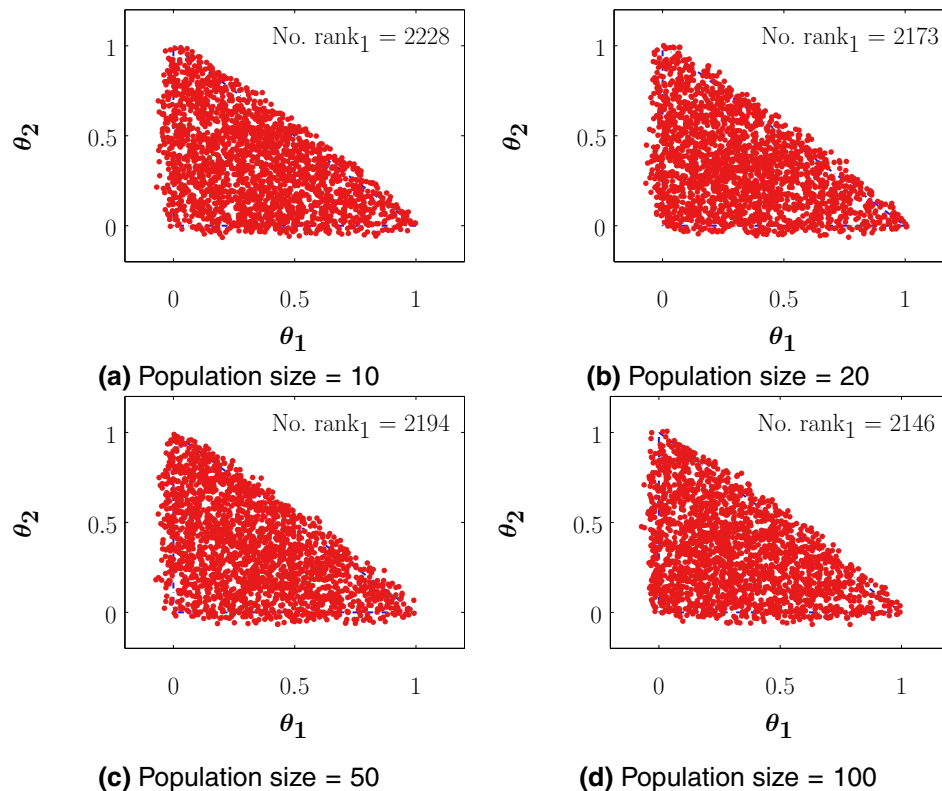


Fig. 4. Scatterplots for the rank 1 points (red dots) for 5000 function evaluations with a population size of (a) 10, (b) 20, (c) 50, and (d) 100 respectively. The boundaries of the triangle enclosing the true front are the dotdashed blue lines.

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

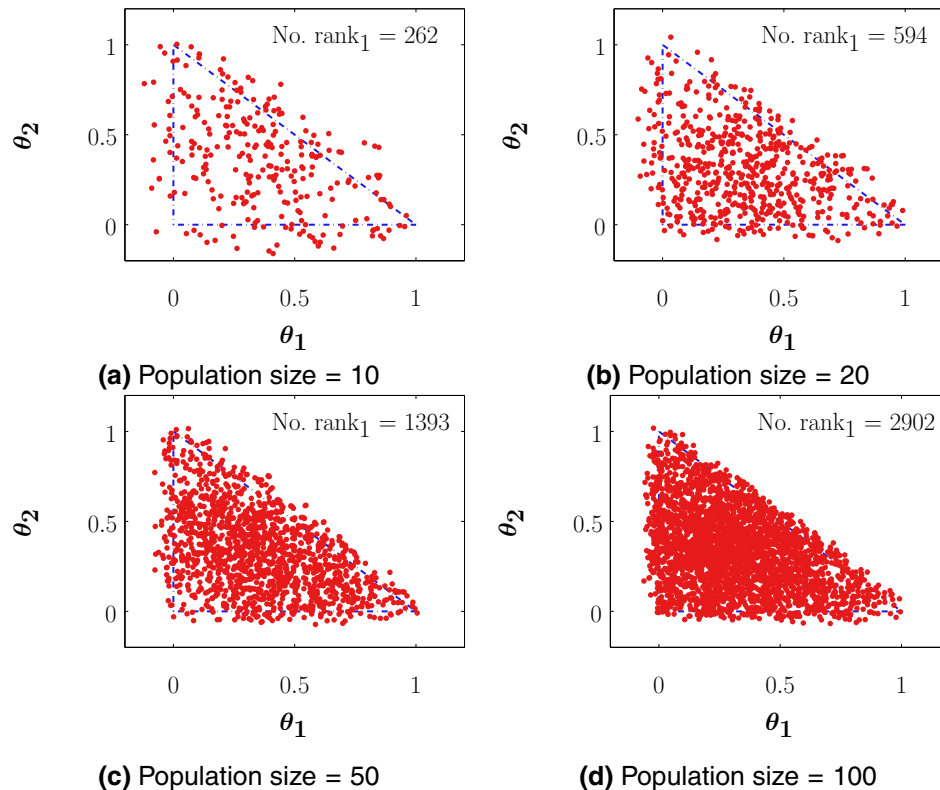


Fig. 5. Scatterplots for the rank 1 points (blue dots) for 50 evaluation generations with a population size of **(a)** 10, **(b)** 20, **(c)** 50, and **(d)** 100 respectively. The boundaries of the triangle enclosing the true front are the blue dotted lines.

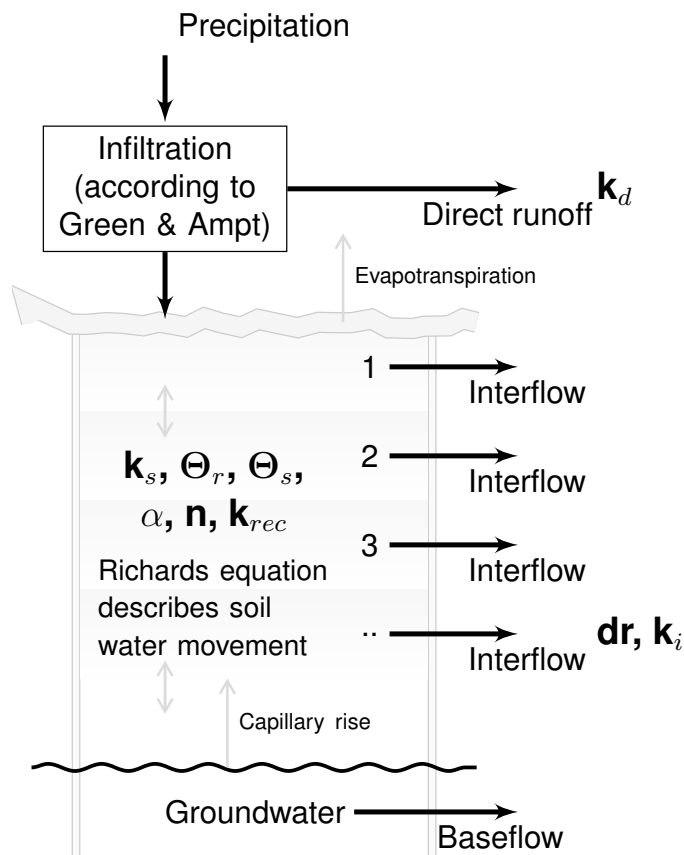


Fig. 6. Scheme of the WaSiM soil module with location of impact of soil hydraulic and conceptual model parameters (bold).

**Robust
multi-objective
calibration strategies**

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

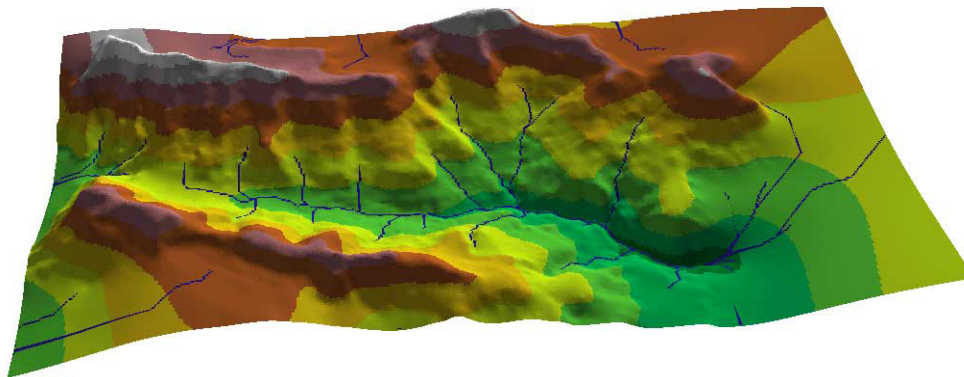


Fig. 7. A 3-D-view of the Rietholzbach catchment with the potential rivers flowpaths.

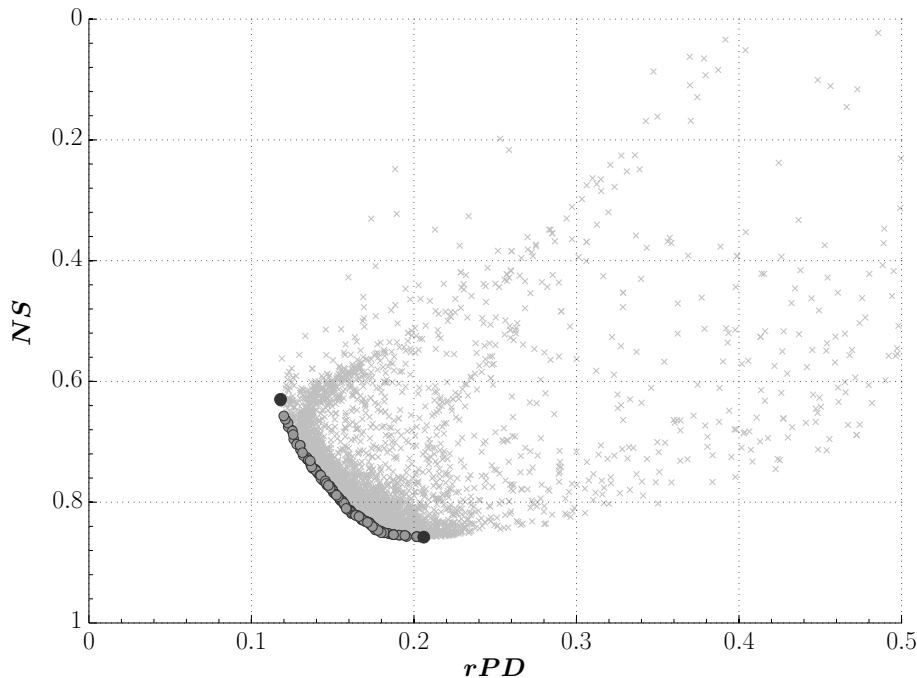


Fig. 8. Evolution of the trade-off surface of the WaSiM parameter vectors in the two-dimensional objective space. The grey dots represent the identified approximation of the Pareto optimal set $\tilde{A}$, whereas the gray crosses represent other dominated parameter vectors that were evaluated by the algorithm. The best single-criteria solutions (NS and rPD respectively) are indicated by black dots.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



T. Krauße et al.

Interactive Discussion



Fig. 9. Comparison of the identified approximation of the Pareto optimal set $\tilde{A}$ (gray dots) and the subsequently sampled deep parameter vectors $\tilde{A}_d$ (red dots in the objective space; additionally the sampled parameter vectors with low depth, i.e. the hull $\tilde{A}_h$, are indicated with blue dots; the scatter plot on the left shows the calibration objective space, the one on the right the validation objective space).

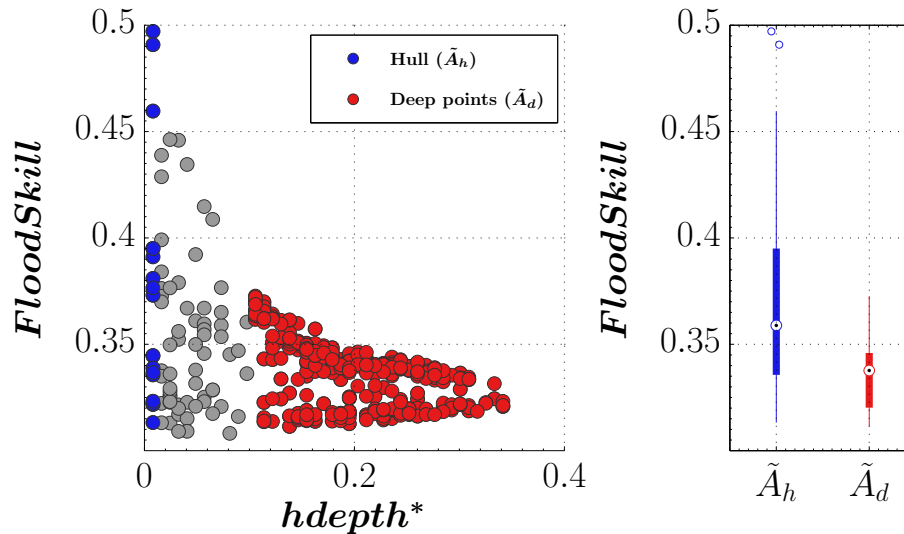


Fig. 10. Correlation between data depth and overall validation performance in terms of the FloodSkill criterion of all robust parameter vectors $\tilde{A}_d$ estimated during the 2-objective calibration. The blue dots indicate the hull, i.e. the parameter vectors with low data depth, the red dots indicate the ones with high data depth, and the grey ones are members of the Pareto set that are neither deep nor in the hull.

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



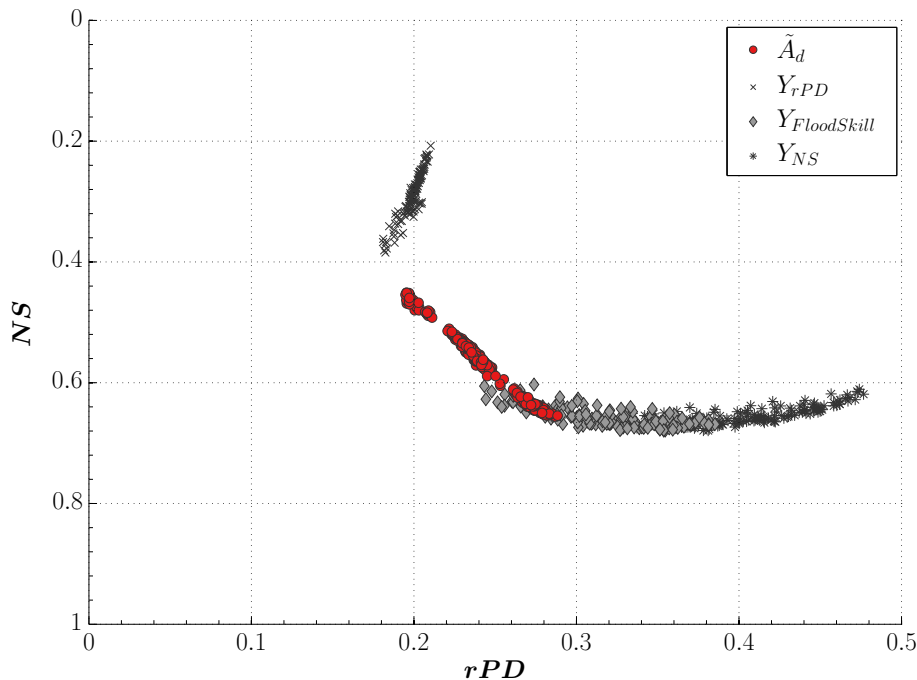


Fig. 11. Comparison of the robust model parameter vectors estimated by the multi-objective MO-ROPE ($\tilde{A}_d$) and the single-objective ROPE_{PSO} (Y_{rPD} , $Y_{FloodSkill}$, and Y_{NS} respectively) in the objective space based on the validation events.

Robust multi-objective calibration strategies

T. Krauß et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



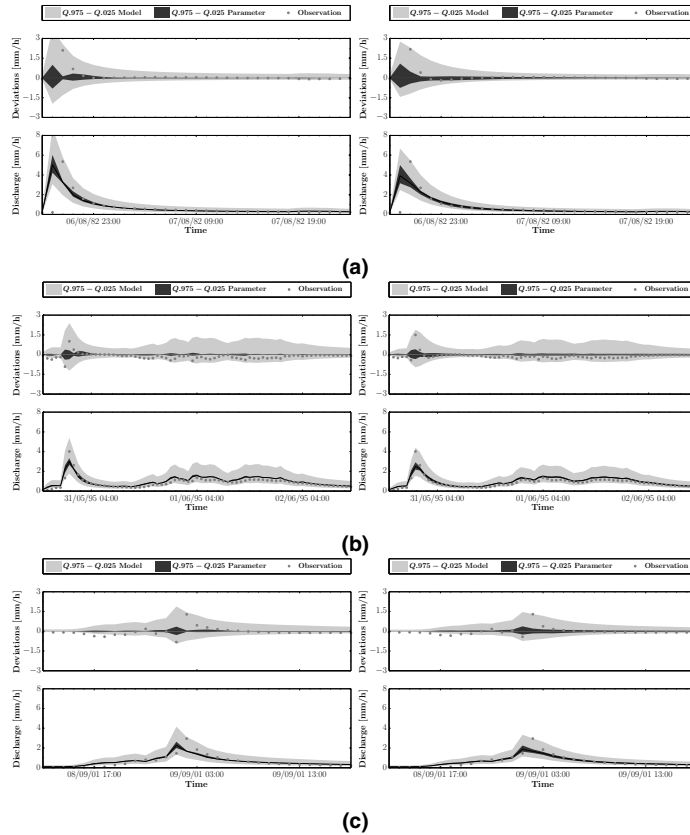


Fig. 12. Hydrograph prediction uncertainty associated with the uncertainty in the model (lighter shading) and parameter estimates (darker shading) for the flood events 6 **(a)**, 8 **(b)** and 13 **(c)**, estimated by MO-ROPE (left column) in terms of rPD and NS and ROPE_{PSO} in terms of the FloodSkill (right column). The dots correspond to the observed streamflow data. The shaded areas of uncertainty correspond to the 95% confidence intervals.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



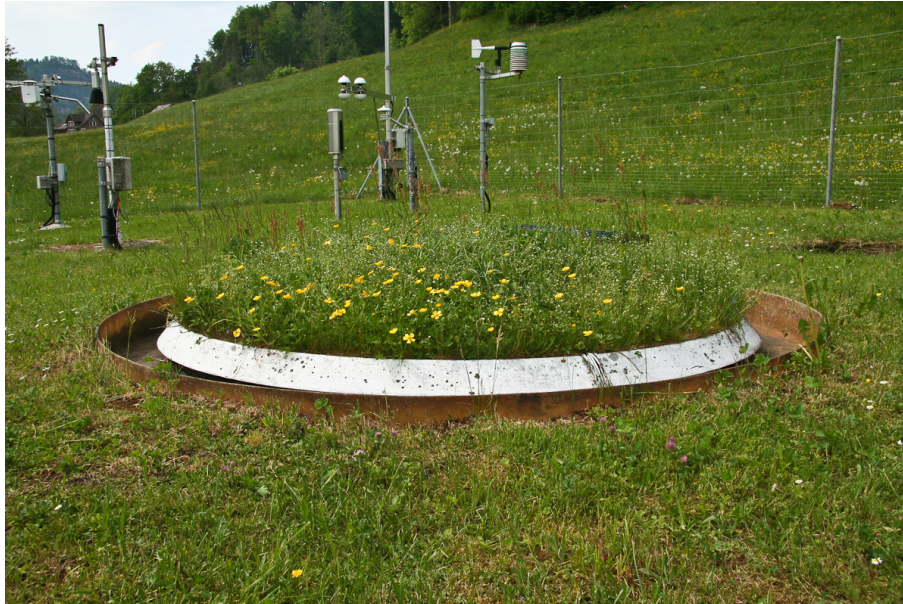


Fig. 13. Lysimeter station located in the centre of the Rietholzbach catchment.

HESSD

8, 3693–3741, 2011

Robust multi-objective calibration strategies

T. Krauße et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Robust
multi-objective
calibration strategies**

T. Krauß et al.

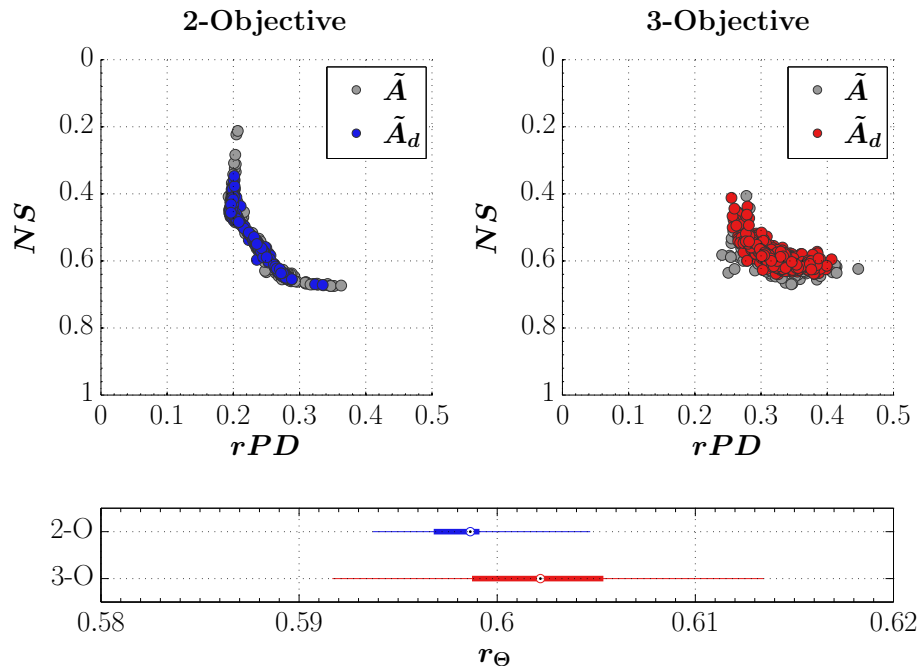


Fig. 14. Comparison of the Pareto optimal sets $\tilde{A}$ and the subsequently sampled set of deep parameter vectors $\tilde{A}_d$ for the 2-objective (left) and 3-objective (right) calibration run in the objective space based on the validation events; below is a boxplot of the corresponding soil moisture correlation r_Θ for the sets $\tilde{A}_d$ each.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Robust
multi-objective
calibration strategies**

T. Krauß et al.

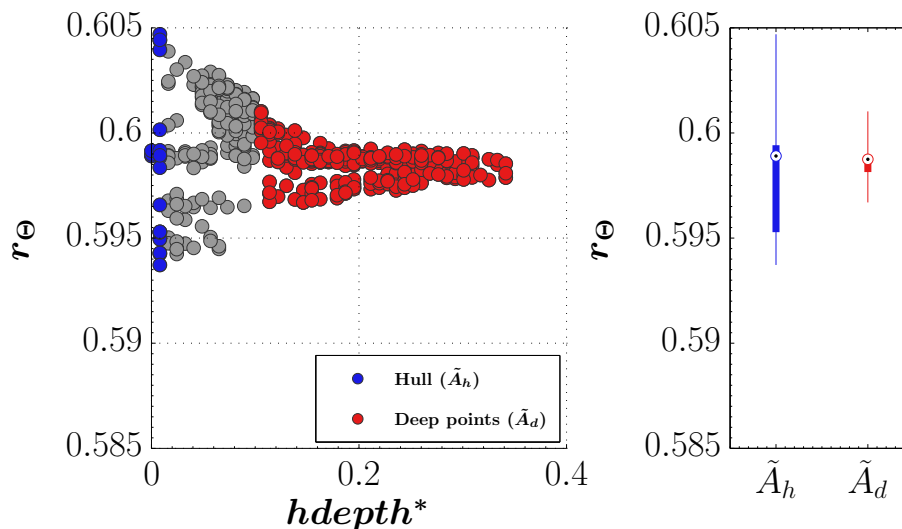


Fig. 15. Correlation between data depth and overall validation performance in terms of the soil moisture correlation r_{Θ} of all robust parameter vectors $\tilde{A}_d$ estimated during the 2-objective calibration. The blue dots indicate the hull, i.e. the parameter vectors with low data depth and the red dots indicate the ones with high data depth.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

