Hydrol. Earth Syst. Sci. Discuss., 8, 2373–2422, 2011
www.hydrol-earth-syst-sci-discuss.net/8/2373/2011/
doi:10.5194/hessd-8-2373-2011

Hydrology and Earth System Sciences Discussions

Towards a more representative parametrisation of hydrological models via synthesizing the strengths of particle swarm optimisation and robust parameter estimation

T. Krauße[1] **and J. Cullmann**[2]

[1]Institute of Hydrology and Meteorology, Department of Hydrology, Dresden University of Technology, Bergstr. 66, 01062 Dresden, Germany
[2]2IHP/HWRP Secretariat, Federal Institute of Hydrology Am Mainzer Tor 1, 56068 Koblenz, Germany

Received: 20 December 2010 – Accepted: 15 February 2011 – Published: 7 March 2011

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

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

Abstract

The development of methods for estimating the parameters of hydrological models considering uncertainties has been of high interest in hydrological research over the last years. In particular methods which understand the estimation of hydrological model parameters as a geometric search of a set of robust performing parameter vectors by application of the concept of data depth found growing research interest. Bárdossy and Singh (2008) presented a first proposal and applied it for the calibration of a conceptual rainfall-runoff model with daily time step. Krauße and Cullmann (2011) further developed this method and applied it in a case study to calibrate a process oriented hydrological model with hourly time step focussing on flood events in a fast responding catchment. The results of both studies showed the potential of the application of the principle of data depth. However, also the weak point of the presented approach got obvious. The algorithm identifies a set of model parameter vectors with high model performance and subsequently generates a set of parameter vectors with high data depth with respect to the first set. These both steps are repeated iteratively until a stopping criterion is met. In the first step the estimation of the good parameter vectors is based on the Monte Carlo method. The major shortcoming of this method is that it is strongly dependent on a high number of samples exponentially growing with the dimensionality of the problem. In this paper we present another robust parameter estimation strategy which applies an approved search strategy for high-dimensional parameter spaces, the particle swarm optimisation in order to identify a set of good parameter vectors with given uncertainty bounds. The generation of deep parameters is according to Krauße and Cullmann (2011). The method was compared to the Monte Carlo based robust parameter estimation algorithm on the example of a case study in Krauße and Cullmann (2011) to calibrate the process-oriented distributed hydrological model focussing for flood forecasting in a small catchment characterised by extreme process dynamics. In a second case study the comparison is repeated on a problem with higher dimensionality considering further parameters of the soil module.

1 Introduction

Hydrological models are designed to approximate the general physical mechanism which govern the rainfall-runoff process within a specific catchment. This is why these models have found favour with many hydrologists and engineers in practice and research. Most of the hydrologic models are driven by a vector of model parameters. Those parameter are supposed to be estimated in order to approximate the general system behaviour which govern the rainfall-runoff process within a specific catchment. In most cases the model parameters cannot be related to measurements in a direct way, but are supposed to be estimated through indirect methods such as calibration. In the process of calibration, the modeller adjusts the values of the model parameters such that the model is able to closely match the behavior of the real system it is intended to represent. Hence the success of a model application is strongly dependent on a good estimation of the model parameters.

In the past models were calibrated by hand. This is very labour-intensive and requires an experienced modeller with profound hydrological knowledge. Thus recently automatic methods for model calibration have evolved significantly (e.g. Duan et al., 1992; Gupta et al., 1998; Vrugt et al., 2003; Kuczera et al., 2006) and have found a common acceptance and broad use in the hydrological community (e.g. Hogue et al., 2000; Cullmann, 2006; Kunstmann et al., 2006; Marx, 2007; Grundmann, 2010). The parameter estimation of hydrological models is affected by numerous uncertainties. Beven and Binley (1992) described the probability to estimate the same model performance for different estimated parameter vectors as the equifinality problem. Recently developed approaches address this problem by estimating the uncertainty of the model parameter vectors considering uncertainties in the observations and the model structure. The uncertainty is often expressed by providing a set of optimal parameter vectors. Besides the meanwhile established Markov Chain Monte Carlo (MCMC) methods (e.g. Vrugt et al., 2003; Kuczera et al., 2006) the robust parameter estimation approach (ROPE) has recently attracted rising scientific interest (see Bárdossy and

Singh, 2008; Krauße and Cullmann, 2011). In this approach the parameter estimation is based upon application of the principle of data depth in order to sample robust model parameter vectors. Data depth is a statistical method used for multivariate data analysis which assigns a numeric value to a point with respect to a set of points based on its centrality. This approach provides center-outward orderings of points in Euclidean space of any dimension and provides the possibility of a new non-parametric multivariate statistical analysis in which no distributional assumptions are needed. Recent studies of computational geometry and multivariate statistics (e.g. Liu et al., 2006; Bremner et al., 2008) showed that members geometrically deep within a set, are more robust in order to represent the whole set. These points can be estimated by the concept of data depth, which has recently attracted a lot of research interest in multivariate statistics and robust modelling (e.g. Cramer, 2003; Liu et al., 2006). The outcome of recent studies (see Bárdossy and Singh, 2008; Krauße and Cullmann, 2011) showed that this concept can also be very useful for the estimation of robust hydrological model parameters. We call parameter vectors robust which

1. lead to good model performance over the selected time period,
2. lead to a hydrologically reasonable representation of the corresponding processes,
3. are not sensitive: small changes of the parameters should not lead to very different results,
4. are transferable: they perform well for other time periods and might also perform well on other catchments (i.e. they can be regionalised).

In a simplified form the ROPE approach consists of two steps. In a first step a set of model parameters with good model performance is identified. According to Bárdossy and Singh (2008) these parameter vectors are from now on called the good parameter vectors. Thereafter a set of parameter vectors 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.

The ROPE approaches proposed by Bárdossy and Singh (2008); Krauße and Cullmann (2011) identify a set of good parameter vectors by iterative application of the Monte Carlo method and the sampling of deep parameter vectors. Hence those kind of approaches also suffer from the main disadvantage of the Monte Carlo method, i.e. that it is slow. That means that many samples may be required to identify a set of good parameter vectors with acceptable precision. That is in particular a problem for computationally intensive process-oriented models where the number of samples is strictly limited by the available computational capacity. However, the effectiveness of the depth based sampling of parameter vectors is highly dependent on the quality of the identified set of good parameter vectors (see Krauße and Cullmann, 2011).

To overcome the shortcomings of the Monte Carlo method in terms of parameter estimation, many evolutionary search strategies for high-dimensional parameter spaces have been developed. One approved method is the particle swarm optimisation (PSO) which bases upon the concept of swarm intelligence. PSO optimizes a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search-space according to simple mathematical formulae. The movements of the particles are guided by the best found positions in the search-space which are updated as better positions are found by the particles. We modified the particle swarm optimization in order to identify a set of good parameter vectors with given tolerance. Afterwards the second step of the ROPE procedure, the depth based parameter sampling can be applied. The new approach is entitled Robust Parameter Estimation with Particle Swarm Optimisation (ROPE_{PSO}).

This paper is organised as follows: after the introduction, the case study area and the hydrological model are introduced. Section 2.1 briefly discusses an approach first presented by Grundmann (2010) which allows considering the uncertainty in soil hydraulic parameters for the calibration of hydrological models. Section 2.2 gives a brief explanation of particle swarm algorithms and presents the new ROPE_{PSO} approach

merging the strength of PSO and depth based sampling. Afterwards the approach is applied for the calibration of the hydrological model WaSiM-ETH focussing on flood events. The results are compared with the results of the Monte Carlo based robust parameter estimation approach AROPE_{MC}, published in Krauße and Cullmann (2011). In a further case study the performance of both algorithms on calibration problems with higher dimensionality is studied. In this study the number of model parameters considered for calibration is increased by taking soil hydraulic parameters into account.

2 Materials and methods

In this section, we will introduce some essential concepts used in this paper and afterwards we present the developed single-objective robust parameter estimation approach ROPE_{PSO}, and describe the algorithmic steps. ROPE_{PSO} evolves previous robust parameter estimation algorithms by means of performance and efficiency.

2.1 Accounting for uncertain soil information on hydrological parameter estimation

The soil hydraulic parameters determine the water retention and conductivity curves and thus govern the process of water movement in the unsaturated zone. For this reason they also influence the generation of direct runoff and interflow in a hydrological model. In many studies the soil hydraulic parameters are considered as physically based parameters and are used as fixed values. Often those values are simply estimated by applying a pedotransfer function to physical soil properties, e.g. the distribution of the grain-size fractions, humus content and bulk density. Typically the soil information is given in a classified form which provides a possible range of the physical soil properties referring to the used classification system. This information is often visualized in a soil texture triangle. However, in most cases the pedotransfer function is just applied to the mean value for the considered soil type. The uncertainty due to classified soil information is typically not considered. However, neglecting this uncertainty can influence the accuracy and uncertainty of the estimation of other conceptual model parameters.

Grundmann (2010) studied this problem and presented a method to account for those uncertainties. In principle the proposed approach is independent of the used soil classification system and the used pedotransfer function. The following procedure is suggested:

1. Identify the lower and upper boundaries of the grain-size fractions for each predominant soil type in the catchment, according to the given soil information and classification system.
2. For each considered soil type, draw a set of possible samples of the grain-size fractions by uniformed sampling (uniform distribution) over the identified range.
3. Apply a suitable pedotransfer function to each sample in order to estimate a set of soil hydraulic parameters describing their prior distribution.
4. The estimated parameters can be scaled to a scaling parameter β using a similar media concept in order to reduce their dimensionality. A suitable approach is presented in Warrick et al. (1977).

The estimated distribution can be used to study the influence of the uncertain soil information on the simulation results of hydrological models. Furthermore this information can be used as a well-founded prior distribution for a subsequent model calibration considering the soil hydraulic parameters. In a further case study (Grundmann, 2010) the estimated prior distribution was used to account for the uncertainty of the soil hydraulic parameters in the context of a Bayesian framework. For further reading and all details of the proposed approach refer to Grundmann (2010).

2.2 Merging the strengths of swarm intelligence and depth based parameter sampling

Particle swarm optimisation

An approved search and optimisation strategy for high-dimensional parameter spaces is the particle swarm optimisation (PSO) which was first presented by Kennedy and Eberhart (1995). It is a population based search algorithm which tries to solve an optimisation problem with arbitrary dimensionality by having a population (swarm) of candidate solutions, here called particles. The performance of each particle is computed and afterwards these particles are moved around in the search-space. The movement is guided by the best found positions of each of the particles and the currently found global best solution. Often those optimisation problems are formulated as the problem of finding the minimum of a function. Algorithm 2.1 gives a simple version of a PSO algorithm for the minimisation of a function f with upper and lower boundaries $\boldsymbol{x}_{\text{lb}}$ and $\boldsymbol{x}_{\text{ub}}$ respectively. A good introduction into the ideas of swarm intelligence and further reading is given in Kennedy et al. (2001)

The basis of our algorithm is a modified version of the PSO presented by Settles and Soule (2005), which is actually a hybrid between a genetic algorithm (GA) and PSO. The algorithm behaves as a normal PSO algorithm besides an additional parameter, the breeding ratio. This parameter determines the proportion of the population which are not moved according to PSO but will undergo breeding in the current generation. From the pool of possible breeding particles candidates are nominated by tournament selection and recombined. In order to do this they introduced the Velocity Propelled Averaged Crossover (VPAC) operator. The goal is to create two child particles whose position is between the parent's position, but accelerated away from the parent's current direction (negative velocity) in order to increase diversity in the population. Algorithm 2.2 shows how the new child position vectors and velocities are calculated using VPAC. The child particles retain their parent's velocity vector. The previous best vector is set to the new position vector, restarting the child's memory. Towards the end of a typical

Algorithm 2.1 Standard PSO algorithm

for all particles i **do**
  initialise position $\boldsymbol{x}_i \in \mathcal{U}[\boldsymbol{x}_{\text{lb}}, \boldsymbol{x}_{\text{ub}}]$
  and velocity $\boldsymbol{v}_i = 0$
end for
while stop criteria not met **do**
  for all particles i **do**
    set personal best $\hat{\boldsymbol{x}}_i$ as best position found so far from the particle
    set global best $\hat{\boldsymbol{g}}$ as best position found so far from the whole swarm
  end for
  for all particles i **do**
    update velocity using equation
    $\boldsymbol{v}_i(t+1) \leftarrow \varpi \boldsymbol{v}_i(t) + \phi_1 R_1(\hat{\boldsymbol{g}}(t) - \hat{\boldsymbol{x}}_i(t)) + \phi_2 R_2(\hat{\boldsymbol{x}}_i(t) - \boldsymbol{x}_i(t))$
    update position using equation
    $\boldsymbol{x}_i(t+1) \leftarrow \boldsymbol{x}_i(t) + \boldsymbol{v}_i(t+1)$
  end for
end while

PSO run, the population tends to be highly concentrated in a small portion of the search space, effectively reducing the search space. With the addition of the VPAC crossover operator, a portion of the population is always pushed away from the group, increasing the diversity of the population and the effective search space. For further details refer to Settles and Soule (2005).

Data depth

The algorithm applies a new approach used for multivariate data analysis that provides the possibility to analyze, quantify and visualize data sets. Most proposed metrics used in data depth function are inherently geometric, with a numeric value assigned to each data point that represents its centrality within the given data set. The concept of data depth is illustrated in Fig. 1 by a small 2-dimensional example. Bárdossy and Singh

Algorithm 2.2 VPAC operator

pick random numbers $\phi_1, \phi_2 \sim U(0,1)$
update positions using equations
$\boldsymbol{x}_1 \leftarrow \frac{x_1 + x_2}{2} - \phi_1 \boldsymbol{v}_1$
$\boldsymbol{x}_2 \leftarrow \frac{x_1 + x_2}{2} - \phi_2 \boldsymbol{v}_2$
reset the particles memory $\boldsymbol{p}_1 \leftarrow \boldsymbol{x}_1$ and $\boldsymbol{p}_2 \leftarrow \boldsymbol{x}_2$
update the velocities: $\boldsymbol{v}_1 \leftarrow \boldsymbol{v}_1$ and $\boldsymbol{v}_2 \leftarrow \boldsymbol{v}_2$

(2008) proposed the application of the principle of data depth in order to estimate a set of robust model parameter vectors. This idea is the basis of the developed ROPE$_{\text{PSO}}$ algorithm presented in this paper.

The following concepts apply to the data depth methodology and distinguish it from other statistical methods.

- Non-parametric methodology: scientific measurements can be viewed as sample points drawn from some unknown probability distribution, where the analysis of the measurements involves computation of quantitative characteristics of the probability distribution (estimators), based on the data set. If the underlying distribution is known (for example normal distribution, log-normal distribution, Cauchy, etc.), the characteristics of the data can be computed using methods from classical statistics. However, in most real life experiments the underlying distribution is not known. The concept of data depth requires no assumption about the underlying distribution and data is analyzed according to the relative position of the data points.

- Center-outward ordering of points: the data depth concept allows the creation of a multivariate analog to the univariate statistical analysis tool of *rank statistics*. Rank statistics is based on the ordering of one-dimensional observations, where the order reflects extremeness, contiguity, variability or the effect of external contamination. In higher dimensions the order of multivariate data is not well defined, and several ordering methods were suggested. The data depth concept provides a method of extending order statistics to any dimension by ordering the points according to their depth values.

- Application to multivariate (high-dimensional) data sets: the concept of data depth is defined with respect to points in Euclidean space in any dimension, thus enabling the derivation of multivariate distributional characteristics of a data set.

- Robustness: in the statistical analysis of data sets, observations that deviate from the main part of the data (outliers) can have an undesirable influence on the analysis of the data. Many depth functions are robust against the possibility of several outliers that may occur in the data and yield nevertheless reasonable results.

Tukey (1975) introduced this concept first with the definition of the halfspace depth. According to Donoho and Gasko (1992) the halfspace depth of an arbitrary point $\boldsymbol{\theta} \in \mathbb{R}^d$ with respect to a d-dimensional data set $Z = \{\boldsymbol{z}_i = (z_{i1}, \cdots, z_{id});\ i = 1, \cdots, n\}$ is defined as the smallest number of data points in any closed halfspace with boundary through $\boldsymbol{\theta}$. This is also called the Tukey or location depth, and it can be written as

$$\mathrm{hdepth}(\boldsymbol{\theta}|Z) := \min_{\|\boldsymbol{u}\|=1} \#\{i, \boldsymbol{u}^\top \boldsymbol{z}_i \geq \boldsymbol{u}^\top \boldsymbol{\theta}\} \tag{1}$$

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

Very often the halfspace depth is normalized by division with the number of points in the set Z:

$$\mathrm{hdepth}^*(\boldsymbol{\theta}|Z) := \frac{\mathrm{hdepth}(\boldsymbol{\theta}|Z)}{\#\{Z\}} \tag{2}$$

The first publication of Tukey (1975) was then followed by many generalizations and other definitions of this concept, e.g. convex-hull peeling depth, simplicial depth, regression depth and L1 depth. A good overview of a broad range of different definitions

of the concept of data depth and its application for multivariate data analysis is given by Hugg et al. (2006) and Liu et al. (2006). A comprehensive study of different data depth measures in robust parameter estimation is provided in Krauße (2011).

ROPE$_{\mathbf{PSO}}$

The so far presented robust parameter estimation algorithms based on depth based sampling (see Bárdossy and Singh, 2008; Krauße and Cullmann, 2011) rely on the Monte Carlo method in order to identify a set of good parameter vectors with given tolerance. Hence the proposed method suffers from the shortcomings of the Monte Carlo method a slow convergence and therefore a large number of samples are needed to estimate a stable solution. This is a major disadvantage for the calibration of computationally intensive process-oriented models. Thus in real-world application the maximum number of model runs has to be limited to a computationally feasible maximum. We try to overcome this problem by substituting the Monte Carlo based estimation of a set of good model parameter vectors with a particle swarm based approach.

The proposed algorithm, called PSO-GA$_\mathrm{u}$ is given in Algorithm 2.3. The general approach follows the breeding swarm algorithm PSO-GA, presented by Settles and Soule (2005). It is a normal PSO algorithm which produces a certain ratio ψ of the population in the next iteration by the genetic VPAC operator. ψ is called breeding rate. Contrary to the normal PSO-GA algorithm the algorithm does not just account for one global optimum, but for a set of good parameter vectors. Good parameter vectors are all points evaluated so far which have a model performance better than the global optimum plus a tolerance $\mathrm{tol_f}$ which is to be set by the user. For environmental model calibration $\mathrm{tol_f}$ should be set according to the accuracy of the used observations and other sources of uncertainty to be considered.

The algorithm accounts for a set of good parameter vectors following an idea used in some multi-objective particle swarm algorithms: all so far found good parameter vectors are stored in an archive X^*. For the movement at the end of each iteration, each particle in the population is assigned to one random member of the archive X^*

Algorithm 2.3 PSO-GA$_\text{u}$

1: initialise archive of good solutions with $\boldsymbol{X}^* \leftarrow \emptyset$
2: **for all** particles i **do**
3: initialise position $\boldsymbol{x}_i \in \mathcal{U}[\boldsymbol{x}_\text{lb}, \boldsymbol{x}_\text{ub}]$
4: and velocity $\boldsymbol{v}_i \leftarrow 0$
5: **end for**
6: **while** stop criteria not met **do**
7: **for all** particles i **do**
8: set personal best $\hat{\boldsymbol{x}}_i$ as best position found so far from the particle
9: set global best $\hat{\boldsymbol{g}}$ as best position found so far from the whole swarm
10: **end for**
11: remove all solutions with a performance worse than $f(\hat{\boldsymbol{g}}) + tol_f$ from the archive $\boldsymbol{X}^*$
12: add all current positions with a performance better than $f(\hat{\boldsymbol{g}}) + tol_f$ to the archive
13: assign to each particle a random member of the archive $\hat{\boldsymbol{x}}_\text{g} \in \boldsymbol{X}^*$ as personal global best
14: discard the worst $n_\psi = \psi \cdot \#\{\text{particles}\}$ from the population
15: initialise genetic offspring $o_\text{ga} \leftarrow \emptyset$
16: **for** $i = 1$ **to** $\frac{n_\psi}{2}$ **do**
17: select a pair $\{\boldsymbol{x}_1, \boldsymbol{x}_2\}$ from the population by tournament selection
18: apply the VPAC operator to generate new offspring $\{\boldsymbol{x}_{1_o}, \boldsymbol{x}_{2_o}\} \leftarrow vpac(\{\boldsymbol{x}_1, \boldsymbol{x}_2\})$
19: $o_\text{ga} \leftarrow o_\text{ga} \cup \{\boldsymbol{x}_{1_o}, \boldsymbol{x}_{2_o}\}$
20: **end for**
21: **for all** particles i **do**
22: update velocity using equation $\boldsymbol{v}_i(t+1) = \omega \boldsymbol{v}_i(t) + \phi_1 R_1(\hat{\boldsymbol{x}}_g(t) - \hat{\boldsymbol{x}}_i(t)) + \phi_2 R_2(\hat{\boldsymbol{x}}_i(t) - \boldsymbol{x}_i(t))$
23: update position using equation $\boldsymbol{x}_i(t+1) = \boldsymbol{x}_i(t) + \boldsymbol{v}_i(t+1)$
24: **end for**
25: merge new population with genetic offspring particles $\leftarrow$ particles $\cup\ o_\text{ga}$
26: **end while**

as personal global best. That ensures that the algorithm not just searches into the direction of the so far found global optimum but searches the whole region within the given tolerance.

Algorithm 2.4 ROPE$_\text{PSO}$

1: Execute the particle swarm based PSO-GA$_\text{u}$ procedure to estimate a set of good model parameter vectors X^* with a model performance within a given tolerance tol$_\text{f}$
2: Apply the GenDeep algorithm to sample a set of parameter vectors Y with high data depth w.r.t. X^*
3: **return** Y

The presented PSO-GA$_\text{u}$ approach can be used to substitute the Monte Carlo based approach in a robust parameter estimation algorithm. We propose a new algorithm, we call ROPE$_\text{PSO}$ which applies PSO-GA$_\text{u}$ to estimate a set of good parameter vectors X^*. Afterwards a set of deep parameter vectors with respect to X^* is sampled by the GenDeep function provided in Krauße and Cullmann (2011). A pseudocode listing of the developed ROPE$_\text{PSO}$ approach is given in Algorithm 2.4. The algorithm was implemented in a robust parameter estimation framework which comprises other published approaches. The implementation was done in the MATLAB programming language. It is open source and available from the author.

3 Case studies

3.1 Preliminary case study: comparison of ROPE$_\text{PSO}$ and AROPE$_\text{MC}$ on synthetical calibration problems

In a preliminary case study we compared the results to the Monte Carlo based approach for test problems, typically used for benchmarking of optimisation algorithms. i.e. the Rosenbrock's function (f_a) and the Rastrigin's function (f_b). The function definitions n the d-dimensional Euclidean space $\mathbb{R}^\text{d}$ are given in Eqs. (3 and 4).

$$f_a(\boldsymbol{x}) = \sum_{i=1}^{N-1} \left[(1-x_i^2)^2 + 100(x_{i+1}-x_i^2)^2\right] \tag{3}$$

$$f_b(\boldsymbol{x}) = \mathcal{A}n + \sum_{i=1}^{N} \left[x_i^2 - \mathcal{A}\cos(2\pi x_i)\right] \tag{4}$$

$\forall \boldsymbol{x} \in \mathbb{R}^n,\ \mathcal{A} = 10$

The global minima for f_a in $\mathbb{R}^d$ is the vector $\mathbf{1^d}$, for f_b it is $\mathbf{0^d}$. Those functions were already used to illustrate the Monte Carlo based robust parameter estimation algorithm AROPE$_{MC}$, presented in Krauße and Cullmann (2011). We applied both the AROPE$_{MC}$ and the ROPE$_{PSO}$ in order to minimise both test problems for $\mathbb{R}^2, \mathbb{R}^2$ and $\mathbb{R}^4$. The used tolerance tol$_f$ of the fitness values was 0.1. For both algorithms we set the maximum number of function evaluations to 10 000.

The statistics of the estimated parameter samples and corresponding fitness values are given in Table 1. It is evident that both algorithms estimate a similar set of parameter vectors that means that both sets have approximately the same statistics for the estimated set and the corresponding fitness values, as long as a stable solution can be found by both algorithms. That holds true for low-dimensional, simpler problems, e.g. f_a in $\mathbb{R}^2$ and $\mathbb{R}^3$, f_b in $\mathbb{R}^2$. For more complicated, higher-dimensional problems, e.g. f_a in $\mathbb{R}^4$, f_b in $\mathbb{R}^3$ and $\mathbb{R}^4$, both algorithms estimate a set of parameter vectors with similar deviation of the corresponding fitness values but different mean fitness. The statistics of the estimated parameter vectors can differ significantly. In those cases the particle swarm based ROPE$_{PSO}$ algorithm outperforms the Monte Carlo based AROPE$_{MC}$ due to a more efficient search of the parameter space. Furthermore ROPE$_{PSO}$ needed much less function evaluations than the particle swarm based approach, i.e. AROPE$_{MC}$ took fully advantage of the maximum number of 10 000 function evaluations whereas ROPE$_{PSO}$ achieved a stable solution after approximately 4000 up to 6000 function evaluations.

3.2 Calibrating of the hydrological model WaSiM focussing on flood events

In two real world case studies the developed ROPE$_{PSO}$ approach is tested on the calibration of the hydrological model WaSiM-ETH/6.4 (in the further referred to as WaSiM). WaSiM is a spatial distributed process-oriented rainfall-runoff model and was developed by Schulla (1997) at the ETH Zurich. An overview of the model structure is given in Fig. 2. WaSiM has been used successfully for modeling the rainfall-runoff processes in several studies in catchments located within mid mountain ranges (e.g. Grundmann, 2010) and especially also in the pre-alpine Rietholzbach catchment (Gurtz et al., 1999, 2003b,a; Krauße and Cullmann, 2011; Krauße, 2011). Furthermore WaSiM-ETH has been used for extrapolation of extreme flood events by Cullmann (2006). For this study we used the version with the approach according to Richards for the simulation of the unsaturated zone (Fig. 2). For further details of the model refer to Schulla (1997); Schulla and Jasper (2007) and the official website of the model http://www.wasim.ch.

The model was calibrated focussing on flood events in the small prealpine Rietholzbach catchment (3.18 km^2). As a sub-basin of the Thur catchment it is located in the north-east of Switzerland. A 3-D-view of the catchment area is given in Fig. 3. This basin has been observed as a research catchment by the ETH Zurich since 1975. Continuous hourly measurements have taken place since 1981. For the case studies presented in this paper we used time series consisting 27 years of meteorological (temperature, precipitation, global radiation, and wind speed) and discharge measurements. Out of this time series we selected a set of 24 flood events with a peak flow of at least 1 mm h^{-1}. We just selected events which were not affected by snow or snowmelt in order to avoid the problem of snow modelling. For further details and a more comprehensive overview refer to Krauße and Cullmann (2011). 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.

Table 2 gives the model parameters considered for calibration. Those are the storage coefficients of direct runoff and interflow, k_d and k_i, and the drainage density dr which is a scaling parameter of interflow generation. In previous studies (Cullmann, 2006; Pompe, 2009; Grundmann, 2010) these three parameters have been proven to be sensitive with respect to modelling flood events. Besides the specified upper and lower boundaries of the model parameters, the additional constraint $k_i \geq 1.05\ k_\mathrm{d}$ was introduced in order to account for the basic consideration that the direct runoff from a cell has a shorter travel time to the catchment outlet than the generated interflow in the unsaturated zone.

The used objective is an aggregated criterion between the global efficiency criterion according to Nash and Sutcliffe (1970) (NS) and the relative peakflow deviation (rPD). We call this criterion FloodSkill, because it quantifies the model's ability to provide both a good estimate of the peak flow values and a minimum reasonable representation of the catchment behavior. It is computed by

$$\mathrm{FloodSkill} \leftarrow 0.5(-(\mathrm{NS}-1)) + 0.5(\mathrm{rPD}) \tag{5}$$

The lower the FloodSkill the better is the model's ability to represent the catchment's behaviour focussing on flood events. A perfect fit corresponds to a FloodSkill of zero.

3.2.1 Case study I: comparison of ROPE$_\mathrm{PSO}$ and AROPE$_\mathrm{MC}$ considering just the conceptual model parameters

In a first case study we compared the developed approach with the Monte Carlo based robust parameter estimation AROPE$_\mathrm{MC}$ (s. Krauße and Cullmann, 2011) on the basis of the second case study presented in Krauße and Cullmann (2011). The conceptual model parameters of the model WaSiM were calibrated focussing on flood events. Out of the mentioned set of 24 flood events, three events were used for calibration, three for overfitting control and the remaining events were used for validation (see Table 4). The used objective was the proposed FloodSkill criterion. For further details refer to Krauße and Cullmann (2011).

Before deep parameters were generated we validated all good parameter vectors, estimated by the PSO-GA$_\mathrm{u}$ algorithm and studied the relationship between the data depth and the model performance on the validation data. The correlation between the optimization objective and the data depth is significantly negative (see Fig. 6) which is a first hint that the estimated set is a stable solution and the application of depth based sampling makes sense in order to improve the calibration results. The scatterplot given in Fig. 4 where each parameter vector of the estimated set is shaded according to its validation performance shows that the parameter vectors with worse model performance are particularly located at the boundary of the estimated set. However, this conclusion does not hold true for the single performance criteria the FloodSkill consists of, i.e. NS and rPD. The optimal regions for this criteria are even located on opposite sides of the estimated set (Fig. 5). This is a hint of a tradeoff between the two objective criteria rPD and problems of the model to represent the global system behaviour of the catchment and a good representation of the peak flow values. This may also be due to the relatively coarse time step considering the small catchment size. In this case a multi-objective calibration might be useful, but this is in the scope of future research.

In a second step deep parameter vectors were sampled with respect to the estimated set of good parameter vectors. The statistics of the estimated samples in comparison with those estimated by AROPE$_\mathrm{MC}$ are given in Table 3. Although both sets overlap each other, it is evident that the mean value of the storage coefficients k_d and k_i of the ROPE$_\mathrm{PSO}$ estimates are significantly lower than those estimated by AROPE$_\mathrm{MC}$. However, the correlation between the two most sensitive model parameters focussing on flood events, k_d and dr is the approximately the same for both algorithms. The standard deviation for the ROPE$_\mathrm{PSO}$ estimates is higher for all considered model parameters. We assume that the iterative Monte Carlo based sampling was not able to sample from the outermost part of the set due to small sample size (2500 samples per iteration) and a fixed boundary in the iteration steps (the 10% best parameter vectors are selected).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

The validation results of the ROPE_{PSO} estimates are slightly better than those estimated by AROPE_{MC} and also outperforms all other compared algorithms compared in the case study of Krauße and Cullmann (2011). Furthermore, regardless of the higher standard deviation of the parameter values with respect to the AROPE_{MC} estimates, the standard deviation of the validation performance is smaller. These results indicate that the estimated solution is robust and transferable. The proposed algorithm also converges faster with less parameter vectors to be evaluated. The ROPE_{PSO} algorithm evaluated 3.574 model parameter vectors during the calibration whereas the Monte Carlo based AROPE_{MC} needed 10 000 parameter evaluations. This advantage gets more even more weight considering that one parameter evaluation [1] takes approximately three minutes on a standard CPU.

3.2.2 Case study II: calibrating WaSiM for flood forecasting considering the uncertainty of the soil hydraulic parameters

In a second case study we calibrated WaSiM again, however this time we additionally considered the uncertainty in the soil hydraulic parameters in the model calibration. As already introduced in Sect. 2.1 the uncertainty due to coarse soil information and the resulting uncertainty in the soil hydraulic parameters can have a tremendous influence on the model uncertainty in the case of flood events. In a preliminary study with WaSiM in the Rietholzbach catchment (see Seifert, 2010) we could prove this conclusion. From the 5 predominant soil types in the basin (Table 4), we found the soil hydraulic parameters of SL and SiL and the soil parameter k_{rec} to be sensitive referring to the simulated discharge for flood events. Hence we considered those parameters together with the conceptual model parameters for model calibration.

According to the approach of Grundmann (2010) we estimated the prior uncertainty in the soil hydraulic parameters of the soils SL and SiL and mapped the computed set of parameter vectors to two scaling parameters β_{LS} and β_{SiL}, one for each soil.

[1] That means that 3 model runs have to be carried out, one for each calibration event.

The prior uncertainty of the soil hydraulic parameters with best fits for Gaussian ($\mathcal{N}$), logarithmic Gaussian ($\log\mathcal{N}$), Gamma (Γ) and bimodal Gaussian (GM) distribution is given in Fig. 7. Consider that the residual water content θ_{r} was constantly 0.01 for both SL and SiL with a deviation of less than 10e-14 and thus is not shown in the plots. The distribution of the saturated conductivities has the maximum density in the range of the lowest possible values and is characterised by a high skewness. The other MVG parameters have distributions which can be approximated by a normal distribution. The distribution of the corresponding scaling parameters is given in Fig. 8. It is evident that the distribution of the scaling parameters is strongly influenced by the distribution of the saturated conductivities. This is due to the relatively high spread of this parameter.

The conceptual and the soil parameters form a six dimensional calibration problem with the model parameters $\{k_{\text{d}}, k_i, dr, \beta_{\text{SL}}, \beta_{\text{SiL}}, k_{\text{rec}}\}$ to be estimated. Considering that the error surface of WaSiM is very bumpy this is already a challenging calibration problem, especially for Monte Carlo based approach. Again we estimated the parameters of WaSiM with both AROPE_{MC} and ROPE_{PSO}. Both calibration algorithms were limited to a maximum of 10 000 model parameter vector evaluations.

The distribution of the parameter vectors estimated by AROPE_{MC} and ROPE_{PSO} is given in Figs. 9 and 10 respectively. It is evident that the particle swarm based parameter estimates are distributed over a much smaller region than those estimated by AROPE_{MC}. That region forms a subset (in geometrical sense) of the large region estimated by AROPE_{MC}. Thus ROPE_{PSO} can more precisely identify a region of good model parameter vectors within a defined measurement uncertainty. The distribution of the soil hydraulic parameters corresponding to the scaling parameters β_{SL} and β_{SiL} estimated by both algorithms are given in Figs. 11 and 12 respectively.

By means of comparison of observed and simulated discharge, the parameter estimation algorithms try to reject soil hydraulic parameter vectors that are not suitable to represent the catchment's behaviour and identify a distribution with the most suitable model parameters. It is obvious that the spread of the distributions of the soil hydraulic parameters compared to their prior uncertainty gets smaller for both algorithms.

Furthermore it stands out that the mean ks of the ROPE_{PSO} estimates is smaller than the prior expectancy whereas the mean ks of the AROPE_{MC} values is higher than the prior value. In terms of the model that means that ROPE_{PSO} identifies parameter vectors that try to simulate just the slow matrix flow in the unsaturated zone whereas faster runoff processes in the unsaturated zone, e.g. preferential flow, are approximated by a fit of the conceptual model parameters controlling direct runoff. Probably that is also the reason why the less sensitive conceptual parameter k_i can by much better identified than in the previous case study. In contrast AROPE_{MC} identifies parameter sets which try to represent the faster components by a higher saturated conductivity in the soil model. Preferential flow in macropores can become a dominant process within the Rietholzbach catchment (Germann, 1981). However, the Richards equation implemented in WaSiM just describes the process of matrix flow in the unsaturated zone. The process of preferential flow cannot be directly represented by this equation. Thus this process can either be represented by in trend higher saturated conductivities used in the soil model or by representing this process by the conceptual model parameters, in particular the parameters k_{d} and k_{rec}. From a process-oriented point of view it might be better to "blame" the preferential flow on the conceptual model parameters instead of trying to describe a physically completely different process by a physically based model, i.e. trying to fit the Richards equation to represent preferential flow and matrix flow instead of just matrix flow.

The calibration performance of both algorithms is given in Fig. 13. The results of PSO-GA_{u} (before deep samples were drawn) are given for an additional comparison. The calibration performance results are better than the results for the calibration in the previous case study just considering the conceptual model parameters. That result is not surprising. The better fit in the calibration might just be due to a larger number of free model parameters and has to be confirmed on the validation data in order to be a considered as improvement. A better model fit is always possible with more free model parameters, but just makes sense if the estimated parameters can be transferred on other time periods. Although the parameter vectors estimated by AROPE_{MC} have a

significantly larger deviation, the corresponding model performances on the calibration data just slightly differ, i.e. ROPE_{PSO} achieves a mean FloodSkill of 0.26 whereas the mean FloodSkill for the AROPE_{MC} results is 0.27. Besides the better performance with respect to the population mean performance, the population estimated by ROPE_{PSO} also has less outliers with worse model performance on the calibration data. Furthermore it is evident that the deep parameter vectors estimated by ROPE_{PSO} do not have a better model performance on the calibration data.

The validation results calibrated on the validation events given in Table 4 are given in Fig. 14. Regardless of the used parameter estimation procedure, the achieved model performance on the validation data is better than the calibration with the conceptual model parameters only. Furthermore these results show the advantages of the ROPE_{PSO} approach merging the strengths of the depth based parameter sampling and the particle swarm optimisation. Referring to the results of PSO-GA_{u} and ROPE_{PSO} it is obvious that the parameters with high data depth do not have just a marginal better model performance on the validation data but also much less bad outliers, e.g. the worst overall FloodSkill for PSO-GA_{u} is 0.41 whereas it is 0.38 for the ROPE_{PSO} after the depth based sampling. Consider for instance that a change of the FloodSkill value of just 0.03 means an improvement of the Nash and Suttcliffe efficiency in mean of 0.06 considering constant ability of peak flow prediction. Nevertheless already the application of the particle swarm based parameter estimation achieves good validation results. In comparison the AROPE_{MC} estimates have a worse mean performance. That is due to high positive skewness, i.e. the best achieved validation performances are as good as those estimated by ROPE_{PSO} but there are more and worse outliers on the negative side of the distribution of the model performances. For example the worst parameter vector estimated by AROPE_{MC} has an overall FloodSkill of 0.47 in comparison to 0.35 estimated by ROPE_{PSO}. The higher spread is also expressed by the significantly higher standard deviation referring to all compared performance criteria. The better accuracy together with less uncertainty of the ROPE_{PSO} estimates is also reflected by a reduced model uncertainty. That means that not just the parameter uncertainty but also the complete model uncertainty can be tremendously reduced. Figure 15 shows the hydrographs and the corresponding parameter and model uncertainties for both algorithms. The model errors were computed by two normal distribution 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).

4 Discussion and conclusions

- This paper presents a depth based parameter estimation method, which is well suited for the robust calibration of hydrological models considering uncertainties. We presented a particle swarm based robust parameter estimation algorithm algorithm, entitled Robust Parameter Estimation with Particle Swarm Optimisation (ROPE_{PSO}). The major difference between ROPE_{PSO} and the algorithms presented by Bárdossy and Singh (2008) and Krauße and Cullmann (2011) is that it substitutes the Monte Carlo based approach for the identification of good parameter vectors by a particle swarm based approach PSO-GA_{u}. We compare the effectiveness of the newly developed algorithm for estimating robust model parameter vectors with the Monte Carlo based variant in three case studies.
- A first preliminary case study with 2 synthetical test problems shows that both algorithms estimate approximately the same set of good parameter vectors for simpler problems in lower dimensions. However, due to the more efficient search of the particle swarm based approach, the results of the ROPE_{PSO} algorithm outperform the Monte Carlo based approach.
- The results of the second case study in this paper compared AROPE_{MC} and ROPE_{PSO} estimating three conceptual model parameters of the model WaSiM using an example presented in Krauße and Cullmann (2011). ROPE_{PSO} estimates a similar set of robust model parameter vectors which slightly outperforms the estimates of the Monte Carlo based approach.

- The case studies in this paper revealed that the used hydrological model WaSiM is not able to represent the correct peak flow values and the basin's system behaviour with respect to the discharge at the catchment outlet with the same parameter vectors. 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 near the boundary and are sensitive to small changes and do transfer to other time periods less well as high depth ones. However the model performance of the sampled deep parameters is also dependent from the quality of the estimated good parameter vectors.
- In this paper, model performance was expressed by just one aggregated objective function. The presented algorithm can be easily altered to a general multi-objective parameter estimation procedure.

The robust parameter estimation approach is a relatively new method which was applied to a limited number of case studies. We strongly propose the comparison with established uncertainty estimation methods, e.g. MCMC, GLUE or multi-objective calibration, in further research. Furthermore, due to the probably high tradeoff between the model's ability to represent both the peak flow values and the global system behaviour equally well, we propose the development and application of a multi-objective version of the presented approach.

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 Rietholzbach 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, http://dx.doi.org/10.1002/hyp.3360060305, 1992. 2375

Brakensiek, D., Rawls, W., and Stephenson, G.: Modifying SCS hydrological soil groups and curve numbers for rangeland soils, ASAE Paper, 203, 1984. 2407

Bremner, D., Chen, D., Iacono, J., Langerman, S., and Morin, P.: Output-sensitive algorithms for Tukey depth and related problems, Statistics and Computing, 18, 259–266, 2008. 2376

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. 2374, 2375, 2376, 2377, 2381, 2384, 2395

Cramer, K.: Multivariate Ausreißer und Datentiefe, Ph.D. thesis, Wirtschafts- und Sozialwissenschaftliche Fakultát, Universitát zu Köln, 2003. 2376

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. 2375, 2388, 2389

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, http://www.jstor.org/stable/2242368, 1992. 2383

Duan, Q., Sorooshian, S., and Gupta, V.: Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models, Water Resour. Res., 28, 1015–1031, 1992. 2375

Engeland, K., Renard, B., Steinsland, I., and Kolberg, S.: Evaluation of statistical models for forecast errors from the HBV model, Journal of Hydrology, 384, 142–155, doi:10.1016/j.jhydrol.2010.01.018, http://www.sciencedirect.com/science/article/B6V6C-4Y95RV1-1/2/502b633b47cc1dc3aeb1143beeca7a8c, 2010. 2395

Germann, P.: Untersuchungen über den Bodenwasserhaushalt im hydrologischen Einzugsgebiet Rietholzbach, no. 51 in Mitteilungen der Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie, ETH Zürich, Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie, ETH Zürich, 1981. 2393

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. 2375, 2377, 2379, 2388, 2389, 2391, 2407

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. 2375

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. 2388

Gurtz, J., Verbunt, M., Zappa, M., Moesch, M., Pos, F., and Moser, U.: Long-term hydrometeorological measurements and model-based analyses in the hydrological research catchment Rietholzbach, J. Hydrol. Hydromech., 51, 1–13, 2003a. 2388

Gurtz, J., Zappa, M., Jasper, K., Lang, H., Verbunt, M., Badoux, A., and Vitvar, T.: A comparative study in modelling runoff and its components in two mountainous catchments, Hydrol. Process., 17, 297–311, http://dx.doi.org/10.1002/hyp.1125, 2003b. 2388

Hogue, T. S., Sorooshian, S., and Gupta, H.: A Multistep Automatic Calibration Scheme for River Forecasting Models, J. Hydrometeorol., 1, 524–542, 2000. 2375

Hugg, J., Rafalin, E., Seyboth, K., and Souvaine, D.: An Experimental Study of Old and New Depth Measures, in: Workshop on Algorithm Engineering and Experiments (ALENEX06), 51–64, 2006. 2384

Kennedy, J. and Eberhart, R.: Particle Swarm Optimization, in: Proceedings of IEEE International Conference on Neural Networks, vol. IV, 1942–1948, 1995. 2380

Kennedy, J., Eberhart, R. C., and Shi, Y.: Swarm Intelligence, Morgan Kaufmann Series in Evolutionary Computation, Morgan Kaufman, 2001. 2380

Krauße, T.: Data depth – parsimonious sampling of robust parameters in hydrological models, Appl. Math. Modelling, submitted, 2011. 2384, 2388

Krauße, T. and Cullmann, J.: Identification of hydrological model parameters for flood forecasting using data depth measures, Hydrol. Earth Syst. Sci., submitted, 2011. 2374, 2376, 2377, 2378, 2384, 2386, 2387, 2388, 2389, 2391, 2395, 2405

Krzysztofowicz, R.: Transformation and normalization of variates with specified distributions, J. Hydrol., 197, 286–292, doi:10.1016/S0022-1694(96)03276-3, http://www.sciencedirect.com/science/article/B6V6C-3T7JKM3-15/2/8920eb68d8b37686a66c044f66d0b0ad, 1997. 2395

Kuczera, G., Kavetski, D., Franks, S., and Thyer, M.: Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using stormdependent parameters, J. Hydrol., 331, 161–177, http://www.sciencedirect.com/science/article/B6V6C-4KBVX50-2/2/b066c6156e2b79860d5e07dd57111ff2, 2006. 2375

Kunstmann, H., Krause, J., and Mayr, S.: Inverse distributed hydrological modelling of Alpine catchments, Hydrol. Earth Syst. Sci., 10, 395–412, doi:10.5194/hess-10-395-2006, 2006. 2375

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. 2376, 2384

Marx, A.: Einsatz gekoppelter Modelle und Wetterradar zur Abschätzung von Niederschlagsintensitäten und zur Abflussvorhersage, vol. 160 of *Mitteilungen des Instituts für Wasserbau, Universität Stuttgart*, Institut für Wasserbau, Universität Stuttgart, http://elib.uni-stuttgart.de/opus/volltexte/2007/3016/, 2007. 2375

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, http://www.sciencedirect.com/science/article/B6V6C-487FF7C-1XH/1/75ac51a8910cad95dda46f4756e7a800, 1970. 2389

Pompe, K.: Untersuchungen zur optimalen Parametrisierung von Prozessmodellen in der Hydrologie, Diplomarbeit, Institut für Hydrologie und Meteorologie, Technische Universitüt Dresden, 2009. 2389

Schulla, J.: Hydrologische Modellierung von Flussgebieten zur Abschötzung der Folgen von Klimaänderungen, Ph.D. thesis, Eidgenössische Technische Hochschule Zürich, diss. ETH Nr. 12018, 1997. 2388

Schulla, J. and Jasper, K.: Model Description WaSiM-ETH, Eidgenösische Technische Hochschule Zürich, 2007. 2388

Seifert, M.: Untersuchungen zum Einfluss unsicherer Bodeninformationen im Schweizer Testeinzugsgebiet Rietholzbach, Master's thesis, Institut für Hydrologie und Meteorologie, Technische Universität Dresden, 2010. 2391

Settles, M. and Soule, T.: Breeding swarms: a ga/pso hybrid, in: GECCO '05: Proceedings of the 2005 conference on Genetic and evolutionary computation, 161–168, ACM Press, 2005. 2380, 2381, 2384

Tukey, J. W.: Mathematics and the picturing of data, in: Proceedings of the International Congress of Mathematicians (Vancouver, B. C., 1974), vol. 2, 523–531, Canad. Math. Congress, Montreal, Que., 1975. 2383

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(8), 1201, doi:10.1029/2002Wr001642, 2003. 2375

Warrick, A. W., Mullen, G. J., and Nielsen, D. R.: Scaling Field-Measured Soil Hydraulic Properties Using a Similar Media Concept, Water Resour. Res., 13, 355–362, doi:10.1029/WR013i002p00355, 1977. 2379

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, http://www.sciencedirect.com/science/article/B6V67-3WWRF2R-C/2/50af110f999083a308c55eedc997ed4c, 1999. 2407

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. 2388

Table 1. Mean value, standard deviation and correlation coefficients between the generated samples in the preliminary case study for the parameters estimated by $ROPE_{PSO}$ and $AROPE_{MC}$ and corresponding fitness values.

Problem		Algorithm	Parameter	Mean	Std	x_1	x_2	x_3	x_4	mean($f(x)$)	std($f(x)$)
2-D	f_a	$AROPE_{MC}$	x_1	1.00	0.06	1.00	0.99	–	–	0.050	0.041
			x_2	1.02	0.13	…	1.00	–	–		
		$ROPE_{PSO}$	x_1	1.01	0.06	1.00	0.98	–	–	0.049	0.039
			x_2	1.02	0.12	…	1.00	–	–		
	f_b	$AROPE_{MC}$	x_1	0.00	0.01	1.00	0.05	–	–	0.038	0.029
			x_2	0.00	0.01	…	1.00	–	–		
		$ROPE_{PSO}$	x_1	0.00	0.01	1.00	0.02	–	–	0.041	0.030
			x_2	0.00	0.01	…	1.00	–	–		
3-D	f_a	$AROPE_{MC}$	x_1	1.01	0.02	1.00	0.93	0.91	–	0.071	0.037
			x_2	1.01	0.04	…	1.00	0.98	–		
			x_3	1.03	0.08	…	…	1.00	–		
		$ROPE_{PSO}$	x_1	1.03	0.02	1.00	0.92	0.91	–	0.064	0.034
			x_2	1.00	0.04	…	1.00	0.98	–		
			x_3	1.02	0.09	…	…	1.00	–		
	f_b	$AROPE_{MC}$	x_1	0.02	0.03	1.00	0.10	−0.10	–	0.211	0.054
			x_2	0.01	0.03	…	1.00	0.77	–		
			x_3	0.01	0.03	…	…	1.00	–		
		$ROPE_{PSO}$	x_1	0.00	0.01	1.00	−0.10	−0.03	–	0.032	0.025
			x_2	0.00	0.01	…	1.00	0.01	–		
			x_3	0.00	0.01	…	…	1.00	–		
4D	f_a	$AROPE_{MC}$	x_1	0.95	0.01	1.00	0.52	0.39	0.35	0.104	0.027
			x_2	0.90	0.01	…	1.00	0.71	0.63		
			x_3	0.80	0.02	…	…	1.00	0.89		
			x_4	0.65	0.03	…	…	…	1.00		
		$ROPE_{PSO}$	x_1	1.01	0.01	1.00	0.86	0.85	0.84	0.075	0.025
			x_2	1.02	0.02	…	1.00	0.95	0.93		
			x_3	1.04	0.04	…	…	1.00	0.98		
			x_4	1.07	0.09	…	…	…	1.00		
	f_b	$AROPE_{MC}$	x_1	0.02	0.03	1.00	−0.08	−0.03	−0.07	5.169	0.074
			x_2	0.01	0.03	…	1.00	0.70	0.03		
			x_3	0.01	0.03	…	…	1.00	−0.20		
			x_4	1.03	0.08	…	…	…	1.00		
		$ROPE_{PSO}$	x_1	0.00	0.01	1.00	−0.02	−0.11	−0.09	0.043	0.032
			x_2	0.00	0.01	…	1.00	0.01	0.00		
			x_3	0.00	0.01	…	…	1.00	0.03		
			x_4	0.00	0.01	…	…	…	1.00		

Table 2. Overview of the used model parameters considered for calibration; the reference parameter vector $\boldsymbol{\theta}_{wb}$ was estimated in order to use WaSiM for water-balance simulations in the Rietholzbach catchment.

parameter	unit	reference ($\boldsymbol{\theta}_{wb}$)	upper and lower boundary		description
k_d	[h]	7	0.01	25	storage coefficient of direct runoff
k_i	[h]	20	0.01	40	storage coefficient of interflow
dr	[–]	2.1	0.01	80	drainage density

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Table 3. Mean value (Mean), standard deviation (Std), coefficient of variation (CV), minimum, maximum and correlation coefficients between the generated samples for the conceptual model parameters estimated by $\mathrm{ROPE_{PSO}}$ (above) and $\mathrm{AROPE_{MC}}$ (below).

Parameter	Mean	Std	CV	Min	Max	k_d	k_i	dr
k_d	2.30	0.41	0.18	1.47	3.54	1.00	0.05	−0.39
k_i	4.60	1.24	0.27	2.22	8.06	...	1.00	−0.30
dr	5.13	1.52	0.30	2.04	9.91	...	...	1.00

Parameter	Mean	Std	CV	Min	Max	k_d	k_i	dr
k_d	2.78	0.30	0.11	2.04	3.67	1.00	−0.20	−0.38
k_i	6.54	1.36	0.21	3.12	12.65	...	1.00	−0.64
dr	5.23	0.96	0.18	3.08	8.20	...	...	1.00

Table 4. Sub-division of all flood events in a calibration, control and validation set.

Set	Used flood events
Calibration	$\{6,11,20\}$
Control	$\{14,18,24\}$
Validation	$\{1,2,3,4,5,7,8,9,10,12,\dots$ $13,15,16,17,19,21,22,23\}$

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Table 5. Mean overall validation performance for the parameter vectors estimated by ROPE_{PSO} and the three compared algorithms in Krauße and Cullmann (2011); the model performance was calculated on the validation set according to Table 4.

	FloodSkill				**NS**				**rPD**			
	μ	σ	worst	best	μ	σ	worst	best	μ	σ	worst	best
ROPE_{PSO}	0.40	0.013	0.44	0.37	0.55	0.036	0.44	0.61	0.35	0.021	0.41	0.30
AROPE_{MC}	0.41	0.015	0.44	0.37	0.55	0.036	0.47	0.61	0.38	0.030	0.44	0.30
IPM	0.43	0.044	0.58	0.38	0.52	0.049	0.33	0.58	0.37	0.048	0.52	0.32
GA	0.42	0.024	0.46	0.38	0.53	0.033	0.43	0.61	0.38	0.030	0.46	0.32

Table 6. Correlation between data depth and overall validation performance of all parameter vectors estimated with AROPE_{MC} and ROPE_{PSO} for the multiple event calibration.

Algorithm	FloodSkill
AROPE_{MC}	−0.25
ROPE_{PSO}	−0.44

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Table 7. Expectation values of the physical properties of the prevailing soil types in the Rietholzbach catchment, classified according to USDA, and corresponding soil hydraulic parameters; the parameterisation of the soil hydraulic parameters is done for each soil according to the approach provided in Grundmann (2010) by the help of the pedotransfer functions provided in Wösten et al. (1999) and Brakensiek et al. (1984); the expectation values are the mean over 10 000 realisations.

	L loam	SL sandy loam	SiCL silty clay loam	SiL silt loam	LS loamy sand
catchment area [%]	15	20	3	51	11
clay [%]	20	10	33.5	13.5	7.5
silt [%]	39	25	56.5	69	15
sand [%]	41	65	10	17.5	77.5
humus content [%]			2.5		
k_s [m s^{-1}]	1.81×10^{-6}	1.45×10^{-5}	8.61×10^{-8}	2.85×10^{-7}	4.26×10^{-5}
α [1 m^{-1}]	3.49	4.48	2.00	1.35	5.52
θ_r			0.01		
θ_s	0.42	0.41	0.43	0.42	0.41
n	1.18	1.27	1.13	1.24	1.32

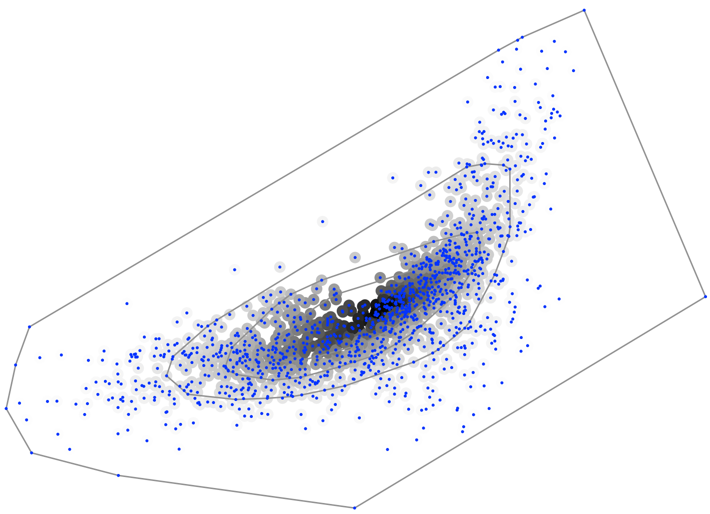


Fig. 1. 2-dimensional point set shaded according to assigned depth. A darker point represents higher depth. The lines indicate convex hulls enclosing the 25%, 50%, 75% and 100% deepest points. The used depth function was halfspace depth.

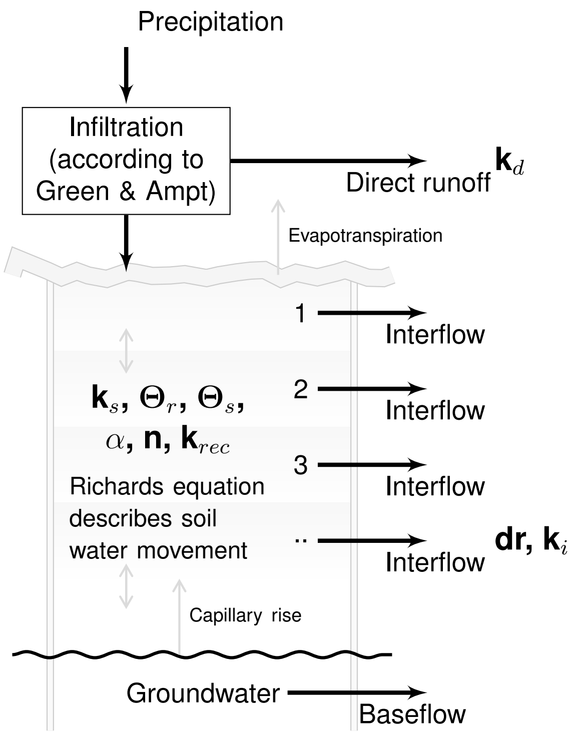


Fig. 2. Scheme of the WaSiM soil module with location of impact of soil hydraulic and conceptual model parameters (bold).

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

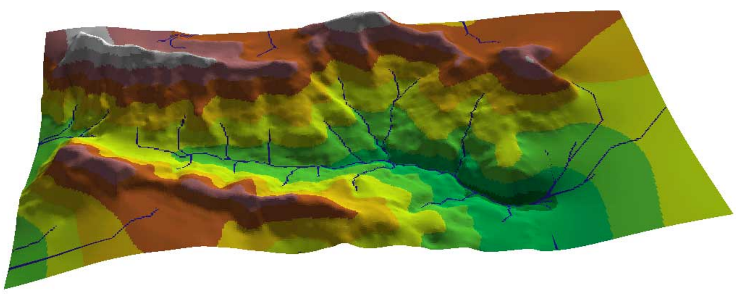


Fig. 3. A 3-D-view of the catchment with the potential rivers flowpaths.

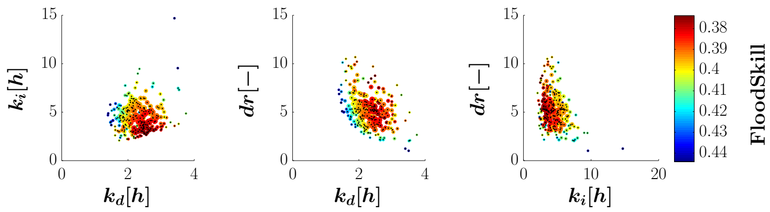


Fig. 4. Scatter plot of the good parameter vectors shaded according to their validation performance. Red points have a good validation performance, blue points are worse (see colorbar). The size of the shades is proportional to the data depth of each point with respect to the whole cloud.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

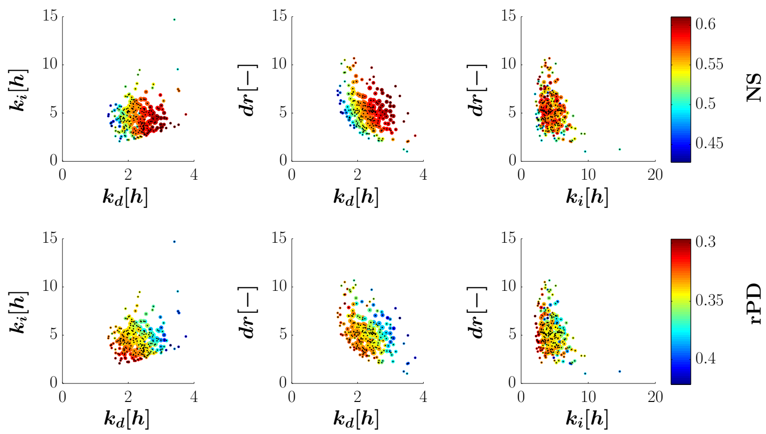
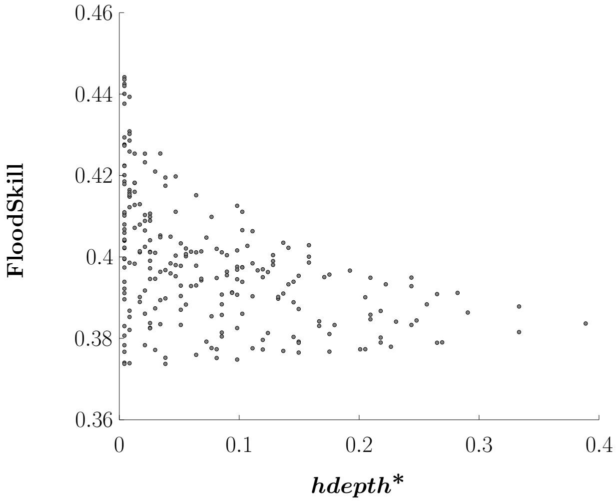


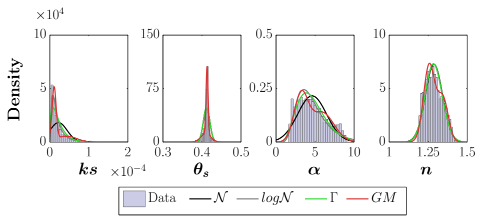
Fig. 5. Scatter plot of the good parameter vectors shaded according to their validation performance for the criteria NS and rPD. Red points have a good validation performance, blue points are worse (see colorbar). The size of the shades is proportional to the data depth of each point with respect to the whole cloud.



Algorithm	FloodSkill
AROPE$_{MC}$	-0.25
ROPE$_{PSO}$	-0.44

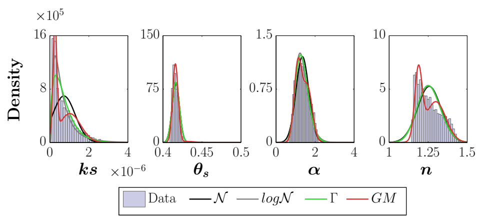
Fig. 6. Correlation between data depth and overall validation performance of all parameter vectors estimated with AROPE$_{MC}$ and ROPE$_{PSO}$ for the multiple event calibration.

	ks	θ_s	α	n
Mean	2.227e-05	0.4120	4.62	1.29
Std	2.185e-05	0.0083	1.85	0.05
CV	0.98	0.02	0.40	0.04



(a)

	ks	θ_s	α	n
Mean	7.207e-07	0.4165	1.35	1.26
Std	5.726e-07	0.0048	0.33	0.08
CV	0.79	0.01	0.24	0.06



(b)

Fig. 7. Prior distribution of the soil hydraulic parameters for the soils SL **(a)** and SiL **(b)**.

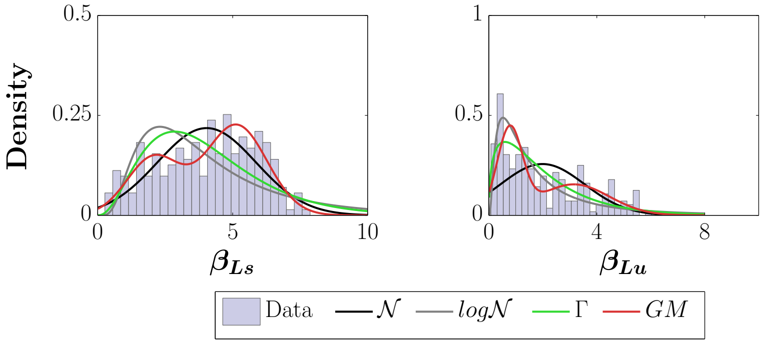


Fig. 8. Prior distribution of the scaling paramters β_{SL} and β_{SiL}.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

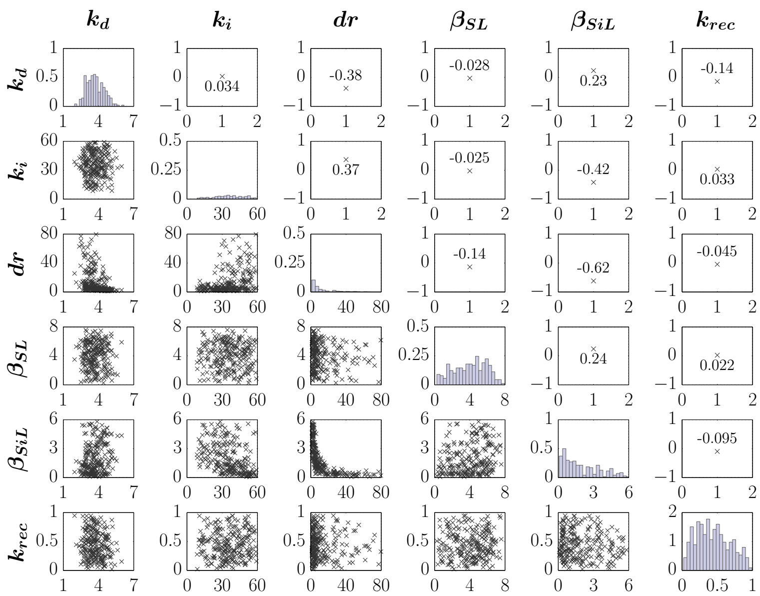


Fig. 9. Distribution of the model parameter vectors estimated by $\mathrm{AROPE_{MC}}$.

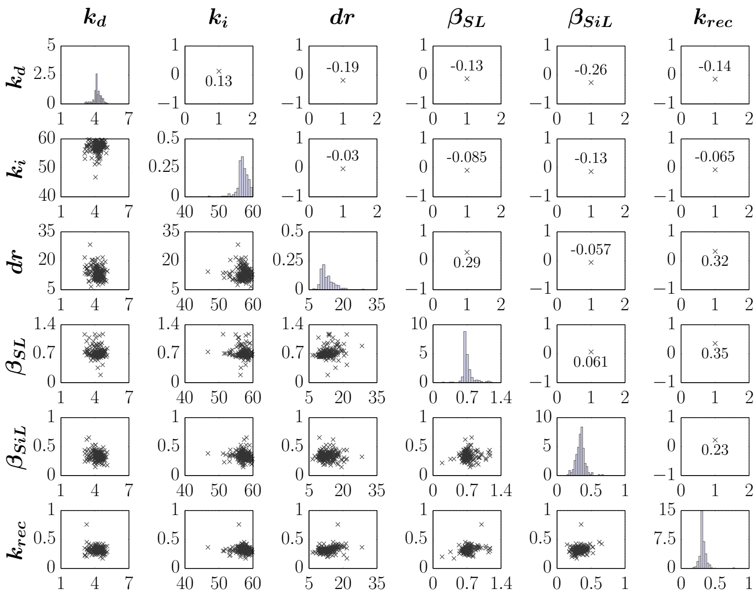
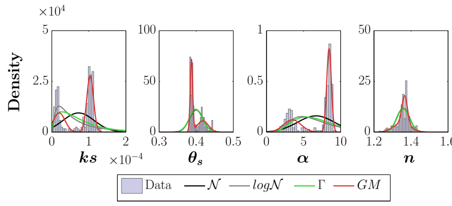


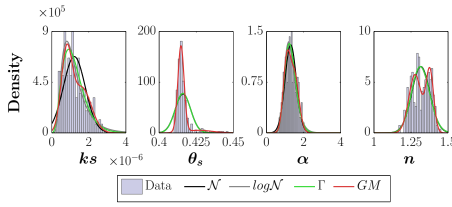
Fig. 10. Distribution of the model parameter vectors estimated by ROPE_{PSO}.

	ks	θ_s	α	n
Mean	7.231e-05	0.400	6.66	1.36
Std	4.224e-05	0.018	2.49	0.03
CV	0.58	0.05	0.37	0.03



(a)

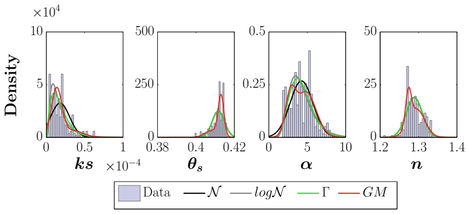
	ks	θ_s	α	n
Mean	1.215e-06	0.416	1.30	1.32
Std	5.893e-07	0.005	0.31	0.06
CV	0.48	0.01	0.24	0.05



(b)

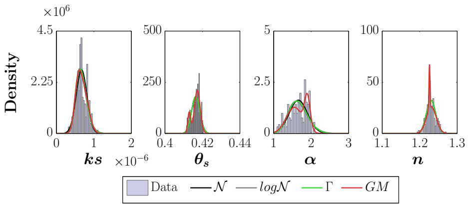
Fig. 11. Distribution of the soil hydraulic parameters estimated by AROPE_{MC} for the soils SL **(a)** and SiL **(b)**.

	ks	θ_s	α	n
Mean	1.801e-05	0.412	4.24	1.29
Std	1.228e-05	0.003	1.49	0.02
CV	0.68	0.01	0.35	0.02



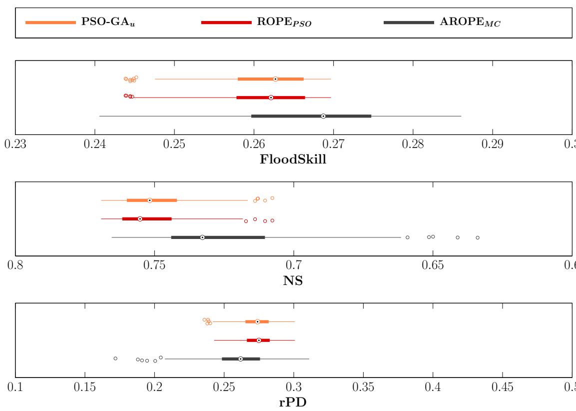
(a)

	ks	θ_s	α	n
Mean	6.707e-07	0.416	1.66	1.23
Std	1.427e-07	0.002	0.25	0.01
CV	0.21	0.01	0.15	0.01



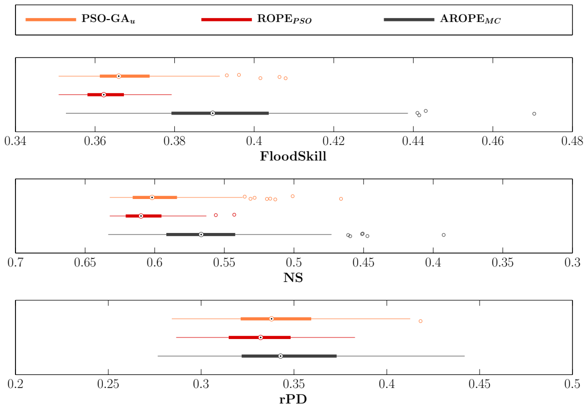
(b)

Fig. 12. Distribution of the soil hydraulic parameters estimated by ROPE_{PSO} for the soils SL **(a)** and SiL **(b)**.



	FloodSkill				NS				rPD			
	μ	σ	worst	best	μ	σ	worst	best	μ	σ	worst	best
PSO-GA_u	0.26	0.006	0.27	0.24	0.75	0.013	0.71	0.77	0.27	0.013	0.30	0.24
ROPE_{PSO}	0.26	0.006	0.27	0.24	0.75	0.014	0.71	0.77	0.27	0.013	0.30	0.24
AROPE_{MC}	0.27	0.010	0.29	0.24	0.73	0.025	0.63	0.77	0.26	0.024	0.31	0.17

Fig. 13. Calibration performance for the parameter vectors estimated by PSO-GA_u, ROPE_{PSO} and AROPE_{MC}.



	FloodSkill				NS				rPD			
	μ	σ	worst	best	μ	σ	worst	best	μ	σ	worst	best
PSO-GA$_u$	0.37	0.008	0.41	0.35	0.60	0.021	0.50	0.63	0.33	0.024	0.41	0.28
ROPE$_{\text{PSO}}$	0.36	0.006	0.38	0.35	0.61	0.018	0.54	0.63	0.33	0.023	0.38	0.29
AROPE$_{\text{MC}}$	0.39	0.019	0.47	0.35	0.56	0.039	0.39	0.63	0.35	0.033	0.44	0.28

Fig. 14. Validation performance for the parameter vectors estimated by PSO-GA$_{\text{u}}$, ROPE$_{\text{PSO}}$ and AROPE$_{\text{MC}}$.

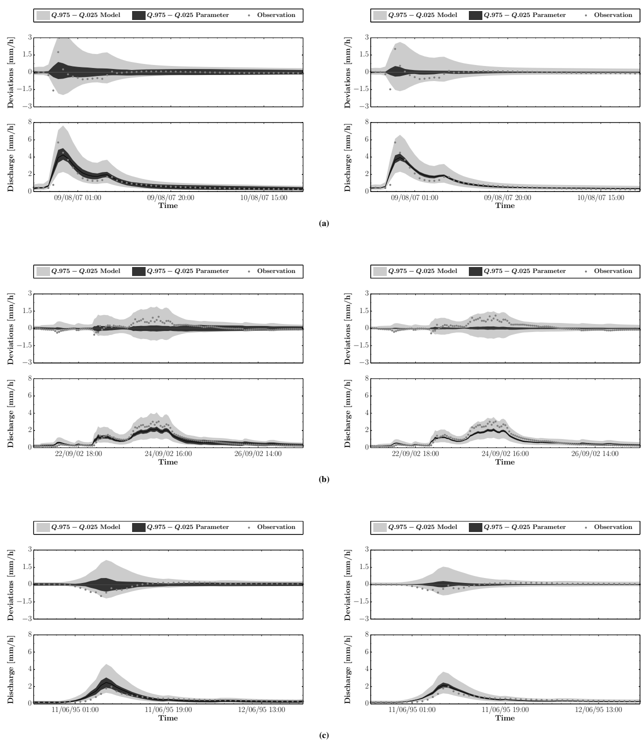


Fig. 15. Hydrograph prediction uncertainty associated with the uncertainty in the model (lighter shading) and parameter estimates (darker shading) for the flood events 4 **(a)**, 12 **(b)** and 19 **(c)**, estimated by AROPE$_{\text{MC}}$ (left column) and ROPE$_{\text{PSO}}$ (right column). The dots correspond to the observed streamflow data. The shaded areas of uncertainty correspond to the 95% confidence intervals.