ELSEVIER

Application of stochastic parameter optimization to the Sacramento Soil Moisture Accounting model

Jasper A. Vrugt [a,*], Hoshin V. Gupta [b], Stefan C. Dekker [c], Soroosh Sorooshian [d], Thorsten Wagener [e], Willem Bouten [f]

[a] *Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[b] *Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721, USA*
[c] *Environmental Sciences, Universiteit Utrecht, P.O. Box 80115, 3508 TC Utrecht, The Netherlands*
[d] *Department of Civil and Environmental Engineering, E/4130 Engineering Gateway, University of California, Irvine, CA 92697, USA*
[e] *Department of Civil and Environmental Engineering, Sackett Building, University Park, PA 16802, USA*
[f] *Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands*

Abstract

Hydrological models generally contain parameters that cannot be measured directly, but can only be meaningfully inferred by calibration against a historical record of input–output data. While considerable progress has been made in the development and application of automatic procedures for model calibration, such methods have received criticism for their lack of rigor in treating uncertainty in the parameter estimates. In this paper, we apply the recently developed Shuffled Complex Evolution Metropolis algorithm (SCEM-UA) to stochastic calibration of the parameters in the Sacramento Soil Moisture Accounting model (SAC-SMA) model using historical data from the Leaf River in Mississippi. The SCEM-UA algorithm is a Markov Chain Monte Carlo sampler that provides an estimate of the most likely parameter set and underlying posterior distribution within a single optimization run. In particular, we explore the relationship between the length and variability of the streamflow data and the Bayesian uncertainty associated with the SAC-SMA model parameters and compare SCEM-UA derived parameter values with those obtained using deterministic SCE-UA calibrations. Most significantly, for the Leaf River catchments under study our results demonstrate that most of the 13 SAC-SMA parameters are well identified by calibration to daily streamflow data suggesting that this data contains more information than has previously been reported in the literature.
© 2006 Elsevier B.V. All rights reserved.

---

* Corresponding author.
E-mail addresses: vrugt@lanl.gov (J.A. Vrugt), hoshin.gupta@hwr.arizona.edu (H.V. Gupta), s.dekker@geog.uu.nl (S.C. Dekker), soroosh@uci.edu (S. Sorooshian), thorsten@engr.psu.edu (T. Wagener), w.bouten@science.uva.nl (W. Bouten).

1. Introduction and scope

Since, the early 1960s, hydrologists have concentrated their efforts on the development and application of models of the rainfall-runoff process. These models

can be classified as conceptual, when water balance dynamics are represented by heuristic or empiric equations deemed to be qualitatively reasonable, or as physically based in which the model equations are based on scientifically accepted principles (Kuczera, 1997). Most operational conceptual rainfall-runoff (CRR) models typically have on the order of 10 or more parameters that link transfer functions of several interconnected water stores. It is assumed that these conceptual storages correspond to physically identifiable control volumes in real space, even though the boundaries of these control volumes are generally not known. While some of the values of parameters in CRR models can be derived directly from knowledge of physical watershed characteristics, others are effective quantities that cannot, in practice, be measured in the field, and therefore have to be estimated through calibration against a measured streamflow hydrograph using either a trial-and-error 'manual approach' or an automated search algorithm (Boyle et al., 2000; Madsen, 2000). The parameters, which are estimated in this manner, represent effective conceptual representations of spatially and temporally heterogeneous watershed properties. A model calibrated by such means can be used for the simulation or prediction of hydrologic events outside of the historical record used for model calibration, if it can be reasonably assumed that the physical characteristics of the watershed and the hydrologic/climate conditions remain similar (Gupta et al., 2002).

Until the early 1990s, the available automated optimization algorithms could not be relied upon to find the actual global optimum in a prescribed objective function. Advances in computational resources finally enabled Duan et al. (1992) to conduct an exhaustive computer based evaluation of the response surface of the objective function. This research revealed the existence of multiple local optima in the response surface, nested within several larger regions of attractions, and explained the convergence problems reported by previous studies. These insights in the response surface led to the development of the Shuffled Complex Evolution (SCE-UA) optimization algorithm (Duan et al., 1992; Sorooshian et al., 1993), whose strength and reliability has been proven by numerous researchers worldwide throughout the years.

While considerable progress has been made in the development and application of automated procedures for watershed model calibration, such methods have received criticism for their lack of rigor in properly treating parameter uncertainty (Beven and Binley, 1992; Gupta et al., 1998; Thiemann et al., 2001; Vrugt et al., 2005a,b). In a previous paper (Vrugt et al., 2003), we presented a general-purpose code, entitled the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm, which is especially designed to provide an estimate of the most likely parameter set and underlying posterior distribution within a single optimization run. As, the SCEM-UA algorithm thoroughly exploits the global parameter space and explicitly accounts for parameter interdependence and non-linearity of the employed CRR model, the algorithm generates an accurate representation of parameter uncertainty, and its antithesis parameter identifiability. The adaptive capabilities of the SCEM-UA algorithm significantly reduces the number of model simulations needed to infer the posterior distribution of the parameters when compared with traditional Metropolis–Hastings samplers.

In the same paper, we evaluated the identifiability of the parameters of a rather parsimonious five-parameter CRR model, consisting of a relatively simple rainfall excess model, connected with two series of linear reservoirs. Those results indicated that the entire model structure was well identifiable by calibration to runoff data, thereby supporting the statements of Beven (1989), and results of Jakeman and Hornberger (1993). They argued that simple RR models with four to five parameters provide an adequate fit to the streamflow data and that the addition of more model structure and its associated parameters leads to no significant improvement in fit yet introduces poorly identified parameters.

The purpose of the present paper is to extend the analysis in Vrugt et al. (2003) by applying the SCEM-UA algorithm to stochastic optimization of the parameters of the Sacramento Soil Moisture Accounting (SAC-SMA) model of the National Weather Service River Forecast System (NWSRFS). In previous work, Yapo et al. (1996) conducted a large number of calibration runs of the SAC-SMA model with the SCE-UA global optimization algorithm (Duan et al., 1992) using different lengths of data from different sections of a 40 years historical record,

and concluded that approximately 8 years of stream-flow data are required to obtain calibrations that are rather insensitive to the period selected. This conclusion was drawn based on the performance of the SAC-SMA model structure, rather than the statistical uncertainty associated with the final parameter estimates. In this paper, we extend the work by Yapo et al. (1996) and investigate the uncertainty and stationarity of the SAC-SMA model parameters as function of the length and variability of the streamflow data. Moreover, we provide a comparison between SCEM-UA derived values of the SAC-SMA model parameters and those obtained using classical deterministic SCE-UA calibrations (Sorooshian et al., 1993).

2. Treatment of parameter uncertainty

2.1. Classical model calibration approach

The fundamental problem with which we are concerned is to estimate parameter values and their associated uncertainty in the SAC-SMA model using a historical record of streamflow data. The formulation of this resulting inverse problem can be expressed in a generic form if we assemble the state variables in the SAC-SMA model at time t into the state vector ψ_t. The evolution of this state vector is described with:

$$\psi_{t+1} = \eta(\psi_t, \tilde{X}_t, \theta) \tag{1}$$

where ψ is a vector of m unknown state variables, $\eta(\cdot)$ represents the SAC-SMA model used to simulate the state evolution, $\tilde{X}$ is an observed forcing field (precipitation and evapotranspiration), θ is a set of p model parameters, and t denotes time. Let $\tilde{Y} = \{\tilde{y}_1, \dots, \tilde{y}_n\}$ denote the vector of streamflow measurement data available at time steps $1,\dots,n$ and let $Y(\theta) = \{y_1(\theta), \dots, y_n(\theta)\}$ represent the corresponding vector of SAC-SMA streamflow predictions using the parameter set θ. The SAC-SMA output predictions are related to its internal state according to:

$$y_t = H(\psi_t) \tag{2}$$

where $H(\cdot)$ is the measurement operator, which maps the state space into the measurement or model output space.

The classical approach to estimating the parameters in Eq. (1) is to ignore input uncertainty ($\tilde{X} = X$) and to assume that the predictive model η is a correct, or at least accurate, representation of the underlying physical data-generating process. In line with classical statistical estimation theory, the traditional 'best' parameter set in Eq. (1) can then be found by minimizing the following lumped simple least square (SLS) objective function with respect to θ:

$$F_{\text{SLS}}(\theta) = \sum_{t=1}^{n} w_t (y_t - \tilde{y}_t)^2 \tag{3}$$

where w_t denote weights for particular data points. Minimization of Eq. (3) will result in a single 'best' parameter set. However, given the presence of errors in the input, output and model structure, over conditioning of the model to a single parameter set is unreasonable and cannot be justified (Vrugt et al., 2005a,b).

2.2. Bayesian model calibration

One way to directly address this problem of overconditioning is to abandon the Frequentists approach of believing that the model parameters in Eq. (1) are fixed but unknown, and to adopt a Bayesian viewpoint which allows the identification of a plausible set of values for the parameters of the model given the available streamflow data. The Bayesian approach treats the SAC-SMA model parameters in Eq. (1) as probabilistic variables having a joint posterior probability density function (pdf), which summarizes the state of knowledge about the model parameters given available streamflow data $\tilde{Y}$. It uses probability distributions to describe this state of knowledge, which is why the parameters are treated as random variables. The posterior pdf, $p(\theta|\tilde{Y})$, is proportional to the product of the likelihood function and the prior pdf. The prior pdf with probability density (or mass) function $p(\theta)$ summarizes information about θ before any data are collected. In this paper, we use the following form of the posterior pdf (Box and Tiao):

$$p(\theta|\tilde{Y}) \propto F_{\text{SLS}}(\theta)^{-1/2n} \tag{4}$$

which is derived when assuming a non-informative (uniform) prior distribution with uncorrelated,

homoscedastic, and Gaussian distributed error residuals. With the implementation of Eq. (4), the belief in the existence of a single optimal parameter set is discarded in favour of the search for a set of plausible parameter values.

2.3. The Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm

To generate samples from Eq. (4), and to summarize the posterior parameter pdf using statistical moments and histograms, we use an implementation of the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm. The SCEM-UA algorithm is a general-purpose global optimization algorithm that provides an efficient estimate of the most likely parameter set (mode) and its underlying posterior probability distribution within a single optimization run (see Vrugt et al., 2003). The algorithm is a Markov Chain Monte Carlo (MCMC) sampler, which generates multiple sequences of parameter sets $\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(k+1)}\}$ that converge to the stationary posterior distribution for a large enough number of simulations k. The SCEM-UA algorithm is related to the successful SCE-UA global optimization method, but uses the Metropolis–Hastings (MH) search strategy (Metropolis et al, 1953; Hastings, 1970) instead of the Downhill Simplex method for population evolution, and is therefore able to simultaneously infer both the most likely parameter set and its underlying posterior probability distribution within a single optimization run. A detailed description and explanation of the method appears in Vrugt et al. (2003), and so will not be repeated here.

The SCEM-UA method involves the initial selection of a population of points distributed randomly throughout the p-dimensional feasible parameter space. In the absence of prior information about the location of the maximum likelihood value, a uniform sampling distribution is used. For each point, the posterior density criterion in Eq. (4) is computed. The population of parameter sets is subsequently partitioned into a number of complexes, and in each complex a parallel sequence is launched from the point that exhibits the highest posterior density. A new candidate point in each sequence is generated using a multivariate normal distribution either centered on the current draw of the sequence or the mean of the points in the complex augmented with the covariance structure induced between the points in the complex. The Metropolis-annealing (Metropolis et al., 1953) criterion is used to test whether the candidate point should be added to the current sequence. Finally, the new candidate point is shuffled into the original population of complexes. The evolution and shuffling procedures are repeated until the Gelman–Rubin convergence diagnostic for each of the parameters demonstrates convergence to a stationary posterior target distribution (Gelman and Rubin, 1992). Although general convergence proofs for non-homogenous Markov Chain algorithms are still a matter of ongoing advanced mathematical research, experiments conducted using standard mathematical test problems have shown that the SCEM-UA derived posterior distribution closely approximates the target distribution (Vrugt et al., 2003).

Implementation of the SCEM-UA algorithm requires a large number of SAC-SMA model evaluations. Fortunately, there has been considerable progress in the development of distributed computer systems using the power of multiple processors to efficiently solve large computational problems. In this study, we implemented the SCEM-UA algorithm using a LAM/MPI distributed computing interface for the Octave programming environment (Vrugt et al., in press). Our parallel implementation takes better advantage of the computational power of a distributed computer system. Based on recommendations in our previous work (Vrugt et al., 2003), the stationary posterior distribution corresponding to the density criterion defined in Eq. (4) was estimated using a population size of 1.000 points in combination with a total of 60.000 SAC-SMA model evaluations and 25 Pentium IV 3.40 GHz parallel processors. The average CPU time required for 60.000 model evaluations was approximately 5 min. To avoid problems with heteroscedastic and non-Gaussian error distributions, we applied a Box–Cox power transformation (Box and Cox, 1974) with $\lambda = 0.3$ to the measured and SAC-SMA predicted streamflow (Misirli, 2003). We do not consider uncertainty in λ-values in our analysis because the goal of the current paper is not to demonstrate how to deal with model uncertainty, but to demonstrate the merits of the SCEM-UA methodology for model calibration.

In another recent paper, we present a combined parameter and state estimation method that provides a better treatment of model structural errors (Vrugt et al., 2005a,b). The estimated pdfs of parameter values are evaluated using a historical record of streamflow data outside the calibration period.

3. Case study

3.1. The Sacramento Soil Moisture Accounting (SAC-SMA) model

The CRR model used throughout this study is the Sacramento Soil Moisture Accounting (SAC-SMA) model used by the National Weather Service River Forecast System (NWSRFS) for flood forecasting throughout the United States. The model is one of the components of the NWSRFS used to convert precipitation input into streamflow outputs (Burnash et al., 1973; Peck, 1976; Kitanidis and Bras, 1980a,b; Brazil and Hudlow, 1981; Sorooshian et al., 1993). The SAC-SMA model has 16 parameters whose values must be specified (Table 1).

Following the recommendation of Peck (1976), three parameters SIDE, RIVA, and RSERV, were fixed at pre-specified values. The remaining 13 parameters were selected for stochastic optimization using the SCEM-UA algorithm, outlined in Section 3.1. The feasible parameter space is specified in Table 1 by fixing the lower and upper parameter bounds at the Level Zero ranges recommended by Boyle et al. (2000). These ranges are defined conservatively based on the maximum plausible ranges for the parameters based on physical reasoning.

3.2. The study basin and hydrological data

We illustrate the usefulness and applicability of the SCEM-UA algorithm for stochastic optimization of the parameters in the SAC-SMA model using historical streamflow data from the Leaf River watershed. The basin is located north of Collins, Mississippi, with an area of approximately 1950 km^2. The data, obtained from the Hydrologic Research Laboratory (HL), consists of 6 hourly mean areal precipitation (mm/day), daily potential evapotranspiration (mm/day), and streamflow (m^3/s). Forty

Table 1
Parameters of the SAC-SMA model and their initial uncertainty ranges

Parameter	Description	Units	Initial ranges
Capacity thresholds			
UZTWM	Upper zone tension water maximum storage	mm	1.0–150.0
UZFWM	Upper zone free water maximum storage	mm	1.0–150.0
LZTWM	Lower zone tension water maximum storage	mm	1.0–500.0
LZFPM	Lower zone free water primary maximum storage	mm	1.0–1000.0
LZFSM	Lower zone free water supplemental maximum storage	mm	1.0–1000.0
ADIMP	Additional impervious area	–	0.0–0.40
Recession parameters			
UZK	Upper zone free water lateral depletion rate	day^{-1}	0.1–0.5
LZPK	Lower zone primary free water depletion rate	day^{-1}	0.0001–0.025
LZSK	Lower zone supplemental free water depletion rate	day^{-1}	0.01–0.25
Percolation and other			
ZPERC	Maximum percolation rate	–	1.0–250.0
REXP	Exponent of the percolation equation	–	0.0–5.0
PCTIM	Impervious fraction of the watershed area	–	0.0–0.1
PFREE	Fraction percolating from upper to lower zone free water storage	–	0.0–0.1
Not optimized			
RIVA	Riparian vegetation area	–	0.0
SIDE	Ratio of deep recharge to channel base flow	–	0.0
RSERV	Fraction of lower zone free water not transferable to tension water	–	0.3

consecutive years of data (1948–1988) are available for this watershed, representing a wide variety of hydrological conditions. In keeping with previous studies (Boyle et al., 2000; Thiemann et al., 2001; Vrugt et al., 2003), the hydrological data used for model calibration and uncertainty assessment of the parameters of the SAC-SMA model consists of approximately 36 years (28 July, 1952–30 September, 1988) of streamflow data. The mean annual precipitation for the entire period is 1324 mm, and the mean runoff is 27.13 m^3/s.

3.3. Parameter uncertainty as function of length calibration set

In the hydrological literature, several contributions are found which examine the relationship between data and the statistical uncertainty associated with the parameter estimates. For instance, Sorooshian et al. (1983) suggested that rather than length, the quality of information contained in the data is important. They also stated that the data sequences that contain greater hydrologic variability are more likely to sufficiently activate the various operational modes of the model, resulting in reliable parameter estimates. Here, reliability is characterized in terms of stability and consistency of parameter estimates.

To study the relationship between the length of the calibration set and the Bayesian uncertainty associated with the parameter estimates, the first 10 years of data on record (28 July, 1952–28 September, 1962) for the Leaf River watershed were used. A traditional split sample technique was applied to divide this data set into 10 different subsets. Starting at 28 July, 1952, for each run with the SCEM-UA algorithm the length of the calibration set was sequentially increased with 1 consecutive year of measurements to finally arrive at the total set 10 years of calibration data. To reduce sensitivity to state value initialisation in any of the optimizations, a 65 days warm-up period was used, during which no updating of the posterior density was performed.

Fig. 1 illustrates the relationship between the length of the calibration set of measured streamflows in years and the Bayesian confidence intervals associated with the parameters in the SAC-SMA model.

The parameters were scaled according to their prior uncertainty ranges defined in Table 1 to yield normalized ranges. The gray-shaded area in each parameter plot represents the evolution of the 95% Bayesian confidence intervals of the HPD region, whereas the marked squares and dotted line correspond to the SCE-UA solution, and the 'most likely' SCEM-UA solution, respectively. Without any calibration, the ranges reflect the initial parameter uncertainty, not conditioned on any input–output time series of measured streamflows. After assimilating and processing streamflow data with the SCEM-UA algorithm, the uncertainty associated with the SAC-SMA model parameters decreases. The important results depicted in Fig. 1 are as follows:

1. The location and size of the HPD region of the posterior probability distribution of the parameters is fairly stable with increasing length of the calibration set. The parameter estimates and their Bayesian confidence intervals appear to be relatively unaffected by the length of the calibration set.
2. There is only a marginal reduction of the size of the Bayesian confidence intervals, centered on the most likely SAC-SMA parameter values, with the availability of more calibration data.
3. In relation to the Level Zero or prior parameter ranges, most of the SAC-SMA parameters are well identified by calibration to streamflow data. In particular, the capacity parameters UZTWM, UZFWM, LZTWM, LZFSM and LZFPM are well determined, while parameters ZPERC and REXP (that control percolation), ADIMP (additional impervious area), and the rate parameters LZSK and LZPK are less well determined. These results suggest that most uncertainty in the model structure is associated with percolation from the upper zone and depletion from the lower zone. Notice, also that with the use of relatively short calibration time series (<5 years) the HPD region of the LZPK parameter traverses through the feasible parameter space. This characteristic jumping behavior suggests at least some correlation with the characteristic of the hydrologic year. We will further elaborate on this issue in Sections 3.8 and 3.9 of this paper.
4. There is an excellent agreement between the most likely value of the posterior distribution, identified using the SCEM-UA algorithm, and the
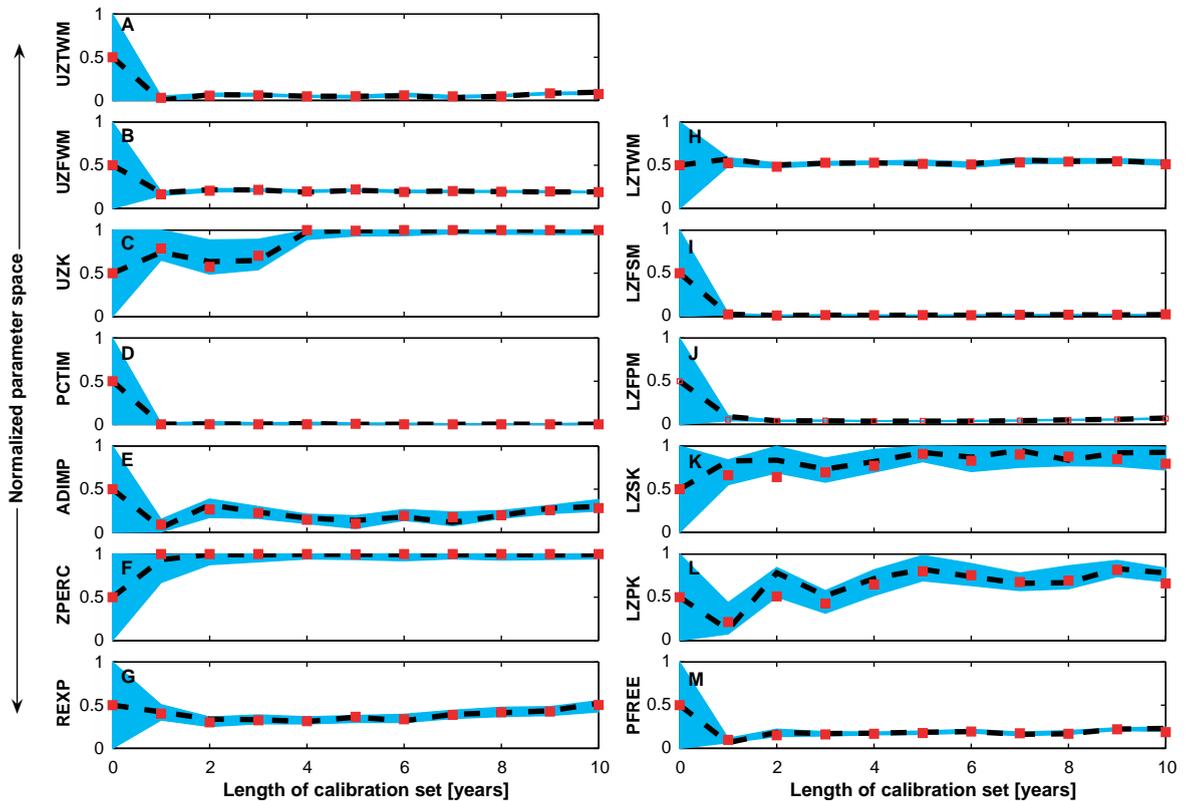
Fig. 1. Evolution of the 95% confidence intervals of the SAC-SMA model parameters as function of the length of the calibration set of measured streamflows in years. The evolution of the most likely parameter values is indicated with the solid line, whereas the SCE-UA solutions are indicated with the squared symbols.

optimal parameter set found by separately fitting the SAC-SMA model to the probability density criterion, previously defined in Eq. (4) using the SCE-UA global optimization algorithm (Duan et al., 1992). This result is important as it confirms that the SCEM-UA algorithm correctly locates the high probability density region of the parameter space, also for a relatively high-dimensional optimization problem. This provides strong evidence that the employed MCMC search strategy in the SCEM-UA algorithm contains the desirable properties to deal with the specific peculiarities of the response surface of hydrological models.

Based on the first two arguments, it appears that 2–3 years of streamflow data are sufficient to provide

a stable estimate for most of the SAC-SMA model parameters. Here, stability is referred to as parameter estimates with a small uncertainty. The slow convergence of the uncertainty bounds with increasing number of calibration years is consistent with results derived from classical statistical Theorems using linear approximations to parameter uncertainty (Gupta and Sorooshian, 1985). However, the use of 2–3 years of streamflow data does not necessarily lead to consistent parameter estimates. Consistency implies that different calibration sets lead to the same approximate parameter values. For instance, if the analysis in this section would have been done with a different 10-years section from the 36-years historical record of streamflow data, the conclusions with respect to parameter stability would have been the same. This was concluded from additional

optimization runs. However, the SCEM-UA method would have assigned the highest probability to a different region in the parameter space. The rationales on data requirements adopted in this section are thus based on the stability of the parameter estimates. Summarizing, if over 2–3 years of daily streamflow data there is a good number of forcing events then little reduction in parameter uncertainty is realized using more than 3 years of data. Hence, these results would vary depending on the forcing climate. In Section 3.7 of this paper, we address the issue of parameter consistency by evaluating the pdf of the parameters over an independent period outside the historical record used for model calibration.

3.4. Marginal and joint probability distributions SAC-SMA model parameters

Fig. 2 presents the marginal posterior probability distributions (histograms) for 12 of the 13 SAC-SMA model parameters when using 3 years of calibration data (WY 1952–1954) with the SCEM-UA algorithm.

These histograms were derived from the SCEM-UA samples that were generated after convergence had been achieved to a stationary posterior distribution. The values of the most likely parameter estimates, derived with the SCE-UA algorithm, are separately indicated with squared symbols in each of the graphs. For display convenience, we decided not
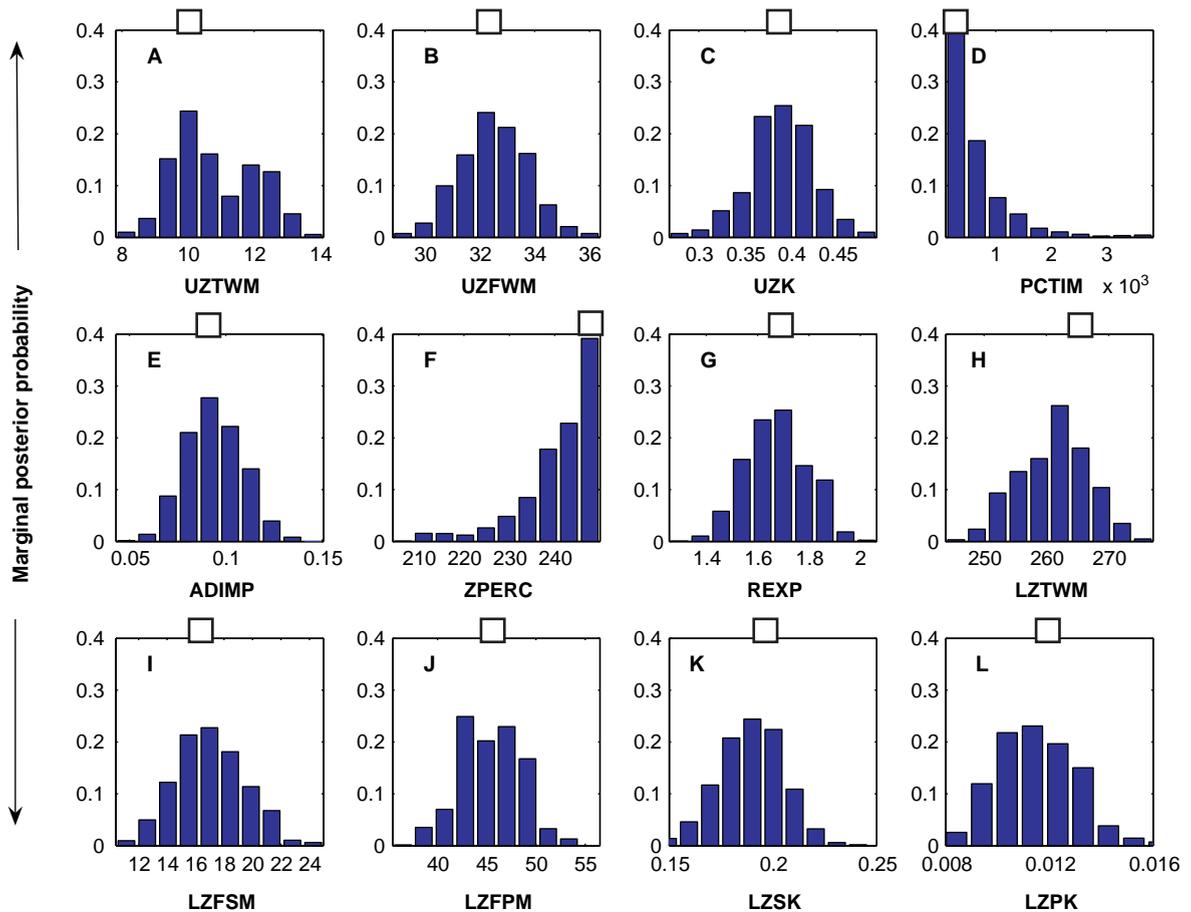

Fig. 2. Marginal posterior probability functions for the SAC-SMA model parameters derived when processing 3 years of streamflow data of the Leaf River with the SCEM-UA algorithm. The SCE-UA derived optimal parameter values are separately indicated with squared symbols in each of the graphs.

to include a graph for the parameter PFREE. While the parameters UZFWM (B), ADIMP (E), REXP (G), LZFSM (I), LZFPM (J) and LZPK (K) are well described by a normal distribution, the remaining SAC-SMA parameters UZTWM (A), UZK (C), PCTIM (D), ZPERC (F), LZTWM (H), and LZSK (K), are better described with bi-modal or lognormal distributions, respectively. The SCEM-UA algorithm retains many desirable features as it not only correctly infers the most likely parameter set and its underlying posterior distribution within a single optimization run, but also generates useful information about the nature of the response surface in the vicinity of the optimum. For most of the parameters, Fig. 2 shows excellent agreement between the modes of the histograms and the SCE-UA solutions within this high-density region. This indicates that previous deterministic calibrations that have used the SCE-UA algorithm are likely to have yielded robust parameters.

Table 2 presents the posterior mean, standard deviation, coefficient of variation (CV), and correlation structure induced between the SAC-SMA parameters in the HPD region of the posterior probability distribution, when using 3 years of calibration data (WY 1952–1954).

The posterior moments in Table 2 demonstrate that most of the SAC-SMA model parameters are well determined by calibration to streamflow data. The correlation between the parameters in the posterior distribution is typically low, further confirming that the SAC-SMA parameters are well identified. Fig. 3 shows how the mean of the absolute values in the parameter correlation matrix, behaves during the evolution of the SCEM-UA algorithm to the stationary posterior target distribution using 2 and 3 years of streamflow data.

Also included in this figure is the total number of parameter combinations in the parameter correlation matrix that are correlated higher than the specified threshold value of 0.50. Without any calibration (left hand side of Fig. 3A), the correlation structure between the parameters resemblances the uniform prior sampling distribution. When subsequently applying the various algorithmic steps of the SCEM-UA sampler, the parallel sequences evolve towards a region with higher posterior density, thereby increasing the correlation between the parameters (location I in Fig. 3A). We call this 'large-scale' correlation

structure. Finally, after convergence of the SCEM-UA sampler has been achieved to the posterior distribution, the correlation structure between the parameters reduces and remains stable with increasing number of SAC-SMA model evaluations. Now, the parameters are fully conditioned on the measured time series of streamflows and the remaining parameter correlation is on the 'small-scale' (location II in Fig. 3A). Clearly, erroneous conclusions about parameter interaction and identifiability can be made when having insufficient sampling close to the optimum of the prescribed objective function. Fig. 3B also illustrates these considerations in a two-dimensional plot of the sampled LZTWM–LZFSM parameter space. For each of the other calibration runs summarized and reported in Fig. 1, similar correlation structures between the parameters in the HPD region were found.

3.5. The information content of streamflow data

The results presented in Figs. 1–3 do not suggest overparametrization (Hooper et al., 1988; Beven, 1989) or equifinality (Beven and Binley, 1992). Also, the results contradict the work by Jakeman and Hornberger (1993), who argued that if daily streamflow data are available, only two linear storages in parallel driven by excess rainfall, corresponding to four parameters, are warranted by the data.

To verify the consistency of this result for the Leaf River dataset, we separately fitted the HYMOD model (Boyle, 2000), a parsimonious model structure consisting of a relatively simple two-parameter rainfall excess model, described in detail by Moore (1985), connected with two series of linear reservoirs to the probability density criterion, previously defined in Eq. (4) using the 10 different years of calibration data with the SCEM-UA algorithm. The Average Relative Parameter Error, ARPE, denoting the average of the diagonal entries of the covariance matrix of the SCEM-UA derived posterior pdf, with each entry normalized with the square of its estimated mean value (Jakeman et al., 1989, 1990), for each of the 10 calibration data sets using the HYMOD and SAC-SMA model, as well, as their predictive capabilities in terms of root mean squared error (RMSE) are summarized in Table 3.

Table 2
Posterior mean, standard deviation, coefficient of variation (CV (%)), and correlation structure between the SAC-SMA model parameters, derived when processing 3 years of streamflow data (1952–1954) of the Leaf River watershed with the SCEM-UA algorithm

| Parameter	Mean	Standard deviation	CV	UZTW-M	UZFW-M	UZK	PCTI-M	ADIMP	ZPERC	REXP	LZTW-M	LZFSM	LZFPM	LZSK	LZPK	PFREE
UZTWM	10.70	1.22	11.39	1.00	−0.02	−0.04	0.19	0.36	−0.01	0.11	−0.84	−0.17	0.28	0.17	0.16	0.57
UZFWM	32.55	1.23	3.78	–	1.00	−0.53	0.05	0.01	0.15	0.11	−0.01	−0.08	0.22	−0.15	−0.18	−0.33
UZK	0.39	0.03	8.91	–	–	1.00	−0.01	−0.41	−0.13	−0.03	0.07	0.37	−0.11	−0.01	−0.02	0.08
PCTIM	5.11×10^{-4}	4.93×10^{-4}	96.30	–	–	–	1.00	−0.16	0.01	0.03	−0.14	0.01	0.02	−0.08	0.02	−0.04
ADIMP	0.10	0.01	14.99	–	–	–	–	1.00	0.07	0.04	−0.21	−0.15	0.09	0.17	0.15	0.31
ZPERC	241.46	7.91	3.27	–	–	–	–	–	1.00	0.05	0.06	−0.06	−0.10	−0.13	0.02	0.00
REXP	1.66	0.12	7.23	–	–	–	–	–	–	1.00	−0.03	0.65	0.68	0.24	−0.10	−0.12
LZTWM	261.46	5.43	2.08	–	–	–	–	–	–	–	1.00	0.28	−0.16	−0.14	−0.23	−0.63
LZFSM	16.96	2.39	14.08	–	–	–	–	–	–	–	–	1.00	0.41	−0.33	−0.55	−0.36
LZFPM	45.08	3.01	6.68	–	–	–	–	–	–	–	–	–	1.00	0.06	−0.43	−0.26
LZSK	0.19	0.02	8.53	–	–	–	–	–	–	–	–	–	–	1.00	0.60	0.41
LZPK	0.01	1.50×10^{-3}	12.65	–	–	–	–	–	–	–	–	–	–	–	1.00	0.56
PFREE	0.10	6.80×10^{-3}	6.82	–	–	–	–	–	–	–	–	–	–	–	–	1.00
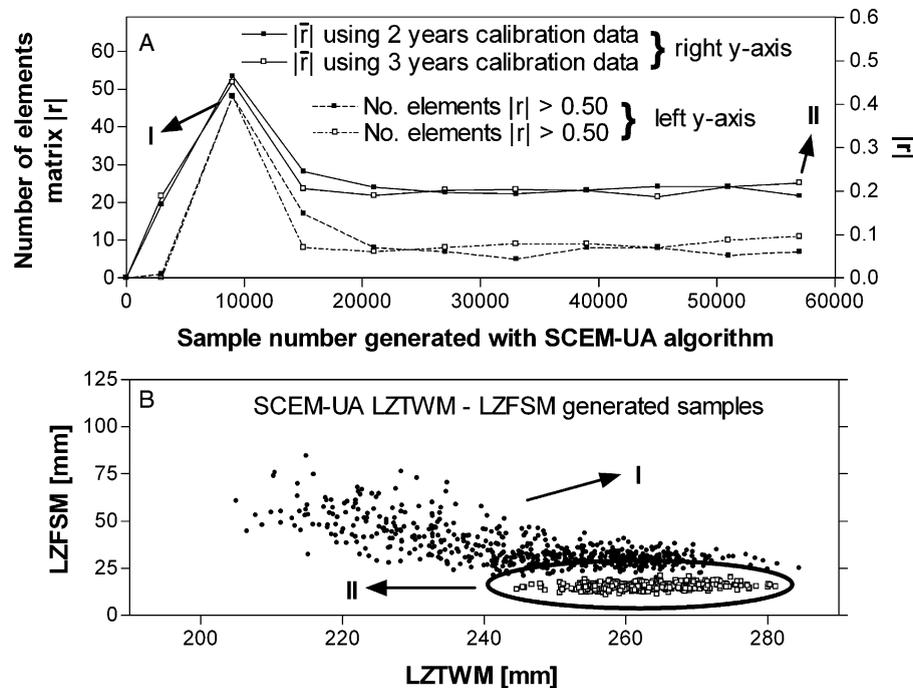
Fig. 3. (A) Evolution of the mean of the absolute values in the parameter correlation matrix (on right y-axis), as function of the number of samples generated with the SCEM-UA algorithm using 2 and 3 years of streamflow data. For completeness, also included are the number of parameter combinations in the parameter correlation matrix (on left y-axis) that are correlated higher than the specified threshold value of 0.50; (B) two-dimensional scatterplot of LZTWM–LZFSM sampled parameter space with the SCEM-UA algorithm, using sample numbers 6.000–12.000 (black dots) and sample numbers 54.000–60.000 (open squares).

Jakeman and Hornberger (1993) have used this ARPE measure, derived from the first-order covariance matrix, to estimate how many parameters are needed to describe the transformation from mean areal precipitation to streamflow emanating from the catchments. The results presented in Table 3, illustrate that a significant improvement in performance can be achieved by using the more complex SAC-SMA model. When only a few years of streamflow data are available for model calibration, the optimum model complexity would be a simple two-parameter rainfall-excess model connected with two series of linear reservoirs. This supports the statements made by Beven (1989, p.159), that '…it appears that 3–5 parameters should be sufficient to reproduce most of the information in a hydrological record' and Jakeman and Hornberger (1993), who based on investigations of 1 year of daily streamflow data, argued that simple RR models with four to five parameters provide an adequate fit to the streamflow

Table 3
ARPE identification statistics for the parsimonious HYMOD and more complex SAC-SMA conceptual rainfall-runoff models, as function of the length of the calibration set in years

Length calibration set Years	HYMOD		SAC-SMA	
	ARPE (%)	RMSE (m^3/s)	ARPE (%)	RMSE (m^3/s)
1	28.91	13.46	5.57	9.80
2	0.66	16.16	1.25	12.01
3	0.59	21.92	0.77	15.34
4	0.42	19.99	0.42	15.28
5	0.46	21.05	0.48	16.89
6	0.37	19.95	0.56	18.08
7	0.80	20.00	0.72	18.34
8	0.78	25.18	0.59	19.12
9	3.32	27.32	0.44	19.65
10	6.67	26.33	0.61	19.40

Also included are the RMSE statistics of the residuals over the respective calibration period. For more information, please refer to the text.

data and that the addition of more model structure and its associated parameters leads to no significant improvement in fit yet introduces poorly identified parameters. However, with the availability of more data, the ARPE values show that larger conceptual configurations, like the SAC-SMA model, are warranted by the streamflow data. Note that the ARPE rapidly climbs for HYMOD as the calibration hits 9–10 years, indicating inconsistent information about the HYMOD model parameters with increasing length of the calibration time series. This is primarily caused by an extreme large rainfall event in calibration year 9 resulting in an average daily streamflow of about 1300 m^3/s, which is by far the largest event on record. Due to this rainfall event the uncertainty of the HYMOD parameters significantly increased compared to situations in which the parameters were fitted without this extreme rainfall-runoff event. Although not explicitly demonstrated here, comparison of the SCEM-UA estimated ARPE values with the theoretical ARPE values derived from the first-order covariance matrix, evaluated in the vicinity of the global optimum in the parameter space, revealed that erroneous conclusions can be drawn

about the ideal model complexity using the first-order ARPE values, when the underlying assumptions of model linearity and a Gaussian posterior density of the parameters are violated.

3.6. SAC-SMA prediction uncertainty ranges

Fig. 4A and B present the residuals from the most probable parameter set and the hydrograph prediction uncertainty intervals for the SAC-SMA simulated streamflows associated with the posterior parameter estimates (dark-gray region) and the residual model uncertainty (light-gray region), for a portion of the wet calibration year 1953.

The solid circles correspond to the observed streamflow data. The model uncertainty is computed by adding the model error to the SAC-SMA model prediction. Note that the streamflow prediction uncertainty ranges (light-gray) bracket the observed flows during almost the entire period, but are quite large. Further, the prediction uncertainty associated with the posterior parameter estimates (dark-gray) does not include the observations and displays systematic error on the long recessions. This indicates
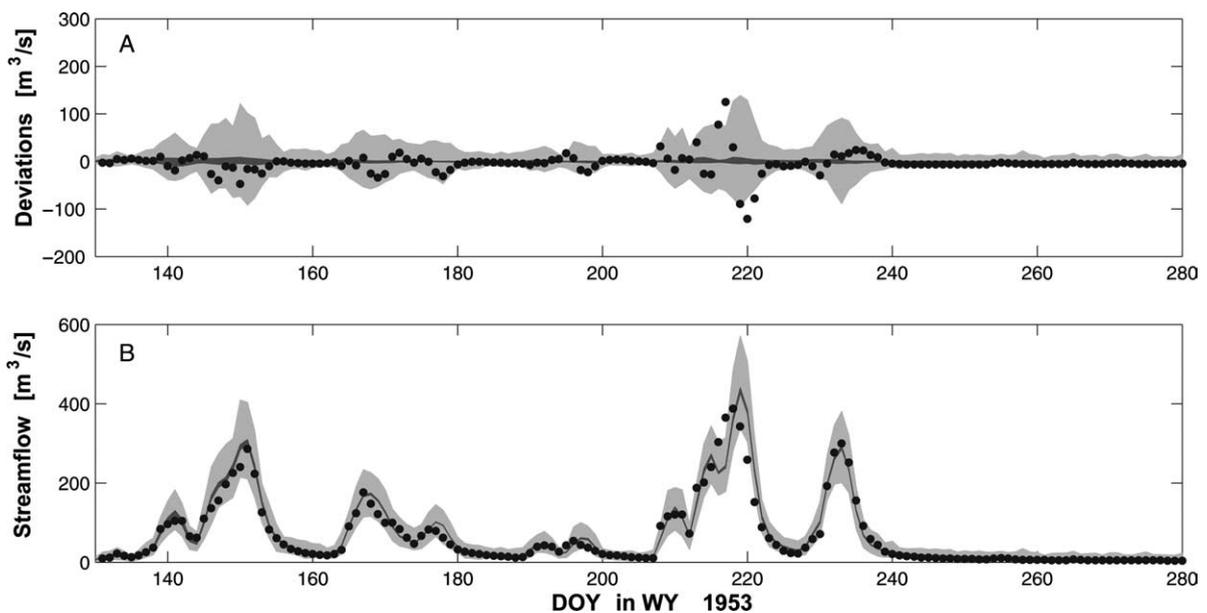

Fig. 4. (A) Streamflow uncertainties associated with the most probable parameter set derived using the SCEM-UA algorithm. The light-gray region denotes model uncertainty, whereas parameter uncertainty is indicated with the dark-gray region. The dots correspond to the observed streamflow data; (B) hydrograph prediction uncertainty associated with the uncertainty in the model (light-gray) and parameter estimates (dark-gray region) for the WY 1953.

that the estimation procedure is placing too much confidence in the validity of the SAC-SMA model. In another paper (Vrugt et al., 2005a,b), we have recently presented a combined parameter and state estimation method that provides a better treatment of model errors, and therefore results in model prediction uncertainty bounds that are more reasonable and tend to bracket the observations. Nevertheless, the results presented here, indicate that for this particular watershed and dataset, it is possible to reasonably identify values for most of the SAC-SMA parameters, using a single objective function.

3.7. Consistency evaluation pdf parameter values

To verify the consistency and reliability of the calibration results, the performance for each of the samples in the HPD region for each of the 10 calibration runs, reported in the previous sections, were evaluated for the remaining 26 years of data not included in the calibration set. In this particular instance, the cross-validation is not based on a single 'best' parameter set, but is based on an ensemble of parameter sets, each having a different likelihood. Fig. 5 summarizes the results of this analysis for the case of using 2, 4, 6, 8, and 10 years of calibration data, in terms of normalized probability distributions of the root mean square error (RMSE), and percent bias (BIAS) statistics of the residuals for the evaluation period.

The two important results are as follows:

1. There is a clear improvement in average model performance measured in terms of RMSE statistic of the residuals with the use of longer time series for calibration purposes. Notice, however, that the dispersion around the mean RMSE is significant, illustrating that even with the use of relatively short calibration sets (2 years), solutions are found in the HPD region, that generate very similar forecasts in terms of RMSE value as parameter solutions obtained using 8–10 years of calibration data.
2. The improvement in model performance measured in terms of percentage BIAS can be considered marginal with the use of longer datasets; the BIAS criterion is rather insensitive to the length and thus variability of the dataset.
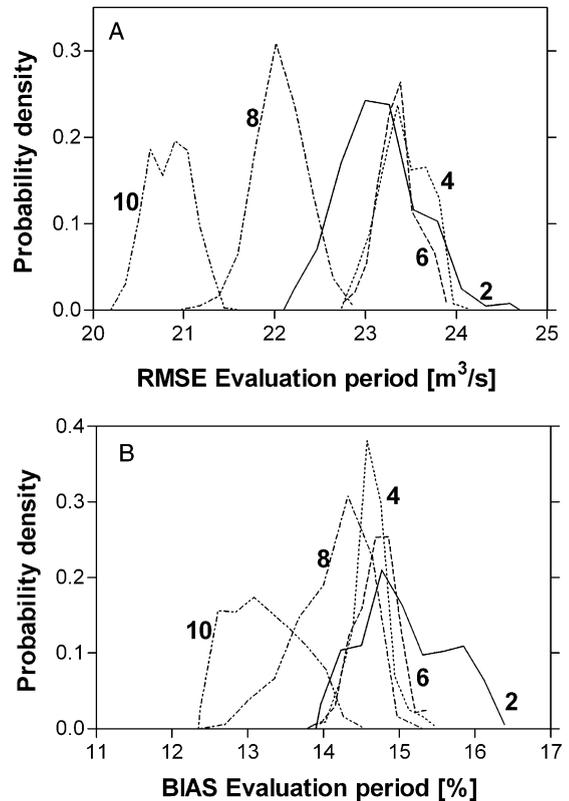

Fig. 5. Normalized probability density functions of the (A); root mean square error (RMSE), and, (B) percent bias (BIAS) statistics over the evaluation period, corresponding to the samples in the HPD region of the posterior distribution using 2, 4, 6, 8, and 10 years of calibration data.

The results presented here indicate that the use of longer calibration sets does not reduce the uncertainty associated with the final parameter estimates as demonstrated in Fig. 1, but improves the average performance (and thus consistency) of the SAC-SMA model structure during the evaluation period. This is not surprising, as longer calibration time series generate more reliable parameter estimates, as a larger variety of hydrologic events are presented to the model during the calibration phase. Consequently, the parameter sets obtained in this way extrapolate better during the independent evaluation period.

3.8. Stationarity of the SAC-SMA model parameters

As the SCEM-UA algorithmic procedure success-fully infers the underlying posterior distribution of the

model parameters, the method is suited to investigate whether the SAC-SMA parameters are statistically stationary over the entire historical record of the Leaf River dataset or if they are correlated to varying characteristics of the watershed. Based on our earlier results presented in Fig. 1, we conducted a traditional split sample test to divide the entire 36 years of measurements available for the Leaf River watershed into calibration sets consisting of 6 chronologically consecutive water years for the SAC-SMA model. Hence, there are 30 possible calibration sets constituting 6 consecutive water-years (using WY 1–6, 2–7, 3–8, and so forth up to 30–36). In each calibration run with the SCEM-UA algorithm, a 65-days warm-up period was used to minimize initialisation errors of state variables.

The evolution of the marginal HPD regions for each of the SAC-SMA parameters in normalized parameter space is illustrated in Fig. 6.

The gray-shaded area in each parameter plot represents the HPD region, whereas the marked squares and dotted line refer to the SCE-UA solution, and most likely SCEM-UA solution within the HPD region, respectively. The results presented in Fig. 6 denote averages over a window of 6 years (1953 denotes calibration results over 1953–1959, respectively and so forth). While the SAC-SMA parameters UZFWM, PCTIM, ZPERC, LZTWM, LZFSM, and LZFPM show little variation over the 36-year historical data record, the HPD region for the other SAC-SMA parameters traverses through the feasible parameter space. Especially, there is considerable
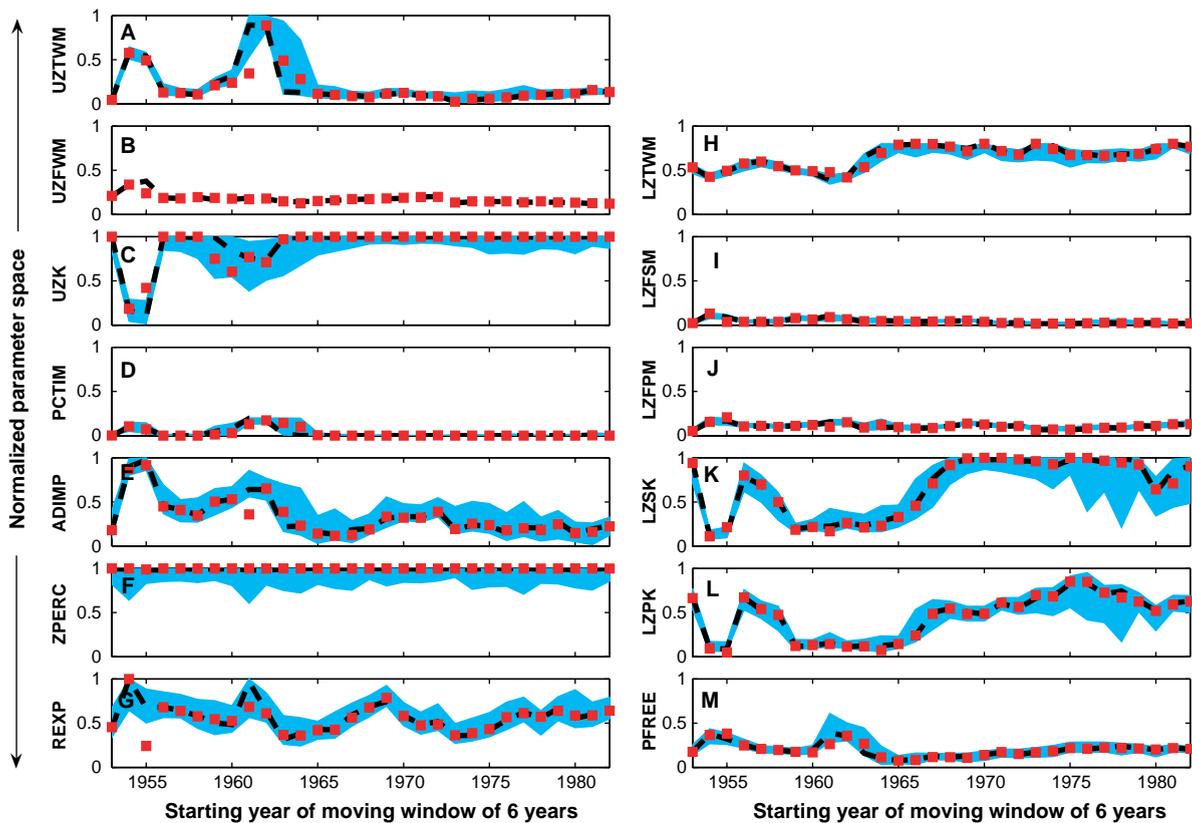

Fig. 6. Evolution of the 95% confidence intervals for each of the SAC-SMA parameters over the 36-year historical record of the Leaf River watershed. The gray shaded area in each parameter plot represents the HPD region, whereas the marked squares and dotted line refer to the SCE-UA solution and most likely SCEM-UA solution, respectively. Results denote averages over a window of 6 years.
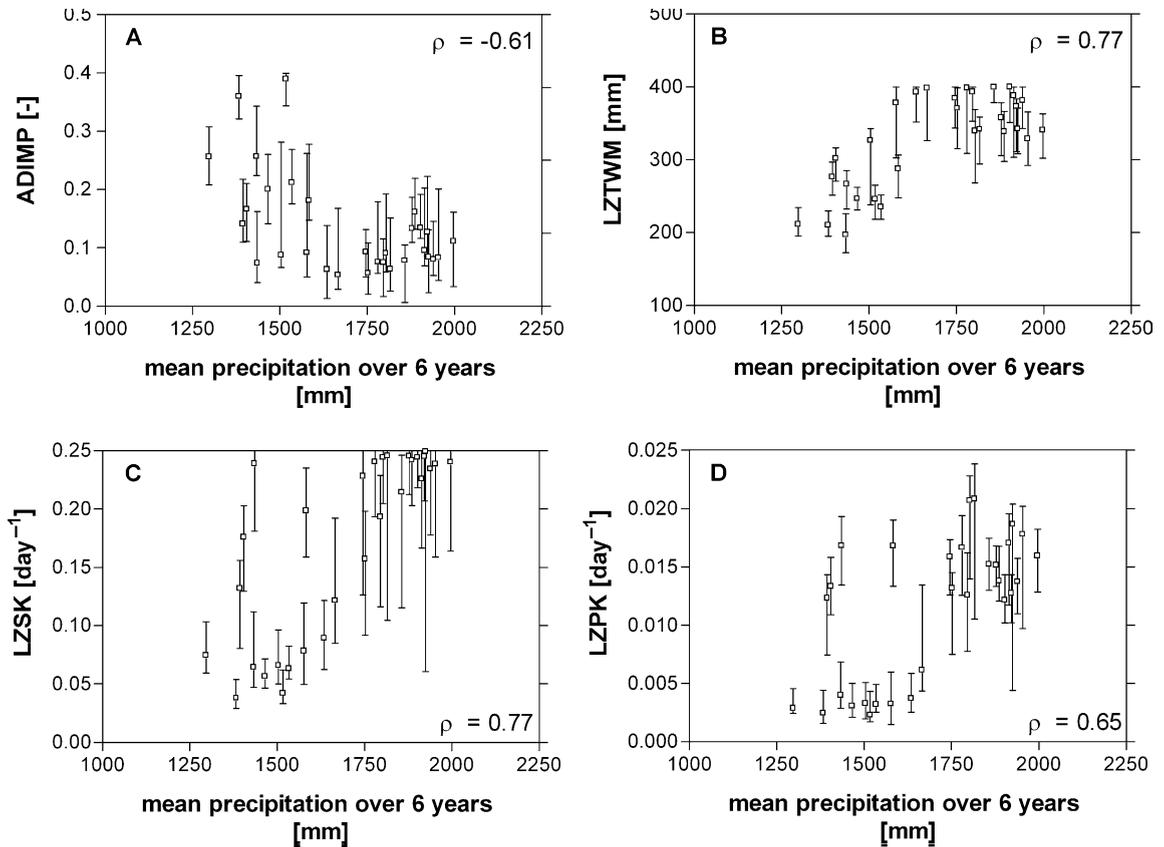
Fig. 7. Two-dimensional plot of the most likely parameter value, indicated with a squared symbol, versus the mean areal rainfall for the SAC-SMA parameters ADIMP (A), LZTWM (B), LZSK (C), and LZPK (D). The bars around the most likely parameter value denote the size of the HPD region.

variation and uncertainty associated with the parameters UZK, LZSK, and LZPK, which primarily determine the shape of the hydrograph during the recession periods, and the percolation parameters REXP and PFREE. The apparent systematic variation of the parameters ADIMP, LZSK and LZPK with time, might suggest that the watershed has undergone hydrologic changes. When some of the parameters in the SAC-SMA model are plotted against the mean areal rainfall over the calibration periods, as done in Fig. 7, a relationship becomes apparent.

To be able to match the observed hydrograph with increasing wetness of the years, the additional impervious fraction is decreased (ADIMP), while the maximum capacity of the lower zone tension water storage (LZTWM) and depletion rate from the lower zone need to be increased. Seemingly,

parameters calibrated for relatively dry years, result in sub-optimal forecasts for the wettest years on record and vice versa. This non-stationarity with increasing wetness of the years for some of the SAC-SMA parameters, point towards aspects of the model structure that needs to be further refined. Similar issues have previously been reported by Gan and Burges (1990), who used the SAC-SMA model in combination with a Nelder–Mead optimization scheme.

3.9. Which hydrologic years to use for model calibration?

The final analysis presented in this paper focuses on the relationship between the characteristics of a hydrologic year and the Bayesian confidence intervals
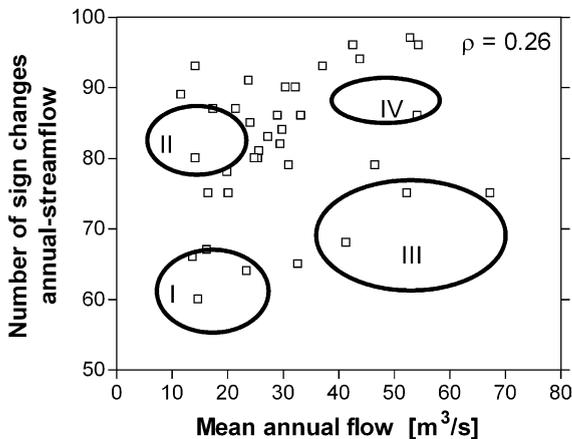
Fig. 8. Two-dimensional bi-criterion plot of the mean annual flow versus the number of sign changes of the measured annual streamflow for each of the 36 years of the historical record. The four different clusters are indicated with a circle.

of the SAC-SMA parameter estimates. For this, the 36 years of available hydrologic data were clustered into four different groups using two different measures for each year, the mean annual flow and the number of sign changes of the measured streamflow (see Fig. 8).

The first criterion, measures the wetness of the year, while the second criterion characterises the variability of the hydrologic year. Among several other criteria, including median and standard deviation of annual flow, these criteria were found to be most uncorrelated. Hence, it is to be expected that the wettest years on record, which simultaneously also exhibit a large variability in streamflows, contain the most information for the parameters, as these type of years should activate the different modes of model operation the most. Table 4 summarizes the results for these calibration runs with the SCEM-UA algorithm for each of the four clusters of years distinguished in Fig. 8.

The important result demonstrated in Table 4 is that none of the identified clusters of hydrologic data is superior in terms of the final identifiability of the parameters, as measured with the standard deviation and CV-values of the parameter estimates. However, the use of the wettest hydrologic years on record simultaneously exhibiting a high intra-annual variability in streamflows for calibration purposes (cluster IV), does increase the average performance of the parameter estimates in the posterior distribution over

the evaluation period (in terms of RMSE and BIAS statistics), thereby confirming the results presented in earlier work (Yapo et al., 1996).

3.10. Efficiency comparison of SCEM-UA with original SCE-UA algorithm

While, the SCE-UA algorithm developed by Duan et al. (1992) converges to a single 'best' solution in the feasible parameter space, the SCEM-UA algorithm converges to a distribution of parameter sets within the HPD region of the parameter space, thereby containing the most optimal (SCE) solution. This feature enables hydrologists to generate consistent model predictions, along with estimates of the underlying model and parameter uncertainty as depicted in Fig. 4. To further illustrate the efficiency of the SCEM-UA algorithm in searching the feasible parameter space for candidate solutions, Fig. 9 presents the evolution of the 'best' parameter set, measured in terms of RMSE of the non-transformed flow space, for the SCE-UA and SCEM-UA algorithms denoted with the solid and dotted line, respectively, using 2 (location A), 5 (location B) and 10 (location C) years of calibration data.

Notice, that the SCE-UA and SCEM-UA algorithm need a nearly identical number of model simulations (approximately 7.500) for identifying the most optimal solution in the high-dimensional parameter space. The additional runs (>7.500) performed with the SCEM-UA algorithm, are needed to construct a large enough sample from which correct kernel density estimates of the HPD region can be estimated.

4. Summary and discussion

Non-linear structural equations (including threshold type discontinuities) in conceptual watershed models have historically posed a challenge to the application of computer-based systems methods for automated parameter estimation. Progress in stochastic optimization has helped to diminish the difficulties associated with parameter estimation. This paper has demonstrated the applicability and usefulness of the Shuffled Complex Evolution Metropolis (SCEM-UA) global optimization algorithm for stochastic optimization of the parameters in the Sacramento Soil

Table 4
Posterior mean, standard deviation, and coefficient of variation (CV (%)), of the SAC-SMA model parameters, for each of the identified clusters of hydrologic years for the Leaf River watershed

Parameter	Cluster I			Cluster II			Cluster III			Cluster IV		
	Mean	Standard deviation	CV	Mean	Standard deviation	CV	Mean	Standard deviation	CV	Mean	Standard deviation	CV
UZTWM	98.23	0.93	0.94	18.85	0.25	1.36	15.39	0.60	3.78	22.45	1.07	4.77
UZFWM	45.77	0.70	1.53	26.92	0.46	1.71	23.92	0.41	1.67	34.33	1.14	3.31
UZK	0.33	1.1×10^{-2}	3.37	0.49	2.8×10^{-2}	0.57	0.49	6.3×10^{-3}	1.26	0.50	2.2×10^{-3}	0.44
PCTIM	1.4×10^{-2}	8.5×10^{-4}	5.87	6.7×10^{-5}	4.1×10^{-5}	61.28	1.94×10^{-4}	1.2×10^{-4}	67.53	4.0×10^{-5}	2.3×10^{-5}	58.00
ADIMP	0.19	1.1×10^{-2}	5.53	8.7×10^{-3}	4.2×10^{-3}	48.79	0.12	8.7×10^{-3}	7.40	0.21	8.2×10^{-3}	3.97
ZPERC	236.54	6.09	2.58	243.72	4.34	1.79	239.47	5.93	2.35	239.40	5.41	2.26
REXP	3.33	0.15	4.50	4.14	0.27	6.63	2.97	0.11	3.79	3.61	0.12	3.44
LZTWM	189.91	2.17	1.14	337.35	2.71	0.80	320.43	4.37	1.35	360.61	6.18	1.71
LZFSM	78.73	3.41	4.33	115.36	10.35	8.97	32.87	1.78	5.33	48.12	2.74	5.70
LZFPM	118.46	4.75	4.02	176.26	3.92	2.22	100.61	2.73	2.71	166.08	4.83	2.91
LZSK	5.2×10^{-2}	1.7×10^{-3}	3.35	7.7×10^{-2}	2.8×10^{-3}	3.73	0.20	8.6×10^{-3}	3.99	0.24	5.4×10^{-3}	2.27
LZPK	1.8×10^{-3}	1.5×10^{-4}	8.38	2.1×10^{-3}	1.33×10^{-4}	6.37	1.5×10^{-2}	4.8×10^{-4}	3.45	1.3×10^{-2}	3.0×10^{-4}	2.33
PFREE	0.18	4.3×10^{-3}	2.37	3.8×10^{-2}	4.0×10^{-3}	10.61	0.11	3.9×10^{-3}	3.64	0.10	4.7×10^{-3}	4.65
RMSE		22.19			20.10			19.30			17.99	
BIAS		11.05			6.24			8.61			5.82	
$|\bar{r}|$		0.21			0.25			0.21			0.18	

Also included are the RMSE (m^3/s) and BIAS (%) statistics of the most optimal parameter set, found by the SCEM-UA algorithm for each of the clusters, when evaluated over the independent 26-year evaluation period (WY, 1961–1988), and the average of the absolute values in the parameter correlation matrix, $|\bar{r}|$.
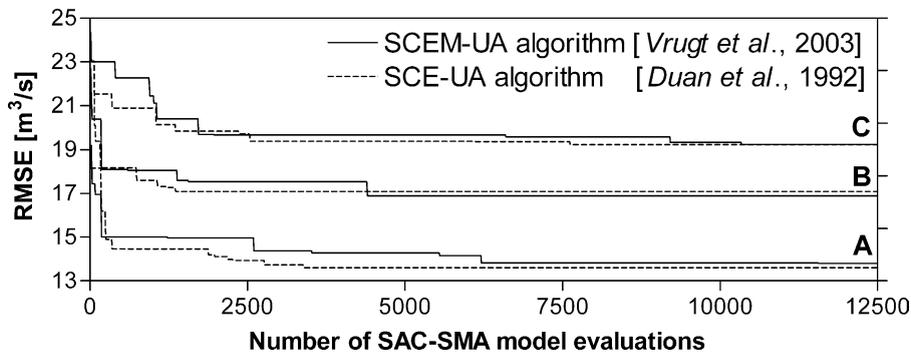
Fig. 9. Evolution of the 'best' parameter set, measured in terms of RMSE of the non-transformed flow for the SCE-UA and SCEM-UA algorithms using 2 (A), 5 (B) and 10 (C) years of calibration data.

Moisture Accounting model. The SCEM-UA algorithm is a Markov Chain Monte Carlo sampler that provides an efficient estimate of the most-likely parameter set and its underlying distribution within a single optimization run. The algorithm was implemented using a LAM/MPI distributed processing interface and Octave programming environment to maximize computing efficiency, using 25 Pentium IV 3.40 GHz processing nodes.

To demonstrate the usefulness and applicability of the SCEM-UA algorithm for watershed model calibration, we performed a variety of case studies. In particular, we investigated the uncertainty and stationarity of the SAC-SMA model parameters as function of the length and variability of the streamflow data. In addition, in each considered case study a comparison was made between SCEM-UA derived parameter values and those obtained using a deterministic SCE-UA calibration. The results may be summarized as follows:

(1) The location and size of the HPD region of the posterior probability distribution appears to be relatively unaffected by the length of the calibration set. It seems that 2–3 years of streamflow data are sufficient to obtain stable estimates for most of the SAC-SMA model parameters. Nevertheless, longer calibration data sets increase the average performance of the SAC-SMA model structure during the evaluation period.

(2) No systematic relationship was found between the characteristics of the hydrologic year and the final identifiability of the SAC-SMA model parameters. However, the use of the wettest years on records simultaneously exhibiting a high intra-annual variability in streamflows increases the average performance of the parameter estimates in the posterior distribution over the evaluation period

(3) Irrespective of the data period used, most of the SAC-SMA parameters are well defined by calibration to streamflow data. In particular, the capacity parameters are precisely determined, while parameters that control percolation and recession are less well determined. These results suggest that much of the uncertainty in the model performance is associated with percolation from the upper zone and depletion from the lower zone.

(4) There is considerable time variation and uncertainty associated with the percolation parameters in the SAC-SMA model and the parameters that determine the shape of the hydrograph during the recession periods. This significant parameter nonstationarity demonstrates that improvements can be made to the SAC-SMA model structure. For instance, when some of the SAC-SMA parameter values were plotted against the mean areal rainfall over the calibration periods, a relationship became apparent.

The results presented in this paper indicate that the prediction uncertainty associated with the parameter estimates is typically small compared to the hydrograph prediction uncertainty ranges associated with the model uncertainty, indicating that the main part of the uncertainty stems from the residuals between model predictions and observations. To be able to

better understand the limitations of our models and improve our understanding and theory of hydrologic processes, we need to develop strategies, which can better distinguish between input, output, parameter, and model structural uncertainty. The SODA framework presented in our most recent work (Vrugt et al., 2005a,b) has been designed to facilitate this task. Software used in this and related work can be found at http://www.science.uva.nl/ibed/cbpg/software/.

Acknowledgements

References

Beven, K.J., 1989. Changing ideas in hydrology—the case of physically based model. J. Hydrol. 105, 157–172.

Beven, K.J., Binley, A.M., 1992. The future of distributed models: model calibration and predictive uncertainty. Hydrol. Process. 6, 279–298.

Box, G.E.P., Cox, D., 1974. An analysis of transformations. J. R. Stat. Soc. 26, 211–252.

Boyle, D.P., 2000. Multicriteria calibration of hydrological models. PhD Dissertation, Department Hydrology Water Resources, University of Arizona, Tucson, AZ.

Boyle, D.P., Gupta, H.V., Sorooshian, S., 2000. Toward improved calibration of hydrological models: combining the strengths of manual and automatic methods. Water Resour. Res. 36 (12), 3663–3674.

Brazil, L.E., Hudlow, M.D., 1981. Calibration procedures used with the national weather service forecast system. In: Haimes, Y.Y., Kindler, J. (Eds.), Water and Related Land Resources. Pergamon, New York, pp. 457–466.

Burnash, R.J.C., Ferral, R.L., McGuire, R.A., 1973. A generalized streamflow simulation system: conceptual models for digital computers. Joint Federal-State River Forecast Center, Sacramento, CA.

Duan, Q., Gupta, V.K., Sorooshian, S., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28, 1015–1031.

Gan, T.Y., Burges, S.J., 1990. An assessment of a conceptual rainfall-runoff model's ability to represent the dynamics of small hypothetical catchments, 2: hydrologic responses for normal and extreme rainfall. Water Resour. Res. 26 (7), 1605–1619.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7, 457–472.

Gupta, V.K., Sorooshian, S., 1985. The relationship between data and the precision of parameter estimates of hydrological models. J. Hydrol. 81, 57–77.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Towards improved calibration of hydrologic models: multiple and non-commensurable measures of information. Water Resour. Res. 34, 751–763.

Gupta, H.V., Bastidas, L., Vrugt, J.A., Sorooshian, S., 2002. Multiple criteria global optimisation for watershed model calibration. In: Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte R., (Eds.), AGU Monograph: Advances in Automatic Calibration of Watershed Models.

Hastings, W.K., 1970. Monte-Carlo sampling methods using Markov Chains and their applications. Biometrika 57, 97–109.

Hooper, R.P., Stone, A., Christophersen, N., de Grosbois, E., Seip, H.M., 1988. Assessing the Birkenes model of stream acidification using a multisignal calibration methodology. Water Resour. Res. 24, 1308–1316.

Jakeman, A., Hornberger, G.M., 1993. How much complexity is warranted in a rainfall-runoff model? Water Resour. Res. 29 (8), 2637–2649.

Jakeman, A.J., Dietrich, C.R., Thomas, G.A., 1989. Solute transport in a stream aquifer system, 2, application of model identification to the river Murray. Water Resour. Res. 25, 2177–2185.

Jakeman, A.J., Littlewood, I.G., Whitehead, P.G., 1990. Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. J. Hydrol. 117, 275–300.

Kitanidis, P.K., Bras, R.L., 1980a. Adaptive filtering through detection of isolated transient errors in rainfall-runoff models. Water Resour. Res. 16 (4), 740–748.

Kitanidis, P.K., Bras, R.L., 1980b. Real-time forecasting with a conceptual hydrological model, 1, analysis of uncertainty. Water Resour. Res. 16 (6), 1025–1033.

Kuczera, G., 1997. Efficient subspace probabilistic parameter optimization for catchment models. Water Resour. Res. 33 (1), 177–185.

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

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equations of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1091.

Misirli, F., 2003. Improving efficiency and effectiveness of Bayesian recursive parameter estimation for hydrological models. PhD Dissertation, Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ.

Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. Hydrol. Sci. J. 30 (2), 273–297.

Peck, E.L., 1976. Catchment modeling and initial parameter estimation for the national weather service river forecast system, NOAA Technology Memorandrom, NWS HYDRO-31. National Weather Service, Silver Spring, MD.

Sorooshian, S., Gupta, V.K., Fulton, J.L., 1983. Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: influence of calibration data variability and length on model credibility. Water Resour. Res. 19 (1), 251–259.

Sorooshian, S., Duan, Q., Gupta, V.K., 1993. Calibration of rainfall-runoff models: application of global optimization to the Sacramento Soil Moisture Accounting model. Water Resour. Res. 29, 1185–1194.

Thiemann, M., Trosset, M., Gupta, H., Sorooshian, S., 2001. Bayesian recursive parameter estimation for hydrological models. Water Resour. Res. 37 (10), 2521–2535.

Vrugt, J.A., Gupta, H.V., Bouten, W., Sorooshian, S., 2003. 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.

Vrugt, J.A., Diks, C.G.H., Gupta, H.V., Bouten, W., Verstraten, J.M., 2005a. Improved treatment of uncertainty in hydrologic modelling. combining the strengths of global optimization and data assimilation. Water Resour. Res. 41 (1), W01017. doi:10.1029/2004WR003059.

Vrugt, J.A., Robinson, B.A., Vesselinov, V.V., 2005b. Improved inverse modelling for transport in subsurface media: combined parameter and state estimation. Geophys. Res. Let., L18408. doi:10.1029/2005GL023940.

Vrugt, J.A., Ó Nualláin, B., Robinson, B.A., Bouten, W., Dekker, S.C., Sloot, P.M.A., in press. Application of parallel computing to stochastic parameter estimation in environmental models. Computers & Geosciences.

Yapo, P.O., Gupta, H.V., Sorooshian, S., 1996. Calibration of conceptual rainfall-runoff models: sensitivity to calibration data. J. Hydrol. 181, 23–48.