



available at [www.sciencedirect.com](http://www.sciencedirect.com)



journal homepage: [www.elsevier.com/locate/jhydrol](http://www.elsevier.com/locate/jhydrol)



# Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model

Luc Feyen <sup>a,\*</sup>, Jasper A. Vrugt <sup>b</sup>, Breannán Ó Nualláin <sup>c</sup>,  
Johan van der Knijff <sup>a</sup>, Ad De Roo <sup>a</sup>

<sup>a</sup> Land Management and Natural Hazards Unit, Institute for Environment and Sustainability, DG Joint Research Centre, European Commission, Via E. Fermi 1, TP 261, 21020 Ispra, Varese, Italy

<sup>b</sup> Earth and Environmental Sciences Division, Los Alamos National Laboratory, United States

<sup>c</sup> Applied Logic Laboratory, Institute for Logic Language and Computation, Universiteit van Amsterdam, The Netherlands

Received 22 September 2005; received in revised form 3 July 2006; accepted 6 July 2006

## KEYWORDS

Distributed modelling;  
Automatic calibration;  
Parameter uncertainty;  
Markov chain Monte Carlo  
methods

**Summary** This work addresses the calibration of the distributed rainfall-runoff model LISFLOOD and, in particular, the realistic quantification of parameter uncertainty and its effect on the prediction of river discharges for large European catchments. LISFLOOD is driven by meteorological input data and simulates river discharge in large drainage basins as a function of spatial information on topography, soils and land cover. Even though LISFLOOD is physically based to a certain extent, some processes are only represented in a lumped conceptual way. As a result, some parameters lack physical basis and cannot be directly inferred from quantities that can be measured. In the current LISFLOOD version five parameters need to be determined by calibration. We employ the Shuffled Complex Evolution Metropolis (SCEM-UA) global optimisation algorithm to automatically calibrate the model against daily discharge observations. The resulting posterior parameter distribution reflects the uncertainty about the model parameters after taking into account the discharge observations, and forms the basis for making probabilistic flow predictions. To overcome the computational burden the optimisation has been implemented using parallel computing. As an illustrative example, we demonstrate the methodology for the Meuse catchment upstream of Borgharen, covering approximately 21,000 km<sup>2</sup>. Results demonstrate the capabilities of the SCEM-UA algorithm to efficiently evolve to the target posterior distribution and to identify, except for the lower groundwater zone time constant,

\* Corresponding author. Tel.: +39 332 789 258; fax: +39 332 786 653.  
E-mail address: [luc.feyen@jrc.it](mailto:luc.feyen@jrc.it) (L. Feyen).

the LISFLOOD calibration parameters using daily discharge observations. It should be noted that the posterior parameter distributions are based on the assumption of independent errors. Analysis of the residuals revealed there is a strong evidence to reject this assumption. This will impact on the parameter posterior distributions and also lead to underestimation of the prediction limits.

© 2006 Elsevier B.V. All rights reserved.

## Introduction

During the past decade, several catastrophic floods have occurred in Central Europe (including Odra, Rhine, Danube, Elbe and Vistula), causing several fatalities and huge damages. Within this regard, the Joint Research Centre of the European Commission is developing the European Flood Alert System (EFAS). The aim of EFAS is to provide the national hydrological services with continuous real-time river discharge forecasts across Europe with a lead-time of 3–10 days, complementary to the forecasting at national level typically done for the first day(s). The system currently runs on a pre-operational basis using a grid resolution of 5 km for all European river basins larger than 2000 km<sup>2</sup>. The predictions of river discharge are obtained with LISFLOOD, a distributed, (partly) physically based rainfall-runoff model that was designed to simulate rainfall-runoff processes in large catchments (De Roo et al., 2000, 2001). The accuracy of the real-time forecasts depend on the ability of the model to capture the dominate hydrological processes that transfer precipitation into river runoff at the catchment scale, and on its ability to reproduce historical time series of observed river discharges.

A crucial step which contributes significantly to the accuracy of the EFAS discharge forecasts is the calibration of the LISFLOOD model for all European catchments. Owing to the general nature of the LISFLOOD model its application to any given river basin requires that certain parameters of conceptual functions be identified for the particular basin. In the process of calibration, the values of unknown model parameters are tuned such that the model matches the observed predictions as closely as possible. This can be done by manually adjusting the parameters while visually inspecting the agreement between the observed and simulated discharges. However, the subjective and time-consuming nature of the trial-and-error method renders this method unappealing for use on a European scale. The large number of catchments for which the model needs to be calibrated within EFAS calls for an automatic parameter estimation procedure. Besides shortening the implementation time this will also enhance the reliability of the calibrated parameters due to a more exhaustive exploration of the parameter space.

The application of automatic parameter estimation techniques has received considerable attention over the last decades (e.g., Sorooshian and Dracup, 1980; Kuczera, 1983; Duan et al., 1992; Thyer et al., 1999; Vrugt et al., 2003). Several studies have reported difficulties in obtaining unique global parameter estimates because of the presence of multiple local optima, non-linear interaction between model parameters, and the shape and roughness of the response surface defined by the selected objective function. This

has led to the development of global optimisation techniques such as the Shuffled Complex Evolution method (Duan et al., 1992) and simulated annealing (Sumner et al., 1997).

Often, the optimal parameter estimates are used without taking due account of the uncertainty in the parameter estimates in subsequent predictions. However, without a realistic assessment of parameter uncertainty it is not possible to undertake with any confidence tasks such as evaluating prediction/confidence limits on future hydrological responses, assessing the significance of deviations in split-sample tests, and assessing the value of regional relationships between model parameters and catchment characteristics (Kuczera and Parent, 1998).

Several approaches have been developed to assess parameter uncertainty and its effect on subsequent predictions. These include the use of multinormal approximations (Kuczera and Mroczkowski, 1998), simple uniform random sampling (URS) over the feasible parameter space (Uhlenbrook et al., 1999), parametric bootstrapping and Markov chain Monte Carlo (MCMC) methods (Kuczera and Parent, 1998; Campbell et al., 1999; Bates and Campbell, 2001; Vrugt et al., 2003). Traditional statistical theory based on first-order approximations and multinormal distributions is typically unable to cope with the non-linearity of complex hydrological models. The URS method allows easy exploration of the parameter space but is computationally inefficient. Unless a large number of random samples are drawn from the highly dimensioned parameter space, uniform sampling can produce seriously misleading results. MCMC methods on the other hand generate samples from a Markov chain that adapts to the stationary posterior parameter distribution, and can be applied to complex inference, search and optimisation problems. The posterior parameter distribution quantifies the uncertainty about the model parameters after considering the observed catchment responses, and forms the basis for making probabilistic predictions.

Other methods have been developed that focus on assessing global uncertainty in rainfall-runoff modelling, such as the generalised likelihood uncertainty (GLUE) method (Beven and Binley, 1992) and the meta-Gaussian approach (Krzysztofowicz and Kelly, 2000; Montanari and Brath, 2004). These methods estimate the aggregated model uncertainty without attempting to separate the individual effects of input, parameter and model uncertainty. We specifically focus on estimating confidence limits for the rainfall-runoff model parameters to quantify the effect of parameter uncertainty on river flow simulation.

In this work, we employ the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm (Vrugt et al., 2003) to infer the posterior distributions for the parameters of the LISFLOOD model. The SCEM-UA algorithm is a modified version of the original SCE-UA global optimisation algorithm (Duan

et al., 1992). The algorithm is Bayesian in nature and operates by merging the strengths of the Metropolis algorithm, controlled random search, competitive evolution, and complex shuffling to continuously update the proposal distribution and evolve the sampler to the posterior target distribution. The SCEM-UA algorithm has been successfully applied to calibrate the five-parameter conceptual rainfall-runoff model HYMOD for a 1944 km<sup>2</sup> watershed (Vrugt et al., 2003) and to estimate vadose zone properties for a small watershed using the distributed fully coupled surface – vadose zone – groundwater model MODHMS (Vrugt et al., 2004).

LISFLOOD is a hybrid between a conceptual and fully physically based distributed rainfall-runoff model that simulates the spatial and temporal patterns of catchment responses in large (European) river basins. These type of models have the potential to evaluate the effects of land use or climate changes on catchment behaviour or to meet the request for more accurate spatial predictions. To date, the capabilities of such models have not yet been fully exploited, primarily due to their computational, distributed input and parameter estimation requirements. However, the rapid increase in computer power, the development of efficient computational methods, the availability of sophisticated Geographical Information Systems, and the increasing wealth of spatial data of different types facilitate the proper calibration of these models, and consequently their use in making reliable predictions. The aim of this paper is to explore the capacity of SCEM-UA to identify the posterior parameter distributions for the LISFLOOD model applied to a large catchment (>20,000 km<sup>2</sup>), and to evaluate the effect of parameter uncertainty on the predictive capabilities of the model. To handle the significant CPU time needed to perform the large number of model iterations the SCEM-UA method has been implemented using parallel computing.

The paper is organised as follows. Section ‘‘Methodology’’ presents details on the Bayesian inverse methodology and the SCEM-UA algorithm used to infer the posterior model parameter distributions. Sections ‘‘The hydrological model’’ LISFLOOD is presented. Sections ‘‘Case study’’ and ‘‘Model parameterisation’’ describe the study area and model parameterisation, respectively. Results are presented in section ‘‘Results’’ and we end with conclusions.

## Methodology

In this section we describe the Bayesian inference method adopted for the calibration and parameter uncertainty assessment of the LISFLOOD rainfall-runoff model. First, a generic mathematical formulation of the Bayesian procedure for hydrological inversion is provided. The second part of this section details the SCEM-UA optimisation code used to perform the computations. We also provide some details on the parallel implementation of SCEM-UA.

In Bayesian inference knowledge and uncertainty about variables is summarised in probability distributions. In what follows, we use the notation  $p()$  for probability density functions,  $L()$  for likelihood functions, and a vertical bar to indicate conditioning. Arguments on the left side of the vertical bar denote the variables of the density; arguments on the right side of the bar denote the fixed values on which the density is conditioned.

## Bayesian framework for hydrological parameter estimation

Calibration or inverse problems arise anywhere data are collected that are related to the unknown quantities by a mathematical model. The unknown quantities here are the hydrological properties of the river basin under study (e.g., hydraulic properties of the soil, storage or surface routing properties), with the collected data given by observed responses of the system (e.g., river stage or discharge, soil moisture content, groundwater table elevation). The Bayesian approach to the non-linear calibration or inverse problem is the transfer of information from the observed system responses to the unknown quantities, hereby updating the probability density functions that describe the uncertainty about the unknown variables.

The functional form, here the non-linear hydrological model LISFLOOD, relating the system responses with the unknown quantities can be written as

$$\mathbf{y}_F = F(\boldsymbol{\theta}, \boldsymbol{\xi}) \quad (1)$$

where  $\mathbf{y}_F = (y_{F,1}, y_{F,2}, \dots, y_{F,n_y})^T$  is the functional output,  $\boldsymbol{\theta} \in \Theta \subset \mathfrak{R}^{n_\theta}$  is the vector of model parameters representing the unknown hydrological river basin properties, and  $\boldsymbol{\xi} \in \Xi \subset \mathfrak{R}^{n_\xi}$  comprises the forcing inputs to the hydrological model, such as precipitation or ET<sub>0</sub> rates, recharge or discharge fluxes at boundaries, and withdrawal/injection rates.

Omitting from the notation the inputs  $\boldsymbol{\xi}$  to the hydrological model that are assumed to be known and fixed, the relation between the observed system responses  $\mathbf{y} = (y_1, y_2, \dots, y_{n_y})^T$  and the model predictions is given by

$$\mathbf{y} = \mathbf{y}_F + \boldsymbol{\varepsilon}_y(\boldsymbol{\theta}) \quad (2)$$

where  $\boldsymbol{\varepsilon}_y = (\varepsilon_{y,1}, \varepsilon_{y,2}, \dots, \varepsilon_{y,n_y})^T$  is the vector of modelling residuals, which contains various sources of potential prediction errors. The functions we employ, even the most elaborate physically based models, cannot reflect the true complexity and are necessarily simplifications of the processes occurring in the field. Aggregation and lumping of processes in space and time leads to parameter identification errors. Also, observations of system responses are prone to measurement errors.

To implement Bayesian inference, a probability density function with parameters  $\boldsymbol{\psi} \in \Psi \subset \mathfrak{R}^{n_\psi}$  needs to be specified for the residuals that is consistent with the available information about the errors. The joint conditional distribution  $p(\boldsymbol{\varepsilon}_y | \boldsymbol{\theta}, \boldsymbol{\psi})$  describes the distribution of the residuals, given  $\boldsymbol{\theta}$  and the parameters  $\boldsymbol{\psi}$  of the assumed error model. This expression, seen as a function of  $\boldsymbol{\theta}$  and  $\boldsymbol{\psi}$ , is called the likelihood function, and expresses the likelihood of observing the residuals given  $\boldsymbol{\theta}$  and  $\boldsymbol{\psi}$ . Since the structure of  $F(\boldsymbol{\theta})$  is known, the likelihood function is actually proportional to the probability distribution of the observed system responses, i.e.,  $L(\boldsymbol{\theta}, \boldsymbol{\psi} | \mathbf{y}) \equiv p(\mathbf{y} | \boldsymbol{\theta}, \boldsymbol{\psi})$ . Parameter values that closely reproduce the observed system responses will be characterised by high likelihood values.

The information contained in the observed system responses, by means of the likelihood function  $L(\boldsymbol{\theta}, \boldsymbol{\psi} | \mathbf{y})$ , is used to update the prior information of the parameters  $\boldsymbol{\theta}$ , expressed by the prior distribution  $p(\boldsymbol{\theta})$ , and the prior information about the error model, expressed by the prior

distribution  $p(\psi)$ . The prior distributions may be defined based on other data sets or a modeller's experience and physical intuition. Assuming conditional independence between  $p(\theta)$  and  $p(\psi)$ , Bayes' theorem gives

$$p(\theta, \psi | \mathbf{y}) = C_y^{-1} L(\theta, \psi | \mathbf{y}) p(\theta) p(\psi) \quad (3)$$

where  $C_y = \int_{\theta} \int_{\psi} L(\theta, \psi | \mathbf{y}) p(\theta) p(\psi) d\theta d\psi$  is the normalising constant, provided the integral exists. The marginal posterior distribution  $p(\theta | \mathbf{y})$  is then obtained by integrating the joint posterior distribution  $p(\theta, \psi | \mathbf{y})$  over the nuisance parameters  $\psi$  of the error model

$$p(\theta | \mathbf{y}) = \int_{\psi} p(\theta, \psi | \mathbf{y}) d\psi \quad (4)$$

The conditional distribution  $p(\theta | \mathbf{y})$  reflects the uncertainty about the model parameters after the observations of the system responses have been considered. It forms the basis for making predictions about the system responses with the model. For any future time step  $t_i$  the predictive distribution of the system responses is then given by the expression

$$p(\mathbf{y}_{t_i} | \mathbf{y}) = \int_{\theta} p(\mathbf{y}_{t_i} | \mathbf{y}, \theta) p(\theta | \mathbf{y}) d\theta \quad (5)$$

Summarising statistics about the flow predictions, such as measures of the central tendency or spread, or the probability of exceeding a user-defined threshold, are readily obtained from this probability density function.

### Shuffled complex evolution metropolis algorithm

Typically, for hydrological problems the joint posterior parameter distribution  $p(\theta | \mathbf{y})$  is highly dimensional and complex, with strong non-linear parameter interdependencies. Hence,  $p(\theta | \mathbf{y})$  is not easily amenable to direct sampling or analytical integration and it is necessary to resort to Monte Carlo methods to approximate the distribution. Since we do not know the form of the joint posterior distribution in our context, we adopt a Markov chain Monte Carlo (MCMC) approach to compute  $p(\theta | \mathbf{y})$ . In particular, we employ the Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm (Vrugt et al., 2003), which uses the Metropolis-Hastings (Metropolis et al., 1953; Hastings, 1970) search strategy to generate a sequence of parameter sets  $\{\theta_1, \theta_2, \dots, \theta_n\}$  that adapts to the target posterior distribution.

The SCEM-UA algorithm starts with generating an initial population of  $s$  parameter sets sampled from the joint prior parameter distribution  $p(\theta)$ . The latter constrains the parameter space and represents the belief about the parameters before any data are collected. Independent uniform prior parameter distributions between realistic lower and upper bounds are typically adopted for each parameter in  $\theta$ . This implies that the information in the data, expressed by the likelihood function, should dominate the form of the resulting posterior distribution. Note that the assumption of prior independence among the parameters is by no means necessary and can be relaxed only at the cost of making computations more burdensome.

Assuming that the residuals  $\boldsymbol{\varepsilon}_y = (\varepsilon_{y,1}, \varepsilon_{y,2}, \dots, \varepsilon_{y,n_y})^T$  are mutually independent, normally distributed with a constant

variance  $\sigma_{\varepsilon}$ , the likelihood of each parameter set  $\theta$  given the observations  $\mathbf{y}$  is computed using (Box and Tiao, 1973)

$$L(\theta, \psi | \mathbf{y}) = \exp \left[ -\frac{1}{2} \sum_{i=1}^{n_y} \left| \frac{v_i(\theta)}{\sigma_{\varepsilon}} \right|^2 \right] \quad (6)$$

where

$$v_i(\theta) = G(\mathbf{y}_{F,i}) - G(\mathbf{y}_i) \quad (7)$$

The transformation  $G(\cdot)$  of the simulated and observed system responses allows to handle non-normality, lack of variance homogeneity and autocorrelation of error terms in the residuals. The parameter of the error model, i.e.,  $\psi = \sigma_{\varepsilon}$ , is treated as a hyperparameter whose uncertainty is accounted for by marginalisation. Assuming a non-informative prior  $p(\sigma_{\varepsilon}) \sim 1/\sigma_{\varepsilon}$ , and given the uniform prior parameter distributions adopted, the posterior density for each parameter set  $\theta$  given the observations  $\mathbf{y}$  is obtained using (Box and Tiao, 1973)

$$p(\theta | \mathbf{y}) = C^{-1} \left[ \sum_{i=1}^N |v_i(\theta)|^2 \right]^{-\frac{1}{2} n_y} \quad (8)$$

where  $C = \int \left[ \sum_{i=1}^N |v_i(\theta)|^2 \right]^{-\frac{1}{2} N} d\theta$  is the normalising constant.

Once the posterior density has been computed for the  $s$  parameter combinations of the initial sample using Eq. (8), the population is partitioned into  $q$  complexes. In each complex a parallel sequence is launched from the point with the highest posterior density. New candidate points are generated employing a multivariate normal distribution centred around 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. Eq. (8) is used to compute the posterior density of new candidate points, which are added (by random replacement) to the current sequence based on the Metropolis-annealing (Metropolis et al., 1953) criterion. After a predefined number of iterations the complexes are shuffled to share information gained independently in the parallel sequences. For a variety of problems of increasing complexity, including a parameter identification problem for a five-parameter conceptual rainfall-runoff model, Vrugt et al. (2003) showed empirical evidence that this series of operations yields a robust MCMC sampler that efficiently and effectively searches the parameter space and converges to the target distribution for a sufficiently large number of iterations (typically > 5000 iterations).

### Implementation of SCEM-UA using parallel computing

Due to the computational demands of the LISFLOOD model and the large number of iterations typically needed to obtain a stable posterior parameter distribution it was required to implement the SCEM-UA algorithm using parallel computing. We employed a Local Area Multicomputer – Message Passing Interface (LAM/MPI) distributed computing interface for the Octave programming environment (Vrugt et al., 2006). LAM/MPI is a high-quality open-source implementation of the Message Passing Interface specification that includes a rich set of features for parallel computing.

GNU Octave (Eaton, 1998, 2001) is a high-level language, compatible with MATLAB™, that is intended primarily for numerical computations. It provides a convenient command line interface for the numerical solution of linear and non-linear problems, and for performing numerical experiments. A detailed description and explanation of the software appear in Fernández et al. (2003, 2004). The parallel implementation scheme of the SCEM-UA algorithm is presented

in Fig. 1. The calculations reported in this paper were performed using 11 Pentium IV 3.40 GHz processors of the LISA Linux cluster belonging to the SARA parallel computing centre (SARA Computing and Networking Services, the Netherlands). The CPU time required for the stochastic calibration of the LISFLOOD model for a 3-year simulation period with a daily time step and 10,000 SCEM-UA generated parameters combinations was approximately 35 h.

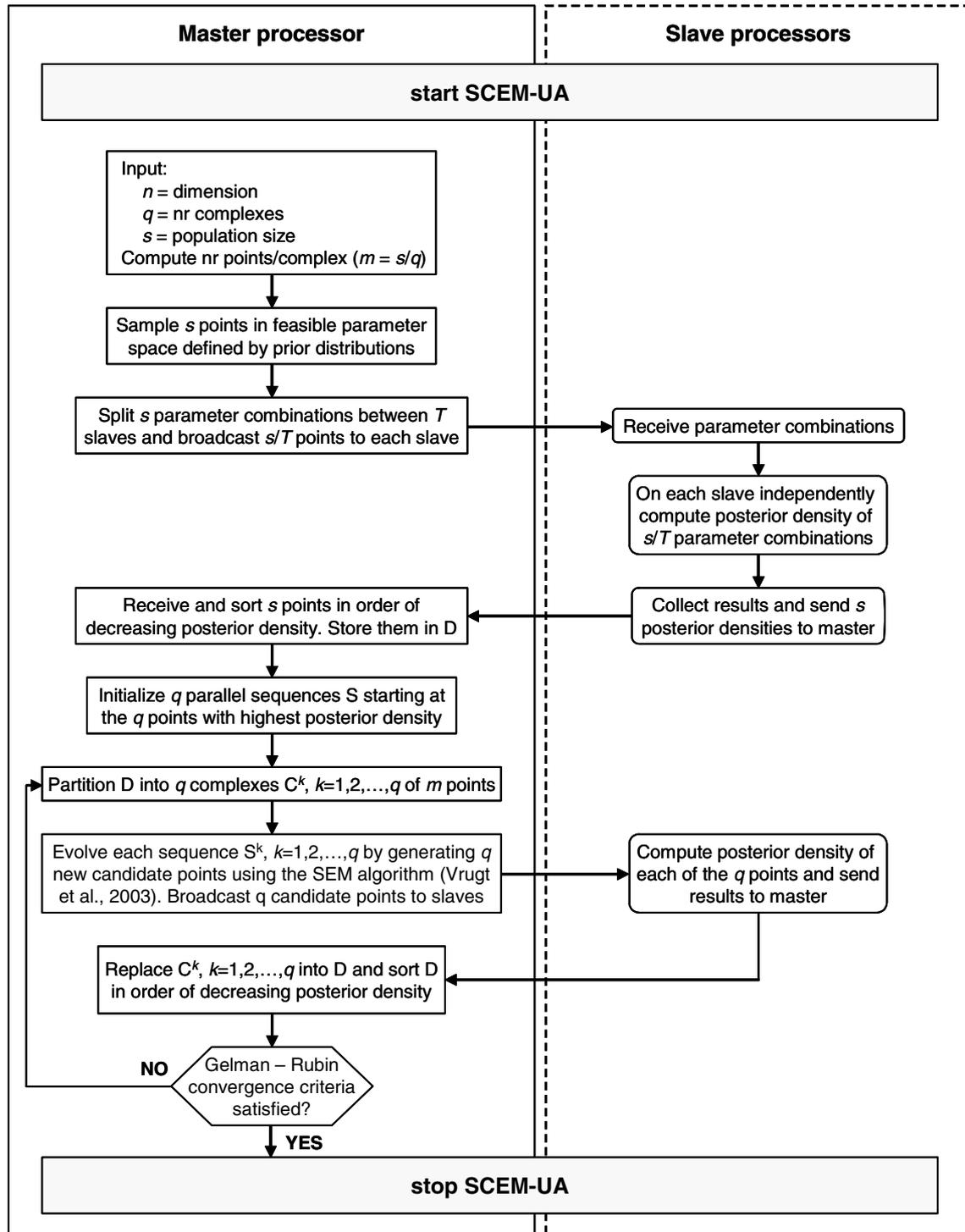
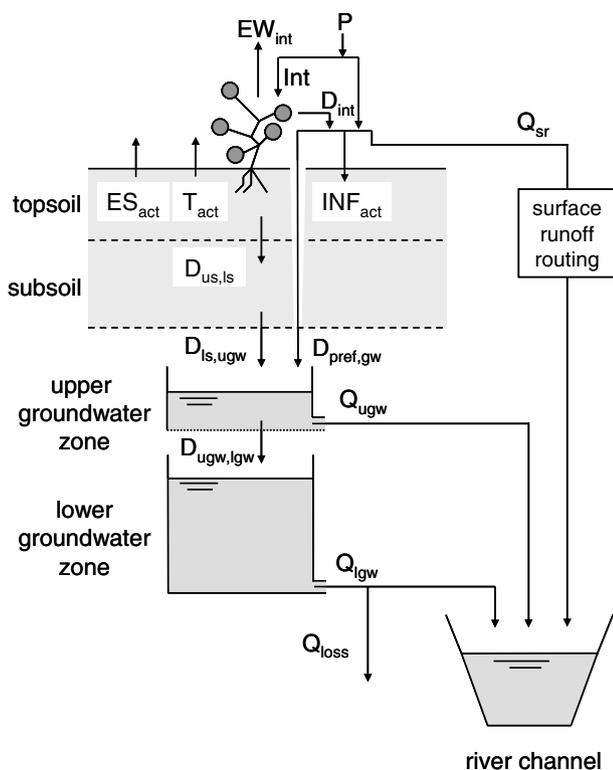


Figure 1 Flow chart of parallel implementation of SCEM-UA.

## The hydrological model

In this study we employ LISFLOOD (De Roo et al., 2000, 2001), a rainfall-runoff model that has been devised to predict floods in large European river basins. To meet this objective the model (i) is distributed, (ii) has physical basis, (iii) uses readily available inputs and (iv) is computationally efficient. The model is raster based and is embedded in a dynamic modelling GIS-environment (PCRaster), which facilitates the handling of large European spatial data sets such as the CORINE Land Cover and the European Soils Database.

A schematic outline of the model with its different components is presented in Fig. 2. The model is driven by meteorological input time series such as precipitation, temperature, wind speed, sunshine duration, cloud cover and actual vapour pressure. Processes simulated for each grid cell include snowmelt, soil freezing, interception of rainfall by vegetation, evaporation from the soil surface, infiltration into the soil, water uptake and transpiration by plants, surface runoff, redistribution of soil moisture within



**Figure 2** Schematic overview of the LISFLOOD model. P = precipitation; Int = interception;  $EW_{int}$  = evaporation of intercepted water;  $D_{int}$  = leaf drainage;  $ES_{act}$  = evaporation from soil surface;  $T_{act}$  = transpiration (water uptake by plant roots);  $INF_{act}$  = infiltration;  $Q_{sr}$  = surface runoff;  $D_{us,ls}$  = drainage from upper to lower soil zone;  $D_{ls,ugw}$  = drainage from lower soil zone to upper groundwater zone;  $D_{pref,gw}$  = preferential flow to upper groundwater zone;  $D_{ugw,lgw}$  = drainage from upper to lower groundwater zone;  $Q_{ugw}$  = outflow from upper groundwater zone;  $Q_{lgw}$  = outflow from lower groundwater zone;  $Q_{loss}$  = loss from lower groundwater zone. (Note that snowmelt is not included in the figure, even though it is simulated by the model.)

the soil profile, drainage of water to the groundwater system, preferential flow, groundwater flow and 1D channel routing.

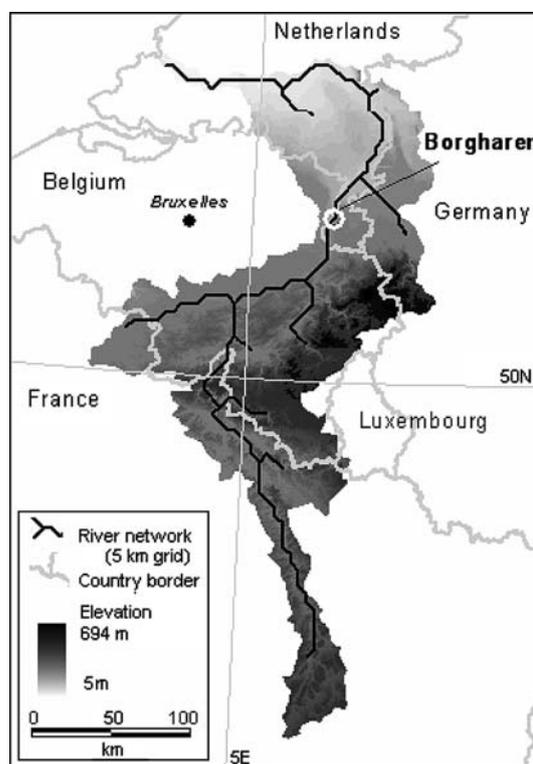
## Case study

We present our analysis and results for the Meuse catchment upstream of Borgharen (see Fig. 3). This part of the Meuse catchment covers an area of approximately 21,000 km<sup>2</sup> and is situated in Belgium, France and the Netherlands. The Meuse is fed mainly by rain all year round; hence flows are generally highest in winter, with relatively low flows during the summer. The topography of the area is hilly with the elevation varying from 50 m to 700 m. The substrata are largely impervious, resulting in precipitation that is discharged quickly into the river. The predominant land use types are forest, agriculture (cultivated patterns and pasture), moor and heath.

In the model, the area was discretised in 5 by 5 km grid blocks. Daily observed discharges are available for the Borgharen gauging station. The model was run with a daily time step. The simulation period in the calibration spanned 1/10/1992 to 30/09/1995. The first year was used as a warming-up period, hence only predicted discharges for the last 2 years were used for the calibration. For validating the model, observations from 10/1/1990 to 30/9/1992 were used.

## Model parameterisation

Distributed models require a vast amount of data to represent the spatial distribution of the meteorological and



**Figure 3** Location of the Meuse catchment, with grey-scale overlay of the topography.

hydrological characteristics of large river basins. A rigorous parameterisation procedure is crucial to avoid methodological problems during model calibration. Spatial patterns of the parameter values have to be specified such that parameters reflect only significant and systematic spatial variations inherent in the available data. As such, the parameterisation process can effectively reduce the number of free parameters to be adjusted during calibration (Refsgaard, 1997).

To avoid problems of over-parameterisation and to reduce the dimensionality of the model calibration, input parameters and variables of LISFLOOD are estimated a priori from available data bases as much as possible. For example, soil physical properties are derived from the European Soil Geographical Database (King et al., 1994). The HYPRES database (Wösten et al., 1999) is used to estimate porosity, saturated hydraulic conductivity and moisture retention properties for each texture class. Vegetation and land use information are obtained from the CORINE Land Cover database (EEA, 2000). Meteorological parameters are extracted from the MARS Meteorological Database. Digital elevation data are obtained from the Catchment Information System, which has a spatial resolution of 1 km (Hiederer and De Roo, 2003).

Although LISFLOOD is based on physics to a certain extent, some processes are only represented in a lumped conceptual way. As a result, some parameters lack physical basis and cannot be directly obtained from field data. In the current version of LISFLOOD, there remain five parameters that need to be estimated by calibration against measured stream flow records. The calibration parameters are tabulated in Table 1, with the upper and lower bounds of the prior distributions used in the inverse procedure. The Upper Zone Time Constant (UZTC) and Lower Zone Time Constant (LZTC) reflect the residence time of water in the upper and lower groundwater zone, respectively. As such, they control the amount and timing of outflow from the respective groundwater reservoirs. The Groundwater Percolation Value (GWPV) controls the flow from the upper to the lower groundwater zone. The Xinanjiang parameter  $b$  ( $X_b$ ) is an empirical shape parameter in the Xinanjiang model (Zhao and Liu, 1995) that is used to simulate infiltration. It controls the fraction of saturated area within a grid cell that is contributing to runoff, hence it is inversely related to infiltration. The Power Preferential Bypass Flow parameter (PPBF) is an empirical shape parameter in the power function relating preferential flow with the relative saturation of the soil.

**Table 1** Calibration parameters of the LISFLOOD model with upper and lower bounds of the prior uniform distributions

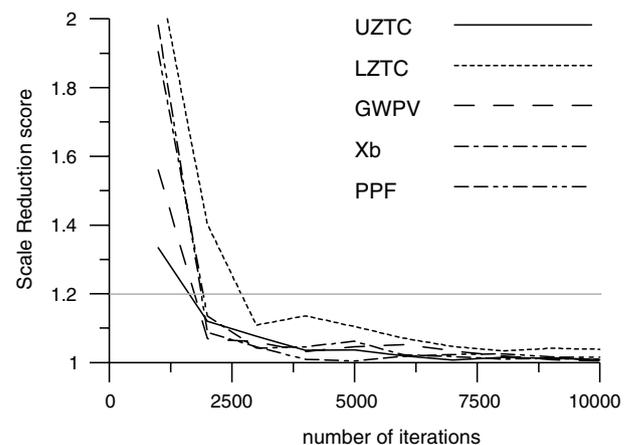
Parameter	Lower bound	Upper bound
Upper zone time constant (UZTC)	1	10
Lower zone time constant (LZTC)	10	5000
Ground water percolation value (GWPV)	0	0.5
Xinanjiang parameter $b$ ( $X_b$ )	0.05	0.5
Power preferential bypass flow (PPF)	5	15

## Results

Several important assumptions underlie the use of Eq. (8) as a likelihood function in the derivation of the posterior parameter distribution, namely that the residuals are uncorrelated in time and normally distributed with constant variance. If these assumptions are not met, the posterior parameter distribution may not adequately describe parameter uncertainty, and the derived predictive uncertainty bounds may be erroneous. The transformation  $G(\cdot)$  in Eq. (7) of the simulated and observed system responses allows to handle non-normality, lack of variance homogeneity and autocorrelation of error terms in the residuals. In this work, we only account for non-normality and heteroscedastic errors by applying a Box–Cox transformation with  $\lambda = 0.4$  to the observed and simulated discharge series. The validity of these assumptions is evaluated later in this section using diagnostic checks. Correlation of the residuals in time can be accounted for by fitting an autoregressive (AR), moving average (MA) or mixed (ARMA) error model to the residuals (see e.g., (Sorooshian and Dracup, 1980; Kuczera, 1983; Bates and Campbell, 2001)).

The SCEM-UA algorithm was run with a population size  $s = 250$  and  $q = 10$  complexes, or 25 points in each complex. Convergence of the MCMC sampler to a stationary distribution was evaluated using the Scale Reduction score ( $\sqrt{SR}$ ) defined by Gelman and Rubin (1992). If the scale reduction score is less than 1.2, the Markov chain is considered to be converged; otherwise, more runs are needed. In Fig. 4 calculated values of  $\sqrt{SR}$  are plotted against the number of MCMC iterations. The line plots indicate that for all parameters the parallel sequences converged to the target distribution after approximately 3000 iterations.

The rapid convergence is confirmed by the evolution of the samples generated in the  $q = 10$  sequences presented in Fig. 5. Initially, parameter values are sampled from the feasible parameter space defined by the prior parameter distributions. After this initial exploration of the parameter space, the sampler discards parameter regions with low posterior probabilities. For the UZTC, GWPV,  $X_b$  and PPF parameters, the SCEM-algorithm rapidly occupies only a



**Figure 4** Evolution of the Gelman and Rubin Scale Reduction score for LISFLOOD calibration parameters.

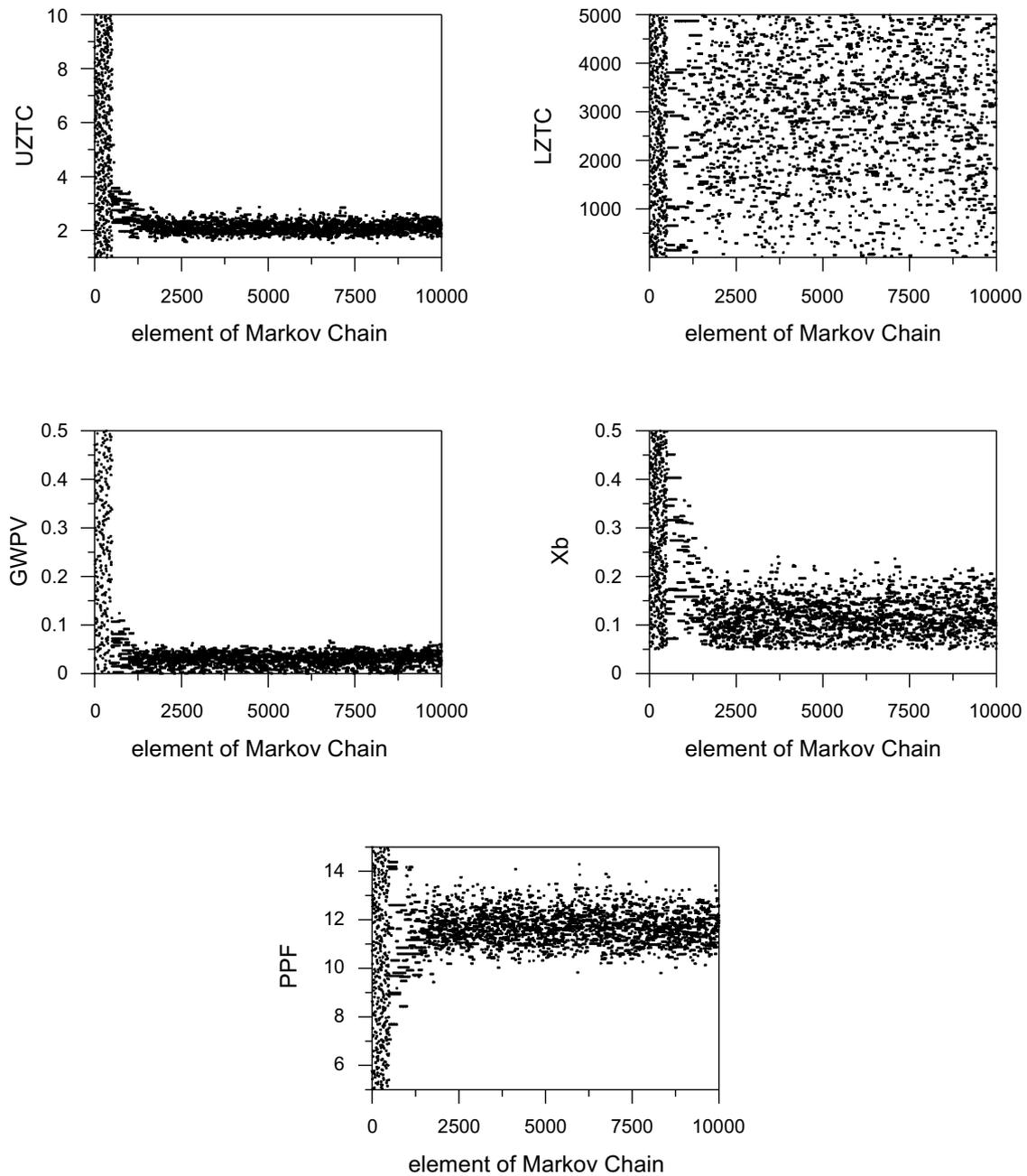
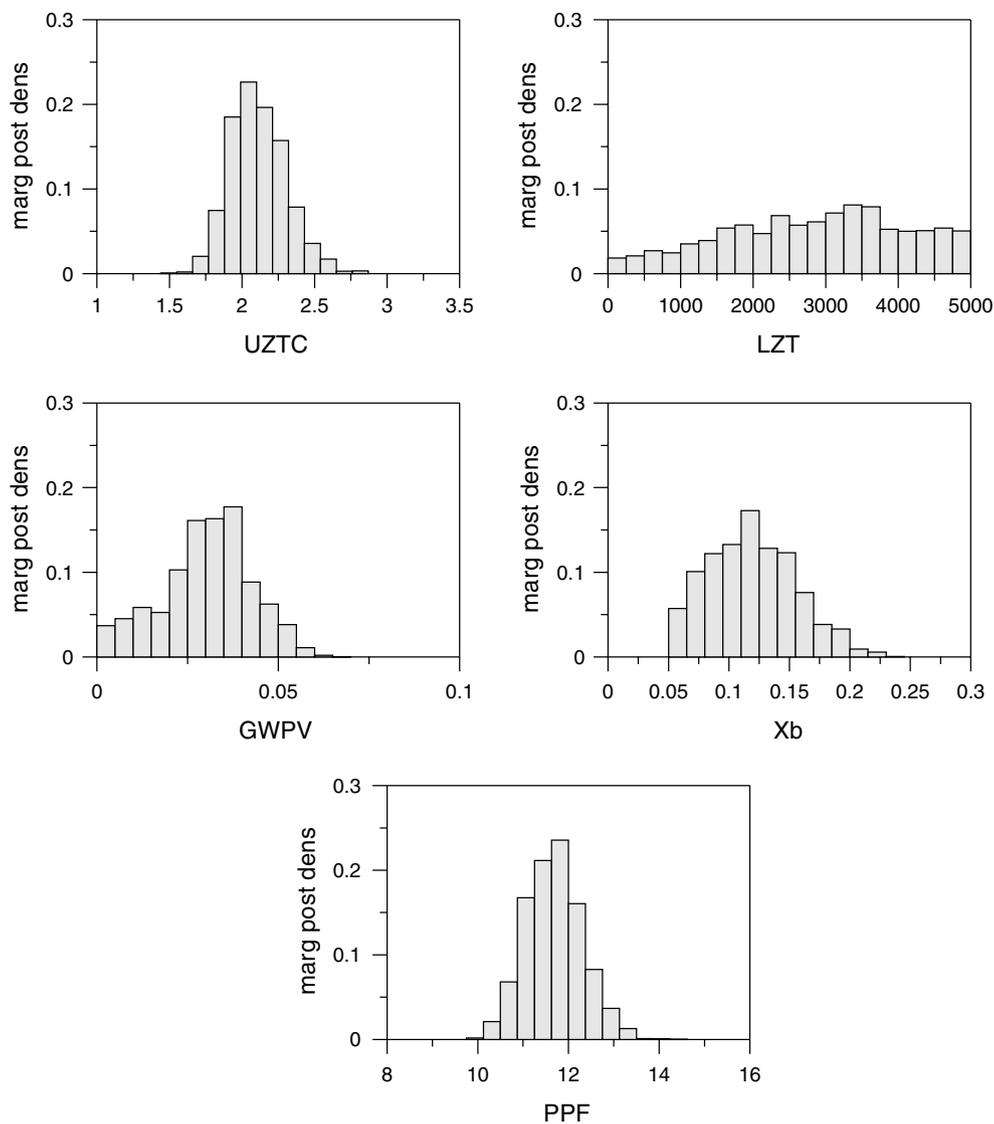


Figure 5 Markov chain Monte Carlo samples generated in the  $q = 10$  sequences for the LISFLOOD calibration parameters.

small range of the initial parameter ranges. For the LZTC no clear defined region of attraction exists.

Fig. 6 presents the marginal posterior probability distributions for the LISFLOOD calibration parameters constructed using 7000 samples generated after convergence of the SCEM-UA algorithm. Note that for all parameters but the LZTC the limits of the x-axis in Fig. 6 do not correspond to the range specified for the respective prior uniform distributions. The summarising statistics of the posterior parameter distributions together with the most likely parameter combination are presented in Table 2. The posterior density for the UZTC and PPF approximate a normal distribution centred around the optimal parameter values; hence, the posterior mean is close to the optimal parameter

value. The well defined region of small values for UZTC reflects the short residence times in the upper groundwater zone. This means that water that has passed the soil zone is quickly discharged into the river channel, which can be related to the impervious subsoil strata present in large parts of the catchment. The high values for the PPF indicate that the contribution of preferential flow bypassing the soil zone is only significant under wet conditions. For the Xb parameter the posterior density also approximates a normal distribution, but it is truncated at the lower boundary of the prior distribution. The low posterior values for the Xb parameter show that runoff in the catchment is small unless the soil is nearly fully saturated. The posterior density for the GWPV parameter is concentrated on very low values,



**Figure 6** Marginal posterior probability distributions of the 5 LISFLOOD calibration parameters.

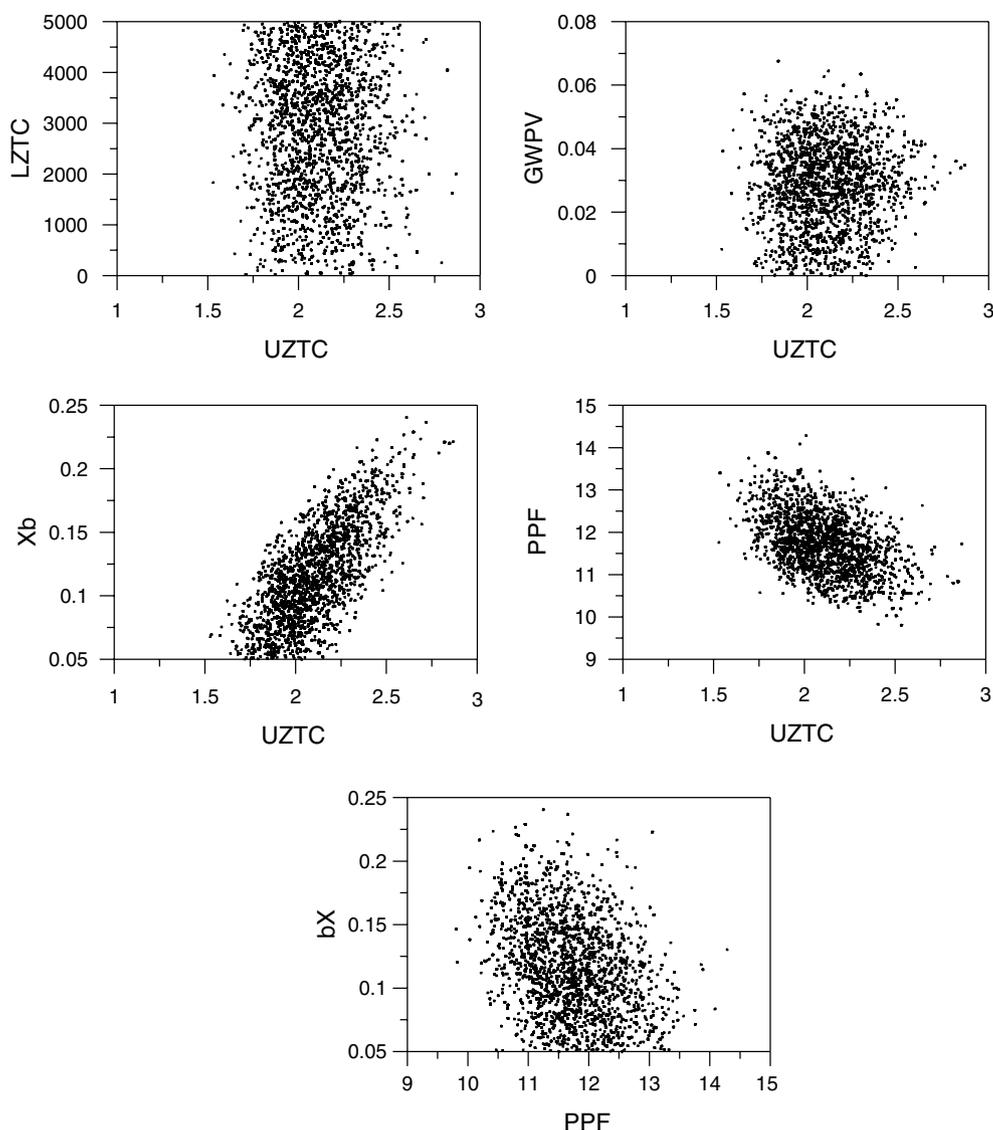
**Table 2** Summarising statistics of the posterior parameter distribution (optimal parameter set, posterior mean, standard deviation and correlation coefficients between the generated samples)

Parameter	Optimal	Mean	Standard deviation	UZTC	LZTC	GWPV	Xb	PPF
UZTC	2.04	2.11	0.19	1	-0.07	0.07	0.71	-0.45
LZTC	3142	2837	1251	—	1	-0.03	-0.01	0.01
GWPV	0.03	0.03	0.01	—	—	1	0.07	0.00
Xb	0.11	0.12	0.04	—	—	—	1	-0.34
PPF	11.8	11.6	0.62	—	—	—	—	1

indicating that the amount of flow from the upper to the lower groundwater zone is limited. Indeed, analysis of the different components contributing to the total discharge showed that the outflow of the lower groundwater zone only accounts for 1.6%. The nearly negligible contribution of the lower groundwater zone also explains the approximately flat response surface for the LZTC parameter. The posterior

distributions and standard deviations depict that the calibration parameters of the LISFLOOD model, except for the lower zone time constant, are well identifiable for the Meuse catchment using 2 years of measured daily discharges.

Fig. 7 presents scatter plots in two dimensions of the parameter space of the 7000 parameters sets sampled from

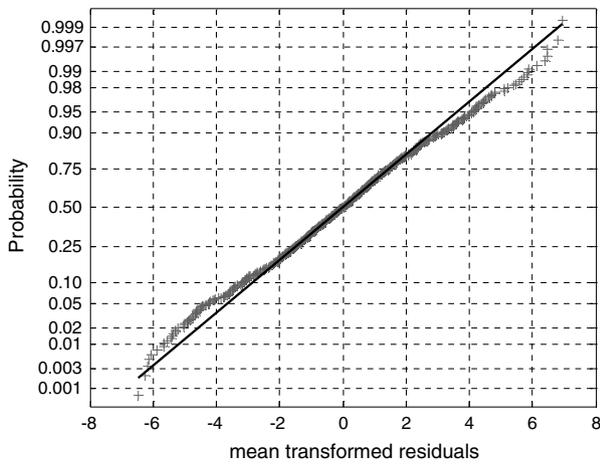


**Figure 7** Scatter plots in two dimensions of parameter space of 7000 MCMC sampled parameter sets for different combinations of parameters.

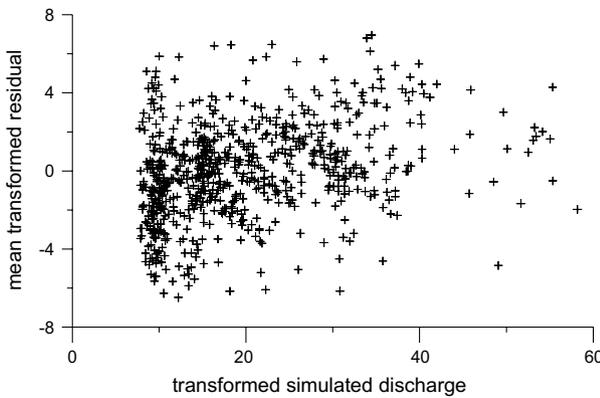
the posterior parameter distribution. The top two plots are representative for the two parameters (LZTC and GWPV) that determine the slow groundwater response of the model. Due to the small contribution of the lower groundwater zone, these parameters show no correlation with the other calibration parameters, which is confirmed by the summarising correlation coefficients presented in Table 2. The bottom three scatter plots and the corresponding correlation coefficients reveal that the parameters that affect the fast response of the model (UZTC,  $X_b$ , and PPF), either through surface runoff or fast groundwater contribution, are correlated. This correlation can be explained as follows. A decrease in the value of  $X_b$  results in (i) more infiltration, hence an increase in soil moisture content and a higher flux out of the soil zone to the groundwater zone and (ii) less surface runoff. The increased flux to the groundwater zone through infiltration can be compensated for by increasing the value of PPF, resulting in less preferential flow to the groundwater zone for the same moisture

content (i.e., infiltration and preferential flow are ‘competing’ processes that recharge the upper groundwater zone). A decrease in surface runoff on the other hand is offset by smaller residence times in the upper groundwater zone, or lower values for the UZTC.

To ensure that the posterior parameter distribution adequately describes parameter uncertainty we perform some diagnostic checks on the modeling residuals. Simulations of the hydrograph were obtained by running the LISFLOOD model for 7000 parameter combinations sampled from the posterior parameter distribution. For each simulation the transformed residuals were obtained as the difference between the Box–Cox transformed observed and simulated discharge series. The normal probability plot of the mean transformed residuals (transformed residuals averaged over the 7000 simulations) shown in Fig. 8 reveals that the residuals closely conform to a normal distribution. This was confirmed by the Kolmogorov–Smirnov and Lilliefors tests at a 5% significance level. The plot of the mean transformed

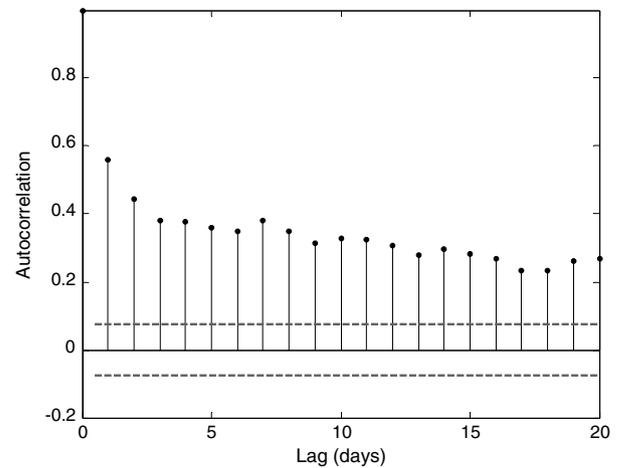


**Figure 8** Normal probability plot for the mean transformed residuals.



**Figure 9** Plot of mean transformed residuals against the transformed predicted runoff obtained with the most likely parameter set.

residuals versus the transformed predicted runoff obtained with the most likely parameter set is shown in Fig. 9, and does not display a strong dependence of the variability of the residuals on predicted runoff. However, the plot reveals that there is a tendency for small discharges to be overestimated. Fig. 10 presents the autocorrelation function (ACF) of the mean transformed residuals. Autocorrelation is a measure of the serial dependency of the error residuals between model predictions and corresponding data. Ideally, this measure should be centred around zero at all lags, implying a lack of bias (systematic errors). Close inspection of this plot, however, reveals that there is a significant autocorrelation between the error residuals at various lags, questioning the reliability of the SCEM-UA derived parameter PDFs of the LISFLOOD model for accurate discharge simulation. This autocorrelation is primarily caused by errors in the forcing data and the presence of structural inadequacies in the LISFLOOD model. Various methods have been proposed in the hydrologic literature to properly handle these errors during parameter estimation, including approaches that use a modified likelihood function to explicitly account for correlation between the error residuals (e.g., Sorooshian

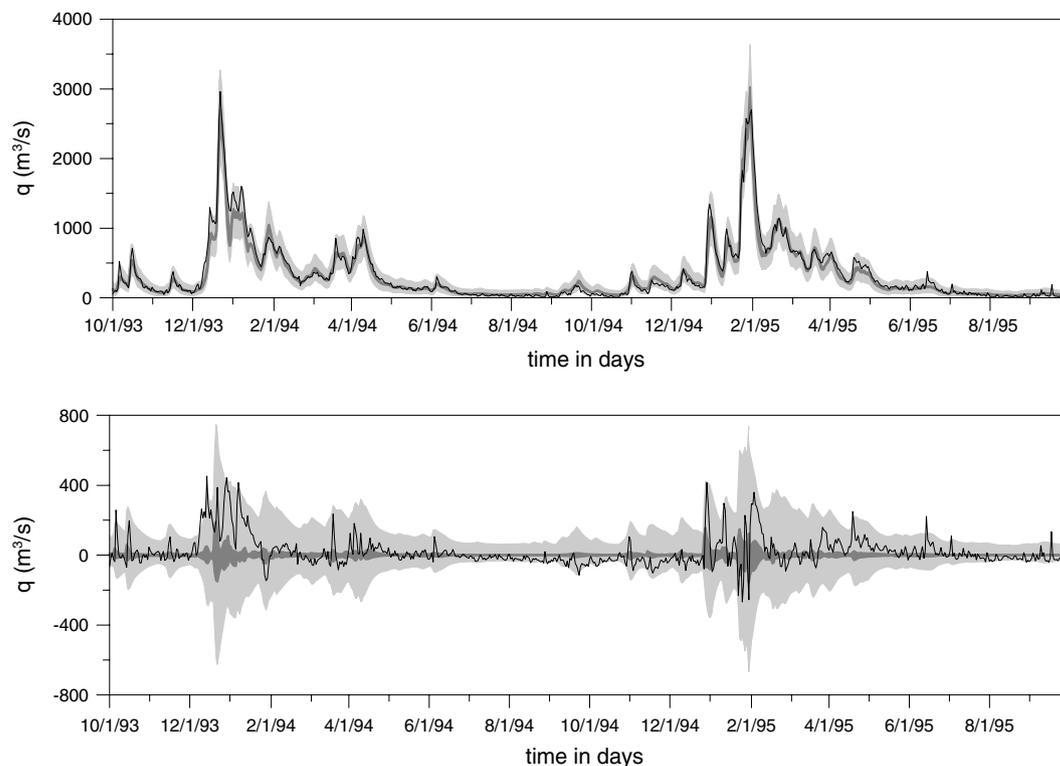


**Figure 10** Autocorrelation function plot of the mean transformed residuals. The dashed lines depict the 95% confidence intervals about zero.

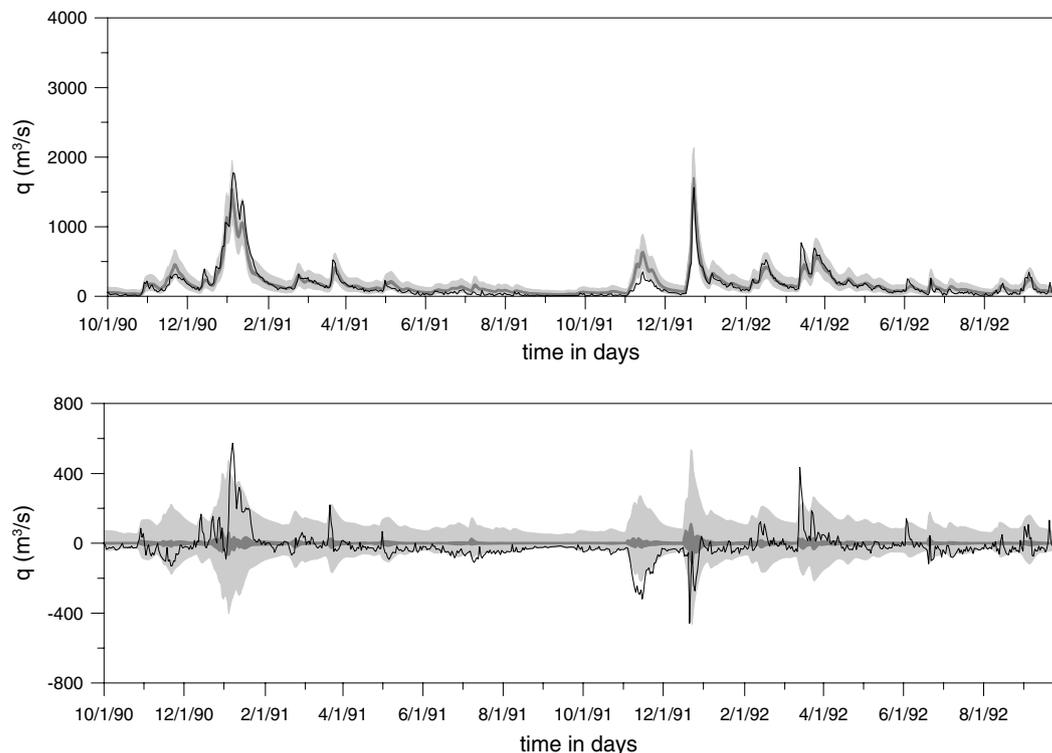
and Dracup, 1980) or combined parameter and state estimation methods (Vrugt et al., 2005), that treat input, output, and model structural error by continuously updating the model states as new observations become available. Such implementations probably impact the posterior parameter PDFs. However, this is beyond the scope of the current paper, and will be explored in future studies.

Probabilistic predictions of the hydrograph were obtained from the discharge series simulated by the LISFLOOD model for 7000 parameter combinations sampled from the posterior parameter distribution. Results for the calibration (10/1/1993 to 9/30/1995) and validation (10/1/1990 to 9/30/1992) period are presented in Figs. 11 and 12, respectively. The top plates in these figure show a plot of the observed discharges (dark line), the 95% prediction uncertainty associated with only parameter uncertainty (dark shaded area), and the 95% prediction uncertainty associated with the total error in terms of modeling residuals (light shaded area). The latter are computed as follows. The standard deviation of the error model, which is assumed constant in the transformed space, is obtained from the RMSE between the transformed observed discharges and the transformed simulated discharges using the most likely parameter set. For each of the 7000 simulations a constant error term equal to  $\pm 1.96 \cdot \text{RMSE}$  is added to the transformed discharges at each time step. The obtained prediction uncertainty limits in the transformed space are then transformed back to the original output space, which explains the varying width of the total prediction uncertainty limits with time. The bottom plates in Figs. 11 and 12 show the observed discharges (dark line) and prediction uncertainty expressed as a deviation from the hydrograph simulated with the most likely parameter set.

The results show that the model predictions reproduce the observed discharges reasonably well during the calibration and validation period. The noticeable over-prediction for the summers of 1991 and 1992 is likely caused by the extraction of water upstream of Borgharen, which can amount to 40 m<sup>3</sup>/s and is not represented in the model. The 95% total prediction uncertainty bounds brackets the



**Figure 11** Hydrograph prediction uncertainty for the calibration period (10/1/1993 to 9/30/1995). Observed discharges are represented by the black line. The dark shaded area denotes the prediction uncertainty that results from parameter uncertainty. The light shaded area denotes the additional prediction uncertainty that results from model and measurement uncertainty. Top plate shows absolute values; bottom plate shows deviations from the hydrograph simulated with the most likely parameter set.



**Figure 12** Hydrograph prediction uncertainty for the validation period (10/1/1990 to 9/30/1992). Observed discharges are represented by the black line. The dark shaded area denotes the prediction uncertainty that results from parameter uncertainty. The light shaded area denotes the additional prediction uncertainty that results from model and measurement uncertainty. Top plate shows absolute values; bottom plate shows deviations from the hydrograph simulated with the most likely parameter set.

observations most of the time. For some periods the total prediction uncertainty is quite large, indicating that the model and/or measurement uncertainty is considerable. The 95% uncertainty region associated with parameter uncertainty is narrow and does not always bracket the observations, indicating that the model structure or the model input data may be in need of further improvement.

## Summary and conclusions

Within the framework of the European Flood Alert System the calibration of the LISFLOOD model is an important and huge task. It involves tuning parameters that can not be obtained by direct measurements against observed discharges for all catchments in Europe, and assessing the parameter uncertainty and its effect on river flow predictions. In this work we demonstrated the capability of SCEM-UA, an automatic Bayesian parameter inference algorithm based on Markov chain Monte Carlo methods, to calibrate the LISFLOOD model for the 21,000 km<sup>2</sup> Meuse catchment upstream of Borgharen. Results showed that the SCEM-UA algorithm was able to efficiently and effectively explore the feasible parameter space and to converge to the target posterior parameter distributions after less than 2500 iterations. The calibration parameters of the LISFLOOD model were well identifiable using 2 years of measured daily discharges, except for the lower zone time constant parameter due to the small contribution of slow groundwater response. The three calibration parameters that affect the fast response of the model showed to be correlated. Hydrograph predictions based on the posterior parameter distributions demonstrated that LISFLOOD is able to reproduce the observed discharges with reasonable accuracy for the Meuse catchment. The narrow parameter uncertainty bounds did not always bracket the results, indicating that improvements in the model or input data may result in more accurate predictions. It should be noted that the posterior parameter distributions were inferred based on the assumption of independent errors. Analysis of the residuals revealed there is a strong evidence to reject this assumption. This will impact on the parameter posterior distributions and also lead to underestimation of the prediction limits. Future research will aim to address this issue.

## Acknowledgement

The hydrological and meteorological data for this study have been provided by the RIZA (Arnhem), Rijkswaterstaat Directie Limburg (Maastricht), Dienst Hydrologisch Onderzoek (Brussels), DIREN (Nancy), KMI (Brussels) and KNMI (De Bilt). We are grateful to the SARA centre of parallel computing at the University of Amsterdam for the computational support. The second author is supported by the LANL Director's funded postdoctoral program. We also wish to acknowledge Katalin Bódis, Rado Bonk and Gianluigi Franchello of the Land Management and Natural Hazards Unit at the Joint Research Centre for their input to this paper. We also wish to thank three anonymous reviewers for their constructive comments to improve the manuscript.

## References

- Beven, K.J., Binley, A., 1992. The future of distributed models: model calibration and uncertainty analysis. *Hydrological processes* 6, 279–298.
- Bates, B.C., Campbell, E.P., 2001. A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. *Water Resources Research* 37 (4), 937–947.
- Box, G.E.P., Tiao, G.C., 1973. *Bayesian Inference in Statistical Analysis*. Addison-Wesley-Longman, Reading, MA.
- Campbell, E.P., Forx, D.R., Bates, B.C., 1999. A Bayesian approach to parameter estimation and pooling in nonlinear flood event models. *Water Resources Research* 35 (1), 211–220.
- De Roo, A.P.J., Wesseling, C.G., Van Deurzen, W.P.A., 2000. Physically based river basin modelling within a GIS: the LISFLOOD model. *Hydrological Processes* 14, 1981–1992.
- De Roo, A.P.J., Odijk, M., Schmuck, G., Koster, E., Lucieer, A., 2001. Assessing the effects of land use changes on floods in the Meuse and Oder catchment. *Physics and Chemistry of the Earth (B)* 26 (7–8), 593–599.
- Duan, Q., Sorooshian, S., Gupta, V.K., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* 28 (4), 1015–1031.
- Eaton, J.W., 1998. Web: <http://www.octave.org/>, University of Wisconsin, Department of Chemical Engineering, Madison WI 53719.
- Eaton, J.W., 2001. Octave: past, present and future. In: *Proceedings of the 2nd International Workshop on Distributed Statistical Computing*, March 2001, Vienna, Austria.
- European Environment Agency, 2000. The European Topic Centre on Terrestrial Environment: CORINE Land Cover raster database 2000 – 100 m.
- Fernández, J., Cañas, A., Díaz, A.F., González, J., Ortega, J., Prieto, A., 2003. Performance of Message-Passing MATLAB toolboxes. In: *Fifth International Conference on High Performance Computing in Computational Sciences*, Lecture notes in computer science, vol. 2565, pp. 228–241.
- Fernández, J., Anguita, M., Mota, S., Cañas, A., Ortigosa, E., Rojas, F.J., 2004. MPI toolbox for octave. In: *Proceedings of the 6th International Conference on High Performance Computing for Computational Science*, June 2004, Valencia, Spain.
- Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7 (4), 457–472.
- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika* 57, 97–109.
- Hiederer, R., de Roo, A., 2003. A European flow network and catchment data set. Report of the European Commission, Joint Research Centre, EUR 20703 EN.
- King, D., Daroussin, J., Tavernier, R., 1994. Development of a soil geographical database from the soil map of the European Communities. *Catena* 21, 37–56.
- Krzysztofowicz, R., Kelly, K.S., 2000. Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resources Research* 36 (11), 3265–3277.
- Kuczera, G., 1983. Improved parameter inference in catchment models. 1. Evaluating parameter uncertainty. *Water Resources Research* 19 (5), 1151–1162.
- Kuczera, G., Mroczkowski, M., 1998. Assessment of hydrologic parameter uncertainty and the worth of multi-response data. *Water Resources Research* 34 (6), 1481–1490.
- Kuczera, G., Parent, E., 1998. Monte-Carlo assessment of parameter uncertainty in conceptual catchment models. *Journal of Hydrology* 211 (1–4), 69–85.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equations of state calculations by fast

- computing machines. *Journal of Chemical Physics* 21, 1087–1091.
- Montanari, A., Brath, A., 2004. A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. *Water Resources Research* 40, W01106. doi:10.1029/2003WR002540.
- Sorooshian, S., Dracup, J.A., 1980. Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: correlated and heteroscedastic errors. *Water Resources Research* 16 (2), 430–442.
- Sumner, N.R., Fleming, P.M., Bates, B.C., 1997. Calibration of a modified SFB model for twenty-five Australian catchments using simulated annealing. *Journal of Hydrology* 197, 166–188.
- Thyer, M., Kuczera, G., Bates, B.C., 1999. Probabilistic optimization for conceptual rainfall-runoff models: a comparison of the shuffled complex evolution and simulated annealing algorithms. *Water Resources Research* 35 (5), 767–773.
- Uhlenbrook, S., Seibert, J., Leibundgut, C., Rodhe, A., 1999. Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrological Sciences Journal* 44, 779–797.
- Vrugt, J.A., Gupta, H.V., Bouten, W., Sorooshian, S., 2003. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic parameter estimation. *Water Resources Research* 39 (8), 1201. doi:10.1029/2002WR001642.
- Vrugt, J.A., Schoups, G.H., Hopmans, J.W., Young, C.H., Wallender, W., Harter, T., Bouten, W., 2004. Inverse modeling of largescale spatially distributed vadose zone properties using global optimization. *Water Resources Research* 40 (6), W06503. doi:10.1029/2003WR002706.
- Vrugt, J.A., Diks, C.G.H., Gupta, H.V., Bouten, W., Verstraten, J.M., 2005. Improved treatment of uncertainty in hydrologic modeling: combining the strengths of global optimization and data assimilation. *Water Resource Research* 41, W01017. doi:10.1029/2004WR003059.
- Vrugt, J.A., Ó Nualláin, B., Robinson, B.A., Bouten, W., Dekker, S.C., Sloot, P.M.A., 2006. Application of parallel computing to stochastic parameter estimation in environmental models. *Computers and Geosciences* 32 (8), 1139–1155. doi:10.1016/j.cageo.2005.10.015.
- Wösten, J.H.M., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. *Geoderma* 90 (3–4), 169–185.
- Zhao, R.J., Liu, X.R., 1995. The Xinanjiang model. In: Singh, V.P. (Ed.), *Computer Models of Watershed Hydrology*. Water Resources Publications, Littleton, Colorado, pp. 215–232.