Stoch Environ Res Risk Assess (2015) 29:847–859
DOI 10.1007/s00477-014-0958-4

ORIGINAL PAPER

# A method to reduce the computational requirement while assessing uncertainty of complex hydrological models

**P. Athira · K. P. Sudheer**

Published online: 10 October 2014


**Abstract** The quantification of uncertainty in the simulations from complex physically based distributed hydrologic models is important for developing reliable applications. The generalized likelihood uncertainty estimation method (GLUE) is one of the most commonly used methods in the field of hydrology. The GLUE helps reduce the parametric uncertainty by deriving the probability distribution function of parameters, and help analyze the uncertainty in model output. In the GLUE, the uncertainty of model output is analyzed through Monte Carlo simulations, which require large number of model runs. This induces high computational demand for the GLUE to characterize multi-dimensional parameter space, especially in the case of complex hydrologic models with large number of parameters. While there are a lot of variants of GLUE that derive the probability distribution of parameters, none of them have addressed the computational requirement in the analysis. A method to reduce such computational requirement for GLUE is proposed in this study. It is envisaged that conditional sampling, while generating ensembles for the GLUE, can help reduce the number of model simulations. The mutual relationship between the parameters was used for conditional sampling in this study. The method is illustrated using a case study of Soil and Water Assessment Tool (SWAT) model on a watershed in the USA. The number of simulations required for the uncertainty analysis was reduced by 90 % in the proposed method compared to existing methods. The proposed method also resulted in an uncertainty reduction in terms of reduced average band width and high containing ratio.

**Keywords** Generalized likelihood uncertainty estimation · Physically based distributed models · Mutual information · SWAT · Uncertainty

## 1 Introduction

The physically based distributed models, which are mostly based on the actual physics behind the hydrologic processes, are ideal for hydrologic predictions at basin level because it can make realistic simulations (Hughes 2010). However, these physically based models are complex in structure, and are also characterized by a multitude of parameters. While some of the model parameters possess physical meaning and can be measured directly from the field, rest of them may require estimation of their values through calibration (Beven 1989; Refsgaard et al. 2007; Wu and Liu 2014). The calibration of these models involve, identifying an appropriate value for the parameters in such a way that the model simulated values closely match with the measured values of the variable of interest (Duan et al. 1992; Deb et al. 2002; Zitzler 2001; Vrugt et al. 2003; Doherty 2004; Razavi and Tolson 2013). Generally, the methods of calibration aim at estimating a single optimal parameter set (combination of parameters) with an objective to minimize the residual variance of simulation. The search for an optimal parameter set from a high dimensional parameter space is a difficult task because of the interrelationship between the parameters and the complexity of the model (Beven and Binley 1992). The interrelation between the parameters will be more in a physically based distributed model with large number of

P. Athira · K. P. Sudheer (✉)
Department of Civil Engineering, Indian Institute of Technology Madras, Chennai 600036, India
e-mail: sudheer@iitm.ac.in

parameters (Cibin et al. 2010; Beven 2006). This will possibly generate many parameter combinations within a model structure with equally acceptable level of performance (equifinality—Beven 2000, 2006; Beven and Freer 2001; Wu and Liu 2014). These parameter combinations (also known as behavioral parameter sets) may be located in different regions in the high dimensional parameter space. Consequently, selection of a single parameter set after calibration may be difficult, probably due to lack of proper basis for differentiation among the parameter sets (Freer et al. 1996). Therefore, one cannot expect that a set of parameter value that is obtained from a calibration procedure will represent a true parameter set of the model (Beven and Binley 1992). The acceptance of all possible behavioral parameter sets and quantification of associated uncertainty will make much more sense than accepting a single calibrated parameter set (Cibin 2014). Accordingly, generating ensemble of good performing parameter sets so as to develop ensemble simulations may be a better approach for building confidence in model outputs. This can be achieved easily if the probability distribution of parameters is known a priori. The influence of parameter probability distribution functions on the uncertainty assessment is clear from Shen et al. (2012). However, the statistical characteristics of the parameters are generally unknown in most of the hydrologic models.

Generally the uncertainty in model simulations are considered to be caused from three different components: (i) structural uncertainty—this source of uncertainty include processes that are not accounted in the model, and model inaccuracy due to over simplification of the processes considered in the model; (ii) input uncertainty—often related to imprecise measurements of model inputs or initial conditions, such as elevation, land use, precipitation, and temperature; and (iii) parameter uncertainty—attributed to a number of unknown parameters in the model (Beven and Freer 2001; Srivastav et al. 2007; Wang et al. 2009; Renard et al. 2010; Wu and Liu 2012; Brigode et al. 2013). While all the sources of uncertainty should be addressed in any modeling effort, the focus of this paper is to quantify the uncertainty caused by the parameters of the model. The techniques that are generally employed for parametric uncertainty evaluation can be categorized as follows: (i) probability theory method (Wiwatenadate and Claycamp 2000), (ii) Taylor series expansion (Rastetter et al. 1992), (iii) Monte Carlo simulation (Gardner and O'Neill 1983), (iv) generalized likelihood uncertainty estimator (GLUE, Beven and Binley 1992), (v) Bayesian statistics (Katz 2002), and (vi) sequential partitioning (Rastetter et al. 1992).

The GLUE is a most commonly used uncertainty estimation method in hydrology. The GLUE helps reduce the parametric uncertainty by deriving the probability distribution function of parameters, and help analyze the uncertainty in model output. The major advantages of GLUE technique are: (i) it is easy to implement in different systems and (ii) it can be applied to non-linear systems also (Jin et al. 2010). Application of GLUE method involves subjective selection of three of its components, the threshold value, the prior distribution and the likelihood index (Blasone et al. 2008a). The threshold value is used for screening behavioral parameter sets from initial population of parameter sets that are generated from the prior probability distribution of parameters, which is generally unknown. Ideally, the prior distribution should be consistent with the effectiveness of changes in a model parameter on changes in model response (Beven and Binley 1992). However, a uniform distribution function over a known range of parameter values will be appropriate in the absence of prior knowledge about the characteristics of parameters. The selection of likelihood in the GLUE methodology, in a very general sense, is a possibilistic measure about the proximity of the model prediction and the observed one. The essential characteristics of the likelihood index are: (i) its value should be equal to zero for simulations that have behavior dissimilar to the observed, and (ii) its value should monotonically increase as the match between observed and simulated increase. Many researchers have employed Nash-Sutcliff efficiency (NSE) as the likelihood in GLUE (Smith et al. 2008).

The major constraint associated with application of GLUE to complex hydrologic models is its high computational demand (Blasone et al. 2008a; Stedinger et al. 2008; Kuczera and Parent 1998; Jia and Culver 2008). The Latin Hypercube sampling (LHS) is generally used in GLUE (Beven and Freer 2001; Blasone et al. 2008a), and it calls for a large number of simulations for getting stable and consistent estimates of behavioral parameters from high dimensional parameter space. While there are a lot of variants of GLUE (e.g., Vrugt et al. 2003; Cibin 2014) that derive the probability distribution of parameters, none of them have addressed the issue of high computational requirement in the analysis. Cibin (2014) used GLUE in an iterative manner with LHS to get a converged actual probability distribution function of model parameters. While this was an improvement over the traditional GLUE, their method also required a large number of model simulations.

This study proposes an approach to reduce the computational demand of the method introduced by Cibin (2014). It is envisaged that conditional sampling, while generating ensembles for the GLUE, can help reduce the number of model simulations. The mutual relationship between the parameters in terms of mutual information (MI) was used for conditional sampling in this study. The proposed approach is illustrated using a real world case study that

uses Soil and Water Assessment Tool (SWAT) model. A comparative analysis of the computational requirement of the three methods viz. the traditional GLUE, Cibin (2014), the currently proposed method, is performed in this study, along with a discussion of the uncertainty estimate from all the three approaches.

## 2 Descrption of the model and study area

### 2.1 Soil and water assessment tool (SWAT)

Soil and Water Assessment Tool is a most widely accepted watershed scale model developed by USDA Agricultural Research Service (ARS) (Arnold et al. 1998). More than 250 applications of SWAT model is presented in-depth in Gassman et al. (2007). This is a continuous time model that operates on a daily time step. This is developed to predict the impact of land management practices on water, sediment and agricultural chemical yield in watersheds with different characteristics. SWAT is a semi-distributed physically based model, and hence it can be applied for simulations in ungauged basins also. The major model components include weather, hydrology, soil temperature, plant growth, land management practices, and reactive and non-reactive contaminant transport modeling.

The SWAT model divides the watershed into sub-watersheds based on the topography of the area. The sub-watersheds are further divided into hydrologic response units (HRUs) that consist of homogenous landuse, soil characteristics and slope. The fundamental SWAT simulations are carried out at HRU level. Water balance (Eq 1) is the driving force behind all simulations in SWAT.

$$SWt = SW_0 + \sum_{i=1}^{t} \left(R_{\mathrm{day}_i} - Q_{\mathrm{surf}_i} - E_{\mathrm{a}_i} - W_{\mathrm{seep}_i} - Q_{\mathrm{gw}_i}\right) \quad (1)$$

where, $SW_t$ is the final soil moisture content (mm), $SW_0$ is the initial moisture content (mm), $t$ is the time (days), $R_{\mathrm{day}_i}$ is the amount of precipitation on day $i$ (mm), $Q_{\mathrm{surf}_i}$ is the amount of surface runoff on day $i$ (mm), $E_{\mathrm{a}_i}$ is the amount of evapotranspiration on day $i$ (mm), $W_{\mathrm{seep}_i}$ is the amount of percolation and bypass flow exiting soil profile on day $i$ (mm), and $Q_{\mathrm{gw}_i}$ is the amount of return flow on day $i$ (mm). The major hydrological processes modeled in SWAT are precipitation, surface runoff, soil and root zone infiltration, evapotranspiration, soil and snow evaporation and base flow (Arnold et al. 1998). Surface runoff is estimated using a modified version of the Soil Conservation Service (SCS) Curve Number method. A kinematic storage model is used for predicting lateral flow. The Muskingum method is used for channel routing. The evapotranspiration can be estimated using three methods, Priestly Taylor, Hargreaves and Penman- Montieth. While SWAT is a comprehensive watershed simulations model, the current study focuses on the stream flow modeling component of SWAT, and assessing the uncertainty. The parameters of the SWAT model that affect the stream flow simulations are identified through a detailed literature review and are presented in Table 1, along with their recommended range of perturbations (Neitsch et al. 2002; Arabi et al. 2007).

### 2.2 Study area and data

The SWAT model has been applied to Illinois River basin in Arkansas, USA, which is one of the major watersheds of the Northwest Arkansas. Illinois River, flowing west across the Arkansas-Oklahoma border into Oklahoma, crosses the state line just south of Siloam Springs at the Arkansas Highway 59 Bridge. This watershed is used in this study for the demonstration of the proposed method. The outlet of the watershed is the USGS gauging site 07195430 on Illinois River, South of Siloam Springs, Arkansas. The geo-reference for the gauging site is 36°06′33.32″ Latitude and 94°32′04.3″ Longitude (NAD83). The drainage area of the watershed up to this gauging site is 1,490 km$^2$. The watershed boundary, delineated using SWAT model (with ArcView interface), is shown in Fig. 1 for reference.

The elevation of the watershed varies from 279.6 to 600.0 m with a mean elevation of about 380.5 m. The digital elevation map obtained from United States Geological Survey (USGS) at 30-m resolution is used to provide the GIS file of elevation in the SWAT model. The land use information of the Illinois watershed has been obtained from the 'Arkansas Land-use/Land-cover, 1999' data prepared by Center for Advanced Spatial Technologies (CAST), University of Arkansas. Major land use categories of the watershed are pasture under tall fescue and Bermuda followed by forests and residential areas. United States Department of Agriculture (USDA), Natural Resources Conservation Service (NRCS) database, Soil Survey Geographic (SSURGO), for Benton County and Washington County, Arkansas are used for extracting soil information in the watershed. Major soil types in the watershed are Nixa, Captina, Clarksville, and Enders covering an area of more than 5 %. There are several minor soil types having a share of less than 5 % in the watershed. Weather data from stations within the region, Fayetteville Experiment Station, and Bentonville, are incorporated to provide the most representative precipitation and temperature data available. Other meteorological data required by SWAT (solar radiation, wind speed, and relative humidity) are estimated using the SWAT weather generator.

The SWAT model is setup for the Illinois River watershed for 9 years, 1995–2003, out of which first 3 years

**Table 1** The SWAT parameters that influence stream flow simulation in the model and their range of perturbation

| Symbol | Description | Unit | Min | Max | Process | Sensitivity index |
|---|---|---|---|---|---|---|
| ALPHA_BF | Base flow recession coefficient | Days | 0 | 1 | Groundwater | 0.015 |
| CN_f* | Curve number | % | −25 | 15 | Surface runoff | 0.385 |
| ESCO | Soil evaporation coefficient | – | 0.001 | 1 | Evapotranspiration | 0.421 |
| GW_DELAY | Groundwater delay time | Day | 1 | 500 | Groundwater | 0.000 |
| GW_REVAP | Revap coefficient | – | 0.02 | 0.2 | Groundwater | 0.000 |
| GWQMN | Depth of water in shallow aquifer | mm | 0 | 5,000 | Groundwater | 0.000 |
| OV_N | Manning's N | – | 0.1 | 0.3 | Overland flow | 0.000 |
| SFTMP | Snowfall temperature | °C | −5 | 5 | Snow | 0.002 |
| SLOPE* | Slope | % | −0.5 | 1 | Surface runoff | 0.018 |
| SLSUBBSN* | Slope sub basin | % | −0.5 | 1 | Surface runoff | 0.000 |
| SOL_AWC* | Available water capacity | % | −0.3 | 2 | Groundwater, evaporation | 0.162 |
| SOL_K* | Saturated hydraulic conductivity | % | −0.5 | 1 | Groundwater | −0.003 |
| SURLAG | Surface lag | Day | 1 | 12 | Surface runoff | 0.001 |

* These parameters were changed as a percentage of their default values to maintain heterogeneity

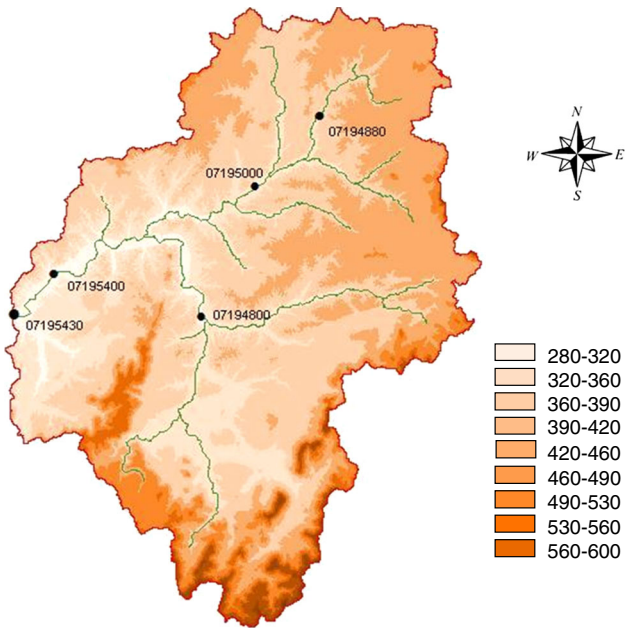


**Fig. 1** Map of the Illinois river watershed area with DEM description (in meters) and location of USGS gauging stations

were considered as warm up period. Thus, effectively 6 years' data were considered for the analysis. The measured daily stream flow values from USGS gauging station 07195430 (Fig. 1) were used for the analysis. Illinois basin experiences an average annual rainfall of 90.5 cm. The daily flow ranged from a minimum of 2.4 $m^3/s$ to a maximum of 538 $m^3/s$ during the period of analysis. The mean flow during the period was 16.5 $m^3/s$ with a standard deviation of 33.7 $m^3/s$. The Illinois River basin experiences a subtropical climatic condition with high evaporative demand; average annual potential evaporation in this area is 105 cm. The SWAT setup for Illinois River watershed had 26 sub basins and 286 HRUs.

## 3 Methodology for uncertainty assessment

As discussed earlier, the proposed methodology works on the concept of the GLUE. Instead of the LHS sampling from probability distribution function of the parameters, the current study proposes a conditional sampling using mutual information between the parameters so as to reduce the computational burden. The method starts with Monte Carlo simulations of the model with ensemble parameter sets generated by Latin Hypercube Sampling (LHS). The original probability distribution function of the parameters is unknown, so a uniform distribution was assumed initially. The sensitive parameters were identified using a global sensitivity analysis method, SOBOL (Freer et al. 1996; Manache and Melching 2008). The details of the SOBOL's sensitivity analysis method for SWAT are described by Cibin et al. (2010), and are not included here.

The sensitive parameters are considered for further analysis of deriving the posterior probability distribution function. An informal likelihood is estimated for each ensemble parameter set that is considered. The Nash–Sutcliffe efficiency is used as a likelihood in this study as it is one of the most commonly used informal likelihood in GLUE with a value close to zero to indicate dissimilarity

between the prediction and observed, and the value 1 to indicate the closeness of prediction and observed. The Nash–Sutcliffe efficiency is defined as,

$$NS = 1 - \frac{\sum_{t=1}^{n} (y_t(\boldsymbol{\theta}) - y_t)^2}{\sum_{t=1}^{n} (y_t - \bar{y})^2} \quad (2)$$

where, $n$ is the number of observed data points, $y_t$ and $y_t(\theta)$ represent respectively the observation and the model simulation with parameters $\theta$ at any given time $t$, and $\bar{y}$ is the average observed value.

A threshold value of the likelihood is used to screen better performing parameter sets from the Monte Carlo simulations. The selection of the threshold is an important step in GLUE, a higher threshold value with acceptance sampling ratio less than 1 % is considered in this study (Jin et al. 2010; Li et al. 2010). The accepted parameter sets are further considered for deriving the posterior probability distribution function of sensitive parameters. The selection of these behavioral parameter sets will help to reduce the effective range of sensitive parameters, and thereby reduce the predictive uncertainty. The posterior probability distribution function of the parameters is derived using Bayes' theorem.

$$p(\boldsymbol{\theta}|\mathbf{x}, \mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) * p(\boldsymbol{\theta})}{p(\mathbf{y}|\mathbf{x})} = \frac{p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) * p(\boldsymbol{\theta})}{\int p(\mathbf{y}|\mathbf{x}, \boldsymbol{\theta}) * p(\boldsymbol{\theta}) * d(\boldsymbol{\theta})} \quad (3)$$

Here $p(\theta \mid x, y)$ denotes the posteriori probability distribution of the parameter set $\theta$, the term $p(y \mid x, \theta)$ is usually known as the likelihood function (Yang et al. 2008) and $p(\theta)$ denotes the prior probability density of the parameter set $\theta$.

The behavioral parameter sets are again used to estimate the nonlinear dependence, mutual information (MI), between the parameters. The MI is a quantitative measure of how much one variable is dependent on the other one (MacKay 2003). For a set of N bivariate measurements $\mathbf{z}_i = (\mathbf{x}_i, \mathbf{y}_i)$, $i = 1, 2, \ldots, N$, which are assumed to be interdependent and identically distributed realizations of a random variable $\mathbf{z} = (\mathbf{x}, \mathbf{y})$, mutual information (MI) is defined as (Tourassi et al. 2001)

$$\mathbf{MI}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{N} f_{x,y}(x_i, y_i) \log_e \frac{f_{x,y}(x_i, y_i)}{f_x(x_i) f_y(y_i)} \quad (4)$$

where $f_x(x)$ and $f_y(y)$ are the marginal probability density functions of $\mathbf{x}$ and $\mathbf{y}$, respectively, and $f_{x,y}(x,y)$ is the joint probability density function of $\mathbf{x}$ and $\mathbf{y}$. Herein, the marginal probability density function of the parameters is the prior distribution and the joint probability density function is derived from the likelihood.

A set of 1,000 parameter values for each sensitive parameter is sampled from the posterior probability distribution functions using LHS sampling. The advantage of using LHS sampling is that it will cover the entire range of the parameter distribution with more parameters from the high probability region. According to Cibin (2014), all of these 1,000 parameter sets are to be used for simulations in the next iteration. However, the current study uses an optimization algorithm to select 200 parameter sets from the LHS sampled pool of 1,000 values for each parameter with an objective function that helps preserve the mutual relationship between the parameters. Genetic Algorithm (GA) is used in this study during this step. This 200 parameter set is considered for simulations in the next iteration. This will continue as a loop until we get a converged probability distribution function for all sensitive parameters in successive iterations. The convergence of the probability distribution functions was checked with the Kullback–Leibler divergence criterion (Carota et al. 1996). It is a non-negative, non-symmetric measure of difference between two distributions, which is an expectation of the logarithmic difference between two distributions. The sum of the probabilities should be unity for estimating the K–L divergence criteria. The K–L divergence of the distribution P on Q is estimated as (Kullback 1959)

$$D_{KL}(P||Q) = \sum \ln\left(\frac{P(i)}{Q(i)}\right) P(i) \quad (5)$$

where, $n$ is the number of data used for the analysis and $P(i)$ and $Q(i)$ are the corresponding probabilities. The converged probability distribution functions can be used for ensemble simulations in the basin and the uncertainty analysis can be done. The flowchart of the methodology is presented in Fig. 2.

## 4 Results and discussion

### 4.1 The sensitivity analysis of SWAT parameters

The SOBOL sensitivity analysis (Sobol 1993; Cibin et al. 2010) has been done for the Illinois basin, and the results are presented in the Table 1. As mentioned earlier, 13 SWAT parameters were considered for the study, with a base sample of 2,000, making it a total of 28,000 simulations for 13 parameters. The value of the sensitivity index for each parameter is also given in Table 1, which shows that ESCO is the most sensitive parameter (sensitivity index = 0.421) in this basin. The Illinois basin is located at the Southern part of USA and experiences higher temperature and radiation (Migliaccio and Chaubey 2008); consequently the evaporation process could be a significant one in the water cycle. The parameter ESCO controls the soil evaporative demand that is to be met from different depths of the soil, and therefore ESCO could be sensitive in

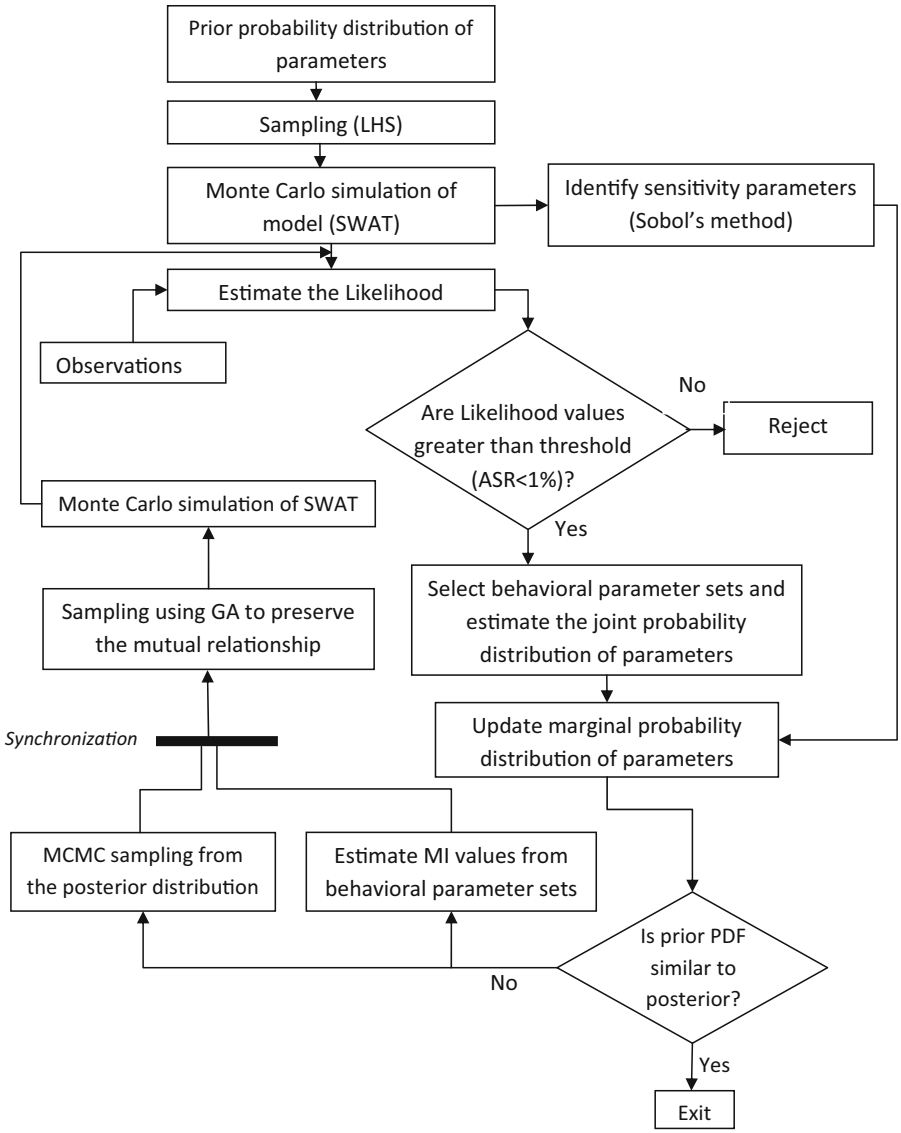


**Fig. 2** Flowchart of the methodology—(flow of computation after the *synchronization block* will take place only after all the information up to this block is arrived). Note: *ASR* acceptance sampling rate; *GA* genetic algorithm; *MI* mutual information; *PDF* probability distribution functions

this basin. The parameters CN_f (Curve Number) and SOL_AWC (available soil water) were the second and third sensitive parameters in the Illinois basin; both of them are closely related to the streamflow generation process. The detailed description about these model parameters and their physical meaning is available in Chen and Wu (2012), and are not reproduced herein for brevity.

The parameters SULRLAG, SLOPE, SOL_K, SFTMP and ALPHA_BF were also slightly sensitive in this basin. However, except SURLAG all other parameters were non-identifiable. The non-identifiable parameters can produce similar order values of NSE for any value of it within its full range, and plausibly cause difficulty in selecting a specific value for the parameter. The identifiability of the parameters is important for the derivation of probability distribution function (Cibin et al. 2010). Consequently, only the SURLAG was considered as the other sensitive parameter. These results are consistent with the observation of Athira et al. (2011), where it is reported that the four sensitive parameters for streamflow generation are ESCO, CN_f, SOL_AWC and SURLAG with varying ranks depending on the watershed. These four parameters were considered for the further analysis in this study. The remaining five parameters, though were found to be slightly sensitive, were kept at the recommended values since they were not identifiable. A calibration of non-identifiable parameters is difficult as they can lead to equifinality of parameters (Demaria et al. 2007; Cibin et al., 2010). In this study these parameters were assigned values considering their importance in the concerned hydrologic processes and also considering the watershed characteristics.

### 4.2 Performance of the model during base simulations

The range of model performance in terms of NSE in the base simulation was −0.2412 to 0.5282. A threshold value of 0.45 was considered in this study with an acceptance sampling ratio less than 1 % (Jin et al. 2010). The selected threshold value was used to identify the behavioral parameter sets, and 243 behavioral parameter sets were found out off 28,000 parameter sets. The range of ESCO was reduced to 0.65–0.99 in the behavioral parameter set from an original range of 0–1. Except SURLAG, all other parameters had experienced reduction in the effective range (−0.1173 to 0.1490 for CN_f and −0.2871 to 1.5064 for SOL_AWC). Even though the parameter SURLAG was identifiable with higher density of parameter sets towards the lower values of the possible range (Table 1), there was no reduction in the effective range.

### 4.3 Probability distribution function of sensitive parameters

The probability distribution functions over the entire effective range of parameters were derived using the identified behavioral parameter sets. The likelihood values along with the corresponding prior probability distribution function were used to update the posterior probability distribution using Bayes' theorem. Thousand parameter values for each parameter were sampled from these derived distributions using LHS. The Mutual Information between the parameters was estimated using the behavioral parameter sets. The joint probability distribution function of parameters was accounted in the estimation of mutual information by considering the likelihood value as the joint probability. A conditional sampling was performed using Genetic Algorithm with an objective function to minimize the sum of squared difference between the MI values that estimated from the GA samples and already estimated MI values using the behavioral parameter sets. The sensitivity of the GA parameters was analyzed and fixed as follows: population size = 10, number of generations = 150, mutation rate = 0.01; and crossover rate = 0.8.

Two hundred parameter sets, which preserved the mutual relationship between the parameters, were selected from the LHS sampled pool. These 200 parameter sets were considered for the ensemble simulations in the next iteration. The same procedure was followed until the probability distribution functions of all the four parameters converged in successive iterations. The convergence of the probability distributions was checked with the Kullback–Leibler divergence criterion (Carota et al. 1996) and the corresponding value of the distribution for all the four parameters after the first iteration was observed to be 0.026, 0.0009, 0.012 and 0.0084 respectively for ESCO, CN_f, SOL_AWC and SURLAG. The Kullback–Leibler divergence criterion value was very close to zero for all the four parameters after the first updating in the proposed method (GLUE with LHS + MI), which clearly indicate convergence of prior and posterior probability distribution.

The converged probability distribution functions were quantified using a Best fit program, @Risk (Palisade Corp. CA, USA). This program considers 28 different distributions to the data, and ranked them according to the specified goodness of fit criterion. The Chi square goodness of fit criterion was used to evaluate and rank the distributions and the parameters of the distributions were estimated using the maximum likelihood estimator. The analysis suggested that all the four parameters followed Beta General Distribution with different shape and location parameters (Table 2). The characteristics of the derived probability distribution for all the parameters are given in Table 2 (GLUE with LHS + MI). It is noted from Table 2 that the specified condition of preservation of mutual information during sampling helps reduce the effective range of parameters in the ensemble simulation. For instance, the effective range for SURLAG has reduced to 1.0197–2.0955 from 1.0–12.0. A similar reduction can be seen in the case of other parameters also (see Tables 1, 2). The probability distribution functions derived are given in Fig. 3 for visual inspection, from which the convergence of the distributions in the first and second updating is visible.

### 4.4 Progressive improvement in the uncertainty reduction

The output uncertainty can be measured in a probability distribution based approach rather than calibrating a model. The simulations with maximum possible range for all the parameters showed a high output uncertainty, NSE varying from -0.2412 to 0.5282. The probability distribution function was unknown in the base simulation, and therefore a uniform distribution function was assumed, and the output uncertainty was high. The updating of the probability distribution function reduced the effective range of the parameters. This helped generate more samples for ensembles from the higher probability region of each parameter, which induced reduction in output uncertainty. The likelihood value plotted against its normalized frequency count is given in Fig. 4a. The frequency plot of base simulation showed a peak at 0.35 NSE. The performance range for the first updating was 0.15–0.59 with a peak at 0.50 NSE in the method proposed in this study. The peak frequency of the distribution is shifted to a higher value from 0.35 to 0.50 NSE; this indicates a reduction in the parametric and output uncertainty.

**Table 2** Location and shape characteristics of the probability distribution of SWAT parameters derived from all the three approaches

| Parameter | Probability distribution function characteristics | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | GLUE with LHS + MI | | | GLUE with LHS | | | GLUE | | |
| SURLAG | Beta general | a1 | 0.4441 | Not identifiable | | | Not identifiable | | |
| | | a2 | 0.9090 | | | | | | |
| | | Min | 1.0197 | | | | | | |
| | | Max | 2.0955 | | | | | | |
| ESCO | Beta general | a1 | 1.0037 | Beta general | a1 | 1.0542 | Beta general | a1 | 1.0625 |
| | | a2 | 0.5059 | | a2 | 0.9074 | | a2 | 0.9005 |
| | | Min | 0.7027 | | Min | 0.7748 | | Min | 0.3211 |
| | | Max | 0.9729 | | Max | 0.9852 | | Max | 0.9957 |
| CN | Beta general | a1 | 1.5526 | Beta general | a1 | 0.9260 | Beta general | a1 | 0.9436 |
| | | a2 | 1.0585 | | a2 | 0.9240 | | a2 | 0.9263 |
| | | Min | 0.0426 | | Min | −0.0575 | | Min | −0.1816 |
| | | Max | 0.1496 | | Max | 0.1481 | | Max | 0.1481 |
| AWC | Beta general | a1 | 2.7859 | Beta general | a1 | 0.9143 | Beta general | a1 | 0.9301 |
| | | a2 | 1.5593 | | a2 | 0.8451 | | a2 | 0.9153 |
| | | Min | −0.2219 | | Min | −0.2466 | | Min | −0.2785 |
| | | Max | 0.1974 | | Max | 1.6905 | | Max | 1.9148 |

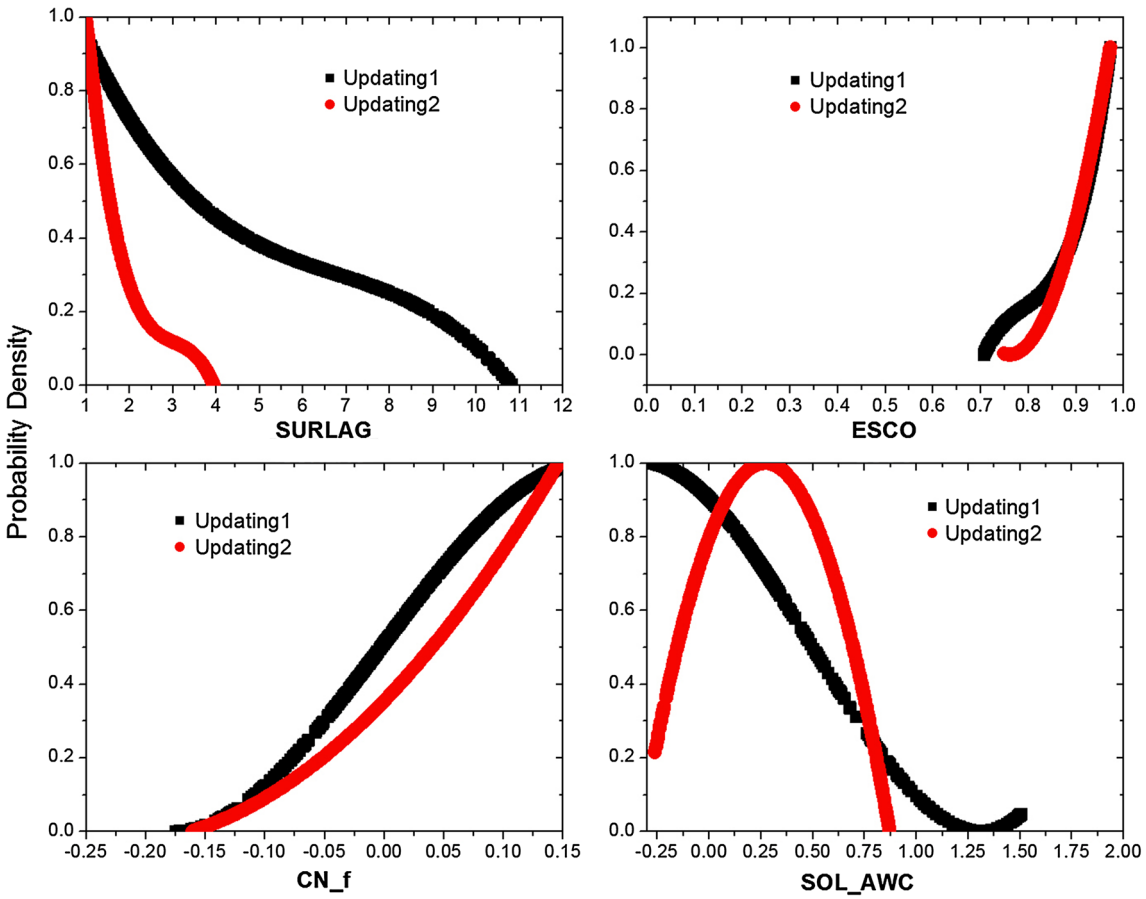


**Fig. 3** Plot of the derived probability distribution functions of the parameters in the proposed method

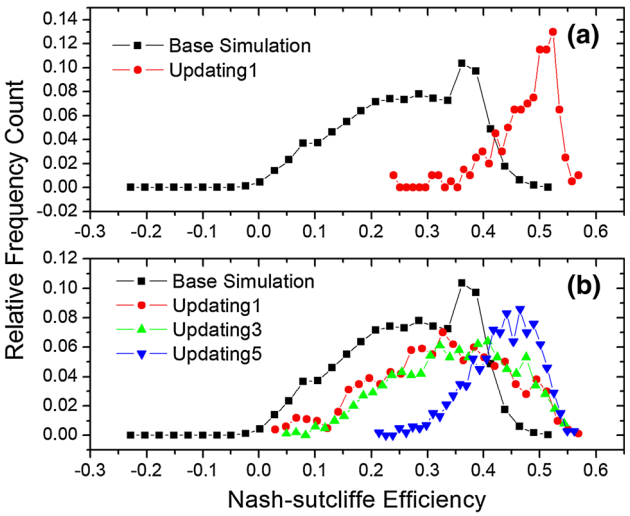


**Fig. 4** Plot of frequency histogram of Nash-Sutcliff efficiency in different iterations for both the approaches **a** from proposed method **b** from Cibin (2014)

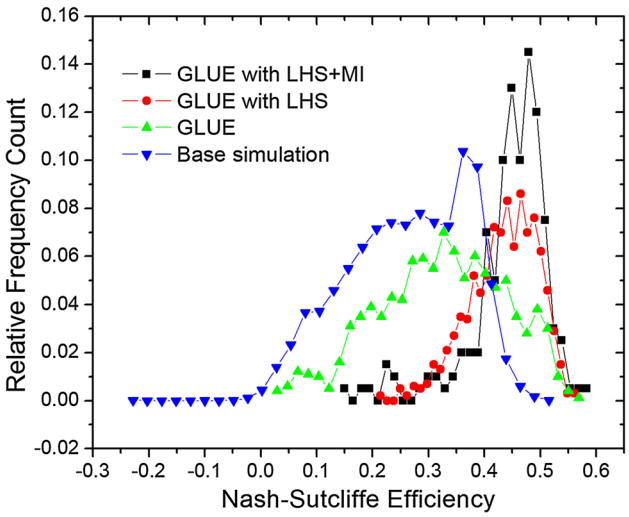


**Fig. 6** Plot of frequency histogram of Nash-Sutcliff efficiency in the final ensemble simulation for different approaches

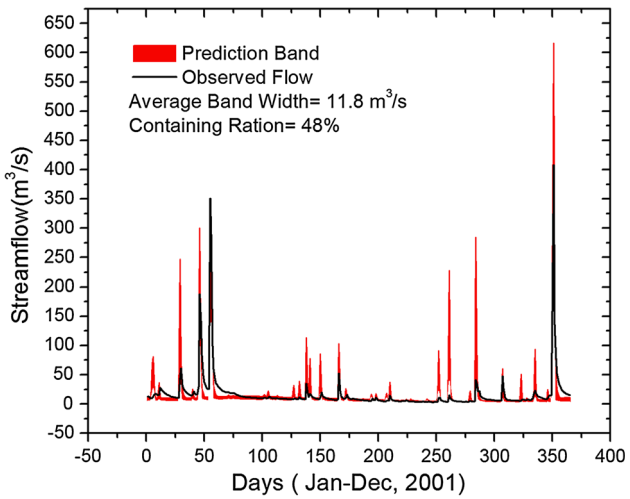


**Fig. 5** The streamflow prediction interval estimated from the ensembles generated using the proposed method for Illinois river basin

## 4.5 Ensemble simulation in Illinois basin

The probability distribution functions of the parameters derived in the proposed method were used for ensemble simulations in Illinois River watershed. The prediction band from the method along with observed data is given in Fig. 5. It is clear from Fig. 5 that the prediction band from the proposed method is narrow, and almost all points of observed data is falling inside the band. The average band width and containing ratio are the two indices commonly employed for analyzing the ensemble simulation (Xiong et al. 2009). The proposed method showed a better prediction with an average bandwidth of 11.8 $m^3/s$ and a containing ratio of 48 %. The improvement in the reduction of parametric uncertainty is visible in the ensemble predictions also. The results indicate that the proposed method simultaneously improves the model predictions and reduces the predictive uncertainty, with less number of iterations (Fig. 6).

## 4.6 Comparison of the proposed method with Cibin (2014) and standard GLUE

The base simulations were same for both the approaches with a model performance of −0.2412 to 0.5282 NSE. A threshold of zero was considered for the first updating in the GLUE with LHS (Cibin 2014) method; however the proposed method (GLUE with LHS + MI) considered a higher threshold 0.45 NSE in the first updating itself. One of the sensitive parameter SURLAG was not identifiable in the GLUE with LHS method, while all the four parameters were identifiable in the current method. The probability distribution functions of the other three parameters were derived using the Bayes' theorem, and checked for the convergence of the distributions (prior and posterior) after every iteration. If the distributions are not converging, 1,000 parameter sets were generated from these derived distributions for the simulations in the next iteration. The threshold value has increased in each iteration and the probability distribution functions got converged after fifth iteration with a threshold of 0.45 NSE in GLUE with LHS method. The best fit standard distribution was identified using a Best Fit program (Palisade Corp. CA, USA). In this case also, all the parameters followed a beta general

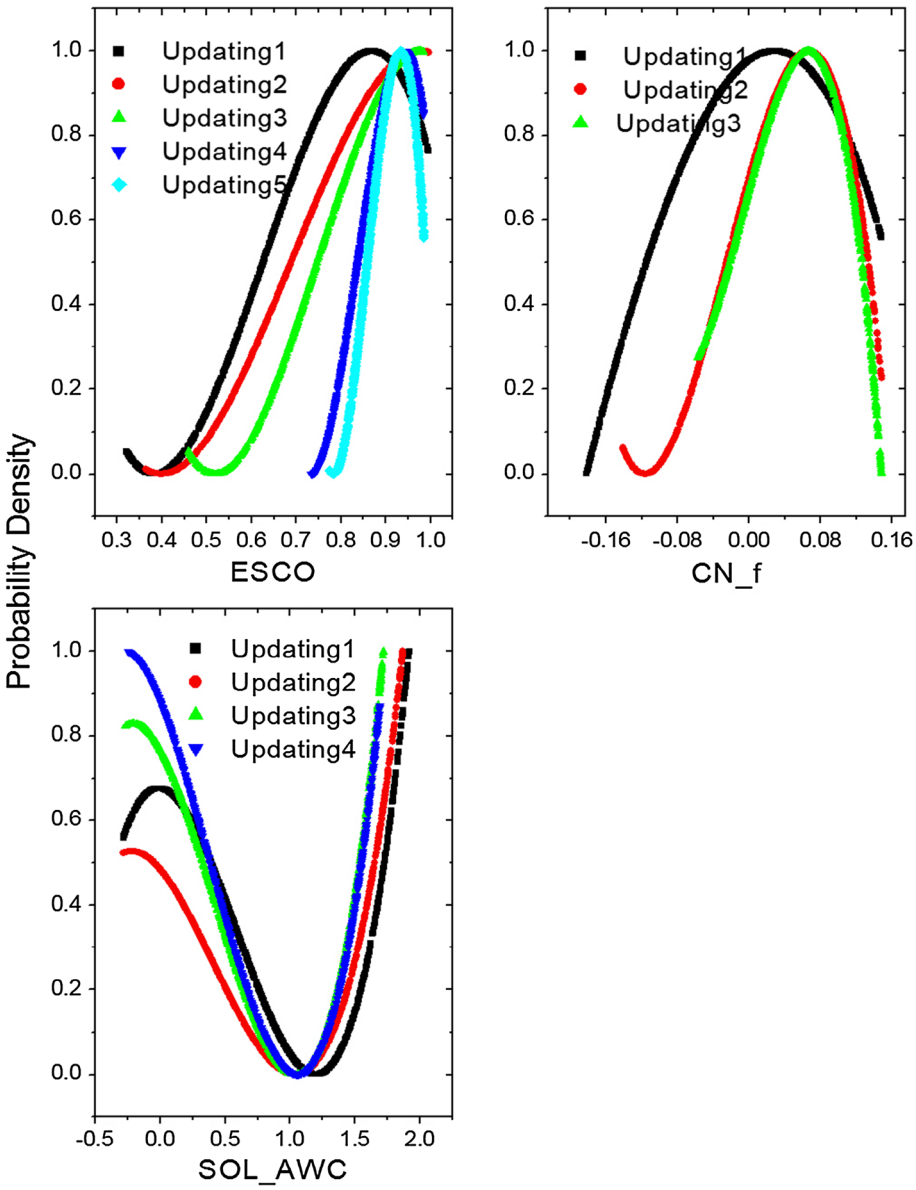


**Fig. 7** Plot of the probability distribution functions of parameters in the GLUE with LHS method

distribution with different shape and location parameters. The characteristics of these distributions are also given in Table 2, and are presented in Fig. 7 for visual inspection. The probability distribution functions derived from both the methods (proposed method and Cibin 2014) are almost similar except in the case of SOL_AWC. The probability distribution function for the parameter SOL_AWC is with multiple peaks, and hence the derivation of original probability distribution function is a difficult task.

The relative frequency count of the Nash–Sutcliffe efficiency in different iterations of the GLUE with LHS method is given in Fig. 4b. Note that the performance frequency count of only three iterations is included in Fig. 4b: base simulation, third updating and fifth updating for the clarity of the figure. The performance range in terms of NSE for the third iteration was 0.05–0.56 with a peak at 0.40 NSE. The range for the last (fifth) updating was 0.21–0.56 with a peak at 0.48 NSE. Though uncertainty reduction is happening in each iteration, the improvement rate is less in this method compared to the proposed method. Therefore, the uncertainty reduction in the proposed GLUE with LHS + MI method is much more fast and effective than the GLUE with LHS method. Further, the proposed method considered less number of iterations for the analysis compared to Cibin (2014).

A comparison of the reduction in parametric uncertainty in different approaches is given in Fig. 6. Different cases are (i) base simulation, which is assumed that all parameters follow uniform distribution (ii) Standard GLUE (iii) GLUE with LHS method (Cibin 2014) (iv) GLUE with

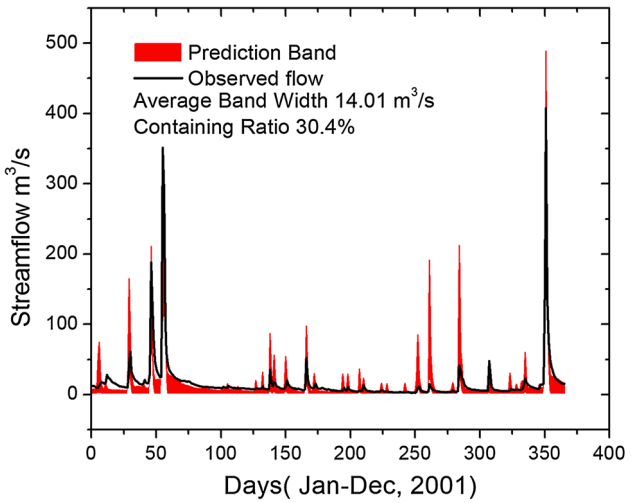


**Fig. 8** The streamflow prediction interval estimated from the ensembles generated using the GLUE with LHS method for Illinois river basin

LHS + MI, proposed method. As mentioned earlier a uniform distribution was assumed for all parameters in base simulation, so the uncertainty is maximum in this case (Case i) which shows an NSE range of −0.2412 to 0.5282. The peak frequency was at a value of 0.35 NSE in base simulation. Case ii is the Standard GLUE approach; here the initial uniform distribution was conditioned with the likelihood values of simulation. The result indicates an improvement in the reduction of uncertainty with an improved performance range of 0.02–0.578 NSE with a peak frequency at 0.35 NSE. Case iii is GLUE with LHS method; here the Standard GLUE was used in an iterative manner and derived the original probability distribution function of parameters. The output uncertainty in terms of NSE was reduced to 0.21–0.566, with a maximum count at 0.48 NSE after five iterations. Case iv is the proposed approach; here the mutual relationship between the parameters was also considered while deriving the original probability distribution function of parameters. The prediction uncertainty has considerably reduced in this approach with an NSE range varying from 0.15 to 0.59, with maximum frequency count at 0.5 NSE. These results clearly indicate the advantage of the proposed method over other methods in terms of the reduction in uncertainty.

The probability distribution functions derived using Cibin (2014) method was used for ensemble prediction in the Illinois basin. The prediction band along with the observed data points are given in Fig. 8. The prediction band was wide with an average band width of 14.01 $m^3/s$ and a containing ratio of 30.4 %. The proposed method was giving a better prediction with less average band width and more containing ratio (Fig. 5). This result indicates the competency of the proposed approach over Cibin (2014) while performing the uncertainty analysis in complex physically based distributed models. This may plausibly be due the incorporation of the MI between parameters for sampling in the procedure for probability distribution function derivation.

### 4.7 Reduction in computational requirement

The proposed method required only one updating to get converged probability distribution functions of parameters. The likelihood values after first updating in the proposed method showed that 60 % of the parameter sets are performing above the threshold. On the other hand, in the GLUE with LHS method (Cibin 2014), 44 % of the parameter sets were above the threshold in the simulations that used the converged probability distribution functions. The number of simulations required to get converged probability distribution functions of parameters were more than 5,000 in GLUE with LHS method. While the number of simulations required in the proposed method (GLUE with LHS + MI) was only 200. The number of simulations required for the uncertainty analysis was reduced by 90 % in the proposed method compared to Cibin (2014). The proposed method achieved 59 % efficiency amongst the 200 ensembles generated, while Cibin (2014) required 5,000 simulations even to achieve 56.6 % efficiency. This shows that the conditional sampling with MI is effective for reducing the computational demand and parametric uncertainty in complex physically based distributed models.

## 5 Summary and conclusions

The physically based distributed models are process based models with a complex model structure. The over-parameterization is the major skepticism involved in the application of these models. The calibration is difficult in these models because of equifinality and over-parameterization. An ensemble based approach will be better than calibrating the model because of the various uncertainties involved in these complex models. The ensemble based methods require the probability distribution functions of the parameters, which is unknown. Cibin (2014) proposed a method for deriving the probability distribution functions of the parameters in a GLUE framework. However the computational demand required in this method is very high. The current study proposed a method for deriving the probability distribution functions of the parameters by considering the mutual relationship between the parameters that help reduce the computational requirement for analysis.

The inclusion of MI during sampling increased the behavioral parameter sets in the Monte Carlo simulation, and it required less number of simulations in each iteration. The proposed method was showing a better prediction band with less average band width and more containing ratio. The number of simulations required was reduced by 90 % in the MI based approach with improvement over the reduction in uncertainty. The LHS in the GLUE can be replaced by other advanced sampling methods like importance sampling, Marko Chain Monte Carlo sampling etc. (Blasone et al. 2008a, b; Vrugt et al. 2003, 2009; Shafii et al. 2014). These sampling methods may improve the reduction in the predictive uncertainty; however one cannot assure the reduction in the computational burden. The inclusion of mutual information along with improved sampling methods may give better predictions with less computational burden.

## References


Arabi M, Govindaraju RS, Engel B, Hantush M (2007) Multiobjective sensitivity analysis of sediment and nitrogen processes with a watershed model. Water Resour Res 43(6):W06409. doi:10.1029/2006WR005463

Arnold JG, Srinivasan R, Muttiah RS, Williams JR (1998) Large area hydrologic modeling assessment: part I model development. J Am Water Resour Assoc 34(1):73–89

Athira P, Sudheer KP, Cibin R, Chaubey I (2011) Sensitivity analysis of stream flow generation parameters of SWAT model. Paper presented at the annual conference of the ASABE, Louisville, KY, paper No. 1111731

Beven KJ (1989) Changing ideas in hydrology: a case of physically based models. J Hydrol 105:157–172

Beven KJ (2000) On the future of distributed modeling in hydrology. Hydrol Process 14(16–17):3183–3184

Beven KJ (2006) A manifesto for the equifinality thesis. J Hydrol 320:18–36

Beven KJ, Binley AM (1992) The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process 6:279–298. doi:10.1002/hyp.3360060305

Beven K, Freer J (2001) Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology. J Hydrol 249(14):11–29

Blasone RS, Madsen H, Rosbjerg D (2008a) Uncertainty assessment of integrated distributed hydrological models using GLUE with Markov chain Monte Carlo sampling. J Hydrol 353:18–32

Blasone RS, Vrugt JA, Madsen H, Rosbjerg D, Robinson BA, Zyvoloski GA (2008b) Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. Adv Water Resour 31:630–648

Brigode P, Oudin L, Perrin C (2013) Hydrological model parameter instability: a source of additional uncertainty in estimating the hydrological impacts of climate change? J Hydrol 476:410–425. doi:10.1016/j.jhydrol.2012.11.012

Carota C, Parmigiani G, Polson NG (1996) Diagnostic measures for model criticism. J Am Statist Assoc 91:753–762

Chen J, Wu Y (2012) Advancing representation of hydrologic processes in the soil and water assessment tool (SWAT) through integration of the topographic model (TOPMODEL) features. J Hydrol 420–421:319–328

Cibin R, Sudheer KP, Chaubey I (2010) Sensitivity and dentifiability of stream flow generation parameters of the SWAT model. Hydrol Process 24(9):1133–1148

Cibin R, Athira P, Sudheer KP, Chaubey I (2014) Application of distributed hydrological models for predictions in ungauged basins: a method to quantify predictive uncertainty. Hydrol Process. doi:10.1002/hyp.9721

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6:182–197

Demaria EM, Njissen B, Wagener T (2007) Monte Carlo sensitivity analysis of land surface parameters using the variable infiltration capacity model. J Geophys Res 112:D11113

Doherty J (2004) PEST model—independent parameter estimation. Watermark Numerical Computing, Brisbane

Duan Q, Sorooshian S, Gupta VK (1992) Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour Res 28(4):1015–1031

Freer J, Beven KJ, Ambroise B (1996) Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. Water Resour Res 32(7):2161–2173

Gardner RH, O'Neill RV (1983) Parameter uncertainty and model predictions: a review of Monte Carlo results. In: Berk MB, Straten GV (eds) Uncertainty and forecasting of water quality. Springer, New York, pp 245–257

Gassman PW, Reyes MR, Geen CH, Arnold JG (2007) The soil and water assessment tool: historical development, applications and future research directions. Trans ASABE 50(4):1211–1250

Hughes DA (2010) Hydrological models: mathematics or science? Hydrol Proces 24:2901–2201

Jia Y, Culver TB (2008) Uncertainty analysis for watershed modeling using generalized likelihood uncertainty estimation with multiple calibration measures. J Water Resour Plan Manage 134(2):97–106

Jin X, Chong-Yu Xu, Zhang Q, Singh VP (2010) Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. J Hydrol 383:147–155

Katz RW (2002) Techniques for estimating uncertainty in climate change scenarios and impact studies. Climate Res 20:167–185

Kuczera G, Parent E (1998) Monte Carlo assessment of parameter uncertainty catchment models: the Metropolis algorithm. J Hydrol 211:69–85

Kullback S (1959) Information theory and statistics. Wiley, New York

Li L, Xia J, Xu C-Y, Singh VP (2010) Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models. J Hydrol 320:132–154

MacKay DJC (2003) Information theory, inference, and learning algorithms. Cambridge University Press, Cambridge (ISBN 0-521-64298-1)

Manache G, Melching CS (2008) Identification of reliable regression- and correlation-based sensitivity measures for importance ranking of water-quality model parameters. Environ Model Softw 23:549–562

Migliaccio KW, Chaubey I (2008) Spatial distributions and stochastic parameter influences on SWAT flow and sediment predictions. J Hydrol Eng 13(4):258–269

Neitsch SL, Arnold JG, Kiniry JR, Williams JR, King KW (2002) Soil water assessment tool theoretical documentation. Version 2000. Texas Water Resource Institute, College Station, Texas. TWRI Report, TR-191

Rastetter EB, King AW, Cosby BJ, Hornberger GM, O'Neill RV, Hobbie JE (1992) Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecol Appl 2:55–70

Razavi S, Tolson BA (2013) An efficient framework for hydrologic model calibration on long data periods. Water Resour Res 49:8418–8431. doi:10.1002/2012WR013442

Refsgaard JC, Sluijs JP, Hojberg AL, Vanrolleghem PA (2007) Uncertainty in the environmental modeling process—a framework and guidance. Environ Model Softw 22:1543–1556

Renard B, Kavetski D, Kuczera G, Thyer M, Franks SW (2010) Understanding predictive uncertaintyin hydrologic modeling: the challenge of identifying input and structural errors. Water Resour Res 46:W05521. doi:10.1029/2009WR008328

Shafii M,Tolson B, Matott LS (2014) Uncertainty based multi-criteria calibration of rainfall-runoff models: a comparative study. Stochast Environ Res Risk Assess. doi:10.1007/s00477-014-0855-x

Shen ZY, Chen L, Chen T (2012) The influence of parameter distribution uncertainty on hydrological and sediment modeling: a case study of SWAT model applied to the Daning watershed of the three gorges reservoir region, China. Stochast Environ Res Risk Assess. doi:10.1007/s00477-012-0579-8

Smith PJ, Tawn J, Beven KJ (2008) Informal likelihood measures in model assessment: theoretic development and investigation. Adv Water Resour. doi:10.1016/j.advwatres.2008.04.012

Sobol IM (1993) Sensitivity estimates for nonlinear mathematical models. Math Model Comput Exp 1:404–414

Srivastav RK, Sudheer KP, Chaubey I (2007) A simplified approach to quantify predictive and parametric uncertainty in artificial neural network hydrologic models. Water Resour Res 43(10) (Art. No. W10407). doi:10.1029/2006WR005352

Stedinger JR,Vogel RM, Lee SU, Batchelder R (2008) Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. Water Resour Res 44. doi:10.1029/2008WR006822

Tourassi GD, Frederick ED, Markey MK, Carey E, Floyd J (2001) Application of the mutual information criterion for feature selection in computer-aided diagnosis. Med Phys 28:2394–2402

Vrugt JA, Gupta HV, 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 JA, ter Braak CJF, Gupta HV, Robinson BA (2009) Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stoch Env Res Risk Assess 23(7):1011–1026. doi:10.1007/s00477-008-0274-y

Wang G, Xia J, Chen J (2009) Quantification of effects of climate variations and human activities on runoff by a monthly water balance model: a case study of the Chaobai river basin in northern China. Water Resour Res 45:W00A11. doi:10.1029/2007WR006768

Wiwatenadate P, Claycamp HG (2000) Error propagation of uncertainties in multiplicative models. Human Ecol Risk Assess 6:355–368

Wu Y, Liu S (2012) Automating calibration, sensitivity and uncertainty analysis of complex models using the R package flexible modeling environment (FME): SWAT as an example. Environ Model Softw 31:99–109

Wu Y, Liu S (2014) Improvement of the R-SWAT-FME framework to support multiple variables and multi-objective functions. Sci Total Environ 466–467:455–466

Xiong LH, Wan M, Wei XJ, O'Connor KM (2009) Indices for assessing the prediction bounds of hydrological models and application by generalised likelihood uncertainty estimation. Hydrol Sci J (J Sci Hydrologiques) 54(5):852–871

Yang J, Reichert P, Abbaspour KC, Xia J, Yang H (2008) Comparing uncertainty analysis techniques for a SWAT application to the Chaohe basin in China. J Hydrol 358:1–23

Zitzler E, Laumanns M, Thiele L (2001) SPEA2: improving the performance of the strength Pareto evolutionary algorithm. Tech. Report 103. Zurich, Switzerland: Swiss Federal Institute of Technology (ETH), Computer Engineering and Communication Networks Lab (TIK)