Evaluation of Estimation of Distribution Algorithm to Calibrate Computationally Intensive Hydrologic Model

Zejun Li[1]; Pan Liu[2]; Chao Deng[3]; Shenglian Guo[4]; Ping He[5]; and Caijun Wang, Ph.D.[6]

Abstract: The estimation of distribution algorithm (EDA) is a new evolutionary algorithm developed as an alternative to the traditional genetic algorithm (GA). The EDA guides the search by avoiding the crossover and mutation operators of the GA in favor of building and sampling probabilistic distributions of promising candidate solutions. By increasing the probability of generating solutions with better fitness values, the EDA locates the region of the global optimum or its accurate approximation. In this study, EDA was used to calibrate the parameters of the soil and water assessment tool hydrologic model for the Xunhe River Basin in China. The EDA was compared with three other algorithms: (1) the Multistart Local Metric Stochastic Radial Basis Function algorithm (a surrogate optimization method), (2) the Shuffled Complex Evolution algorithm, and (3) the GA. Four metrics are presented to assess the performance of the algorithms: (1) efficiency in terms of the average best objective function value in a limited number of function evaluations, (2) variability in terms of standard deviation and the box plot, (3) reliability in terms of the empirical cumulative distribution function, and (4) accuracy in terms of the Nash–Sutcliffe efficiency coefficient and overall volume error. Results indicated that the EDA is more efficient and could provide more accurate solutions with a relatively high probability, at least for this case study. **DOI: 10.1061/(ASCE)HE.1943-5584.0001350.** *© 2016 American Society of Civil Engineers.*

Author keywords: Estimation of distribution algorithm; Hydrologic model; Parameter calibration.

Introduction

The International Association of Hydrological Sciences Scientific Decade 2013–2022, entitled "Panta Rhei—Everything Flows," is dedicated to research activities on change in hydrology and society (Montanari et al. 2013). Hydrologic models are needed to better understand the behavior of hydrologic and social systems (Deng et al. 2015b), but the effective use of these models is challenging because parameters have to be appropriately calibrated. Calibration can be implemented manually or automatically. With recent increases in computational power, most studies are adopting automatic procedures for model calibration. In the process of automatic hydrologic model calibration, parameters are adjusted using an optimization strategy according to a specific goodness-of-fit measure, the so-called objective function, which describes the discrepancy between the simulated and observed streamflow for a basin outlet (Deng et al. 2015a).

[1]Ph.D. Student, State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan Univ., Wuhan 430072, China.

[2]Professor, State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan Univ., Wuhan 430072, China (corresponding author). E-mail: liupan@whu.edu.cn

[3]Ph.D. Student, Hubei Provincial Collaborative Innovation Center for Water Resources Security, College of Water Resources and Hydropower, Wuhan Univ., No. 8, Donghu Rd. (South), Wuhan 430072, China.

[4]Professor, Hubei Provincial Collaborative Innovation Center for Water Resources Security, College of Water Resources and Hydropower, Wuhan Univ., No. 8, Donghu Rd. (South), Wuhan 430072, China.

[5]Professor, China International Engineering Consulting Corporation, Zhongzi Bldg., No. 32, Chegongzhuang Rd. (West), Beijing 100048, China.

[6]China International Engineering Consulting Corporation, Zhongzi Bldg., No. 32, Chegongzhuang Rd. (West), Beijing 100048, China.

Most early automatic calibration strategies use local search methods. However, local search methods can get trapped in the first local optimum they find. For hydrologic models, there may be multiple local optima on the objective function surface. Therefore, the performance of a local search method depends on its initial solution.

To overcome this limitation, global search methods have been used. One important branch of heuristic optimization techniques for the global search is the population-evolution-based search strategy, which includes the widely-used genetic algorithm (GA) (Goldberg 1989) and shuffled complex evolution (SCE) algorithm (Duan et al. 1992). The basic concept of the SCE and GA algorithms is to maintain a population of solutions, and to develop potential solutions toward the region of the global optimum of the objective function. A large number of studies have been conducted demonstrating that the SCE and GA algorithms are suitable for optimization problems in water resources (Duan et al. 1994; Liu et al. 2011; Sorooshian et al. 1993; Wang 1991; Yapo et al. 1996).

However, as well as satisfying calibration results, the computational burdens involved in using traditional global search methods must be considered because these results usually require a large number of function evaluations. Furthermore, running complex hydrologic models is time consuming, which makes it impractical to implement large numbers of model simulations. In such situations, algorithms that yield better solutions with a limited number of function evaluations would be preferred.

The function approximation method, also known as the surrogate response surface method, uses a response surface to approximate the objective function, consequently reducing the computational budget. The method has been applied widely to hydrologic model calibration problems (Razavi et al. 2012). There are many functional approximation models including radial basis function (RBF) models (Gutmann 2001; Ishikawa et al. 1999) and kriging models (Booker et al. 1999; Simpson et al. 2001). For example, the Multistart Local Metric Stochastic Radial Basis Function (MLMSRBF) algorithm, in which the approximation is built with the RBF (Regis and Shoemaker 2007), has been

effectively used for groundwater bioremediation model calibration (Mugunthan et al. 2005; Razavi et al. 2012; Regis and Shoemaker 2007).

The estimation of distribution algorithm (EDA) is a novel population-based evolutionary algorithm that differs from the GA in the way it generates new candidate solutions and propagation between generations (Fu et al. 2006). The GA uses genetic operators of crossover and mutation to generate/update the population of candidate solutions, whereas the EDA replaces the genetic operators with modeling and exploiting the probability distribution of the promising solutions. The model is sampled generating new candidate solutions, which are reincorporated to update the model. The probability distribution model in the EDA captures the features of promising solutions to increase the probability of generating the global optimum.

The potential of the EDA has been demonstrated across a broad spectrum of optimization problems, such as environmental monitoring network design (Kollat et al. 2008), environmental economic dispatch (Li et al. 2015), military antenna design (Santarelli et al. 2006), and flowshop scheduling problems (Jarboui et al. 2009). However, few attempts have been made to tailor the EDA for hydrologic watershed model calibration and to measure its performance.

In this study, four different automated strategies, the EDA, MLMSRBF algorithm, SCE algorithm, and GA, were used to calibrate parameters of the soil and water assessment tool (SWAT) model (Arnold et al. 1998). The efficiency, variability, reliability, and accuracy of these methods were evaluated, and the performance of the EDA was examined to determine whether it could represent an improvement over other methods.

The remainder of this paper is organized as follows: (1) a description of the EDA and the other three automatic calibration algorithms; (2) the algorithms are applied to calibrate the parameters of the SWAT model for the Xunhe River Basin in China, and the results are analyzed and compared; and (3) the conclusions are presented.

Methodology

The parameter calibration of hydrologic models can be generalized as a global minimization problem with box constraints

$$\min f(x_i) \qquad (1)$$

subject to

$$x_{i,\min} < x_i < x_{i,\max}, \qquad i = 1, \ldots, n \qquad (2)$$

where x_i = model parameter value that needs to be calibrated, varying from $x_{i,\min}$ to $x_{i,\max}$. The bounds $x_{i,\min}$ and $x_{i,\max}$ should be on the basis of a physically feasible range of parameters. The $f(x)$ is the objective function on the basis of the error between the observed and simulated streamflow of the basin outlet. In most cases, $f(x)$ is a multimodal function without prior information.

Four automatic search methods, the EDA, MLMSRBF algorithm, SCE algorithm, and GA, were tailored for this minimization problem and compared on the basis of their performance. These search methods can be generalized because no problem-dependent information is needed.

Estimation of Distribution Algorithm

The EDA was developed as a novel class of evolutionary optimization algorithm (Mühlenbein et al. 1996). The main advantage of

the EDA over the GA is the absence of genetic operator parameters to be tuned (Armananzas et al. 2008) (Fig. 1).

Selection Method

After the generation of initial population and entering the iteration for solution search, promising solutions were selected to construct the probability model on the basis of a selection method. The selection method employed in this study was the truncation method (Lima et al. 2007), which is one of the most commonly used methods in the EDA. In the truncation method, solutions are ranked according to the fitness values (i.e., the objective function values). A proportion of the best solutions in the current population were selected.

Probabilistic Model

The presentation of the probability model is crucial when using an EDA because of its impact on EDA efficiency and applicability. Probabilistic graphical models (PGMs) combine probability theory with graph theory and can provide a flexible framework to model the interactions among variables. In this study, a PGM with a Gaussian network structure was used to present the probability model. Gaussian network models are the most commonly used probabilistic distribution models for continuous optimization.

A Gaussian network is a graphical representation of the probabilistic relationship among a set of random variables. It consists of two components, the structure and corresponding parameters (Lima et al. 2007). The structure is a directed acyclic graph (DAG), in which nodes correspond to the random variables $X = (X_1, \ldots, X_n)$ and edges represent conditional dependencies among variables. The DAG encodes a joint probability distribution in which the factorization can be written as

$$p(x) = \prod_{i=1}^{n} p(x_i | pa_i) \qquad (3)$$

where x = value of X; pa_i = set of variables corresponding to the initial node of the directed edge in DAG, which is probabilistically depended upon by the variable x_i corresponding to the terminal node of this edge; and $p(x_i | pa_i)$ = conditional probability of x_i given its parents pa_i.

The parameters are local density functions that present the conditional probabilities for each variable. The local density function in a Gaussian network is a linear regression model (Larranaga and Lozano 2001)

$$f(x_i | pa_i) \equiv \mathcal{N}\left[x_i; m_i + \sum_{x_j \in pa_i} b_{ji}(x_j - m_j), v_i\right] \qquad (4)$$

where $\mathcal{N}(x; \mu, \sigma^2)$ = univariate normal distribution with mean μ and variance σ^2; m_i = unconditional mean of X_i; b_{ji} = linear coefficient reflecting the strength of the relationship between X_i and X_j; and v_i = conditional variance of X_i given pa_i. The Bayesian information criterion (BIC) (Cooper and Herskovits 1992; Schwarz 1978) was used to evaluate the Gaussian network structure in the greedy network construction algorithm (K2 algorithm) (Cooper and Herskovits 1992).

Sampling Method

Once the structure of the Gaussian network is built and its parameters estimated, new solutions are generated by sampling from this distribution. In this study, sampling was conducted by using the probabilistic logic sampling (PLS) method (Henrion 1986), which is a forward sampling strategy.
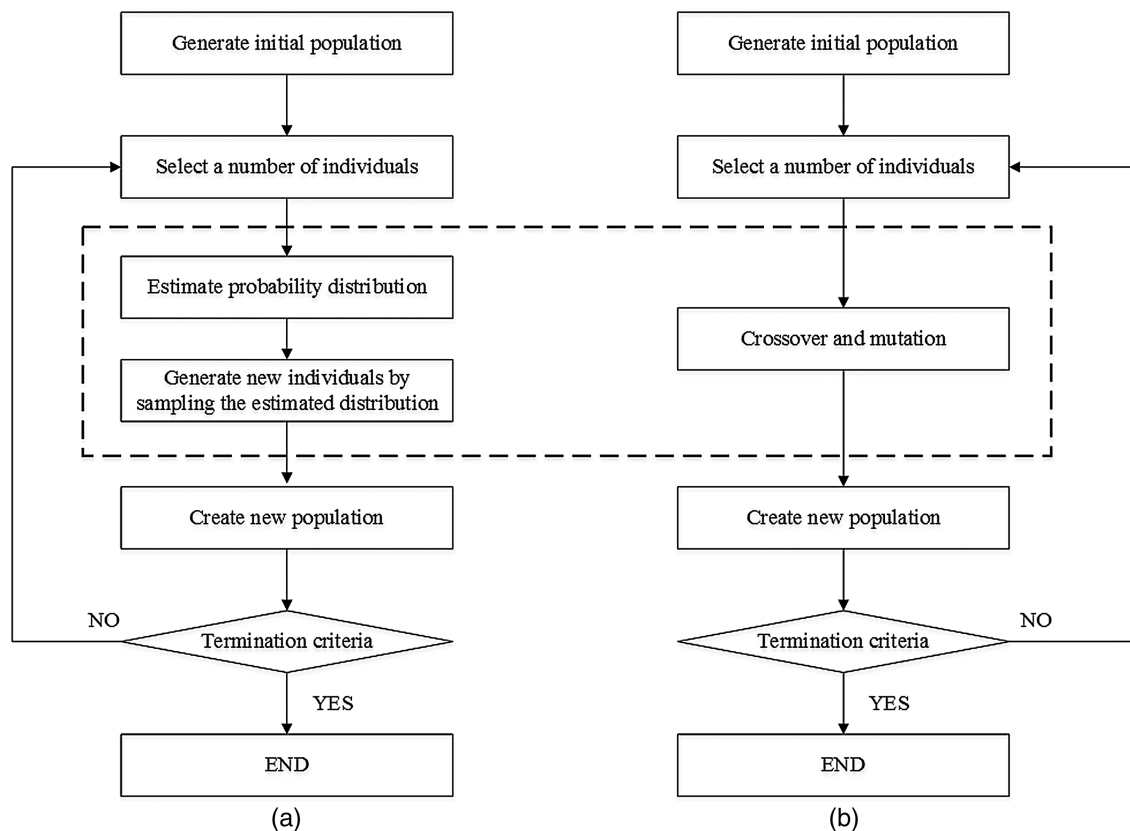
Fig. 1. Flow chart of (a) the estimation of distribution algorithm (EDA); (b) the genetic algorithm (GA)

Replacement Method

The replacement method determines how new solutions are incorporated into the previous population. It is an effective way to increase the diversity in the population. In this study, the elitist replacement method was used, which is appropriate for the truncation method in the EDA (Lima et al. 2007). In the elitist replacement method, solutions are ordered according to the ranking of their fitness values and a proportion of the worst are replaced with new solutions. The number of replaced solutions is equal to that of new solutions, keeping the population size unchanged.

Genetic Algorithm

The GA is a loosely modeled optimization methodology, and is the most widely referenced and used automatic optimization method (Goldberg 1989). It is a population-based heuristic search algorithm, simulating the evolutionary process of natural selection and genetics. As shown in Fig. 1, the GA method employs a population of individuals evolving through selection and recombination, which incorporates mutation and crossover operators. The process iterates, and individuals with better fitness values are selected more easily to the next generation.

Shuffled Complex Evolution Algorithm

The SCE algorithm was developed for global optimization problems by simulating the natural evolution. It has been applied widely to automatic calibration of hydrologic models and has been proven to be effective and reliable (Duan et al. 1994; Sorooshian et al. 1993). In the SCE algorithm, the initial population is generated randomly. The population is then divided into several communities represented by complexes in the algorithm. Every community evolves independently and each point has the probability to

generate offspring. To ensure the competitiveness of evolution, each individual is assessed by its corresponding objective function value. Better individuals have higher competitiveness contributing to reproduction. After several generations, communities are updated through a process of shuffling. This procedure shares the information contained in every community and the searching space is thoroughly explored.

Multistart Local Metric Stochastic Radial Basis Function Algorithm

The response surface surrogates use data-driven function approximation techniques to approximate the response of the original model (Razavi et al. 2012), which emulate the relationship between model parameters and model outputs. The MLMSRBF algorithm (Regis and Shoemaker 2007), one of the response surface surrogate methods, is a global (multistart global) optimization algorithm in which a RBF model is employed as the response surface to approximate the objective function and to identify promising solutions. To balance the global search and local search, two criteria are considered: the RBF criterion (the objective function value predicted by the RBF model) and the distance criterion (the distance between the new and previously evaluated points) (Espinet and Shoemaker 2013). The RBF model is constructed and updated in each iteration.

Case Studies

Xunhe River Basin

The study area, Xunhe River Basin, is located between 108°24′–109°26′E and 32°52′–33°55′N, and has a surface area of
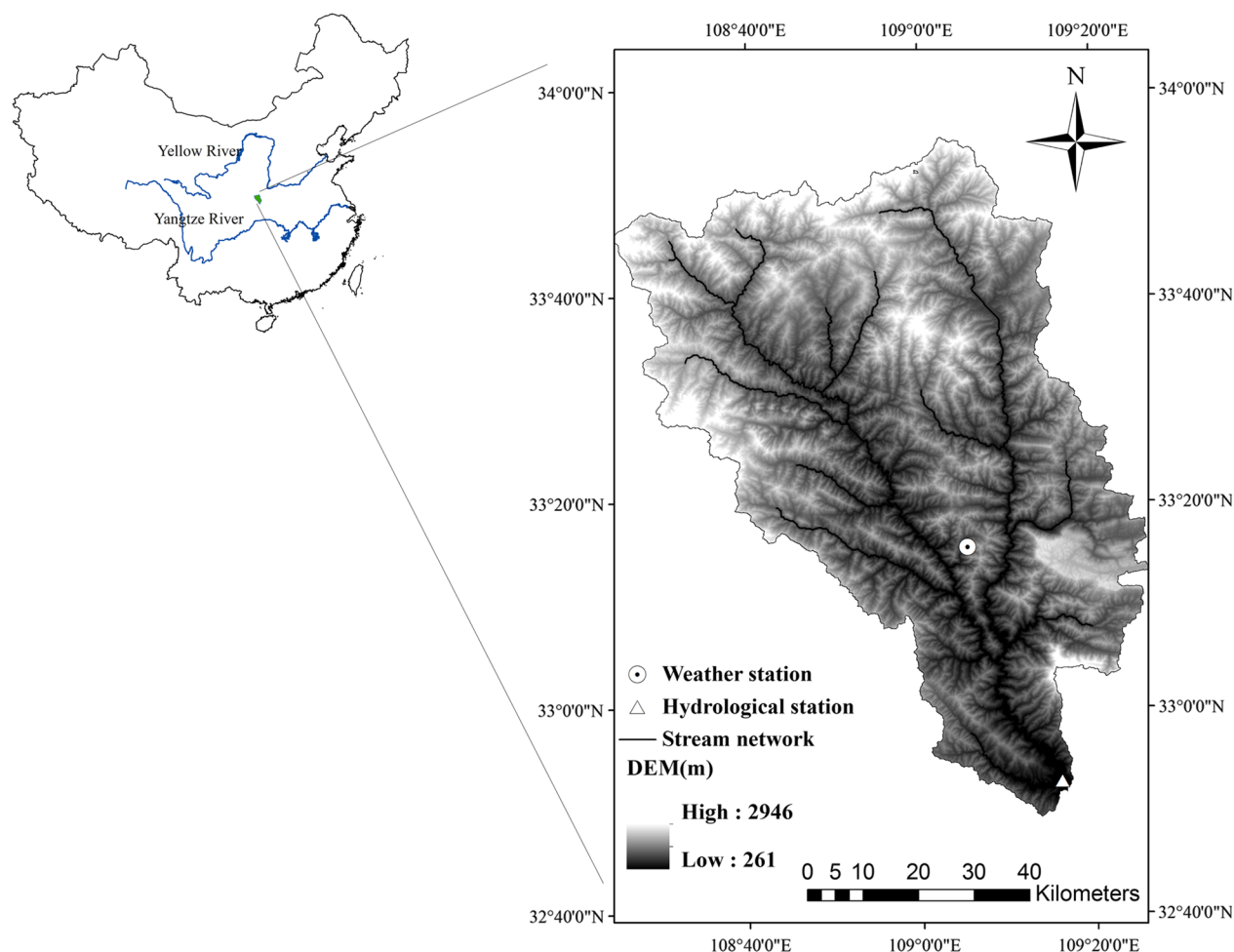
© ASCE

04016012-3

J. Hydrol. Eng.

Fig. 2. Location of the Xunhe River Basin in China

approximately 6,300 km² (Fig. 2). The elevation ranges from 200 to 3,000 m above mean sea level. The Xunhe River is about 218 km long and has an average annual flow of 73 m³/s. The average annual precipitation and temperature are 850 mm and 15°C, respectively.

In this study, the topography data were determined on the basis of the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) global digital elevation model (GDEM) with a cell size of 30 × 30 m. A digital land cover/land use map was derived from the large-scale (1:100,000) land use map of China, and a digital soil map was generated from the 1:1,000,000 scale soil map of China. A total of 61 subbasins and 487 hydrologic response units (HRUs) were delineated in the study area. Meteorological data (rainfall, temperature, solar radiation, wind speed, and relative humidity) at Zhen'an station were obtained from the China Meteorological Data Sharing Service System. The streamflow data were obtained from the Xiangjiaping hydrologic station, which is located at the basin outlet. Meteorological and hydrological data were based on daily observations and were collected for a period of 11 years from 1980 to 1990. The first eight years and the remaining three years were used for calibration and validation periods, respectively.

The hydrologic model for the study area was established with the SWAT. The SWAT is a continuous-time, distributed parameter model designed to simulate flow, sediment, and the transport of nutrients and pesticides through the watershed system on a daily time step.

Design of Study Cases

Calibrated Parameters in the Soil and Water Assessment Tool Model
Ten flow-related parameters in this study were selected for model calibration. Of these, three governed surface water response, five governed subsurface water response, and two governed basin response. The ranges of these parameters were determined according to van Griensven et al. (2006). General definitions of the parameters are listed in Table 1. The parameters across all spatial units share the same value in the calibration process.

Control Parameters of Algorithms
Algorithms like the SCE and the GA require some user defined algorithm parameters. The most important parameter in the SCE algorithm is the number of complexes p (Duan et al. 1994). In this study, p was set to equal the number of model parameters, as suggested by Kuczera (1997). For other parameters in the SCE algorithm, the default values provided by Duan et al. (1994) were used. In the GA, the probabilities of mutation and crossover were 0.01 and 0.9, respectively. The EDA and GA were both operated with a population size of 20.

Objective Function
The simulation performance of a model is usually determined by pairwise comparisons of model simulation results with observations.

Table 1. Parameters for Calibration in the Soil and Water Assessment Tool (SWAT) Model

Parameters	Definition	Ranges
	Parameters governing surface water response	
CN2	SCS runoff curve number	35–98
ESCO	Soil evaporation compensation factor	0.01–1
SOL_AWC	Available water capacity of the soil layer (mm/mm soil)	0–1
	Parameters governing subsurface water response	
GW_REVAP	Groundwater re-evaporation coefficient	0.02–0.2
GW_DEALY	Groundwater delay (days)	0–500
ALPHA_BF	Baseflow alpha factor (days)	0–1
GWQMN	Threshold depth of water in the shallow aquifer for re-evaporation to occur (mm)	0–5,000
REVAPMN	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	0–500
	Parameters governing basin response	
CH_K2	Effective hydraulic conductivity in main channel alluvium (mm/h)	−0.01–500
SURLAG	Surface runoff lag coefficient	0.05–24

The Nash–Sutcliffe efficiency (NSE) is a commonly used indicator to evaluate the performance of hydrologic models (Legates and McCabe 1999; Nash and Sutcliffe 1970). It is a normalized statistic that determines the relative magnitude of the residual variance compared with the measured data variance. Therefore, the NSE evaluates how well the simulated hydrograph fits the observations.

The NSE is defined as follows:

$$\text{NSE} = 1 - \frac{\sum_{t=1}^{N}(Q_t^{\text{obs}} - Q_t^{\text{sim}})^2}{\sum_{t=1}^{N}(Q_t^{\text{obs}} - Q_{\text{avg}}^{\text{obs}})^2} \tag{5}$$

where Q_t^{obs} and Q_t^{sim} = observed and simulated discharges at time step t, respectively; $Q_{\text{avg}}^{\text{obs}}$ = mean discharge for the entire time period of the evaluation; and N = total number of time steps in the evaluation period. The NSE ranges from minus infinity to 1.0, with higher values indicating better agreement. Because it is a minimization problem, the objective function is selected as 1–NSE.

Results

Generally, algorithm efficacy can be measured by the number of function evaluations required to find the global optimum. However, because of limited time and computing resources, it is impossible to perform large numbers of simulations for a computationally intensive hydrologic model. In this study, on average, one execution of the SWAT model required 63 s on a Core i5 3GHZ processor (Intel, China). Furthermore, previous investigations (Duan et al. 1992; Fu et al. 2006; Goldberg 1989; Mugunthan et al. 2005; Paul and Iba 2002; Vose 1999) have proven that the EDA, MLMSRBF algorithm, SCE algorithm, and GA can converge to the global optimum if adequate computational effort can be implemented. Therefore, in this study the maximum number of function evaluations was limited to 500. As a result, the time consumed by one execution was 6.6 h for EDA, 6.8 h for MLMSRBF, 6.5 h for SCE, and 6.3 h for GA when calibrating the SWAT model for the Xunhe basin. Meanwhile, in consideration of the randomness of these algorithms, each algorithm was implemented 10 times. In each trial, the best function value reached by the algorithm after every function evaluation was recorded.

To assess the performance of the algorithms, the following were compared: (1) efficiency, which is indicated by the average best objective function value; (2) variability, which can be evaluated by the standard variance and interquartile range; (3) reliability, which is shown by the means of the empirical cumulative distribution function; and (4) accuracy, which is estimated by the NSE, overall
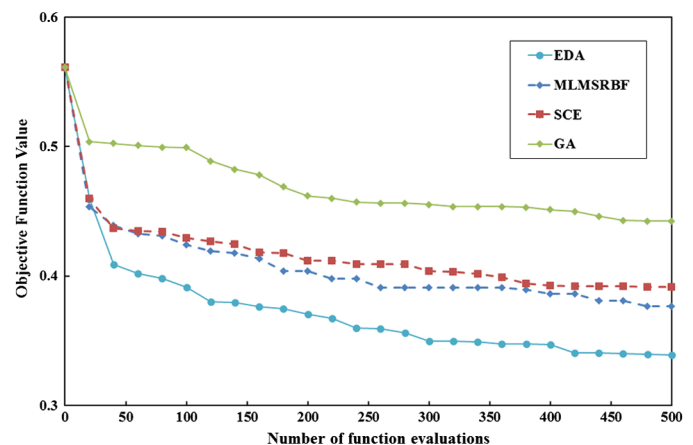
Fig. 3. Comparison of algorithm performance

Table 2. Average and Standard Deviation of the Objective Function Values for Each Algorithm

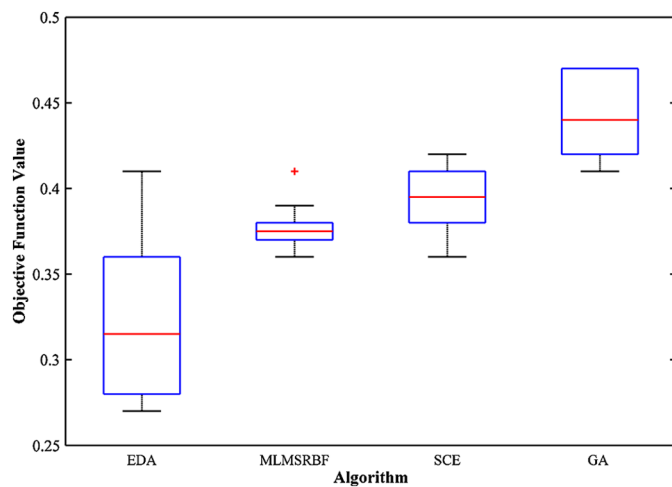| Algorithm	Average	SD
EDA	0.34	0.053
MLMSRBF	0.38	0.016
SCE	0.39	0.022
GA	0.44	0.025

volume error (R) and peak percent threshold statistics (PPTS) (Lohani et al. 2014).

Algorithm Efficiency

The algorithm efficiency is denoted by the average (over 10 trials) best objective value versus the number of model simulations.

The minimization progress for the four algorithms is shown in Fig. 3. The points in the figure represent the average best values of the objective function. The average best value indicates the average objective value for the best solution obtained until that point in the optimization iterates over 10 trials.

The results show that EDA was the most efficient algorithm because it had the lowest best objective value (Table 2). The MLMSRBF algorithm and the SCE algorithm had relatively similar convergence speed in the minimization progress, but the MLMSRBF algorithm reached lower objective values than the

Fig. 4. Box plot of the best solution for each algorithm based on 10 trials showing medians, interquartile ranges, and outliers

Fig. 5. Comparison of the empirical cumulative distribution function for each algorithm

SCE algorithm. The GA was the least efficient method because of its slowest convergence speed and highest average objective values.

Variability in Multiple Trials

Variability describes the stability of the algorithm. In general, the standard deviation and interquartile range can be used to measure the variability. The EDA had the largest standard deviation (SD) of 0.053, which means it had the worst stability among the four algorithms. The MLMSRBF algorithm had the lowest SD of 0.016.

Fig. 4 shows that the EDA had the largest dispersion (interquartile range) and the lowest median in 10 trials, as indicated by the graphical box plot comparisons. The MLMSRBF algorithm had the second lowest median for the objective function value and the smallest interquartile range. Similar to the SD summarized in Table 2, the SCE algorithm had the second smallest range and its median was the second highest after the GA. Obviously, the GA performed worst among all four algorithms because of its large dispersion and comparatively high median.

Even though the EDA performed worst in terms of algorithm stability, its worst objective value was only slightly larger than that of the MLMSRBF algorithm, which means the EDA is capable of producing better solutions (compared with other algorithms) consistently.

Empirical Cumulative Distribution Function Plots

To compare the reliability of each algorithm, the empirical cumulative distribution function (Devore 2011) was used, which is calculated using the ordered best objective values. The empirical distribution function is a direct estimate of the true underlying cumulative distribution function, and conveys the frequency of occurrence for objective function values less than a reference value.

The empirical cumulative distribution functions of the four algorithms are shown in Fig. 5. Because the optimization problem is to find the minimum in a limited number of function evaluations, the empirical cumulative distribution function is better when it has a high probability for a low objective value. Focusing on this criterion, the EDA performed best among all algorithms, because it had 90.9% probability to generate objective values less than 0.41. Moreover, it had 63.6% probability to generate objective values less than 0.35, which means it has the ability to generate some good solutions that other algorithms cannot. The range of objective values produced by the MLMSRBF algorithm was less than the other three algorithms and it is capable of producing objective values
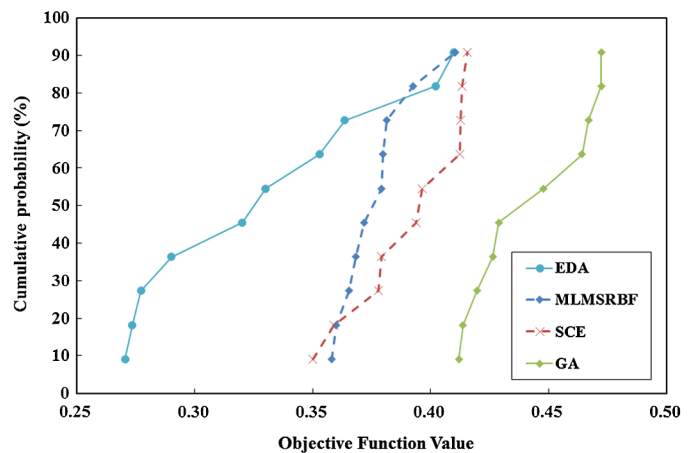
ranging from 0.36 to 0.41 with a probability of 90.9%. The GA was inferior to all other algorithms and the interval of its solutions was fairly broad. High reliability cannot be expected from the GA because it finds poor solutions on some trials.

Rainfall-Runoff Simulation Results

The previous comparisons of efficiency, variability, and reliability focused on the advantages and disadvantages of the algorithms themselves. To verify the calibration effects in hydrologic models, three criteria, the NSE, R and PPTS, were employed. R is defined as follows:

$$R = \frac{|\sum_{t=1}^{N} Q_t^{\text{obs}} - \sum_{t=1}^{N} Q_t^{\text{sim}}|}{\sum_{t=1}^{N} Q_t^{\text{obs}}} \tag{6}$$

R calculates the relative difference between average simulated and average observed runoff, which predicts the water balance error during the entire calibration period. $\text{PPTS}_{(l,u)}$ represents the average absolute error of simulated flows lying in the band of top $u\%$ and $l\%$ data. PPTS being introduced was determined on the basis of the consideration that NSE and R do not evaluate the prediction ability of the model from higher to low region. By arranging the observed data in descending order, Lohani et al. (2014) peak percent threshold statistic of prediction between top $u\%$ and $l\%$ data ($\text{PPTS}_{(l,u)}$) is given by

$$\text{PPTS}_{(l,u)} = \frac{1}{(k_l - k_u + 1)} \sum_{t=k_l}^{k_u} |\xi_t| \tag{7}$$

in which $k_l = l \times N/100$, $k_u = u \times N/100$ and $\xi_t = Q_t^{\text{obs}} - Q_t^{\text{sim}}/Q_t^{\text{sim}}$ where N = number of data; ξ_t = average relative error of the tth data; and l and u are respectively lower and higher limits in percentage, representing the frequency of occurrence of values of flow data higher than a reference value. Therefore, peak percent threshold statistic of top $l\%$ data can be presented as $\text{PPTS}_{(l.0)}$ or simply $\text{PPTS}_{(l)}$. In this study, PPTS values for highest 2, 3, 5, 10, and 20% flows were computed, which are denoted by $\text{PPTS}_{(2)}$, $\text{PPTS}_{(3)}$, $\text{PPTS}_{(5)}$, $\text{PPTS}_{(10)}$, and $\text{PPTS}_{(20)}$.

The NSE, R, and PPTS values were calculated using the parameters that were the best objective value for each algorithm in 10 trials. Table 3 shows their values in the calibration and validation periods at the end of 500 function evaluations. In the calibration period, the EDA performed best among the four algorithms in terms of the NSE with a value of 0.73, but performed worst in terms of R.

Table 3. Best Nash–Sutcliffe Efficiency (NSE), Overall Volume Error (R) and Peak Percent Threshold Statistic (PPTS) Values after 500 Model Evaluations

Simulation period	Criteria	EDA	MLMSRBF	SCE	GA
Calibration period	NSE	0.73	0.64	0.61	0.56
	R	0.28	0.25	0.23	0.26
	$PPTS_{(2)}$	0.09	0.31	0.42	0.49
	$PPTS_{(3)}$	0.08	0.29	0.40	0.45
	$PPTS_{(5)}$	0.06	0.27	0.35	0.40
	$PPTS_{(10)}$	0.12	0.23	0.26	0.29
	$PPTS_{(20)}$	0.22	0.14	0.16	0.17
Validation period	NSE	0.58	0.56	0.56	0.54
	R	0.43	0.28	0.27	0.27
	$PPTS_{(2)}$	0.15	0.20	0.31	0.32
	$PPTS_{(3)}$	0.21	0.22	0.26	0.26
	$PPTS_{(5)}$	0.24	0.18	0.21	0.20
	$PPTS_{(10)}$	0.29	0.16	0.15	0.14
	$PPTS_{(20)}$	0.33	0.20	0.14	0.13

Table 4. Optimized Parameters Using the Estimation of Distribution Algorithm (EDA)

Parameter	Value
ESCO	0.25
CN2	53.87
SOL_AWC	0.02
GW_DEALY	0.39
ALPHA_BF	0.80
GWQMN	20.55
GW_REVAP	0.05
REVAPMN	430.33
CH_K2	492.86
SURLAG	0.05

Note: See Table 1 for definitions of the parameters.
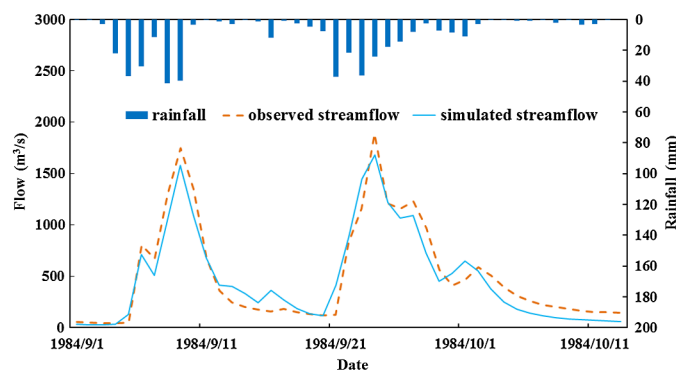
Fig. 6. Observed and simulated hydrographs obtained by the EDA

Fig. 7. Scatterplot of the logarithm observed ($\ln Q^{obs}$) and simulated ($\ln Q^{sim}$) daily streamflow during the entire calibration period for the EDA

This may have resulted from the single objective function in the calibration period. Conversely, lower values of $PPTS_{(2)}$, $PPTS_{(3)}$, $PPTS_{(5)}$, and $PPTS_{(10)}$, and higher values of $PPTS_{(20)}$ for EDA indicated the capability of simulating higher flow values in the calibration period. Legates and McCabe (1999) pointed out that because of the squared differences in calculation formula, NSE is overly sensitive to extreme values, which means the results of PPTS statistics agree with the performance of NSE and R. Similar performance criteria appeared in the validation period. A set of parameters optimized by the EDA is shown in Table 4.

The observed and simulated hydrographs are shown in Fig. 6. The simulated hydrographs were calculated using the best set of parameters calibrated by the EDA at the end of 500 evaluations given in Table 4. Because of the large number of time steps, a short period of about one month was selected to clearly show the rainfall-runoff process simulation performance. The comparison showed that the simulated streamflow curve matched the observed streamflow curve.

The scatterplot of the simulated and observed series is shown in Fig. 7. The plot indicates that the parameters calibrated by the EDA at the end of 500 function evaluations could underestimate the streamflow a little bit.

Conclusions

In this study, the EDA was used to calibrate parameters of the SWAT model, in which a Gaussian network was employed as the probabilistic model. Results from the EDA were compared with those obtained from three other automatic optimization methods: (1) the MLMSRBF algorithm, (2) SCE algorithm, and (3) GA. The results indicated that the EDA performed best, in some respects, in a limited computational budget of only 500 function evaluations. The EDA had higher efficiency compared with the other algorithms. Even though a larger standard deviation, interquartile range, and range of the empirical cumulative distribution function value indicated lower stability, they revealed the ability of the EDA to produce better solutions in the same probability compared with other algorithms. The optimization results indicated that EDA could provide solutions with better performance of NSE and PPTS statistics for higher magnitude flows, while performance of R and PPTS statistics for low magnitude flows were relatively poor. This may have resulted from the single objective function.

The current research focused on applying the EDA to an optimization problem for a computationally intensive distributed hydrologic model. In the future, this approach could be extended to more optimization problems in hydrology and water resources, and to explicitly inspect the probabilistic relationships of problem variables.

Acknowledgments

References

Armananzas, R., et al. (2008). "A review of estimation of distribution algorithms in bioinformatics." *BioData Min.*, 1(6), 1–12.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Williams, J. R. (1998). "Large area hydrologic modeling and assessment. Part I: Model development1." *J. Am. Water Resour. Assoc.*, 34(1), 73–89.

Booker, A. J., Dennis, J. E., Frank, P. D., Serafini, D. B., Torczon, V., and Trosset, M. W. (1999). "A rigorous framework for optimization of expensive functions by surrogates." *Struct. Optim.*, 17(1), 1–13.

Cooper, G., and Herskovits, E. (1992). "A Bayesian method for the induction of probabilistic networks from data." *Mach. Learn.*, 9(4), 309–347.

Deng, C., Liu, P., Guo, S., Wang, H., and Wang, D. (2015a). "Estimation of nonfluctuating reservoir inflow from water level observations using methods based on flow continuity." *J. Hydrol.*, 529(Part 3), 1198–1210.

Deng, C., Liu, P., Liu, Y., Wu, Z., and Wang, D. (2015b). "Integrated hydrologic and reservoir routing model for real-time water level forecasts." *J. Hydrol. Eng.*, 10.1061/(ASCE)HE.1943-5584.0001138, 05014032.

Devore, J. (2011). *Probability and statistics for engineering and the sciences*, Duxbury Press, Belmont, CA.

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

Duan, Q., Sorooshian, S., and Gupta, V. K. (1994). "Optimal use of the SCE-UA global optimization method for calibrating watershed models." *J. Hydrol.*, 158(3–4), 265–284.

Espinet, A. J., and Shoemaker, C. A. (2013). "Comparison of optimization algorithms for parameter estimation of multi-phase flow models with application to geological carbon sequestration." *Adv. Water Resour.*, 54, 133–148.

Fu, M. C., Hu, J., and Marcus, S. I. (2006). "Model-based randomized methods for global optimization." *Proc., 17th Int. Symp. on Mathematical Theory of Networks and Systems*, Kyoto Univ., Kyoto, Japan, 355–363.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley Professional, Boston.

Gutmann, H. M. (2001). "A radial basis function method for global optimization." *J. Global Optim.*, 19(3), 201–227.

Henrion, M. (1986). "Propagating uncertainty in bayesian networks by probabilistic logic sampling." *Proc., 2nd Annual Conf. on Uncertainty in Artificial Intelligence*, Elsevier, Philadelphia, 149–164.

Ishikawa, T., Tsukui, Y., and Matsunami, M. (1999). "A combined method for the global optimization using radial basis function and deterministic approach." *IEEE Trans. Magn.*, 35(3), 1730–1733.

Jarboui, B., Eddaly, M., and Siarry, P. (2009). "An estimation of distribution algorithm for minimizing the total flowtime in permutation flowshop scheduling problems." *Comput. Oper. Res.*, 36(9), 2638–2646.

Kollat, J. B., Reed, P. M., and Kasprzyk, J. R. (2008). "A new epsilon-dominance hierarchical Bayesian optimization algorithm for large multiobjective monitoring network design problems." *Adv. Water Resour.*, 31(5), 828–845.

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

Larranaga, P., and Lozano, J. A. (2001). *Estimation of distribution algorithms: A new tool for evolutionary computation*, Kluwer Academic Publishers, Norwell, MA.

Legates, D. R., and McCabe, G. J. (1999). "Evaluating the use of" goodness-of-fit" Measures in hydrologic and hydroclimatic model validation." *Water Resour. Res.*, 35(1), 233–241.

Li, Y., He, H., Wang, Y., Xu, X., and Jiao, L. (2015). "An improved multi-objective estimation of distribution algorithm for environmental economic dispatch of hydrothermal power systems." *Appl. Soft Comput.*, 28, 559–568.

Lima, C. F., Pelikan, M., Goldberg, D. E., Lobo, F. G., Sastry, K., and Hauschild, M. (2007). *Influence of selection and replacement strategies on linkage learning in BOA*, Vols. 1–10, IEEE, New York.

Liu, P., Cai, X., and Guo, S. (2011). "Deriving multiple near-optimal solutions to deterministic reservoir operation problems." *Water Resour. Res.*, 47(8), W08506.

Lohani, A. K., Goel, N. K., and Bhatia, K. K. S. (2014). "Improving real time flood forecasting using fuzzy inference system." *J. Hydrol.*, 509, 25–41.

Montanari, A., et al. (2013). "'Panta Rhei—Everything flows': Change in hydrology and society—The IAHS Scientific Decade 2013–2022." *Hydrol. Sci. J.*, 58(6), 1256–1275.

Mugunthan, P., Shoemaker, C. A., and Regis, R. G. (2005). "Comparison of function approximation, heuristic, and derivative-based methods for automatic calibration of computationally expensive groundwater bioremediation models." *Water Resour. Res.*, 41(11), W11427.

Mühlenbein, H., Bendisch, J., and Voigt, H. M. (1996). "From recombination of genes to the estimation of distributions: II. Continuous parameters." *Parallel problem solving from nature—PPSN IV. Lecture notes in computer science*, H. M. Voigt, W. Ebeling, I. Rechenberg, and H. P. Schwefel, eds., Springer, Berlin, 188–197.

Nash, J. E., and Sutcliffe, J. V. (1970). "River flow forecasting through conceptual models: Part I—A discussion of principles." *J. Hydrol.*, 10(3), 282–290.

Paul, T. K., and Iba, H. (2002). "Linear and combinatorial optimizations by estimation of distribution algorithms." *9th MPS Symp. on Evolutionary Computation*, IPSJ, Japan.

Razavi, S., Tolson, B. A., and Burn, D. H. (2012). "Review of surrogate modeling in water resources." *Water Resour. Res.*, 48(7), W07401.

Regis, R. G., and Shoemaker, C. A. (2007). "A stochastic radial basis function method for the global optimization of expensive functions." *INFORMS J. Comput*, 19(4), 497–509.

Santarelli, S., et al. (2006). "Military antenna design using simple and competent genetic algorithms." *Math. Comput. Model.*, 43(9–10), 990–1022.

Schwarz, G. (1978). "Estimating the dimension of a model." *Ann. Statist.*, 6(2), 461–464.

Simpson, T. W., Mauery, T. M., Korte, J. J., and Mistree, F. (2001). "Kriging models for global approximation in simulation-based multidisciplinary design optimization." *AIAA J.*, 39(12), 2233–2241.

Sorooshian, S., Duan, Q., and Gupta, V. K. (1993). "Calibration of rainfall-runoff models: Application of global optimization to the Sacramento soil moisture accounting model." *Water Resour. Res.*, 29(4), 1185–1194.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, A., and Srinivasan, R. (2006). "A global sensitivity analysis tool for the parameters of multi-variable catchment models." *J. Hydrol.*, 324(1–4), 10–23.

Vose, M. D. (1999). *The simple genetic algorithm: Foundations and theory*, MIT Press, Cambridge, MA.

Wang, Q. J. (1991). "The genetic algorithm and its application to calibrating conceptual rainfall-runoff models." *Water Resour. Res.*, 27(9), 2467–2471.

Yapo, P. O., Gupta, H. V., and Sorooshian, S. (1996). "Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data." *J. Hydrol.*, 181(1–4), 23–48.