scientific reports

OPEN

Uncertainty analysis of model inputs in riverine water temperature simulations

Babak Abdi[1], Omid Bozorg-Haddad[1✉] & Xuefeng Chu[2]

Simulation models are often affected by uncertainties that impress the modeling results. One of the important types of uncertainties is associated with the model input data. The main objective of this study is to investigate the uncertainties of inputs of the Heat-Flux (HFLUX) model. To do so, the Shuffled Complex Evolution Metropolis Uncertainty Algorithm (SCEM-UA), a Monte Carlo Markov Chain (MCMC) based method, is employed for the first time to assess the uncertainties of model inputs in riverine water temperature simulations. The performance of the SCEM-UA algorithm is further evaluated. In the application, the histograms of the selected inputs of the HFLUX model including the stream width, stream depth, percentage of shade, and streamflow were created and their uncertainties were analyzed. Comparison of the observed data and the simulations demonstrated the capability of the SCEM-UA algorithm in the assessment of the uncertainties associated with the model input data (the maximum relative error was 15%).

Mathematical modeling is subject to uncertainties from multiple sources, such as measurement errors and inputs, and model structure[19]. Each of these sources of uncertainties contributes to the generation of uncertain outputs and results from the modeling. Uncertainty analysis for the input data and model parameters can be performed by specifying the ranges of their variations and identifying the most probable values from these ranges. For example, in hydrologic modeling, uncertainties associated with the amount and timing of streamflow can be presented by incorporating the full range of simulation results of a hydrologic model[7]. Sun et al.[14] evaluated the uncertainty of the Storm Water Management Model (SWMM) by using a generalized likelihood uncertainty estimation (GLUE). The GLUE based uncertainty analysis generated a probability distribution of possible outcomes from SWMM, which was compared with the Gaussian distribution developed by using the observed streamflow data. Feyen et al.[5] used the GLUE algorithm to analyze the uncertainty of the hydraulic conductivity parameter in their MODFLOW modeling for a groundwater capture zone. They found that the delineated capture zones were most sensitive to the mean hydraulic conductivity. Zheng et al.[22] analyzed the parameter uncertainty in an irrigated and rainfed agroecosystem model. They used the DayCent agroecosystem model and examined the role of parameter uncertainties in characterizing production functions. Their study indicated that incorporating rigorous estimates of uncertainty significantly enhanced the use of water production functions for effective water management. Yan et al.[16] evaluated the impact of parameter uncertainty and water stress parameterization on wheat growth simulations by using the GLUE algorithm in the CERES-Wheat modeling. They found that GLUE-calibrated parameters were significantly different from the observations and concluded that this disagreement resulted mainly from the unrealistic water stress parameterization, which strongly affected the GLUE algorithm in selection of the calibrated parameters.

The Monte Carlo simulation has been used for uncertainty analysis as it provides multiple benefits over the conventional uncertainty analysis methods[13]. The Markov chain Monte Carlo (MCMC) combines the Monte Carlo approach with the Markov Chain to estimate the uncertainty bound of a model output. MCMC-based algorithms update the characteristics of the proposed probability distribution by maintaining the ergodicity[21]. Although the uncertainties associated with model parameters have been examined in numerous studies, the MCMC-based algorithms are able to account for input errors and/or model structural errors. Also, such algorithms use time-dependent parameters for exploring model deficits and considering the uncertainties of inputs and model structure[17]. Among the MCMC algorithms, the Shuffled Complex Evolution Metropolis algorithm (SCEM-UA) has been used in various hydrologic and environmental studies[20]. The SCEM-UA algorithm

[1]Department of Irrigation and Reclamation Engineering, Faculty of Agricultural Engineering and Technology, College of Agriculture and Natural Resources, University of Tehran, Karaj, Tehran, Iran. [2]Department of Civil and Environmental Engineering, North Dakota State University, Dept 2470, Fargo, ND 58108-6050, USA. ✉email: obhaddad@ut.ac.ir

combines the strengths of the Metropolis algorithm, controlled random search, competitive evolution, and complex shuffling to infer the posterior target distribution by a continuous updating process[20]. Due to these advantages, the SCEM-UA algorithm was selected for the current study. Lin et al.[9] used the GLUE algorithm and the SCEM-UA to address the issue of parameter uncertainty for conceptual hydrological modeling. Their results showed that when setting the threshold value at the interior sites, the simulated runoff series by the Xinanjiang model with the behavioral parameter sets fitted better with the observed runoff series. Liu et al.[10] used the SCEM-UA to calibrate a hydrological model. They represented the posterior distribution functions of the hydrologic model output by the SCEM-UA. The calibration of the model, using 16 different parameters, was performed using two procedures of SCEM-UA and the genetic algorithm (GA). Results indicated that both methods were performing very well, but the SCEM-UA was better. Ajami et al.[3] proposed the Integrated Bayesian Uncertainty Estimator (IBUNE) to analyze the uncertainty of model parameters. They used the SCEM algorithm to analyze the parameter uncertainty of a rainfall-runoff model. By utilizing the Bayesian model average (BMA) and the developed SCEM algorithm, they examined the total uncertainty of the model. Tang et al.[15] used the SCEM-UA algorithm to analyze the uncertainties for nonlinear structural systems and demonstrated that the algorithm effectively estimated the parameters with uncertainties. Liu et al.[11] evaluated the uncertainty of urban flood modeling. Due to the lack of reliable discharge data, they combined experimental data and modelling to characterize the floods and map the inundated areas in Xiamen Island, China.

Quantification of the uncertainties for river temperature models based on heat fluxes is considered in evaluating the effectiveness of ecological restoration alternatives. In ecological restoration studies such as aquatic green infrastructure[4] and fish habitat[2], small changes in driving parameters could cause faulty understanding from simulated water temperature, a fundamental factor affecting water quality and ecosystems (e.g.,[12,18]).

Thus, it is important to select the most efficient method to assess the uncertainties associated with the input data and model parameters, as most methods require longer running time and are computationally intensive for specific models. Neglecting uncertainty analysis can lead to a series of problems such as overdesign, higher costs, reduced reliability, and failure to achieve optimal benefits. Examining the uncertainties of model inputs provides a comprehensive insight into their influences on the model outputs, which potentially improves the modeling efficiency and lowers the cost in their measurements.

The objective of this study is to investigate the uncertainties of the inputs of the HFLUX model using the SCEM-UA algorithm. As a new effort, the SCEM-UA algorithm is used for assessing the HFLUX model inputs, and its performance is evaluated in this study. The HFLUX[8] is an efficient and useful river water quality model for 1D simulation of the spatio-temporal distribution of stream water temperatures. In the simulation of temperature at each discretized node and each time step, the HFLUX considers the heat fluxes from the environment and lateral inflows of water to the node. It is also flexible in choosing the solution methods. Thus, the HFLUX model is selected for simulating water temperatures in this study. The simulation results are compared with the observed data to evaluate the performance of the SCEM-UA algorithm in the analysis of the uncertainties of the model inputs.

Materials and methods

In this study, the HFLUX model was coupled with the SCEM-UA algorithm for analyzing the uncertainties of the model inputs. The specific procedures started with selecting the inputs of the HFLUX model. With the linked HFLUX and SCEM-UA model and implementation of an iteration scheme, the uncertainty of each of the selected inputs was obtained based on the ranges (minimum and maximum values) of the input data/parameters and the Latin hypercube sampling. The simulations were then compared against the observed data to evaluate the performance of the SCEM-UA algorithm. These steps are depicted in Fig. 1.

River water temperatures simulated by the HFLUX model. River water temperature affects the water quality and the ecosystem health, and hence control of river water temperature is important to mitigation of its adverse effects[1]. The HFLUX model was used to simulate the streamflow temperatures at different locations and times. The model is highly flexible in terms of choosing the solution methods for solving the governing equations and selecting the energy budget terms such as shortwave solar radiation, latent heat flux, and sensible heat transfer flux. The model input data include the initial spatial and temporal temperature conditions, stream geometry data, discharge data, and meteorological data[8]. The water balance and energy balance equations are respectively given by[8]:

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = q_L \tag{1}$$

$$\frac{\partial (A\,T_w)}{\partial t} + \frac{\partial (Q\,T_w)}{\partial x} = q_L\,T_L + R \tag{2}$$

$$R = \frac{B\,\varphi_{total}}{\rho_w\,C_w} \tag{3}$$

where A is the cross section area of the stream (m²), x is the distance along the stream (m), t is the time (s), Q is the discharge of the stream (m³/s), q_L is the lateral inflow per unit stream length (m²/s), T_w is the stream temperature (°C), T_L is the temperature of the lateral inflow (°C), R is the energy flux (source or sink) per unit stream length (°C m²/s), B is the width of the stream (m), φ_{total} is the total energy flux to the stream per surface area (W/m²), ρ_w is the density of water (kg/m³), and C_w is the specific heat of water (J/kg °C). Equation (3) is
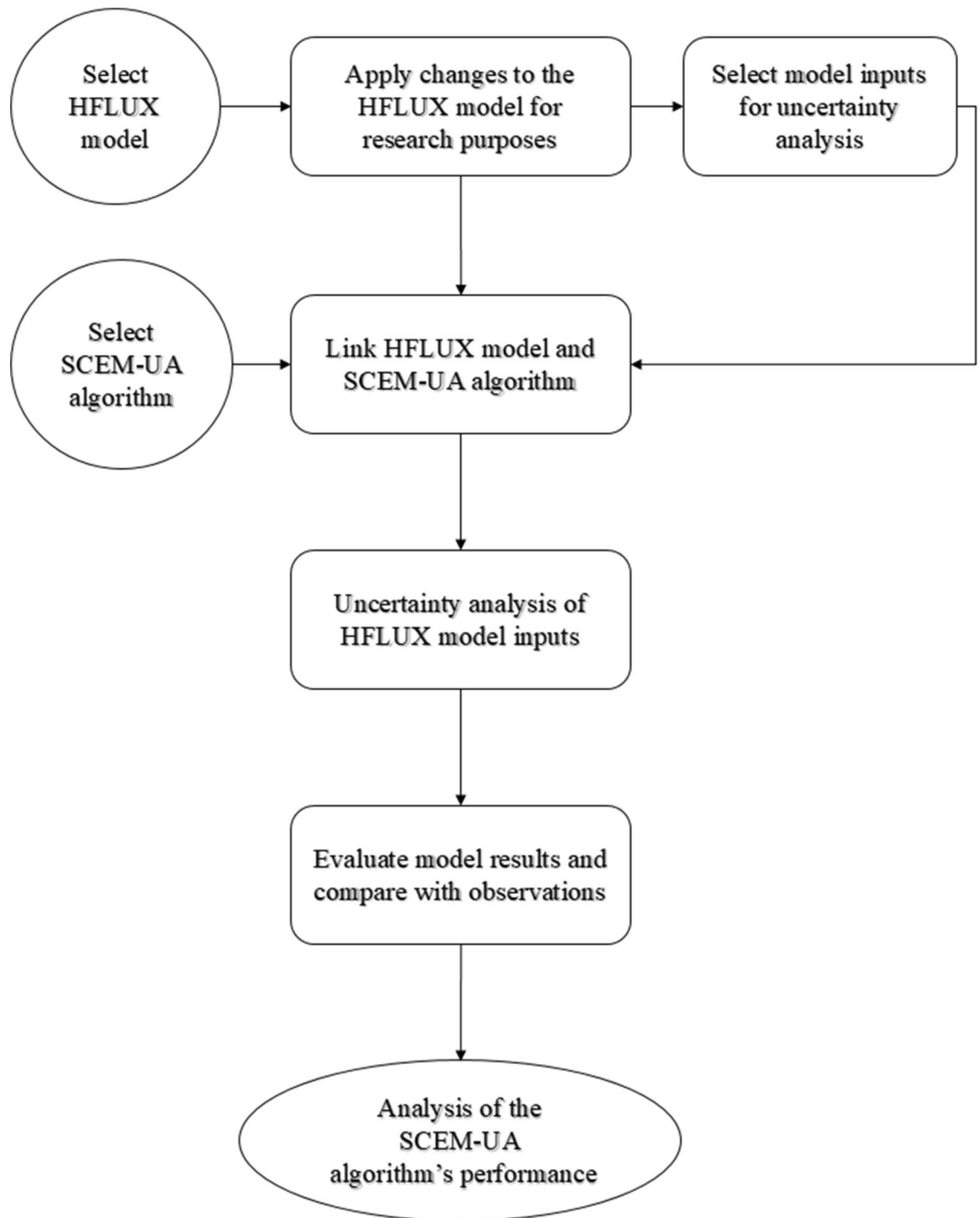
Figure 1. Flowchart for the uncertainty analysis.

based on a thermal datum of $0\,^{\circ}C$ and the impact on the absolute value of the advective heat flux term. In Eq. (2), if q_L is negative, the first term on the right-hand side of the equation becomes a loss of $q_L T_w$. Also, dispersive heat transport that is omitted in Eq. 2 is negligible when the longitudinal change in water temperature is small in comparison to the temporal changes[8].

SCEM-UA algorithm. The SCEM-UA algorithm provides posterior distribution functions for the model parameters and input data by generating an initial sample from the parameter space. First, the indicators of n, q, and s that are respectively dimension (the number of investigate inputs), number of complexes (the population to be divided), and population (the number of sample points) are determined for the algorithm. Then, the algorithm searches the sampling points in the feasible space and sorts the points according to the density. The algorithm determines the sequence and complexes based on those points. The sequence is the first q points of

the population and complexes are a collection of m points from the population. Note that $m = s/q$. In the next step, the points of each complex are sorted based on the density, which can be mathematically expressed as[20]:

$$\begin{cases} \alpha^k \leq T & \theta^{t+1} = N(\theta^t, C_n^2 \Sigma^k) \\ \alpha^k > T & \theta^{t+1} = N(\mu^k, C_n^2 \Sigma^k) \end{cases} \tag{4}$$

where $k = 1,2,\ldots,q$, α is the ratio of the mean posterior density of the m points of complexes to the mean posterior density of the last m generated points of sequences, θ is the points of complexes, $c_n = \frac{2.4}{\sqrt{n}}$, $T = 10^6$, μ is the mean, and Σ denotes the covariance. To investigate the new points created by the algorithm, the points of complexes are replaced by[20]:

$$\begin{cases} \Omega \geq Z & replace\ best\ member\ of\ C^k\ with\ \theta^{t+1} \\ \Omega < Z & \theta^{t+1} = \theta^t \end{cases} \tag{5}$$

where C^k is the Kth complex, Z is drawn from the uniform distribution in the range of 0–1, and Ω is calculated by[20]:

$$\Omega = \frac{P(\theta^{t+1}|y)}{P(\theta^t|y)} \tag{6}$$

where $P(\theta^{t+1}|y)$ and $P(\theta^t|y)$ are the posterior probability distributions for θ^{t+1} and θ^t, respectively. Then, the algorithm examines the following condition for each complex. If it is rejected, the algorithm replaces the worst member c^k (the point with the lowest density) with θ^{t+1} [20].

$$\Gamma^k \leq T\ and\ P(\theta^{t+1}/y) < P(C_m^k/y) \tag{7}$$

where Γ^k is the ratio of the posterior density of the best (the point with the highest density) to the posterior density of the worst member of c^k. The last step is to examine β and L. Note that $\beta = 1$ and $L = m/10$. If $\beta < L$, $\beta = \beta + 1$ and the algorithm returns to sort complex points. Otherwise, the algorithm examines the Gelman and Rubin convergence[6], and eventually provides the posterior distribution functions[20]. The value of the Gelman and Rubin convergence should be less than 1.2. The Gelman and Rubin convergence is examined by:

$$R = \sqrt{\frac{g-1}{g} + \frac{q+1}{q \cdot g} \frac{B}{W}} \tag{8}$$

where g is the number of iterations within each sequence, B is the variance between the q sequence means, and W is the average of the q within-sequence variances for the parameter under consideration[20].

Study AREA. Meadowbrook Creek was selected to test the methods proposed in this study[8]. The creek flows through the City of Syracuse in New York. Thus, this catchment consists of high residential and industrial land covers, which contribute runoff to the main channel. The creek is about 4 km long. A portion of this creek (475 m long) was selected for the modeling for a period of June 13–19, 2012 in this study. The upstream boundary condition in the HFLUX model was set based on the water temperature of the creek observed at the upstream station[8]. The uncertainty of the model inputs was examined at three selected points as shown in Fig. 2. Note that the input values at these three points had greater relative changes than the changes at other locations, which provided the possibility to improve the evaluation of the algorithm performance. In addition, these three locations had the same sampling of the selected input data. During the simulation period, the streamflow velocity varied within a range of 0.06–0.63 (m/s). The daily temperature changed between 8.9 and 28.2 °C. The relative humidity, used to calculate the total energy flux to the stream per surface area, changed from 36 to 93%. The creek bed mainly consisted of clay, cobbles, sand, and gravel materials. The basic statistics of the data/variables used in the HFLUX model are presented in Table 1. Figure 2 shows the study area, the creek, and the three selected points for analysis.

Ethical approval. All authors accept all ethical approvals.

Consent to participate. All authors consent to participate.

Consent to publish. All authors consent to publish.

Results and discussion

In the uncertainty analysis, the inputs at the three selected points on the main channel include: the depth and width of the creek, the percentage of shade, and the streamflow. Note that the shade value at a location ranges from 0 to 1 with 0 being no shade and 1 being total shade. The values of these inputs estimated by the SCEM-UA algorithm at different locations along the creek are shown in Fig. 3.

To implement the SCEM-UA, the posterior distribution functions for the selected inputs of the model were first developed. The initial ranges of the four selected inputs, according to the related literature and field observations, were selected to form the first generation of the SCEM-UA population using the Latin hypercube sampling. Table 2 presents the values of the maximum and minimum values of these inputs.
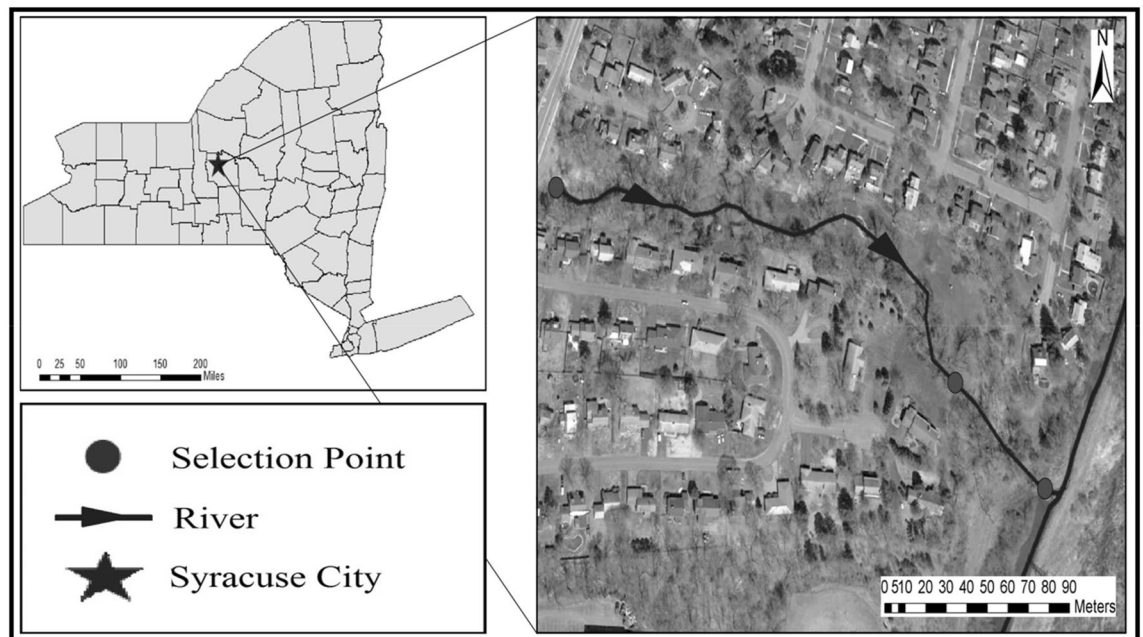
Figure 2. Study area and the locations of three evaluation sections (the gray enlarged map shows the State of New York), the map in this Figure is created by Google Earth 7.0.2.8415 (https://google.com/earth/versions).

Variable	Mean	Std	Min	25%	50%	75%	Max
Water temperature at the upstream boundary (°C)	16.98	1.61	13.68	15.87	16.88	18.19	20.26
Discharge (m³/s)	0.0670	0.0039	0.0603	0.0634	0.0672	0.0705	0.0734
Air temperature (°C)	20.08	5.10	8.90	16.40	20.50	24.10	28.20
Relative humidity (%)	64.89	16.30	36.00	50.00	65.00	77.00	93.00
Shortwave solar radiation (W/m²)	251.23	364.63	0.00	0.00	24.00	469.36	1208.00

Table 1. Basic statistics of the data/variables used in the HFLUX model.

The SCEM-UA simulations were performed for a spatial interval of 10 m and a time step of 30 min. The number of the SCEM-UA iterations considered for this study was 30,000, depending upon the Gelman and Rubin convergence criterion[6], a statistical indicator used for examining the convergence of the chains. The value should be less than 1.2[20]. Figure 4 shows the changes in the Gelman and Rubin convergence, indicating that the values for all inputs are less than 1.2.

The posterior distribution functions for the selected points of the HFLUX model were obtained based on the SCEM-UA algorithm in the form of histograms for the inputs. The width or range of the histograms indicates the uncertainty of the inputs and the average value or the most likely value of the histograms is the most likely prediction value of the SCEM-UA algorithm for that input. Figures 5, 6, 7 and 8 show the histograms of the posterior distribution functions of the selected inputs. As shown in Fig. 5, the uncertainty for the creek width at the first point/location ranges from 5 to 6 and the most likely value predicted by the SCEM-UA is 5.6 with a probability of 50%. The range of the uncertainty for this input at the second point/location is from 3.6 to 4.05 and the most likely value is 3.7 with a probability of approximately 60%. For this input at the third point/location, the uncertainty ranges from 2.5 to 5.2 and the most likely value is 3.1 with a probability of 50%. According to Fig. 6, the uncertainty ranges for the creek depth at the first, second, and third points/locations are 0.08–0.125, 0.07–0.25, and 0.26–0.44, respectively. The most likely values for this input at the first, second, and third points/locations are 0.085 with a probability of nearly 70%, 0.13 with a probability of almost 70%, and 0.36 with a probability of almost 55%, respectively. Figure 7 shows the uncertainty ranges and the most likely values for the percentage of shade at the first, second, and third points/locations. The ranges for the tree points are respectively 0.2–0.47, 0.17–0.44, and 0.15–0.42, while the most likely values are respectively 0.26 with a probability of almost 60%, 0.23 with a probability of almost 40%, and 0.18 with a probability of almost 50%. Similarly, Fig. 8 shows the uncertainty ranges and the most likely values for the streamflow at the first, second, and third points/locations. The ranges for the three points respectively are 0.06018–0.06027, 0.0716–0.0725, and 0.073–0.0739, while the most likely values respectively are 0.06023 with a probability of almost 30%, 0.0719 with a probability of almost 60%, and 0.0735 with a probability of almost 55%.

Based on the statistical results related to the inputs, their sensitivity can be identified. The smaller the coefficient of variation (CV) of an input, the more sensitive the input. Figure 9 indicates the order of sensitivity of
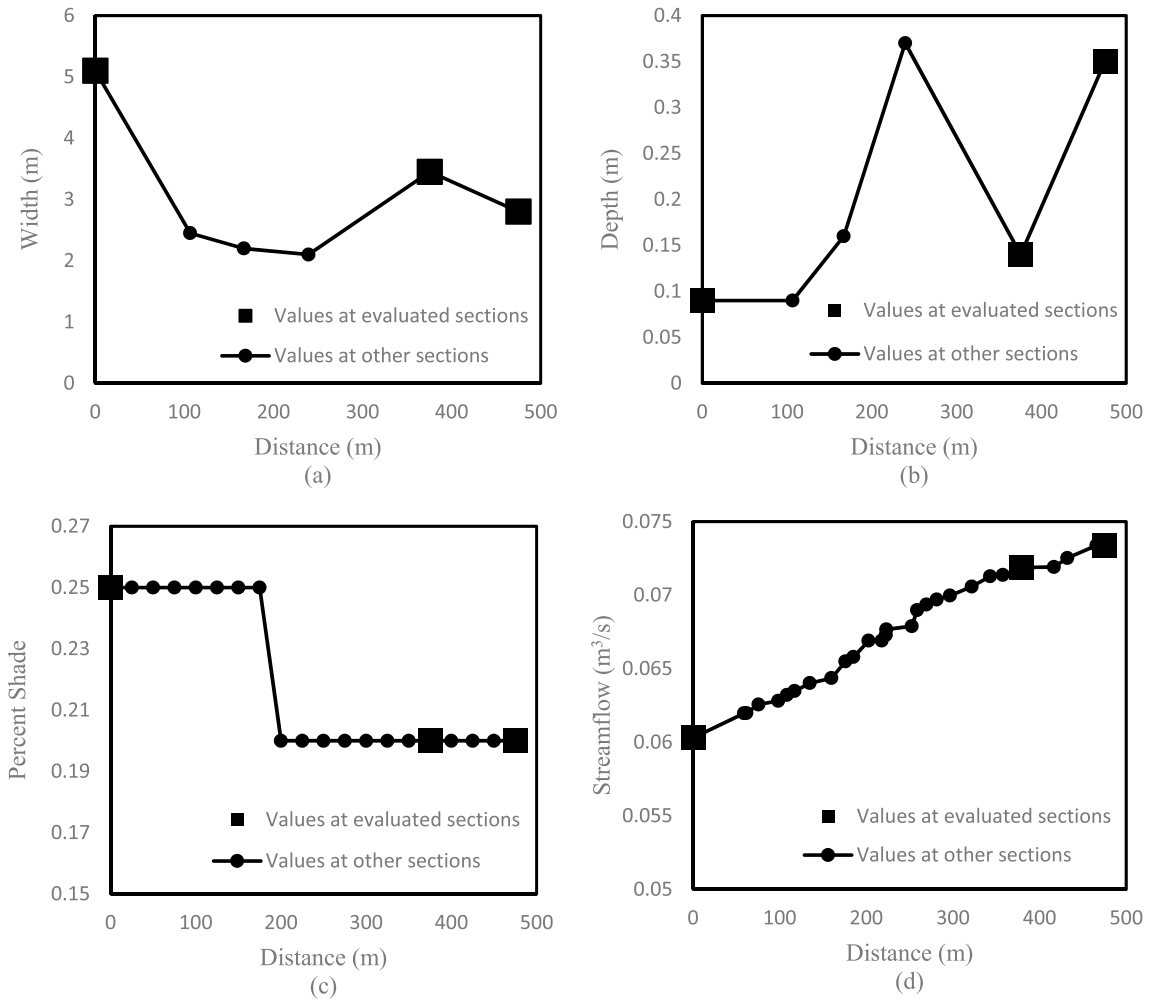
Figure 3. Characteristics of Meadowbrook Creek: (**a**) width, (**b**) depth, (**c**) percent shade, and (**d**) streamflow.

Input	Unit	Minimum	Maximum	Description
W_1	m	1	7	Creek width at the beginning of the study reach (0 m)
W_2	m	1	7	Creek width at 375 m from the beginning of the study reach
W_3	m	1	7	Creek width at the end of the study reach (475 m)
D_1	m	0.001	1	Creek depth at the beginning of the study reach (0 m)
D_2	m	0.001	1	Creek depth at 375 m from the beginning of the study reach
D_3	m	0.001	1	Creek depth at the end of the study reach (475 m)
Sh_1	-	0	1	Percent shade coefficient at the beginning of the study reach (0 m)
Sh_2	-	0	1	Percent shade coefficient at 375 m from the beginning of the study reach
Sh_3	-	0	1	Percent shade coefficient at the end of the study reach (475 m)
Q_1	m³/s	0.05	0.1	Streamflow at the beginning of the study reach (0 m)
Q_2	m³/s	0.05	0.1	Streamflow at 375 m from the beginning of the study reach
Q_3	m³/s	0.05	0.1	Streamflow at the end of the study reach (475 m)

Table 2. The selected HFLUX inputs and their upper and lower limits considered in the uncertainty analysis.

the inputs. Accordingly, inputs Q_2, Q_3, and Q_1 with CV values of 1.94%, 2.57%, and 3.18%, respectively, have the higher sensitivity, implying very small changes in the histograms of these three points. Inputs D_1, Sh_3, and Sh_2 with CV values of 78.14%, 65.08%, and 60.48%, respectively, have the lower sensitivity. Table 3 shows the CV
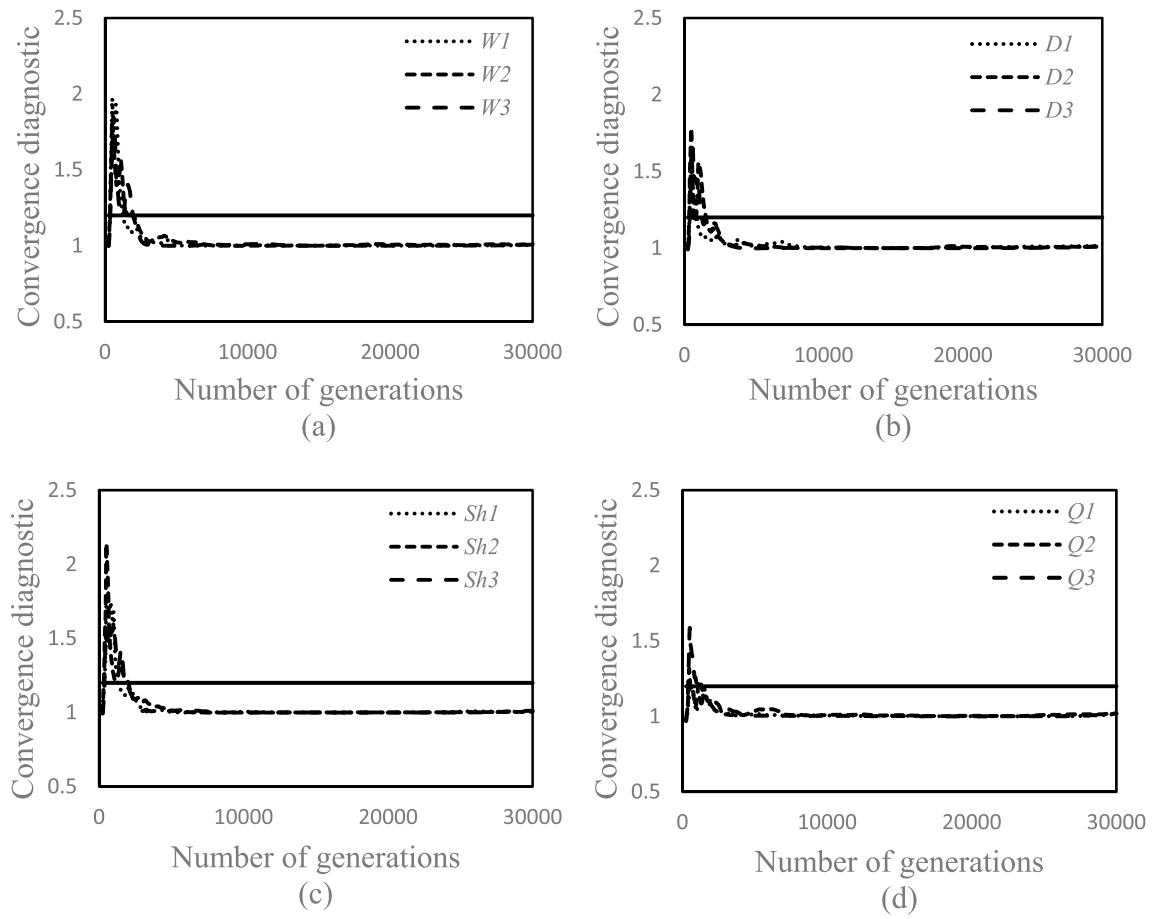
Figure 4. Gelman and Rubin convergence criterion chart: (**a**) width, (**b**) depth, (**c**) percent shade, and (**d**) streamflow.

and the most likely values of all inputs. Knowing the most likely value of each input is important in evaluating the performance of the uncertainty analysis algorithm. Figure 10 shows the comparison of the observed data and the most likely values estimated by the SCEM-UA algorithm. According to Fig. 10 the SCEM-UA algorithm overestimated the inputs of W_1, W_2, W_3, D_3, Sh_1, Sh_2, Q_2, and Q_3 and underestimated the inputs of D_1, D_2, Sh_3, and Q_1. The most likely values of the selected inputs simulated by the SCEM-UA algorithm were compared with the observed data and their relative errors are shown in Table 4. It can be observed that the maximum relative error (15%) is related to the percentage of shade for the second point/location. Thus, it is suitable to use the SCEM-UA algorithm for evaluating the uncertainties of the HFLUX inputs.

Concluding remarks

This study focused on investigating the uncertainties of the HFLUX model inputs by using the SCEM-UA algorithm. Meadowbrook Creek in the City of Syracuse, New York was selected as an application of the proposed methods. The histograms of the selected model inputs were obtained, based on which the uncertainty of the inputs and their most likely values were determined. Specifically, the width of each histogram indicated the uncertainty of the corresponding input. It was found that the creek depth at the beginning of the study reach with a CV of 78.14% was the most uncertain and thus the least sensitive input. The streamflow at 375 m from the beginning of the study reach with a CV of 1.94% was the least uncertain and thus the most sensitive input. The mean of each histogram indicated the most likely value for the corresponding input. The performance of the SCEM-UA algorithm was evaluated by comparing the observed data and the most likely values from the SCEM-UA algorithm. Based on the comparisons, the streamflow at 375 m from the beginning of the study reach with the smallest relative error (0.03%) was the most accurately estimated input, while the percent shade coefficient at 375 m from the beginning of the study reach with the largest relative error (15%) was the least accurately estimated input. These results demonstrated that the SCEM-UA algorithm was suitable for analyzing the uncertainties associated with the inputs of the HFLUX model.
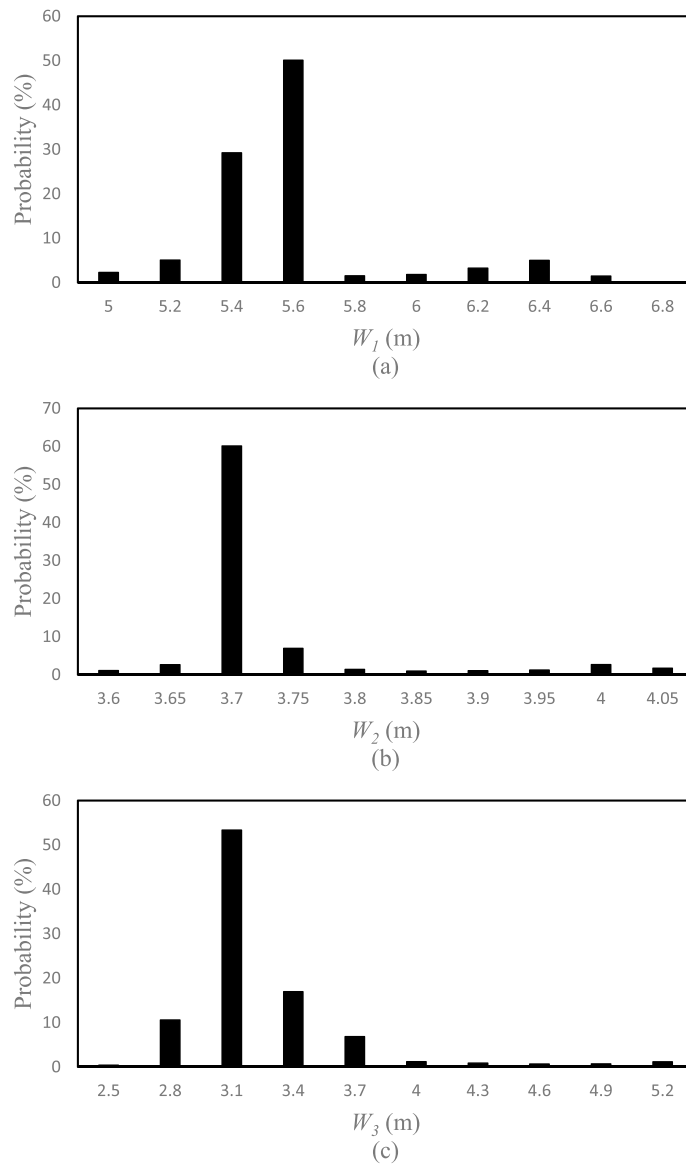
Figure 5. Histograms of the HFLUX inputs at three sections (**a**) Creek width at the beginning of the reach study range (0 m), (**b**) Creek width at 375 m from the beginning of the reach study range, and (**c**) Creek width at the end of the reach study range (475 m).
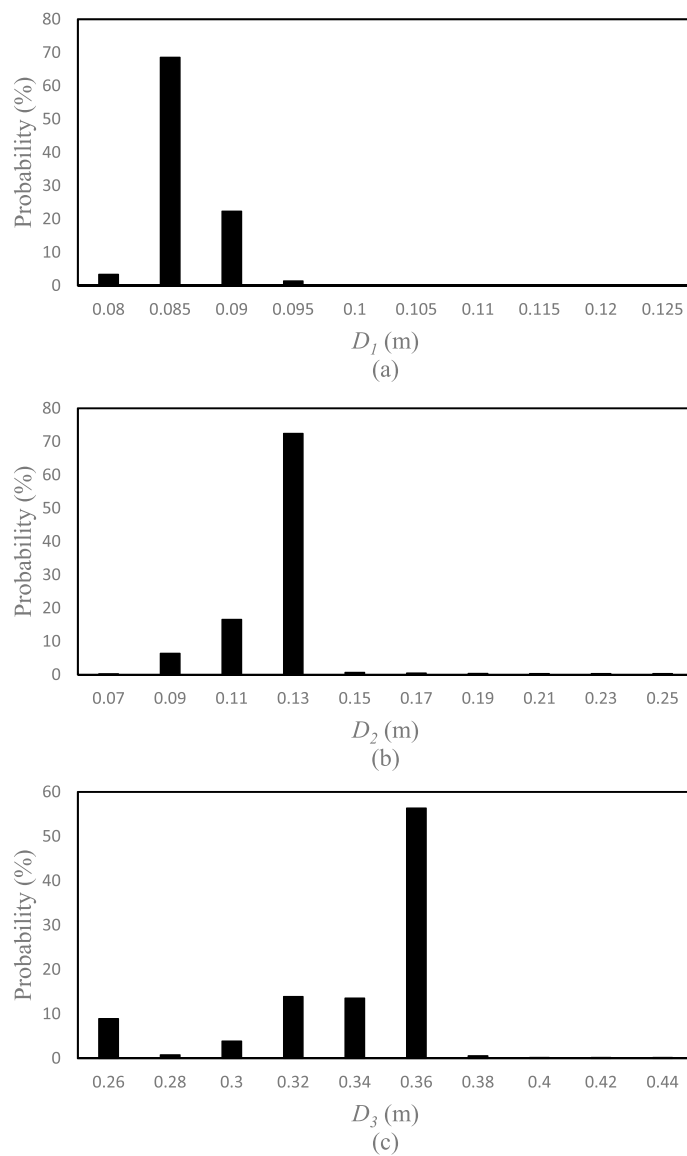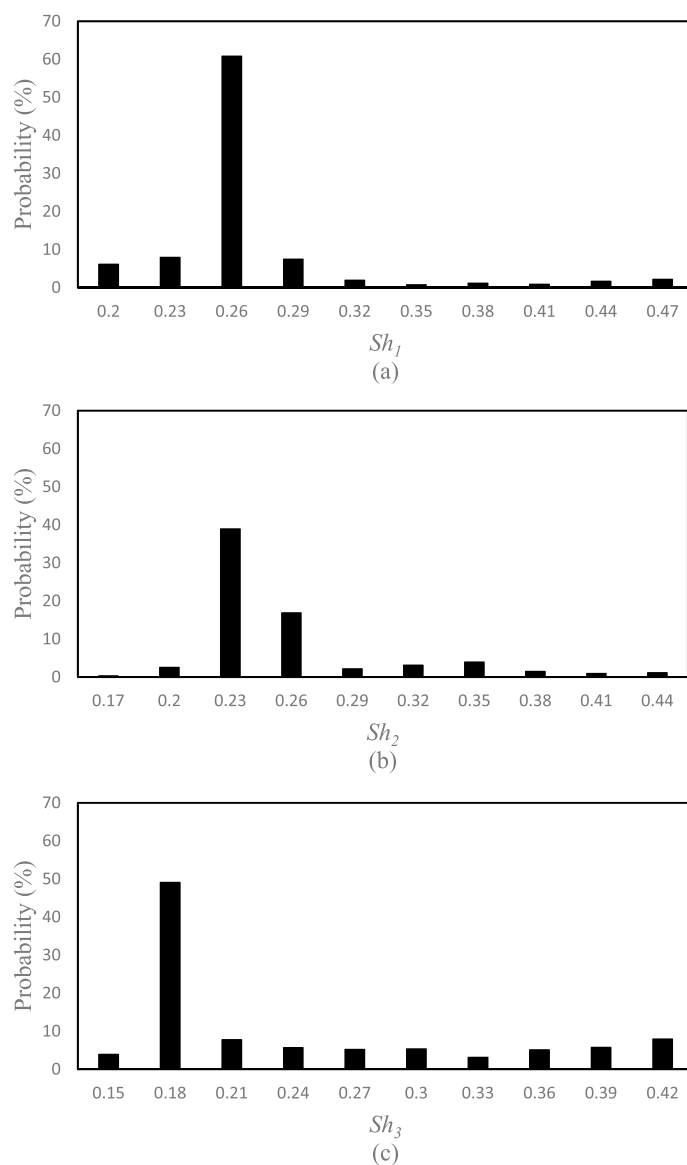
8

Figure 6. Histograms of the HFLUX inputs at three sections (**a**) Creek depth at the beginning of the reach study range (0 m), (**b**) Creek depth at 375 m from the beginning of the reach study range, and (**c**) Creek depth at the end of the reach study range (475 m).

Figure 7. Histograms of the HFLUX inputs at three sections (**a**) percent shade coefficient at the beginning of the reach study range (0 m), (**b**) percent shade coefficient at 375 m from the beginning of the reach study range, and (**c**) percent shade coefficient at the end of the reach study range (475 m).
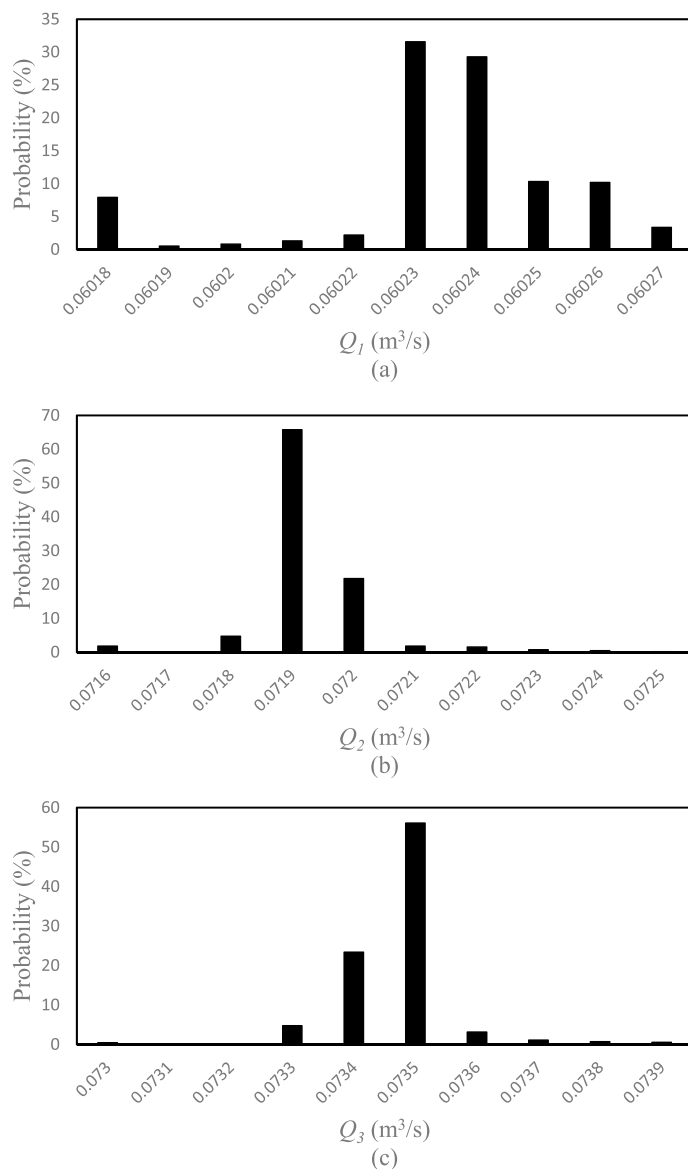
Figure 8. Histograms of the HFLUX inputs at three sections (**a**) streamflow at the beginning of the reach study range (0 m), (**b**) streamflow at 375 m from the beginning of the reach study range, and (**c**) streamflow at the end of the reach study range (475 m).
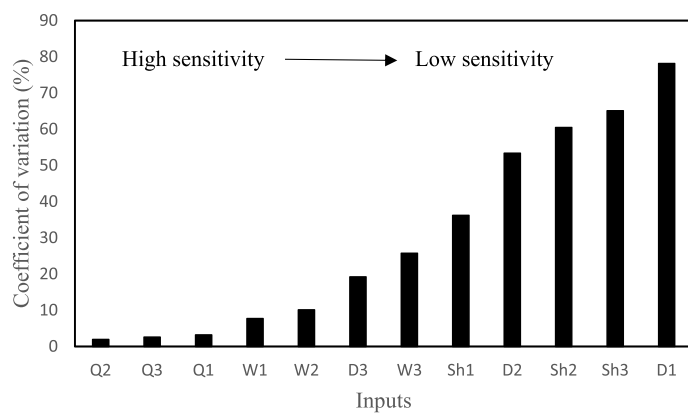
Figure 9. Order of sensitivity of the inputs quantified by the CV values.

Input	CV (%)	Most likely value	Input	CV (%)	Most likely value
$W1$ (m)	7.71	5.6	$Sh1$ (dimensionless)	36.20	0.26
$W2$ (m)	10.10	3.7	$Sh2$ (dimensionless)	60.46	0.23
$W3$ (m)	25.74	3.1	$Sh3$ (dimensionless)	65.08	0.18
$D1$ (m)	78.14	0.085	$Q1$ (m³/s)	3.18	0.06023
$D2$ (m)	53.35	0.130	$Q2$ (m³/s)	1.94	0.07190
$D3$ (m)	19.19	0.360	$Q3$ (m³/s)	2.57	0.07350

Table 3. Coefficient of variation (CV) and most likely values of inputs.
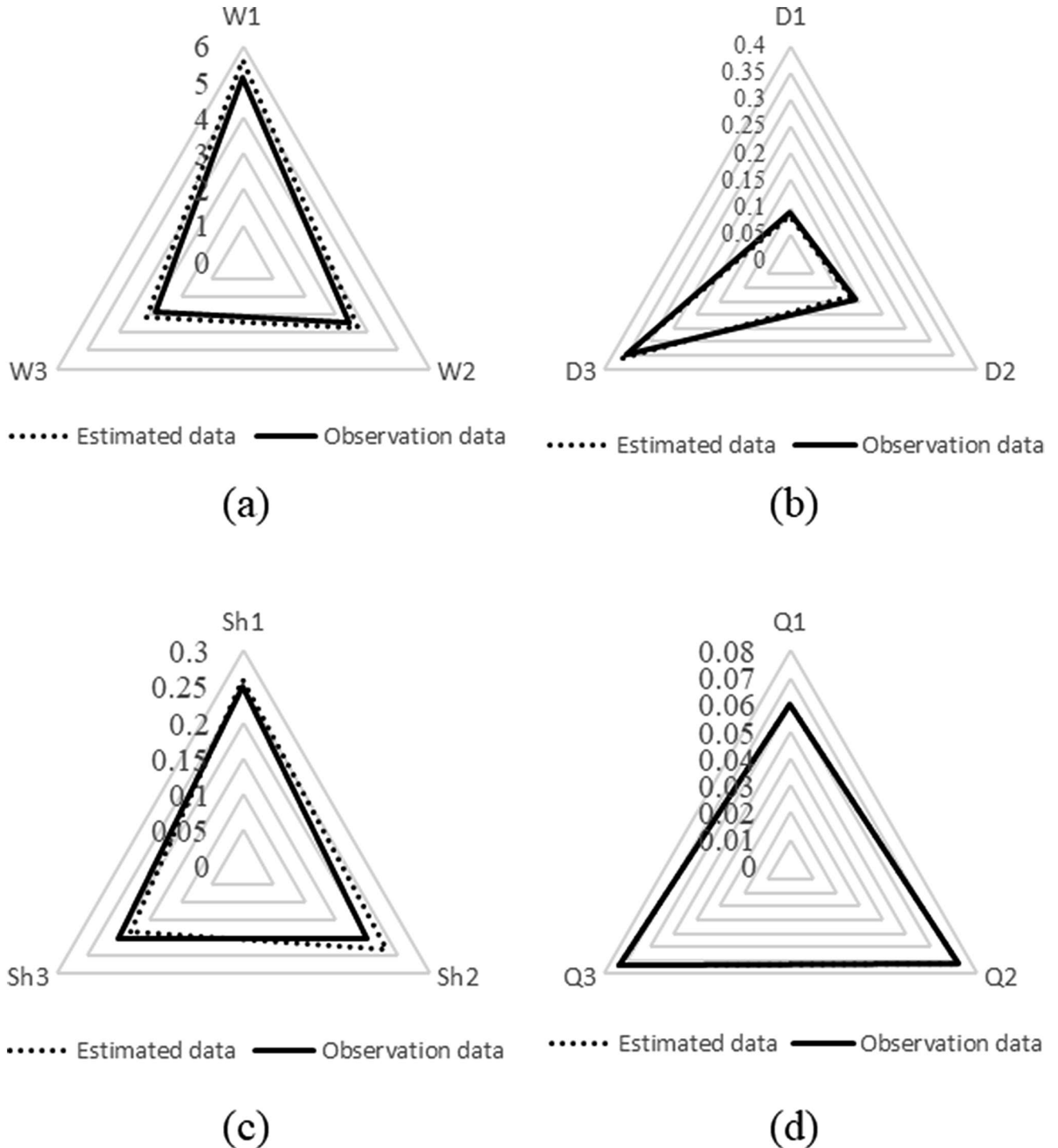
Figure 10. Comparison of the observed data and the most likely values estimated by the SCEM-UA algorithm: (**a**) width, (**b**) depth, (**c**) percent shade, and (**d**) streamflow.

Input	Relative error (%)	Input	Relative error (%)
W_1	9.80	Sh_1	4.00
W_2	7.24	Sh_2	15.00
W_3	10.71	Sh_3	10.00
D_1	5.56	Q_1	0.11
D_2	7.14	Q_2	0.03
D_3	2.86	Q_3	0.16

Table 4. Relative errors of the inputs in the SCEM-UA algorithm.

Data availability
All of the required data have been presented in our article.

References
1. Abdi, R. & Endreny, T. A river temperature model to assist managers in identifying thermal pollution causes and solutions. *Water* **11**(5), 1060. https://doi.org/10.3390/w11051060 (2019).
2. Abdi, R., Endreny, T. & Nowak, D. A model to integrate urban river thermal cooling in river restoration. *J. Environ. Manage.* **258**, 110023. https://doi.org/10.1016/j.jenvman.2019.110023 (2020).
3. Ajami, N. K., Duan, Q. & Sorooshian, S. An integrated hydrologic Bayesian multi-model combination framework: confronting input, parameter and model structural uncertainty in hydrologic prediction. *Water Resour. Res.* **43**, 1. https://doi.org/10.1029/2005WR004745 (2007).
4. Bernhardt, E. S. *et al.* Synthesizing river restoration efforts. *Science* **308**(5722), 636–637. https://doi.org/10.1126/science.1109769 (2005).
5. Feyen, L., Beven, K. J., De Smedt, F. & Freer, J. Stochastic capture zone delineation within the generalized likelihood uncertainty estimation methodology: Conditioning on head observations. *Water Resour. Res.* **37**(3), 625–638. https://doi.org/10.1029/2000WR900351 (2001).
6. Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **7**(4), 457–472. https://doi.org/10.1214/ss/1177011136 (1992).
7. Georgakakos, K. P., Seo, D. J., Gupta, H., Schaake, J. & Butts, M. B. Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. *J. Hydrol.* **298**(1–4), 222–241. https://doi.org/10.1016/j.jhydrol.2004.03.037 (2004).
8. Glose, A. M., Lautz, L. K. & Baker, E. A. Stream heat budget modeling with HFLUX: Model development, evaluation, and applications across contrasting sites and seasons. *Environ. Model. Softw.* **92**, 213–228. https://doi.org/10.1016/j.envsoft.2017.02.021 (2017).
9. Lin, K., Liu, P., He, Y. & Guo, S. Multi-site evaluation to reduce parameter uncertainty in a conceptual hydrological modeling within the GLUE framework. *J. Hydroinf.* **16**(1), 60–73. https://doi.org/10.2166/hydro.2013.204 (2014).
10. Liu, J., Dong, X. & Li, Y. Automatic calibration of hydrological model by shuffled complex evolution metropolis algorithm. *Int. Conf. Artif. Intell. Comput. Intell. Sanya* **3**, 256–259. https://doi.org/10.1109/AICI.2010.291 (2010).
11. Liu, J., Shao, W., Xiang, C., Mie, C. & Li, Z. Uncertainties of urban flood modeling: Influence of parameters for different underlying surface. *Environ. Res.* **182**, 108929. https://doi.org/10.1016/j.envres.2019.108929 (2020).
12. Morrill, J. C., Bales, R. C. & Conklin, M. H. Estimating stream temperature from air temperature: implications for future water quality. *J. Environ. Eng.* **131**(1), 139–146 (2005).
13. Papadopoulos, C. E. & Yeung, H. Uncertainty estimation and Monte Carlo simulation method. *Flow Meas. Instrum.* **12**(4), 291–298. https://doi.org/10.1016/S0955-5986(01)00015-2 (2001).
14. Sun, N., Hong, B. & Hall, B. Assessment of the SWMM model uncertainties within the generalized likelihood uncertainty estimation (GLUE) framework for a high-resolution urban sewershed. *Hydrol. Process.* **28**(6), 3018–3034. https://doi.org/10.1002/hyp.9869 (2013).
15. Tang, H., Guo, X., Xie, L. & Xue, S. Experimental validation of optimal parameter and uncertainty estimation for structural systems using a shuffled complex evolution metropolis algorithm. *Appl. Sci.* **9**(22), 4959. https://doi.org/10.3390/app9224959 (2019).
16. Yan, L., Jin, J. & Wu, P. Impact of parameter uncertainty and water stress parameterization on wheat growth simulations using CERES-wheat with GLUE. *Agric. Syst.* **181**, 102823. https://doi.org/10.1016/j.agsy.2020.102823 (2020).
17. Yang, J., Reichert, P., Abbaspour, K. C., Xia, J. & Yang, H. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *J. Hydrol.* **358**(1–2), 1–23. https://doi.org/10.1016/j.jhydrol.2008.05.012 (2008).
18. Yang, X. E., Wu, X., Hao, H. L. & He, Z. L. Mechanisms and assessment of water eutrophication. *J. Zhejiang Univ. Sci.* **9**(3), 197–209. https://doi.org/10.1631/jzus.B0710626 (2008).
19. Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W. & Verstraten, J. M. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resour. Res.* **41**, 1. https://doi.org/10.1029/2004WR003059 (2005).
20. Vrugt, J. A., Gupta, H. V., Bouten, W. & Sorooshian, S. A Shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.* **39**, 8. https://doi.org/10.1029/2002WR001642 (2003).
21. Vrugt, J. A., TerBraak, C. J. F., Clark, M. P., Hyman, J. M. & Robinson, B. A. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resour. Res.* **44**, 12. https://doi.org/10.1029/2007WR006720 (2008).
22. Zheng, Y., Arabi, M. & Paustian, K. Analysis of parameter uncertainty in model simulation of irrigated and rainfed agroecosystems. *Environ. Model. Softw.* **126**, 104642. https://doi.org/10.1016/j.envsoft.2020.104642 (2020).

Author contributions

B.A.; software, formal analysis, writing—original draft. O.B.-H.; conceptualization, supervision, project administration. X.C.; validation, writing—review and editing.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to O.B.-H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.