Journal of Hydrology 579 (2019) 124195

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

A bootstrap approach for the parameter uncertainty of an urban-specific rainfall-runoff model

Saritha Padiyedath Gopalan[a,*], Akira Kawamura[a], Hideo Amaguchi[a], Tadakatsu Takasaki[b], Gubash Azhikodan[a]

[a] Dept. of Civil and Environmental Eng., Tokyo Metropolitan University, 1-1 Minami Osawa, Hachioji, Tokyo 192-0397, Japan
[b] The Bureau of Construction, Tokyo Metropolitan Government, 1-9-15 Shinsuna, Koto-ku, Tokyo 136-0075, Japan

ARTICLE INFO

This manuscript was handled by Marco Borga, Editor-in-Chief, with the assistance of Alessio Domeneghetti, Associate Editor

Keywords:
Model parameter uncertainty
Urban storage function model
Residual-based bootstrap method
Model parameter uncertainty indices
Model simulation uncertainty indices

ABSTRACT

The predictions made using rainfall-runoff models are inherently uncertain and it is important to recognize and account for this uncertainty, especially in urban watersheds due to the high flood risk in these areas. Recent studies on hydrological model uncertainty mostly refer to the identification of model parameter uncertainty. However, such studies are somewhat limited using the bootstrap approach, a nonparametric method which makes less prior assumptions on the model structure and thus is more flexible. Hence, a residual-based bootstrap approach associated with the SCE-UA global optimization algorithm is demonstrated in this study for the analysis of calibrated parameter uncertainty and its subsequent effect on the model simulation of an urban-specific rainfall-runoff model, urban storage function (USF) model, under two different data scenarios of individual event-based and whole data-based scenarios. Initially, the parameter uncertainty was expressed by estimating the confidence interval (CI) of the USF model parameters obtained from bootstrapping and then the parameters from the highest to the lowest uncertainties were derived by utilizing two newly proposed parameter uncertainty indices which can make the best use of CI. Moreover, investigations on the effect of calibrated parameter uncertainty on model simulations revealed that the model was able to bracket most of the observations within the prediction range of considered scenarios. This further indicates that the residual-based bootstrap approach along with the SCE-UA method reasonably well predicted the uncertainty range of the USF model. For a better understanding of simulation uncertainty, we defined and demonstrated two model simulation uncertainty indices and these indices could be useful in future studies to analyze the simulation uncertainty of different rainfall-runoff models in the watersheds worldwide.

1. Introduction

In urban watersheds, the urban areas occupy most of the basin from upstream to downstream and are under constant development in terms of buildings, roads, other infrastructures, etc. (Kawamura, 2018). There is no specific definition of urban watersheds quantitatively in terms of the threshold urban area. However, according to Kjeldsen (2010) and Salavati et al. (2016), the watersheds with an urban area percentage greater than 15% can be considered as urban watersheds. The modeling of urban watershed processes is complicated due to the increasing complexities of the urban hydrologic system, that can be attributed to urbanization, rapid population growth, model scale, etc. (McPherson and Schneider, 1974). Urbanization is a radical form of land-use change and will replace the natural land-use with impervious surfaces and storm drainage system (Kjeldsen, 2010). This further inhibits the natural infiltration capacity and reduces the lag time. The other hydrological impacts of urbanization can be listed as the increase of low return period floods more than the high return period floods, increase in drainage density and channel cross-sectional area, flashiness of storm flow, etc. (Graf, 1977; Hollis, 1975). These changes alter the runoff process significantly and accelerate the rainfall-runoff transformation process which could lead to higher and rapid flood flows (Hollis, 1975; Salavati et al., 2016). In addition, the urban watersheds face constant and drastic changes in terms of frequent occurrence of high intensity rainfall due to the heat island phenomena (Bornstein and Lin, 2000), leakage from the water distribution system, increased flood and inundation risk (Suriya and Mudgal, 2012), increase in human settlement and associated activities, removal of vegetation and natural storage,

* Corresponding author.
E-mail addresses: charu666@gmail.com (S. Padiyedath Gopalan), kawamura@tmu.ac.jp (A. Kawamura), amaguchi@tmu.ac.jp (H. Amaguchi), takasaki@doboku.metro.tokyo.jp (T. Takasaki), gubash@tmu.ac.jp (G. Azhikodan).

https://doi.org/10.1016/j.jhydrol.2019.124195
Received 24 March 2019; Received in revised form 10 August 2019; Accepted 27 September 2019
Available online 28 September 2019

etc. (Amaguchi et al., 2012; WMO, 2008).

In contrast to urban watersheds, the percentage impervious areas as well as the installed sewer system density is either very low or absent in rural watersheds and will undergo slow and steady changes of hydrologic and hydraulic characteristics. They will retain a significant amount of rainfall due to the presence of natural land-use (Hollis, 1975; Sheng and Wilson, 2009) and will substantially decrease the peak discharge and increases the flood duration (Sheng and Wilson, 2009). Further, Oudin et al. (2018) analyzed the effect of distance of urban areas from the catchment outlet on runoff response and the results revealed that the location of urban areas has a secondary effect compared with the percentage impervious surfaces. Hence, the corresponding flood risk will be moderate in rural watersheds although it includes urban areas at the outlet due to the relatively low population and settlements it accommodates. Therefore, it is very important to detect urban flood flows compared to those in rural areas as they are associated with increased risks and costs (Mason et al., 2012). For this purpose, the rainfall-runoff models are important tools and they play a central role in urban watersheds (Padiyedath et al., 2018a).

However, the predictions made using rainfall-runoff models are inherently uncertain. The uncertainties in the modeling mainly arise from parameter uncertainties, measurement errors associated with the input data, and from model structure errors arising from the aggregation of spatially distributed watershed processes into a relatively simple runoff model (Sivakumar and Berndtsson, 2010). The models which are not parameterized properly or lack input data of reasonable quality could provide an inaccurate representation of the actual hydrological processes (Shrestha, 2009). Generally, the effectiveness of a model in providing a good prediction of the hydrological processes is mainly determined by its parameter values (Jeremiah et al., 2012). Therefore, the model parameter uncertainty has received a prime recognition over other sources of uncertainties in the field of hydrological modeling and recent studies on hydrological model uncertainties mostly refer to the identification of parameter uncertainty (Uhlenbrook et al., 1999) or parameter calibration (Ajami et al., 2004) and their impacts on the model simulation results (Freer et al., 1996). Traditionally, model parameters have been quantified from watershed properties or theoretically analyzed by comparing with similar models (Li et al., 2010). However, since the introduction of computer-intensive statistics, the parameters are being estimated by calibrating the models against observed data even though it increases the efforts and costs taken for the data measurements required for calibration and validation. However, this calibrated parameters are influenced by factors such as quantity and quality of input data, model error, correlation between the parameters, etc. (Duan et al., 1992) which will lead the parameter estimates to deviate from their underlying true values (Ebtehaj et al., 2010) thus leading to a great level of uncertainty of parameters, and cannot fully characterize the actual processes.

Different quantity and quality of input data used for calibration may provide quite different optimum parameter sets, and the resulting parameter distributions would reflect the uncertainty in the parameter estimates and the interaction between the individual parameters (Beven, 1993). Generally, the hydrologists calibrate their models with all the available data (e.g. Vaze and Teng, 2011) to get the more robust parameter set. However, several studies analyzed the effect of input data quantity on parameter uncertainty by calibrating the model for different sub periods of all available data. KlemeŠ (1986) proposed a framework to evaluate the model transposability over time by calibrating the model for a selected sub period and validating in another period different from that used for parameter calibration. Later, Jakeman et al. (1993) investigated the parameter fluctuations and associated uncertainties in model predictions for variable climatic conditions by calibrating a lumped rainfall-runoff model for sub periods. Recently, Vaze et al. (2010) analyzed the transposability of four rainfall-runoff models to a changing climatic scenario different from those used for parameter calibration and they found that the model performance was affected by the calibrated parameters. Subsequently, Poulin et al. (2011) studied the effects of model structure and parameter equifinality on the model simulation uncertainty in a climate change impact perspective. Soon after, Merz et al. (2011) examined the parameter time stability of a conceptual model by calibrating the model in different sub periods. Moreover, Coron et al. (2012) also analyzed the parameter time variability of three hydrological models using the split sample test and its effect on the model simulation. Thereafter, Brigode et al. (2013) investigated the dependence of the optimal parameter set on the climate characteristics of the calibration period and the results revealed that the dependence can contribute variability in streamflow projections.

However, the calibrated parameters often have no measurable reference in nature (Bellprat, 2013). Therefore, the major concern which arises is that what confidence bounds can be placed on the calibrated parameters for a given period, where the different sources of uncertainty is mainly arising from observational (with different input data quality) and model parameterization errors and how do these uncertainties affect the hydrologic simulations (Ebtehaj et al., 2010). This parameter uncertainty further contributes to model simulation uncertainties, but to what extent is unknown. Hence, its quantitative evaluation is critical in reducing the uncertainty of these simulations. Recent researchers have paid more attention to these uncertainties and many uncertainty analysis techniques have been developed and applied to different catchments in the past decades (Yang et al., 2008). Most of these techniques rely on either parametric methods or Bayesian methods (Selle and Hannah, 2010; Yang et al., 2008). The most traditional parametric methods used for the assessment of model parameter uncertainty are the linear analysis (first-order approximation) by providing a rough confidence interval (CI) of parameters (Kuczera, 1988), nonlinear constrained maximization or minimization, etc. (Gallagher and Doherty, 2007). However, in the parametric method, the structure of the model is specified a priori and the number and nature of the parameters are generally fixed in advance, and there is a little flexibility (Sivakumar, 2017). Since the advances in computer technology, the Monte Carlo approaches with a Bayesian inference have become popular due to their ability to handle nonlinearity and interdependency of parameters in complex hydrological models (Li et al., 2010). The other methods based on Bayesian approaches are Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992; Hornberger and Spear, 1981), Sequential Uncertainty Fitting algorithm (SUFI) (Abbaspour et al., 1997, Abbaspour et al., 1999), etc. Still, the Bayesian technique requires the form specification of the error distribution for response variables (Selle and Hannah, 2010). Hence, the nonparametric methods have the advantage that they make less prior assumptions on error structures and, thus, are potentially more flexible.

The bootstrap method, a nonparametric technique, has been developed by Efron (1979) for random resampling of the original data set to develop replicate data sets from which the underlying distribution of the statistics of interest such as mean, variation, correlation, etc. can be estimated (Sivakumar, 2017). This resampling technique has applications in diverse fields like hydrology, groundwater hydrology, air pollution modeling, toxicology, etc. (Dixon, 2006), where it has been successfully used in hydrological modeling to design storms from exceedance series, to develop artificial neural network (ANN) model, to estimate the sampling variability of reconstructed runoff, etc. (Jeong and Kim, 2005; Sun et al., 2013; Zucchini and Adamson, 1989) by utilizing non-time series data. However, the outcome from the rainfall-runoff models is the time-series hydrograph, and hence the time series application of the bootstrap method becomes necessary. Sophisticated approaches have been developed for this purpose and has been extensively used for the trend analysis of temperature and streamflow time series, generation of synthetic streamflow sequences that are used in simulation studies, forecasting of low flow frequency, uncertainty assessment of water quality trends, etc. (Hirsch et al., 2015; Lall and Sharma, 1996; Önöz and Bayazit, 2012; Sonali and Nagesh Kumar,

2013; Srinivas and Srinivasan, 2005; Tasker and Dunne, 1997). However, use of the bootstrap technique for model parameter uncertainty analysis by employing the time series data appears to be quite narrow until recently and very limited studies have been conducted to quantify the calibrated parameter uncertainty of rainfall-runoff models using this technique. Ebtehaj et al. (2010) introduced a nonparametric block bootstrapping approach coupled with global optimization to estimate the parameter uncertainty resulting from uncertainty in the forcing data and evaluate its impacts on the resulting streamflow simulations. Later, Selle and Hannah (2010) demonstrated a model based bootstrap and compared the results with the block bootstrap approach for the parameter uncertainty of abc hydrological model and a conceptual salt load model. The bootstrap approach also seems to have been used for the analysis of parameter uncertainty of SWAT model (Li et al., 2010; Zhang et al., 2014). Further, Brigode et al. (2014, Brigode et al., 2015) analyzed the effect of different rainfall-runoff calibration periods and information contained in the calibration period on extreme flood estimations using the block bootstrap method. The uncertainty studies conducted using the SWAT model for different catchments reported that the parameter uncertainty and its effect on model simulation uncertainty vary from catchment to catchment even for the same model (Li et al., 2010; Zhang et al., 2014). Therefore, there is a need to carry out such studies in different types of watersheds worldwide under varying agro-climatic conditions with different rainfall-runoff models.

In light of the aforementioned discussions, it is apparent that very few studies have been conducted for model parameter uncertainty analysis using the bootstrap method. Among these, none of the studies have been carried out in urban watersheds, particularly using the urban storage function (USF) model (Takasaki et al., 2009), a relatively new storage function (SF) model specially developed for urban watersheds where combined sewer systems are in use. Hence, an attempt has been made to explore the use of bootstrap resampling to evaluate the uncertainty of optimal parameter estimates that arise due to uncertainties in the input data using a case study in the upper Kanda river basin, a typical small to medium sized urban watershed in Tokyo, Japan. Previous studies on model parameter uncertainty analysis using bootstrap approach were conducted using all the available data instead of individual flood events since the hydrologists were mainly interested in the estimation of catchment hydrological variables such as peak flow, flood volume, etc. with utmost accuracy and reliability. The flood-runoff analysis in urban watersheds is generally event based due to the relevance of flash flood peak estimation and short time of concentration. Therefore, the bootstrap approach was applied in this study to both the individual flood events and the whole events in order to demonstrate the impact of different available data scenarios on the uncertainty behavior of calibrated parameters. Additionally, two types of new indices have been proposed for the detailed analysis of uncertainty involved in the model parameters and model simulations.

2. Methodology

2.1. USF model

The USF model is an SF model (Hoshi and Yamaoka, 1982; Kimura, 1961; Prasad, 1967) developed for the urban watersheds by incorporating the outflow from the basin to the treatment plant through the combined sewer system. The model is based on the relationship between rainfall over the basin and runoff at the outlet point and is governed by the following equation (Padiyedath et al., 2018a; Takasaki et al., 2009):

$$s = k_1(Q + q_R)^{p_1} + k_2\frac{d}{dt}(Q + q_R)^{p_2} \tag{1}$$

where s is the watershed storage (mm); Q is the river discharge (mm/min); q_R is the storm drainage from the basin through the combined sewer system (mm/min); t is the time (min); and k_1, k_2, p_1, p_2 are the model parameters. Combining the above expression of storage with the following continuity equation yields the nonlinear expression of the USF model.

$$\frac{ds}{dt} = R + I - E - O - (Q + q_R) - q_l \tag{2}$$

where R is the rainfall (mm/min); I is the urban-specific and groundwater inflows from other basins (mm/min); E is the evapotranspiration (mm/min); O is the water intake from the basin for intended purposes such as water supply, agricultural needs, etc. (mm/min); and q_l is the groundwater-related loss (mm/min). Groundwater-related loss (q_l) was defined by considering the infiltration hole height (z) and is given by the following equation (Padiyedath et al., 2018a; Takasaki et al., 2009):

$$q_l = \begin{Bmatrix} k_3(s - z) & (s \geqslant z) \\ 0 & (s < z) \end{Bmatrix} \tag{3}$$

where k_3 and z are the parameters. The expression for storm drainage q_R from the combined sewer system discharged out of the basin was developed by assuming a linear relationship between total discharge $Q + q_R$ and the storm drainage q_R immediately after the rainfall. The q_R is defined (Padiyedath et al., 2018a; Takasaki et al., 2009) as follows:

$$q_R = \begin{Bmatrix} \alpha(Q + q_R - Q_0) & \alpha(Q + q_R - Q_0) < q_{R\ max} \\ q_{R\ max} & \alpha(Q + q_R - Q_0) \geqslant q_{R\ max} \end{Bmatrix} \tag{4}$$

where α is the slope of the linear relationship between total discharge $Q + q_R$ and the drainage q_R, and Q_0 is the initial river discharge just before the rain starts. The maximum volume of q_R cannot exceed the sewer maximum carrying capacity q_{Rmax}. Substituting Eqs. (1) and (3) into (2) leads to a second-order ordinary differential equation (ODE) and can be numerically solved after transforming into a first-order ODE. The river discharge Q was obtained as the solution after subtracting q_R, which was calculated using Eq. (4), from the total discharge. Overall, the USF model is a seven-parameter model with parameters k_1, k_2, k_3, p_1, p_2, z, and α used in the rainfall-runoff modeling. For a detailed description of the USF model, see Takasaki et al. (2009) and Padiyedath et al. (2018a).

2.2. Model calibration and validation

The Shuffled Complex Evolution-University of Arizona (SCE-UA) method proposed by Duan et al. (1992) was used to calibrate the USF model. It is a well-known global optimization strategy developed for the effective and efficient calibration of the watershed models by optimizing a single objective function for up to 16 parameters (Duan et al., 1992; Green and Van Griensven, 2008). This method combines a simplex method with the concept of controlled random search for the competitive evolution of the population with complex shuffling. The algorithmic parameters of SCE-UA method were selected as per the recommendations of Duan et al. (1993). In the first step of the method, it generates an initial population as the first generation by random sampling from the feasible parameter space, which was defined by setting the lower and upper search range for p number of parameters to be optimized. From the second generation onwards, this population is partitioned into several complexes, each of which is permitted to evolve independently. Size of the population produced in each generation was decided based on the number of parameters in the target model. The number of complexes, C, was set equal to 20 and the number of populations in each complex, $r = 2p + 1$. The objective function to be minimized using the SCE-UA method was selected as the root mean squared error (RMSE) between the observed and simulated discharge. Common use in hydrological modeling and simplicity were the reasons for the selection of RMSE as the objective function (Ebtehaj et al., 2010; WMO, 1992). The calibrated parameters are functionally dependent on the length and properties of the calibration data, objective function used for calibration, etc. and these factors subsequently affect the

Table 1
Description and search range of USF model parameters.

Parameter	Definition	Search range
k_1	Physical watershed characteristics (Sugiyama et al., 1997)	[10, 500]
k_2	Loop relationship between the storage and discharge (Prasad, 1967)	[100, 5000]
k_3	Groundwater related loss	[0.001, 0.05]
p_1	Index of flow regime (Sugiyama et al., 1997)	[0.1, 1]
p_2	Non-linear unsteady flow effects (Hoshi and Yamaoka, 1982)	[0.1, 1]
z	Infiltration hole height	[0, 50]
α	Effect of storm drainage diverted to the treatment plant	[0.1, 1]

model simulations. The seven parameters of USF model are shown in Table 1 with their descriptions and search range used in the SCE-UA parameter optimization method (Padiyedath et al., 2018a; Takasaki et al., 2009). A narrow search range will constrain the parameters and the calibrated parameters will not reflect the actual watershed characteristics. Therefore, a wide search range was defined, instead of a narrow one by considering the possible physical minimum and maximum parameter values. The model calibration was conducted using three data scenarios: (i) whole data-based scenario where all the available events were used for the model calibration, (ii) individual event-based scenario where individual flood events were used, (iii) leave-one-out scenario where leaving out each flood event at a time from the all available events and calibrating on the remaining events. The model performance during calibration was analyzed by simulating the flood events using the whole data-based and individual event-based parameters.

The performance of a model derived from the calibration data set is insufficient evidence for its satisfactory performance since no simulation model is intended merely to show how well it fits the data used for its development. Thus the data used for model validation should be different as those used for calibration but must represent a situation similar to that for which the data are to be generated (KlemeŠ, 1986). Therefore, to test the operational performance of the model, we have used the calibrated parameters from the individual event-based scenario and the leave-one-out scenario. The available flood events are split into two segments in which the first segment consists of one flood event at a time and the second segment is comprised of the remaining flood events. The calibrated parameters of the individual event-based and leave-one-out scenarios from the second segment are used to validate the model on the first segment. For example, event 1 forms the first segment and the second segment is composed of the remaining flood events (events 2–5). The calibrated parameters of individual events 2–5 (four parameter sets) and the whole events except 1 (one parameter set) are used to validate the event 1. This calibration was repeated by leaving out the subsequent flood event (one at a time) from the available events.

2.3. Residual-based bootstrap method

The classical idea behind the bootstrap method is the resampling and extraction of M samples from the original data with an unknown distribution having a size of N. This method makes no assumptions concerning the distribution of data or model being used. The M bootstrap samples can provide the best knowledge regarding the underlying true distribution of the sample. However, the use of bootstrap for time series data was limited because the classical bootstrap technique assumes that the data set is independent and identically distributed (*iid*) (Efron, 1982), which means each data of the data set will be mutually independent and selected from the same population. Generally, the time series data sets are highly dependent in nature and it is quite unreasonable to perform classical bootstrapping as it destroys the original dependency structure of the time series. In order to overcome the problem of dependence of time series, the bootstrap method can be extended either as a block bootstrap (Davison and Hinkley, 1997; Künsch, 1989), in which the time series is divided into different blocks and these blocks are resampled instead of the individual data value or as a model based bootstrap (Lahiri, 2003; Selle and Hannah, 2010), which adopts a specific time-series form of dependence. In Selle and Hannah (2010), they considered a first-order autoregressive model to consider the dependence of model error rather than assuming it is independent. However, a modification of this approach, which can be simpler to apply in practice, is to construct the bootstrap samples as fitted values plus residuals, where the residuals are sampled with replacement from the observed distribution of residuals. This method is known as the resampling of residuals (Shalizi, 2016; Shao and Tu, 1995) (hereafter residual-based bootstrap).

The residual-based bootstrap approach was used to generate sample estimates of the hydrologic model parameters corresponding to the calibrated data set and to quantify the associated uncertainties. In this method, first we calibrate the model, and then simulate by resampling residuals to that estimate and adding them back to the fitted values (Shalizi, 2016). This surrogate data set is then re-analyzed like a new data set. By repeating the procedure M times, bootstrapped time series are generated and the hydrologic model is then calibrated using each bootstrapped time series, to arrive at bootstrapped estimates of the calibrated parameter sets. These estimates are further used for the confidence interval analysis of the model parameter estimators. However, the residual-based bootstrap method is usually based on some model assumptions although it requires no theoretical formula for the quantity to be estimated and is less model-dependent than the traditional approach (Shao and Tu, 1995). The general assumptions for performing the residual-based bootstrap are given as,

(1) The residuals are independent of each other.
(2) The residuals are identically distributed.
(3) The residuals are homoscedastic.
(4) The residuals are a good approximation of the observed data and model structure error.
(5) The residuals are randomly selected from their population with equal probability.

The procedure for the residual-based bootstrap, explained by Stine (1985) and others (Shalizi, 2016; Shao and Tu, 1995), is described as follows. Consider the original data set $\{X(N), Q(N)\}$; where $X(N)$ is the input data, $Q(N)$ is the observed discharge data, and N is the data length. The observed discharge can be written as a function, $Q(N) = F(\boldsymbol{X}, \theta) + \varepsilon(N)$ where $\boldsymbol{X} = X(1), ...,X(N)$, θ is the parameter vector $\theta_1, ...\theta_p$ with p being the number of model parameters, and $\varepsilon(N)$ is the model residuals. Initially, the model was calibrated using the SCE-UA method to obtain the calibrated parameter vector $\hat{\theta}$ which was further used along with the input data to compute the model calibrated discharge data, $\hat{Q}(N) = F(\boldsymbol{X}, \hat{\theta})$, which can be demonstrated as $Q(N) = F(\boldsymbol{X}, \hat{\theta}) + \hat{\varepsilon}(N)$. Thereafter, the model residuals can be expressed using the following equation.

$$\hat{\varepsilon}(N) = Q(N) - \hat{Q}(N) = Q(N) - F(\boldsymbol{X}, \hat{\theta}) \quad (5)$$

The model residuals, $\hat{\varepsilon}(N)$, were assumed to be *iid* for all N, which is one of the assumptions made for the bootstrapping (Stine, 1985). Julian and Gardner (2014) examined the effect of land cover on runoff patterns and the results revealed that increases in urbanization caused a decrease in long-term hydrologic memory. In urban watersheds, the impervious surfaces decrease water storage, which is the predominant factor that affecting long-term hydrologic memory, and the runoff became flashier (Julian and Gardner, 2014). This flash flood decreases the watershed hydrologic memory. Therefore, the resulting model residuals of urban watersheds can be assumed to hold low memory due to the

quick runoff response resulted from the increased percent of impervious surfaces. The detailed residual-based bootstrapping procedure is outlined as follows:

(1) Bootstrap resampling of the residual time series $\hat{\varepsilon}(N)$ with replacement to form new bootstrapped residual series, $\hat{\varepsilon}^j(N)$, where $j = 1, ..., M$.
(2) Add the new bootstrapped residual series, $\hat{\varepsilon}^j(N)$ to the calibrated discharge data $\hat{Q}(N)$ to form the bootstrapped discharge series as $Q^j(N) = \hat{Q}(N) + \hat{\varepsilon}^j(N)$. This bootstrapped discharge series will form the replication of the observed discharge series.
(3) Calibrate the bootstrapped discharge series, $Q^j(N)$ with input data set, $\boldsymbol{X}$ and obtain the j^{th} bootstrapped parameter vector $\hat{\theta}^j$ using SCE-UA method and the associated simulated discharge series $\hat{Q}^j(N) = F(\boldsymbol{X}, \hat{\theta}^j)$. By this way, the model can be re-fitted to each bootstrapped discharge series, $Q^j(N)$ yielding 'bootstrap estimates' of model parameters.
(4) Derive the ordered bootstrap estimates $\hat{\theta}^j = \hat{\theta}^j(1), ..., \hat{\theta}^j(p)$ obtained after the bootstrap resampling method. Then, the 95% CI for $\hat{\theta}^j$ was estimated from the ordered bootstrap samples.

The essential idea of this bootstrap approach is that the pseudo replicate samples (bootstrap samples) drawn at random with replacement from the data can be used to furnish information about the uncertainty of quantities estimated from the data (Selle and Hannah, 2010). The precision of a bootstrap estimate depends on the number of times the original data set is randomly resampled (i.e., how many bootstrap replicates). The bootstrap estimate converges to a consistent range of values as the number of resamples become large by the law of large numbers (Meyer et al., 1986). Efron and Tibshirani (1993) suggested that the number of bootstrap samples should be at least 1000. Therefore, all the computations performed in this study were based on 1000 bootstrap replicates ($M = 1000$) from which the confidence intervals for the original corresponding parameter estimates can be calculated. This method has asymptotic convergence properties which means that by increasing the number of simulated bootstrapped time series, estimation error will be reduced (Ebtehaj et al., 2010). The whole data-based and individual event-based scenarios described in Section 2.2 were used to perform the residual-based bootstrapping in order to identify the parameter fluctuation under the varying conditions of available data scenarios.

2.4. Model parameter uncertainty quantification

The performance of hydrological models is significantly affected by the calibrated parameter uncertainty. The parameter uncertainty is generally expressed by estimating the CI of the parameters. However, the CI gives the uncertainty range of each parameter from which it is difficult to identify the parameter with the highest and least uncertainty due to the different ranges of parameter values. Therefore, it is necessary to propose certain indices which can interpret the CI of parameters with different ranges and clearly differentiate them based on their contribution to the uncertainty. Therefore, in addition to the statistical estimators of the mean ($\bar{\theta}$), median ($\hat{\theta}_{50}$), and the coefficient of variation (CV) of the M parameter sets from bootstrapping method, two indices were proposed for assessing the uncertainty of model parameters (model parameter uncertainty indices), which can elucidate the CI in a better way. Here, the indices are considering the individual parameters rather than the parameter vector in order to derive the parameters from the highest to the lowest uncertainty (Selle and Hannah, 2010). The first parameter uncertainty index (PUI_1) utilizes the width of the CI and hence the parameter with a small index value will be less uncertain as compared with other parameters. The second parameter uncertainty index (PUI_2) compares the median value from the confidence region with the calibrated parameter vector $\hat{\theta}$ and it should be minimum for the stability of the parameters. The model parameter uncertainty indices are given as,

$$PUI_1(i) = \left(\frac{\hat{\theta}_{97.5}(i) - \hat{\theta}_{2.5}(i)}{\hat{\theta}(i)} \right) \times 100 \tag{6}$$

$$PUI_2(i) = \frac{\hat{\theta}(i) - \hat{\theta}_{50}(i)}{\hat{\theta}(i)} \times 100 \tag{7}$$

where $\hat{\theta}_{97.5}$ is the 97.5th percentile; $\hat{\theta}_{2.5}$ is the 2.5th percentile; and $\hat{\theta}_{50}$ is the 50th percentile (median) for the i^{th} parameter obtained from bootstrapping. $\hat{\theta}$ is the calibrated parameter.

2.5. Model simulation uncertainty

The model simulation uncertainty is referred to as the uncertainty of simulated discharge, $\hat{Q}^j$, which occurs due to the calibrated parameter uncertainty, and is illustrated as the 95% CI of simulated discharge series by the model. This 95% CI should envelope most of the observations and at the same time, it is desirable to have a narrow envelope (Swain and Patra, 2017). *P*-factor is a statistical term used for the assessment of model simulation uncertainty and is calculated as the percentage of original discharge data at each time step that lies within the 95% CI (Yang et al., 2008). The value of the *P*-factor ranges between 0 and 100% and the goodness of model simulation uncertainty is judged based on the closeness of *P*-factor to 100% (i.e., all observations bracketed within the 95% CI). The *P*-factor is computed as follows:

$$P - \text{factor} = (n/N) \times 100 \tag{8}$$

where n is the number of original observed discharge values at each time step that are bracketed within the 95% CI. In addition to the *P*-factor, two other indices have been proposed for assessing the uncertainty of model simulated discharge (model simulation uncertainty indices) by utilizing the 95% CI similar to the model parameter uncertainty indices. The model simulation uncertainty indices (*SUI*) are given below.

$$SUI_1 = \frac{1}{N} \sum_{t=1}^{N} \left(\frac{\hat{Q}_{97.5}(t) - \hat{Q}_{2.5}(t)}{Q(t)} \right) \times 100 \tag{9}$$

$$SUI_2 = \frac{1}{N} \sum_{t=1}^{N} \left(\frac{Q(t) - \hat{Q}_{50}(t)}{Q(t)} \right) \times 100 \tag{10}$$

where $\hat{Q}_{97.5}(t)$, $\hat{Q}_{2.5}(t)$, and $\hat{Q}_{50}(t)$ are the 97.5%, 2.5%, and 50% levels of the cumulative distribution of the model simulated discharge series respectively. $Q(t)$ is the observed discharge data.

2.6. Rainfall spatial variability

Usually, the characteristics of rainfall events are spatially distributed and different from others even for the same watershed due to the effect of several meteorological factors which will result in the different parameter values. Moreover, there will be an interaction between the spatial variability in rainfall and the spatial storage distribution which controls the discharge. The discharge response at the outlet point to an averaged input will differ significantly from that to a distributed input (Shah et al., 1996). High spatial variability of rainfall in the basin can affect the runoff prediction capability of USF model since it uses the average basin rainfall. Therefore, further analysis was carried out to have a clear understanding about the extent of spatial variability of rainfall in the watershed by computing the percentage variation of total rainfall obtained from each rain gauge with respect to the mean rainfall from all the gauges. The percentage variation can be computed using the following formula.

$$\%\ variation(i) = \frac{TR_i - \bar{TR}}{\bar{TR}} \tag{11}$$

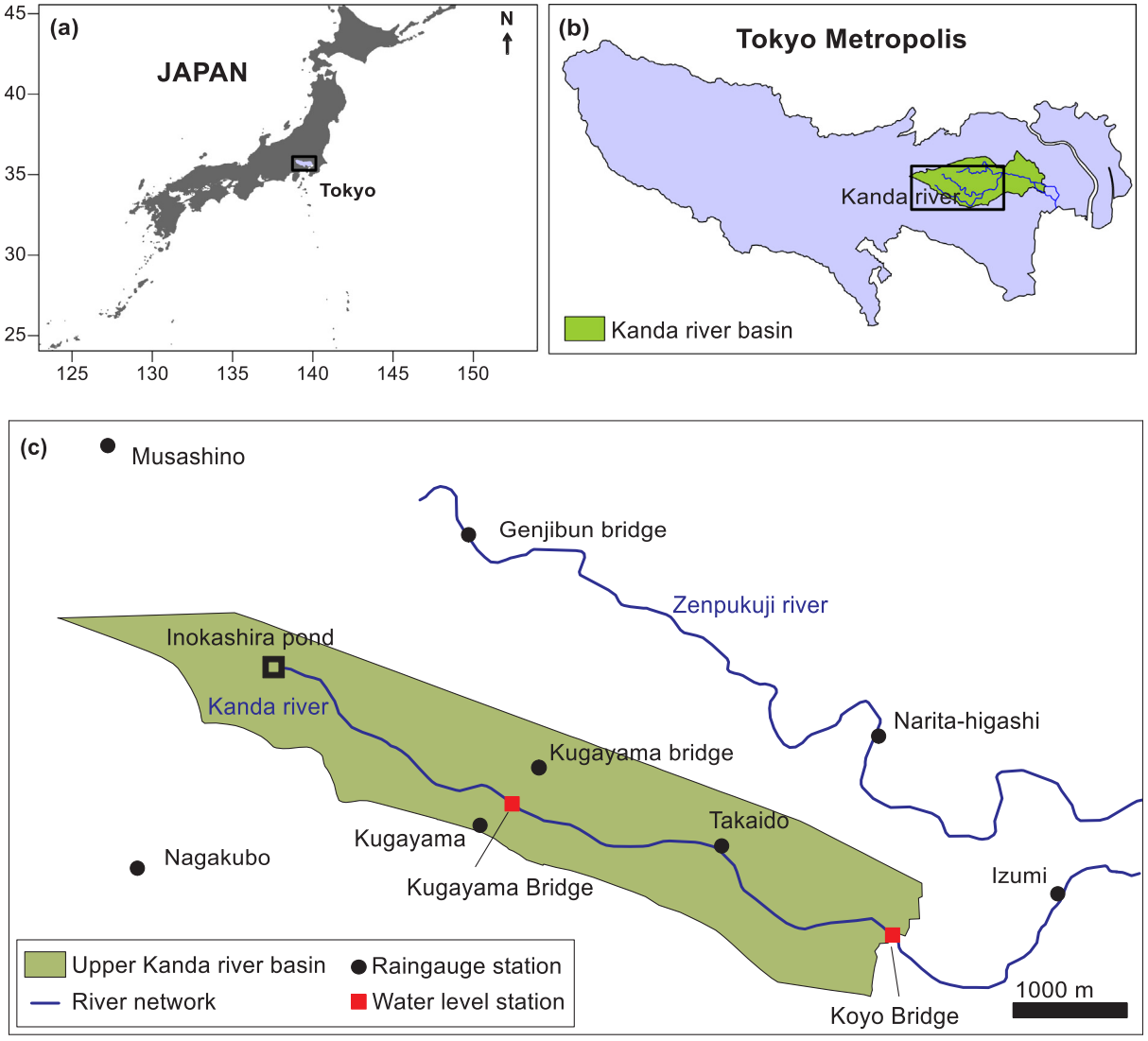

Fig. 1. Index map of (a) Japan, (b) Kanda river basin in Tokyo and (c) target watershed – upper Kanda river basin at Koyo Bridge.

where TR_i is the total rainfall from gauge i (mm), and $\bar{TR}$ is the mean rainfall from all the gauges (mm).

3. Study area and data used

The urban watershed of the upper Kanda River basin, having an area of 7.7 km^2 at Koyo Bridge as shown in Fig. 1, was selected for the study. The basin lies between latitudes 35.70° N and 35.64° N and longitudes 139.56° E and 139.64° E in Tokyo, Japan, with an urbanization rate of 97% since 2003 (TMG, 2016). The river originates from the Inokashira pond and joins the Zenpukuji River and flows east (Ando and Takahasi, 1997). The drainage pattern follows the combined sewer system and 100% of the population is connected to the sewer. The urban landscape GIS delineation was used to precisely estimate the impervious area percentage as 68% (Koga et al., 2016), that significantly reduced the water retention capacity of the basin. The computed time of concentration of surface runoff from the upstream reaches to the watershed outlet was about 30 min. The reduced time of concentration indicated that the river discharge will occur immediately after the rainfall within a short period and it is desirable to use the hydrological data at very short time intervals for the rainfall-runoff analysis. Therefore, rainfall and water level data at one-minute intervals were collected from the Bureau of Construction, Tokyo Metropolitan Government (TMG) during 2003–2006 for the present study.

The average rainfall of the basin was determined using the Thiessen polygon method from the eight rain gauges scattered over the basin, as shown in Fig. 1. Five target events, whose 60-minute maximum rainfall (R_{60}) was greater than 30 mm and were capable of producing flash floods, were selected from the data. Table 2 shows the characteristics of the five selected rainfall events. The inflow component I in the continuity equation was fixed at 0.0012 mm/min, based on the annual report of the Bureau of Construction, TMG. The water intake O from the basin and evapotranspiration E were set at 0 as there was no intake from the target basin and the evapotranspiration during heavy rainfall

Table 2
Characteristics of the five selected events.

Event No.	Event date	R_{60}[†] (mm)	Total R[‡] (mm)	Meteorological factors
1	13-10-2003	53.9	57.5	Intensive localized storm
2	25-06-2003	42.6	46.2	Frontal rainfall
3	8~10/10/2004	42.0	261.1	Typhoon
4	11-09-2006	32.7	37.9	Frontal rainfall
5	15-07-2006	31.5	31.5	Frontal rainfall

[†] R_{60} is the 60-minute maximum rainfall.
[‡] R is the rainfall.

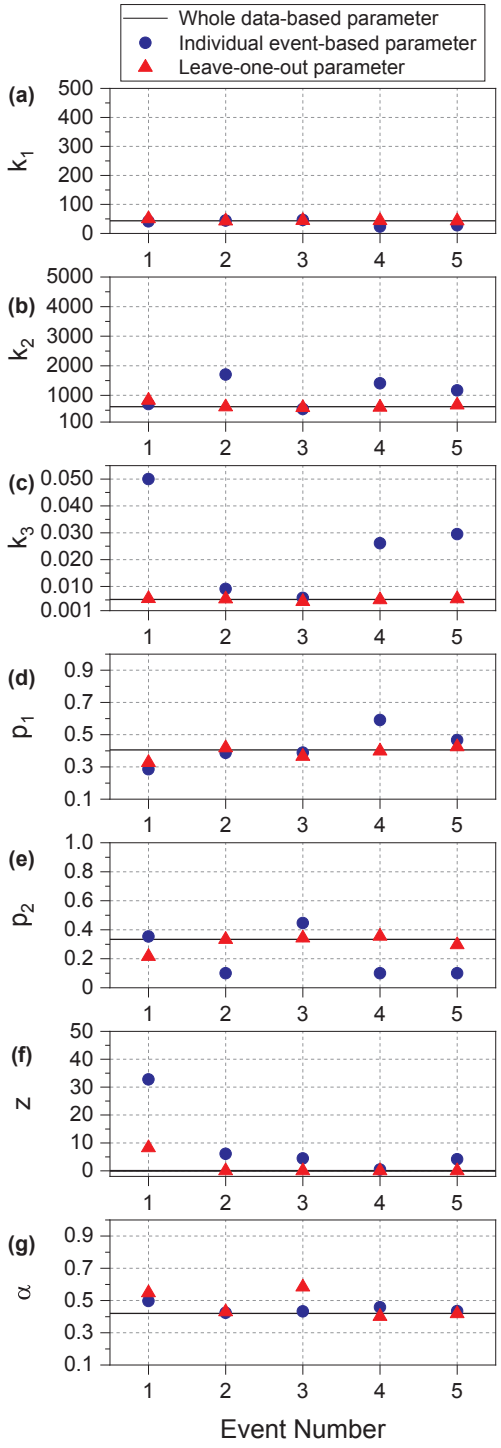

Fig. 2. The calibrated parameters of USF model from the selected data scenarios. The leave-one-out parameters at each event number represent the calibrated parameters from the flood events except that particular event.

is insignificant (Padiyedath et al., 2018b). The maximum storm drainage, $q_{R\ max}$ was estimated as 0.033 mm/min using Manning's equation.

4. Results and discussion

4.1. Model calibration and performance

The SCE-UA optimization method was applied for calibrating the USF model in the target watershed with RMSE as the objective function for the considered data scenarios discussed in Section 2.2. In the whole data-based scenario, all the selected events were considered for model calibration and a single set of parameters was derived from all the events. On the other hand, in the individual event-based scenario, each event was considered for model calibration. The leave-one-out scenario leaves out one flood event at a time and calibrating on the remaining events. The convergence of parameters was checked and the parameters were found to converge before the 50th generation in each SCE-UA application run. Thereafter, the best parameter set, $\hat{\theta}$, among the population at the 50th generation with a minimum RMSE value was used for the further simulations.

Fig. 2 shows the calibrated model parameters using the considered data scenarios in which the whole data-based parameters are represented by a black line, and the individual event-based and leave-one-out parameters are depicted by the blue circles and red triangles respectively. It is clear from Fig. 2(a) and (g) that the individual event-based parameters k_1 and α are identical in all the events and are similar to the whole data-based parameter values. Also, the parameters k_2 and p_1, as shown in Fig. 2(b) and (d), respectively, have similar values in the data scenarios even though they exhibit slight variations between events. However, the remaining individual event-based parameters k_3, p_2, and z are varying significantly between events and are similar to the whole data-based parameters only during certain events. The parameters k_3 and z represent the loss to the groundwater (Takasaki et al., 2009). The z values close to zero indicate a high rate of recession and higher z values represent a higher river flow at the outlet point instead of contributing to the groundwater. The higher value of z in event 1 can be attributed to its meteorological factor which is an intensive localized storm, as shown in Table 1. Parameter p_2 portrays the change in storage during the rising and recession limbs of a hydrograph based on the type of rainfall event (Hoshi and Yamaoka, 1982). The higher p_2 values exhibited during events 1 and 3, as shown in Fig. 2(e), may be possibly due to differences in meteorological factors from other events. The leave-one-out parameters at each event number represent the calibrated parameters from the flood events except that particular event. For example, the parameters at event number 1 in Fig. 2 show the calibrated parameters from all the flood events except event 1. The leave-one-out parameters exhibit close resemblance with the whole data-based parameters except for leaving out event 1.

The results confirm that the parameters k_3, p_2, and z will have a prominent effect on the estimation of discharge due to their high variability. Also, the meteorological factors can significantly affect parameter values during the model calibration. However, the considered data sets are not sufficient to generalize the above discussions despite providing a brief description of parameter uncertainty. Therefore, to have an elaborative idea of calibrated parameter uncertainty, the bootstrap approach was employed for generating samples which could be further utilized for conducting the uncertainty analysis.

Thereafter, the model performance on estimating discharge using the whole data-based and individual event-based parameters was analyzed. Table 3 presents the detailed performance evaluation using the statistical indicators of RMSE, Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), percentage error in peak (PEP), and percentage error in volume (PEV) (Padiyedath et al., 2018a). Additionally, Fig. 3 provides a visual representation of the hydrograph reproduced by the model for both the data scenarios. It is apparent from Table 3 that the model with individual event-based parameters has lower values of RMSE and higher values of NSE in all the events as compared to the whole data-based parameters. It is evident from the table that the PEP values estimated by the model using the individual event-based parameters were very low (close to zero) compared to that from the whole data-based parameters, even though the model exhibited a higher PEP value during event 1, and was not greater than 10% during any of the events. Fig. 3 also shows that the model was able to reproduce the peak discharge with utmost accuracy using the individual event-based parameters. On

Table 3
Performance evaluation of USF model using different statistical indicators of Root mean squared error (RMSE), Nash-Sutcliffe efficiency (NSE), percentage error in peak (PEP), and percentage error in volume (PEV).

Statistical index	RMSE (mm/min)		NSE (%)		PEP (%)		PEV (%)	
	Whole data	Individual event	Whole data	Individual event	Whole data	Individual event	Whole data	Individual event
Event 1	0.013	0.006	99.16	99.86	0.13	0.86	11.68	−0.8
Event 2	0.010	0.004	98.76	99.83	−6.18	−2.13	1.67	0.76
Event 3	0.010	0.009	98.41	98.62	−3.95	0.33	4.39	0.61
Event 4	0.012	0.006	92.91	98.25	17.63	6.18	−9.90	3.60
Event 5	0.014	0.005	91.14	98.88	16.72	8.09	−18.02	−0.20

the other hand, the model with whole data-based parameters exhibited remarkable anomaly in the reproduction of peak values, especially in events 4 and 5. Likewise the PEP, USF model shows the best performance in PEV values using the individual event-based parameters as shown in Table 3 which is close to zero as compared to that of the whole data-based parameters. The results further confirmed that the difference between the values of statistical indicators shown in Table 3 for the considered data scenarios is quite large and significantly greater for events 4 and 5. It can also be envisaged from Fig. 3 that the calibrated discharge using the individual event-based parameters by USF model nearly overlaps with the observed river discharge and reproduces the shape of the observed hydrograph with only slight variations. On the contrary, the model with whole data-based parameters deviated significantly while reproducing the shape of the hydrograph. Therefore, the results indicate that the individual event-based parameters can exactly reproduce the shape of the observed hydrograph as well as the peak discharge by significantly reducing the RMSE by 50% compared to that of the whole data-based parameters during the calibration. Hence, it is necessary to consider the calibration based on individual flood events for the parameter uncertainty analysis rather than considering the whole data-based calibration alone.

4.2. Model validation

The model validation on independent sub events was carried out to assess the operational performance of the USF model. Fig. 4(a1–a5) shows the reproduced hydrographs during the validation using the leave-one-out parameters as well as the individual event-based parameters. It is clear from Fig. 4(a1–a5) that the simulated discharge using the calibrated parameters of individual event-based scenario has highly deviated from the observed discharge, except event 3. On the other hand, simulated discharge using the parameters of leave-one-out scenario and event 3 was close to the observed discharge during validation. However, it is not easy to clearly portray the difference between the simulated discharge hydrographs of these cases from Fig. 4(a1–a5). Hence, we evaluated the performance further using the same performance evaluation criteria of RMSE, NSE, PEP, and PEV that used during the calibration and is shown in Fig. 4(b1–b5). The left y-axis of Fig. 4(b) represents RMSE and NSE, while the right y-axis depicts the PEP and PEV values. It can be envisaged from the figure that the parameters from event 3 and the leave-one-out scenario consistently generated the lowest RMSE and highest NSE for reproducing the hydrographs in validation. The remaining individual event-based parameters exhibited low performance in terms of RMSE and NSE during validation on the selected sub events. In the same way, the PEP and PEV values were also close to zero in the case of event 3 and the leave-one-out scenario parameters, and the rest of the cases exhibited varying values of PEV and PEV. The validation based on event 3 and leave-one-out scenario performed equally in terms of RMSE and other performance evaluation criteria, similar to that exhibited in calibration (Table 3).

The calibration of USF model over specific flood events and validation on independent flood events revealed that the model with leave-one-out parameters has consistent performance compared with the individual event-based parameters. This leave-one-out parameters represent the robust parameter set of the model and can be implemented for operational use in the context of flood simulation. However, there arises a question that which calibration approach is the most performant, subject to validation. Based on the above discussions, it is recommended that the model should be calibrated with all the available flood event data to get a more robust parameter set for the flood simulation. Vaze et al. (2010) studied the effect of calibration periods on model simulation and their study revealed that the model calibration on a portion of the record with conditions similar to those of the future period to simulate can provide a more reasonable set of parameters. This result was supported by de Vos et al. (2010), who suggested clustering time series according to climate similarities during calibration. Therefore, the USF model could be calibrated using the available flood events since USF is a flood event-based model for the flood simulation in urban watersheds with combined sewer system. These calibration techniques will provide a minimum standard for operational validation of the model. However, it needs more such calibrations at different sub-periods including more flood events to arrive at a conclusion. Moreover, Moore et al. (2007) suggest that in many situations it is hard for the lumped conceptual models to outperform in operational use for flood simulation. Therefore, it is recommended that further work is undertaken on alternative formulations which will describe the operational adequacy of the model.

4.3. Model parameter uncertainty analysis

The calibration parameter uncertainty analysis was conducted using the five available flood events under the whole data-based and individual event-based scenarios. The computed model residual series (Eq. (5)) was used to perform the resampling process for 1000 times by employing the residual-based bootstrap approach. Then, associated bootstrapped discharge series were generated by adding 1000 bootstrapped residual series to the calibrated discharge series as described in Section 2.3. These bootstrapped discharge series, $\hat{Q}^j(t)$, were used to obtain the j^{th} bootstrapped parameter vector, $\hat{\theta}^j$ of the USF model for both the whole data-based and individual event-based scenarios. Fig. 5 shows the scatter plots of the bootstrapped parameter vectors with their 95% CI in grey shading. The search range shown in Table 1 is illustrated as the left y-axis and the percentage contribution of 95% CI to the search range is depicted on the right y-axis of Fig. 5.

It is apparent from Fig. 5(a1)–(f1) that the k_1 values lie close to the lower limit of the search range and converged to a reduced range between 20 and 70 in all the cases. The 95% CI of parameter k_1 have a comparable width in all the cases and constitute about 1–5% of the search range, which is quite narrow. The scatter plot of parameter k_2 was narrow as well as close to the lower search range with values clustered around 500–2000, as shown in Fig. 5(a2)–(f2). Similarly, the 95% CI of k_2 was also narrow for the whole data-based parameter and was quite similar to the values of events 1 and 3. As compared to this pattern, the CI was relatively wide during the rest of the events. The parameter k_3 has a widespread pattern as compared to parameters k_1 and k_2, as shown in Fig. 5(a3)–(f3). The 95% CI of k_3 was wide and its

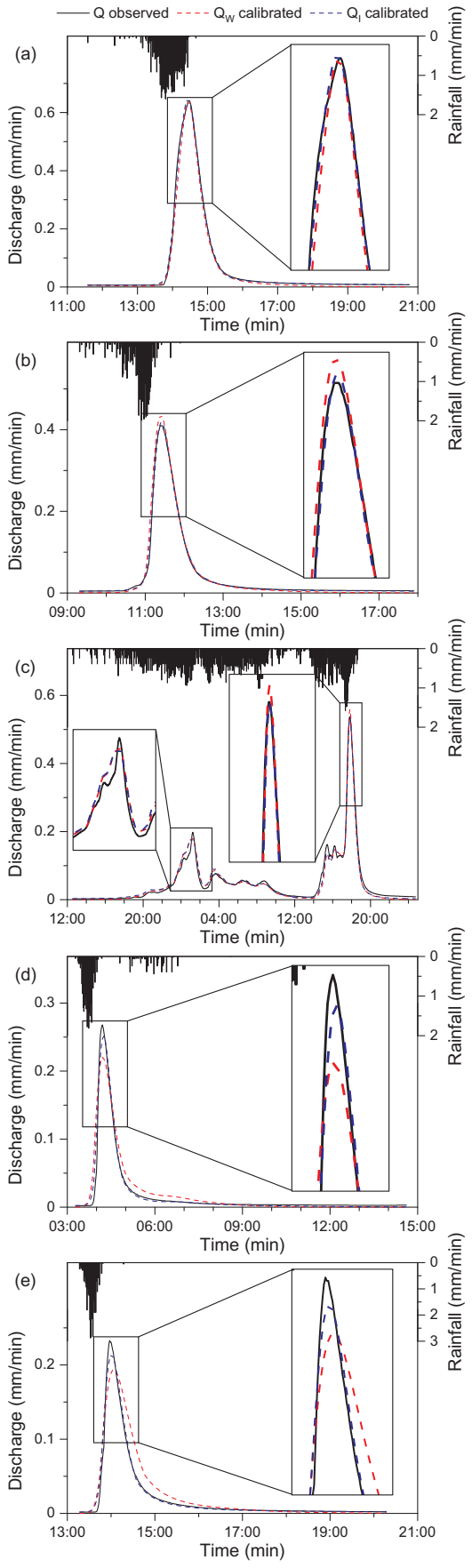

Fig. 3. The reproduced hydrographs by USF model for (a) Event 1; (b) Event 2; (c) Event 3; (d) Event 4; and (e) Event 5 using the whole data-based and individual event-based parameters. (Q_W and Q_I represent the calibrated discharge using the whole data-based and individual event-based parameters respectively).

contribution to the search range was high, except for event 3 and the whole data-based scenario. There is a well-spread pattern for parameter p_1 from 0.1 to 1 within the search range whereas most of the p_2 values accumulated near the lower search range between 0.1 and 0.7, as shown in Fig. 5(a4)–(f4) and (a5)–(f5), respectively. The 95% CI of these parameters are wide as compared to k_1 and k_2 with a high percentage contribution to the search range. It is evident from Fig. 5(a6)–(f6) that most of the z values are gathered near the lower limit of the search range and the 95% CI is relatively wide. Parameter α demonstrated a very widespread pattern with values close to the upper search range, as depicted in Fig. 5(a7)–(f7). The 95% CI of α was wide in all the scenarios and constituted about 25–50% of the search range. It can be envisaged from Fig. 5 that the scatter plot of whole data-based parameters is similar to the parameter pattern exhibited in event 3, which further indicate that the events with a large number of observations have domination on the whole data-based model calibration. Even though a wide search range was considered, certain parameters converged to a very narrow range. On the other hand, some parameters showed a widespread pattern from the lower to the upper limit of the search range. This could be a reflection of the equifinality concept (Beven and Freer, 2001) and reconsideration of this search range could enhance the performance of the model in cases where parameter values accumulate on the search range boundaries. Additionally, a larger number of bootstrap resamples can lead to a narrow 95% CI due to the convergence of parameters as observed in the study of Selle and Hannah (2010).

Fig. 6 shows the statistical representation of Fig. 5 in terms of the box-whisker plot of 1000 bootstrapped parameter vectors for both the data scenarios in which the bottom and the top lines of the boxes show the 25th and 75th percentiles, respectively and the line passing through the box represents the median ($\hat{\theta}_{50}$). The whiskers extend to the 2.5th and the 97.5th percentiles and the highest and lowest observations are plotted as asterisks. Additionally, Fig. 6 represents the mean ($\bar{\theta}$ as the blue square) and CV (values at the top) of the bootstrapped parameter vectors along with the calibrated parameter values ($\hat{\theta}$ as the red circle). Detailed descriptions of the plot are given in the figure caption. It is clear from Fig. 6(a) and (b) that the bootstrap estimates of $\hat{\theta}_{50}$ and $\bar{\theta}$ of parameters k_1 and k_2 have similar values and are very close to the calibrated $\hat{\theta}$ values in all the events. However, the parameter k_3 exhibited differences in the $\bar{\theta}$ values during event 1, as depicted in Fig. 6(c). It is apparent from Fig. 6(d) that the $\hat{\theta}_{50}$ and $\hat{\theta}$ values of p_1 are similar except for event 4 whereas the $\bar{\theta}$ and $\hat{\theta}$ values are identical except for events 4 and 5. On the other hand, the $\hat{\theta}_{50}$, $\bar{\theta}$, and $\hat{\theta}$ values of parameter p_2 were close enough except for event 5, as shown in Fig. 6(e). Similar to the parameter p_2, the $\hat{\theta}_{50}$, $\bar{\theta}$, and $\hat{\theta}$ values of parameter z in Fig. 6(f) were identical, except in event 1. Subsequently, the $\bar{\theta}$ values of α exhibited minor discrepancy only in event 5 as compared to the $\hat{\theta}_{50}$ and $\hat{\theta}$ values, as illustrated in Fig. 6(g). Overall, the $\hat{\theta}$ values were in accordance with the $\hat{\theta}_{50}$ and $\bar{\theta}$ values in most of the cases even though the $\bar{\theta}$ values show minor deviations. The CV values also varied between scenarios, as shown in Fig. 6. During the whole data-based analysis, the highest CV value of around 133% was observed for parameter z and the least CV value was noted for parameter k_2. Further, the parameters were ordered based on their CV values as follows: $z > \alpha > k_3 > p_1 > p_2 > k_1 > k_2$. However, for the individual event-based analysis, during event 1, the highest CV value was exhibited by the parameter k_3 which was followed by parameter z. The order of parameters based on CV values are $k_3 > z > p_1 > \alpha > p_2 > k_2 > k_1$. The same order for event 2 was $p_2 > k_3 > z > p_1 > k_2 > \alpha > k_1$. During events 3 and 4, based on the CV

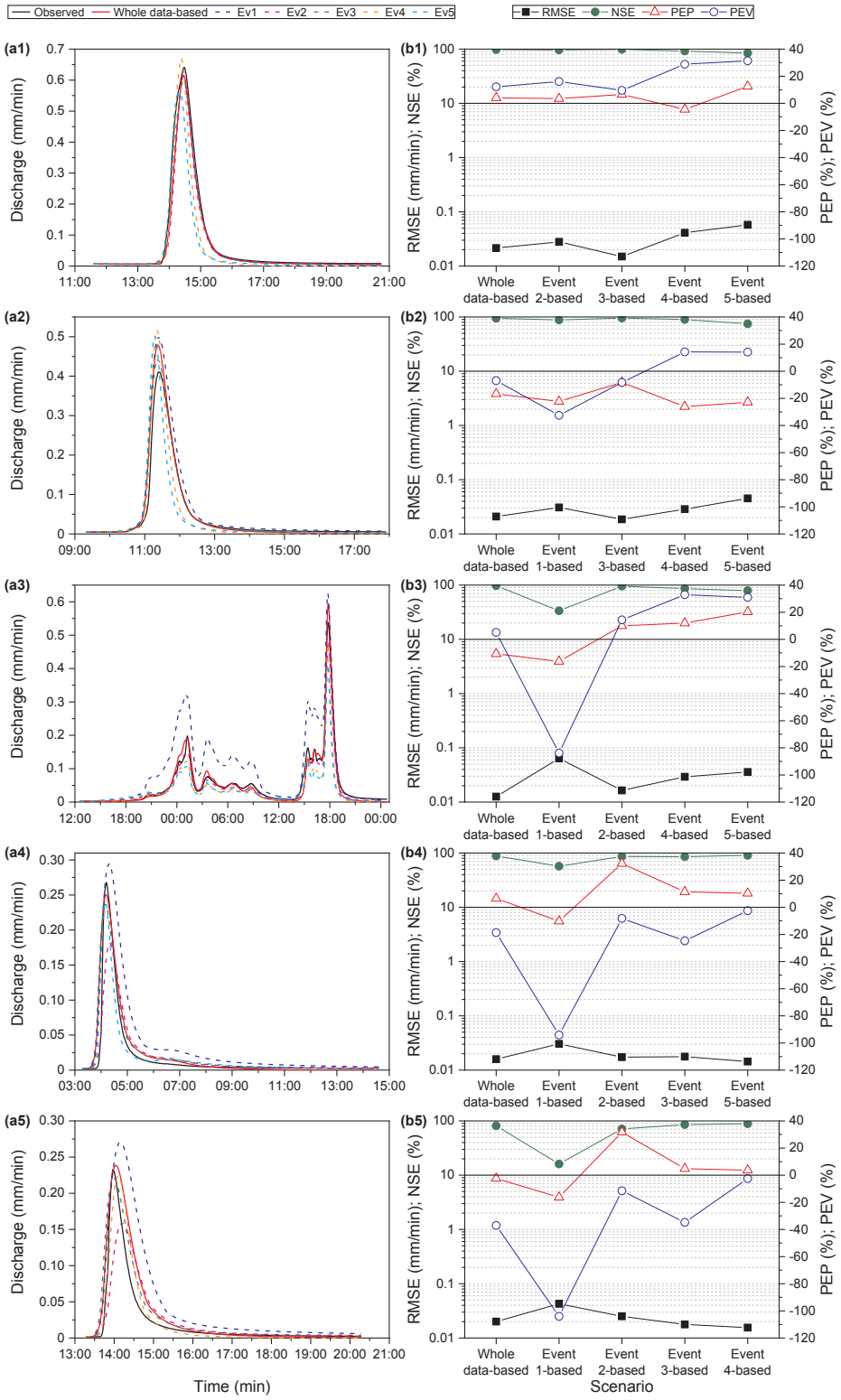

Fig. 4. The USF model validation for selected sub events of (i) Event 1 (a1–b1); (ii) Event 2 (a2–b2); (iii) Event 3 (a3–b3); (iv) Event 4 (a4–b4); and (v) Event 5 (a5–b5) using the leave-one-out and individual event-based parameters. The plots (a1–a5) and (b1–b5) show the reproduced hydrographs and performance evaluation respectively in each scenario (Ev represents event).

values, the parameter z had the highest uncertainty. On the contrary, parameters p_2 and z had very high CV values as compared to other parameters during event 5. Therefore, the results reveal that the parameter with the highest and lowest uncertainty varies from case to case. However, z and k_1 were the parameters with the highest and lowest uncertainty respectively based on their CV values in most of the cases.

Selle and Hannah (2010) identified the parameter with the highest uncertainty of a conceptual salt load model using the CV value as the only index. However, in this study, two proposed indices were computed, in addition to the CV, for objectively assessing the parameter

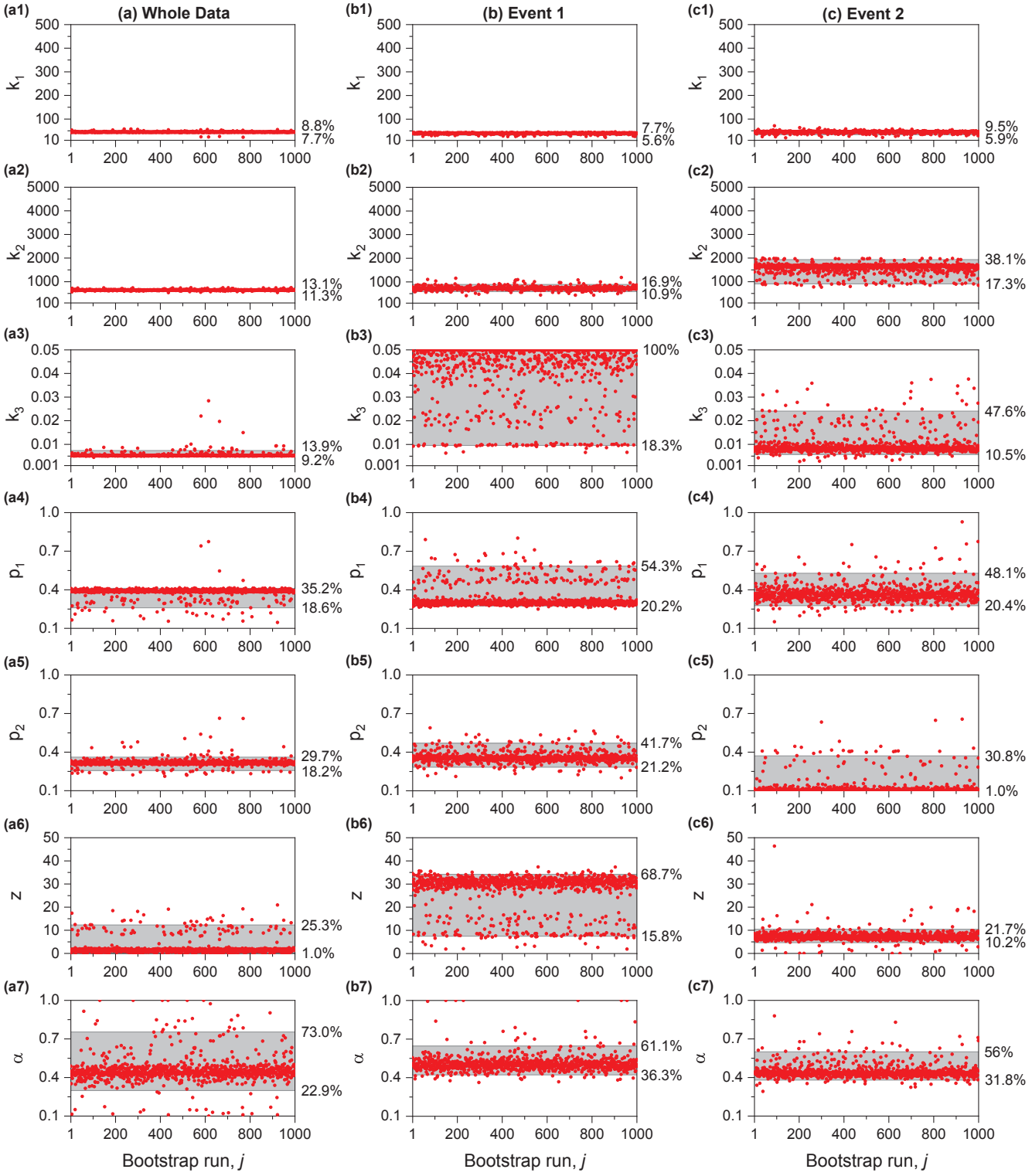

Fig. 5. Scatter plots of bootstrapped parameter samples by the USF model for (a) whole data (a1–a7), (b) event 1 (b1–b7), (c) event 2 (c1–c7), (d) event 3 (d1–d7), (e) event 4 (e1–e7), and (f) event 5 (f1–f7) with their 95% CI in grey shading. The percentage contribution of CI to the search range is depicted in the right y-axis.

uncertainty in order to have a clear understanding of the 95% CI as well as the median values of the parameters, as shown in Fig. 7. In this figure, the PUI_1 and PUI_2 index values of parameter z for the whole data-based scenario is represented at the figure boundaries with their values. Since the $\hat{\theta}$ value of parameter z was very close to zero in the whole data-based scenario, as shown in Fig. 2(f), the calculated PUI_1 and PUI_2 values of z were very high and represented at the boundaries.

It is evident from Fig. 7(a) that the parameter z has the highest value of PUI_1 during events 3 and 4 and also in the whole data-based case, which further indicates that z had the highest uncertainty in most of the considered scenarios based on PUI_1. This higher PUI_1 value of z can be interpreted as to its wide 95% CI and calibrated parameter values close to zero during these scenarios. However, the parameter p_2 demonstrated a higher value of PUI_1 during events 2 and 5, which could also

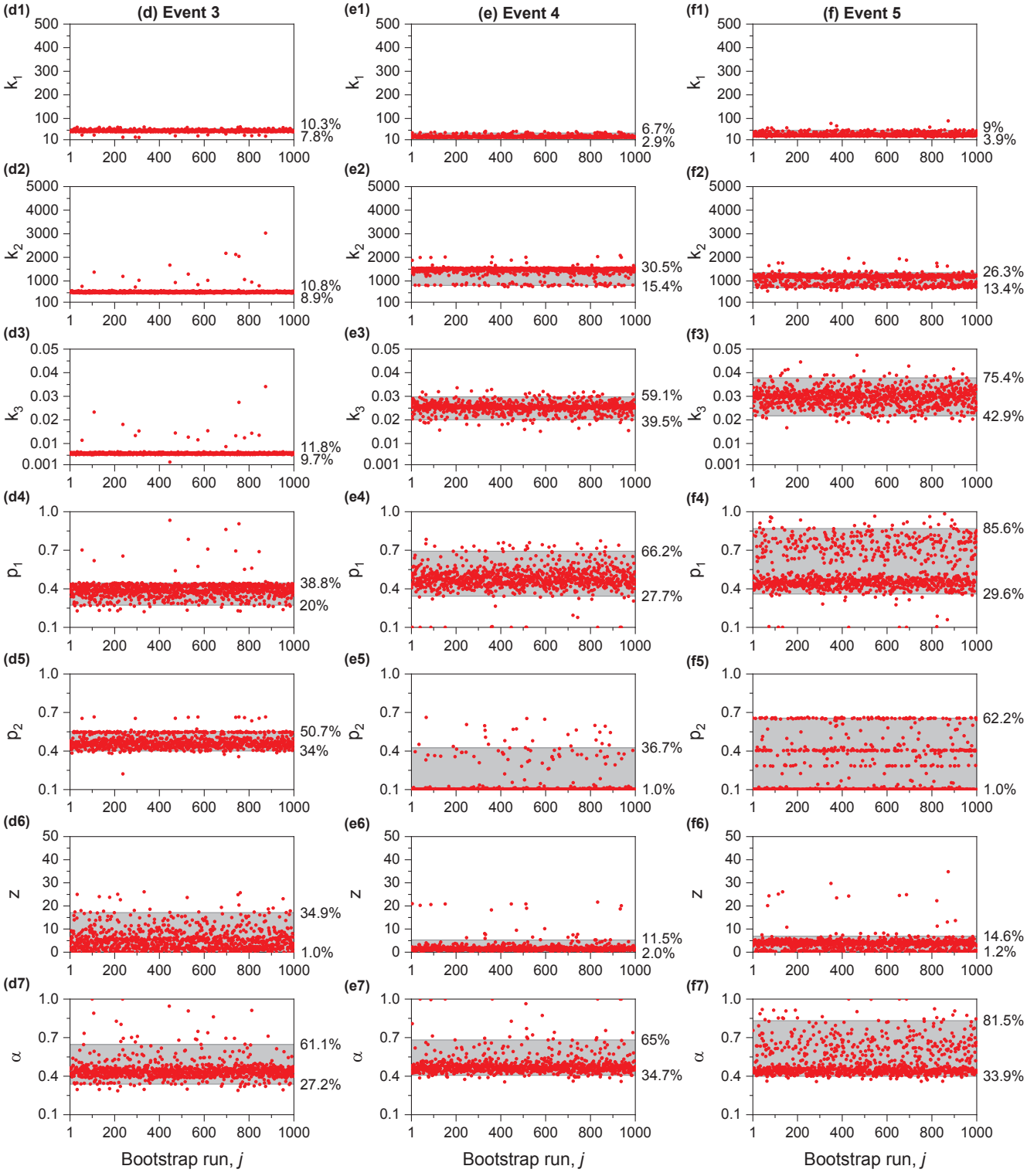

Fig. 5. (*continued*)

be attributed to its very wide 95% CI as compared to other parameters, as shown in Fig. 5(c5) and (f5) while parameter p_1 has got the highest PUI_1 value during event 1. The parameters k_2 and k_1 demonstrated the lowest uncertainty from the cases whose 95% CI was very narrow as compared to the rest of the parameters. The remaining parameters showed a different order of PUI_1 values in the considered scenarios, based on their 95% CI width and the calibrated parameter values. The high uncertainty, in terms of PUI_2, was exhibited by the parameter z in all the cases except for event 3, as illustrated in Fig. 7(b). During event 3, all the parameters exhibited a small magnitude of uncertainty and were comparable to each other, even though the PUI_1 portrayed z as the most uncertain parameter in event 3, which can be further interpreted as the very similar median and calibrated values of parameters. Also, for a narrow CI, the median and calibrated parameter values will come closer and further the PUI_2 values will approach zero. The parameter z exhibited negative PUI_2 values with a high magnitude in the whole data-based scenario and event 4, which was significantly high in the whole data-based scenario. These high magnitude negative PUI_2 values

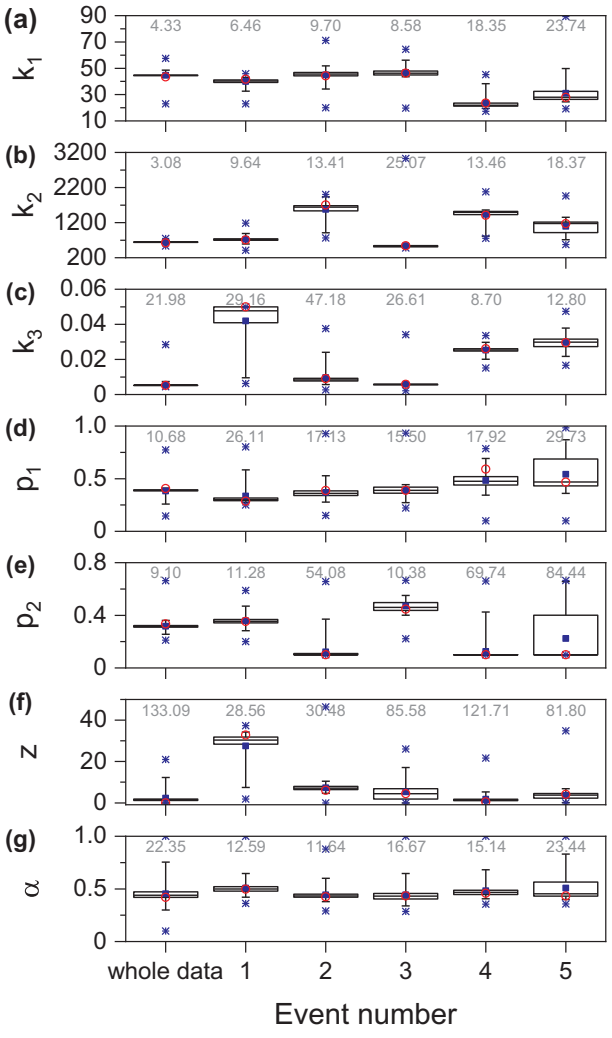

Fig. 6. The box plot of bootstrapped parameter vector (a) k_1, (b) k_2, (c) k_3, (d) p_1, (e) p_2, (f) z, and (g) α for both the data scenarios. The bottom and the top lines of the boxes show the 25th and 75th percentiles respectively and the line passing through the box represents the median ($\hat{\theta}_{50}$). The whiskers extend to the 2.5th and 97.5th percentiles and the highest and lowest observations are plotted as asterisks. The red circle and blue square within the box indicate calibrated parameter ($\hat{\theta}$) and mean ($\bar{\theta}$) respectively. The written values represent the CV for each data scenario. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

indicate that the bootstrap median values are much higher than the calibrated parameters. The results revealed that the difference in both the calculated indices, PUI_1 and PUI_2, is because both represent different aspects of uncertainty analysis.

The parameter z represents the infiltration hole height in the USF model (Takasaki et al., 2009), which depends upon parameters like basin storage and rainfall intensity, and will vary greatly from event to event with the highest uncertainty. Additionally, the optimum value of z estimated from the model calibration is close to zero with a very wide range of bootstrapped parameter values, as shown in Fig. 5(a6)–(f6) which also makes it the most uncertain parameter. The parameter with higher uncertainty after z varied from case to case. However, based on the PUI_1 and PUI_2 values, p_1, p_2, α, k_3, and k_2 had higher uncertainty values after parameter z for most of the cases. The 95% CI of these parameters are relatively wide which can be transformed into higher index values. The parameter k_1 exhibited the least uncertainty. One of the reasons for this could be the fact that parameter k_1 describes features of the watershed (Sugiyama et al., 1997) and there is a very low chance of a change in watershed features within a short span. This further indicated that parameter k_1 remains reasonably stable under varying input data scenarios. Besides, the equifinality concept can also derive parameter uncertainty by generating non-unique parameter sets during the calibration process and there will be a lot of different parameter combinations that lead to multiple optimal solutions (Beven and Freer, 2001; Yang et al., 2008). However, this parameter uncertainty can be overcome to some extent by using global searching techniques during calibration.

4.4. Model simulation uncertainty

The uncertainty in model simulation due to the calibrated parameter uncertainty was estimated by computing the 95% CI of the 1000 simulated discharge series generated by the bootstrapped parameters from the whole data-based and the individual event-based scenarios. Fig. 8 shows the 95% CI of simulated discharge (uncertainty range) for each event from both the whole data-based and the individual event-based scenarios during calibration. It is desirable to have a narrow range and Fig. 8(a1–b1) shows that the uncertainty range from the whole data-based parameters at the peak flows of event 1 is wider than the range simulated from the individual event-based parameters. However, the width of the uncertainty range at low flows was almost identical for event 1 from both the scenarios. Fig. 8(a2–b2) shows that the uncertainty range was unable to capture the observed values during the flood peak of the whole data-based scenario, whereas the uncertainty range of individual event-based scenario was able to bracket a large amount of the observed values, including the flood peaks. The uncertainty range of the whole data-based scenario illustrated in Fig. 8(a3–b3) included the highest flood peak value with a wide uncertainty range. On the other hand, the uncertainty range of the

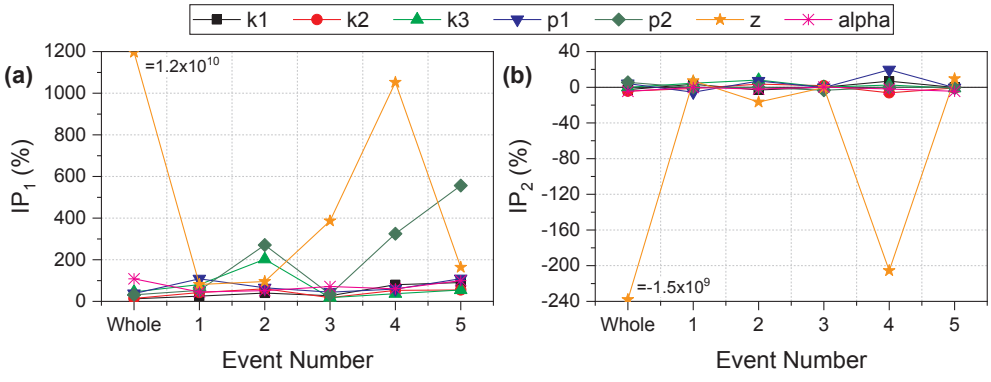

Fig. 7. Two proposed indices of (a) PUI_1, and (b) PUI_2 for analyzing the model parameter uncertainty.

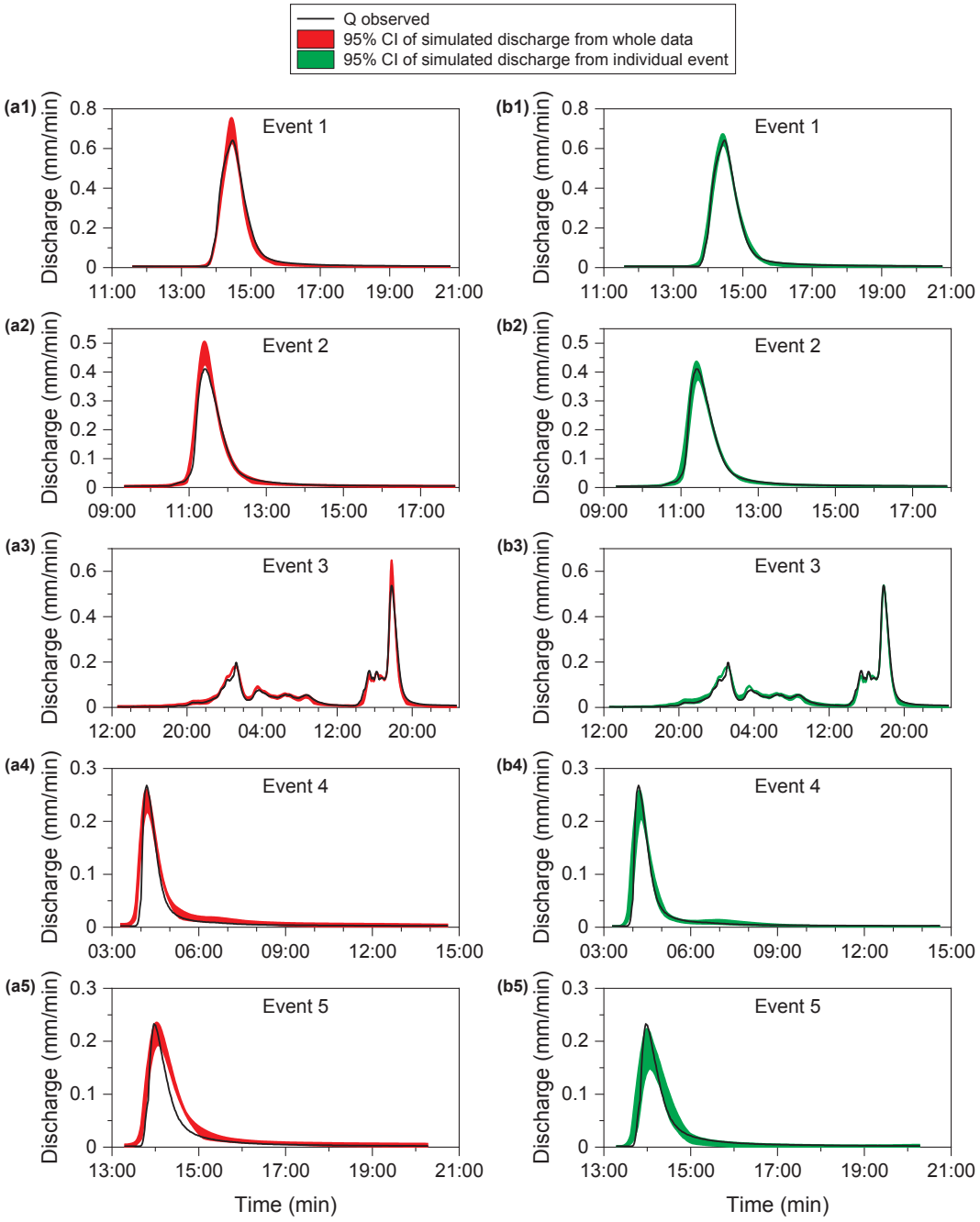

Fig. 8. The simulation uncertainty of USF model for the selected events using the whole data-based (a1-a5) and individual event-based (b1-b5) scenarios.

individual event-based scenario was very narrow and was close to the observed peak flows and hence was not able to capture the flood peaks. During events 4 and 5 from the whole data-based parameters, most of the flood peak values were falling inside the uncertainty range even though the low flows were not well captured as shown in Fig. 8(a4) and (a5). Concurrently, the uncertainty range of the individual event based scenario was able to capture almost all the flows, except the peak value during events 4 and 5 as shown in Fig. 8(b4) and (b5). Overall, the whole data-based scenario captured the observed discharge with a very wide uncertainty range whereas the individual event-based scenario bracketed observations within a very narrow uncertainty range during calibration. Also, as can be seen from Fig. 8, that the uncertainty range is very narrow at the low flows for both the scenarios. Hence, it can be concluded that the model simulates peak discharge with higher uncertainty as compared to low flows. High uncertainty in the model simulation during flood peaks can be attributed to the influence of low flows as they may have dominated the parameter estimation process due to their greater numbers as compared to the peak flow (Gallagher and Doherty, 2007). However, it is essential to estimate flood peaks with lesser uncertainty as compared to the low flows due to the high-risk factor associated with them. It is possible to do so by calibrating the model parameters using a specific objective function which can initiate the type of simulation that the model is required to make.

In order to portray the differences in the model simulation uncertainty during calibration, a detailed uncertainty analysis was further conducted using the P-factor and the two proposed model simulation uncertainty indices, SUI_1 and SUI_2. Fig. 9 shows the model simulation uncertainty for the whole data-based and the individual event-based scenarios using the P-factor, SUI_1, and SUI_2. The P-factor value, as defined by Eq. (8), close to 100% represents the capability of the model to

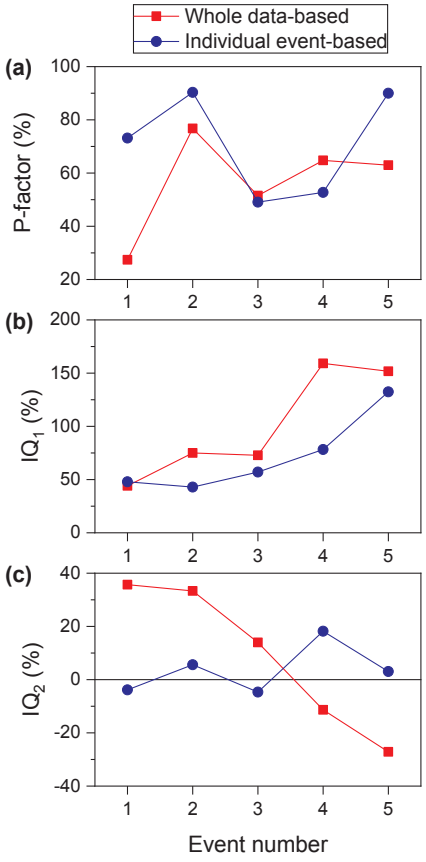

Fig. 9. The model simulation uncertainty analysis using the (a) P-factor, and the two proposed model simulation uncertainty indices of (b) SUI_1, and (c) SUI_2 for the whole data-based and individual event-based scenarios.

reasonably capture almost all the observed discharge values within the uncertainty range. Similarly, the values of SUI_1 and SUI_2 close to zero indicate low uncertainty of the model in simulating the discharge. It is clear from Fig. 9(a) that the obtained value of P-factor is higher for events 1, 2, and 5 in the individual event-based scenario and the whole data-based scenario showed higher values of P-factor during events 3 and 4. Even though the P-factor of individual event-based analysis is lower than the whole data-based analysis in events 3 and 4, the difference between the values is quite small. The individual event-based scenario captured an almost equal number of observations with a very narrow uncertainty range, especially at the peak flows, in event 3 compared with the whole data-based scenario that resulted in an almost equal P-factor. The high uncertainty of parameter z during event 4 in terms of parameter uncertainty indices, as shown in Fig. 7, could be a reason for the low P-factor exhibited by the individual event-based scenario in event 4. Subsequently, values of the proposed index SUI_1, derived from the individual event-based analysis were less than the values derived from whole data-based analysis in all the events except for event 1, as shown in Fig. 9(b). During event 1, the SUI_1 value of both the scenarios was similar and the whole data-based scenario received the least value. The results revealed that the width of the uncertainty range is narrow relative to the observed discharge values in the individual event-based analysis as compared to the whole data-based analysis. The individual event-based values of SUI_2 were also close to zero in all the events except for event 4, as illustrated in Fig. 9(c). The SUI_2 values were negative for events 4 and 5 in the whole data-based scenario, which indicates that the observed discharge values were lesser than the median values and lied in the lower confidence region. In the same way, values from the individual event-based analysis were mostly positive, except for events 1 and 3, and were found in the upper

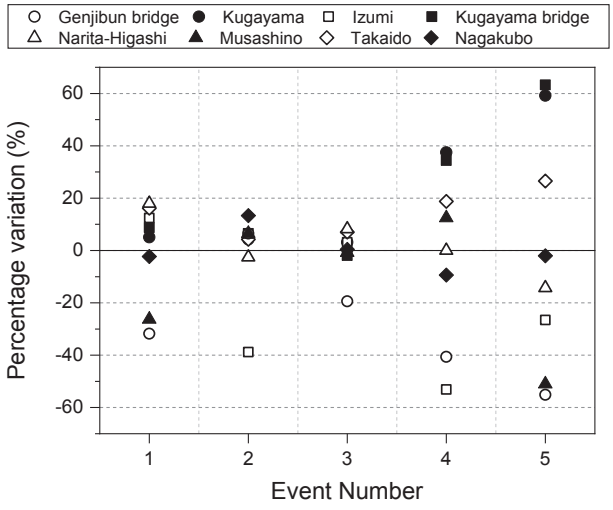

Fig. 10. The percentage variation of total rainfall obtained from each rain gauge with respect to the mean rainfall from all the gauges.

confidence region. Overall, considering all the indices, the simulation uncertainty was lower during the individual event-based analysis as compared to the whole data-based analysis.

4.5. Rainfall spatial variability

The results revealed that the model simulation uncertainty varies from event to event as well as from the considered data scenarios. This can be ascribed to the difference in parameter values in each event and the whole data-based scenario resulted from the spatial variability in rainfall. Therefore, further analysis was carried out to have a clear understanding of the extent of spatial variability of rainfall in the watershed. Fig. 10 shows the percentage variation of total rainfall obtained from each rain gauge with respect to the mean rainfall from all the gauges. It is clear from Fig. 10 that during event 1, the variation of two gauges are around −30%, whereas the same in event 2 for one gauge is about −40%. These high percentage variation values of rainfall exhibited by several rain gauges indicate that there is a relatively high spatial variability in rainfall during events 1 and 2. However, all the gauges showed relatively low variability except one gauge during event 3 which resulted in an overall low spatial variability in this event. Moving to events 4 and 5, almost all the gauges portrayed high percentage variation values ranging from −60% to 60% and exhibited the highest spatial variability. This high spatial variability exhibited by the rain gauges in each event can be attributed to their completely different rainfall pattern and this could be a cause of the high uncertainty in simulations.

During the whole data-based scenario, the model parameters were averaged spatially as well as temporally over the watershed without considering the spatial variability of rainfall as well as the meteorological factors that caused the rainfall events, and the estimated model parameters will be different from the true watershed parameters (Chaubey et al., 1999). This could be a possible reason for the high simulation uncertainty exhibited by the model in the whole data-based scenario. However, during the individual event-based scenario, the catchment properties are only spatially averaged, not temporally. This will lead to a reduced simulation uncertainty in the event-based scenario compared with the whole data-based scenario. At the same time, lumping up of the complex, spatially varying catchment properties such as rainfall, inflow, etc. in a model will induce considerable errors associated with the spatially averaged input data (Cooley, 2004) and further affect the model simulation uncertainty in both the data scenarios. Notwithstanding the problems associated with the spatial averaging of the watershed processes, the USF model was able to

simulate the discharge with reasonable accuracy in the individual event-based and the whole data-based scenarios using the bootstrap approach associated with the SCE-UA method. Therefore, the bootstrap approach will contribute to the development of parsimonious hydrologic models (Selle and Hannah, 2010) with reliable estimates of parameter uncertainty. In this study, we found that the range of simulation uncertainty due to the parameters is relatively small. Apart from parameter uncertainty, input data measurement errors from all sources and model structure errors also cause model simulation uncertainty (Sivakumar and Berndtsson, 2010). However, it is not practicable to define the extent to which the other sources of uncertainties will affect the model simulation uncertainty based on our present study.

5. Conclusions

The residual-based bootstrap technique was utilized to analyze the calibration parameter uncertainty of the USF model to assess its impact on the model simulation in the upper Kanda River basin, an urban watershed in Tokyo. The bootstrap approach was applied to the individual flood events for the first time due to the relevance of flood-runoff analysis in urban watersheds along with the available whole data in order to demonstrate the impact of different available data scenarios on the calibration uncertainty behavior of the parameters. The 95% CI of certain parameters converged to a very narrow range as compared to the search range while some parameters showed a wider confidence range from the lower to the upper limit of the search range in both the scenarios. Further, the parameter uncertainty was scrupulously analyzed and the parameters with the highest and the lowest uncertainty were identified by utilizing two newly proposed indices that are based on the width of the confidence interval and the median value. Although the order of the model parameters assigned during the uncertainty analysis based on the proposed indices differs, the significant parameters yielded was the same. Hence, the proposed indices could be useful in future studies to derive the parameters from the highest to the lowest uncertainty. Additionally, the effect of parameter uncertainty on the model simulation was investigated by computing the 95% CI of 1000 simulated discharge series generated from the bootstrapped parameters, and by utilizing the P-factor and two other proposed indices for assessing the model simulation uncertainty. The results revealed that the simulation uncertainty is low in the individual event-based analysis compared with the whole data-based analysis based on all the considered indices. Also, the uncertainty range was wider at the peak flows and hence the model simulated peak discharge values had higher uncertainty than low flows. As a conclusion, the parameter uncertainty and its effect on model simulation uncertainty were successfully evaluated and the characteristics of an urban specific rainfall-runoff model (USF model) were explained in detail using the bootstrap approach.

The residual-based bootstrap approach used in this study assumed an independent and identically distributed model residual series. However, this assumption does not seem to be appropriate after looking at the autocorrelation plot of residuals. The residuals demonstrated a short-term persistence although the residual autocorrelation function decayed exponentially and was statistically insignificant beyond a lag of 6 min. One solution to tackle this issue is the use of block bootstrap resampling of the residual with sufficiently long blocks to preserve the time dependence of the residuals. Another possibility is the use of the autoregressive component to reconstruct auto correlated bootstrap residuals. Selle and Hannah (2010) have already carried out such an improvement and obtained comparable results from both bootstrap approaches. It is also important to conduct comparative studies between the bootstrap and other techniques for parameter uncertainty analysis which will reveal the strength and weaknesses of the bootstrap approach. The violation of the assumptions made during uncertainty analysis by the bootstrap method may lead to the inadequate characterization of simulation uncertainty. Therefore, it is desirable to validate the simulation uncertainty using observations which will further provide some empirical evidence that the bootstrap method provides a good estimation of parameter and simulation uncertainties. Further, the parameter variance obtained from this uncertainty analysis can be used in the data assimilation approaches for the real-time prediction of flood and will give confidence to the hydrologist who uses the model in an operational context. However, this study primarily focused on the residual-based bootstrap approach for the calibration parameter uncertainty analysis and its subsequent effect on model simulation uncertainty. We will carry out further research on the aforementioned areas by considering the residual correlation and heteroscedasticity.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This study was carried out as a part of the research project entitled “Study on guerrilla rainstorm, flood, and water pollution in megacity urban watersheds – Countermeasures against megacity urban water-related disasters bipolarized by climate change” supported by Tokyo Metropolitan Government, Japan (Represented by Prof. Akira Kawamura).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2019.124195.

References

Abbaspour, K.C., Sonnleitner, M., Schulin, R., 1999. Uncertainty in estimation of soil hydraulic parameters by inverse modeling: example lysimeter experiments. Soil Sci. Soc. Am. J. 63, 501–509.

Abbaspour, K.C., Van Genuchten, M.T., Schulin, R., Schläppi, E., 1997. A sequential uncertainty domain inverse procedure for estimating subsurface flow and transport parameters. Water Resour. Res. 33, 1879–1892. https://doi.org/10.1029/97WR01230.

Ajami, N.K., Gupta, H., Wagener, T., Sorooshian, S., 2004. Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. J. Hydrol. 298, 112–135. https://doi.org/10.1016/j.jhydrol.2004.03.033.

Amaguchi, H., Kawamura, A., Olsson, J., Takasaki, T., 2012. Development and testing of a distributed urban storm runoff event model with a vector-based catchment delineation. J. Hydrol. 420–421, 205–215. https://doi.org/10.1016/j.jhydrol.2011.12.003.

Ando, Y., Takahasi, Y., 1997. Recent flood control measures for urban rivers in japan: case study of the Kanda river in Tokyo. Water Int. 22, 245–251. https://doi.org/10.1080/02508069708686714.

Bellprat, O., 2013. Parameter uncertainty and calibration of regional climate models. Diss. ETH No. 21147, ETH Zurich, Zurich, Switzerland. p. 144. doi: 10.3929/ethz-a-009915303.

Beven, K., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. Adv. Water Resour. 16, 41–51. https://doi.org/10.1016/0309-1708(93)90028-E.

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process. 6, 279–298. https://doi.org/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 (1–4), 11–29. https://doi.org/10.1016/S0022-1694(01)00421-8.

Bornstein, R., Lin, Q., 2000. Urban heat islands and summertime convective thunderstorms in Atlanta: three case studies. Atmos. Environ. 34 (3), 507–516. https://doi.org/10.1016/S1352-2310(99)00374-X.

Brigode, P., Bernardara, P., Paquet, E., Gailhard, J., Garavaglia, F., Merz, R., Mićović, Z., Lawrence, D., Ribstein, P., 2014. Sensitivity analysis of SCHADEX extreme flood estimations to observed hydrometeorological variability. Water Resour. Res. 50, 353–370. https://doi.org/10.1002/2013wr013687.

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. https://doi.org/10.1016/j.jhydrol.2012.11.012.

Brigode, P., Paquet, E., Bernardara, P., Gailhard, J., Garavaglia, F., Ribstein, P., Bourgin, F., Perrin, C., Andréassian, V., 2015. Dependence of model-based extreme flood estimation on the calibration period: case study of the Kamp River (Austria). Hydrol.

Sci. J. 60, 1424–1437. https://doi.org/10.1080/02626667.2015.1006632.
Chaubey, I., Haan, C.T., Grunwald, S., Salisbury, J.M., 1999. Uncertainty in the model parameters due to spatial variability of rainfall. J. Hydrol. 220, 48–61. https://doi.org/10.1016/S0022-1694(99)00063-3.
Cooley, R.L., 2004. A Theory for Modeling Ground-Water Flow in Heterogeneous Media. USGS Professional Paper 1679. U.S. Department of the Interior. U.S. Geological Survey.
Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M., Hendrickx, F., 2012. Crash testing hydrological models in contrasted climate conditions: an experiment on 216 Australian catchments. Water Resour. Res. https://doi.org/10.1029/2011WR011721.
Davison, A.C., Hinkley, D.V., 1997. Bootstrap Methods and their Application. Cambridge University Press.
De Vos, N.J., Rientjes, T.H.M., Gupta, H.V., 2010. Diagnostic evaluation of conceptual rainfall-runoff models using temporal clustering. Hydrol. Process. 24, 2840–2850. https://doi.org/10.1002/hyp.7698.
Dixon, P.M., 2006. Bootstrap resampling. In: Encyclopedia of Environmetrics. John Wiley & Sons, Ltd, Chichester, UK. https://doi.org/10.1002/9780470057339.vab028.
Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76, 501–521. https://doi.org/10.1007/BF00939380.
Duan, Q.Y., Sorooshian, S., Gupta, V.K., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28, 1015–1031. https://doi.org/10.1029/91WR02985.
Ebtehaj, M., Moradkhani, H., Gupta, H.V., 2010. Improving robustness of hydrologic parameter estimation by the use of moving block bootstrap resampling. Water Resour. Res. 46. https://doi.org/10.1029/2009WR007981.
Efron, B., 1979. Bootstrap methods: another look at the jackknife. Ann. Stat. 7, 1–26. https://doi.org/10.1214/aos/1176344552.
Efron, B., 1982. The Jackknife, the Bootstrap and Other Resampling Plans. CBMS-NSF regional conference series in applied mathematics. Philadelphia. doi: 10.1137/1.9781611970319.
Efron, B.J., Tibshirani, R., 1993. An Introduction to the Bootstrap. Chapman and Hall, London.
Freer, J., Beven, K., Ambroise, B., 1996. Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. Water Resour. Res. https://doi.org/10.1029/95WR03723.
Gallagher, M., Doherty, J., 2007. Parameter estimation and uncertainty analysis for a watershed model. Environ. Model. Softw. 22, 1000–1020. https://doi.org/10.1016/J.ENVSOFT.2006.06.007.
Graf, W.L., 1977. Network characteristics in suburbanizing streams. Water Resour. Res. 13, 459–463. https://doi.org/10.1029/WR013i002p00459.
Green, C.H., Van Griensven, A., 2008. Autocalibration in hydrologic modeling: using SWAT2005 in small-scale watersheds. Environ. Model. Softw. 23, 422–434. https://doi.org/10.1016/j.envsoft.2007.06.002.
Hirsch, R.M., Archfield, S.A., De Cicco, L.A., 2015. A bootstrap method for estimating uncertainty of water quality trends. Environ. Model. Softw. 73, 148–166. https://doi.org/10.1016/j.envsoft.2015.07.017.
Hollis, G.E., 1975. The effect of urbanization on floods of different recurrence interval. Water Resour. Res. 11, 431–435. https://doi.org/10.1029/WR011i003p00431.
Hornberger, G.M., Spear, R.C., 1981. An approach to the preliminary-analysis of environmental systems. J. Environ. Manage. 12 (1), 7–18.
Hoshi, K., Yamaoka, H., 1982. A relationship between kinematic wave and storage routing models. Proc. 26th Japanese. Conf. Hydraul. JSCE 273–278. https://doi.org/10.2208/prohe1975.26.273.
Jakeman, A.J., Chen, T.H., Post, D.A., Hornberger, G.M., Littlewood, I.G., Whitehead, P.G., 1993. Assessing uncertainties in hydrological response to climate at large scale. Macroscale model. Hydrosphere 214, 37–47.
Jeong, D., Kim, Y., 2005. Rainfall-runoff models using artificial neural networks for ensemble streamflow prediction. Hydrol. Process. 19, 3819–3835. https://doi.org/10.1002/hyp.5983.
Jeremiah, E., Sisson, S.A., Sharma, A., Marshall, L., 2012. Efficient hydrological model parameter optimization with Sequential Monte Carlo sampling. Environ. Model. Softw. 38, 283–295. https://doi.org/10.1016/j.envsoft.2012.07.001.
Julian, J.P., Gardner, R.H., 2014. Land cover effects on runoff patterns in eastern piedmont (USA) watersheds. Hydrol. Process. 28, 1525–1538. https://doi.org/10.1002/hyp.9692.
Kawamura, A., 2018. Status quo and perspectives of flood runoff analysis for urban watersheds. J. Jpn Soc. Hydrol. Water Resour. 31 (6), 451–466.
Kimura, T., 1961. The Flood Runoff Analysis Method by the Storage Function Model. The Public Works Research Institute, Ministry of Construction.
Kjeldsen, T.R., 2010. Modelling the impact of urbanization on flood frequency relationships in the UK. Hydrol. Res. 41, 391–405. https://doi.org/10.2166/nh.2010.056.
KlemeŠ, V., 1986. Operational testing of hydrological simulation models. Hydrol. Sci. J. 31, 13–24. https://doi.org/10.1080/02626668609491024.
Koga, T., Kawamura, A., Amaguchi, H., Tanouchi, H., 2016. Assessing impervious area ratios of grid-based land-use classifications on the example of an urban watershed. Hydrol. Sci. J. 61, 1728–1739. https://doi.org/10.1080/02626667.2015.1133909.
Kuczera, G., 1988. On the validity of first-order prediction limits for conceptual hydrologic models. J. Hydrol. 103, 229–247.
Künsch, H.R., 1989. The jackknife and the bootstrap for general stationary observations. Ann. Stat. 17, 1217–1241.
Lahiri, S.N., 2003. Resampling Methods for Dependent Data. Springer-Verlag, New York.
Lall, U., Sharma, A., 1996. A nearest neighbor bootstrap for resampling hydrologic time series. Water Resour. Res. 32, 679–693. https://doi.org/10.1029/95WR02966.
Li, Z., Shao, Q., Xu, Z., Cai, X., 2010. Analysis of parameter uncertainty in semi-distributed hydrological models using bootstrap method: a case study of SWAT model applied to Yingluoxia watershed in northwest China. J. Hydrol. 385, 76–83. https://doi.org/10.1016/j.jhydrol.2010.01.025.
Mason, D.C., Davenport, I.J., Neal, J.C., Schumann, G.J.-P., Bates, P.D., 2012. Near Real-time flood detection in urban and rural areas using high-resolution synthetic aperture radar images. IEEE Trans. Geosci. Remote Sens. 50, 3041–3052. https://doi.org/10.1109/TGRS.2011.2178030.
McPherson, M.B., Schneider, W.J., 1974. Problems in modeling urban watersheds. Water Resour. Res. 10, 434–440. https://doi.org/10.1029/WR010i003p00434.
Merz, R., Parajka, J., Blöschl, G., 2011. Time stability of catchment model parameters: implications for climate impact analyses. Water Resour. Res. 47, W02531. https://doi.org/10.1029/2010WR009505.
Meyer, J.S., Ingersoll, C.G., McDonald, L.L., Boyce, M.S., 1986. Estimating uncertainty in population growth rates: jackknife vs. bootstrap techniques. Ecology 67, 1156–1166. https://doi.org/10.2307/1938671.
Moore, R.J., Bell, V.A., Cole, S.J., Jones, D.A., 2007. Rainfall-runoff and other modelling for ungauged/low-benefit locations: operational guidelines. R&D Technical Report SC030227/SR2 Product Code: SCHO0307BMEU-E-P.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I – A discussion of principles. J. Hydrol. 10, 282–290. https://doi.org/10.1016/0022-1694 (70)90255-6.
Önöz, B., Bayazit, M., 2012. Block bootstrap for Mann-Kendall trend test of serially dependent data. Hydrol. Process. 26, 3552–3560. https://doi.org/10.1002/hyp.8438.
Oudin, L., Salavati, B., Furusho-Percot, C., Ribstein, P., Saadi, M., 2018. Hydrological impacts of urbanization at the catchment scale. J. Hydrol. 559, 774–786. https://doi.org/10.1016/j.jhydrol.2018.02.064.
Padiyedath, S.G., Kawamura, A., Takasaki, T., Amaguchi, H., Azhikodan, G., 2018. An effective storage function model for an urban watershed in terms of hydrograph reproducibility and Akaike information criterion. J. Hydrol. 563, 657–668. https://doi.org/10.1016/j.jhydrol.2018.06.035.
Padiyedath, S.G., Kawamura, A., Takasaki, T., Amaguchi, H., Azhikodan, G., 2018. Performance evaluation of urban storage function (USF) model compared with various conventional storage function models for an urban watershed. J. JSCE, Ser. B1 74 (4), 973–978.
Poulin, A., Brissette, F., Leconte, R., Arsenault, R., Malo, J.-S., 2011. Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snowdominated river basin. J. Hydrol. 409, 626–636. https://doi.org/10.1016/j.jhydrol.2011.08.057.
Prasad, R., 1967. A nonlinear hydrologic system response model. Proc. ASCE 93 (HY4), 201–221.
Salavati, B., Oudin, L., Furusho-Percot, C., Ribstein, P., 2016. Modeling approaches to detect land-use changes: urbanization analyzed on a set of 43 US catchments. J. Hydrol. 538, 138–151. https://doi.org/10.1016/j.jhydrol.2016.04.010.
Selle, B., Hannah, M., 2010. A bootstrap approach to assess parameter uncertainty in simple catchment models. Environ. Model. Softw. 25, 919–926. https://doi.org/10.1016/j.envsoft.2010.03.005.
Shah, S.M.S., O'Connell, P.E., Hosking, J.R.M., 1996. Modelling the effects of spatial variability in rainfall on catchment response. 2. Experiments with distributed and lumped models. J. Hydrol. 175, 89–111. https://doi.org/10.1016/S0022-1694(96)80007-2.
Shalizi, C.R., 2016. Advanced Data Analysis From an Elementary Point of View. Cambridge University Press, Cambridge, England, UK.
Shao, J., Tu, D., 1995. The Jackknife and Bootstrap. Statistics Soc., Springer Series in Statistics. J. R. Stat. Soc. Ser. A. https://doi.org/10.1007/978-1-4612-0795-5.
Sheng, J., Wilson, J.P., 2009. Watershed urbanization and changing flood behavior across the Los Angeles metropolitan region. Nat. Hazards 48, 41–57. https://doi.org/10.1007/s11069-008-9241-7.
Shrestha, D.L., 2009. Uncertainty Analysis in Rainfall-Runoff Modelling: Application of Machine Learning Techniques. CRC Press.
Sivakumar, B., 2017. In: Chaos in Hydrology: Bridging Determinism and Stochasticity, Chapter 3. Springer Science + Business Media, Dordrecht, pp. 63–110. https://doi.org/10.1007/978-90-481-2552-4.
Sivakumar, B., Berndtsson, R., 2010. In: Advances in Data-Based Approaches for Hydrologic Modeling and Forecasting. Chapter 3. World Scientific Publishing Co. Pte. Ltd., Singapore. https://doi.org/10.1142/9789814307987.
Sonali, P., Nagesh Kumar, D., 2013. Review of trend detection methods and their application to detect temperature changes in India. J. Hydrol. 476, 212–227. https://doi.org/10.1016/j.jhydrol.2012.10.034.
Srinivas, V.V., Srinivasan, K., 2005. Matched block bootstrap for resampling multiseason hydrologic time series. Hydrol. Process. 19, 3659–3682. https://doi.org/10.1002/hyp.5849.
Stine, R.A., 1985. Bootstrap Prediction Intervals for Time Series. J. Am. Stat. Assoc. 80 (392), 1026–1031. https://doi.org/10.2307/2288570.
Sugiyama, H., Kadoya, M., Nagai, A., Lansey, K., 1997. Evaluation of the storage function model parameter characteristics. J. Hydrol. 191, 332–348. https://doi.org/10.1016/S0022-1694(96)03026-0.
Sun, J., Liu, Y., Wang, Y., Bao, G., Sun, B., 2013. Tree-ring based runoff reconstruction of the upper Fenhe River basin, North China, since 1799 AD. Quat. Int. 283, 117–124. https://doi.org/10.1016/j.quaint.2012.03.044.
Suriya, S., Mudgal, B.V., 2012. Impact of urbanization on flooding: the Thirusoolam sub watershed – A case study. J. Hydrol. 412–413, 210–219. https://doi.org/10.1016/j.jhydrol.2011.05.008.
Swain, J.B., Patra, K.C., 2017. Streamflow estimation in ungauged catchments using regionalization techniques. J. Hydrol. 554, 420–433. https://doi.org/10.1016/j.jhydrol.2017.08.054.
Takasaki, T., Kawamura, A., Amaguchi, H., Araki, K., 2009. New storage function model

considering urban runoff process. J. JSCE B 65 (3), 217–230. https://doi.org/10.2208/jscejb.65.217.
Tasker, G.D., Dunne, P., 1997. Bootstrap position analysis for forecasting low flow frequency. J. Water Resour. Plan. Manage. 123, 359–367. https://doi.org/10.1061/(ASCE)0733-9496(1997) 123:6(359).
Tokyo Metropolitan Government – TMG, 2016. River development plan in Kanda watershed, Tokyo.
Uhlenbrook, S., Seibert, J., Leibundgut, C., Rodhe, A., 1999. Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. Hydrol. Sci. J. 44, 779–797. https://doi.org/10.1080/02626669909492273.
Vaze, J., Post, D.A., Chiew, F.H.S., Perraud, J.M., Viney, N.R., Teng, J., 2010. Climate non-stationarity – validity of calibrated rainfall-runoff models for use in climate change studies. J. Hydrol. 394 (3–4), 447–457. https://doi.org/10.1016/j.jhydrol.2010.09.018.
Vaze, J., Teng, J., 2011. Future climate and runoff projections across New South Wales, Australia: results and practical applications. Hydrol. Process. 25 (1), 18–35. https://doi.org/10.1002/hyp. 7812.
World Meteorological Organisation – WMO, 1992. Simulated Real-time Intercomparison of Hydrological Models. WMO Publications, Geneva.
World Meteorological Organization – WMO, 2008. Urban flood risk management - A Tool for Integrated Flood Management, APFM Technical Document No. 11, Flood Management Tools Series.
Yang, J., Reichert, P., Abbaspour, K.C., Xia, J., Yang, H., 2008. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. J. Hydrol. 358, 1–23. https://doi.org/10.1016/j.jhydrol.2008.05.012.
Zhang, Z., Lu, W., Chu, H., Cheng, W., Zhao, Y., 2014. Uncertainty analysis of hydrological model parameters based on the bootstrap method: a case study of the SWAT model applied to the Dongliao River Watershed, Jilin Province, Northeastern China. Sci. China Technol. Sci. 57, 219–229. https://doi.org/10.1007/s11431-013-5385-0.
Zucchini, W., Adamson, P.T., 1989. Bootstrap confidence intervals for design storms from exceedance series. Hydrol. Sci. J. 34, 41–48. https://doi.org/10.1080/02626668909491307.