SCIENCE CHINA
Technological Sciences

• Article •

January 2014 Vol.57 No.1: 219–229
doi: 10.1007/s11431-013-5385-0

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

ZHANG Zheng, LU WenXi*, CHU HaiBo, CHENG WeiGuo & ZHAO Ying

Key Laboratory of Groundwater Resources and Environment, Ministry of Education, College of Environment and Resources, Jilin University, Changchun 130021, China

Received June 4, 2013; accepted September 25, 2013; published online October 24, 2013

As an important tool for the description and analysis of hydrological processes, the watershed hydrological model has been increasingly applied to watershed hydrological simulations and water resource management. However, in most cases, model parameters are only determined in a calibration scheme which fits the modeled data to observations, thus significant uncertainties exist in the model parameters. How to quantitatively evaluate the uncertainties in model parameters and the resulting uncertainty impacts on model simulations has always been a question which has attracted much attention. In this study, two methods based on the bootstrap method (specifically, the model-based bootstrap and block bootstrap) are used to analyze the parameter uncertainties in the case of the SWAT (Soil and Water Assessment Tool) model applied to a hydrological simulation of the Dongliao River Watershed. Then, the uncertainty ranges of five sensitivity parameters are obtained. The calculated variation coefficients and the variable parameter contributions show that, among the five parameters, *ESCO* and *CN*2 have relatively high uncertainties: the variation coefficients and contribution rates are 23.98 and 70%, 14.43 and 18%, respectively. The three remaining parameters have relatively low uncertainties. We compare the two uncertainty ranges of parameters acquired by the two bootstrap methods, and find that the uncertainty ranges of parameters acquired by the block bootstrap are narrower than those acquired by the model-based bootstrap. Further analysis of the effects of parameter uncertainties on the model simulation reveals that the parameter uncertainties have great impacts on results of the model simulation, and in the model calibration stage 60%−70% of runoff observations were within the corresponding 95% confidence interval. The uncertainty in the model simulation during the flood season (i.e. the wet period) is relatively higher than that during the dry season.

parameters, uncertainty analysis, hydrological model, bootstrap, SWAT (Soil and Water Assessment Tool)

Citation: Zhang Z, Lu W X, Chu H B, et al. 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 Tech Sci, 2014, 57: 219–229, doi: 10.1007/s11431-013-5385-0

1 Introduction

As an important tool for the description and analysis of hydrological processes, hydrological models have been increasingly applied to watershed hydrological simulations and water resource management [1–3]. Hydrological models often contain a large number of parameters, and these parameters are used to describe the important characteristics of the watershed's sub-surface and hydrological processes. The accuracy of the parameter values has a significant impact on the model's ability to truly reflect the hydrological processes of the real system [4–6]. However, due to various practical constraints, these parameters are typically not measured directly, but instead are often indirectly estimated via model calibration, where the modeled values are fitted

*Corresponding author (email: luwenxi@jlu.edu.cn)

to observations. Model calibration can be affected by a variety of factors. When the residual between the observed and simulated values reaches its minimum, the model parameter value is considered as the optimum value, which may differ from the true parameter value inferred by the actual physical process [7, 8]. Parameters obtained in this manner cannot fully characterize the true state of their corresponding physical process in the real system, thus leading to a great level of uncertainty. Parameter uncertainty will inevitably have an impact on the model simulations, by introducing uncertainties in simulation results [9, 10]. How to assess the uncertainties in hydrological model parameters and their impacts on the uncertainty of model simulations has always been a topic of great interest. The quantitative evaluation of parameter uncertainty and its influence on the uncertainty of hydrological model simulations is critical in reducing the uncertainty of these simulations, and in assessing their effectiveness [11–13].

The uncertainties in hydrological model parameters are of great importance and have therefore received considerable attention [14–18]. Many methods have been applied to the analysis of parameter uncertainty, of which the Generalised Likelihood Uncertainty Estimation (GLUE) proposed by Beven and Binley [19] is one of the most common methods [20–22]. In comparison with other methods, GLUE may be easily implemented, using simple principles, and does not require any changes to the source code of the simulation model. However, GLUE also has very evident shortcomings, such as subjective choice of the likelihood function and truncation threshold used to separate behavioral and non-behavioral models. Moreover, the computational efficiency of this method is rather low. Sampling in the parameter space cannot be guaranteed to have a sufficient density (thus Monte Carlo sampling is usually performed in GLUE), which will lead to incorrect estimation results [23]. Markov Chain Monte Carlo (MCMC) simulation is another relatively common method used for uncertainty analysis [24–27]. This method emphasizes the use of a random step to generate a sample in the parameter space, thereby building a Markov chain with a smooth distribution for deriving the posterior distribution of the parameters. This method has a significant advantage in the model as it does not require any linear assumptions to be made, nor does it require differentiability of the output values of the model with respect to the parameter values [28]. With the increasing attention toward the MCMC method, many researchers have integrated this method with others [29–32], in order to improve the effectiveness and efficiency of the original method by using the powerful spatial search functions of the MCMC method. Blasone et al. [33] demonstrated that a combination of the MCMC sampling scheme and GLUE was capable of greatly improving the effectiveness and efficiency of the GLUE method.

In addition to the first two methods mentioned above, the Bayesian method is another commonly used method for uncertainty analysis, in which the fixed but unknown values of the parameters are taken as random variables to derive the posterior distribution of the parameters, based on the prior distribution, in the Bayesian theory. This method possesses high simulation and computational flexibility [34, 35], but despite being widely used, it has remained controversial due to the necessity of assuming that the parameters obey a particular prior distribution, when such an assumption is often unrealistic or incorrect.

Compared with these methods, a non-parametric statistical technique termed the "bootstrap" may be easily implemented using simple principles and has several advantages [36]: (i) only a small amount of programming is required; (ii) compared with other relevant methods, the amount of computation required is relatively small; and (iii) there is no need to make any assumptions regarding the prior distribution of variables. As the bootstrap method has gained popularity, many novel methods have been successively developed based on the concepts of the bootstrap [37–39]. In the present study, two different bootstrap methods are used to assess the parameter uncertainty in a semi-distributed hydrological model: the Soil and Water Assessment Tool (SWAT), and the model's impacts on the uncertainty of a model simulation are analyzed. By comparing the parameter uncertainty analysis results of the two methods, the model parameters with the greatest uncertainty are identified, and possible reasons for the differences in the parameter uncertainty analysis results of the two methods are discussed.

The overall structure of this study is as follows. Section 2 describes the basic principles of the bootstrap and the steps used in the two different bootstrap methods; Section 3 provides a basic overview of the Dongliao River Watershed and the data collection required for this study; Section 4 describes several major sensitivity parameters obtained through model sensitivity analysis; Section 5 addresses the parameter uncertainty analysis and its impacts on the uncertainty of the model simulation results, and discusses possible reasons for differences in the results of the two methods; and finally, Section 6 summarizes some of the main conclusions drawn by the study.

2 Bootstrap

The bootstrap is a nonparametric resampling method, first proposed by Eforn [40]. Its basic concept is the extraction of a "bootstrap sample" of size n from a sample of an unknown distribution, also of size n, by sampling with replacement. First, a number of bootstrap samples are extracted from the original sample sequentially and independently. Then these bootstrap samples are used to perform statistical tests regarding the statistical characteristics of the unknown distribution, such as the variance, standard deviation and confidence interval. Compared with the classical statistical methods, the advantage of this method is that the calculation process is relatively simple, with no need to make assumptions regarding the unknown distribution. The bootstrap method has become more and more widely applied to uncertainty analysis [41–43]. Sampling with replacement is performed directly on the independent samples with the same distribution, leading to the generation of bootstrap samples. These bootstrap samples may then be used for deducing the statistical characteristics of the variables. However, for sample data with a certain dependency structure (e.g. time-series data), it is unreasonable to perform direct resampling, which would destroy the original dependency structure. Maintaining the dependency structure of the data throughout the resampling process yields a variety of methods based on the bootstrap, including the model-based bootstrap and block bootstrap. This study attempts to apply these latter two methods to solve the issue of maintaining the dependency structure of the data during destroying by direct resampling in the uncertainty analysis.

2.1 Model-based bootstrap

In the model-based bootstrap, the data are defined as follows [44]:

$$\left\{\boldsymbol{X}(t), Y^{\mathrm{obs}}(t)\right\}(t=1,2,\cdots,N),$$

where $\boldsymbol{X}(t)$ indicates the input data (e.g. precipitation), $\boldsymbol{Y}^{\mathrm{obs}}(t)$ indicates the observation data (e.g. runoff), and N indicates the sequence length. In this study, the hydrological model is expressed as

$$\boldsymbol{Y}(t)=f\left[\boldsymbol{X}(t),\boldsymbol{\beta}\right], \tag{1}$$

where $\boldsymbol{\beta}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m)$ indicates the m parameter vectors in the model. Through model calibration, the estimated value $\hat{\boldsymbol{\beta}}$ of the model parameter vector $\boldsymbol{\beta}$ and the estimated value $\hat{\boldsymbol{Y}}(t)$ of model output vector $\boldsymbol{Y}(t)$ may be obtained, after which the model is expressed as

$$\hat{\boldsymbol{Y}}(t)=f\left[\boldsymbol{X}(t),\hat{\boldsymbol{\beta}}\right]. \tag{2}$$

The residual of the model is expressed as

$$\boldsymbol{\varepsilon}(t)=\boldsymbol{Y}^{\mathrm{obs}}(t)-\hat{\boldsymbol{Y}}(t)=\boldsymbol{Y}^{\mathrm{obs}}(t)-f\left[\boldsymbol{X}(t),\hat{\boldsymbol{\beta}}\right], \tag{3}$$

where $\boldsymbol{\varepsilon}(t)$ is typically considered to be mutually independent and $t=1,2,\cdots,N$. The basic concept of the model-based bootstrap is to conduct bootstrap sampling of the residuals to create a set of new residuals; this new set is then used for model fitting, yielding the bootstrap estimated sample $\hat{\boldsymbol{\beta}}^*$ of parameter $\boldsymbol{\beta}$. The model-based bootstrap is applied to parameter uncertainty analysis using the following specific steps:

(i) Replacement sampling is randomly conducted on the residual sequence $\boldsymbol{\varepsilon}(t)$ to form a new residual sequence $\boldsymbol{\varepsilon}^*(t)$ (i.e. the bootstrap sample generated through the bootstrap method).

(ii) The bootstrap data $\boldsymbol{Y}^*(t)=\hat{\boldsymbol{Y}}(t)+\boldsymbol{\varepsilon}^*(t)$ are generated, where $t=1,2,\cdots,N$.

(iii) The model $\boldsymbol{Y}^*(t)=f\left[\boldsymbol{X}(t),\boldsymbol{\beta}\right]$ is fitted to obtain the bootstrap estimated value $\hat{\boldsymbol{\beta}}^*=(\hat{\boldsymbol{\beta}}_1^*,\hat{\boldsymbol{\beta}}_2^*,\cdots,\hat{\boldsymbol{\beta}}_m^*)$ of parameter vector $\boldsymbol{\beta}$ and new runoff data $\hat{\boldsymbol{Y}}^*(t)=f\left[\boldsymbol{X}(t),\hat{\boldsymbol{\beta}}^*\right]$.

(iv) Steps 1–3 are repeated 1000 times.

(v) The confidence interval of parameter vector $\boldsymbol{\beta}$ is obtained. Taking $\boldsymbol{\beta}_m$ as an example, the bootstrap sampling method is used to obtain the ordered bootstrap estimation sequence $[\hat{\boldsymbol{\beta}}_{m1}^*,\hat{\boldsymbol{\beta}}_{m2}^*,\cdots,\hat{\boldsymbol{\beta}}_{m1000}^*]$ of $\boldsymbol{\beta}_m$, so as to derive the confidence interval $[\boldsymbol{\beta}_{m1000(\alpha/2)}^*,\cdots,\boldsymbol{\beta}_{m1000(1-\alpha/2)}^*]$ of parameter $\boldsymbol{\beta}_m$ at the confidence level of α.

2.2 Block bootstrap

The basic concept of the block bootstrap is no longer the bootstrap sampling of a single data point, but the division of the entire data sequence into a number of "blocks". Then, bootstrap sampling is carried out on these blocks. The specific steps of the block bootstrap method are as follows:

(1) The residual sequence $\boldsymbol{\varepsilon}(t)=\boldsymbol{Y}^{\mathrm{obs}}(t)-\hat{\boldsymbol{Y}}(t)$ is obtained, where $t=1,2,\cdots,N$.

(2) N data sequence points are divided into B blocks of lengths $L(N=BL)$. Then, random bootstrap sampling is conducted on these blocks, and the block bootstrap samples are connected to form the bootstrap sample $\boldsymbol{\varepsilon}^*(t)$ of $\boldsymbol{\varepsilon}(t)$. Referring to the studies performed by Hall et al. [45], in this study the length L is set to 3 months.

(3) The bootstrap data $\boldsymbol{Y}^*(t)=\hat{\boldsymbol{Y}}(t)+\boldsymbol{\varepsilon}^*(t)$ are generated, where $t=1,2,\cdots,N$.

(4) The model $\boldsymbol{Y}^*(t)=f\left[\boldsymbol{X}(t),\boldsymbol{\beta}\right]$ is fitted to obtain the bootstrap estimated value $\hat{\boldsymbol{\beta}}^*=[\hat{\boldsymbol{\beta}}_1^*,\hat{\boldsymbol{\beta}}_2^*,\cdots,\hat{\boldsymbol{\beta}}_m^*]$ of parameter vector $\boldsymbol{\beta}$ and the new runoff time series data $\hat{\boldsymbol{Y}}^*(t)=f\left[\boldsymbol{X}(t),\hat{\boldsymbol{\beta}}^*\right]$.

(5) Steps 1–4 are repeated 1000 times.

(6) The confidence interval of parameter vector $\boldsymbol{\beta}$ is obtained. Taking $\boldsymbol{\beta}_m$ as an example, the bootstrap sampling method is used to obtain the ordered bootstrap estimation sequence $[\hat{\boldsymbol{\beta}}_{m1}^*,\hat{\boldsymbol{\beta}}_{m2}^*,\cdots,\hat{\boldsymbol{\beta}}_{m1000}^*]$ of $\boldsymbol{\beta}_m$, so as to derive the confidence interval $[\boldsymbol{\beta}_{m1000(\alpha/2)}^*,\cdots,\boldsymbol{\beta}_{m1000(1-\alpha/2)}^*]$ of parameter $\boldsymbol{\beta}_m$ at the confidence level of α.

3 Study area and data

3.1 Study area

Located in the center of Jilin Province of China, the Dongliao

River originates from Sahaling Mountain near Liaoyuan City, and has a total length of 448 km and drainage area of 11306 km^2. The geographic coordinate range is 123°39′–125°32′E, 42°37′–44°09′N. The Dongliao River Watershed has a semi-humid temperate continental monsoon climate with an average temperature of 5.2°C, annual extreme maximum temperature of 36°C, annual extreme minimum temperature of –41.2°C, multi-year average rainfall of 450–700 mm, and evaporation of 700–900 mm. The rainfall is very unevenly distributed within a year, the June-September period accounts for 75% of the annual total. The upper reaches of the Dongliao River Watershed are mainly dominated by hills with altitudes of 258–598 m. The central and lower reaches of the Dongliao River are mainly composed of alluvial-proluvial and valley alluvial plains with altitudes of 155–258 m [46]. There are a variety of soil types in the study area, among which brown soil (22.59%), black soil (14.65%) and meadow soil (13.13%) are the main types [47].

3.2 Data preparation

When applying the SWAT model to the runoff simulation, the topography, climate, soil type, land use and other input data related to the study area are required as support for the model. In the present study, terrain data for the Dongliao River Watershed comprised a digital elevation model (DEM) with a resolution of 30 m×30 m downloaded from the international scientific data services platform (http://datamirror.csdb.cn/); the soil type map and land use map were developed in ArcGIS with a 1∶500000 paper chart from the year 2000 as the base maps. The maximum temperature, minimum temperature, solar radiation, relative humidity, average wind speed and other meteorological data were derived from the 60 consecutive years of daily observations recorded from 1951 to 2010 at three meteorological stations: Changchun, Siping and Shuangliao. The rainfall and runoff data were derived from daily observational data recorded from 2008 to 2009 at the Wangben, Shiwu, Erlongshan, Shuangliao and Siping hydrological stations. The respective locations of the study area and hydrometeorological stations are shown in Figure 1.

4 Model parameters

The SWAT model includes many parameters, but not all the parameters have significant impacts on the model simulation results. Therefore, in this study the LH-OAT (Latin Hypercube-One factor At a Time) method was used for sensitivity analysis of the model parameters, so as to obtain five sensitive parameters for the model output results. These were *CN2*, *ESCO*, *SOL_AWC*, *ALPHA_BF* and *SURLAG*. Here, we focused on analyzing the uncertainties in these five parameters, and assessing the impacts of their uncertainties on the model simulation results. However, in the SWAT model, the study area is divided into different sub-watersheds and hydrological response units (HRUs), and it is necessary to calibrate the parameters in every sub-watershed and HRU. Therefore, in order to simplify the calibration and calculation processes, the parameters of the model were expressed in the form of aggregate parameters, whereby the new parameters were created by making some modifications to the original parameters. The aggregate parameters were constructed using the iSWAT interface program; for the specific steps please refer to ref. [48]. The construction of the aggregate parameter space is shown in Table 1. In this study, the least squares method was used for calibration of the aggregate parameters, and the Nash-Suttcliffe coefficient E_{ns} and correlation coefficient R^2 were applied to assess the model simulation results.

5 Results and discussion

5.1 Uncertainty analysis of model parameters

The model-based bootstrap and block bootstrap methods were used to generate 1000 bootstrap estimated values for each of the five aggregate parameters in the SWAT model. Then, the ordered arrangement was performed on the 1000 bootstrap estimated values of each parameter, with thresholds of 2.5% and 97.5% as the lower and upper limits, respectively, in order to obtain the 95% confidence interval of the parameters (the uncertainty range of the parameters is typically expressed by the confidence interval). The confidence intervals and marginal distributions of the five parameters are shown in Table 2 and Figure 2.

Table 2 shows the uncertainty range of each parameter revealing that the value ranges of the parameters obtained either by model-based bootstrap or block bootstrap were narrower than those of most of the initial parameters (Table 1). *CN2* indicates the curve number in the SWAT model. Before the uncertainty analysis was performed, the aggregate parameter *a_CN2.mgt* had the value range of [5, 20], but the confidence intervals obtained by the bootstraps were [8.521, 12.359] (model-based bootstrap) and [8.682, 12.253] (block bootstrap). *ESCO* indicates the soil evaporation compensation coefficient factor in the SWAT model. Parameter *v_ESCO.hru* had the initial value range of [0, 1], but after the bootstrap uncertainty analysis, the confidence intervals were obtained as [0.072, 0.282] (model-based bootstrap) and [0.085, 0.215] (block bootstrap). The other parameters were as follows: parameter *r_SOL_AWC.sol* is the soil available water content, with confidence intervals of [0.283, 0.581] (model-based bootstrap) and [0.366, 0.452] (block bootstrap); parameter *v_ALPHA_BF.gw* represents the response intensity of underground runoff to precipitation recharge, with uncertainty ranges of [0.154, 0.384] (model-based bootstrap) and [0.182, 0.331] (block bootstrap); and parameter *v_SURLAG.bsn* is the coefficient of surface runoff

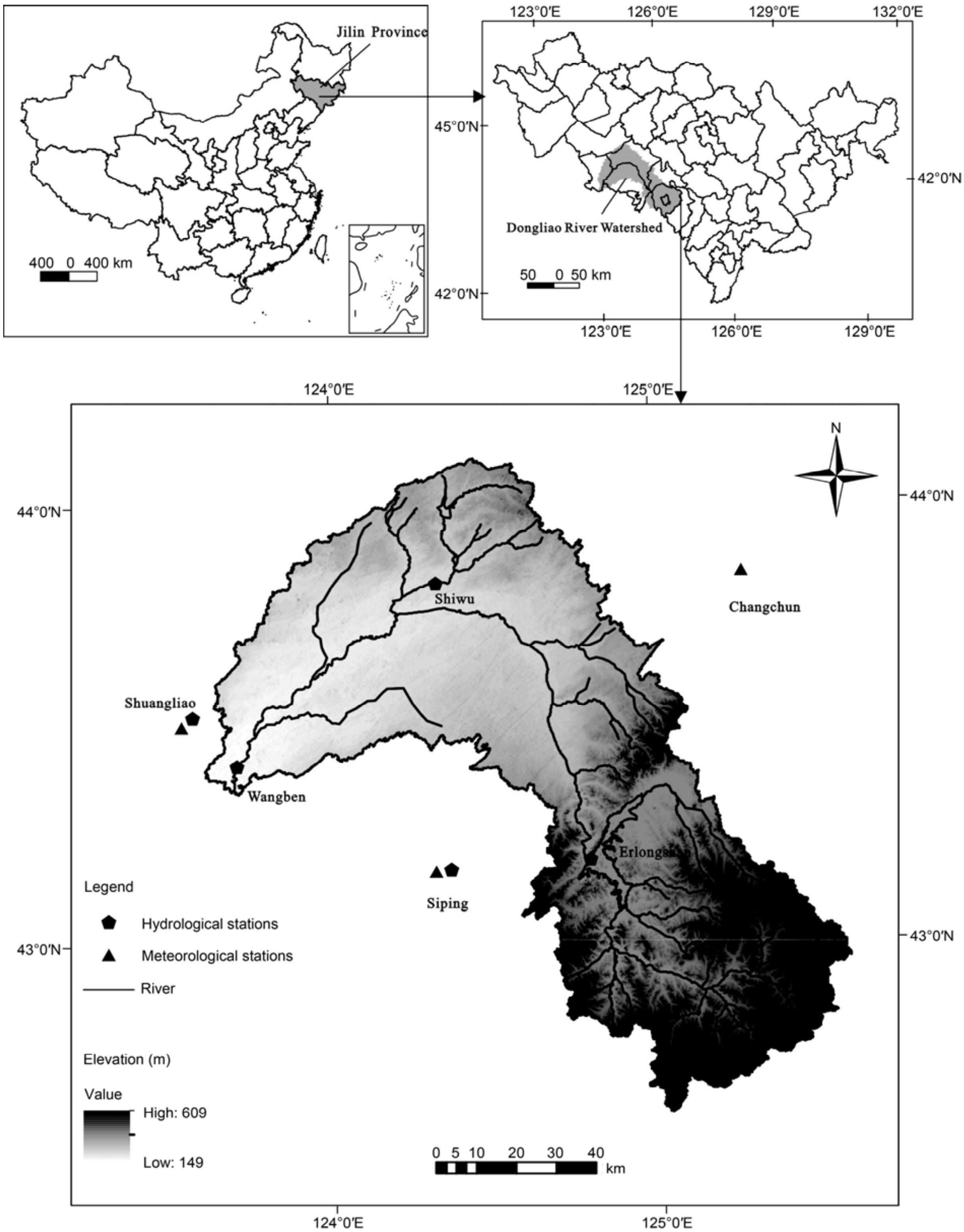

Figure 1 Locations of Dongliao River Watershed and hydrometeorological stations.

lag time, with uncertainty ranges of [2.821, 4.124] (model-based bootstrap) and [2.636, 3.851] (block bootstrap). Therefore, after the uncertainty analysis, the parameter confidence intervals were significantly smaller than the initial ranges, indicating that the bootstrap approach is both feasible and useful in parameter uncertainty analysis. The method not only reduces the confidence intervals of the parameter value ranges, but also provides a reference for parameter

Table 1 Value ranges of aggregate parameters in the calibration stage

Aggregate parameter[a)]	Definition of parameter in model	Initial value	Range of value[b)]
a_CN2.mgt	initial SCS curve number for moisture condition II	10	[5, 20]
v_ESCO.hru	soil evaporation compensation factor	0.15	[0, 1]
r_SOL _AWC.sol	soil available water capacity (mmH_2O/mm soil)	0.2	[0.1, 0.75]
v_ALPHA _BF.gw	base flow alpha factor, 1/d	0.06	[0, 1]
v_SURLAG.bsn	surface runoff lag coefficient	2.0	[0, 15]

a) The modifications of *a*, *v_* and *r_* refer to an absolute increase, a replacement and a relative change to the initial parameter value, respectively; b) the value ranges of the aggregate parameters are determined according to those of the initial values and corresponding original parameters of the SWAT model user manual.

Table 2 95% confidence intervals and variation coefficients of model parameters

Aggregate parameter	Model-based bootstrap	CV[a)]	Block bootstrap	CV
a_CN2.mgt	[8.521, 12.359]	13.02%	[8.682, 12.253]	14.43%
v_ESCO.hru	[0.072, 0.282]	26.21%	[0.085, 0.215]	23.98%
r_SOL _AWC.sol	[0.283, 0.581]	9.01%	[0.366, 0.452]	7.65%
v_ALPHA _BF.gw	[0.154, 0.384]	14.89%	[0.182, 0.331]	13.86%
v_SURLAG.bsn	[2.821, 4.124]	11.27%	[2.636, 3.851]	8.51%

a) CV represents coefficient of variation

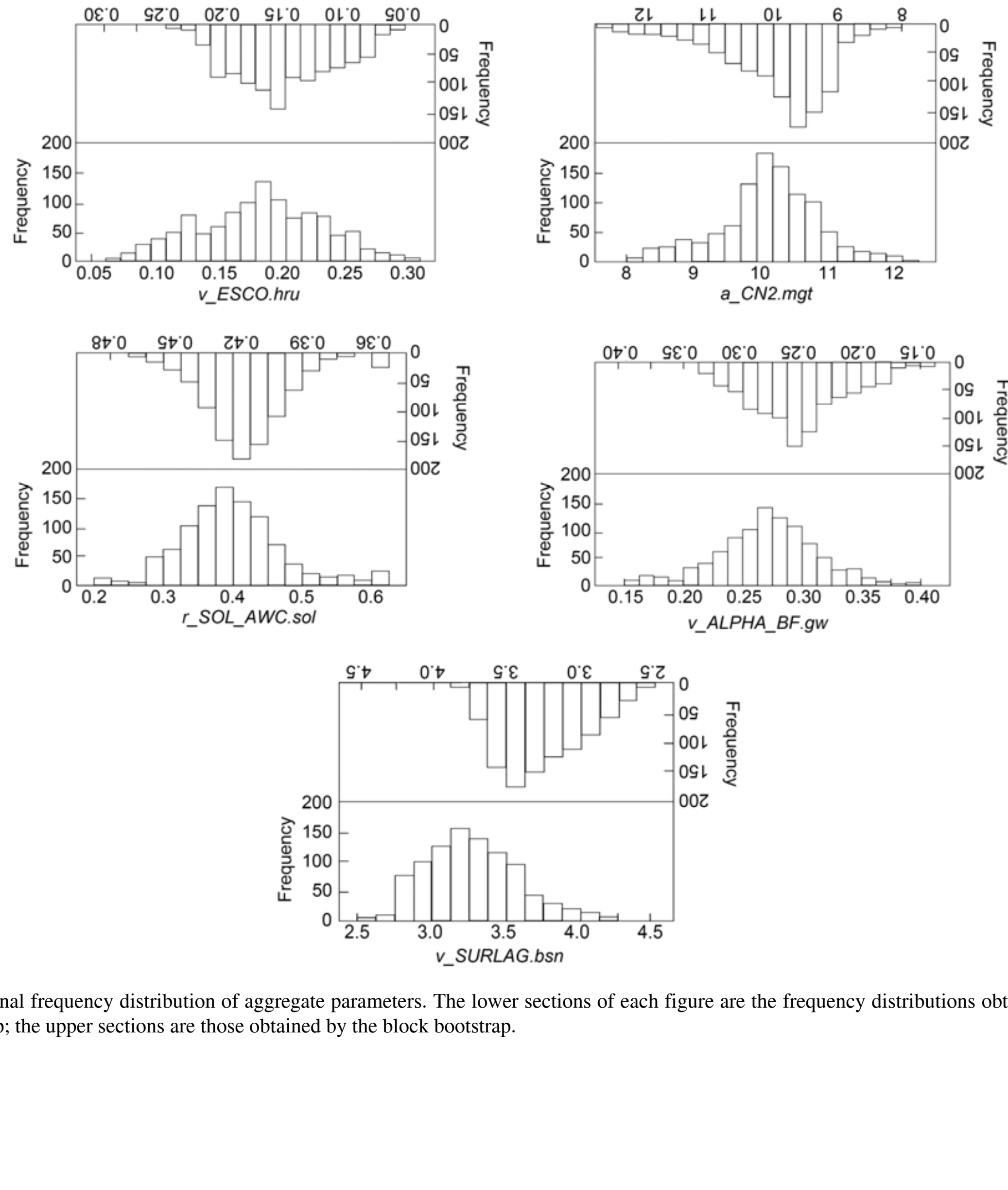

Figure 2 Marginal frequency distribution of aggregate parameters. The lower sections of each figure are the frequency distributions obtained by the model-based bootstrap; the upper sections are those obtained by the block bootstrap.

calibration in the SWAT model.

After obtaining the confidence interval for each parameter, we also calculated the corresponding variation coefficients and each parameter's contribution to variations in the model outputs. Through analysis of the variability of the parameters themselves and their impacts on the model outputs, we next discuss the uncertainty in each parameter. The variation coefficient is generally used to characterize the level of dispersion of data, with a larger variation coefficient indicating a higher degree of uncertainty. It can be seen from Table 2 that the variation coefficients of *r_SOL_AWC.sol* and *v_SURLAG.bsn* were relatively small, at 7.65% and 8.51%, respectively; those of *v_ALPHA_BF.gw* and *a_CN2.mgt* were relatively large, at 13.86% and 14.43%, respectively, both being close to 15%; and the variation coefficient of *v_ESCO.hru* was the largest, at 23.98%. We then ranked the calculated uncertainties in the parameters as follows: *v_ESCO.hru*>*a_CN2.mgt*>*v_ALPHA_BF.gw*>*v_SURLAG.bsn*>*r_SOL_AWC.sol*. The contribution rate is defined as the contribution of the variance of a single parameter to the total variance of the model outputs. In this study, the contribution rate was selected to identify the parameters with the greatest uncertainties, and it is generally held that parameters with contribution rates exceeding 5% have relatively high uncertainties [49]. It can be seen from Table 3 that the contribution rate of parameter *v_ESCO.hru* was the largest, at 70%; parameter *a_CN2.mgt* ranked second, at 18%; and those of *r_SOL_AWC.sol*, *v_SURLAG.bsn* and *v_ALPHA_BF.gw* were relatively low, each less than 5%. We ranked the uncertainties in the parameters based on the contribution rate as: *v_ESCO.hru*>*a_CN2.mgt*>*r_SOL_AWC.sol*>*v_SURLAG.bsn*>*v_ALPHA_BF.gw* (note the parameters ranked in the same order with both bootstrap methods). The calculated variation coefficients and contribution rates show that parameters *v_ESCO.hru* and *a_CN2.mgt* had high degrees of uncertainty, and the remaining three parameters had low degrees of uncertainty. This is consistent with the results of Shen et al. [50]. The calculation results also show that parameters *v_ESCO.hru* and *a_CN2.mgt*, which had the highest uncertainties, had a strong impact on the overall uncertainty of the watershed hydrological simulation. Of the five main parameters, *CN2* is an important factor that affects runoff production; it is affected by various factors, such as soil moisture conditions and rainfalls; however, its high uncertainty strongly affected the model simulation. The large uncertainty in *ESCO* was likely to be related to the natural characteristics of the Dongliao River Watershed, since this parameter is affected by vegetation coverage and air temperature. The low vegetation coverage and high air temperature in the research area may aggravate soil evaporation and further impact the runoff production. This is supported by the low value of *ESCO*, since lower values indicate more intense watershed evaporation. However, some studies have found the opposite, for example, Wu and Liu [51] found an uncertainty range of [0.86, 0.98] for parameter *ESCO*, and the estimated value of the parameter was relatively high. This discrepancy may be caused by the different study areas covered by model, or by the different analysis methods.

It can be seen from Figure 2 that while the forms of the marginal distributions of parameters acquired in the model-based bootstrap and those acquired in the block bootstrap were not completely consistent with each other, they did follow similar overall trends, and the corresponding confidence intervals of parameters acquired by the two methods were similar to each other. As described above, one of the advantages of the bootstrap method is that it does not require any assumption of the distribution of variables, and instead uses random resampling based on the existing samples. The marginal distributions of the parameters acquired by the two methods were similar but not consistent, possibly because even though the two methods are based on the same principle, the specific sampling strategies are different when resampling is conducted based on the existing samples. After comparing the lengths of the confidence intervals (the difference between the upper limit and the lower limit of the confidence intervals) acquired by the two methods (Table 2), we found that the confidence interval of parameters acquired by the block bootstrap was narrower than that acquired by the model-based bootstrap. Specifically, the length of the confidence interval for parameter *a_CN2.mgt* acquired by the model-based bootstrap was 3.838, while that acquired by the block bootstrap was 3.571; the length of the confidence interval of parameter *v_ESCO.hru* acquired by the model-based bootstrap was 0.21, while that acquired by the block bootstrap was 0.13; the results of the

Table 3 Contribution rates to streamflow uncertainty (model-based bootstrap)

Parameter	MV	SD	Sensitivity coefficient $S=DY/DX$	Variance component S^2×SD^2	Contribution rate (%)	Rank
v_ESCO.hru	0.16	0.04	16.47	271.28	70.00	1
a_CN2.mgt	10.5	1.37	8.24	69.76	18.00	2
r_SOL _AWC.sol	0.38	0.03	4.40	19.38	4.9	3
v_SURLAG..bsn	3.2	0.36	3.92	15.50	3.9	4
v_ALPHA_BF.gw	0.26	0.04	3.41	11.63	3.2	5

Note. MV represent mean value; SD represent standard deviation; "Sensitivity coefficient" was acquired by first-order error analysis (FOEA), Contribution=S^2×SD^2/Variance, where Variance=387.54 (sum of variances of monthly flow) remaining three parameters were consistent. This may be because the block bootstrap more effectively retains the inherent structure of the data sequence when conducting resampling than the model-based bootstrap, as suggested in the previous work by Benny and Murray [52]. However, in their study they applied an ABC hydrological model, which uses relatively simple parameters. In order to understand the interaction among the model parameters, we constructed a pair-wise correlation scatter diagram of the five model parameters (Figure 3), which shows moderate correlation between parameters *CN2* and *ESCO*, but low or even no correlation among the other parameters. Parameters *CN2* and *ESCO* significantly affected the watershed runoff production; and the two parameters are related with their physical significance, as also found in the parameter correlation analysis of Xie and Lian [53].

5.2 Model uncertainty

In this study, the determination coefficient R^2 and the Nash-Suttcliffe coefficient E_{ns} were selected for evaluation of the SWAT model simulation performance. The monthly runoff data for 2008 and 2009 were used for the model calibration and validation, respectively. Figure 4 shows that the simula-

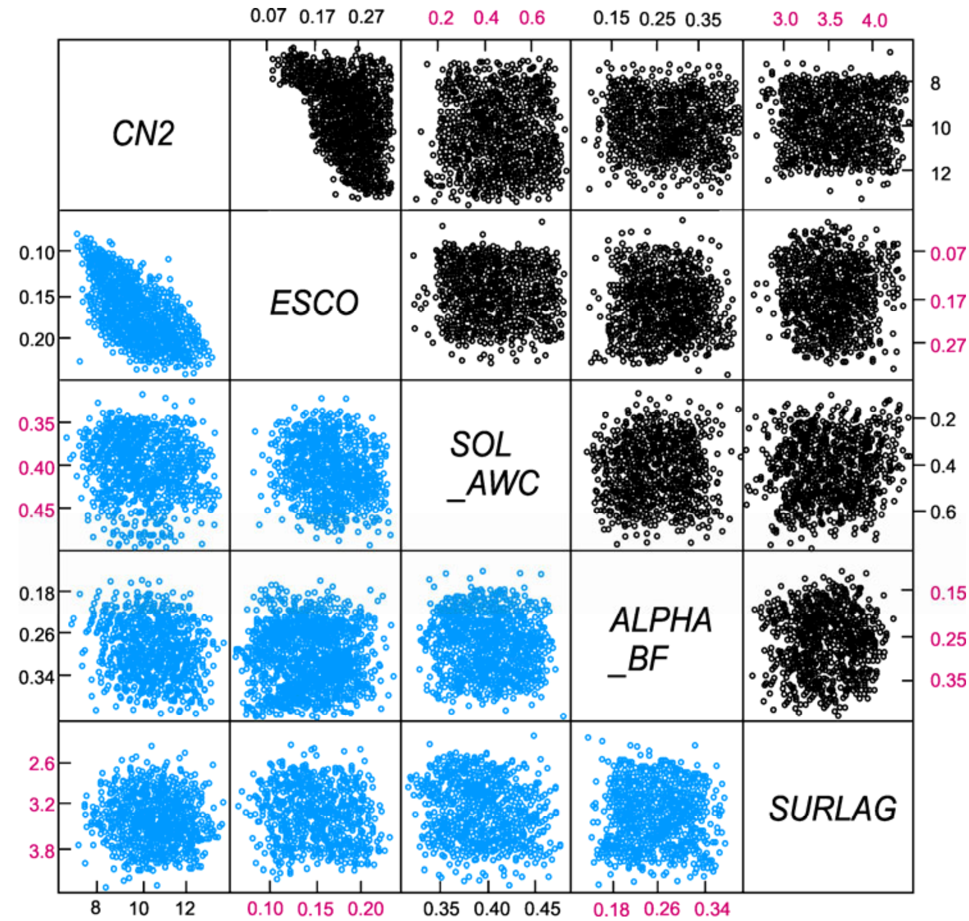

Figure 3 Pair-wise correlations among the model parameters (black represents model-based bootstrap; blue represents block bootstrap).

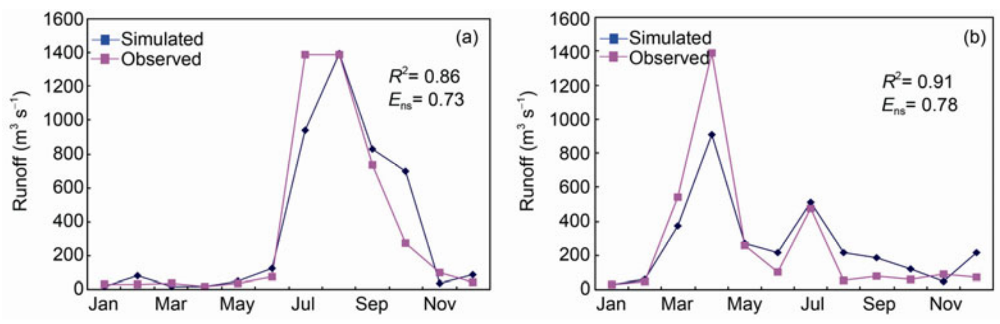

Figure 4 (a) Comparison of simulated and observed values of runoff for 2008; (b) comparison of simulated and observed values of runoff for 2009.

tion results of the SWAT model are reasonable. In the model calibration phase, the determination coefficient R^2 and Nash-Suttcliffe coefficient E_{ns} were 0.86 and 0.73, respectively, but in the model validation phase they were 0.91 and 0.78.

Parameter uncertainty is very likely to have an impact on model simulation. In order to assess the impact of parameter uncertainty on the uncertainty of the model simulation, the 95% confidence intervals of the 1000 groups of corresponding runoff values were calculated. The greater the number of observations contained in the uncertainty interval, the greater the impact of parameter uncertainty on the uncertainty of the model simulation, and vice versa. Figure 5 shows the uncertainty intervals in the model simulation using the model-based bootstrap. The solid line in the figure represents the lower limit of the confidence interval, and the dotted line represents the upper limit of the confidence interval. The results show that 62% of the observations lay within this confidence interval. Meanwhile, 71% of the observations lay within the 95% confidence interval of runoff generated by the block bootstrap (figure omitted). This indicates that parameter uncertainty has a great impact on the uncertainty of the model simulation. Next, it can be seen from Figure 5 that the uncertainty range of runoff in the flood season (wet period) is broader than that of the dry season; in the wet season some of the observation points lie outside the upper and lower limits of the confidence intervals. This shows that the uncertainty of the model simulation is higher in the flood seasons. As is well known, the uncertainty of the model simulation originates from three elements: the uncertainty in the model structure, uncertainties in the parameters and uncertainties in the input data (rainfall). The most likely reason for the high uncertainty of the model simulation in the flood season (given that the model structure is stable in this period, thus the simulation is relatively accurate), is that the uncertainty of the model simulation is mainly affected by the uncertainties in the parameters and rainfall. In contrast, the uncertainty of the model simulation in the non-flood season is mainly affected by the model structure, and the effects of parameter uncertainties are relatively small. Similar inferences were made by Yang et al. [54]. However, based on the results of the present study, we cannot distinguish the effects of the uncertainty in model structure and input data on the model simulation. Some studies [55] have shown that uncertainty in the model structure and input data sometimes has great effects on the overall uncertainty of the model simulation, and must not be ignored. The results of the present research show that when conducting a hydrological simulation by application of the SWAT model to the Dongliao River Watershed, parameter uncertainties have a relatively large effect on the model simulation.

Two prospective applications of bootstrap method in future model uncertainty research are explored in this study. (i) Combination of bootstrap method with other methods, such as BATEA (The Bayesian Total Error Analysis) method [56] or IBUNE (The Integrated Bayesian Uncertainty Estimator) method [57], can be applied to the uncertainty research of model structure. (ii) Considering its advantage, no excessive hypothesis is needed on parameter distribution in estimating distribution characteristics, it only resamples from the limited samples and can reduce the times of model runs, which is helpful to improving the operating efficiencies of some complex models (such as physically-based hydrologic models, land surface models, or even weather and climate models). These two points will be an important aspect for future application of bootstrap method.

6 Conclusions

The bootstrap is a nonparametric statistical method, which requires no assumptions concerning the distributions of the model parameters or their errors. Therefore, this method is more convenient to use than the traditional Bayesian method, and provides a novel method of uncertainty analysis. In this study, both the model-based and block bootstrap methods were used to analyze the impact of parameter uncertainty on the overall uncertainty of a model simulation in the case of the SWAT model applied to a hydrological simulation of

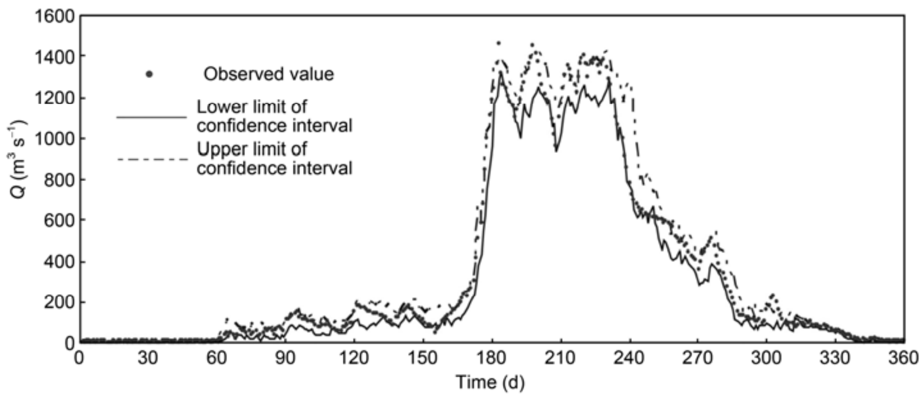

Figure 5 95% confidence interval of runoff obtained by the model-based bootstrap in 2008.

the Dongliao River Watershed. Then, the marginal distribution and uncertainty ranges of five sensitive parameters were obtained. The calculated variation coefficients and variable contribution rates show that among the five key parameters, *ESCO* and *CN2* had relatively high uncertainties while the remaining three parameters had relatively low uncertainties. By comparison we found that the uncertainty ranges of parameters acquired by the block bootstrap were narrower than those acquired by the model-based bootstrap. However, we applied only one block length in the parameter calibration with the block bootstrap method, and further research is required to determine whether the same or better results could be obtained when other block lengths are used.

Next, the impact of individual parameter uncertainties on the overall uncertainty of the model simulation was further analyzed. Calculations revealed that 62% and 71% of the observations lay within the corresponding 95% confidence interval obtained by the model-based bootstrap and the block bootstrap, respectively, indicating that parameter uncertainty had a greater impact on the uncertainty of the model simulation. In addition, due to the uncertainty in model parameters and rainfall, the uncertainty in the model simulation during the flood season was relatively higher than that during the dry season. This is a key point for further research, which needs to address the effects of uncertainties on model structure, parameters and input data. It is hoped that the results of this study may provide a reference for the application of bootstrap methods to parameter uncertainty research with hydrological models, as well as for the application of the SWAT model to water resource management in the Dongliao River Watershed.

This work was supported by the Major Science and Technology Program for Water Pollution and Treatment of China (Grant No. 2012ZX07201-001). The authors are very grateful to Dr SHAO QuanXi for his help in the paper preparation.

1 Zheng J G, Li Y, Han Z Z, et al. Hydrological cycle simulation of an irrigation district based on a SWAT model. Math Comput Model, 2010, 51: 1312–1318
2 Li H Y, Zhang Y Q, Wang B D. Separating impacts of vegetation chang and climate variability on streamflow using hydrological models together with vegetation data. Sci China Tech Sci, 2012, 55: 1964–1972
3 Linden S V D, Woo M K. Application of hydrological models with increasing complexity to subarctic catchments. J Hydrol, 2003, 270: 145–457
4 Dunn S M, Lilly A. Investigating the relationship between a soils classification and the spatial parameters of a conceptual catchment-scale hydrological model. J Hydrol, 2001, 252: 157–173
5 Jeremiah E, Sisson S A, Sharma A, et al. Efficient hyrological model parameter optimization with Sequential Monte Carlo sampling. Environ Modell Softw, 2012, 38: 283–295
6 Shrestha S, Bastola S, Babel M S, et al. The assessment of spatial and temporal transferability of a physically based distributed hydrological model parameters in different physiographic regions of Nepal. J Hydrol, 2007, 347: 153–172
7 Singh S K, Bardossy A. Calibration of hydrological models on hydrologically unusual events. Adv Water Resour, 2012, 38: 81–91
8 Apaydin H, Anli A S, Ozturk A. The temporal transferability of calibrated parameters of a hydrological model. Ecol Model, 2006, 195: 307–317
9 Christiaens K, Feven J. Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework. Hydrol Process, 2002, 16: 373–91
10 Benke K K, Lowell K E, Hamilton A J. Parameter uncertainty, sensitivity analysis and prediction error in a water-blance hydrological model. Math Comput Model, 2008, 47: 1134–1149
11 Hughes D A, Kapangaziwiri E, Sawunyama T. Hydrological model uncertainty assessment in southern Africa. J Hydrol, 2010, 378: 221–232
12 Bastola S, Ishidaira H, Takeuchi K. Regionalisation of hydrological model parameters under parameter uncertainty: A case study involving TOPMODEL and basins across the globe. J Hydrol, 2008, 357: 188–206
13 Christiaens K, Feyen J. The influence of different methods to derive soil hydraulic properties on the uncertainty of various model outputs of a distributed hydrological model. Phys Chem Earth Pt B, 2000, 25: 679–683
14 Jin X L, Xu C Y, Zhang Q, et al. Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. J Hydrol, 2010, 383: 147–155
15 Wang H R, Gao X, Qian L X, et al. Uncertainty analysis of hydrological processes based on ARMA-GARCH model. Sci China Tech Sci, 2012, 55: 2321–2331
16 Mantovan P, Todini E. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. J Hydrol, 2006, 330: 368–381
17 Li Z L, Shao Q X, Xu Z X, et al. Uncertainty issues of a conceptual water balance model for a semi-arid watershed in north-west of China. Hydrol Process, 2013, 27: 304–312
18 Kouchak A A, Nakhjiri N, Habib E. An educational model for ensemble streamflow simulation and uncertainty analysis. Hydrol Earth Syst Sc, 2013, 17: 445–452
19 Beven K, Binley A. The future of distributed models-Model calibration and uncertainty prediction. Hydrol Process, 1992, 6: 279–298
20 Li L, Xia J, Xu C Y, et al. Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models. J Hydrol, 2010, 390: 210–221
21 Rankinen K, Karvonen T, Butterfield D. An application of the GLUE methodology for estimating the parameters of the INCA-N model. Sci Total Environ, 2006, 365: 123–139
22 Gong Y W, Shen Z Y, Hong Q, et al. Parameter uncertainty analysis in watershed total phosphorus modeling using the GLUE methodology. Agr Ecosyst Environ, 2011, 142: 246–255
23 McMichael C E, Hope A S, Loaiciga H A. Distributed hydrological modelling in California semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation. J Hydrol, 2006, 317: 307–324
24 Kanso A, Chebbo G, Tassin B. Application of MCMC-GSA model calibration method to urban runoff quality modeling. Reliab Eng Syst Safe, 2006, 91: 1398–1405
25 Yustres A, Asensio L, Alonso J, et al. A review of Markov Chain Monte Carlo and in formation theory tools for inverse problems in subsurface flow. Computat Geosci, 2012, 16: 1–20
26 Dotto C B S, Mannina G, Kleidorfer M, et al. Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling. Water Res, 2012, 46: 2545–2558
27 Campbell E P, Fox D R, Bates B C. A Bayesian Approach to parameter estimation and polling in nonlinear flood event models. Water Resour Res, 1999, 35: 211–220
28 Gallagher M, Doherty J. Parameter estimation and uncertainty analysis for a watershed model. Environ Modell Softw, 2007, 22: 1000–1020
29 Xiong L H, O'Connor K M. An empirical method to improve the prediction limits of the GLUE methodology in rainfall-runoff modeling. J Hydrol, 2008, 349: 115–124
30 Kuczera G, Parent E. Monte Carlo assessment of parameter uncer-

tainty in conceptual catchment models: the Metropolis algorithm. J Hydrol, 1998, 211: 69–85
31 Yang J, Reichert P, Abbaspour K C. Bayesian uncertainty analysis in distributed hydrological modelling: A case study in the Thur River basin (Switzerland). Water Resour Res, 2007, 43: W10401
32 Vrugt J A, Gupta H V, Bouten W. A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour Res, 2003, 39: 1201–1216
33 Blasone R S, Vrugt J A, Madsen H, et al. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. Adv Water Resour, 2008, 31: 630–648
34 Liang Z M, Li B Q, Yu Z B, et al. Application of Bayesian approach to hydrological frequency analysis. Sci China Tech Sci, 2011, 54: 1183–1192
35 Engeland K, Xu C Y, Gottschalk L. Assessing uncertainties in a conceptual water balance model using Bayesian methodology. Hydrolog Sci J, 2005, 50: 45–63
36 Efron B, Tibshirani R J. An Introduction to the Bootstrap. London: Chapman & Hall, 1993
37 Zucchini W, Adamson P T. Bootstrap confidence intervals for design storms from exceedance series. Hydrolog Sci J, 1989, 34: 41–48
38 Lall U, Sharma A. A nearest neighbor bootstrap for time series resampling. Water Resour Res, 1996, 32: 679–693
39 Srinivas V V, Srinivasan K. Matched block bootstrap for resampling multi-season hydrologic time series. Hydrol Process, 2005, 19: 3659–3682
40 Efron B. Bootstrap methods: another look at the jackknife. Ann Stat, 1979, 7: 1–26
41 Cover K A, Unny T E. Application of computer intensive statistics of parameter uncertainty in streamflow synthesis. J Am Water Resour As, 1986, 22: 495–507
42 Tasker G D, Dunne P. Bootstrap position analysis for forecasting low flow frequency. J Water Res Pl-ASCE, 1997, 123: 359–367
43 Parasuraman K, Elshorbagy A, Si B C. Estimating saturated hydraulic conductivity using genetic programming. Soil Sci Soc Am J, 2007, 71: 1676–1684
44 Li Z L, Shao Q X, Xu Z X, et al. 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, 2010, 385: 76–83
45 Hall P, Horowitz J L, Jing B Y. On blocking rules for the bootstrap with dependent data. Biometrika, 1995, 82: 561–574
46 Zhang L, Lu W X, Yang Q C, et al. Hydrological impacts of climate change on streamflow of Dongliao River watershed in Jilin Province, China. Chinese Geogr Sci, 2012, 22: 522–530
47 Weng B S, Yan D H, Wang H, et al. Study on soil stable infiltration characteristics at different depths in Dongliao River Basin, Northeast China. 2012 International Symposium on Geomatics for Integrated Water Resources Management. Gansu: GIWRM, 2012. 1–5
48 Yang J, Reichert P, Abbaspour K C, et al. Hydrological modelling of the Chaohe basin in China: Statistical model formulation and Bayesian inference. J Hydrol, 2007, 340: 167–182
49 Shen Z Y, Chen L, Chen T. The influence of parameter distribution uncertainty on hydrological and sediment modeling: a case study of SWAT model applied to the Daning watershed of the Three Gorges Reservoir Region, China. Stoch Env Res Risk A, 2013, 27: 235–251
50 Shen Z Y, Hong Q, Yu H, et al. Parameter uncertainty analysis of non-point source pollution from different land use types. Sci Total Environ, 2010, 408: 1971–1978
51 Wu Y P, Liu S G. Automating calibration, sensitivity and uncertainty analysis of complex models using the R package Flexible Modeling Environment (FME): SWAT as an example. Environ Modell Softw, 2012, 31: 99–109
52 Benny S, Murray H. A bootstrap approach to assess parameter uncertainty in simple catchment models. Environ Modell Softw, 2010, 25: 919–926
53 Xie H, Lian Y Q. Uncertainty-based evaluation and comparison of SWAT and HSPF application to the IIlinois River Basin. J Hydrol, 2013, 481: 119–131
54 Yang J, Reichert P, Abbaspour K C, et al. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. J Hydrol, 2008, 358(1-2): 1–23
55 Bates B C, Campbell E P. A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling. Water Resour Res, 2011, 37: 937–947
56 Kavetski D, Kuczera G, Franks S. Bayesian analysis of input uncertainty in hydrological modeling: 2. Application. Water Resour Res, 2006, 42: W03408
57 Ajami N K, Duan Q Y, Sorooshian S. An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resour Res, 2007, 43: W01403