Uncertainty estimation for the Xin'anjiang model parameters

FANGGUI LIU[1,2], XIAOFANG RUI[1,2], HEHAI XIE[1,2], JIE FENG[1,2] & JIYU YU[1,3]

1 *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing 210098, China*
qufuliufanggui@163.com or liufanggui@hhu.edu.cn

2 *College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China*

3 *Shandong Water Conservancy Vacational College, Rizhao, 276826, China*

Abstract Due to the uncertainties in hydrological forecasting mainly derived from hydroclimatic input data, hydrological model structure and model parameters, the investigations on the uncertainty of model parameters are crucial to improve the precision of flood forecasting. The Xin'anjiang (XAJ) model developed by Renjun Zhao is considered an effective conceptual watershed hydrological model, and it has been extensively employed for hydrological modelling problems. In this study, the generalised likelihood uncertainty estimation (GLUE) methodology for XAJ model identification allowing for equifinality was proposed to identify the uncertainty of model parameter sets. Additionally, the GLUE analysis was utilized to determine the Yanduhe catchment, one tributary of the Yangtze River watershed, and the uncertainty of XAJ model parameters. Based on the view of two storm events, we observed *SM* in the XAJ model is very sensitive. For example, its likelihood values displayed the peak value area and its little change will have a large influence on the simulated results; while *K*, *B*, *EX* are not sensitive. Further, we found the observed discharge hydrograph can not be located wholly within the upper and the lower limits of the simulated discharge hydrograph, and some discharge values may fall outside of the 90% uncertainty bounds. This indicates that the XAJ model can not be used to simulate the discharge hydrograph well because of the uncertainty of the model.

Key words uncertainty estimation; GLUE methodology; XAJ model; Yanduhe catchment

INTRODUCTION

Predictions in ungauged basins (PUB) is one of the important and difficult problems in hydrology arousing general concern at home and abroad (Rui *et al.*, 2007). As such, the IAHS decade of PUB (2003–2012) (Sivapalan *et al.*, 2003), focusing on the effective methodology of hydrological simulation to decrease the uncertainty of hydrological prediction and increase its forecasting precision, has been organized, including the establishment of the China PUB Working Groups in 2004 (Yang *et al.*, 2004). According to investigations by Sivapalan *et al.* (2003), Yang *et al.* (2004), etc. the uncertainty of hydrological prediction is mainly caused by the uncertainties of the input data, model structure and model parameters. Therefore, it is definitely necessary to investigate the uncertainty of parameters in hydrological models.

It may be endemic to mechanistic modelling of watershed hydrological systems that there are many different parameter sets within a chosen model structure that may be behavioural or acceptable in reproducing the observed behaviour of that system. This has been called the equifinality concept (Beven & Freer, 2001). The points that produce equifinality include (Rui *et al.*, 2007): (1) The objective function is multivalued; (2) There are numerous interactions among model parameters; (3) Model parameters are stochastic. As such, equifinality makes it uncertain to ultimately find one "optimal" parameter set, so a novel modality should be analysed to fully evaluate such uncertainty. In recent years, although some advances in the uncertainty research of the watershed hydrological model structure and parameter sets have been achieved (Guo *et al.*, 1995; Beven & Freer, 2001), the GLUE methodology of Beven & Binley (1992) is by far one of the most effective uncertainty analysis methods in such a research field (Beven & Binley, 1992; Beven & Freer, 2001).

In this study, the uncertainty of parameters sets in XAJ model was analysed using the GLUE method to investigate the equifinality phenomenon so that the XAJ model can be better applied in hydrological modelling.

XAJ MODEL DESCRIPTION

The XAJ model, first built in 1973, is a conceptual watershed model. Its basic feature is the concept of runoff formation on repletion of storage, which means that runoff is not produced until the soil moisture content of the aeration zone reaches field capacity, and thereafter runoff equals the rainfall excess without further loss. The XAJ model is mainly used for hydrological forecasting in agricultural, forested and pastural land in humid or semi-humid regions (Zhao, 1992).

The schematic diagram of XAJ model is shown in Fig. 1. It demonstrates that the outflow has four major components including the flow routing, the runoff production, the evapotranspiration and the runoff separation.

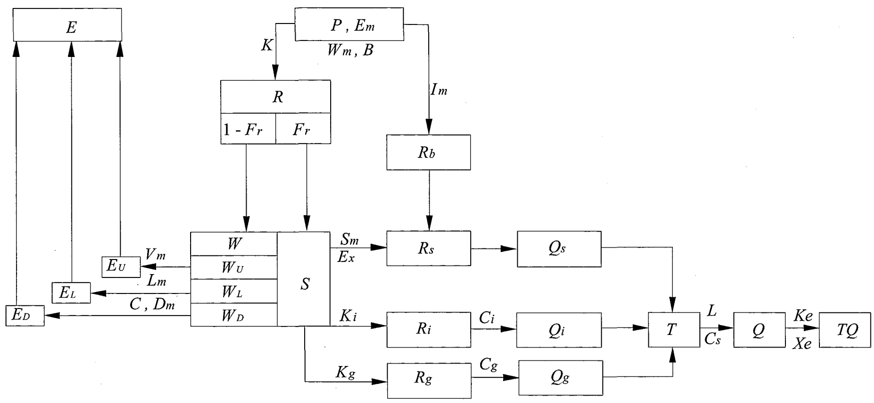

Fig. 1 Flowchart of XAJM (Zhao, 1992).

According to the model structure, the runoff was divided into three components, namely surface runoff, interflow and groundwater flow. The definition of the parameters (Zhao Renjun, 1992) is given in Table 1. Values of a large number of XAJ model parameters cannot be obtained from field measurements and need to be determined through a model calibration procedure. The trial-and-error procedure is used for the optimization of model parameters.

Table 1 Parameters of XAJ model.

notation	definition	notation	definition
K	coefficient for potential evapotranspiration	E_x	exponent of the free water capacity distribution curve
W_U	upper zone tension water capacity	S	mean areal free water storage capacity of the surface soil layer
W_L	lower zone tension water capacity	K_i	daily interflow coefficient
W_D	deep zone tension water capacity	K_g	daily groundwater coefficient
W	average areal tension water capacity	C_i	daily interflow recession coefficient
C	coefficient of deep evapotranspiration	C_g	daily groundwater recession coefficient
I_m	ratio of impervious area to the total area	C_s	recession constant in the “lag and route ” method for routing
B	exponent of the tension water curve	L	corresponding “lag”

THE GLUE METHODOLOGY

Rationale of GLUE

The background of the GLUE methodology has been an attempt to recognize more explicitly the fundamental limitations of hydrological models as simulators of catchment rainfall–runoff processes (Beven, 1989, 1993; Grayson, 1992). One implication of such a recognition is that it should not be assumed that there is one "optimal" model parameter set which can be found to represent a catchment (whether a lumped or distributed representation). For the GLUE methodology, parameter set combinations within the given parameter ranges are calculated using Monte-Carlo simulation. Then the likelihood function is used, and a function value is computed by utilizing both the simulated results and the observed values. As such, once the likelihood weight is computed, the likelihood values of different parameter set combinations are captured. Further, a marginal value is chosen for all of the likelihood values, which of course is a subjective choice to a certain extent. If the likelihood values are smaller than the marginal value, they are specified as zero, since these parameter sets can not reasonably describe the function characteristics of the model. Conversely, if the likelihood values are larger than the marginal value, they are rescaled and sized. So the uncertainty bounds in model prediction below a certain confidence limit are calculated (Beven & Binley, 1992; Beven & Freer, 2001; Xiong & Guo, 2004).

Analysis steps of GLUE

Based on the rationale of the GLUE methodology mentioned above, its estimation steps are described as follows:

(a) The formal definition of the likelihood measure. The presented results employ a coefficient of determination as the basic likelihood measure and basic likelihood measurement is written in the form:

$$L(\theta_i \mid Y) = 1 - \sum_{j=1}^{n} (Q_{ij} - Q_{Oj})^2 \Big/ \sum_{j=1}^{n} (Q_{ij} - \overline{Q_O})^2 \tag{1}$$

in which $L(\theta_i \mid Y)$ is the likelihood measure for the ith parameter set, Q_{ij} the simulated value for jth time step, Q_{Oj} the observed value for jth time step, $\overline{Q}_O$ the average value of the observations and n the sequence number.

(b) Determination of the parameter ranges and an *a priori* distribution. Initially the ranges are assumed as wide as possible based on the rational physical properties. For most of the applications of GLUE *a priori* distribution is replaced by a uniform distribution and uniform random sampling across the specified parameter range is adopted.

(c) Uncertainty estimation of model parameters. Scatter plot of likelihood values for selected model parameters is plotted and uncertainty of model parameters is computed and analysed.

(d) Calculation of uncertainty bounds. In this study, 5% and 95% accumulative likelihood distributions are used as uncertainty limits in the predictions.

CASE STUDY

Description of the Yanduhe catchment

The Yanduhe catchment, located upstream of the Yangtze River watershed, is a first-order tributary of the Yangtze River. It occupies an area of 601 km^2, with the Yanduhe stream gauging station at its outlet. Its exact location is at latitude 31°12'N and longitude 110°18'E. For the catchment, the mean annual rainfall is 1222.2 mm, so it is a typical humid region and has rich precipitation. The soil in the catchment has a strong infiltration capacity and excess rainfall is difficult to generate. Further, as the excess-infiltration runoff is the main mode of runoff generation, the XAJ model can be used to simulate both its runoff-generating process and runoff concentration process. Five rainfall gauges named as Banqiao, Xiagu, Duizi, Songziyuan and Yanduhe, and one evaporation station are located in the catchment.

On the basis of 60 × 60 m DEM data of 1:50 000, the Yanduhe catchment is divided into 540 000 grid cells with 600 rows and 900 columns (see Fig. 2). In this investigation, the distribution curve of the topographic index $\ln(\alpha/\tan\beta)$ is calculated using the method developed by Xie & Huang (2006).

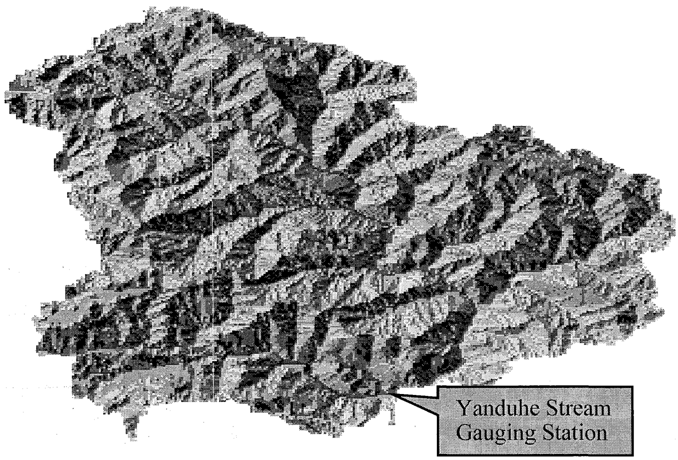

Fig. 2 3-D map of the Yanduhe catchment extracted from the DEM.

THE RESULTS ANALYSIS

The XAJ model is selected as the rainfall–runoff model of the Yanduhe catchment. In this study, only the runoff generation parameters in the XAJ model are considered. The concentration parameters in the XAJ model will be further investigated in the future. Previous studies indicated that *K*, *B*, *SM* and *EX* are very sensitive, and they have large impacts on the simulations. These parameters are therefore used to investigate the uncertainty of the simulated results.

Employing 19810625 and 19860909 storm events of the Yanduhe catchment and the deterministic factor as the likelihood measure, 2 × 5000 parameter sets were chosen to estimate the uncertainty based on uniform random sampling within the specified parameter ranges (see Table 2). The scatter plot of likelihood values for four selected XAJ model parameters of the Yanduhe catchment are displayed in Figs 3 and 4.

Table 2 *a priori* distribution of parameters.

Parameters	Minimum value	Maximum value	Mean value
K	0.4	1.0	0.7
B	0.1	0.4	0.25
SM	20 (mm)	110 (mm)	65 (mm)
EX	1.0	1.7	1.35

According to Figs 3 and 4, it is found that the value ranges of *SM* reduce greatly after its rating through the discharge hydrograph at the catchment outlet. It is also noted that *SM* shows peak value ranges, while *K*, *B*, *EX* are not sensitive at all. We also observed that high likelihood values exist in the whole parameter distribution range. In addition, many parameter value sets are displayed in high likelihood value ranges, namely equifinality, indicating that there are many equivalent parameter sets.

In this study, for parameter sets' deterministic factor is greater than zero, their likelihood values are rescaled and sized. In the following, the 90% uncertainty bounds for simulations of XAJ model, including the observations and the upper bound and the lower bound of uncertainty, are calculated and shown in Figs 5 and 6. The uncertainty bounds change with the discharge process, due to its larger values in the higher discharge zone and fewer values in the lower discharge zone. The simulated discharge limits can not totally encompass the observed discharge hydrograph,

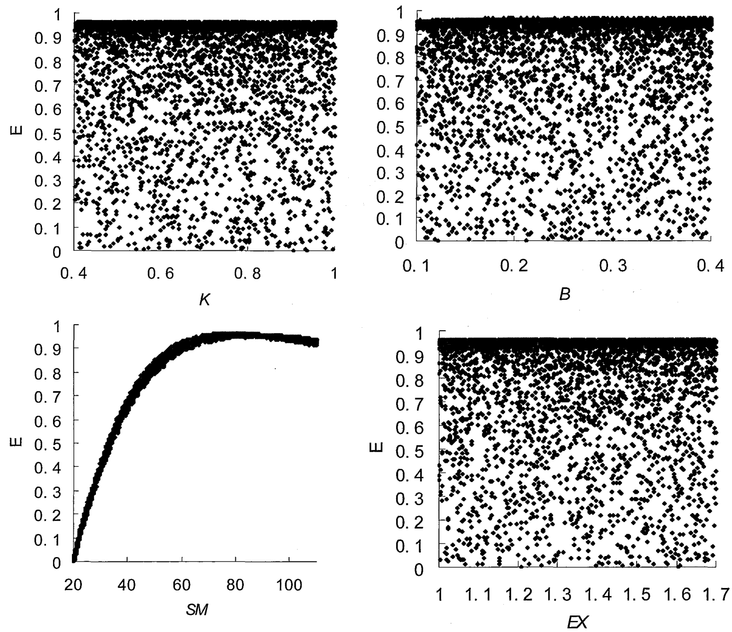

Fig. 3 Scatter plot of likelihood values for four selected parameters of the Yanduhe catchment for the 25 June 1981 storm event..

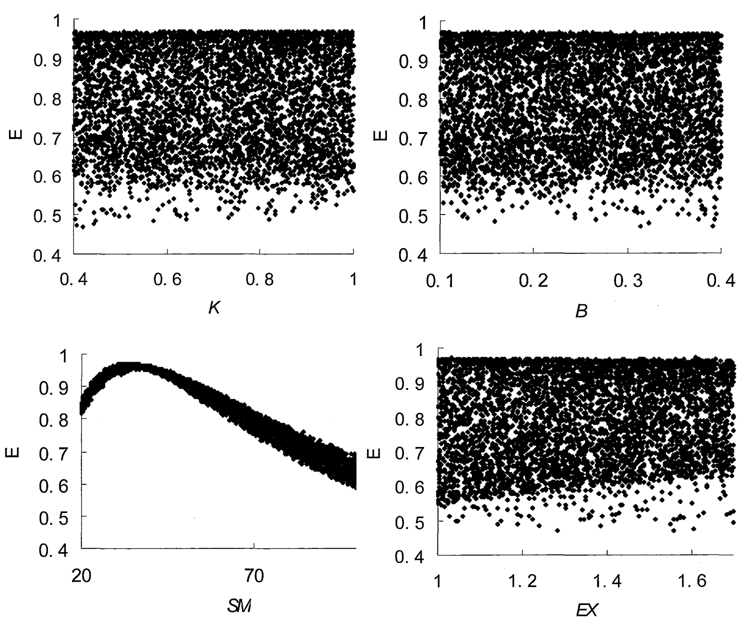

Fig. 4 Scatter plot of likelihood values for four selected parameters of the Yanduhe catchment for the 9 September 1986 storm event.

because some discharge values always fall outside of the 90% confidence interval. This shows that XAJ model can not simulate the discharge hydrograph well due to the uncertainty of the model.

The 90% uncertainty bounds are not wide enough to encompass all of the observed discharges during the calibration period and, the possibilities are listed as follows:

(a) The influence of *a priori* parameter distribution. The uniform distribution is simply used to generate model parameters. However, *a priori* distribution of model parameters under study is not well determined.
(b) The influence of the sampling method for Monte Carlo simulation. Although Monte Carlo simulation may overcome some disadvantages of automatic scan, random scan and trial-and-error scan in higher dimensional parameter space of hydrological models, many combinations of parameters for the model structure with multi parameters require several ten thousand or tens of thousands, even up to more than one million parameter samplings. The computational efficiency is very low. For simulation of the flood process with a long duration, it is time consuming with large samplings so it is better if the samplings are not too large.
(c) The influence of hydrological data. The rainfall interpolation time interval is specified as 1 h for this catchment, while the time interval for the initial step is longer. If the rainfall is divided into the same interpolation time interval, the simulated discharge hydrograph does not agree well with the observed discharge hydrograph.

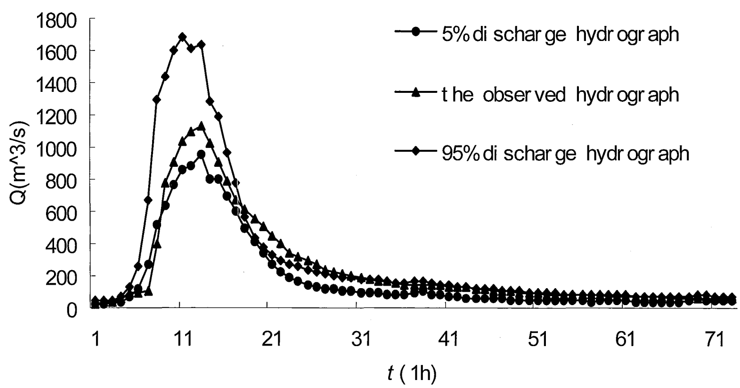

Fig. 5 Ninety percent uncertainty bounds for simulations for the 25 June 1981 storm event in the Yanduhe catchment

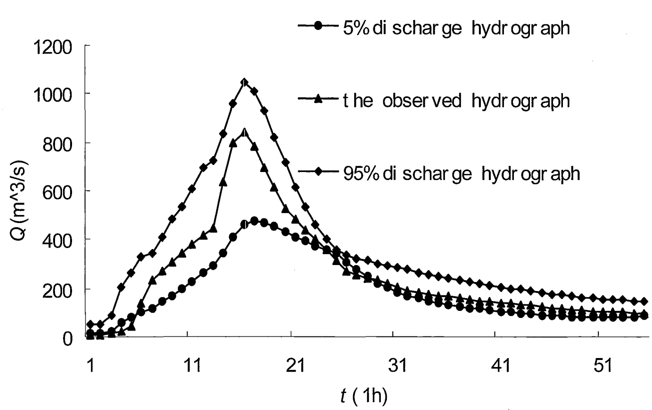

Fig. 6 Ninety percent uncertainty bounds simulations for the 9 September 1986 storm event in the Yanduhe catchment.

DISCUSSION AND SUMMARY

When uncertainty and equifinality are concisely defined, the rationale and the analysis steps of the GLUE methodology are proposed and the methodology is applied to the Yanduhe catchment, located in the Yangtze River watershed. Additionally, in our computation *a priori* distribution of XAJ model parameters is specified via estimation, and the code for XAJ model algorithm was run with 2 × 5000 randomly generated parameter sets. It is noted that the equifinality is easy to get by. *SM* in XAJ model is very sensitive, it shows peak value ranges, while *K*, *B*, *EX* are not sensitive at all. We also observed that high likelihood values exist in the whole parameter distribution range. According to assumed rules, the upper and the lower limits of 90% uncertainty bounds are calculated. Given the parameter distribution and the parameter space, the observed discharge hydrograph is not able to totally fall inside the upper and the lower limits, which also demonstrates that the parameter ranges can not encompass all the values. As a watershed hydrological model is a complex system, the influencing factors causing the uncertainty should be investigated in future. Additionally, using the GLUE methodology is easy to find the equifinality phenomenon, but how to figure out the equifinality problem by using it should be further studied in future.

Acknowledgements This project was supported by the National Natural Science Foundation of China (50609005), the Fok Ying-Tong Education Foundation for Young Teachers in the Higher Education Institutions of China (101075) and the Specialized Research Fund for the Doctoral Program of Higher Education of China (20050294002).

REFERENCES

Beven, K. J. (1989) Changing ideas in hydrology: The case of physically based models. *J. Hydrol.* **105**, 157–172.

Beven, K. J. (1993) Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv. Water Resour.* **16**, 41–51.

Beven, K. J. & Binley, A. M. (1992) The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Processes* **6**(2), 279–298.

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**, 11–29.

Grayson, R. B, Moore, I. D. & McMahon, T. A. (1992) Physically based hydrologic modelling, 2, Is the concept realistic? *Water Resour. Res.* **28**, 2659–2666.

Guo, Shenglian, Li, Lan & Zeng, Guangming (1995) Uncertainty analysis of impact of climate change on hydrology and water resources. *Hydrology* **6**, 1–5.

Rui, Xiaofang, Liu, Fanggui & Xing Zhengxiang (2007) Advances in hydrology and some frontier prolems. *Adv. Sci. & Tech. Water Resour.* **27**(1), 75–79.

Sivapalan, M. K. K., Takeuchi, S. W. & Franks, S. W. (2003) IAHS decade of prediction in ungauged basins (PUB), 2003–2012: shaping an exiting future for the hydrological sciences. *Hydrol. Sci. J.* **48**(6), 857–879.

Xie, Hehai & Huang, Guoru (2006) Some calculation methods for topographic index. *J. Hohai University (Natural Sciences)* **34**(1), 46–50.

Xiong, Lihua & Guo, Shenglian (2004) Research on equifinality of the Xin'anjiang Model of three components. In: *Investigations and Advances in Complexity and Uncertainty of Water Problems* (ed. by Xia Jun). China Waterpower Press, Beijing, China.

Yang, Dawen, Xia, Jun & Zhang, Jianyun (2004) Investigations and advances on China PUB. In: *Investigations and Advances in Complexity and Uncertainty of Water Problems* (ed. by Xia Jun), 47–54. China Waterpower Press, Beijing, China.

Zhao, Renjun (1992) The Xin'anjiang model applied in China. *J. Hydrol.* **135**, 371–381.