Assessment of hydrological model structure based on parameter identifiability

HONGLI ZHAO, XIAOLIU YANG & YUNZHONG JIANG

Department of Water Resources, Institute of Water Resources and Hydropower Research, 100044, China
zhaohl@iwhr.com

Abstract The uncertainty in structures of hydrological models results in uncertainties in estimation of parameter values, and simulation or prediction results of models. One of the most direct ways to reduce model uncertainty is to improve model structure. This improvement has traditionally been done, dependant on modeller's intuitions and experiences, not quantitative indexes. Parameters are always set to correspond to model components. The identifiability of parameters is intimately related to the rationality of model structure. Based on the relationship between them, this paper presents an index of parameter identifiability as a measurement of the rationality of model structure. The measurement can be used as the guiding information for adjusting model structure. The GRJ model is selected here as the case study. Parameters identifiabilities of different versions of GRJ models are calculated as the measurements of evaluating model structures. Finally, some suggestions for improving model structures are presented based on the analysis result.

Key words hydrological model; model assessment; uncertainty; parameter identifiability; GRJ model

INTRODUCTION

Hydrological models are the basic tools for research into water resources and the water environment. Hydrologists have devoted most of their time to improving the precision of models in the past two decades. Their research means that model structures are getting more and more complicated. Meanwhile, in order to solve the problems due to correlativity among model parameters in multidimensional space, more and more emphasis is put on the research of efficiency and precision of optimum algorithms of model parameter, but the efforts were not satisfied in many applications. Compared to the uncertainty of data observed and model parameters, the model structure achieved from the existing knowledge system is the biggest source of uncertainty of hydrological model (Beck, 1987; Liu *et al.*, 2002). The structure identification is a very difficult process. Sometimes models with different structures may yield equally good results. This ambiguity has serious impacts on parameter calibration and predictive accuracy, and therefore limits the application of hydrological models, e.g. for the simulation of land use or climate change scenarios, or regionalization studies.

To address the credit crisis of model calculation results due to uncertainty, two reactions were found in hydrological literature. The first is the increased use of parsimonious model structures (e.g. Jakeman & Hornberger, 1993; Young *et al.*, 1996; Wagener *et al.*, 2002), i.e. structures only containing those parameters, and therefore model components, that can be identified from observed output. Some hydrologists even considered that the general hydrological observation can only satisfy the identification of 3–5 model parameters (Beven, 1989; Jakeman & Horberger, 1993). However, the increase in identifiability of model parameter is at the cost of the decrease in the number of separate hydrological processes that can be described by the model. There is therefore a danger of building a model structure that is too simplistic for the anticipated purpose. The second reaction is the search for better calibration methods to mine more information hidden in time series of observed discharge or groundwater level, etc. Some researchers suggested that multi-objective calibration methods can help to retrieve more information (DYNIA, Wagener *et al.*, 2003).

The most primary solution for model uncertainty is the adjustment of model structure, since the parameters' uncertainty largely results from model structure. Traditionally, the adjustment of model structure is implemented based on intuition and experience of the modeller, which is not an objective criterion. Reducing the uncertainty of the calculation result should be considered as an important aim of structure adjustment. Therefore, it is necessary to quantify the structural causes of model uncertainty and supply hints for model structure adjustment.

This paper makes use of the Regional Sensitivity Analysis (RSA, Spear & Hornberger, 1980; Hornberger & Spear, 1981), the Generalized Likelihood Uncertainty Estimation framework (GLUE, Beven & Binley, 1992), the Dynamic Identifiability Analysis (DYNIA, Wagener *et al.*, 2003) and constructs the assessment criterion of structure rationality and reliability using the measure of information content to help improve the model structure. For the goal that orientates model development in the right direction, this paper focuses on structure assessment for different versions of one model, instead of structure evaluation for different models.

MODEL STRUCTURE ASSESSMENT AND PARAMETER IDENTIFICATION

Hydrologists construct models to describe the behaviour of the real world based on their perception. Due to the limit of human perceiving capability, simplification and generalization are adopted to deal with the parts beyond perception or difficult to describe. It causes the difference between the result and real world behaviour. The enhancement of abilities of perceiving and describing model structure could definitely be improved. The review of hydrological model development indicates that the update period of the latter is always shorter than the former. Therefore, there are some invalid modifications in the "improvement" process. To keep the model development in a healthy and sustainable way, these invalid modifications should be identified.

The ultimate objective of model building is to supply information and remove doubt in the hydrological process step by step. From the view of uncertainty, the process of identification and improvement of model structure is just about the process of reducing uncertainty. So model structure should be modified following the guidance of the assessment of whether or not the structure adjustment reduces the uncertainty of simulation or forecast.

Most uncertainty analysis methods belong to the statistics category, which is a kind of data, and information analysis based on the theory of probability and statistics. Informatics deems that statistics methods can make up the deficiency of mathematics and physics methods. Most research on hydrological uncertainty focus on parameter uncertainty and study to find out the most appropriate parameter set corresponding to some structure, such as the popular GLUE. However, compared to model parameters, the imperfect model structure is more likely a primary cause for model uncertainty. Therefore, it will be more efficient to reduce model uncertainty through structure adjustment.

A model is the combination of structure and parameters. Usually there is an intimate relationship between parameters and structure components. Generally, the unidentifiability of parameters resulted from over-simplification or too much detail to be supported by practical data condition. The parameter uncertainty is the reflection of structure illogicality. So the faultiness of a model structure can be found out using the analysis methods of parameter uncertainty.

PARAMETER INDENTIFIABILITY

The parameter indentifiability can be defined as the confirmation degree of a parameter based on the information which is available and can be validated in practice. The basic steps in the calculation procedure are as follows:

- Set the initial range and prior distribution of parameters on the basis of physical argument or experience. A methodology for sampling the parameter space is required to get enough parameter sets. In most of the analysis methods of parameter uncertainty this has been done by Monte Carlo simulation, using uniform random sampling across the specified parameter range.
- Define the likelihood measure and criteria for acceptance or rejection of parameter sets. The measure describes the coincident degree between simulation and observation with a special parameter set. The measure value will go up monotonously following the increase of coincident degree. The Nash and Sutcliffe efficiency criterion has been chosen as the basic likelihood measure by many studies. Its form is as follows:

$$L(\underline{\theta}_i \mid \underline{Y}) = 1 - \frac{\sigma_i^2}{\sigma_o^2} \tag{1}$$

where $L(\underline{\theta}_i \mid \underline{Y})$ is the likelihood measure for the *i*th parameter set conditional on the observation, σ_i^2 is the associated error variance for the *i*th parameter set, σ_o^2 is the observed variance for the period under consideration.

Calculate the likelihood measure for each parameter set obtained from step 1; reject the parameter sets whose likelihood measures are less than the criterion set in step 2 by setting their likelihood measure as 0.

Following rejection, the likelihood measures associated with the retained parameter sets can be rescaled to give a cumulative sum of 1. Sort by the likelihood measure for each parameter and get their 90% confidence limit.

The identifiability = 1 – the width of 90% confidence limits

The lower parameter identifiability means that the corresponding structure component is difficult to express accurately, can not work actively and should be considered to adjust.

GRJ MODELS

The model selected for this study is the GRJ model series. The parameter identifiabilities of different versions of GRJ models are analysed and the value of structure adjustment in their development progress is assessed.

The GRJ model is a lumped daily rainfall–runoff model developed by Edijatno *et al.* (1999). The original intention of the model development is to build a representation for the process of rainfall–runoff in a manner as parsimonious as possible. The model should have the ability of self-adopting to guarantee the accuracy in different applications. The trace of model development passes through several stages, which have 1, 2, 3, and 4 parameters, accordingly named GR1J, GR2J, GR3J, and GR4J models. The only discussion about this model is the value of structure adjustment from GR3J to GR4J with the measure of parameter identifiability.

The GRJ model series belong to the family of soil moisture accounting models. The presentation of GR3J and GR4J are provided in Figs 1 and 2.

$$Ps = \frac{P_n\left(1-\left(\frac{S}{A}\right)^2\right)}{1+\frac{P_n}{A}\left(\frac{S}{A}\right)} \qquad Es = \frac{E_n\frac{S}{A}\left(2-\frac{S}{A}\right)}{1+\frac{E_n}{A}\left(2-\frac{S}{A}\right)}$$

$$0<t<X3, \qquad SH1(t) = \left(\frac{t}{X3}\right)^3$$

$$t \ge X3, \qquad SH1(t) = 1$$

$$0<t\le X3, \qquad SH2(t) = \frac{1}{2}\left(\frac{t}{X3}\right)^3$$

$$X3<t<2.X3, \qquad SH2(t) = 1-\frac{1}{2}\left(2-\frac{t}{X3}\right)^3$$

$$t \ge 2.X3, \qquad SH2(t) = 1$$

$$F = X1.\left(\frac{R}{X2}\right)^4 \qquad R = \max\,(0\,; R + Q9 + F)$$

$$Qr = R.\left\{1-\left[1+\left(\frac{R}{X2}\right)^4\right]^{-\frac{1}{4}}\right\}$$

$$R = R - Qr \qquad Q = Qr + Qd$$

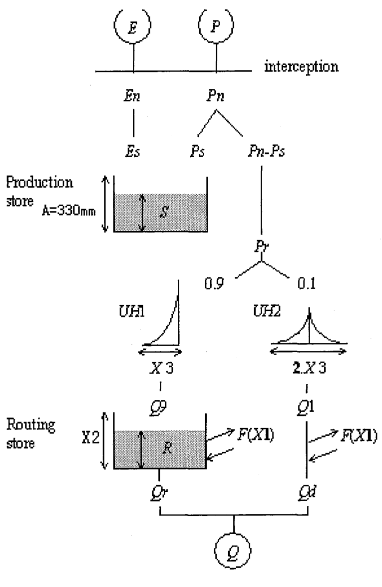

Fig. 1 Scheme of the GR3J rainfall–flow model.

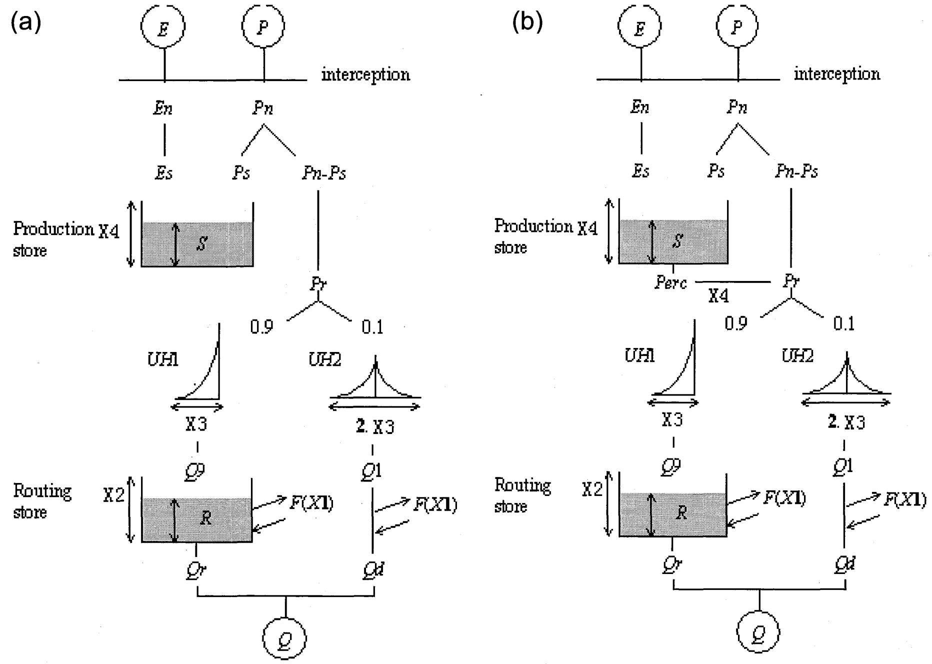

Fig. 2 Scheme of the GR4J rainfall-flow model. (a) GR4J-a model; (b) GR4J-b model.

A summary description of the GR3J model is shown in Fig. 1, as proposed by Edijatno *et al.* (1999). The three parameters in GR3J represent groundwater exchange coefficient (X1), one day ahead maximum capacity of the routing store (X2), and time base of unit hydrograph UH1(X3), respectively. P, the rainfall depth and E, the potential evapotranspiration, are inputs to the model. Pn is net rainfall, En is net evapotranspiration capacity and Ps is the part of Pn that fills the production store. S is the level of the production store, whose maximum capacity A is a fixed parameter with a value of 330 mm. The total quantity Pr of water that reaches the routing functions is divided into two flow components according to a fixed split: 90% of Pr is routed by a unit hydrograph $UH1$ and then a nonlinear routing store, and the remaining 10% of Pr is routed by a single unit hydrograph $UH2$. $SH1$ and $SH2$ are denoted as the ordinates of both unit hydrographs $UH1$ and $UH2$, respectively. F is a groundwater exchange that acts on both flow components. R is the level in the routing store and Qr is the outflow of the reservoir. Q is the total streamflow.

Compared with GR3J, the fixed maximum capacity A of the production store is replaced by parameter X4 in GR4J-a, as shown in Fig. 2(a), a percolation leakage *Perc* from the production store is introduced into GR4J-b in addition, as shown in Fig. 2(b). The change in computation equations from GR3J to GR4J is as follows:

$$Ps = \frac{X4.\left(1-\left(\frac{S}{X4}\right)^2\right).\tanh\left(\frac{Pn}{X4}\right)}{1+\frac{S}{X4}.\tanh\left(\frac{Pn}{X4}\right)} \tag{2}$$

$$Es = \frac{S.\left(2-\frac{S}{X4}\right).\tanh\left(\frac{En}{X4}\right)}{1+\left(1-\frac{S}{X4}\right).\tanh\left(\frac{En}{X4}\right)} \tag{3}$$

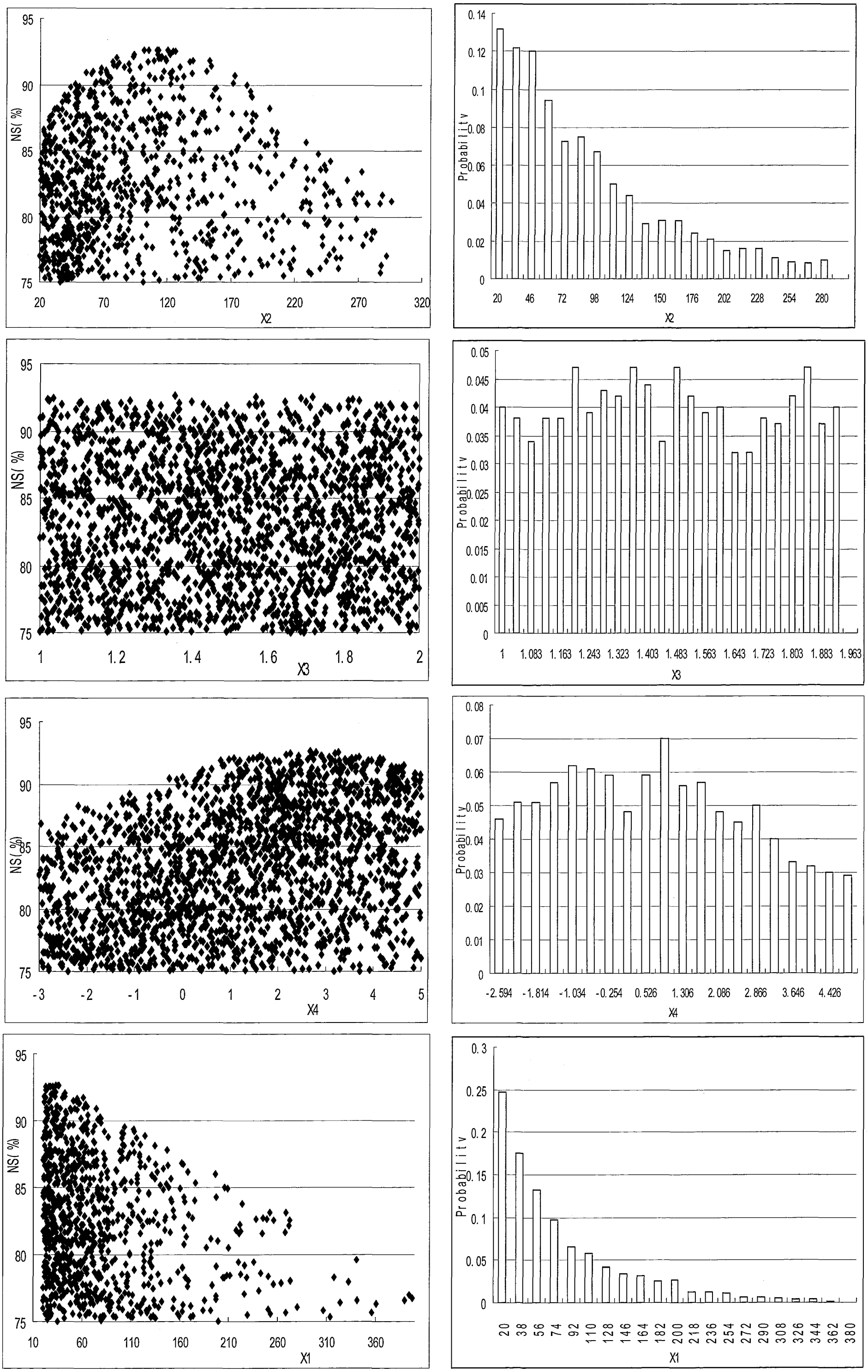

Fig. 3 Parameter distribution and model efficiency of GR4J-b.

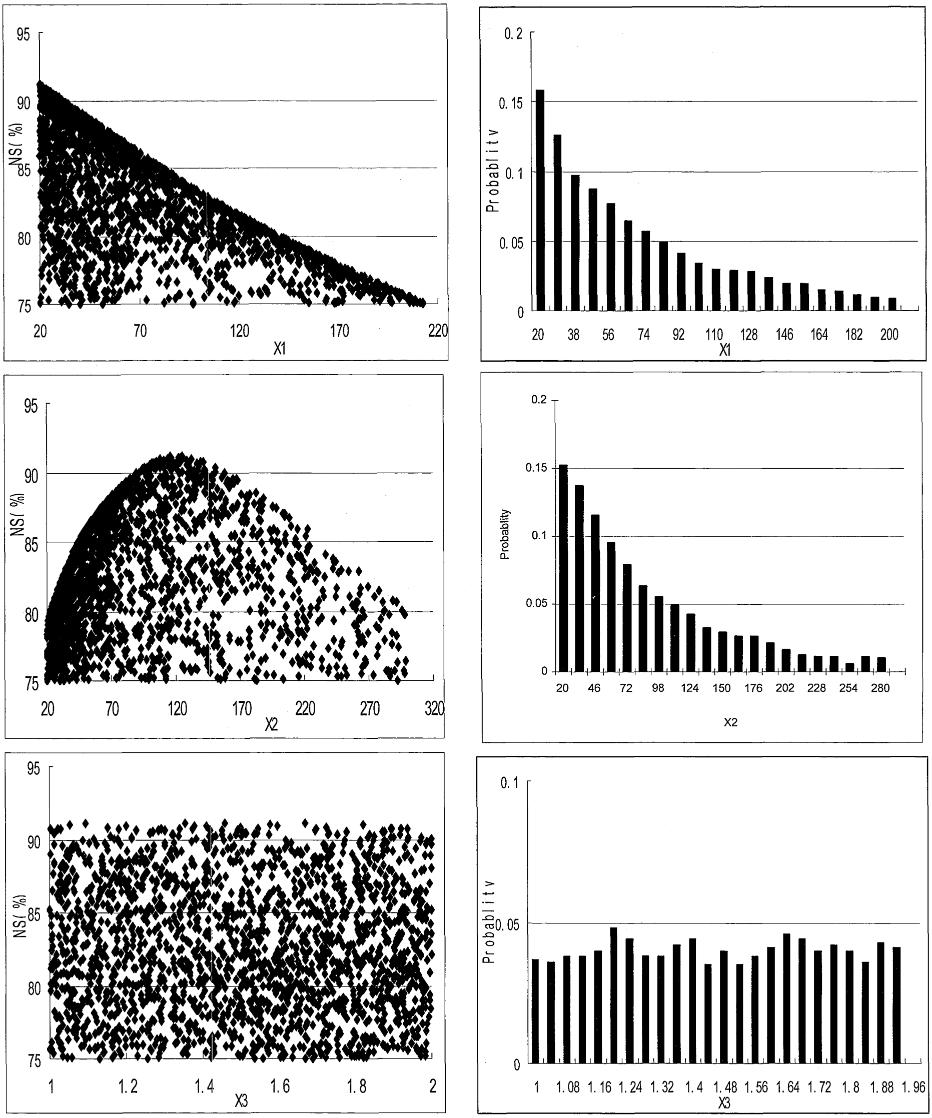

Fig. 4 Parameter distribution and model efficiency of GR3J.

$$Perc = S.\left\{1-\left[1+\left(\frac{4}{9}\frac{S}{X4}\right)^{4}\right]^{-\frac{1}{4}}\right\} \qquad (4)$$

PARAMETERS INDENTIFIABILITY OF GRJ MODELS

With a daily time step, a case study was carried out in Nandian catchment, with an area of 765 km^2, located in the Taizihe River watershed, tributary of Liaohe River, northeast China. There are seven raingauges in this catchment. The daily observed time series of rainfall, evaporation and discharge from 1979 to 1987 were selected for analysis of parameter identifiability of GRJ models. The average annual precipitation in the period reaches 949 mm.

According to experience in parameters setting of GR3J and GR4J in the literature, the initial ranges of parameters were set as follows:

*X*1: –3~5
*X*2: 20~320
*X*3: 1~2
*X*4: 20~120

The Monte Carlo simulation approach using uniform random sampling is applied to the analysis of this case. The Nash and Sutcliffe efficiency (NSE) criterion is chosen as the likelihood measure. The criterion for acceptance or rejection is 75% of NSE value. The parameter sets whose NSE are higher than 75% are called behavioural. The distribution of behavioural parameters and their corresponding NSE, the probability distribution histogram of behavioural parameters of GR3J and GR4J-b, are shown in Figs 3 and 4.

The parameters identifiabilities of GR3J and GR4J models calculated using the method mentioned above are listed in Table 1.

As showed in Table 1, compared with GR3J, the parameters identifiabilities of GR4J have been generally improved. Furthermore, with the additional description of percolation leakage from the production store, the parameter identifiability of *X*4, which corresponds to the component of production reservoir in the model structure, has evidently increased from 0.381 to 0.55. It means that the part of structure change is reasonable, since it results in a lower uncertainty of its parameter. However, it is found that the identifiability of parameter *X*3 in GR4J-b with a more complicated structure is lower than the identifiability of it in GR4J-a.

Table 1 Parameters identifiabilities of three versions of GRJ.

Model	Max NS	*X*1	*X*2	*X*3	*X*4
GR3J	0.92	0.107	0.382	0.077	
GR4J-a	0.92	0.155	0.428	0.125	0.381
GR4J-b	0.93	0.158	0.428	0.083	0.55

In all versions of GRJ models, the identifiability of parameter *X*3 is always low. The structure it describes is the routing progress using a unit hydrograph. This is the most extraordinary design in the model: two unit hydrographs parsimoniously depend on the same time parameter *X*3 expressed in days. However, the measure of parameter identifiability indicates it is difficult to get a satisfactory value for this parameter as its uncertainty is huge. This means that this part of the structure should be adjusted in the future.

CONCLUSIONS AND DISCUSSIONS

For watershed hydrological simulation or prediction, there is not a most optimized model structure or parameter set, which results in the uncertainty of calculation results. The uncertainty is intimately related with model structure. Neither parsimonious nor complicated approaches can be treated as the assessment criteria of structure improvement. The parameter identifiability reveals the rationality of a model structure under the given data condition and it is instructive information for structure adjustment. A measure index of parameter identifiability is presented in this paper and used for assessment of different versions of GRJ model structures. The improvement degree in performance of structure adjustment is demonstrated with this measure and the existing problems in the present version are suggested.

The identification of model structure failure is relatively straightforward and objective. However, the analysis of why a failure occurred and how the model structure can be improved very much depends on the experience and creativity of the modeller himself. As Beck (1985) pointed out: “there is no systematic “algorithm” for changing an inadequate structure that is equivalent to increasing a polynomial order from *n*, say, to ($n + 1$), as would be possible for a class III (data-based) model structure”. The modeller’s task is to draw an inference from the type of failure that has occurred with respect to the hypothesis underlying the specific model component in order to develop an improved version. Obviously, the adjustment of model structure only based on the measure of parameter identifiability is not enough. Many efforts on selecting more information about structure failure have to be done in the next steps.

Acknowledgements The authors would like to thank the data support from Prof. X. M. Xie of IWHR, China. Thanks are also due to the reviewers for their valuable comments.

REFERENCES

Beck, M. B. (1985) Structure, failure, inferences, and prediction. In: *Identification and System Parameter Estimation* (ed. by M. A. Baker & P. C. Young) , 1443–1448. (Proc. 7th Symposium Vol. 2, July 1985), IFAC/IFORS, York, UK.

Beck, M. B. (1987) Water quality modeling: a review of the analysis of uncertainty. *Water Resour. Res.* **23**(8), 1393–1442.

Beven, K. J. (1989) Change ideas in hydrology – the case of physically-based models. *J. Hydrol.* **23**, 1393–1442.

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

Edijatno, N. O., Yang, X. L., Makhlouf, Z. & Michel, C. (1999) GR3J: a daily watershed model with three free parameters. *Hydrol. Sci. J.* **44**(2).

Hornberger, G. M, & Spear, R. C. (1981) An approach to the preliminary analysis of environmental system. *J. Environ. Manage.* **12**, 7–18.

Jakeman, A. J. & Hornberger, G. M. (1993) How complexity is warranted in a rainfall–runoff model? *Water Resour. Res.* **29**, 2637–2649.

Liu, Yi, Chen, Jining & Du, Pengfei, (2002) Parameters identification and uncertainty analysis for environment model. *Environ. Sci.* **23**(6), 6–10 (in Chinese).

Mo, Xingguo & Liu, Suxia (2004) The application of GLUE on hydrological uncertainty analysis, In: *Advances in Research of Complexity and Uncertainty of Water Problem.* China Water publishing Press, China (in Chinese).

Perrin, C., Michel, C. & Andréassian, V. (2003) Improvement of a parsimonious model for streamflow simulation. *J. Hydrol.* **279**, 275–289.

Spear, R. C. & Hornberger, G. M. (1980) Eutrophication in Peel Inlet, II, identification of critical via generalised sensitivity analysis. *Water Resour. Res.* **14**, 43–49.

Wagener, T. Camacho, L A. Lees, M. J. & Wheater, H. S. (2001) Dynamic parameter identifiability of a solute transport model. In: *Advances in Design Sciences and Technology* (ed. by R. Beheshti), 251–264. (Proc. Europa’8 Delft, April).

Wagener, T., Mclntyre, N., Lees, M. J., Wheater, H. S. & Gupter, H. V. (2003) Towards reduced uncertainty in conceptual rainfall–runoff modeling: dynamic identifiability analysis. *Hydrol. Processes* **17**, 455–476.

Yang, X. L. & Michel, C. (2000) Flood forecast with a watershed model: a new method of parameter updating. *Hydrol. Sci. J.* **45**(4).

Young, P. C., Parkinson, S. & Lees, M. J. (1996) Simplicity out of complexity in environmental modeling: Occham’s razor revisited. *J. App. Statistics* **23**, 165–210.

Zhang, Xuewen (1981) *The Information Analysis of Weather Forecast.* Science Publishing Press, China (in Chinese).