A comparison of simple and complex versions of the Xinanjiang hydrological model in predicting runoff in ungauged basins

Peng Bai, Xiaomang Liu, Kang Liang, Xiaojie Liu and Changming Liu

ABSTRACT

There are different views on the selection of hydrological model structural complexity for streamflow prediction in ungauged basins. Some studies suggest that complex models are better than simple models due to the former's prediction capability; whereas some studies favor parsimonious model structures to overcome a risk of over-parameterization. The Xinanjiang (XAJ) model, the most widely used hydrological model in China, has two different versions, as follows: (1) the simple version with seven parameters (XAJ7) and (2) the complex version with 14 parameters (XAJ14). In this study, the two versions of the XAJ model were comprehensively evaluated for streamflow prediction in ungauged basins based on their efficiency, parameter identifiability, and independence. The results showed that the complex XAJ14 model was more flexible than the simple XAJ7 in calibration mode; while the two models have similar performance in validation and regionalization modes. Lack of parameter identifiability and the presence of parameter interdependence most likely explain why the complex XAJ14 cannot consistently outperform the XAJ7 in different modes. Therefore, the simple XAJ7 is a better choice than XAJ14 for streamflow prediction in ungauged basins.

Key words | hydrological modeling, model parsimony, regionalization, ungauged basin, Xinanjiang model

Peng Bai
Xiaomang Liu (corresponding author)
Kang Liang
Xiaojie Liu
Changming Liu
Key Laboratory of Water Cycle & Related Land Surface Process,
Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences,
Beijing 100101,
China
E-mail: *liaoxfxm@163.com*

INTRODUCTION

Streamflow observation records are important for hydrological research and practices, such as flooding forecasting, hydraulic engineering design, drought risk assessment, and water resources planning and management. However, streamflow observations are not always available in many parts of the world, and the publicly available streamflow records are often incomplete or very brief (Sivapalan 2003; Wagener *et al.* 2004a). Missing or incomplete streamflow observations have greatly impeded the operation of hydrological research and practices. Some methods can be used to reconstruct the streamflow time series, including statistic-based methods, artificial intelligence-based methods, and process-based hydrological models (Wagener *et al.* 2004b; Besaw *et al.* 2010; Benito 2012). However, these methods often become invalid, unable to reproduce streamflow given the limited observed data for calibration. Consequently, reliable streamflow prediction in ungauged basins is a great challenge for hydrological research (Sivapalan *et al.* 2003; Seibert & Beven 2009). The International Association of Hydrological Sciences (IAHS) launched the Scientific Decade of IAHS (2003–2012), entitled *Predictions in Ungauged Basins (PUB)*, which focuses on improving scientific understanding and simulation of hydrological processes in ungauged basins (Sivapalan *et al.* 2003; Blöschl *et al.* 2013; Hrachowitz *et al.* 2013).

Several methods have been proposed for estimating streamflow in ungauged basins, and most of them involve the use of hydrological models (McIntyre *et al.* 2005; Castiglioni *et al.* 2010; Ahiablame *et al.* 2012; Booker & Snelder 2012; El-Hames 2012). Parameter estimation is a key step of

doi: 10.2166/nh.2016.094

hydrological modeling, which largely determines the accuracy of streamflow simulation and prediction (Xu 2003; Lee et al. 2005; McIntyre et al. 2005). Due to high heterogeneity in landscape properties, hydrological model parameters cannot be measured directly at the catchment scale, which usually are inferred by a calibration process (Xu 1999; Vaché & McDonnell 2006; Jin et al. 2009). However, parameter calibration in ungauged basins cannot be directly performed because of the lack of observations. An alternative strategy is to transfer model parameters from gauged basins to ungauged basins, i.e., parameter regionalization (Blöschl 2006; Zhang & Chiew 2009; Pechlivanidis et al. 2010; Kizza et al. 2013). An overview of the parameter regionalization methods has been presented by He et al. (2011) and Parajka et al. (2013).

Model complexity is also an important criterion of model selection in addition to the goodness-of-fit. The complexity of hydrological models, which is commonly quantified by the number of model parameters, has significant impacts on parameter regionalization in ungauged basins (Chiew et al. 1993; Perrin et al. 2001). The calibration and parameter regionalization of the hydrological model is not straightforward as the number of model parameters increases (McCabe et al. 2005). It is generally expected that a model with a high degree of freedom (complex model) will perform better in the calibration period in comparison to a simple model (due to higher degrees of freedom). However, this is not always the case in the 'untrained' validation period (Wheater et al. 1993; Wagener et al. 2001b; Hailegeorgis & Alfredsen 2015). Some studies have suggested the model users avoid using complex models to estimate streamflow in ungauged basins (Jakeman & Hornberger 1993; Lee et al. 2005; Bárdossy 2007; Skaugen et al. 2015). A complex model may result in model over-parameterization, which can bring a large degree of uncertainty in streamflow prediction (Jakeman & Hornberger 1993; Bárdossy 2007; Skaugen et al. 2015). Nevertheless, conclusions from other studies were in favor of complex models (Vaché & McDonnell 2006; Li et al. 2015).

The Xinanjiang (XAJ) model, as the most widely used hydrological model in China, has two different versions, as follows: (1) the simple version with seven parameters (XAJ7) and (2) the complex version with 14 parameters (XAJ14). The main objective of this study is to evaluate which of these two models, XAJ7 and XAJ14, is more suitable for streamflow prediction in ungauged basins. In this study, model efficiency is not the only criterion of model evaluation, and parameter identifiability and independence are also considered. This paper is structured as follows: the section below describes the two XAJ model versions, study area and data sources; then the next section introduces methods used for model evaluation; followed by a section focusing on the results and discussion; and, finally the summary and conclusions are presented.

XAJ MODEL DESCRIPTION AND STUDY AREA

XAJ model description

The XAJ14 model has more extensive applications than the XAJ7 model in China, and it has been employed in the China National Flood Forecasting System (WMO 2011; Yao et al. 2014). The XAJ14 model contains the following four modules: runoff generation, three-layer evapotranspiration (*ET*), separation of runoff components, and runoff routing (Table 1). Generally, *ET* occurs in the top three soil layers; the actual evapotranspiration (*AET*) is estimated as a function of potential evapotranspiration (*PET*) and available soil moisture. *AET* first occurs in the upper layer at potential rate until the water storage is exhausted. Then, the water storage in the lower layer begins to supply for *AET*. *AET* occurs in the deepest layer of soil only when the lower storage layer is reduced to a proportion of storage capacity. A similar mechanism is also used to describe the soil moisture repletion process among the three soil layers. The XAJ14 model uses a single parabolic curve to describe the spatial heterogeneity of the soil moisture storage capacity and assumes that the runoff is not produced until the soil moisture storage reaches field capacity (Zhao 1992; Cheng et al. 2006). The generated runoff is separated into the following three components: surface runoff, interflow, and groundwater according to different free water storage structures. The surface runoff directly flows into the river, and the interflow and groundwater is released slowly to river channels through a single linear reservoir. Finally, the Muskingum routing equation is adopted to calculate the discharge at the watershed outlet.

The XAJ7 model has a similar model structure to XAJ14 (Figure 1), and it can be regarded as the simplified version of the XAJ14 model (Zhao et al. 1980). There are three

Table 1 | Descriptions of model parameters in the XAJ14 and XAJ7 models

Module	Pars.	Parameter descriptions	Range and units
Runoff generation	B[b]	Exponential of the distribution to tension water capacity	0–1
	IMP	Percentage of impervious and saturated areas in the catchment	0–0.1 (%)
Evapotranspiration	UM[b,c]	Average soil moisture storage capacity of the upper layer	5–100 (mm)
	LM	Average soil moisture storage capacity of the middle layer	50–300 (mm)
	DM[b,c]	Average soil moisture storage capacity of the deepest layer	5–100 (mm)
	C	ET coefficient of the deepest layer	0.1–0.2
Runoff separation	SM	Areal mean free water capacity of the surface soil layer	10–60 (mm)
	EX	Exponential of the spatial distribution curve of free water storage capacity	1.0–1.5
	KI	Outflow coefficient of free water storage to the interflow	0.3–0.7
	KG	Outflow coefficient of free water storage to the groundwater	0.1–0.2
	FC[a]	Steady recharge constant to groundwater	0.5–15 (mm/d)
Routing	CI	Recession constant of the lower interflow storage	0.1–0.9
	CG[b]	Recession constant of the lower groundwater storage	0.95–0.99
	XE[b]	Muskingum coefficient of geometry factor	0–0.5
	KE[b]	Muskingum coefficient of residence time of water	1–3 (d)

[a]The parameter that only belongs to XAJ7 model.
[b]The parameters that are shared by XAJ7 and XAJ14 models.
[c]The parameter range of UM and DM in the XAJ7 are 10–150 and 50–350, respectively.
Note that the range of parameters comes from the published studies of Li *et al.* (2009) and Zhao *et al.* (1980).

differences between the XAJ7 and XAJ14 models: (1) the XAJ7 model does not consider the impact of impervious surface on runoff generation; (2) the XAJ7 model calculates *AET* using a two-layer rather than three-layer *ET* submodel; and (3) corresponding to the two-layer water storage, the generated runoff is separated into two components (surface runoff and subsurface runoff) using the Horton's concept of steady infiltration rate. In the saturated areas of the watershed, the portion of the generated runoff that exceeds the steady infiltration rate forms surface runoff, and the rest forms subsurface runoff. Detailed descriptions of model structures, parameters, and state variables in the XAJ7 and XAJ14 can be found in Figure 1 and Table 1.

Study area and data used

Twenty-six mountainous catchments with limited anthropogenic influences were selected as test catchments (Figure 2). These test catchments are located in the Poyang Lake basin of Jiangxi Province (113°42′–118°15′ E, 24°30′–29°48′ N), which is the largest freshwater lake in China. The elevations in the study area decrease from south to north, ranging from 1,700 m to 50 m above mean sea level (a.m.s.l). The study area has a subtropical wet climate governed by the East Asian monsoon. Approximately 60–70% of the annual precipitation occurs in the wet season from May to October. The land use type is predominantly forest and grassland, and soil types are dominated by red soil and paddy soil. Table 2 summarizes the physical and climatic characteristics of the catchments.

Daily meteorological data from 42 stations were obtained from the Meteorological Bureau of Jiangxi Province. *PET* was estimated by the Hargreaves and Samani equation (Hargreaves & Samani 1985). Basin average rainfall and *PET* were calculated by the Thiessen polygons method based on the available meteorological stations in and around each catchment. Observed streamflow data were provided by China's Hydrological Year Book, published by the Hydrological Bureau of the Ministry of Water Resources, China. All test catchments have 17 years of continuous streamflow data from 1970 to 1986. The period 1971–1978 was used for model calibration (1970 was used for model warm up), and the period 1979–1986 was used for model validation.

METHODOLOGY

Parameter regionalization methodology

For basins without streamflow records, model parameters must be estimated from other sources, such as neighboring basins or

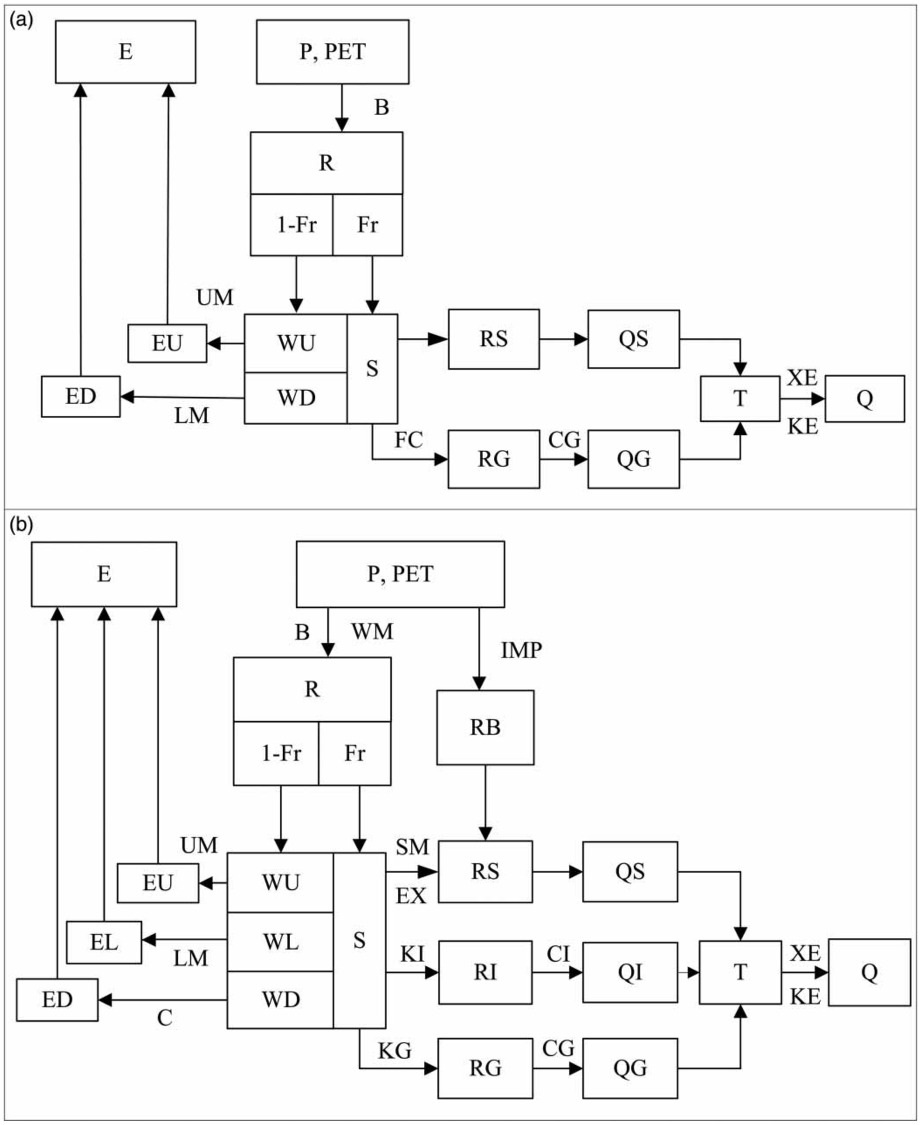

Figure 1 | Model structures for seven-parameter (a) and fourteen-parameter (b) XAJ models.

physically similar basins where streamflow data are available. The methods allowing the transfer of hydrological model parameters from gauged to ungauged basins can be called parameter regionalization approaches (Oudin *et al.* 2010). Parameter regionalization approaches typically include the following steps: (1) identifying the gauged basins (donor basins), which have similar hydrological characteristics as the ungauged basins (target basins); (2) calibrating model parameters in donor basins (calibration mode); (3) validating model performance on an independent period (validation mode); and (4) transferring the model parameters from donor basins into target basins for streamflow simulations and predictions (regionalization mode) if the validation results are satisfactory. In this study, only one donor basin was used as a target basin. Two commonly used parameter regionalization methods were employed, spatial proximity and physical similarity. The difference between the two methods lies in how the donor basin is chosen. The spatial proximity method chooses a donor basin based on spatial distances between the centroids of two neighboring catchments, while the physical similarity method chooses a donor basin based on the similarity of the catchment characteristics between two catchments

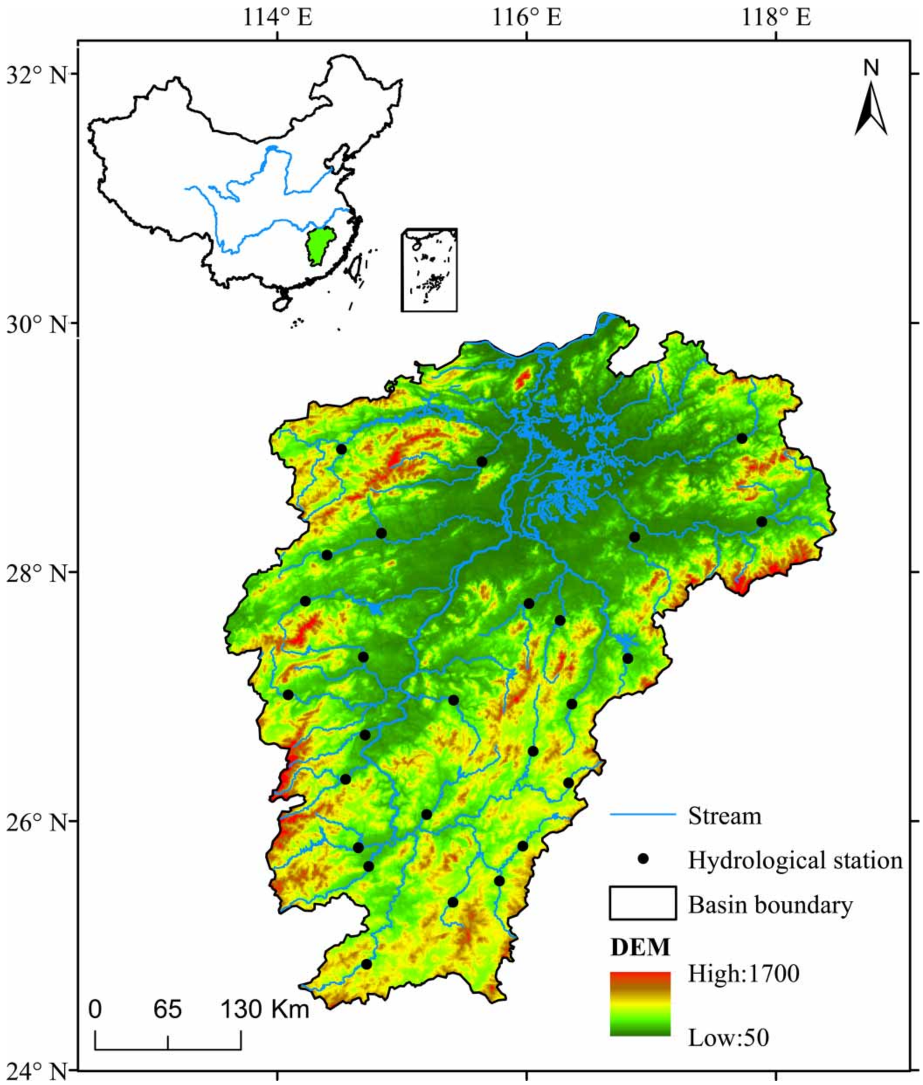

Figure 2 | Location of the study area and spatial distribution of the hydrological stations.

Table 2 | Summary of the catchment physical and climatic characteristics

Catchment characteristics	Min	Median	Max
Catchment area (km^2)	435	1,081	3,548
Mean annual rainfall (mm)	1,341	1,544	1,940
Aridity index	0.58	0.80	0.92
Runoff coefficient	0.48	0.58	0.72
Percent forest cover (%)	48	75	92
Stream length (km)	36	57	99
Mean elevation (m)	76	178	348
Catchment slope (‰)	1.5	5.2	16.7

(Parajka *et al.* 2013). In the regionalization mode, each of the 26 catchments was first considered as a donor basin for calibration and validation, and then each one in turn was considered as a target basin to receive optimized parameter values from the donor basin. The catchment characteristics in Table 2 were considered in the physical similarity method. The Kay approach (Kay *et al.* 2007) was used to combine the seven catchment characteristics into a single similarity index (θ):

$$\theta = \sum_{i=1}^{k} \sqrt{\omega_i \left(\frac{X_i^D - X_i^T}{\sigma X_i} \right)^2} \tag{1}$$

where j represents the number of catchment characteristics. X_i^D and X_i^T are the values of the catchment characteristic for the donor basin and target basin, respectively. σX_i is the standard deviation of the characteristics over all catchments, and w_j is the weight attributed to the jth catchment characteristic. The weight values of each catchment characteristic have a significant impact on donor basin selection (Burn & Boorman 1993; Oudin et al. 2008; Zhang & Chiew 2009; Arsenault et al. 2015). In this study, the entropy method was used to determine the weight coefficient of each catchment characteristic, enabling the measurement of useful catchment characteristic information and avoiding the subjective influence of decision-makers (Yu & Lai 2011). More detailed descriptions about entropy-based weight measures can be found in Wang & Lee (2009) and Zou et al. (2006).

Model performance assessment criteria

The Kling–Gupta efficiency (KGE) (Gupta et al. 2009) was used as an objective function to calibrate model parameters. The KGE incorporates three basic assessment criteria, can provide a reliable measurement of the overall agreement between simulated and observed values (Pechlivanidis et al. 2014), and is defined as:

$$KGE = 1 - \sqrt{(1-\gamma)^2 + (1-\alpha)^2 + (1-\beta)^2} \tag{2}$$

where γ is the correlation coefficient between simulated and observed values; α is the relative variability in simulated and observed values, and β is the ratio of the mean values of simulated and observed values. A value of $KGE = 1$ indicates a perfect fit between observed and simulated values.

In addition to the KGE, a complementary evaluation was carried out based on the following three criteria: percent bias ($PBIAS$), Nash–Sutcliffe efficiency (NSE) (Nash & Sutcliffe 1970), and NSE of log-transformed flows. $PBIAS$ is used to measure the difference between the observed and simulated total runoff. NSE is a valuable means to evaluate high flow simulations (Pushpalatha et al. 2011). NSE calculated on log transformed flows (NSE_{log}) has been proven to be an efficient evaluation criterion for low flow simulations (Oudin et al. 2006; Pushpalatha et al. 2012). $PBIAS = 0$, $NSE = 1$, and $NSE_{log} = 1$ indicate a perfect fit between the observed and simulated values. The three criteria can be expressed as follows:

$$PBIAS = \left| 1 - \frac{\sum_{i=1}^{N} Q_{sim,i}}{\sum_{i=1}^{N} Q_{obs,i}} \right| \tag{3}$$

$$NSE = 1 - \frac{\sum_{i=1}^{N} (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^{N} (Q_{obs,i} - \overline{Q_{obs}})^2} \tag{4}$$

$$NSE_{log} = 1 - \frac{\sum_{i=1}^{N} (log(Q_{obs,i} + \varepsilon) - log(Q_{sim,i} + \varepsilon))^2}{\sum_{i=1}^{N} (log(Q_{obs,i} + \varepsilon) - log(\overline{Q_{obs}}))^2} \tag{5}$$

where Q_{sim} and Q_{obs} are the simulated and observed streamflow, respectively; $\overline{Q_{obs}}$ is the mean observed streamflow; N is the total number of days in the calibration period; and ε is a small constant to avoid having zero in the denominator of Equation (5). Model parameters are automatically optimized using the particle swarm optimization algorithm to maximize the objective function. This optimization algorithm proposed by Eberhart & Kennedy (1995) has been widely applied to optimize the parameters of hydrological models (Gill et al. 2006; Zhang & Chiew 2009; Luo et al. 2012).

Parameter sensitivity and identifiability analysis

Parameter sensitivity analysis is an important way to identify dominant hydrological processes that exhibit major influence on the simulation results (Wagener et al. 2003; Abebe et al. 2010). Here, a modified version of the regional sensitive analysis method was employed to analyze parameter sensitivity (Freer et al. 1996; Wagener & Kollat 2007). Monte Carlo random sampling was performed considering a uniform probability distribution for all parameters in the XAJ7 and XAJ14 models (Table 1). A total of 10,000 simulations were run for each model. The parameter sets were then split into ten groups of equal size based on the evaluation criterion of KGE (Figure 3). In general, the sensitivities of the parameters were determined by comparing the differences in cumulative distribution functions (CDFs) of parameter groups (Freer et al. 1996; Wagener & Kollat 2007). A larger difference between the two CDFs indicates a higher sensitivity of parameter to model performance. The parameter sensitivity can be quantitatively described by calculating the area between the CDFs of the best and the worst groups of parameter populations. Figure 3 presents the scheme map of parameter sensitivity

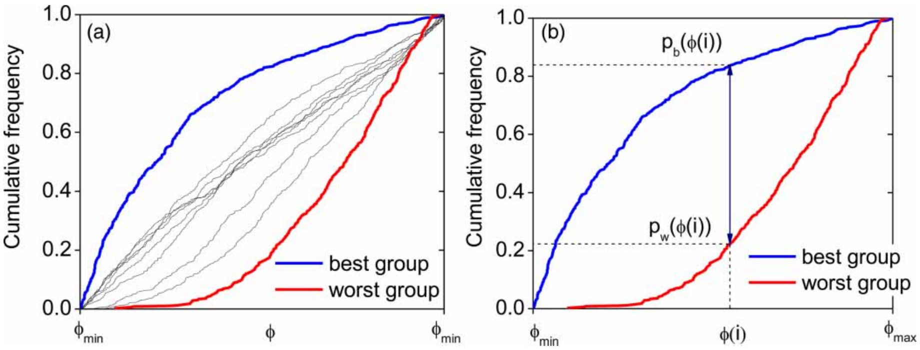

Figure 3 | Scheme map of parameter sensitivity measures.

calculation. For the *j-th* parameter ϕ, the normalized common area (A_j) under the two CDFs can be calculated as follows:

$$A_j = \sum_{i=1}^{N} \frac{|p_b(\phi_i) - p_w(\phi_i)|}{\max\left[p_b(\phi_i),\ pw(\phi_i)\right]} \tag{6}$$

and the sensitivity coefficient (*SC*) of the parameter ϕ is calculated as:

$$SC(j) = \frac{A_j}{\sum_{j=1}^{M} A_j} \tag{7}$$

where p_b is the cumulative frequency of the best group parameter populations; p_w is the cumulative frequency of the worst group parameter populations; N is the number of bins of the parameter range being split ($N = 500$ used in this study), and M is the number of parameters.

Parameter identifiability is also an important concern of model evaluation (Wagener *et al.* 2001a, 2003). Generally, only a few parameters in a hydrological model are identifiable based on automatic parameter calibration. The poorly identifiable parameters result from an insensitivity of model outputs to these parameters, parameter interaction, or both (Doherty & Hunt 2009), which produces a similar model performance within a full range (i.e., parameter equifinality) and negatively affects hydrological predictions (Schoups *et al.* 2008). Parameter identifiability can be measured by calculating the gradient of the cumulative distribution of the top 10% best performing parameter sets (Wagener *et al.* 2001b). Figure 4 shows the schematic diagram for measuring a well and a poorly identifiable parameter. The scatter plots in Figure 4(a) and 4(b) show the two parameter populations against corresponding objective function (*KGE*) values. In this study, the top 10% best performing parameter sets were used to derive a CDF. One can then split the parameter range into ten groups of equal size and calculate the gradient of cumulative distribution for each group (Figure 4(c) and 4(d)). A steeper gradient in the cumulative distribution indicates a more identifiable parameter in the top 10% best performing parameter sets. As suggested by Wagener & Kollat (2007), the largest gradient value (max_id) is used as an indicator for measuring parameter identifiability. More detailed descriptions of parameter identifiability measurement can be found in Wagener & Kollat (2007).

RESULTS AND DISCUSSION

Model evaluation in calibration, validation, and regionalization modes

Figure 5 shows the comparisons of model performance for the XAJ7 and XAJ14 in calibration, validation, and regionalization modes. None of the comparisons show statistically significant differences in model performance. In calibration mode, the XAJ14 model performs better than the XAJ7 model and yields higher *KGE* and lower *PBIAS* values than XAJ7. The difference in model performance decreases significantly when moving from calibration mode to validation mode where the two models achieve similar model performance, indicating the potential inconsistency of complex models to perform similarly under ‘trained’ and ‘untrained’ (spatial and/or temporal) conditions. Similar

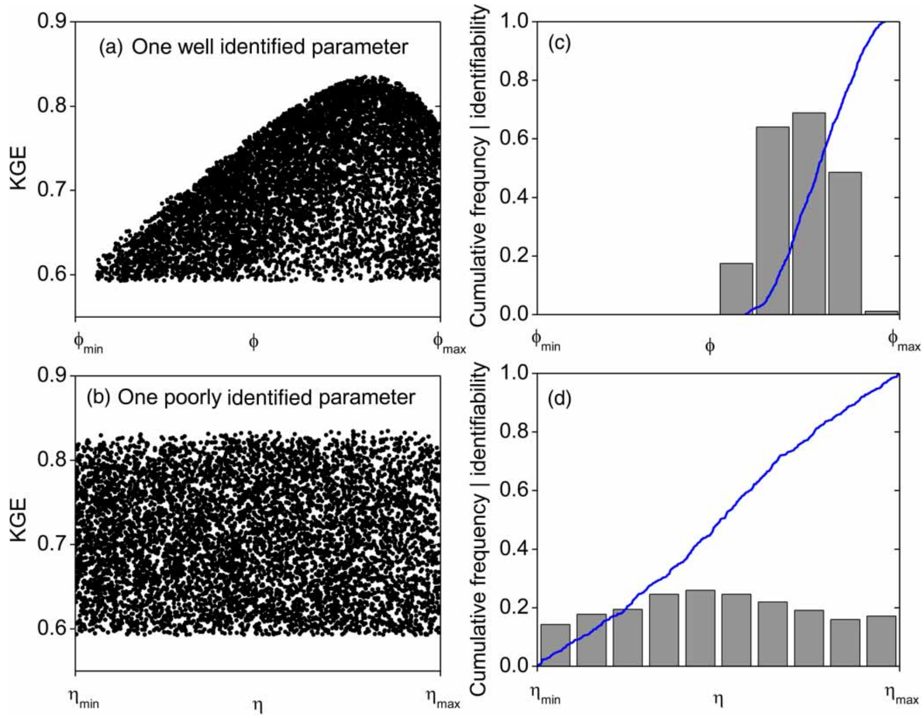

Figure 4 | Sketch map of parameter identifiability measures based on parameter populations conditioned on *KGE*.

findings also have been reported by Michaud & Sorooshian (1994), Yew Gan *et al.* (1997), Perrin *et al.* (2001), and Bai *et al.* (2015) using different hydrological models.

Figure 5 shows model regionalization results based on physical similar and spatial proximity methods. The physical similar method performs slightly better than the spatial proximity method in the regionalization mode. Model regionalization results are poorer than the model calibration and validation results, with median *KGE* values from the regionalization results being approximately 0.10 to 0.15 lower than the calibration and validation results. The XAJ14 and XAJ7 models achieve similar regionalization results whether for the physical similar method or the spatial proximity method, indicating that more complex process representations in the XAJ14 model do not improve streamflow prediction ability in ungauged basins compared with the simple XAJ7 model.

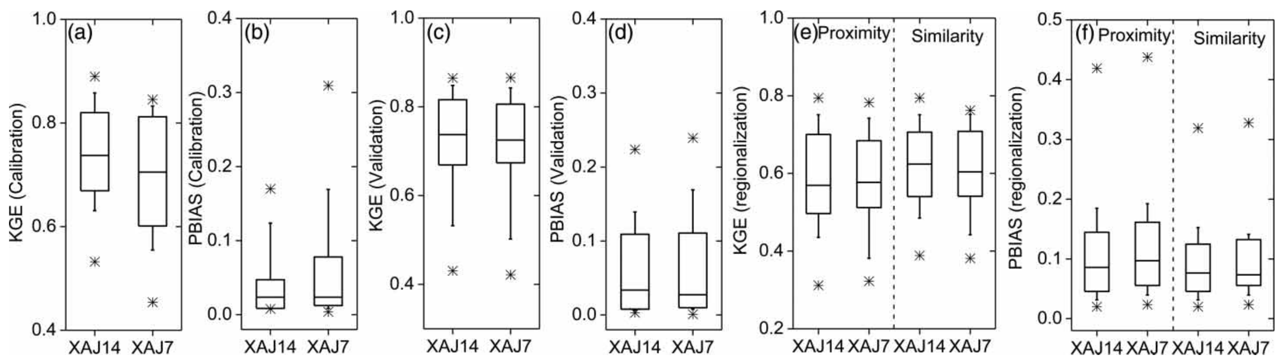

Figure 5 | Model performance for the XAJ14 and XAJ7 models in calibration mode (a) and (b) and validation mode (c) and (d). In each boxplot, the bottom, middle, and top of the box are the 25th, 50th, and 75th percentiles, and the bottom and top whiskers are the 10th and 90th percentiles. Data points inside the boxes are the average values, and data points beyond the whiskers are the maximum and minimum values.

Regionalization under high flow and low flow simulations

The *NSE* and NSE_{log} are used to evaluate high flow and low flow simulations, respectively. The physical similarity method is used for model parameterization due to its better performance than the spatial proximity method. Figure 6 shows the regionalization results under high flow and low flow simulations. The XAJ7 model is slightly better than the XAJ14 model in high flow simulations, and the median *NSE* values from XAJ14 and XAJ7 are 0.62 and 0.60, respectively. However, for low flow simulations, the XAJ7 model performs better than the XAJ14 model, and the median NSE_{log} obtained from XAJ14 and XAJ7 are 0.25 and 0.36, respectively. The two models perform better in high flow simulations than in low flow simulations. Similar conclusions were also drawn by Pushpalatha *et al.* (2011) and Zhang *et al.* (2014), who also reported that rainfall–runoff models have minimal ability to predict low flows. This result could be partly caused by the objective function used in the hydrological model calibration, which is more sensitive to medium and high flows than to low flows. Use of a low flow sensitive objective function may have resulted in different conclusions.

Parameter identifiability and independence

As mentioned in the section 'Parameter sensitivity and identifiability analysis', the max_id was used as an indicator of parameter identifiability. Figure 7 shows the median max_id in XAJ7 and XAJ14 models for over 26 test basins. The median max_id values in XAJ7 range from 0.39 to 0.49, with an average of 0.44, while the values in XAJ14 range from 0.28 to 0.40, with an average of 0.33. This indicates that model parameters in XAJ7 are easier to identify than those in XAJ14. For the parameters shared by the two models, the XAJ7 model has higher max_id values than the XAJ14 model, which indicates that model simplification can enhance parameter identifiability.

We also analyzed the independence of model parameters for the two hydrological models. Table 3 shows the correlation coefficients of calibrated model parameters in XAJ14 and XAJ7 models. The interdependence of the calibrated parameters is weak for XAJ7, and the correlation coefficients between model parameters range from −0.25 and 0.26. The weak correlations between the parameters of XAJ7 probably stem from the parsimony of the model. Compared to XAJ7, the parameter interdependence in XAJ14 is more pronounced, and the correlation coefficients between some parameters are greater than 0.50 or less than −0.50, indicating that some of the parameters in XAJ14 model are covariant with each other and have a similar effect on the streamflow simulations.

Overall, parameter identifiability and independence analysis indicate that model parameters in XAJ7 are easier to identify and have less correlation behavior than those in the XAJ14 model. Lack of parameter identifiability and parameter interaction could lead to large uncertainties in streamflow prediction (Beven 1993; Wheater *et al.* 1993).

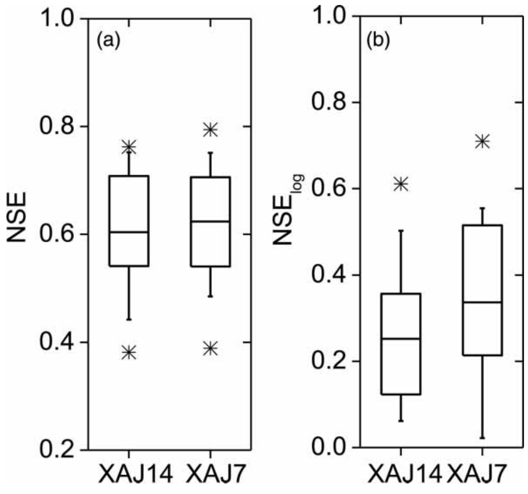

Figure 6 | Model regionalization results for high flow (a) and low flow (b) simulations.

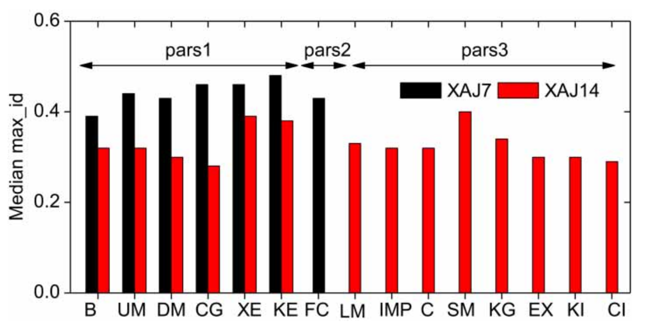

Figure 7 | Comparison of the parameter identifiability between XAJ7 and XAJ14 models based on median max_id over 26 test basins. The pars1 indicates the parameters that are shared by XAJ7 and XAJ14 models; the pars2 indicates the parameter that only belongs to XAJ7; the pars3 indicates the parameters that only belong to XAJ14.

Table 3 | Correlation coefficients between model parameters

	Parameters													
Model	B	UM	DM	CG	XE	KE	LM/FC	IMP	C	SM	EX	KG	KI	CI
XAJ14														
B	1.00													
UM	0.15	1.00												
DM	-0.02	0.21	1.00											
CG	0.02	−0.30	0.18	1.00										
XE	0.07	−0.01	−0.19	0.00	1.00									
KE	−0.13	0.01	−0.31	−0.31	−0.17	1.00								
LM	0.16	**−0.50**	**−0.63**	−0.13	−0.07	0.24	1.00							
IMP	0.21	0.05	0.14	0.17	**−0.70**	0.24	0.06	1.00						
C	−0.04	−0.41	0.10	0.28	0.07	−0.10	−0.16	−0.17	1.00					
SM	0.09	0.13	−0.42	−0.26	0.31	**0.53**	0.07	−0.47	−0.07	1.00				
EX	0.21	−0.03	−0.37	0.01	0.04	−0.38	0.07	0.03	−0.17	0.24	1.00			
KG	0.13	0.36	−0.27	**−0.54**	0.22	−0.14	0.20	−0.46	−0.37	0.16	−0.05	1.00		
KI	0.04	0.42	−0.03	−0.44	−0.32	−0.04	−0.07	0.24	**−0.56**	0.03	0.17	0.25	1.00	
CI	−0.15	−0.14	0.36	0.26	**−0.75**	0.03	0.03	**0.61**	0.03	**-0.54**	−0.23	**−0.58**	0.19	1.00
XAJ7														
B	1.00													
UM	0.07	1.00												
DM	−0.01	−0.15	1.00											
CG	−0.11	−0.19	0.07	1.00										
XE	0.20	−0.09	−0.01	0.10	1.00									
KE	−0.24	0.05	−0.20	−0.21	−0.25	1.00								
FC	0.23	0.17	0.26	−0.19	0.10	0.10	1.00							

Note that the values greater (less) than or equal to 0.50 (−0.50) are shown in bold.

This is likely the reason why the complex XAJ14 model cannot consistently outperform the XAJ7 model in validation and regionalization modes.

Model complexity for prediction in ungauged basins

Streamflow PUB require the development of a model that is able to capture the dominant hydrological processes and avoid parameter interaction and lack of identifiability (Wagener *et al.* 2004b; Young 2006; Reusser *et al.* 2011). Under the premise of similar prediction ability, the simplistic models are more attractive than sophisticated models for regionalization because they are easier to parameterize. Model comparison results suggest that the increased model complexity of the XAJ14 model results in, as expected, a better performance of streamflow simulation than the XAJ7 model in calibration mode. This is mainly due to higher degree of freedom of XAJ14, which allows parameters to compensate for model structure uncertainty and lack of process understanding. However, model selection should additionally focus on the model performance in the validation and regionalization modes, which represent the model's 'untrained' conditions and reflect the actual streamflow prediction capability where the XAJ7 performs comparably or even better (in low flow simulations) than the XAJ14 model (Figure 5). Additionally, parameter sensitivity analysis suggests that not all hydrological modeling processes are active or at the same level of importance (Figure 8). For the XAJ14 model, only four parameters (SM, KI, XE, and KE) related to runoff separation and routing processes are sensitive to simulation results. With respect to the XAJ7 model, it retains the routing process from XAJ14 and simplifies the runoff generation, evapotranspiration, and runoff separation processes. These simplifications enhance parameter identifiability and reduce the parameter interdependence. Generally, the simple XAJ7 is a better choice than XAJ14 for regionalization in ungauged basins.

SUMMARY AND CONCLUSION

The main objective of this study was to determine whether the XAJ7 or XAJ14 model is more suitable for streamflow prediction in ungauged basins. Model evaluation was performed not only based on model performance, but also on the dependence on parameter identifiability. The results showed that the XAJ14 model benefits from the increased complexity and yields better model performance than the simple XAJ7 model in calibration mode. However, the superior performance cannot be sustained in validation and regionalization modes where the simple XAJ7 model performed similarly or even better (in low flow simulations) than the complex XAJ14 model. Parameter identifiability and independence analysis suggested that model parameters in the simple XAJ7 model are more identifiable and have less correlation behavior than those in the complex XAJ14 model, which probably causes the inconsistency of model performance between XAJ14 and XAJ7 models in different modes. Considering model efficiency, parameter identifiability, and independence, the XAJ7 model is a better choice than the XAJ14 model for streamflow prediction in ungauged basins.

In addition, current hydrological research highlights the development of an integrated model of water-related processes and the development of land surface process models (Foley *et al.* 1996; Arnold *et al.* 1998; Liu *et al.* 2008), in which the hydrological model only serves as a sub-model to simulate water balance. These models without exception have a large number of parameters, which makes reliable parameter estimation a challenging task. Compared with the XAJ14 model, the simple XAJ7 model seems more suitable to be coupled with other water-related models, or to be embedded into a land surface model.

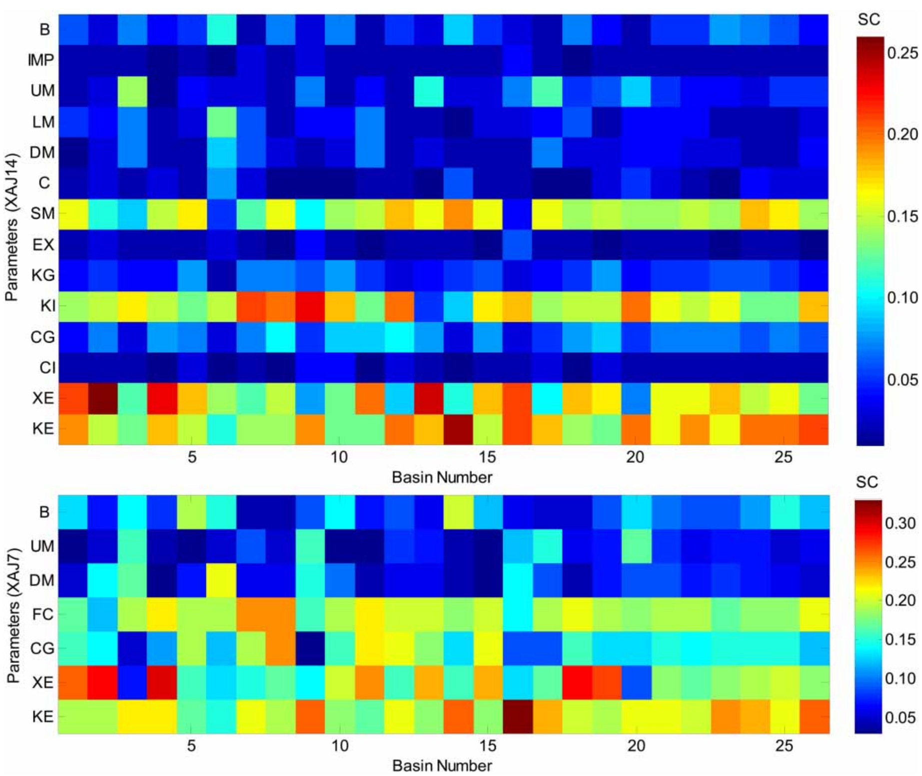

Figure 8 | Sensitivity of model parameters for XAJ14 model and XAJ7 model.

ACKNOWLEDGEMENTS

This research was supported by the Natural Science Foundation of China (41201034, 41330529), the program for 'Bingwei' Excellent Talents in Institute of Geographic Sciences and Natural Resources Research, CAS (Project No. 2013RC202), Chinese Academy of Sciences Visiting Professorship for Senior International Sciences (Grant No. 2013T2Z0014) and Natural Sciences Foundation of Jiangsu Province (BK20141059).

REFERENCES

Abebe, N. A., Ogden, F. L. & Pradhan, N. R. 2010 Sensitivity and uncertainty analysis of the conceptual HBV rainfall–runoff model: implications for parameter estimation. *Journal of Hydrology* **389** (3–4), 301–310. DOI:10.1016/j.jhydrol.2010.06.007.

Ahiablame, L., Chaubey, I., Engel, B., Cherkauer, K. & Merwade, V. 2012 Estimation of annual baseflow at ungauged sites in Indiana USA. *Journal of Hydrology* **476**, 13–27. http://dx.doi.org/10.1016/j.jhydrol.2012.10.002.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. 1998 Large area hydrologic modeling and assessment part I: model development. *Journal of the American Water Resources Association* **34** (1), 73–89. DOI:10.1111/j.1752-1688.1998.tb05961.x.

Arsenault, R., Poissant, D. & Brissette, F. 2015 Parameter dimensionality reduction of a conceptual model for streamflow prediction in Canadian, snowmelt dominated ungauged basins. *Advances in Water Resources* **85**, 27–44. DOI:10.1016/j.advwatres.2015.08.014.

Bai, P., Liu, X., Liang, K. & Liu, C. 2015 Comparison of performance of twelve monthly water balance models in different climatic catchments of China. *Journal of Hydrology* **529** (3), 1030–1040. http://dx.doi.org/10.1016/j.jhydrol.2015.09.015.

Bárdossy, A. 2007 Calibration of hydrological model parameters for ungauged catchments. *Hydrology and Earth System Sciences Discussions* **11** (2), 703–710.

Benito, G. 2012 Rainfall-runoff modelling and palaeoflood hydrology applied to reconstruct centennial scale records of flooding and aquifer recharge in ungauged ephemeral rivers. *Quaternary International* **279–280**, 47. http://dx.doi.org/10.1016/j.quaint.2012.07.204.

Besaw, L. E., Rizzo, D. M., Bierman, P. R. & Hackett, W. R. 2010 Advances in ungauged streamflow prediction using artificial neural networks. *Journal of Hydrology* **386** (1–4), 27–37. DOI:10.1016/j.jhydrol.2010.02.037.

Beven, K. 1993 Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources* **16** (1), 41–51. http://dx.doi.org/10.1016/0309-1708(93)90028-E.

Blöschl, G. 2006 *Rainfall-Runoff Modeling of Ungauged Catchments, Encyclopedia of Hydrological Sciences*. John Wiley & Sons, Chichester, UK. DOI:10.1002/0470848944.hsa140.

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A. & Savenije, H. (eds) 2013 *Runoff Prediction in Ungauged Basins*. Cambridge University Press, Cambridge, UK.

Booker, D. J. & Snelder, T. H. 2012 Comparing methods for estimating flow duration curves at ungauged sites. *Journal of Hydrology* **434–435**, 78–94. DOI:10.1016/j.jhydrol.2012.02.031.

Burn, D. H. & Boorman, D. B. 1993 Estimation of hydrological parameters at ungauged catchments. *Journal of Hydrology* **143** (3–4), 429–454. http://dx.doi.org/10.1016/0022-1694(93)90203-L.

Castiglioni, S., Lombardi, L., Toth, E., Castellarin, A. & Montanari, A. 2010 Calibration of rainfall-runoff models in ungauged basins: a regional maximum likelihood approach. *Advances in Water Resources* **33** (10), 1235–1242.

Cheng, C.-T., Zhao, M.-Y., Chau, K. & Wu, X.-Y. 2006 Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure. *Journal of Hydrology* **316** (1), 129–140.

Chiew, F. H. S., Stewardson, M. J. & McMahon, T. A. 1993 Comparison of six rainfall-runoff modelling approaches. *Journal of Hydrology* **147** (1–4), 1–36. http://dx.doi.org/10.1016/0022-1694(93)90073-I.

Doherty, J. & Hunt, R. J. 2009 Two statistics for evaluating parameter identifiability and error reduction. *Journal of Hydrology* **366** (1–4), 119–127. DOI:10.1016/j.jhydrol.2008.12.018.

Eberhart, R. & Kennedy, J. 1995 A new optimizer using particle swarm theory, MHS'95. In: *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43.

El-Hames, A. S. 2012 An empirical method for peak discharge prediction in ungauged arid and semi-arid region catchments based on morphological parameters and SCS curve number. *Journal of Hydrology* **456–457**, 94–100. DOI:10.1016/j.jhydrol.2012.06.016.

Foley, J. A., Prentice, I. C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S. & Haxeltine, A. 1996 An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochemical Cycles* **10** (4), 603–628.

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 Resources Research* **32** (7), 2161–2173.

Gill, M. K., Kaheil, Y. H., Khalil, A., McKee, M. & Bastidas, L. 2006 Multiobjective particle swarm optimization for parameter estimation in hydrology. *Water Resources Research* **42** (7), W07417. DOI:10.1029/2005WR004528.

Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F. 2009 Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *Journal of Hydrology* **377** (1–2), 80–91. DOI:10.1016/j.jhydrol.2009.08.003.

Hailegeorgis, T. T. & Alfredsen, K. 2015 Comparative evaluation of performances of different conceptualisations of distributed HBV runoff response routines for prediction of hourly streamflow in boreal mountainous catchments. *Hydrology Research* **46** (4), 607–628.

Hargreaves, G. H. & Samani, Z. A. 1985 Reference crop evapotranspiration from ambient air temperature. *American Society of Agricultural Engineers*, (Microfiche collection) (USA). no. fiche no. 85-2517.

He, Y., Bárdossy, A. & Zehe, E. 2011 A review of regionalisation for continuous streamflow simulation. *Hydrology and Earth System Sciences* **15** (11), 3539–3553.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E. & Cuddennec, C. 2013 A decade of predictions in ungauged basins (PUB) – a review. *Hydrological Sciences Journal* **58** (6), 1198–1255.

Jakeman, A. & Hornberger, G. 1993 How much complexity is warranted in a rainfall–runoff model? *Water Resources Research* **29** (8), 2637–2649.

Jin, X., Xu, C.-y., Zhang, Q. & Chen, Y. D. 2009 Regionalization study of a conceptual hydrological model in Dongjiang basin, south China. *Quaternary International* **208** (1), 129–137.

Kay, A., Jones, D., Crooks, S., Kjeldsen, T. & Fung, C. 2007 An investigation of site-similarity approaches to generalisation of a rainfall-runoff model. *Hydrology & Earth System Sciences* **11** (1), 500–515.

Kizza, M., Guerrero, J.-L., Rodhe, A., Xu, C.-y. & Ntale, H. K. 2013 Modelling catchment inflows into Lake Victoria: regionalisation of the parameters of a conceptual water balance model. *Hydrology Research* **44** (5), 789–808.

Lee, H., McIntyre, N., Wheater, H. & Young, A. 2005 Selection of conceptual models for regionalisation of the rainfall-runoff relationship. *Journal of Hydrology* **312** (1–4), 125–147. http://dx.doi.org/10.1016/j.jhydrol.2005.02.016.

Li, H., Xu, C. Y. & Beldring, S. 2015 How much can we gain with increasing model complexity with the same model concepts? *Journal of Hydrology* **527**, 858–871. http://dx.doi.org/10.1016/j.jhydrol.2015.05.044.

Li, H., Zhang, Y., Chiew, F. H. S & Xu, S. 2009 Predicting runoff in ungauged catchments by using Xinanjiang model with MODIS leaf area index. *Journal of Hydrology* **370** (1–4), 155–162.

Liu, C., Wang, Z., Zheng, H., Zhang, L. & Wu, X. 2008 Development of hydro-informatic modelling system and its application. *Science in China Series E: Technological Sciences* **51** (4), 456–466.

Luo, J., Wang, E., Shen, S., Zheng, H. & Zhang, Y. 2012 Effects of conditional parameterization on performance of rainfall-runoff model regarding hydrologic non-stationarity. *Hydrological Processes* **26**, 3953–3961. DOI:10.1002/hyp.8420.

McCabe, M., Franks, S. & Kalma, J. 2005 Calibration of a land surface model using multiple data sets. *Journal of Hydrology* **302** (1), 209–222.

McIntyre, N., Lee, H., Wheater, H., Young, A. & Wagener, T. 2005 Ensemble predictions of runoff in ungauged catchments. *Water Resources Research* **41** (12). DOI:10.1029/2005wr004289.

Michaud, J. & Sorooshian, S. 1994 Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed. *Water Resources Research* **30** (3), 593–605.

Nash, J. E. & Sutcliffe, J. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *Journal of Hydrology* **10** (3), 282–290.

Oudin, L., Andréassian, V., Mathevet, T., Perrin, C. & Michel, C. 2006 Dynamic averaging of rainfall–runoff model simulations from complementary model parameterizations. *Water Resources Research* **42** (7), W07410. DOI:10.1029/2005WR004636.

Oudin, L., Andréassian, V., Perrin, C., Michel, C. & Le Moine, N. 2008 Spatial proximity, physical similarity, regression and ungaged catchments: a comparison of regionalization approaches based on 913 French catchments. *Water Resources Research* **44** (3). DOI:10.1029/2007wr006240.

Oudin, L., Kay, A., Andréassian, V. & Perrin, C. 2010 Are seemingly physically similar catchments truly hydrologically similar? *Water Resources Research* **46** (11). DOI:10.1029/2009WR008887.

Parajka, J., Viglione, A., Rogger, M., Salinas, J. L., Sivapalan, M. & Blöschl, G. 2013 Comparative assessment of predictions in ungauged basins – Part 1: runoff-hydrograph studies. *Hydrology and Earth System Sciences* **17** (5), 1783–1795. DOI:10.5194/hess-17-1783-2013.

Pechlivanidis, I., McIntyre, N. & Wheater, H. 2010 Calibration of the semi-distributed PDM rainfall–runoff model in the Upper Lee catchment, UK. *Journal of Hydrology* **386** (1), 198–209.

Pechlivanidis, I., Jackson, B., McMillan, H. & Gupta, H. 2014 Use of an entropy-based metric in multiobjective calibration to improve model performance. *Water Resources Research* **50** (10), 8066–8083.

Perrin, C., Michel, C. & Andréassian, V. 2001 Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology* **242** (3), 275–301.

Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T. & Andréassian, V. 2011 A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. *Journal of Hydrology* **411** (1–2), 66–76. http://dx.doi.org/10.1016/j.jhydrol.2011.09.034.

Pushpalatha, R., Perrin, C., Moine, N. L. & Andréassian, V. 2012 A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology* **420–421**, 171–182. DOI:10.1016/j.jhydrol.2011.11.055.

Reusser, D. E., Buytaert, W. & Zehe, E. 2011 Temporal dynamics of model parameter sensitivity for computationally expensive models with the Fourier amplitude sensitivity test. *Water Resources Research* **47** (7). DOI:10.1029/2010wr009947.

Schoups, G., van de Giesen, N. C. & Savenije, H. H. G. 2008 Model complexity control for hydrologic prediction. *Water Resources Research* **44** (12). DOI:10.1029/2008WR006836.

Seibert, J. & Beven, K. J. 2009 Gauging the ungauged basin: how many discharge measurements are needed? *Hydrology and Earth System Sciences* **13** (6), 883–892. DOI:10.5194/hess-13-883-2009.

Sivapalan, M. 2003 Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes* **17** (15), 3163–3170. DOI:10.1002/hyp.5155.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S. & Zehe, E. 2003 IAHS Decade on predictions in ungauged basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal* **48** (6), 857–880.

Skaugen, T., Peerebom, I. O. & Nilsson, A. 2015 Use of a parsimonious rainfall-runoff model for prediting hydrological response in ungauged basins. *Hydrological Processes* **29** (8), 1999–2013.

Vaché, K. B. & McDonnell, J. J. 2006 A process-based rejectionist framework for evaluating catchment runoff model structure. *Water Resources Research* **42** (2). DOI:10.1029/2005WR004247.

Wagener, T. & Kollat, J. 2007 Numerical and visual evaluation of hydrological and environmental models using the Monte Carlo analysis toolbox. *Environmental Modelling & Software* **22** (7), 1021–1033.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V. & Sorooshian, S. 2001a A framework for development and application of hydrological models. *Hydrology and Earth System Sciences* **5** (1), 13–26. DOI:10.5194/hess-5-13-2001.

Wagener, T., Lees, M. J. & Wheater, H. S. 2001b A toolkit for the development and application of parsimonious hydrological models. In: *Mathematical Models of Small Watershed Hydrology* (V. P. Singh & D. K. Frevert, eds). Water Resources Publications, LLC, Chelsea, MI, USA, p. 2.

Wagener, T., McIntyre, N., Lees, M., Wheater, H. & Gupta, H. 2003 Towards reduced uncertainty in conceptual rainfall–runoff modelling: dynamic identifiability analysis. *Hydrological Processes* **17** (2), 455–476.

Wagener, T., Sivapalan, M., McDonnell, J., Hooper, R., Lakshmi, V., Liang, X. & Kumar, P. 2004a Predictions in ungauged basins as a catalyst for multidisciplinary hydrology. *Eos, Transactions American Geophysical Union* **85** (44), 451–457. DOI:10.1029/2004EO440003.

Wagener, T., Wheater, H. S. & Gupta, H. V. 2004b *Rainfall-runoff Modelling in Gauged and Ungauged Catchments*. World Scientific.

Wang, T.-C. & Lee, H.-D. 2009 Developing a fuzzy TOPSIS approach based on subjective weights and objective weights. *Expert Systems with Applications* **36** (5), 8980–8985. http://dx.doi.org/10.1016/j.eswa.2008.11.035.

Wheater, H., Jakeman, A. & Beven, K. 1993 *Progress and Directions in Rainfall-runoff Modelling. Modelling Change in Environmental Systems*. Department of Economics, University of Western Australia, Perth, pp. 101–132.

WMO 2011 *Manual on Flood Forecasting and Warning*. World Meteorological Organization, WMO-No. 1072, Geneva, Switzerland.

Xu, C.-Y. 1999 Estimation of parameters of a conceptual water balance model for ungauged catchments. *Water Resources Management* **13** (5), 353–368.

Xu, C.-y. 2003 Testing the transferability of regression equations derived from small sub-catchments to a large area in central Sweden. *Hydrology and Earth System Sciences Discussions* **7** (3), 317–324.

Yao, C., Zhang, K., Yu, Z., Li, Z. & Li, Q. 2014 Improving the flood prediction capability of the Xinanjiang model in ungauged nested catchments by coupling it with the geomorphologic instantaneous unit hydrograph. *Journal of Hydrology* **517**, 1035–1048. http://dx.doi.org/10.1016/j.jhydrol.2014.06.037.

Yew Gan, T., Dlamini, E. M. & Biftu, G. F. 1997 Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. *Journal of Hydrology* **192** (1), 81–103.

Young, A. R. 2006 Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model. *Journal of Hydrology* **320** (1–2), 155–172. DOI:10.1016/j.jhydrol.2005.07.017.

Yu, L. & Lai, K. K. 2011 A distance-based group decision-making methodology for multi-person multi-criteria emergency decision support. *Decision Support Systems* **51** (2), 307–315. http://dx.doi.org/10.1016/j.dss.2010.11.024.

Zhang, Y. & Chiew, F. H. S. 2009 Relative merits of different methods for runoff predictions in ungauged catchments. *Water Resources Research* **45** (7). DOI:10.1029/2008wr007504.

Zhang, Y., Vaze, J., Chiew, F. H. S., Teng, J. & Li, M. 2014 Predicting hydrological signatures in ungauged catchments using spatial interpolation, index model, and rainfall–runoff modelling. *Journal of Hydrology* **517**, 936–948. http://dx.doi.org/10.1016/j.jhydrol.2014.06.032.

Zhao, R. 1992 The Xinanjiang model applied in China. *Journal of Hydrology* **135** (1), 371–381.

Zhao, R., Zhang, Y., Fang, L. & Zhang, Q. 1980 The Xinanjiang model. In: *Hydrological Forecasting Proceedings Oxford Symposium*, IASH, Oxford, pp. 351–356.

Zou, Z.-h., Yun, Y. & Sun, J.-n. 2006 Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. *Journal of Environmental Sciences* **18** (5), 1020–1023. http://dx.doi.org/10.1016/S1001-0742(06)60032-6.

First received 20 March 2016; accepted in revised form 17 August 2016. Available online 6 October 2016