Journal of Hydrology 590 (2020) 125488

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

A rational performance criterion for hydrological model

Dedi Liu

State Key Laboratory of Water Resources and Hydropower Engineering Science, Hubei Key Laboratory of Water System Science for Sponge City Construction, Wuhan University, Wuhan 430072, China

A R T I C L E I N F O

This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Saket Pande, Associate Editor

Keywords:
Nash-Sutcliffe efficiency
Kling-Gupta efficiency
Model performance criterion
Sampling statistic
Hydrological model

A B S T R A C T

Performance criteria are essential for hydrological model identification or its parameters estimation. The Kling-Gupta efficiency (*KGE*), which combines the three components of Nash-Sutcliffe efficiency (*NSE*) of model errors (i.e. correlation, bias, ratio of variances or coefficients of variation) in a more balanced way, has been widely used for calibration and evaluation hydrological models in recent years. However, the *KGE* does not take a reference forecasts or simulation into account and still underestimates of variability of flow time series when optimizing its value for hydrological model. In this study, we propose another performance criterion as an efficiency measure through reformulating the previous three components of *NSE*. Moreover, the distribution function of the new criterion was also derived to analyze uncertainties of the new criterion, which is originated from the distinction between the theoretical or population statistic and its corresponding sampling properties. The proposed criterion was tested by calibrating the “abcd” and XAJ hourly hydrological models at monthly and hourly time scales data for two different case study basins. Evaluation of the results of the case study clearly demonstrates the overall better or comparable model performances from the proposed criterion. The analysis of the uncertainties of the new criterion based on its distribution probability function suggests a rational approach to distinguish between the probabilistic properties and behavior of the theoretical statistics and the rather different sampling properties of estimators of those statistics when computed from data.

1. Introduction

Hydrological model is being increasingly and widely applied in the design, planning and management of water resources systems under changing climates, land use, and other anthropogenic shifts (Farmer and Vogel, 2016). In the context of hydrologic modelling, evaluating the performance of the models is essential for guiding model identification and also for their parameters estimation (Santos et al., 2018). The performance is usually driven by measure of the goodness-of-fit, which provides a quantitative and objective assessment of the agreement between observed and simulated hydrological data (Beven, 2001; Cramér, 1946; Legates and McCabe, 1999; Pechlivanidis et al., 2012). There have been a number of criteria based on the time series error metrics. Jackson et al. (2019) had reviewed over 60 different criteria along with various common modifications with their strengths and weaknesses. The open source HydroErr library of these criteria had also been presented by them to facilitate greater use of them. And most of these original criteria are a function of the residuals in the modeled and measured quantities, and emphasize different systematic and/or dynamic behaviors within the hydrological system (Pechlivanidis et al., 2011, 2012; Jackson et al., 2019; Oreskes et al., 1994). In order to reduce the influences of outliers and the impacts of high-flow regimes on these criteria, various prior transformations (e.g. a logarithmic, inverse or square-root transformation) on the simulated and observed flow time series have also been used for calculating their values. This is commonly done within the Nash-Sutcliffe efficiency criterion (*NSE*, defined by Nash and Sutcliffe, 1970), which has been one of the most popular criteria used in hydrological modelling in the past few decades (Krause et al., 2005; Oudin et al., 2006; De Vos and Rientjes, 2010; Pushpalatha et al., 2012; Santos et al., 2018). However, Gupta et al. (2009) clearly demonstrated that discharge variability is not correctly taken into account in the criterion using a decomposition of *NSE* based on the correlation, bias and ratio of variances. Combining these three components of *NSE* (i.e. correlation, bias, ratio of variances or coefficients of variation) in a more balanced way, the Kling-Gupta efficiency (*KGE*) and its extension *KGE′* (Kling et al., 2012) are gaining dominance for hydrological model calibration in recent literature (Kling et al., 2012; Hirpa et al., 2018; Becker et al., 2019; Quintero et al., 2020).

In the case of the *KGE* or *KGE′* criteria, a tendency towards underestimation of flow variability is not severe as with the *NSE* (Gupta et al. 2009; Kling et al., 2012; Santos et al., 2018), but there still need improvement on this underestimation for its importance of the extreme flow simulation. Prior transformations on flow before computing *KGE* had been adopted to put more weight on low flows for increasing

E-mail address: dediliu@whu.edu.cn.

https://doi.org/10.1016/j.jhydrol.2020.125488
Received 28 March 2020; Received in revised form 24 August 2020; Accepted 28 August 2020
Available online 08 September 2020

variability simulation. The logarithmic transformation are often used (Pechlivanidis et al, 2014; Seeger and Weiler, 2014; Beck et al., 2016; Quesada-Montano et al., 2018) but logarithmic transformation may lead to several numerical flaws, potentially resulting in a biased evaluation of model performance (Santos et al., 2018). Garcia et al. (2017) found the logarithmic transformation is inadequate to low flow simulation and had applied root-squared, inverse transformations on flows and the flow duration curve. In order to take the highly skewed flow time series and model simulation errors into account, a modification of *KGE* towards a non-parametric criterion was proposed by Pool et al. (2018). As the approach widely employed for satisfying the independence hypotheses uses one extreme flow event only within a time interval (such as a year), the mismatch of the magnitude and timing of high or low flows from the non-parametric criterion will bring uncertainties for estimating hydrological frequency even though their proposed modified non-parametric criterion can result in better agreement between simulated and observed flow than the original formulation.

As the *KGE* or its extension *KGE′* and their modified forms are the Euclidian distances to an ideal value in a three-dimensional space defined by their components of the modelling error, there are no reference forecasts or simulation, which can make their values or scores more reasonable and understandable for modeler. Mizukami et al. (2019) further proved the importance of the improvement of flow variability for flood magnitude estimates at the return period. An improved criterion is still worth developing to further correct the underestimation of variability of flow with taking the advantages of *KGE* and a reference forecasts or simulation into account as time series.

Although a large volume of literature exists comparing the advantages and disadvantages of performance criteria (Krause et al., 2005; Jackson et al., 2019), few studies have sought to evaluate and compare the performance criteria as statistics and distinguish between the probabilistic properties and behavior of the theoretical statistics and the rather different sampling (statistical) properties of estimators of those statistics when computed from data (Barber et al., 2019). For example, the coefficient of determination R^2 and Pearson correlation coefficient ρ are distinguished for normal and non-normal distribution data (Barber et al., 2019). Sampling (statistical) properties of *NSE* was also noted for data splitting for the calibration and validation (according to Oreskes et al. (1994), model "evaluation" will be used instead of "validation" in this paper) of hydrological models even there is no a generally agreed probability distribution function of *NSE* (Liu et al., 2018). Therefore, the distinction between the theoretical or population statistic and the sampling properties of a performance criterion should be taken into hydrological modelling identification or evaluation.

The aim of this study was therefore to propose a new performance criterion by reformulating the previous three components of *NSE* or *KGE* in a more rebalanced and rational way from the model skill scores aspects. And the uncertainties associated with distinctions between the theoretical or population statistic and the sampling properties was exploited through deriving the distribution function of the proposed criterion. In the remainder of the paper, the properties of maximizing *NSE* and *KGE* are explored firstly. A rational performance criterion with improving the optimal values of variabilities of simulation flow time series is proposed, and its distribution function is derived. Within the context of the "abcd" and XAJ hydrological model, the proposed criteria is tested with case studies. Finally, conclusions and discussion on some possible ways forward are stated.

2. Theoretical framework

2.1. The conditions for maximizing NSE and KGE

The previous work had decomposed the *NSE* into three distinctive components that are correlation, the bias, and a measure of relative variability in the simulated and observed values (Murphy, 1988; Weglarczyk, 1998; Gupta et al., 2009).

$$NSE = 2\alpha r - \alpha^2 - \beta_n^2 \tag{1}$$

with $\alpha = \frac{\sigma_s}{\sigma_o}$, $\beta_n = \frac{(\mu_s - \mu_o)}{\sigma_o}$, r is the linear correlation coefficient between the observed and the simulated data set.

If $k_1 = \frac{Cov_{so}}{\sigma_s\sigma_o}\frac{\sigma_s}{\sigma_o} = r\frac{\sigma_s}{\sigma_o} = \alpha r$ and $k_0 = \mu_s - r\frac{\sigma_s}{\sigma_o}\mu_o = \mu_s - \alpha r\mu_o = \mu_s - k_1\mu_o$, $NSE = 2k_1 - \left(\frac{k_1}{r}\right)^2 - \left(\frac{k_0 + k_1\mu_0 - \mu_0}{\sigma_0}\right)^2$. The μ_o and σ_o are determined from oberserved data series and are independent of model calibration or evaluation. While the r can preserve the shape of the hydrograph in time, it is insensitive to the magnitude of flows (Pechlivanidis, et al., 2014). The components of k_1 and k_0 in performance criterion are the focus of our study from the hydrological frequency analysis aspect, which is directly estimated by the magnitude of flows. Taking the first derivative of *NSE* with respect to k_1 and k_0, respectively as shown in the following Eqs. (2) and (3), the maximum value of the *NSE* can be obtained when the Eqs. (2) and (3) equal zero. And then $k_1 = \alpha r = r^2$ (resulting in $\alpha = r$), $k_0 = \mu_s - r^2\mu_o = \mu_o(1 - r^2)$ and resulting in $\mu_o = \mu_s$ and $\beta = 1$.

$$\frac{\partial NSE}{\partial k_1} = 2 - \frac{2k_1}{r^2} - \frac{2}{\sigma_o^2}(k_0 + k_1\mu_0 - \mu_0)\mu_0 \tag{2}$$

$$\frac{\partial NSE}{\partial k_0} = -\frac{2}{\sigma_o^2}(k_0 + k_1\mu_0 - \mu_0) \tag{3}$$

As the *KGE* or *KGE′* means the Euclidian distance to an ideal value in a three dimensional space defined by three components of the modelling error:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2},$$
$$KGE' = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{4}$$

with $\beta = \frac{\mu_s}{\mu_o}$, $\gamma = \frac{\alpha}{\beta}$, the maximum value of the *KGE* can also be easily obtained by taking the first derivative of *KGE* and *KGE′* with respective $k1$ and $k0$ (Details of their first derivatives are given by the Eqs. (A.1)–(A.4) in Appendix). Both of them are maximizing when $\mu_s = \mu_o$ (resulting $\beta = 1$) and. $k_1 = r$ (resulting $\alpha = 1$), $k_0 = \mu_s - r\mu_o = \mu_o(1 - r)$.

For a desirable simulation of X_s, on $\hat{X}_s = X_o$ for all observation. This can be written as a regression model $\hat{X}_s = k_1X_o + k_0$, with a desirable model fit for $k_1 = 1$ and $k_0 = 0$ (Murphy, 1988; Gupta et al., 2009). Thus, there are two simulation processes. One is the simulation from hydrological model, the other is the simulation from the line regression model that based on the observation and simulation results from hydrological model. As the optimal value of k_1 and k_0 that maximizes *NSE* and *KGE* or *KGE′* are given by a model simulation for which k_1 equals r^2 (for *NSE*) or r (for *KGE* and *KGE′*) and k_0 equals $\mu_0(1-r^2)$ (for *NSE*) or $\mu_0(1-r)$ (for *KGE* and *KGE′*). As the characteristics of *KGE* and *KGE′* are similar, only *KGE* will be discussed henceforth in this paper). Since r will often be smaller than unity, this means that in maximizing *KGE* we will tend to select values (the means simualtion) of for k_1 (< 1) that underestimating the means observation X_o especially in the high flows. The value of r^2 is smaller than that of r, the underestimation will be more severe for *NSE*. More precisely, the hydrological models or their parameters sets that generate simulated flows that underestimating the high flows after the compensation effect k_0 (> 0). In other words, If $\varepsilon = \hat{X}_s$ - $X_o = k_0 + (k_1 - 1)X_o$ and $k_1 < 1$, ε tends to be negative when X_o is bigger enough for offsetting the positive compensation effect k_0 (> 0). However, the negative k_0 value can also balance the bias to make ε *to be zero.*

2.2. A rational performance criterion derived

As the discussed in Section 2.1, the maximum or potential values of k_1 equal r^2 for *NSE* and r for KGE, respectively. And both of them are often less than unity. One way to overcome this issue is by taking the first derivative of a criterion with respect to k_1 and k_0 that the maximum value of the criterion is obtained when k_1 is approaching one and k_0 is approaching zero (It is defined as Requirement I). As a criterion can be represented by a skill score that generally defined as measures of the relative accuracy of forecasts of interest relative to the accuracy of forecasts produced by standard of reference (Murphy and Daan, 1985; Murphy, 1988), a standard reference is also important for understanding the criterion. And the mean observed data set, which is an easy reference forecast and has been adopted in *NSE*, will be employed in this study (It is defined as Requirement II). Taking the above two requirements, a new rational criterion, called Liu-Mean Efficiency (*LME*) referring to *NSE* and *KGE*, can be derived as the following Eq. (5) (details of the process are given by the Eq. (A.5) in Appendix).

$$LME = 1 - \frac{\sqrt{[(\mu_o k_1 - \mu_o)^2 + (\mu_s - \mu_o)^2]}}{\mu_o} = 1 - \sqrt{[(k_1 - 1)^2 + (\beta - 1)^2]} \tag{5}$$

where $k_1 = r\frac{\sigma_s}{\sigma_o} = \alpha r$ and $\beta = \frac{\mu_s}{\mu_o}$. Note that *LME* in Eq. (5) is a function of the forecasts or simulation results from hydrological model, the standard reference forecasts or simulation is the mean value of observation μ_0. The first term of the *LME* is the square of the difference between the slope of the regression line (k_1) and one, where the k_1 describes the relationship between the expected values of the observations and the forecasts or simulation. The k_1 is equal to unity, which means an obviously desirable characteristic of regression line in the context of forecast or simulation verification. If the k_1 is not equal to unity, it implies that the conditional expected values of the observations are not equal to the corresponding forecasts or simulation. This term represents a nondimensional measure of the conditional bias in the forecast or simulation. The second term of the *LME* is the square of the difference between the mean simulation or forecasting and the mean observation, divided by the mean observation. It is a nondimensional measure of the unconditional or overall bias in the forecasts and vanishes only for unbiased forecasts or simulations. Moreover, this term is related to the constant or intercept term (k_0) scaled by the mean observation in the linear regression model. Thus, it is evident that *LME* consist of the conditional bias in the forecast or simulation, as reflected by k_1, and the unconditional bias in the forecast or simulation, as reflected by the k_0 (scaled by the mean observation).

The criterion can be maximized with respect to k_1 and k_0 by taking the partial derivatives of *LME* with respect to k_1 and k_0 and setting the resulting equations equal to zero. As the value of μ_o is bigger than zero, the solution of the resulting equations in terms of k_1 and k_0 are $k_1 = 1$, and $k_0 = \mu_s - \mu_o$ respectively, which is the smallest among the results from *NSE* and *KGE* (Eq. (A.6) in Appendix A.).

The *LME* can also be represented by the components r, α and β as the following Eq. (6):

$$LME = 1 - \sqrt{[(k_1 - 1)^2 + (\beta - 1)^2]} = 1 - \sqrt{[(r\alpha - 1)^2 + (\beta - 1)^2]}. \tag{6}$$

It is also easy to show, by talking the first derivative of *LME* (in Eq. (6)) with respect to r, α and β that the maximum value of *LME* is obtained when $r\alpha = 1$ and $\beta = 1$ (the Eqs. (A.6) and (A.7)).

Taking the slope k_1 of the regression line and bias k_0 when regression the simulated against the observed values, we note that when $\beta_n = 0$ or $\beta = 1$ (i.e. $\mu_s = \mu_o$) and $k_1 = r^2$ for *NSE* or $k_1 = r$ for *KGE* and $k_1 = 1$ for *LME*, then the maximum (potential) values for *NSE*, *KGE* and *LME* equal their k_1 values. Fig. 1 illustrates the relationships of *NSE* (Fig. 1(a)), *KGE* (Fig. 1(b)) with r and k_1, while assuming that $\beta_n = 0$ or $\beta = 1$ and $k_1 > 0$. For a given r, the optimal k_1 for maximizing these two criteria equals maximizing their k_1 values. The relationship of *LME* with k_1 and k_0/μ_o can be shown as Fig. 1(c), the maximum values of *LME* can only be achieved by k_1 is approaching to unity as β is approaching to zero. Since r will often be smaller than unity, this means maximizing *NSE* and *KGE* we tend to select a value for k_1 that underestimate the slope. The high values (peak flows) and the low values are always the most concerned in hydrological practice. However, the maximum values of *NSE* and *KGE* (i.e. equal their k_1 values) are less than unity (even *KGE* has been improved from r^2 to r), and this smaller slope value k_1 tends to underestimate the mean of high values (peak flows) and overestimate the mean of low values. Although the optimal k_1 value with combining r as shown in Fig. 1 is not always possible with a hydrological model due to restrictions imposed by the model structure, feasible parameter values and input–output data (Gupta et al., 2009), the proposed *LME* provides a possibility to make k_1 value is unity through improving the α value under the r value is not unity condition (Fig. 2). Therefore, maximizing *NSE*, *KGE* and *LME* will ultimately maximize the slope k_1 of the regression line when its corresponding k_0 value is assumed or approaching to be their ideal value. The optimal *LME* value will directly equal k_1 value and its value might be unity, which will reduce the overestimation of the mean of low flow and underestimation of the mean of high flow.

The relationships of *NSE*, *KGE* and *LME* with r and α are also illustrated in Fig. 2, while $\beta_n = 0$ or $\beta = 1$ and $2 > \alpha > 0, r > 0.5$. For a given r the optimal α for maximizing *NSE* lies on the 1:1 line, although the ideal value of α is on a horizontal line at 1.0 (Gupta, et al., 2009). The horizontal line ($\alpha = 1.0$) is also the optimal value and ideal value for maximizing *KGE* while the optimal values for maximizing *LME* is the curve line $\alpha r = 1.0$ with its ideal value $\alpha = 1.0$ and $r = 1.0$. The interplay between α and r shown in Fig. 2 is likely to be of importance for the hydrological model calibration that is optimized with *NSE* or *KGE* or *LME*. The likelihood of underestimating of the variability in the flows will happen by taking the *NSE* criterion. The *KGE* overcomes this underestimation through optimizing the variability error as separating criteria, and the optimal solution with $\alpha r = 1.0$ for *LME* can inflate the variability as the value of r is always smaller than unity while at the same time preserving the means of the observations.

2.3. Confidence intervals of LME

Since values of *LME* are often estimated by $\widehat{LME}$ (as Eq. (7)), which ignores the distinction between the theoretical or population statistic and the sampling properties of its various possible estimators.

$$\widehat{LME} = 1 - \sqrt{[(\hat{k}_1 - 1)^2 + (\hat{\beta} - 1)^2]}. \tag{7}$$

In order to distinguish between probabilistic properties from sampling data and behavior of the theoretical statistics, an interval estimate may be more desirable. If the differences ε_i between $\hat{X}_{s,i}$ (results of the regression model $\hat{X}_{s,i} = \hat{k}_1 X_{o,i} + \hat{k}_0$,) and $X_{s,i}$ (results of the hydrological model) are assumed to be identically and independently distributed as normal distribution with a mean of zero and a variance of σ_ε (A shorthand way of writing this is $\varepsilon_i \sim N(0, \sigma_\varepsilon)$). If $z = \sqrt{(\hat{k}_1 - k_1)^2 + \left[\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right)\right]^2}$ and $\hat{k}_1 - k_1 = Z\sin\theta$, $\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right) = Z\cos\theta (0 \leqslant \theta \leqslant 2\pi)$, the differences between the theoretical or population statistic k_1 and the sampling properties $\hat{k}_1$ are expressed by $k_1 = \hat{k}_1 - Z\sin\theta$, $\beta = \frac{k_0}{\mu_0} = \hat{\beta} + Z(\sin\theta - \cos\theta) - \hat{k}_1$.

According to the Eq. (6), the theoretical or population statistic *LME* can be estimated by

$$LME(Z) = 1 - \sqrt{[(\hat{k}_1 - 1 - Z\sin\theta)^2 + (Z(\sin\theta - \cos\theta) - \hat{k}_1 + \hat{\beta} - 1)^2]},\ 0 \leqslant \theta \leqslant 2\pi \tag{8}$$

The interval estimate of *LME* is dependent on the interval of Z with confidence level $(1-\kappa)$. The density distribution function of Z function is obtained as

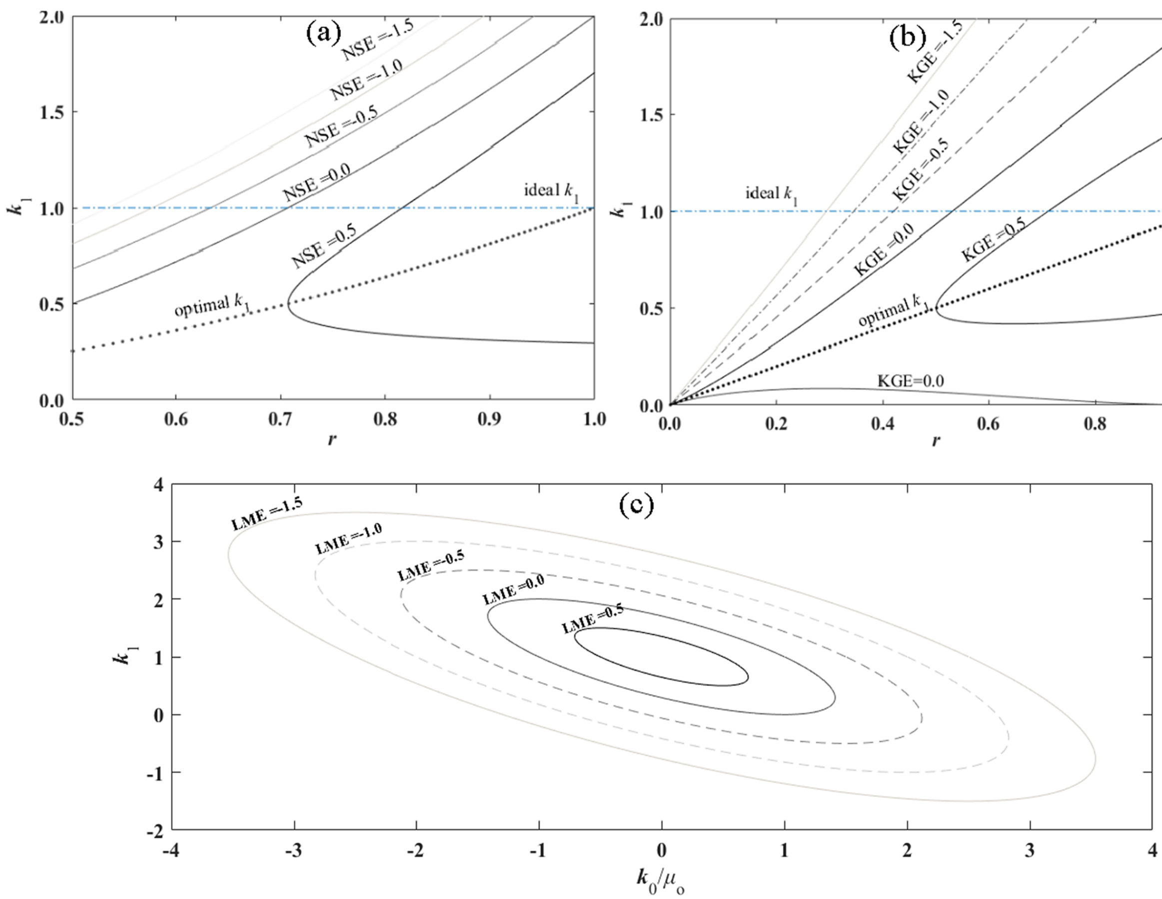

Fig. 1. Theoretical relationships of *NSE* (a), *KGE* (b) with k_1 and r (β_n is assumed to be zero) and *LME* (c) with k_1 and k_0/μ_o.

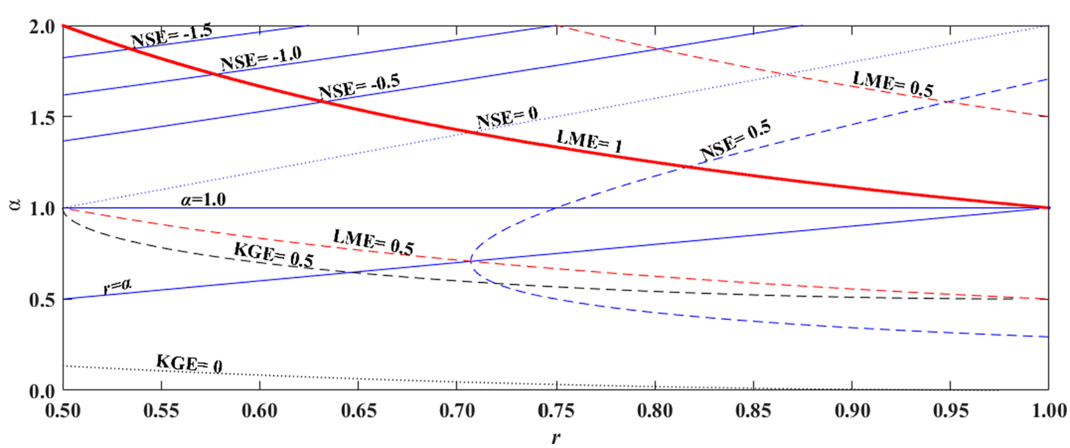

Fig. 2. Theoretical Relationships of *NSE*, *KGE* and *LME* with α and r (β_n is assumed to be zero or β is assumed to be unity).

$$f_Z(z) = \int_0^{2\pi} a'b'z\left[\frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi(n-2)}\,\Gamma\left(\frac{n-2}{2}\right)}\right]^2\left(1+\frac{[a'\sin\theta' z]^2}{n-2}\right)^{\frac{-(n-1)}{2}}\left(1+\frac{(b'\cos\theta' z)^2}{n-2}\right)^{\frac{-(n-1)}{2}}d\theta' \tag{9}$$

where $a' = \frac{\sqrt{S_{xx}}}{\hat{\sigma}}$, $b' = \frac{\mu_0\sqrt{n}}{\hat{\sigma}}$ and $S_{xx} = \sum_{t=1}^{n}(x_{o,t}-\mu_o)^2$, $\hat{\sigma}^2 = \frac{Q_e}{n-2}$, $Q_e = \sum_{i=1}^{n}[x_{s,i}-\mu_s-\hat{k}_1(x_{o,i}-\mu_o)]^2$ (Details of the process are given in Appendix B). Probability density function of Z with various values of a' and b' are illustrated in Fig. 3. Note that for the small a' or b', the probability is very high that the largest value of Z.

If $1-\kappa = \int_0^{Z_\kappa} f_Z(z)dz$, Z_κ can then be determined on the priori chosen probability κ. As the first derivative of *LME* (in Eq. (8)) with respect to Z that the maximum value of *LME* is obtained when $Z_{LME\max} = \frac{\sin\theta(\hat{k}_1-1)+(\sin\theta-\cos\theta)(\hat{k}_1-\hat{\beta}+1)}{\sin^2\theta+(\sin\theta-\cos\theta)^2}$. If $Z_k \leqslant Z_{LME\max}$, the lower and confidence limits are determined by $LME(-Z_\kappa)$, $LME(Z_\kappa)$, respectively. If $Z_k > Z_{LME\max} > -Z_\kappa$, the lower and confidence limits are determined by $LME(-Z_\kappa)$, $LME(Z_{LME\max})$, respectively. If $Z_k > -Z_\kappa > Z_{LME\max}$, the lower and confidence limits are determined by $LME(Z_\kappa)$, $LME(-Z_\kappa)$, respectively.

However, the distribution of *NSE* has not been known exactly (McCuen et al., 2006; Ritter and Muñoz-Carpena, 2013; Liu et al., 2018), and the distribution of *KGE* seems to be impossible to be derived except by the "bootstrap" method. The critical values of these criteria for the goodness-of-fit evaluation are often selected through experience rather than the rigorous derivation. For example, Moriasi et al. (2007) present an excellent review of the ranges of values for *NSE* in

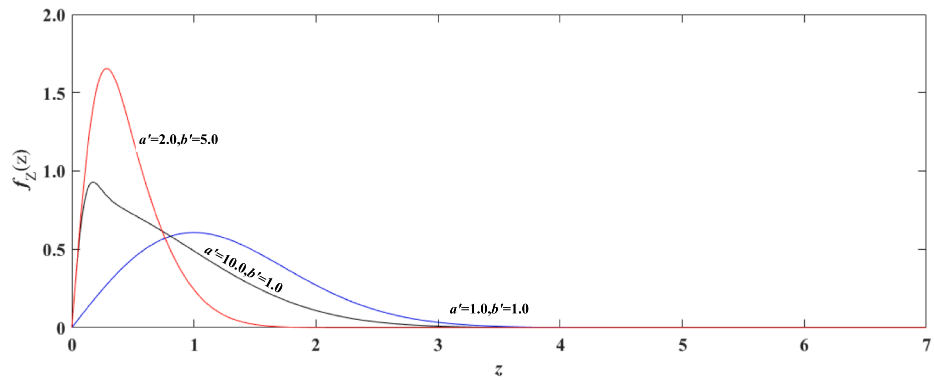

Fig. 3. Probability density function of Z ($n = 200$).

hydrologic model applications based on the studies published before. However, the critical values of the *NSE* are constant and not changed with the number and variation of the observed sample data. Although higher value of *NSE* represents the high confidence level of the model performance, its significant level is still not known, which is important to convince the decision maker. In this study, the critical values of the proposed *LME* for a model performance can be determined by $LME(-Z_\kappa)$, $LME(Z_\kappa)$, $LME(Z_{LME\max})$.

3. Case study

To examine and illustrate the implications of the proposed *LME* in hydrological model for water planning or design, we have chosen "abcd" model as the monthly flow prediction for water utilization, and the Xinanjiang (XAJ) model for the flood forecasting. Using $\hat{LME}$ (Eq. (7)) as model performance criteria, two different sets data series were adopted into calibration and evaluation of the models. In order to compare the effects of $\hat{LME}$ on the discharge simulation, *NSE* and *KGE* are also used as model performance criteria. The confidence intervals of *LME* estimated by sample set were also analyzed through the Eqs. (8) and (9).

3.1. Study area

Two basin were used for test the criteria in our study. One is the Xunhe River basin, a tributary of the Hanjiang River, and the other is the Chongyang River basin, a main upstream tributary of Minjiang River basin. They locations are shown in Fig. 4. The Xunhe River basin has an area of 6314 km^2, with elevation ranging from 200 to 3000 m above sea level. The average annual precipitation and temperature are 850 mm and 15℃, respectively. The Chongyang River basin has an area of 4848 km^2 and a river length of 126 km, with the highest percentage of forest cover in China (Hu et al., 2014).

The mean annual precipitation is 1700 mm, of which 70–80% occurs in the rainy season from April to September due to the typhoon rain and convection rain of short duration and high precipitation intensity (Jie, et al., 2016). The basin consists of highly dissected topography with steep slopes and high stream densities. Therefore, this geography and climate make the basin to be high flood risk, which often greatly threatens the safety of life and property.

3.2. Data sets

There are 23 precipitation gauge stations and runoff gauging station to be used in Xunhe River basin, whose locations are shown in Fig. 4. The precipitation gauges provide a data series comprising 31 years (1980–2010) of monthly precipitation and pan evaporation (taken as potential evaporation) measurements. The Xiangjiaping runoff gauge station, which is located at the outlet of the basin, provide the monthly streamflow from 1980 to 2010.

The hourly precipitation data from 2001 to 2013 were obtained from six precipitation gauge stations in the Chongyang River basin. The hourly potential evaporation measurements and runoff data series at the outlet location of the basin were also applied to drive the hydrological model (shown in Fig. 4.).

3.3. Hydrological models

The "abcd" model was originally proposed by Thomas (1981) for national water assessment and was used as monthly runoff model in the Xunhe River basin. Liu et al. (2018) previously applied this model to the same basin. Inputs to the model are rainfall and potential evaporation. The "abcd" model defines two state variables: W_t as available water and Y_t as evaporation opportunity. Y_t is postulated as a nonlinear function of W_t:

$$Y_t = \frac{W_t + b}{2a} - \sqrt{\left(\frac{W_t + b}{2a}\right)^2 - \frac{W_t b}{a}} \tag{10}$$

The parameter a represents the propensity for runoff to occur before the soils are fully saturated; the parameter b is the upper bound of storage in the unsaturated zone above the groundwater table (Thomas, 1981; Wang and Tang, 2014). The total streamflow is computed as

$$Q_t = (1 - c)(W_t - Y_t) + dG_t + \xi_t \tag{11}$$

The parameter $(1-c)$ represents the portion of (W_t-Y_t) into direct runoff. The parameter d is the linear recession coefficient of groundwater storage (G_t) to Groundwater discharge to the stream channel. ξ_t represents model error in month t.

The XAJ model is a conceptual rainfall-runoff model proposed by Zhao in the 1970s (Zhao et al.1980) and was used for our flood simulation in the Chongyang River basin. The runoff generation of the model is based on the concept of repletion of storage and storage capacity curve, which can solve the problem of the unevenly distributed soil moisture deficit (Zhao, 1992; Zhao et al., 1995). Unit hydrograph and Muskingin methods are used to simulate the runoff routing processes. The model has been widely and successfully implemented for flood forecasting over the globe (Li et al., 2012; Zhuo et al., 2014; Yan et al., 2016; Zhang et al., 2019). The structure of the model can be found in Zhao (1992) and Zhao et al. (1995).

The "abcd" model has four parameters a, b, c, and d (Fernandez et al., 2000) and the XAJ model contains 15 parameters (Jie et al., 2016). The parameters in two models are calibrated using the SCE-UA (Shuffled Complex Evolution method developed at the University of Arizona) algorithm that has a high probability of succeeding in finding the global optimum (Duan et al., 1992, 1993; Jeon et al., 2014). All the codes of two hydrological models are written in Fortran 90, they can be shared by request for no commercial purposes.

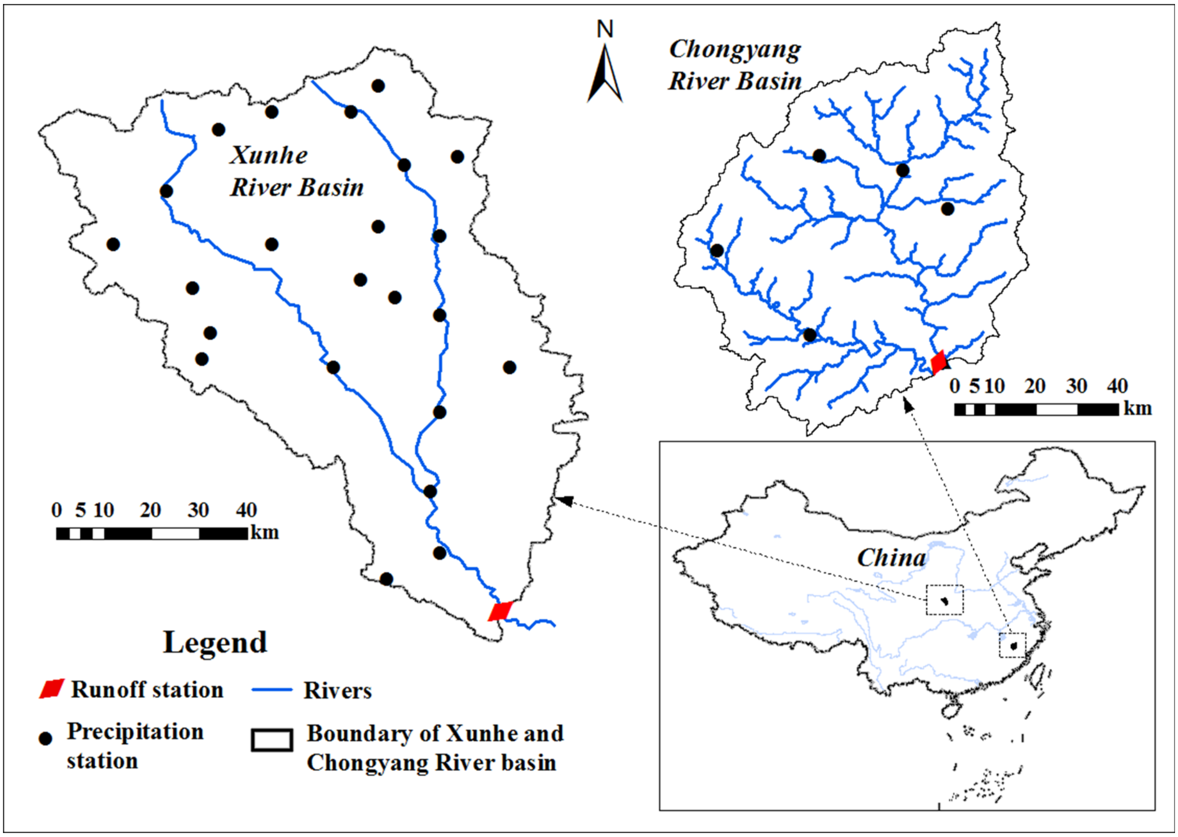

Fig. 4. Locations of the study area and gauges.

4. Results and discussion

4.1. Calibration and evaluation of the "abcd" and XAJ hydrological models

The parameters of hydrological models are calibrated by the running optimization objective functions methods. The "optNSE" method is maximizing *NSE* and yields its optimal runoff simulations, while "optKGE" and "optLME" are maximizing *KGE* and *LME* that yield their optimal runoff simulations, respectively. The available record was equally split into two segments one of which should be used for calibration and the other for evaluation (Klemeš, 1986). As there are 372 months (31 years) data sets in Xunhe River basin, the first 186 months data were used for calibrating the "abcd" hydrological model while the last 186 months data were used for evaluation. In order to test the performance of *NSE*, *KGE* and *LME* on the flood simulation, 30 flood events with different magnitudes were selected and abstracted from the continuous 75, 560 hourly data sets from 2001 to 2013 in Chongyang River basin. These flood events represent various hydrological behaviors and display a wide range of durations and rainfall intensity (Jie et al., 2016). The first 20 flood events were used for XAJ hydrological model calibration and the left 10 flood events were used for evaluation.

The "abcd" hydrological model has been calibrated through "optNSE", "optKGE" and "optLME" methods, respectively. As the available record is sample rather than population, all the three performance criterion can only be estimated by sample and are denoted as $\widehat{NSE}$, $\widehat{KGE}$ and $\widehat{LME}$. Fig. 5 illustrates the scatter plots for monthly runoff simulation. The simulated and observed runoff data fall along the 1:1 line. The optimal runoff simulations by "optNSE" method seem smoother than those by "optKGE" and "optLME" methods. And the $\widehat{LME}$ has the highest variation for both calibration (Fig. 5(a)) and evaluation (Fig. 5(b)) data sets. As the monthly model is always used for low flow simulation, the 10% lowest runoff during the calibration and evaluation period are zooming in the Fig. 5. Most of simulation runoff (≤ 20 m^3/s) is less than the observation. And the results from "optLME" method is the lowest for underestimation. As the three components $\hat{\alpha}$, $\hat{\beta}$ and $\hat{r}$ can also show the features in the simulated and observed runoff values, they are presented by Radar Chart in Fig. 6. The values of $\hat{\alpha}$, $\hat{\beta}$ from the "optLME" method are bigger than those from the "optNSE" and "optKGE"methods while the values of $\hat{r}$ from the three method are slight differences. As the every objective functions have their own merits for the performances of the model, optimization on $\widehat{KGE}$ and $\widehat{LME}$ strongly controls the values that the $\hat{\alpha}$, $\hat{\beta}$ components can achieve, where optimization on $\widehat{NSE}$ constrains these two components only weakly (shown in Fig. 6). Even the $\hat{r}$ is the biggest from the "opt*NSE*" method, the differences of $\hat{\alpha}$, $\hat{\beta}$ from every objective functions are more than that of $\hat{r}$. Therefore, the relative contribution of the criterion components $\hat{r}$ to the final value of the (optimized) model performance is insensitive. When moving from calibration to evaluation period, there are also slight differences of $\hat{\beta}$ and $\hat{r}$ between the two periods. And the bias between all the criterion ($\widehat{NSE}$, $\widehat{KGE}$ and $\widehat{LME}$ value) and their corresponding ideal value (i.e. unity) are mainly deteriorated by their $\hat{\alpha}$. As the slope $\hat{k}_1$ of the regression lines from every optimal methods results is determined by $\hat{\alpha}\hat{r}$, the bigger $\hat{\alpha}$ value (more than unity) can make the $\hat{k}_1$ more close to unity where the $\hat{r}$ always is smaller than unity. Thus, the results from the "optLME" method with bigger $\hat{\alpha}$ value will improve the underestimation and the variability of flows model simulation. While the value of $\widehat{NSE}$ obtained by optNSE is larger than that with optKGE and optLME, the differences are small (show in Fig. 6.). This indicates that by calibrating on $\widehat{KGE}$ and $\widehat{LME}$, we have obtained only a slight deterioration in overall performance as measured by $\widehat{NSE}$. The $\widehat{KGE}$ obtained by optKGE and $\widehat{LME}$ obtained by optLME are also largest among the three optimizing methods, respectively. Their differences with the measurement of $\widehat{KGE}$ and $\widehat{LME}$ are bigger than with the measurement of $\widehat{NSE}$. These indicate calibrating on $\widehat{KGE}$ and $\widehat{LME}$ can bring deterioration in overall performance as measured

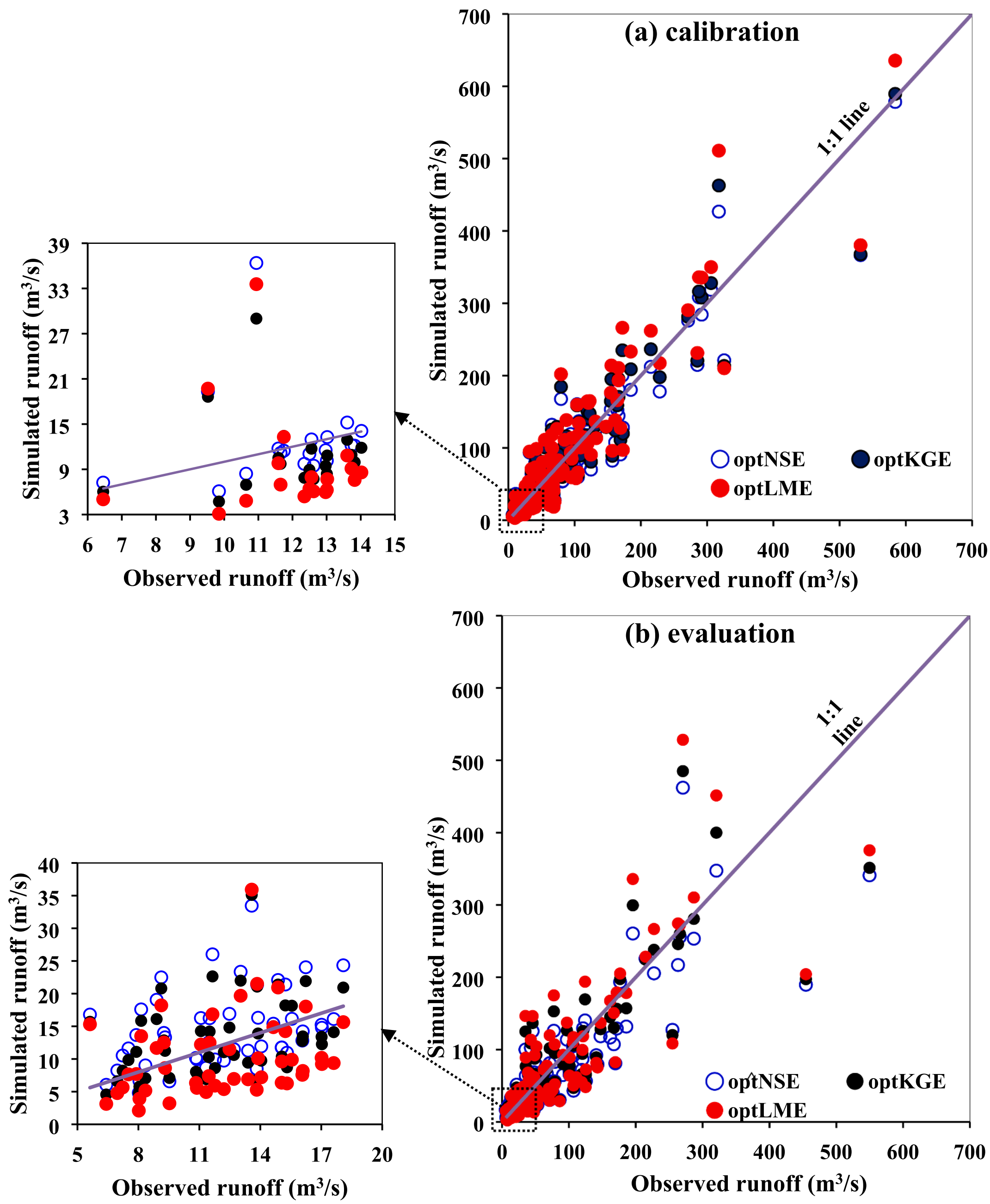

Fig. 5. Scatter plots depicting observed and simulated runoff by the "abcd" hydrological model in Xunhe River basin: (a) calibration period and (b) evaluation period, with the zooming in the 10% lowest values at their left figures.

by $\widehat{KGE}$ and $\widehat{LME}$. There are significant reduction in all the three performance measurements from calibration to evaluation period, the reductions of $\widehat{NSE}$ are smaller than those of $\widehat{KGE}$ and $\widehat{LME}$ by all optNSE, optKGE and optLME.

As theoretical discussed in Section 2.2, the value of $\hat{\alpha}$ is indeed close to $\hat{r}$ and the value of $\hat{k}_1$ is less than unity when optimizing with $\widehat{NSE}$ (shown in Table 1.). When optimizing with $\widehat{KGE}$, the values of $\hat{\alpha}$ are more close to its ideal value of unity, which will make the $\hat{k}_1$ close to unity. In contrast, when optimizing with $\widehat{LME}$, the values of $\hat{k}_1$ are directly more close to its ideal value of unity, which will help to improve low flow simulation (shown in Fig. 5).

The XAJ hydrological model has also been calibrated and evaluated by "optNSE", "optKGE" and "optLME" methods based on the 30 flood events. The scatter plots for hourly runoff simulation are shown in Fig. 7. The simulation for runoff is scattered along the 1:1 line. It is obvious that the variance of runoff simulation from "optLME" method is much bigger than those from the "optNSE" and "optKGE" methods (shown in Fig. 7(a)). Even the means of high flows seems to be overestimated in evaluation period (Fig. 7(b)), the simulation from the "optLME" method are the largest one. And the means of high flow simulation from the "optNSE" method is the smallest one while the results from "optKGE" method are among them. The performances of different optimizing methods are also shown by Radar Chart in terms of criteria (Fig. 8). The values of $\hat{\alpha}$ from the "optLME" method are the biggest during the both calibration and evaluation periods, which represent the highest variability in simulation. However, the values of $\hat{r}$ are smallest. There are no significant differences among the values of $\hat{r}$ from the "optNSE" and "optKGE" methods while the differences among the values of $\hat{\beta}$ are in the middle of $\hat{\alpha}$ and $\hat{r}$. The value of $\hat{k}_1$ from "optLME" method for the XAJ hydrological model is also the most close to unity, which indicates the probability of underestimating the means of extreme high flow is the lowest (shown in Table 2).

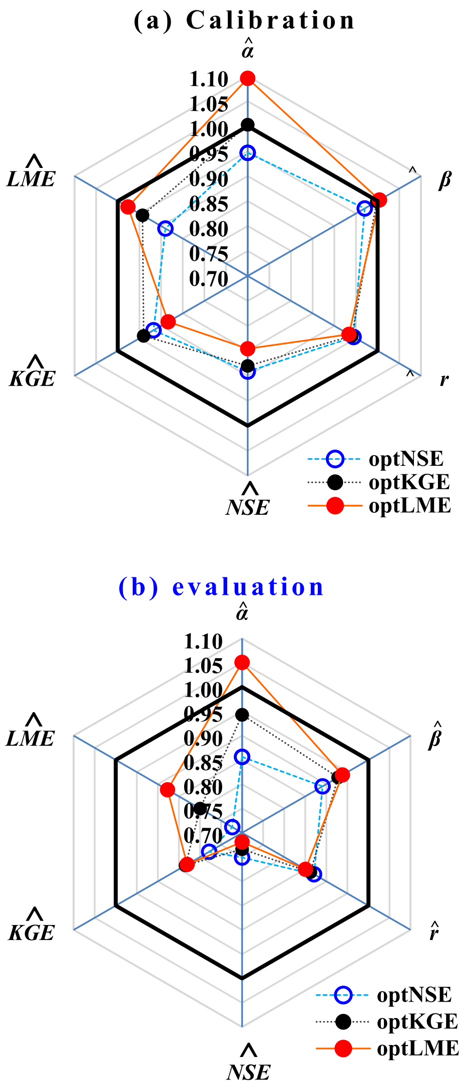

Fig. 6. The criterion values of "abcd" monthly model results by "optNSE", "optKGE" and "optLME" methods: (a) Calibration; (b) evaluation.

4.2. *The confidence intervals estimation of LME*

If the significant level κ is 0.05 (i.e. confidence level is $1-\kappa = 0.95$), the confidence intervals $[-Z_\kappa, Z_\kappa]$ of Z can be determined by the equation $\int_0^{Z_\kappa} f_Z(z)dz = 0.95$. As the parameters a', b' and the sample sizes n of $f_z(z)$ were obtained from the results of "abcd" and XAJ hydrological models, their density and cumulative distribution functions of Z can be determined as shown in Fig. 9. If the parameters a', b' are bigger, the distribution of function will be denser, otherwise will be more dispersive. All the distributions of Z from "optLME" method are the most dispersive while those from "optNSE" method are the densest one in Fig. 9. The distributions during the evaluation periods are more dispersive than those during the calibration periods. Their corresponding cumulative distributions are also shown in Fig. 9 with the subscripts "2". As the density distribution function from the "optLME" method is the most dispersive one, its quantile value Z_κ of the confidence level $(1-\kappa)$ is the largest. In other words, there is the biggest differences between the theoretical or population statistics (i.e. k_1 and β_1) and the sampling properties for the results from the "optLME" method with the same confidences level or significant level. When $Z_k > Z_{LME_{\max}}$, $\frac{\partial LME}{\partial Z} < 0$, and when $Z_k < Z_{LME_{\max}}$, $\frac{\partial LME}{\partial Z} > 0$. As the $Z_{LME_{\max}}$, is the function of variable θ with a period of π, the value of LME is monotonous of variable Z_κ on the condition of θ value. As the first derivative of LME (in Eq. (8)) with respect to θ that the extreme value of LME is obtained when $\theta_{\text{extreme}} = arctg\left(\frac{\hat{k}_1-1}{\hat{\beta}-1}\right) or \left[arctg\left(\frac{\hat{k}_1-1}{\hat{\beta}-1}\right) + K\pi\right]$ (K is an integer). The positions θ of the maximum or the minimum LME is only related to the $\hat{k}_1$ and $\hat{\beta}$, which are not directly depended on the Z_κ. And the confidence interval of LME is not an even distribution function with Z_κ and θ due to the relationship is nonlinear except $Z_\kappa = 0$.

In order to illustrate the confidence intervals of LME resulted from "abcd" and XAJ hydrological models under different optimizing methods (referring the three performance criteria), the performance criterion values had been determined as presented in Table 1 and 2. And the confidence intervals of LME are estimated by the nonlinear Eq. (8) through discretizing the $[-Z_\kappa, Z_\kappa]$ or $[-Z_\kappa, Z_{LME_{\max}}]$ in the first two periods of θ $[0, 2\pi]$ (shown in Fig. 10.). The performances of the "optLME" method for the two hydrological models are the best one in term of the value of $\hat{LME}$, but their confidence intervals of LME might overlap. The ranges of LME confidence interval in the evaluation periods (Fig. 10(b) and (d)) are bigger than those in the calibration period. The maximum and minimum values of LME are also figured out from the Fig. 10. And are listed in the Table 1 and Table 2 as $LME_{\min}$ and $LME_{\max}$. In other words, the uncertainties of LME are larger in the evaluation periods for the both models. It should be noted that the ranges of LME confidence interval or uncertainties are varied with the variable θ rather than constant value except the confidence interval of LME determined by the "optLME" method ($Z_\kappa = 0$) in Fig. 10(c), which is different from the common fixed confidence interval (e.g. $[-Z_\kappa, Z_\kappa]$). Their maximum and minimum values of LME from the different optimizing methods are also found to be at different θ values presented in Tables 1 and 2 as $\theta_{extreme}$. Due to the larger value of a', b' from the objective functions of $\hat{LME}$, the density distribution of Z are dispersive and the value of Z_κ are larger (shown in Fig. 9.) And their uncertainties are larger even they have bigger $\hat{LME}$; however, the uncertainties are varied with the variable θ.

4.3. *Discussion*

The proposed rational performance criterion LME is represented in terms of two components, which measure the slope of the regression line when regressing the observed against the simulated values, and the ratio between the mean simulated and mean observed flows. And LME is reformulated by the previous three components containing in the NSE and KGE. When optimizing the LME, the optimal values of the slope and the ratio (i.e. the bias) are their ideal values, respectively. In contrast, both the optimal values of the slope from the optNSE and optKGE methods are smaller than their ideal values (i.e. unity), which indicates that the means of peak flows will tend to be systematically

Table 1
Performance criterion values resulted from "abcd" hydrological model under different optimal methods.

Method	calibration									evaluation								
	$\hat{k}_1$	$\hat{\beta}$	$\hat{NSE}$	$\hat{KGE}$	$\hat{LME}$	Z_κ	$LME_{\min}$	$LME_{\max}$	θ_{extreme}	$\hat{k}_1$	$\hat{\beta}$	$\hat{NSE}$	$\hat{KGE}$	$\hat{LME}$	Z_κ	$LME_{\min}$	$LME_{\max}$	θ_{extreme}
optNSE	0.894	0.970	0.892	0.917	0.890	0.0623	0.828	0.952	0.412π	0.746	0.891	0.751	0.778	0.723	0.0876	0.636	0.811	0.371π
optKGE	0.943	0.998	0.880	0.940	0.943	0.0683	0.875	0.989	0.489π	0.813	0.928	0.734	0.834	0.800	0.0993	0.701	0.898	0.383π
optLME	1.023	1.004	0.846	0.884	0.976	0.0783	0.899	0.946	0.445π	0.894	0.938	0.681	0.830	0.877	0.1150	0.762	0.992	0.332π

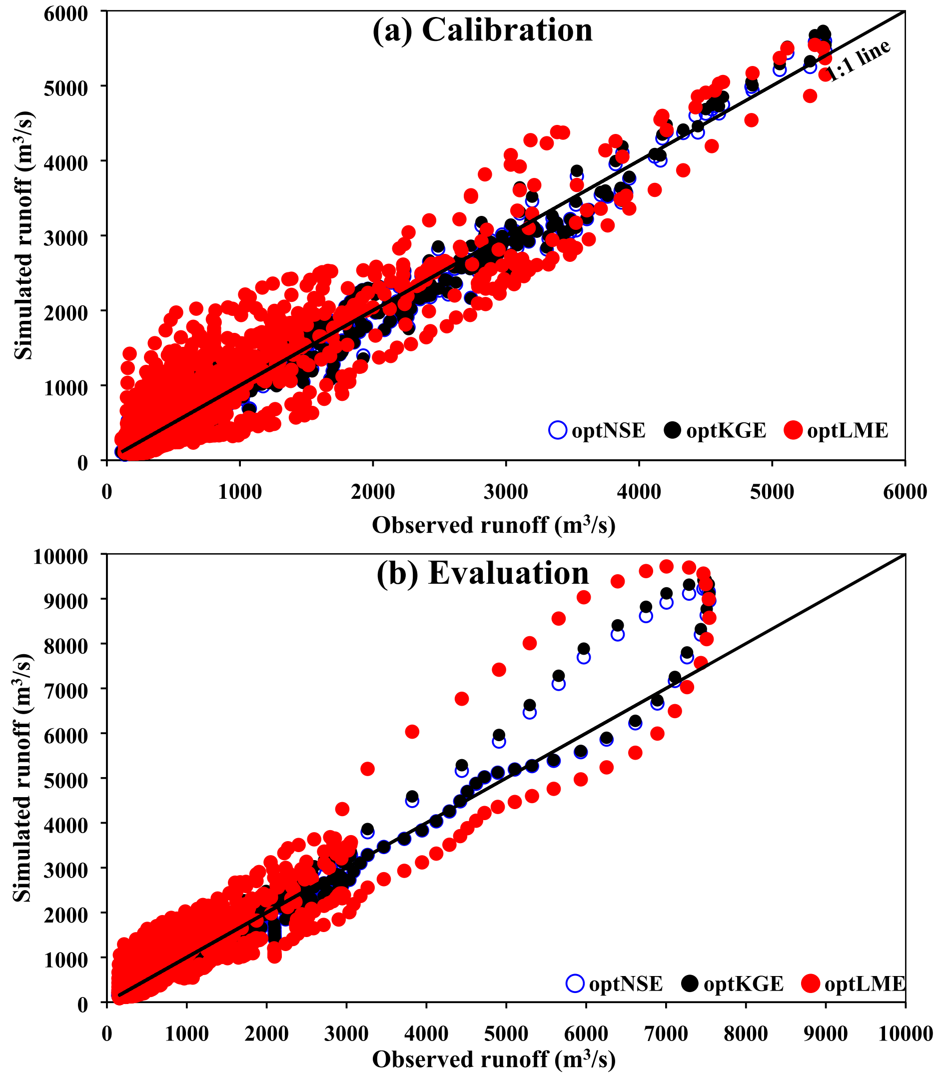

Fig. 7. Scatter plots depicting observed and simulated runoff by the XAJ hydrological model from 2001 to 2013 in Chongyang River basin: (a) Calibration period and (b) Evaluation period.

underestimated and the means of low flows will tend to be overestimated. The optimal values of the variability of flow (α) from the optNSE method is the correlation coefficient (r), which always is smaller than its ideal values (i.e. unity). Even the optimal value of α from the optKGE method is unity, the optimal value of α from the optLME method tend to be bigger than its ideal values where its optimal equation $\alpha r = 1$ and $r \le 1$. By simple theoretical analyzing, the tendency for underestimation in the hydrological model simulation from the optimizing the performance criterion will be improved by our proposed *LME*. These theoretical analyzing were all supported by the results of "abcd" and XAJ hydrological model simulation in our case study. In order to test the impacts of the optNSE, optKGE and optLME methods on the designed flows, Pearson type III distribution was used for low flow and flood hydrologic frequency analysis. And the Linear moment method is used to estimate the parameters of the Pearson type III distribution function based on the simulation results from the two hydrological models and the observed data. The 31 years annual minimum monthly low flow data sets in Xunhe River basin were used for the designed low flow (shown in Fig. 11(a)) while the 12 year annual maximum flood data sets in Chongyang River basin were used for the designed flood (shown in Fig. 11(b)). The designed low flows from the optNSE are bigger than that from those from optKGE. And the results from optLME are the lowest one. In contrast to the results from the observed data, the designed low flows from the optNSE are bigger in both the lower and higher frequencies (e.g. $\le$5% and $\ge$95%). The designed low flows from the optKGE seems to be bigger only in the lower frequencies (e.g. $\le$2%) while the designed low flows from the optLME are also smaller than those from observed data. Thus, it might be more reliable for the designing water supply project by the results from the optLME. The designed floods from the optLME are highest when the frequencies are less 25% (shown in Fig. 11(b)), and these designed flood are often adopted for designing the sizes of flood control project or their operation rules. Although there are many uncertainties from the sample size for the hydrological frequency analysis in our case study, the results from the optLME will be much safer or reliable for water project designing or operation. Therefore, the variability of the flow simulation improved by the optimizing objective functions might decrease the rise or increase the reliability of the designed extreme flow from the hydrological frequency analysis aspect. However, it should be note that the variability of the annual maximum or minimum flow series for the hydrological frequency analysis is different from the variability of the flow simulation. The impact of improvement of variability of the flow simulation on the variability of the annual maximum or minimum flow series is still not clear and also not easily to be deduced. More case studies with different magnitude and sample size flow series might help to figure it out in the future works.

The desirable regression model $\hat{X}_s = X_o$ can only make the means of the simulated and observed series the same rather than can guarantee

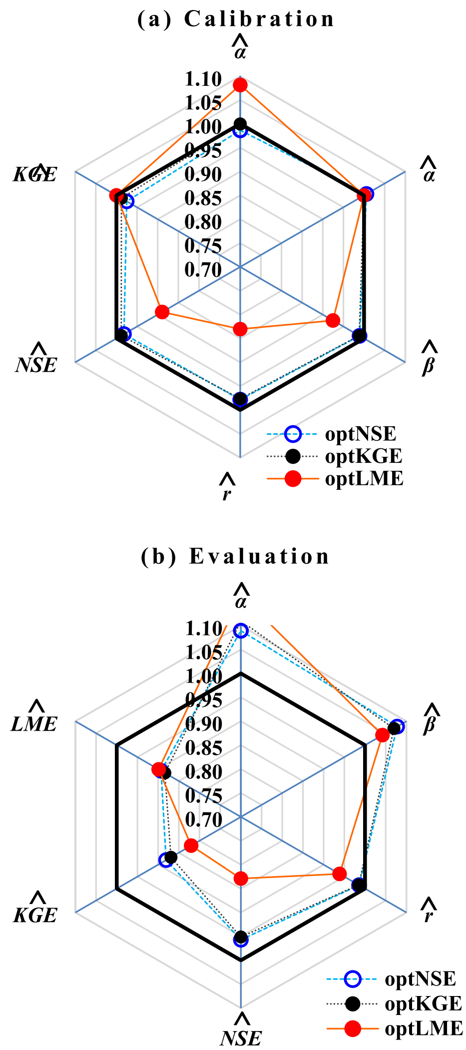

Fig. 8. The criterion values of XAJ hourly model results by “optNSE”, “optKGE” and “optLME” methods:(a) Calibration; (b) Evaluation.

the perfect simulation of whole processes. Actually, the desirable regression model is based on the assumption of a linear relationship between the simulation and observed series. The simulation results from a hydrological model X_s is expressed by $X_s = \hat{X}_s + \varepsilon = \hat{k}_1 X_o + \hat{k}_0 + \varepsilon$ where $\varepsilon_i \sim N(0, \sigma_\varepsilon)$. The linear relationship assumption and the normal distribution with a mean of zero and homoscedasticity assumption for the error or the residuals ε are the two prerequisite conditions for understanding the proposed *LME* value and its confidence intervals. The $\hat{k}_0 = \mu_o(\hat{\beta} - \hat{k}_1)$ and $\hat{k}_0$ is approaching zero as $\hat{\beta}$ and $\hat{k}_1$ are simultaneously optimized to their ideal values (i.e. unity). However, the slope k_1 in *LME* is the combination of α and r where the errors in α and r can compensate each other, the *LME* metric does not guarantee that higher metric values correspond to smaller errors in the hydrological simulation even the ideal value of $\hat{k}_1$ is unity (e.g. the XAJ model application case study as shown in Table 2 and Fig. 7.). Therfore, only high *LME* score cannot reflects smaller simulation error.

As hydrological model performance criterion are often only estimated by sample data rather than population data, the confidences level of criteria should be estimate by its probability distribution. Different from a common way to understand the good or bad score of current criterion (e.g > 0.65 or < 0.5 for *NSE,* Moriasi et al. 2007), a criterion value together with its confidential level is recommended as a goodness-of-fit measure to express the inherent variability in samples of data series. The density distribution function of the criterion can also reflect the goodness-of-fit through its parameters. For example, the parameters a' and b' in Eq. (9) are related to the error or the residual Q_e of regression model and hydrological model simulations. And the bigger value of Q_e, the smaller values of a' and b', which indicate bigger uncertainties for Z and *LME* from the Fig. 3. And vice versa. In XAJ model application case study, the values of a' and b' from optLME method are much smaller than that from optKGE and optNSE, their density functions of Z are more dispersive as shown in Fig. 9(c_1) and (d_1). Their lower bounds of confidential interval might be smaller than that from the optKGE or optNSE even their *LME* scores are bigger (shown in Fig. 9(c_2), (d_2) and Table 2). Only both the *LME* score and its confidential level can represent the goodness-of-fit measure.

Two or more performance criteria are often adopted for model calibration and evaluation, in essence, the proposed *LME* including its score value and confidential level or intervals also reflects more than one performance criteria through the Eq. (6) and the parameters (a' and b') of its density distribution function. These two aspects of LME can not only integrate different aspects of model performances but also can help us understand the meaning of the criterion.

However, the probability distribution functions of *LME* has been derived when the differences ε_i between $\hat{X}_{s,i}$ and $X_{s,i}$ are assumed to be Gaussian distribution while there are no rigorous or analytical probability distribution functions for *NSE* and *KGE*. The sample size and variability will affect the shape of the probability density distribution of *LME*. Different from the common variable with even distribution in its domain, *LME* distribution is varied with an intermedia variable θ that has π periods at a confidence or significant level (i.e. a probability). As the ranges of *LME* at a confidence level are mainly dependent on the values of $\hat{LME}$, there are overlaps for *LME* estimated from optNSE, optKGE and optLME methods. The differences between $\hat{LME}$ from every optimizing objective functions should be taken their confidence levels into account. The sample size n can affect the distribution function of *LME* through the Eq. (9). As the comparison between the results from optNSE, optKGE and optLME methods in our case study, their sample sizes for calibration or evaluation are the same in every study area. The impacts of sample size on the performance criterion will help the sample splitting for the calibration and evaluation of hydrological models (Liu et al., 2018). However, the t distribution in Appendix B is approaching to normal distribution as the n increase (e.g. $n \geq 120$). The impacts of sample size for calibration or evaluation on the distribution function of *LME* or its significant level can be ignored if the sample sizes are enough.

There have been a lot of performance criteria for evaluating the hydrological model. Every criterion has its own peculiarities, so do *LME* proposed in our study. Evaluation of a model performances always

Table 2
Performance criterion values resulted from XAJ hydrological model under different optimal methods.

Method	calibration									evaluation								
	$\hat{k}_1$	$\hat{\beta}$	$\hat{NSE}$	$\hat{KGE}$	$\hat{LME}$	Z_κ	LME_{min}	LME_{max}	$\theta_{extreme}$	$\hat{k}_1$	$\hat{\beta}$	$\hat{NSE}$	$\hat{KGE}$	$\hat{LME}$	Z_κ	LME_{min}	LME_{max}	$\theta_{extreme}$
optNSE	0.976	1.006	0.978	0.982	0.975	0.0078	0.9675	0.9831	0.422π	1.074	1.078	0.957	0.881	0.893	0.0170	0.8755	0.9095	0.242π
optKGE	0.988	1.000	0.976	0.988	0.988	0.0082	0.9798	0.9962	0.500π	1.093	1.070	0.951	0.869	0.884	0.0177	0.8659	0.9013	0.295π
optLME	1.000	1.000	0.830	0.889	1.000	0.0220	0.9780	1.000	[0, π]	1.092	1.042	0.829	0.820	0.899	0.0322	0.8667	0.9311	0.364π

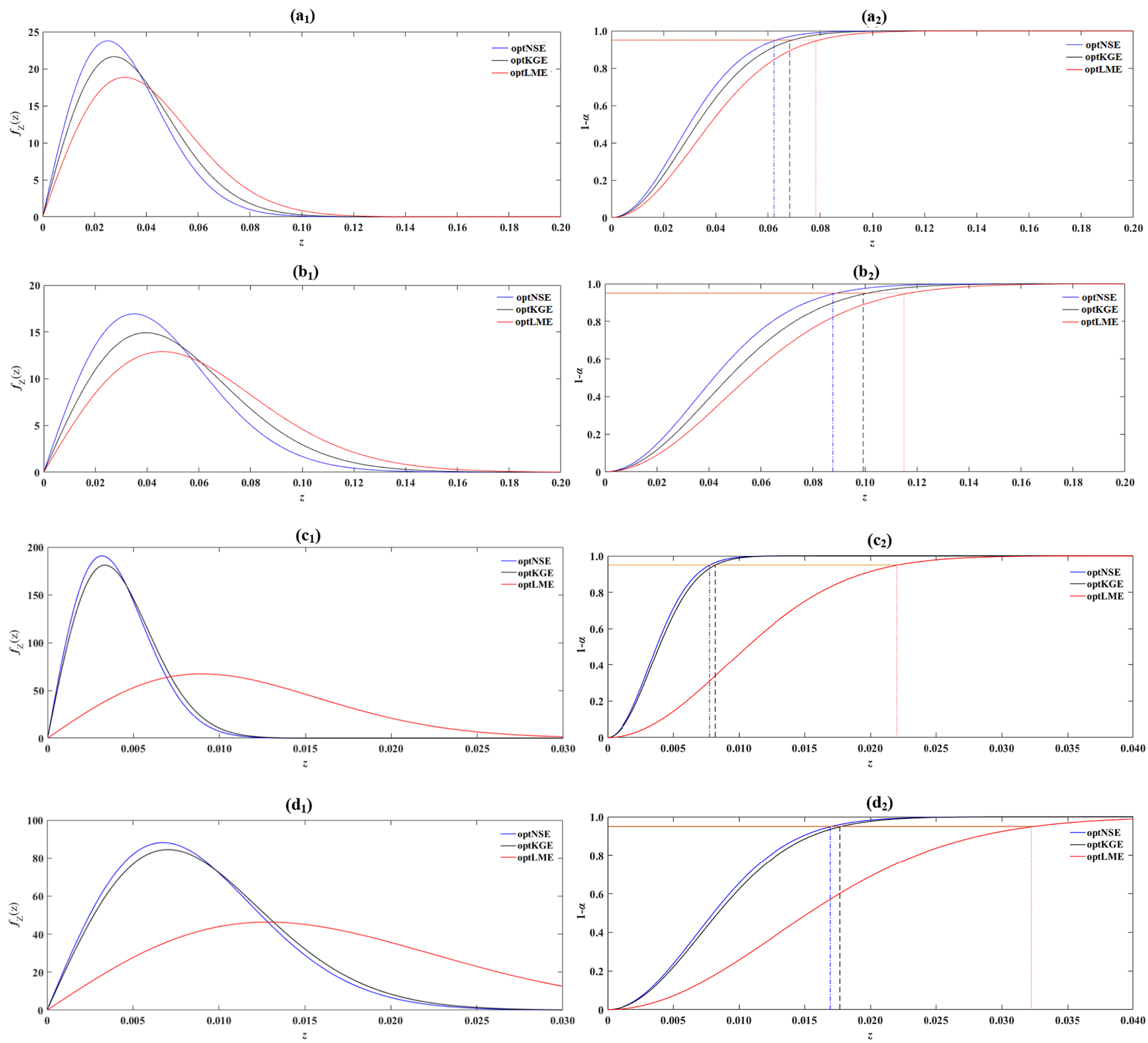

Fig. 9. Density and cumulative distribution functions of Z: (a) and (b) are resulted from the "abcd" hydrological model during the calibration and evaluation periods; (c) and (d) are resulted from the XAJ hydrological model during the calibration and evaluation periods. Subscripts "1" and "2" denote density and cumulative distribution functions, respectively.

depends on the purpose of the application of the model. If a performance criteria is used to evaluate the hydrological performance model, its limitations and advantages of the criteria should be known according to the type of hydrological model application. The primary purpose of our study was not to replace the *NSE* or *KGE*, but to show the advantages of optLME on improvement of the slope of the regression line when regressing the observed against the simulated values and increasing the variability of simulation data series. It is also interesting to note that the performance criteria can be taken as a sampling statistic, which is different from the population statistic and has sampling error. Confidence level of the criteria should be shown or its hypothesis testing should be done for evaluating the model performance. Ultimately, the decision on the performance of a hydrological often depends upon not a single measures but a multiple-criteria framework. Our proposed *LME* not only provides another option of performance criteria with its own advantages on extreme simulation but also its probability distribution function can quantify reliability or risk of the criteria on evaluation.

Although the structure and feasible parameter values of a hydrological model can make the r value not to be the worst (e.g. 0 or negative values), more case studies should be done further to investsigate the significant linear relationship (i.e. r value is big enough) resulted from different hydrological models, time scales, basin charactersitics and so on. Beside the density distribution function of proposed cretia *LME* has been derived, the relationship between *LME* and r (such as jointly distributed function, their trade-off through multiobjective calibration function) is worth to explore their confidence intervals or hypothesis testings. The homoscedasticity assumption ε has been challenged by the heterogeneous non-Gaussian in hydrological simulation (Jiang et al., 2019). Future work can be done to derive a general density distribution function of the proposed creteria (e.g. based on the heteroscedasticity) and simutanously to minimize the variance value during the calibration process.

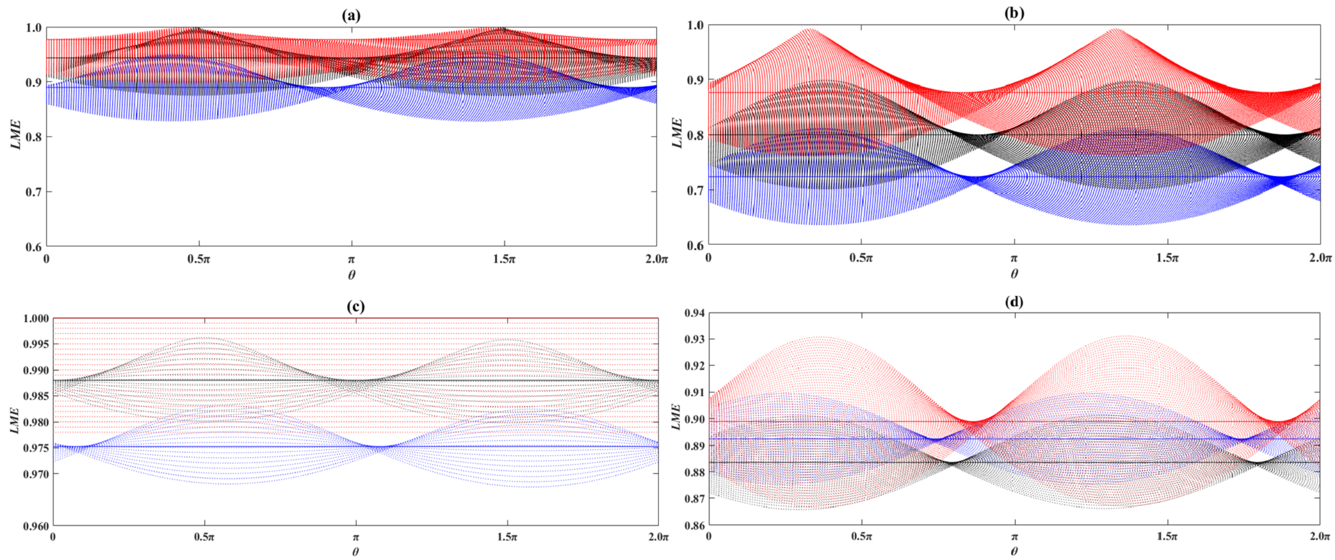

Fig. 10. Confidence intervals of *LME* determined by different optimizing methods: (a) “abcd” model in calibration period; (b) “abcd” model in evaluation period; (c) XAJ model in calibration period and (d) XAJ model in evaluation period. The blue dash curve represents the results from optNSE, the black dash curve line represents the optKGE and the red dash curve line represents the optLME estimated by Eq. (8) with $\kappa = 0.05$. And their straight lines represent $\widehat{LME}$ estimated by the Eq. (7).

5. Summary

In this paper a new rational performance criterion *LME* was derived through its theoretical considerations that serve to highlight advantages associated with variability of hydrological model simulation. As the slopes of the regression lines regressing the observed against the simulated values and the ratio between the mean simulated and mean observed flows are the components of the proposed criterion, these two components value will be approaching to their ideal values when *LME* is used in optimization the criterion. This means that the

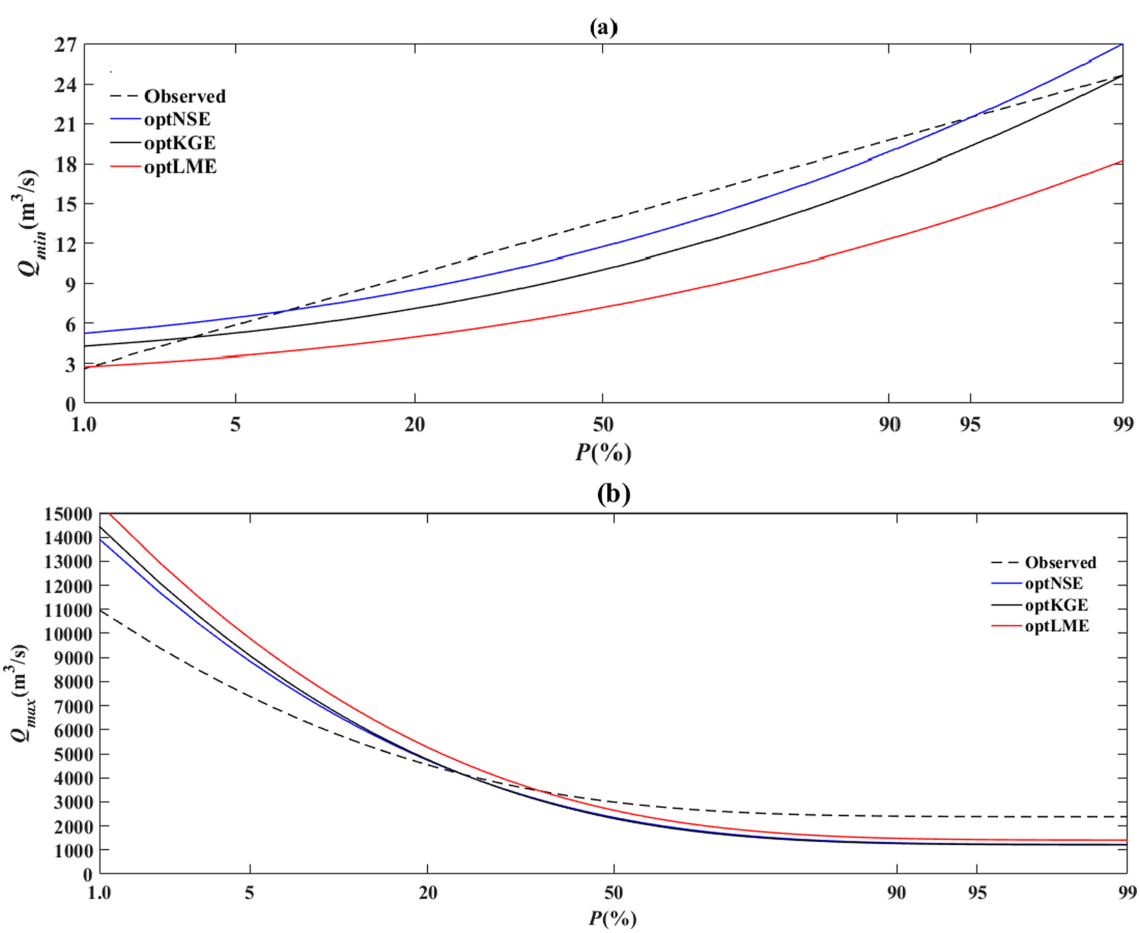

Fig. 11. Impacts of the optNSE, optKGE and optLME methods on the Pearson type III distribution: (a) results of abcd model in Xunhe River basin; (b) results of XAJ model in Chongyang River basin.

underestimation of peak flow or the overestimation of low flow will be mitigated comparing with optimizing on *NSE* and *KGE*. In order to take the impact of sampling error on determining the value of performance criterion into account, its confidence level is estimated after deriving the probability distribution functions of *LME*. As the confidence levels of $\widehat{LME}$ might be overlapped, the differences between every $\widehat{LME}$ should be compared by their confidence level or hypothesis testing through the probability distribution functions of *LME*. Therefore, it may be appropriate to point out that the advantages of our proposed *LME* improving hydrological extreme flow simulation and quantifying the criteria confidence level are general in the sense that it is applicable to all types of forecasts. *LME* was also applied in our case study and had supported its theoretical merits. However, although some aspects of model performance can be reflected in the *LME* score and its confidential level through its parameters (e.g. k_1, β, a', b' and so on), there is no a versatile criterion that can reflect all the concerns of hydrologist at any time and any place. *LME* is also no exception. The performance of hydrological model must be evaluated by an expert hydrologist according to the purposes of hydrological model application, where such an evaluation is best based in the advantages of the criteria. Although the proposed criteria is primarily tested on hydrological model, it can also applicable to dynamic environmental systems models.

Declaration of Competing Interest

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

Acknowledgements

The author gratefully acknowledges the financial support from the National Natural Science Foundation of China (Nos. 51879194 and 51579183). This work is also partly funded by the Ministry of Foreign Affairs of Denmark and administered by Danida Fellowship Centre (File number: 18-M01-DTU). The author thanks the reviewers who makes very constructive comments to improve the quality of this paper.

Appendix A

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} = 1 - \sqrt{(r-1)^2 + \left(\frac{k_1}{r} - 1\right)^2 + \left(\frac{k_0}{\mu_0} + k_1 - 1\right)^2} \tag{A.1}$$

$$\begin{aligned}
\frac{\partial KGE}{\partial k_1} &= -\left[\frac{1}{r}\left(\frac{k_1}{r} - 1\right) + \left(\frac{k_0}{\mu_0} + k_1 - 1\right)\right]\left[(r-1)^2 + \left(\frac{k_1}{r} - 1\right)^2 + \left(\frac{k_0}{\mu_0} + k_1 - 1\right)^2\right]^{-\frac{1}{2}} = 0 \\
\frac{\partial KGE}{\partial k_0} &= -\frac{1}{\mu_0}\left(\frac{k_0}{\mu_0} + k_1 - 1\right)\left[(r-1)^2 + \left(\frac{k_1}{r} - 1\right)^2 + \left(\frac{k_0}{\mu_0} + k_1 - 1\right)^2\right]^{-\frac{1}{2}} = 0 \\
&\Rightarrow k_1 = r,\ k_0 = \mu_0(1-r),\ \mu_o \neq 0 \text{ and } r \neq 0
\end{aligned} \tag{A.2}$$

$$KGE' = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\gamma-1)^2} = 1 - \sqrt{(r-1)^2 + \left(\frac{k_1}{r} - 1\right)^2 + \left(\frac{k_1}{r}\frac{\mu_0}{k_0 + k_1\mu_0} - 1\right)^2} \tag{A.3}$$

$$\begin{aligned}
\frac{\partial KGE'}{\partial k_1} &= -\left[\frac{1}{r}\left(\frac{k_1}{r} - 1\right) + \frac{\mu_0}{r}\frac{k_0}{(k_0 + k_1\mu_0)^2}\left(\frac{k_1}{r}\frac{\mu_0}{k_0 + k_1\mu_0} - 1\right)\right]\left[\sqrt{(r-1)^2 + \left(\frac{k_1}{r} - 1\right)^2 + \left(\frac{k_1}{r}\frac{\mu_0}{k_0 + k_1\mu_0} - 1\right)^2}\right]^{-\frac{1}{2}} = 0 \\
\frac{\partial KGE'}{\partial k_0} &= -\frac{\mu_0 k_1}{r}\frac{-1}{(k_0 + k_1\mu_0)^2}\left(\frac{k_1}{r}\frac{\mu_0}{k_0 + k_1\mu_0} - 1\right)\left[\sqrt{(r-1)^2 + \left(\frac{k_1}{r} - 1\right)^2 + \left(\frac{k_1}{r}\frac{\mu_0}{k_0 + k_1\mu_0} - 1\right)^2}\right]^{-\frac{1}{2}} = 0 \\
&\Rightarrow \frac{k_1}{r}\frac{\mu_0}{k_0 + k_1\mu_0} - 1 = 0 \text{ and} \frac{1}{r}\left(\frac{k_1}{r} - 1\right) = 0 \\
&\Rightarrow k_1 = r,\ k_0 = \mu_0(1-r),\ \mu_o \neq 0 \text{ and } r \neq 0
\end{aligned} \tag{A.4}$$

$$\begin{aligned}
LME &= 1 - \frac{\sqrt{[(\mu_o k_1 - \mu_o)^2 + (\mu_s - \mu_o)^2]}}{\mu_o} = 1 - \frac{\sqrt{[(\mu_o k_1 - \mu_o)^2 + (\mu_o k_1 + k_0 - \mu_o)^2]}}{\mu_o} \\
&= 1 - \sqrt{[(k_1 - 1)^2 + (\beta - 1)^2]} = 1 - \sqrt{[(r\alpha - 1)^2 + (\beta - 1)^2]}
\end{aligned} \tag{A.5}$$

$$\begin{aligned}
\frac{\partial LME}{\partial k_1} &= -\frac{\mu_o^2(k_1 - 1) + (\mu_o k_1 + k_0 - \mu_o)\mu_o}{\mu_o}[(\mu_o k_1 - \mu_o)^2 + (\mu_o k_1 + k_0 - \mu_o)^2]^{-\frac{1}{2}} = 0 \\
\frac{\partial LME}{\partial k_0} &= -\frac{(\mu_o k_1 + k_0 - \mu_o)}{\mu_o}[(\mu_o k_1 - \mu_o)^2 + (\mu_o k_1 + k_0 - \mu_o)^2]^{-\frac{1}{2}} = 0 \\
&\Rightarrow k_1 = 1,\ k_0 = 0,\ \mu_o \neq 0 \text{ or } \mu_o = 0
\end{aligned} \tag{A.6}$$

$$\begin{aligned}
&\left.\begin{aligned}
\frac{\partial LME}{\partial \alpha} &= -(r\alpha - 1)r[(r\alpha - 1)^2 + (\beta - 1)^2]^{-\frac{1}{2}} = 0 \\
\frac{\partial LME}{\partial r} &= -(r\alpha - 1)\alpha[(r\alpha - 1)^2 + (\beta - 1)^2]^{-\frac{1}{2}} = 0
\end{aligned}\right\} \Rightarrow\ \alpha r = 1,\ r \neq 0,\ \alpha \neq 0 \\
&\frac{\partial LME}{\partial \beta} = -(\beta - 1)[(r\alpha - 1)^2 + (\beta - 1)^2]^{-\frac{1}{2}} = 0 \Rightarrow \beta = 1
\end{aligned} \tag{A.7}$$

Appendix B

Distribution function of *Z*:

$$\hat{k}_1 N\left(k_1, \frac{\sigma_\varepsilon^2}{S_{xx}}\right), \hat{\mu}_s N\left(k_1\mu_0 + k_0, \frac{\sigma_\varepsilon^2}{n}\right) \cdot \frac{(n-2)\hat{\sigma}^2}{\sigma_\varepsilon^2} = \frac{Q_e}{\sigma_\varepsilon^2}\chi^2(n-2)$$

where $S_{xx} = \sum_{t=1}^{n}(x_{o,t} - \mu_o)^2$ and $\hat{\sigma}^2(n-2) = Q_e = \sum_{i=1}^{n}(x_{s,i} - \hat{k}_0 - \hat{k}_1 x_{o,i})^2$

$$\frac{\hat{k}_1 - k_1}{\hat{\sigma}_\varepsilon}\sqrt{S_{xx}} = a'(\hat{k}_1 - k_1)t^{\sim}(n-2)$$

$$\frac{\left(\frac{\hat{\mu}_s}{\mu_0}\right) - \left(k_1 + \frac{k_0}{\mu_0}\right)}{\hat{\sigma}_\varepsilon}\mu_0\sqrt{n} = \frac{\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right)}{\hat{\sigma}_\varepsilon}\mu_0\sqrt{n} = \left[\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right)\right]b't^{\sim}(n-2)$$

where $a' = \frac{\sqrt{S_{xx}}}{\hat{\sigma}_\varepsilon}$, $b' = \frac{\mu_0\sqrt{n}}{\hat{\sigma}_\varepsilon}$.

The t distribution with (n-2) degrees of freedom is given by

$$f(t) = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi(n-2)}\Gamma\left(\frac{n-2}{2}\right)}\left(1 + \frac{t^2}{n-2}\right)^{\frac{-(n-1)}{2}}, f(\hat{k}_1 - k_1) = \frac{a'\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi(n-2)}\Gamma\left(\frac{n-2}{2}\right)}\left(1 + \frac{[a'(\hat{k}_1 - k_1)]^2}{n-2}\right)^{\frac{-(n-1)}{2}}$$

$$f\left(\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right)\right) = \frac{b'\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi(n-2)}\Gamma\left(\frac{n-2}{2}\right)}\left(1 + \frac{\left(\left[\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right)\right]b'\right)^2}{n-2}\right)^{\frac{-(n-1)}{2}}$$

As the random variables $\hat{\mu}_s$, $\hat{k}_1$, Q_e are independent, $(\hat{k}_1 - k_1)$ and $\left(\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right)\right)$ are independent. Their joint probability density function is shown by

$$f\left((\hat{k}_1 - k_1)\cdot\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right)\right) = a'b'\left[\frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi(n-2)}\Gamma\left(\frac{n-2}{2}\right)}\right]^2\left(1 + \frac{[a'(\hat{k}_1 - k_1)]^2}{n-2}\right)^{\frac{-(n-1)}{2}}\left(1 + \frac{\left(\left[\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right)\right]b'\right)^2}{n-2}\right)^{\frac{-(n-1)}{2}}$$

If $z = \sqrt{(\hat{k}_1 - k_1)^2 + \left[\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right)\right]^2}$, and then $\hat{k}_1 - k_1 = z\sin\theta'$, $\hat{\beta} - \left(k_1 + \frac{k_0}{\mu_0}\right) = z\cos\theta'$, $0 \leqslant \theta' \leqslant 2\pi$.

$$f(z\cdot\theta') = a'b'z\left[\frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi(n-2)}\Gamma\left(\frac{n-2}{2}\right)}\right]^2\left(1 + \frac{[a'\sin\theta' z]^2}{n-2}\right)^{\frac{-(n-1)}{2}}\left(1 + \frac{(b'\cos\theta' z)^2}{n-2}\right)^{\frac{-(n-1)}{2}}$$

$$f_Z(z) = \int_0^{2\pi} a'b'z\left[\frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{\pi(n-2)}\Gamma\left(\frac{n-2}{2}\right)}\right]^2\left(1 + \frac{[a'\sin\theta' z]^2}{n-2}\right)^{\frac{-(n-1)}{2}}\left(1 + \frac{(b'\cos\theta' z)^2}{n-2}\right)^{\frac{-(n-1)}{2}} d\theta' \tag{B.1}$$

References

Barber, C., Lamontagne, J.R., Vogel, R., 2019. Improved estimators of correlation and R2 for skewed hydrologic data. Hydrolog. Sci. J. https://doi.org/10.1080/02626667.2019.1686639.

Beck, H.E., van Dijk, A.I.J.M., de Roo, A., Miralles, D.G., McVicar, T.R., Schellekens, J., Bruijnzeel, L.A., 2016. Global-scale regionalization of hydrologic model parameters. Water Resour. Res. 52, 3599–3622.

Becker, R., Koppa, A., Schulz, S., Usman, M., aus der Beek, M., Schuth, C., 2019. Spatially distributed model calibration of a highly managed hydrological system using remote sensing-derived ET data. J. Hydrol. 577 123944.

Beven, K., 2001. Rainfall-Runoff modelling. In: The Primer. John Wiley and Sons, Chichester, UK, pp. 1–360.

Cramér, H., 1946. Mathematical Methods of Statistics. Princeton University Press, Princeton.

De Vos, N.J., Rientjes, T.H.M., 2010. Multi-objective performance comparison of an artificial neural network and a conceptual rainfall-runoff model. Hydrol. Sci. J. 52, 397–413.

Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optimiz. Theory App. 76 (3), 501–521.

Duan, Q.Y., Sorooshian, S., Gupta, V.K., 1992. Effective andefficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28 (4), 1015–1031.

Farmer, W.H., Vogel, R.M., 2016. On the deterministic and stochastic use of hydrologic models. Water Resour. Res. 52, 5619–5633.

Fernandez, W., Vogel, R.M., Sankarasubramanian, A., 2000. Regional calibration of a watershed model. Hydrolog. Sci. J. 45 (5), 689–706.

Garcia, F., Folton, N., Oudin, L., 2017. Which objective function to calibrate rainfall-runoff models for low-flow index simulations? Hydrolog. Sci. J. 62 (7), 1149–1166.

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. J. Hydrol. 377, 80–91.

Hirpa, F.A., Salamon, P., Beck, H.E., Lorini, V., Alfieri, L., Zsoter, E., Dadson, S., 2018. Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data. J. Hydrol. 566, 595–606.

Hu, X.S., Wu, C.Z., Hong, W., Qiu, R.Z., Li, J., Hong, T., 2014. Forest cover change and its drivers in the upstream area of the Minjiang River. China. Ecol. Indic. 46, 121–128.

Jackson, E.K., Roberts, W., Nelsen, B., Williams, G.P., Nelson, E.J., Ames, D.P., 2019. Introductory overview: Error metrics for hydrologic modelling-A review of common practices and an open source library to facilitate use and adoption. Environ. Model. Softw. 119, 32–48.

Jeon, J.H., Park, C.G., Engel, B.A., 2014. Comparison of performance between genetic algorithm and SCE-UA for calibration of SCS-CN surface runoff simulation. Water 6 (11), 3433–3456.

Jiang, X.L., Gupta, H.V., Liang, Z.M., Li, B.Q., 2019. Toward Improved probabilistic predictions for flood forecasts generated using deterministic models. Water Resour. Res. 55, 9519–9543.

Jie, M.X., Chen, H., Xu, C.Y., Zeng, Q., Tao, X.E., 2016. A comparative study of different objective functions to improve the flood forecasting accuracy. Hydrol. Res. 47 (4), 718–735.

Klemeš, V., 1986. Operational testing of hydrological simulation models. Hydrol. Sci. J. 31 (1), 13–24. https://doi.org/10.1080/02626668609491024.

Kling, H., Fuchs, M., Paulin, M., 2012. Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. J. Hydrol. 424, 264–277.

Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. 5, 89–97.

Legates, D.R., McCabe, G.J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour. Res. 35 (1), 233–241.

Li, H., Zhang, Y., Vaze, J., Wang, B., 2012. Separating effects of vegetation change and climate variability using hydrological modelling and sensitivity-based approaches. J. Hydrol. 420–421, 403–418.

Liu, D.D., Guo, S.L., Wang, Z.L., Liu, P., Yu, X.X., Zhao, Q., Zou, H., 2018. Statistics for

sample splitting for the calibration and validation of hydrological models. Stoch. Env. Res. Risk. A. 32 (11), 3099–3116.
McCuen, R.H., Knightm, Z., Cutter, A.G., 2006. Evaluation of the Nash-Sutcliffe efficiency index. J. Hydrol. Eng. 11 (6), 597–602.
Mizukami, N., Rakovee, O., Newman, A.J., Clark, M.P., Wood, A.W., Gupta, H.V., Kuman, R., 2019. On the choice of calibration metrics for "high-flow" estimation using hydrologic models. Hydrol. Earth Syst. Sci. 23 (6), 2601–2614.
Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.
Murphy, A.H., 1988. Skill scores based on the mean square error and their relationships to the correlation coefficient. Mon. Weather Rev. 116, 2417–2423.
Murphy, A.H., Daan, H., 1985. Forecast Evaluation. Probability, Statistics and Decision Making in the Atmospheric Sciences. Westview Press, pp. 379–437.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I – a discussion of principles. J. Hydrol. 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 Resour. Res. 42, W07410. https://doi.org/10.1029/2005wr004636.
Oreskes, N., Shrader-Frechette, K., Belitz, K., 1994. Verification, validation and confirmation of numerical models in the earth sciences. Sciences 263 (5147), 641–646.
Pechlivanidis, I.G., Jackson, B., McMillan, H., Gupta, H., 2014. Use of an entropy-based metric in multiobjective calibration to improve model performance. Water Resour. Res. 50, 8066–8083.
Pechlivanidis, I.G., Jackson, B., McIntyre, N., Wheater, H.S., 2011. Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. Global. Nest. J. 13 (3), 193–214.
Pechlivanidis, I.G., Jackson, B.M., Mcmillan, H.K., Gupta, H.V., 2012. Using an informational entropy-based metric as a diagnostic of flow duration to drive model parameter identification. Global. Nest. J. 14 (3), 325–334.
Pool, S., Vis, M., Seibert, J., 2018. Evaluating model performance: towards a non-parametric variant of the Kling-Gupta efficiency. Hydrol. Sci. J. 63 (13–14), 1941–1953.
Pushpalatha, R., Perrin, C., Moine, N.L., Andréassian, V., 2012. A review of efficiency criteria suitable for evaluating low-flow simulations. J. Hydrol. 420–421, 171–182.
Quesada-Montano, B., Westerberg, I.K., Fuentes-Andino, D., Hidalgo, H.G., Halldin, S., 2018. Can climate variability information constrain a hydrological model for an ungauged Costa Rican catchment? Hydrol. Process. 32, 830–846.
Quintero, F., Krajewski, W.F., Seo, B.C., Mantilla, R., 2020. Improvement and evaluation of the Iowa Flood Center Hillslope Link Model (HLM) by calibration-free approach. J. Hydrol. 584, 124686. https://doi.org/10.1016/j.jhydrol.2020.124686.
Ritter, A., Muñoz-Carpena, R., 2013. Performance evaluation of hydrological models: statistical significance for reducing subjectivity in goodness-of-fit assessments. J. Hydrol. 480 (3), 33–45.
Santos, L., Thirel, G., Perrin, C., 2018. Technical note: Pitfalls in using log-transformed flows within the KGE criterion. Hydrol. Earth Syst. Sci. 22, 4583–4591.
Seeger, S., Weiler, M., 2014. Reevaluation of transit time distributions, mean transit times and their relation to catchment topography. Hydrol. Earth Syst. Sci. 18, 4751–4771.
Thomas, H.A., 1981. Improved methods for national water assessment: Final report. U.S. Geol. Surv. Water Resour. Contract WR15249270, 44.PP.
Wang, D., Tang, Y., 2014. A oneparameter Budyko model for water balance captures emergent behavior in darwinian hydrologic models. Geophys. Res. Lett. 41, 4569–4577. https://doi.org/10.1002/2014GL060509.
Weglarczyk, S., 1998. The interdependence and applicability of some statistical quality measures for hydrological models. J. Hydrol. 206, 98–103.
Yan, R.H., Huang, J.C., Wang, Y., Gao, J.F., Qi, L.Y., 2016. Modeling the combined impact of future climate and land use changes on streamflow of Xinjiang Basin, China. Hydrol. Res. 47 (2), 356–372.
Zhang, H., Wang, B., Liu, D.L., Zhang, M.X., Feng, P.Y., Cheng, L., Yu, Q., Eamus, D., 2019. Impacts of future climate change on water resource availability of eastern Australia: a case study of the Manning River basin. J. Hydrol. 573, 49–59.
Zhao, R.J., 1992. The Xinanjiang model applied in China. J. Hydrol. 135, 371–381.
Zhao, R.J., Liu, X.R., Singh, V.P., 1995. The Xinanjiang model. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Highlands Ranch, CO, pp. 215–232.
Zhao, R. J., Zhang, Y. L., Fang, L. R., Liu, X. R., Zhang, Q. S., 1980. The Xinangjiang model. In: Proc. Oxford Symposium on Hydrological Forecasting. International Association of Hydrological Sciences, Wallingford, UK, pp. 351-356.
Zhuo, L., Dai, Q., Han, D., 2014. Meta-analysis of flow modeling performances – to build a matching system between catchment complexity and model types. Hydrol. Process. 29 (11), 2463–2477.