

Article

Bayesian Theory Based Self-Adapting Real-Time Correction Model for Flood Forecasting

Jun Wang ¹, Zhongmin Liang ^{1,2}, Xiaolei Jiang ¹, Binqun Li ^{1,2} and Li Chen ^{2,3,*}

¹ College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China; wangjun_hhu@yahoo.com (J.W.); zmliang@hhu.edu.cn (Z.L.); jxl_hhu@126.com (X.J.); libinqun@hhu.edu.cn (B.L.)

² State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China

³ Division of Hydrologic Sciences, Desert Research Institute, Las Vegas, NV 89119, USA

* Correspondence: lchen@hhu.edu.cn; Tel.: +86-25-83787079

Academic Editor: Ataur Rahman

Received: 11 November 2015; Accepted: 23 February 2016; Published: 26 February 2016

Abstract: Real-time correction models provide the possibility to reduce uncertainties in flood prediction. However, most traditional techniques cannot accurately capture many sources of uncertainty and provide a quantitative evaluation. To account for a wide variety of uncertainties in flood forecasts and overcome the limitations of stationary samples in a changing climate, a Bayesian theory based Self-adapting, Real-time Correction Model (BSRCM) was proposed. BSRCM uses the Autoregressive Moving Average (ARMA (n, m)) model as the prior distribution for the flood hydrograph, and the autoregressive model or order p (AR(p)) as the likelihood function to describe the likelihood relationship between the predicted and observed discharges, on the basis the posterior distribution of real values of discharge at any step can be deduced under the framework of Bayesian theory. Combined with the Xin'anjiang hydrological model, it was applied for flood forecasting in the Misai basin in southern China. Results from this study indicate that: (1) BSRCM can achieve a good precision and perform better than AR(p) in the study region; (2) BSRCM provides not only deterministic results but also rich uncertainty information for real-time correction results, such as the mean, error variance, and confidence intervals of flow discharge at any time during the flood event; (3) BSRCM can achieve better performance with a longer lead time; (4) BSRCM can achieve a good precision even with a small sample for parameter estimates. In addition to good precision, BSRCM can also provide further scientific grounding in flood control, operations and decision making for risk management.

Keywords: Bayesian theory; real-time correction; flood forecast; self-adapting; Xin'anjiang model; Misai basin

1. Introduction

Flood forecasts provide the technical support for reservoir operations and play an important role in flood control [1–4]. Thus, the accuracy of real-time flood forecasts is the technical basis for flood control decision-making. Recently, great progress has been made in flood forecast technology. However, hydrological models used for flood forecasting, including lumped models such as TOPMODEL [5] and distributed models such as TOPKAPI [6], are all approximations of the actual hydrologic processes and may not fully describe the real processes; hence, many types of uncertainties in flood forecasting inevitably exist, which produces biases in the forecasting results. Real-time correction techniques

improve the forecast accuracy by reducing uncertainties in the prediction [7]. Since the development of real-time correction theory, many real-time correction models, including autoregressive models [8,9], Kalman filtering models [10–13], fuzzy models [14,15], and neural network models [16], *etc.*, have been developed. Recently, combined approaches, such as the combination of filtering and error forecasting procedures, and the combination of forecasted errors with time series models and the Kalman filter method [17–19], were developed for real-time correction, and demonstrated their ability to provide improved results. Among these models, the auto-regressive and Kalman filtering models are found to produce better results in real-time correction than others [7]. However, these two models cannot quantitatively describe various possible uncertainties. With increasing demands on flood risk management, research on the uncertainties of flood forecasts has become a hotspot. To improve the accuracy of flood risks' prediction and provide uncertainty information for risk-guided decision-making, quantification of the possible uncertainties in real-time correction appear to be a crucial issue. However, currently available approaches for real-time correction of flood forecasts cannot adequately quantify the uncertainties.

The recent progress in probabilistic flood forecasting provides great opportunities for real-time correction of flood forecasts with treatment of uncertainties. Probabilistic flood forecasting has become important because of the need for quantitative estimation of uncertainties in forecasting results. Among recent research progress, the Bayesian forecasting system [20–26] is a noteworthy approach, which provides a comprehensive estimation of uncertainties using three main procedures, including the separation of hydrological uncertainty and precipitation input uncertainty, estimating these two types of uncertainties independently, and synthesizing the overall uncertainty according to the Bayesian theory. In view of the advantage of Bayesian inference in uncertainty estimation, the Bayesian flood forecast system provides a promising resolution for real-time correction.

Meanwhile, real-time correction methods take advantage of historical information to estimate parameters whenever possible. The traditional methods can adapt well to basins with ample historical data and long sample series. However, the runoff mechanisms of the natural watersheds could undergo changes because of increasing human activities [27]. This change may cause inconsistencies between historical sample series and current data. Thus, traditional real-time correction methods that depend on the historical samples may suffer from certain negative influences and produce more uncertainties.

A promising technique to address these issues is the real-time correction approach for flood forecasting that can capture and quantify the uncertainties and avoid the possible inconsistency in historical information. Inspired by previous hydrological studies applying the Bayes' theorem, a Bayesian theory based Self-adapting Real-time Correction Model (BSRCM) was proposed in the study. In the Bayesian theory framework, BSRCM adopts prior distributions to describe prior information and likelihood functions to describe conditional distributions of discharge series, which is a different approach from other pertinent studies. This study attempted to develop an autoregressive real-time correction model based on the Bayesian theory for flood forecasting. The model can provide quantitative evaluation of uncertainties and further improve the accuracy of real-time correction. The paper is structured as follows. First, Section 2 presents the Bayesian theory based Self-adapting Real-time Correction Model (BSRCM), the error autoregressive model and the hydrological model. Second, Section 3 describes a case study in the Misai basin, a humid catchment in southeastern China. In addition, the results and discussions of the application of BSRCM are provided in Sections 4 and 5 respectively. Finally, Section 6 summarizes the main conclusions drawn by the study.

2. Method

2.1. The BSRCM

Denote the starting time for the flood forecast as t , the observed streamflow before t as $X = \{x_1, x_2, \dots, x_t\}$, and the simulation of X as $\hat{X} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_t\}$. Meanwhile, denote the

streamflow after t as $S = \{s_1, s_2, \dots, s_K\}$ and the simulation of S as $\widehat{S} = \{\widehat{s}_1, \widehat{s}_2, \dots, \widehat{s}_K\}$, where K is the lead time. According to the Bayesian theory, the posterior distribution of forecasting variables can be obtained under the condition that the prior distribution and likelihood function were known:

$$\phi_k(s_k | \widehat{s}_k, X, Y) = \frac{f_k(\widehat{s}_k | s_k, Y) g_k(s_k | X)}{\int_{-\infty}^{\infty} f_k(\widehat{s}_k | s_k, Y) g_k(s_k | X) ds_k} \quad (1)$$

where the probability density distribution, denoted as f_k , is a likelihood function, g_k is the prior distribution, and ϕ_k is the posterior distribution.

In this study, the error Autoregressive model of order p (AR(p)) is used to describe the likelihood relationship between the observed discharge and forecasted discharge. The Autoregressive Moving Average model (ARMA(n, m)) is used to describe the prior information of discharge. According to the Bayesian theory, the posterior distribution of forecasting discharge at any time step can be deduced and the real-time correction analysis can be conducted.

2.1.1. Prior Distribution

ARMA (n, m) is one of the most popular methods of time series analysis which can be used to simulate time series in hydrologic forecast [28]. In this study, it is adopted to describe the prior information about flooding in the study area with the assumption that the gauged discharge curve is stationary during a specific period. The prediction or simulation variable s_k can be expressed as:

$$s_k = \mu + C_k \Delta X^T - O_k \Xi_k + \eta_k \quad (2)$$

where μ is the mean of X ; $\Delta X = \{(x_{t-n+1} - \mu), \dots, (x_t - \mu)\}$ is departure series of X ; $\Xi_k = \{\xi_{k,t-m+1}, \dots, \xi_{k,t}\}$ is the m -dimensional residual error series at time t with lead time X ; n is the degree of auto-regression while m is the degree of moving average; C_k is the coefficient of ΔX while O_k is the coefficient of Ξ_k ; supposing that the error series η_k are normally distributed— $N(0, \chi_k^2)$, and the prior distribution of s_k is also normally distributed. Therefore, the distribution of s_k can be expressed as:

$$s_k \sim N(\mu + C_k \Delta X^T - O_k \Xi_k^T, \chi_k^2) \quad (3)$$

2.1.2. Likelihood Function

Based on the aforementioned method of error forecasting, an AR(k) model can be developed:

$$\widehat{s}_k = s_k + \Theta_k Y^T + \varepsilon_k \quad (4)$$

where $Y = \{y_{t-l+1}, \dots, y_t\}$; $\Theta_k = \{\theta_{k,l}, \dots, \theta_{k,l}\}$ are model parameters, l is the degree of model; ε_k is the residual error series independent of s_k and Y . Assuming that ε_k is normally distributed while its mean is zero and the variance is δ_k^2 , the likelihood function of $\widehat{s}_k$ is normally distributed:

$$\widehat{s}_k \sim N(s_k + \Theta_k Y^T, \delta_k^2) \quad (5)$$

According to Equation (5), $\widehat{s}_k$ is normally distributed with mean $s_k + \Theta_k Y^T$ and variance δ_k^2 .

2.1.3. Posterior Distribution

After the prior distribution and likelihood function are determined, the posterior density function can be obtained according to the Bayesian formula:

$$\phi_k(s_k | \widehat{s}_k, X, Y) = \frac{1}{T_k} q\left(\frac{s_k - A_k \widehat{s}_k - D_k}{T_k}\right) \quad (6)$$

with:

$$A_k = \frac{\chi_k^2}{\chi_k^2 + \delta_k^2}, D_k = \frac{\delta_k^2(\mu + C_k \Delta X^T - O_k \Xi_k^T) - \chi_k^2 \Theta_k Y^T}{\chi_k^2 + \delta_k^2}, T_k^2 = \frac{\chi_k^2 \delta_k^2}{\chi_k^2 + \delta_k^2}$$

where A_k , D_k and T_k are corresponding intermediate variables.

Assuming q and Q are the density function and the distribution function of a standard normal distribution, respectively, the corresponding posterior probability function Φ_k is:

$$\Phi_k(s_k | \hat{s}_k, X, Y) = Q\left(\frac{s_k - A_k \hat{s}_k - D_k}{T_k}\right) \quad (7)$$

2.2. Error Autoregressive Model (AR Model)

An error autoregressive model (AR model) is selected in the study for comparison with the BSRCM. The basic principle of the AR (p) model is described as follows.

Using the self-correlation of the error series, the autoregressive model was developed by incorporating the existing error series. The model can consider the stochastic disturbance elements and further predict the errors. The model reads [29]:

$$H(j) = \varphi_1 H(j-1) + \varphi_2 H(j-2) + \cdots + \varphi_p H(j-p) + \alpha(j) = \sum_{i=1}^p \varphi_i H(j-i) + \alpha(j) \quad (8)$$

where $j = 1, 2, \dots, n$, $p > 0$. Equation (8) is an autoregressive model of p order, namely AR(p). The parameters φ_i ($i = 1, 2, \dots, p$) of this model can be obtained by the least square method. Generally, the series of disturbance variable, $\alpha(j)$, is independent white noise series. In addition, it is normally distributed with a zero mean and a non-zero variance.

2.3. The Xin'anjiang Hydrological Model

Real-time correction models for flood forecast should be implemented by coupling with a deterministic hydrological model. In this study, the Xin'anjiang model is used as the deterministic model, which has been widely used in China since its development in 1973 [30–33]. The model is based on the concept of saturation excess (Dunne) runoff formation mechanism, which means that runoff would not occur until soil moisture reaches the field saturation. The study basin is divided into a set of sub-basins. The outflow from each sub-basin is first simulated and then routed down the channels to the basin outlet. The simulation of outflow from each sub-basin is consisted of four major parts:

- (1) Evapotranspiration. It generates the deficit of soil storage which is divided into the upper, lower and deep layers;
- (2) Runoff production. It produces the runoff according to the rainfall and soil storage deficit;
- (3) Runoff separation. It divides the total runoff into three components: surface, subsurface and groundwater;
- (4) Flow routing. It transfers the local runoff to the outlet of each sub-basin to form the outflow of the sub-basin.

The flow chart of Xin'anjiang model is shown in Figure 1. All symbols inside the blocks are variables, including inputs, outputs, state variables and internal variables, while those outside the blocks are parameters.

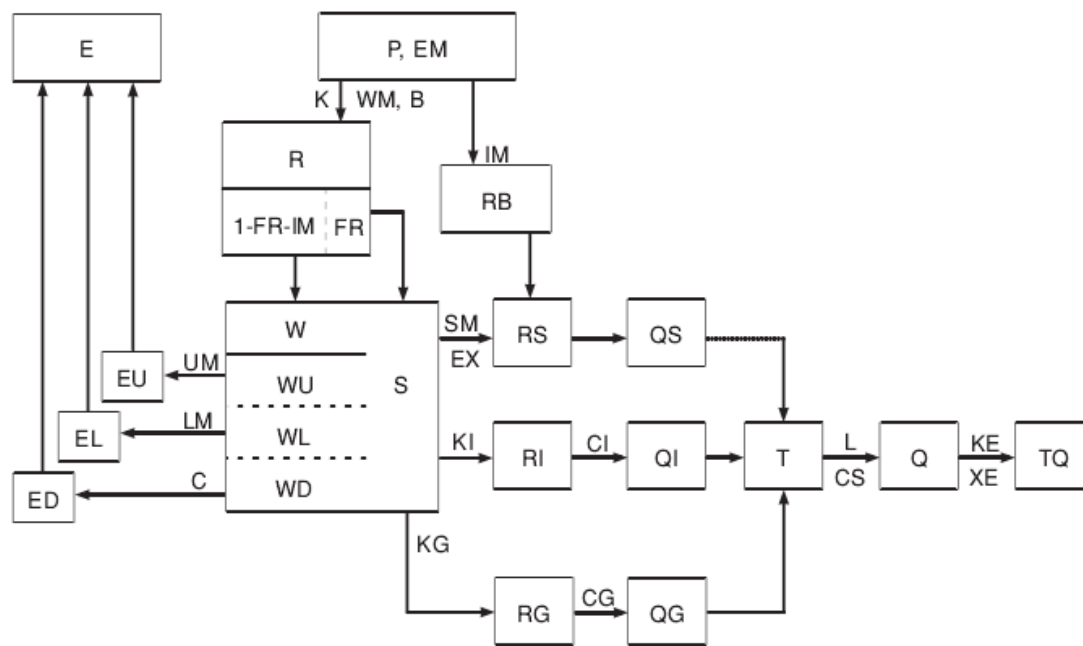


Figure 1. Flow chart of the Xin'anjiang model.

The inputs to the model are areal mean rainfall, P , and measured pan evaporation, EM . The outputs are the discharge, TQ , from the whole basin and the actual evapotranspiration, E , which includes three components, EU , EL and ED . The state variables are the areal mean tension water storage, W , and the areal mean free water storage, S . The areal mean tension water W has three components, WU , WL , and WD , in the upper, lower and deep layer, respectively.

The FR is the factor of runoff contributing area which is related to W . The rest of the symbols inside the blocks are all internal variables. RB is the direct runoff from impervious areas. R is the runoff produced from pervious areas and is divided into three components, *i.e.*, RS , RI , and RG denoting surface runoff, interflow and groundwater runoff, respectively. These three components are further transferred into QS , QI , and QG and combined to form the total inflow to the channel network of the sub-basin. The outflow of the sub-basin is Q .

The attributes outside the blocks in Figure 1 are all parameters. K is the ratio of potential evapotranspiration to pan evaporation if pan evaporation measurements are used as input. WM and B are two parameters describing the tension water distribution. WM is the areal mean tension water capacity which has three components, UM , LM and DM . B is the exponent of the tension water capacity distribution curve. IM is the factor of impervious area. SM and EX are similar to WM and B while they describe the free water capacity distribution. KI and KG are coefficients related to RI and RG . CI , CG , L , CS , KE and XE are parameters for flow routing.

3. Case Study

3.1. Study Basin

The Misai Basin (118.0° E and 29.10° N) located in southeastern China is selected for this study. The watershed drains an area of 797 km^2 . It has a humid climate and the mean annual precipitation is about 1800 mm . There are one stream gauge and six rainfall gauging stations in the basin (Figure 2).

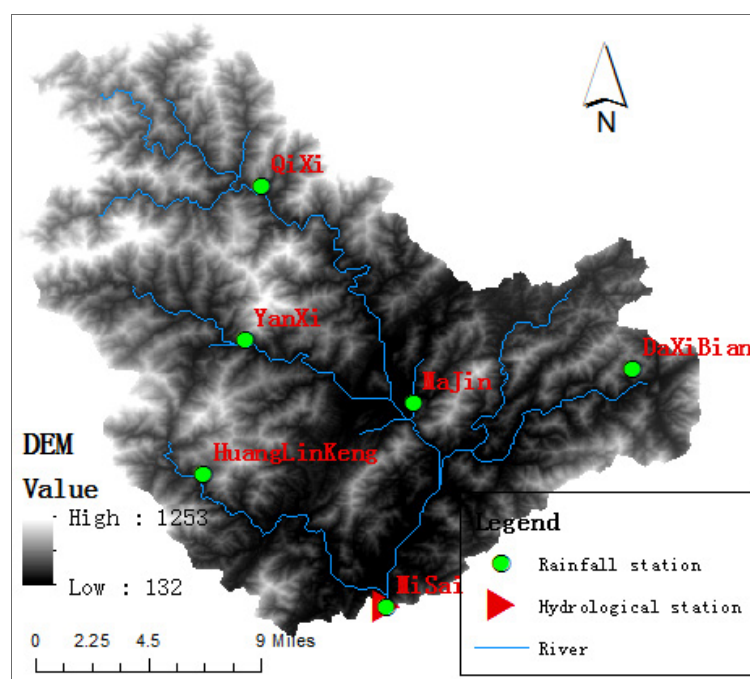


Figure 2. The distribution of gauging stations, river network and DEM of the Misai basin.

3.2. Model Implementation

The observed precipitation of the six rainfall gauging stations shown in Figure 2 and observed streamflow discharge at the Misai station were used in the simulation. The time intervals of the observed precipitation and discharge are shorter than 1 hour, while the network density of rainfall gauging stations is about 1 per 130 km². Nine flood events of the Misai basin were selected for examining the BSRCM. The first five events were used to calibrate the Xin'anjiang model while the remaining four were for validation.

The structure of BSRCM is described by Equations (1)–(7) and the main procedures of BSRCM are as follows: (a) at time step t , the residual error series of the observed and forecast flows by Xin'anjiang model for the previous $t-1$ steps are used to determine the AR(2) model, on the basis, the parameters of the likelihood function can be obtained; Similarly, the parameters of prior distribution can be obtained by Equation (3), while the parameters A_k , D_k and T_k in Equation (7) and the posterior distribution of real discharge at time step t can be deduced; (b) According to the posterior distribution of real discharge at time step t , the 50% quantile can be obtained as the corrected results of time step t ; (c) at time step $t+1$; Similarly, the parameters A_k , D_k and T_k and the posterior distribution of real discharge at time step $t+1$ will be updated by the "new information" of observed discharge at the time step t ; (d) According to the posterior distribution of real discharge at time step $t+1$, the 50% quantile can be obtained as the corrected results of time step $t+1$; (e) and by this analogy, the corrected results of any moment can be obtained.

4. Results

4.1. Lead Time of One Hour

4.1.1. Model Parameters

According to the flood forecasting results by the Xin'anjiang model, the real-time correction model activates at one third of the whole flood event duration. Initial model parameters of nine flood events can be estimated, which are shown in Table 1.

Table 1. Parameters estimation results at starting time.

Flood	Total Duration	Starting Time (h) of the Flood Duration (h)	C_k	O_k	X	θ_1	θ_2	χ	AK	DK	TK
1	192	64	0.97	−0.02	101.47	−0.02	0.97	72.19	0.66	66.40	58.82
2	99	33	1.02	−0.22	78.95	−0.22	1.02	79.17	0.50	745.00	55.90
3	141	47	0.96	−0.08	57.65	−0.08	0.96	33.13	0.75	−25.08	28.72
4	165	55	0.96	0.02	85.49	0.02	0.96	61.85	0.66	26.66	50.11
5	132	44	1.03	−0.42	35.86	−0.42	1.03	24.62	0.68	299.53	20.30
6	156	52	0.95	−0.04	32.71	−0.04	0.95	30.21	0.54	68.67	22.19
7	192	64	1.01	−0.08	16.79	−0.08	1.01	12.65	0.64	84.02	10.10
8	207	69	0.93	0.06	93.94	0.06	0.93	66.55	0.67	63.07	54.30
9	144	48	0.98	0.00	94.35	0.00	0.98	67.87	0.66	429.18	55.10

During the BSRCM model runs, the model parameters update in real-time. When the forecasting discharge is obtained at the next step, the corresponding error is automatically added into the sample used to estimate the parameters of BSRCM. Using the “new information” of the error sample, the BSRCM parameters can be updated at the new time step.

4.1.2. Comparisons

The statistical results of the two models (namely, AR(p) model and BSRCM), including flood peak errors, relative errors of flood volume, the coefficient of determination, and flood peak lag time, are listed in Table 2.

Table 2. Statistical results of AR(p) model and BSRCM (50% quantile is treated as the corrected result of BSRCM).

Flood	Relative Error of Flood Volume Error BSRCM/AR(2)		Relative Error of Flood Peak BSRCM/AR (2)		Coefficient of Determination BSRCM/AR (2)/Xin'anjiang Model			Flood Peak Lag Time BSRCM/AR (2)/Xin'anjiang Model		
1	0.05	0.02	0.16	0.09	0.99	0.96	0.91	1	1	1
2	0.04	0.18	−0.01	−0.01	0.98	0.98	0.89	0	0	−1
3	0.09	0.21	0.02	0.02	0.97	0.85	0.74	−1	−1	−1
4	0.02	0.14	0.01	−0.14	0.99	0.89	0.88	0	0	0
5	0.10	−0.02	0.10	−0.04	0.99	0.98	0.94	1	1	1
6	0.06	0.15	0.01	0.06	0.99	0.95	0.6	0	0	−1
7	0.08	0.16	0.06	0.01	0.99	0.95	0.75	0	0	0
8	−0.02	0.07	−0.02	−0.10	0.99	0.98	0.9	0	1	1
9	0.04	−0.12	0.01	−0.14	0.98	0.9	0.91	0	0	1

Note: In this study, the results of BSRCM is the 50% quantile of posterior distribution of real values of flow discharge at the corresponding moment.

Because BSRCM can generate posterior distribution of real values of flow discharge at every time step, the uncertainty of the results can be quantified. In this study, the 50% quantile is regarded as the value of real-time correction at the corresponding moment and is used to make a comparison with the results of the AR(2) and Xin'anjiang model. According to Table 2, it is noted that the two real-time correction models both improve the precision of flood forecasting while the BSRCM performs better than the AR(2). The coefficient of determination of the nine floods is over 0.95 by the BSRCM while, for four floods predicted by AR(2), it is over 0.95. The relative errors of flood volume and peak discharge of nine floods simulated by the BSRCM are less than those by the AR(2) model. In addition, the results of flood peak lag time simulated by BSRCM are less than those by the AR(2) model.

(1) Flood hydrograph comparison

Flood events are named according to their starting dates with the flood ID in the format YYYYMMDD. The first four digits denote the year, the middle two the month, and the last two the day, respectively. Taking the flood 19820619 as an example, the observed and simulated results by the three models are compared in Figure 3. It indicates that BSRCM predictions could match the observation better than AR(2) predictions.

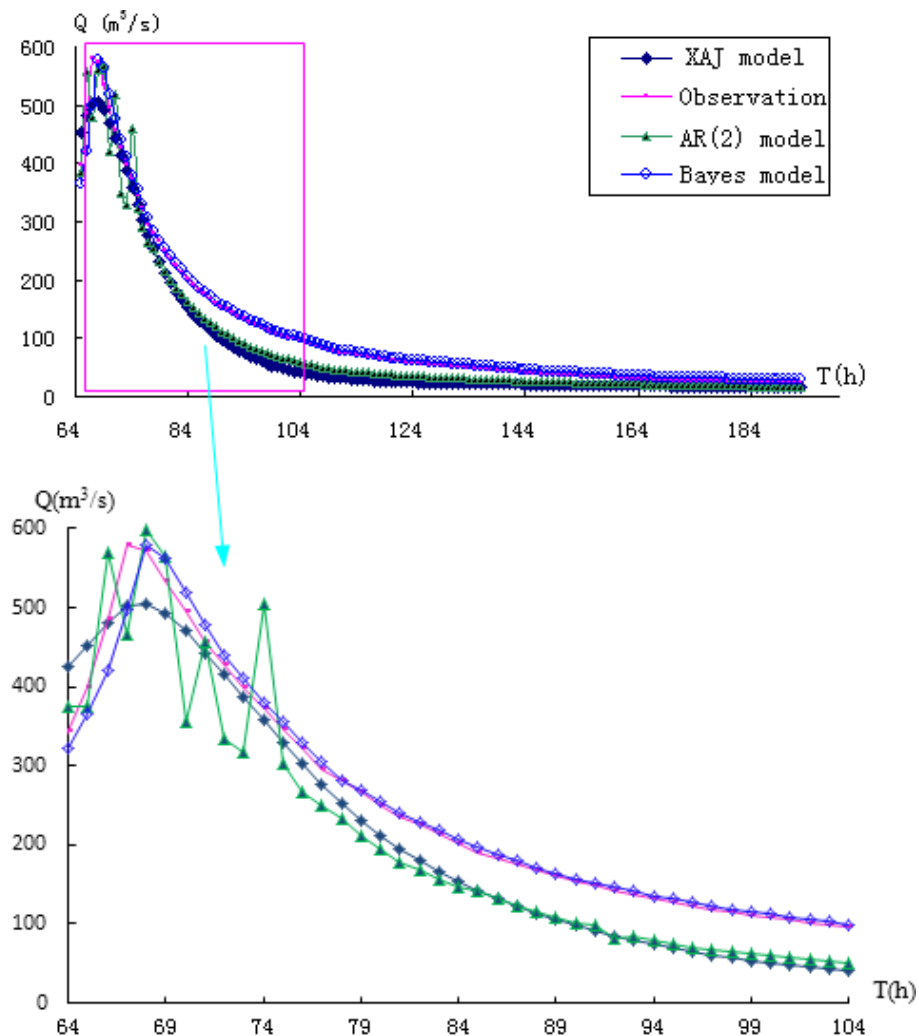


Figure 3. Flood hydrograph comparison of flood 19820619.

(2) Quantile and confidence interval results by BSRCM

The BSRCM provides not only the correction mean results, but also the uncertainty information about corrected results, such as the corresponding quantile of a specific probability, confidence interval of different confidence levels, posterior density and distribution of discharge at any time. Figure 4 shows the 90% confidence interval of the flood 19860707. The 5%Q denotes forecasted hydrograph of 5% quantile, while 95%Q denotes the forecasted starting hydrograph of 95% quantile, and thus the zone between the two lines is the 90% confidence interval.

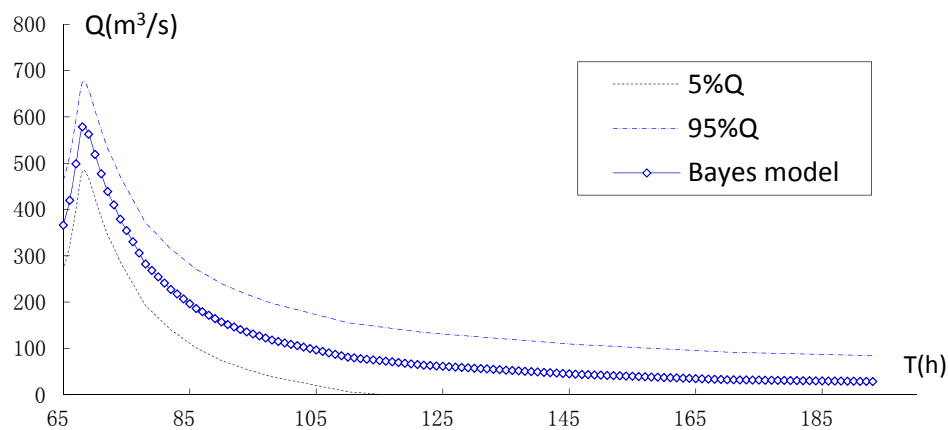


Figure 4. The correction results and 90% confidence interval of flood 19820619.

(3) Posterior density and posterior distribution of flood peak correction results

As mentioned above, posterior density and posterior distribution of flood peak correction result can be obtained. According to the posterior distribution, the uncertainty information can be analyzed straightforwardly. Taking flood 19820619 as an example, Figure 5 shows the Posterior distribution and density of the peak of the flood 19820619. According to the posterior distribution and density of this flood peak, the uncertainty information, such as the mean of peak discharge, the corresponding quantile of any probability, confidence interval of different confidence levels, error variance, and variable coefficient can be obtained.

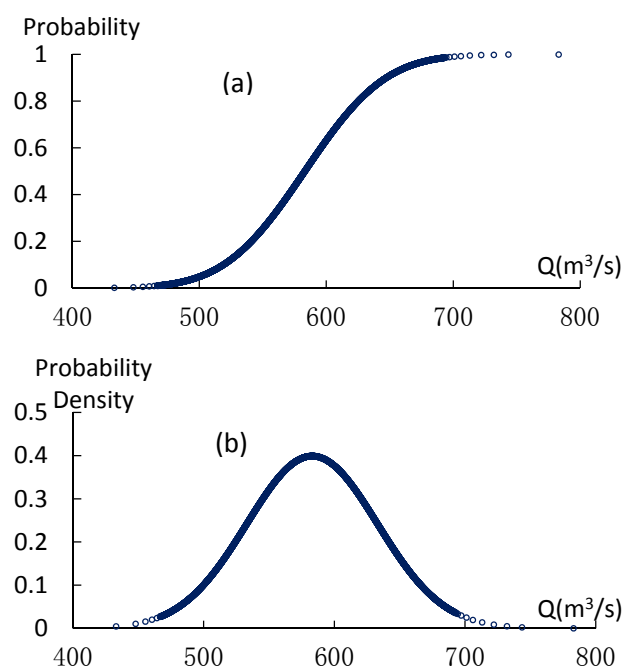


Figure 5. (a) Posterior distribution of peak discharge of flood 19820619; (b) Density of peak discharge of flood 19820619.

4.2. Lead Time $n > 1$ h

4.2.1. Statistical Analysis of Correction Results with Different Lead Times

Table 3 lists the statistical correction results of all flood events with five different lead times. Results show that the deterministic coefficients of real-time correction show a decreasing trend along

with the increase of lead time. All deterministic coefficients are greater than 0.90 except for the flood event 19830529 with a lead time $n = 5$ hour. It indicates that the BSRCM can produce good real-time correction results for flood forecasting even with long lead times.

Table 3. Deterministic coefficients of nine floods with the condition of different lead times.

Flood ID	Lead Time (h)				
	1	2	3	4	5
19820619	0.99	0.98	0.97	0.96	0.93
19830529	0.98	0.95	0.93	0.91	0.85
19830614	0.97	0.98	0.98	0.97	0.94
19830620	0.99	0.99	0.98	0.94	0.96
19840402	0.99	0.98	0.98	0.99	0.99
19850703	0.99	0.98	0.97	0.95	0.91
19860704	0.99	0.98	0.99	0.96	0.98
19870620	0.99	0.98	0.96	0.95	0.92
19880621	0.98	0.95	0.93	0.90	0.93

4.2.2. Flood Hydrographs of Different Lead Times

Figure 6 shows the flood hydrograph of real-time correction for the flood event 19820619. Results show that the real-time correction results of different lead times have the larger biases during the initial running period of the BSRCM. The biases become smaller with the model forecast, and further approach the “observed value”. It suggests that the BSRCM can achieve a good performance using “new information” automatically.

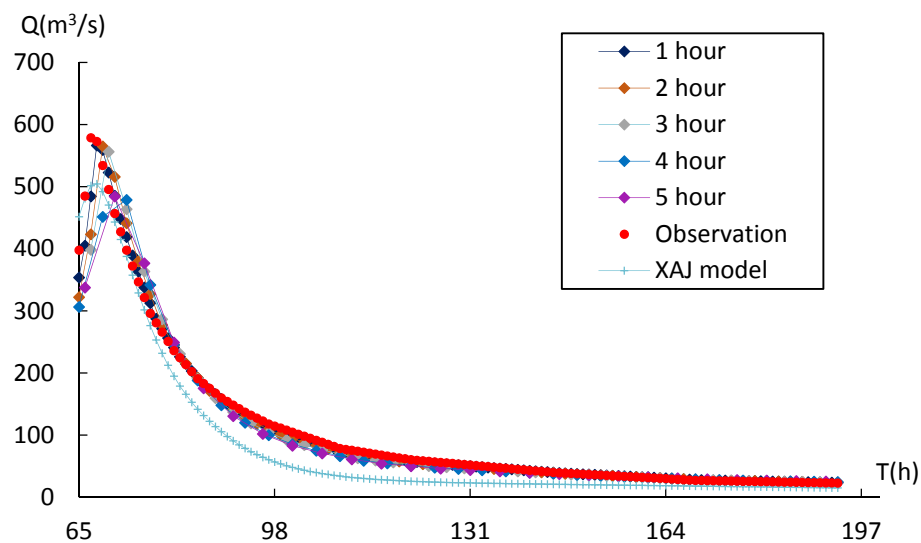


Figure 6. Real-time correction results of flood 19820619 by BSRCM. The legend “1 h” denotes the correction flood hydrograph when lead time $n = 1$ h, and so forth.

Figure 7 is the flood hydrograph of 90% confidence interval of real-time correction results with the lead time $n > 1$ h. The range of confidence intervals gets wider the extension of lead time. It indicates that the uncertainty of real-time correction gets larger with a greater lead time.

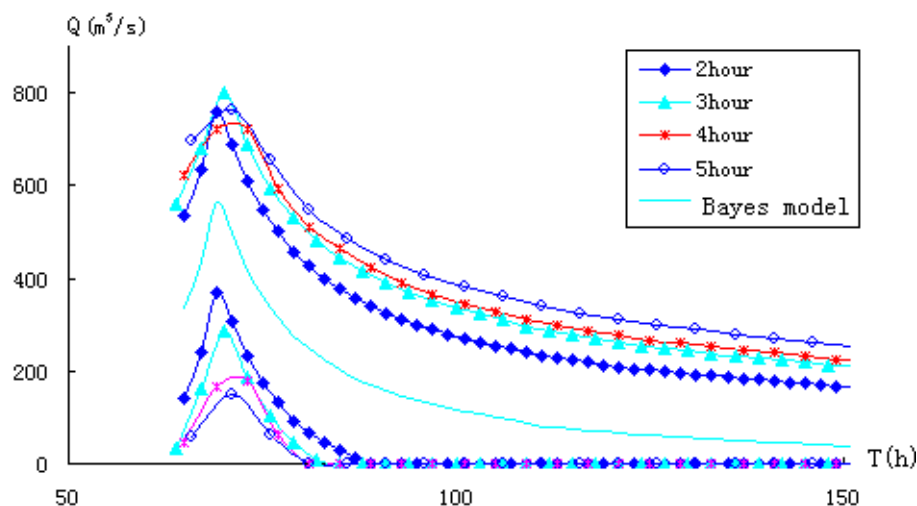


Figure 7. Ninety percent confidence interval of correction results of the flood event 19820619 by BSRCM.

The legend “2 h” denotes the correction flood hydrograph when lead time $n = 2$ h, and so forth.

5. Discussion

5.1. Parameters and Error Varying with Time Step

According to the flood hydrograph of real-time correction and observation, the error of real-time correction is relatively large at the beginning of a model run, but the error tends to decrease with time. The reason for this is examined by revisiting the proposed BSRCM. At the beginning of the model run, the size of sample used to estimate model parameters is too small, which could lead to uncertainty in the parameters' estimation. This uncertainty could be reduced as the modeling continues because the size of the historical sample becomes larger, *i.e.*, the prior information continuously increases.

It is found that A_K , D_K and T_K are the key parameters which determine the posterior distribution of real-time correction; hence, determining the final results of real-time correction. The variation process of A_K , D_K and T_K during the real-time correction of a flood forecast are demonstrated in Figures 8–10. It can be found that the posterior distributions of these three parameters have a tendency to be more stationary over time with the increasing historical information.

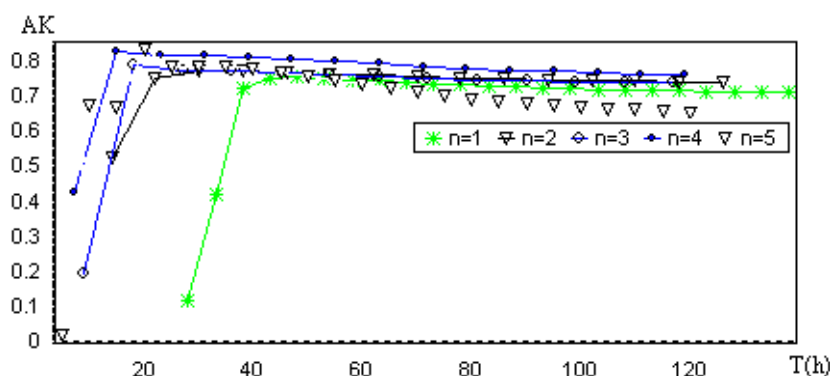


Figure 8. Variation of A_K for different forecast lead times during real-time correction of a flood forecast.

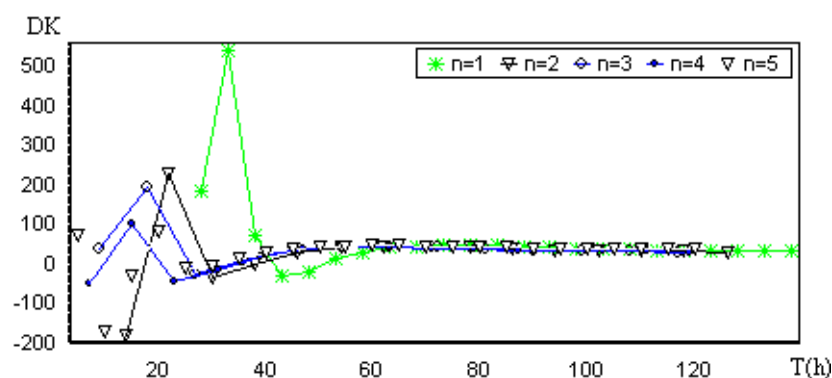


Figure 9. Variation of DK for different forecast lead times during real-time correction of a flood forecast.

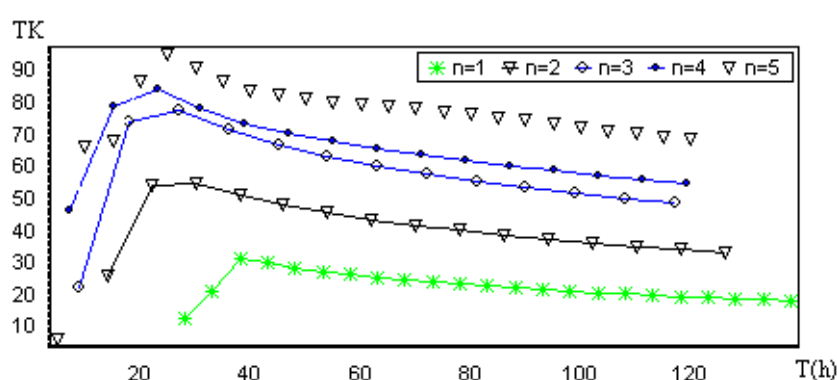


Figure 10. Variation of TK for different forecast lead times during real-time correction of a flood forecast.

The variation of the error and the parameters (A_K , D_K and T_K) as a function of the lead time from one to five hours is shown in Figures 11–15. It indicates that the error of real-time correction reduces with simulation time, and the corrected results approach the observed discharge values closely. Taking Figure 11 as an example, in the early stages of the real-time correction, the real-time correction error is large and reaches almost -300 . By $T = 20$ h, the real-time correction error rapidly reduces and approaches zero. When $T > 20$ h, the real-time correction error essentially equals zero and is stable over time. The other three cases also indicate the similar behaviors of the error.

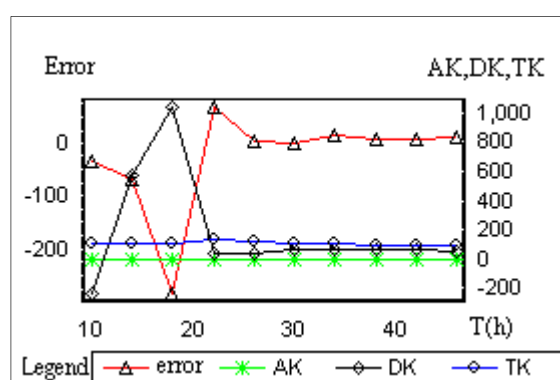
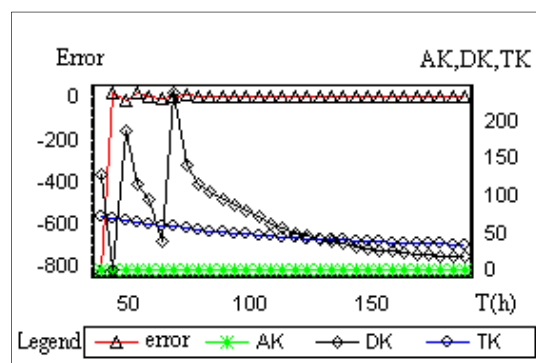
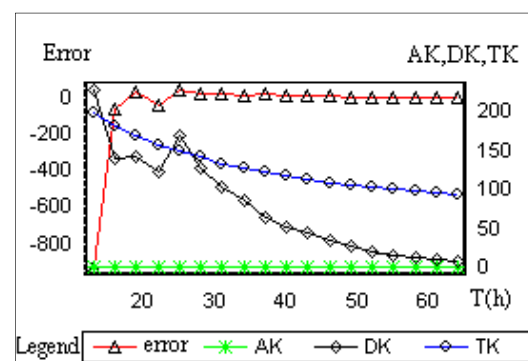
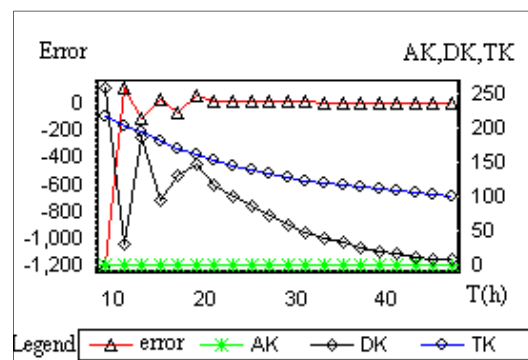
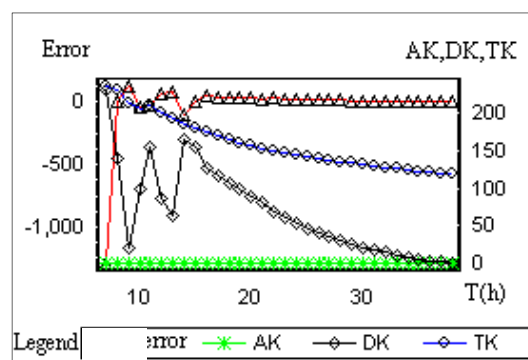


Figure 11. Forecast lead time $n = 1$.

Figure 12. Forecast lead time $n = 2$.Figure 13. Forecast lead time $n = 3$.Figure 14. Forecast lead time $n = 4$.Figure 15. Forecast lead time $n = 5$.

5.2. Selection of Model Start-Up Period

Clearly, the selection of the model start-up time will affect the correction results of the BSRCM. If the starting time is too early, the uncertainties of parameters at the initial period will be large due to the small size of the sample. The uncertainties will further affect the precision of the BSRCM. In addition, if the start running time of BSRCM is too late, the most important flood peak which is significant for flood control and operation may be missed. Hence, how to select the starting time of the BSRCM is crucial. To address this question, a comparison was made among different starting times, including one third, one fourth, one fifth, one sixth of the duration of the entire flood event to the real-time correction of a flood event. Taking the flood event 19820619 as an example, the statistical results for five lead times ($n = 1, 2 \dots \dots, 5$ h), are listed in the Table 4.

Table 4. Statistical correction results of different starting time.

Forecast Lead Time (h)	Starting Time	Relative Error of Flood Volume (%)	Relative Error of Peak Discharge (%)	Error of Flood Peak Time (h)	Deterministic Coefficient
1	One third of flood duration	−0.01	−0.02	1	0.99
2		0.01	−0.08	1	0.98
3		0.01	−0.08	1	0.97
4		0.03	−0.16	1	0.96
5		0.04	−0.22	1	0.93
1	One fourth of flood duration	−0.01	−0.02	1	0.99
2		−0.01	−0.08	1	0.98
3		0.01	−0.08	1	0.96
4		0.02	−0.16	1	0.95
5		0.05	−0.22	1	0.92
1	One fifth of flood duration	0.01	0.01	1	0.99
2		−0.01	0.01	1	0.97
3		−0.02	−0.08	1	0.95
4		0.01	−0.16	1	0.95
5		0.01	−0.19	1	0.93
1	One sixth of flood duration	−0.01	0.01	1	0.98
2		−0.02	0.01	1	0.93
3		0.01	−0.08	1	0.95
4		0.05	−0.16	1	0.93
5		0.01	−0.19	1	0.88

Results of Table 4 indicate that, with all four starting times, the BSRCM can achieve good precision. In contrast to statistical correction results of different starting times, it seems that the selection of model starting times plays an insignificant role in the accuracy of the real-time correction. The influence of different starting times tends to decrease over time, the BSRCM will quickly go into a stationary period to obtain a good effect of real-time correction.

6. Conclusions

In this study, a new approach for flood forecasting, BSRCM, was proposed and applied to a humid basin in southeastern China coupled with the Xin'anjiang hydrological model. Based on the modeling study, the following conclusions can be drawn:

- (1) The BSRCM can increase the precision of flood forecasting and perform well in real-time correction. Results of simulations for nine floods in the Misai basin indicate that, according to the analysis of the determination coefficient, and relative errors of flood volume and flood peak, the BSRCM performed well in the real-time correction and better than the AR(p) model.

- (2) The BSRCM provides not only deterministic results but also rich uncertainty information for the forecasting results. The new model can generate the posterior distribution of discharge at any time during the entire flood period. Therefore, the uncertainty information of forecasting results, such as the mean, error variance, variable coefficient, quantile of a specific probability and confidence interval of different confidence levels, can be determined. The uncertainty information could provide more technical support to the local flood control agencies.
- (3) The BSRCM can achieve good performance with longer lead times. Through the comparisons of the lead time from one to five hours, it was found that the posterior distributions of the three parameters, AK, DK and TK, tend to be more stationary over time with increasing historical information, and the BSRCM with different lead times can predict good results.

As a result, the real-time correction with the BSRCM in this study not only produces high accuracy but also provides rich uncertainty information of real-time correction results. Because the risk evaluations for flood control and operations are based on an uncertainty analysis of forecasts, the BSRCM can provide further scientific grounding for flood control and risk-based decision making. Furthermore, since the BSRCM only depends on information for forecasting floods before the starting point but does not depend on that of other historical floods in parameter estimation, it is suitable for real-time correction of flood forecasting under the background of climate change and human activity.

Acknowledgments: This study was supported by the Public Welfare Industry Special Fund Project of Ministry of Water Resources of China (No. 201301066 & 201401034), the National Natural Science Foundation of China (No. 51109054, 51179046), National Science and Technology Support Plan (No. 2013BAB06B01) and excellent PhD training program (No. 2015B05514).

Author Contributions: All authors contributed extensively to the work. Jun Wang, Zhongmin Liang and Xiaolei Jiang performed the calculation, result analysis, and wrote the article. Binqun Li and Li Chen contributed to the article in discussing the methodology and writing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhuang, Y.L. *Flood Forecasting*; Water Resources and Electric Power Press: Wuhan, China, 1995.
2. Rui, X. New progress of flood forecasting theory and applicability of the existing methods. *Water Conserv. Hydropower Sci. Technol. Prog.* **2001**, *21*, 1–4.
3. Shi, Y.; Davis, K.J.; Zhang, F.; Duffy, C.J.; Yu, X. Parameter estimation of a physically based land surface hydrologic model using the ensemble Kalman filter: A synthetic experiment. *Water Resour. Res.* **2014**, *50*, 706–724. [[CrossRef](#)]
4. Li, B.; Yu, Z.; Liang, Z.; Acharya, K. Hydrologic response of a high altitude glacierized basin in the central Tibetan Plateau. *Glob. Planet. Chang.* **2014**, *118*, 69–84. [[CrossRef](#)]
5. Beven, K.J.; Kirkby, M.J. A physically based variable contributing model of basin hydrology. *Hydrol. Sci. Bull.* **1979**, *24*, 43–69. [[CrossRef](#)]
6. Ciarapica, L.; Todini, E. TOPKAPI: A model for the representation of the rainfall-runoff process at different scales. *Hydrol. Processes* **2002**, *16*, 207–229. [[CrossRef](#)]
7. Guo, S. *Reservoir Operation Integrated System*; Press of Wuhan University of Hydraulic and Electric Engineering: Wuhan, China, 2000.
8. Arheimer, B.; Lindstrom, G.; Olsson, J. A systematic review of sensitivities in the Swedish flood-forecasting system. *Atmos. Res.* **2011**, *100*, 275–284. [[CrossRef](#)]
9. Valipour, M.; Banihabib, M.E.; Behbahani, S.M.R. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *J. Hydrol.* **2013**, *476*, 433–441. [[CrossRef](#)]
10. Kalman, R. A new approach to linear filtering and prediction theory. *Trans. Am. Soc. Mech. Engrs J. Basic Engng Ser.* **1960**, *D82*, 35–45.
11. Kitanidis, P.; Bras, R. Real-time forecasting with a conceptual hydrologic model Analysis of uncertainty. *Water Resour. Res.* **1980**, *16*, 1025–1033. [[CrossRef](#)]

12. Georgakao, K.; Smith, G. On improved hydrologic forecasting results from a WMO real-time forecasting experiment. *J. Hydrol.* **1990**, *114*, 17245.
13. Wu, X.; Wang, C.; Chen, X.; Xiang, X. Kalman filtering correction in real-time forecasting with hydrodynamic model. *J. Hydrodyn.* **2008**, *20*, 391–397. [[CrossRef](#)]
14. Takagi, T.; Sugeno, M. Fuzzy identification of system and its application to modeling and control. *IEEE Trans. Syst. Man Cybern.* **1985**, *15*, 116–132. [[CrossRef](#)]
15. Nguyen, P.K.T.; Chua, L.H.C. The data-driven approach as an operational real-time flood forecasting model. *Hydrol. Process.* **2012**, *26*, 2878–2893. [[CrossRef](#)]
16. Shamseldin, A.; O'Connor, K. A non-linear neural network technique for updating river flow forecasts. *Hydrol. Earth Syst. Sci.* **2001**, *5*, 577–597. [[CrossRef](#)]
17. Madsen, H.; Skotner, C. Adaptive state updating in real-time river flow forecasting—a combined filtering and error forecasting procedure. *J. Hydrol.* **2005**, *308*, 302–312. [[CrossRef](#)]
18. Wu, S.J.; Lien, H.C.; Chang, C.H.; Shen, J.C. Real-time correction of water stage forecast during rainstorm events using combination of forecast errors. *Stoch. Environ. Res. Risk Assess.* **2012**, *26*, 519–531. [[CrossRef](#)]
19. Shen, J.C.; Chang, C.H.; Wu, S.J.; Hsu, C.T.; Lien, H.C. Real-time correction of water stage forecast using combination of forecasted errors by time series models and Kalman filter method. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 1903–1920. [[CrossRef](#)]
20. Coccia, G.; Todini, E. Recent developments in predictive uncertainty assessment based on the model conditional processor approach. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3253–3274. [[CrossRef](#)]
21. Todini, E. A model conditional processor to assess predictive uncertainty in flood forecasting. *Int. J. River Basin Manag.* **2008**, *6*, 123–137. [[CrossRef](#)]
22. Krzysztofowicz, R. Bayesian Forecasting via Deterministic Model. *Risk Anal.* **1999**, *19*, 739–749. [[CrossRef](#)]
23. Krzysztofowicz, R.; Karen, K.S. Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resour. Res.* **2000**, *36*, 3265–3277. [[CrossRef](#)]
24. Krzysztofowicz, R. Bayesian system for probabilistic river stage forecasting. *J. Hydrol.* **2004**, *268*, 16–40. [[CrossRef](#)]
25. Krzysztofowicz, R.; Maranzano, C.J. Bayesian system for probabilistic stage transition forecasting. *J. Hydrol.* **2004**, *269*, 15–44. [[CrossRef](#)]
26. Herr, H.D.; Krzysztofowicz, R. Bayesian ensemble forecast of river stages and ensemble size requirements. *J. Hydrol.* **2010**, *387*, 151–164. [[CrossRef](#)]
27. Li, B.; Yu, Z.; Liang, Z.; Song, K.; Li, H.; Wang, Y.; Zhang, W.; Acharya, K. Effects of climate variations and human activities on runoff in the Zoige alpine wetland in the eastern edge of the Tibetan Plateau. *J. Hydrol. Eng.* **2014**, *19*, 1026–1035. [[CrossRef](#)]
28. Zhang, H. Study on probability hydrological forecast system and application based on Bayes theory. Ph.D. Dissertation, Wuhan University, Wuhan, China, 2005.
29. Wang, W.; Ding, J.; Jin, J. *Statistical Hydrology*; HOHAI University Press: Wuhan, China, 2008.
30. Zhao, R.; Zhuang, Y.; Fang, L.; Liu, X.; Zhang, Q. The Xin'anjiang Model. In *Hydrological Forecasting Proceedings Oxford Symposium*; IAHS: Oxford, England, 1980; Volume 129, pp. 351–356.
31. Singh, V. *Computer Models of Watershed Hydrology*; Water Resources Publications: Littleton, CO, USA, 1995.
32. Zhao, R.J. The Xin'anjiang model applied in China. *J. Hydrol.* **1992**, *135*, 371–381.
33. Alazzy, A.A.; Lü, H.; Zhu, Y.H. Assessing the Uncertainty of the Xin'anjiang rainfall-runoff model: Effect of the likelihood function choice on the GLUE method. *J. Hydrol. Eng.* **2015**, *20*, 04015016. [[CrossRef](#)]

