KSCE Journal of Civil Engineering (2015) 19(3):796-803
DOI 10.1007/s12205-013-0483-x

Water Engineering

pISSN 1226-7988, eISSN 1976-3808
www.springer.com/12205

An Error Updating System for Real-time Flood Forecasting based on Robust Procedure

Li Qian*, Bao Weimin**, and Qian jinglin***

Received August 8, 2013/Revised February 5, 2014/Accepted April 4, 2014/Published Online October 3, 2014

Abstract

An updating technique is a tool to update the forecasts of mathematical flood forecasting model based on observed real-time data, and is an important element in a flood forecasting. The quality of the final updating results depends primarily on the quality of observed data. However, the input flow to reservoirs, or so called "observed input flow", is usually calculated by the observed water stage, outflow, and stage-storage relation curve. In this process, coarse errors which have great negative influence on updating are often created. This paper introduces a robust theory in real-time forecasting updating, and proposes an error updating system with robust procedure to limit the influence of coarse errors. Cases of ten reservoirs are studied to test the proposed system. Results indicate that the proposed robust procedure deflates the influence of outliers. The results of real-time updating with robust procedure in flood forecasting are better than those without robust procedure.

Keywords: *coarse error, observed input reservoir flow, robust estimation, real-time updating*

1. Introduction

A real-time forecasting model normally yields prediction errors due to input errors, imperfections in modeling, and other uncertainties in the hydrological model. Hence, updating schemes must be employed to improve the forecasts of a forecasting model. The updating output variables, also known as error prediction is probably the most popular among hydrologists and has been successfully applied in flood forecasting. The method establishes an error prediction model based on the previous error information derived from the original forecasting model that is based on the error between observed data and simulated data (Borersen and Weerts, 2005; Bao *et al*., 2011).

With normal observed data, traditional error updating methods (Toth *et al*., 1999; Goswani *et al*., 2005) have better ability for error revising and dynamic tracking. However, if the observed data are abnormal and some outlier exists£¨the results are not satisfactory (Zhao *et al*., 2008). For a reservoir, the input flow of reservoir, so called 'observed input flow', is usually calculated by the observed water stage, out flow and stage-storage relation curve. In this process, coarse errors, that are outliers in data set, which are not of normally distributed characteristics, are often created. It is noted that the outliers have an unknown distribution with a much bigger variance, and appear to be inconsistent with the remainder of data set (Barnet and Lewis, 1994; Han and Kamber, 2001) but are relatively large in magnitude. Estimation algorithms such as normal estimation method (least square estimation) used in error updating are greatly influenced by errors. Therefore prevention of outliers and extreme errors in error updating efficiency to improve the stability of updating results is a key issue in flood forecasting research This can be achieved by using a robust procedure. As in every other branch of applied mathematics, such rationalizations or simplifications are vital, and one justifies their use by appealing to a vague continuity or stability principle: a minor error in the mathematical model should cause only a small error in the final conclusion. Unfortunately, this does not always hold. During the past decades one has become increasingly aware that some of the most common statistical procedures are excessively sensitive to seemingly minor deviations from the assumption, and a plethora of alternative "robust" procedures have been proposed. It should be emphasized once more that the occurrence of coarse errors in a small fraction of the observations is to be regarded as a small deviation, and that, in view of the extreme sensitivity of some classical procedures, a primary goal of robust procedures is to safeguard against coarse errors. In the robust estimation theory,

*Associate Professor, Zhengjiang State Key Laboratory of Disaster Prevention and Reduction, Zhejiang Institute of Hydraulics and Estuary, Zhejiang University, Zhejiang 310000, China (Corresponding Author, E-mail: liqianhhu@163.com)
**Professor, State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China (E-mail: wmbao@hhu.edu.cn)
***Professor, Zhengjiang State Key Laboratory of Disaster Prevention and Reduction, Zhejiang Institute of Hydraulics and Estuary, Zhejiang 310000, China (E-mail: qjlhz@163.com)

the concept of contaminated distribution of error is put forward, and the equivalent weighted objective function is constructed to resist the effects of coarse error. Based on these, the estimation method shows stronger anti-error ability. Bao *et al.* (2003) firstly introduce the theory of robust estimation into flood forecasting system and analyzed its application in hydrology (Bao, 2003, 2005). Then some researches apply the robust theory to resist coarse errors existing in hydrology such as: Zhao *et al.* (2009) propose a three-step robust method combining robust statistical theory with distribution features of precipitation for the detection of outliers in a telemetry system. Zhao *et al.* (2008) propose a new robust recursive method of estimating auto-regressive updating model parameters for real-time flood forecasting. Bao (2004) apply the robust estimation to estimate the Muskingum parameters. The results of these reports illustrate that the robust estimation have stronger anti-error ability and suitable for hydrological need.

This paper presents an error updating method combined with the robust estimation theory for minimizing coarse errors that exist in the observed input flow. In our real-time flood forecasting updating system, the error prediction model selected is Auto Regression (AR) error-forecast updating model with the Recursive Least-Squares (RLS) estimation; the observed input flow should be treated by robust process before putting into the AR model. The simulated flow (another input of AR model) is calculated by a widely used rainfall-runoff model---Xinanjiang (XAJ) model. The AR error-updating model provides estimates of the simulation forecast errors that must be added to the corresponding simulated discharges to provide the updated forecasts; the XAJ model provides such forecasts directly.

2. Problem Formulations

2.1 Error Updating in Real-time Forecasting

In proposed real-time flood forecasting updating system, the errors between the observed and calculated values are used in the AR model to forecast the next error for improving the calculated values. The least-square method is adopted to estimate the AR model parameters.

An Auto-Regression (AR) error updating model for forecasting is established.

$$e(t) = \psi_t^T\theta + \xi(t) \tag{1}$$

where, $e(t)$ the error between observed discharge and discharge predicted by the hydrological model; $\psi_t^T = [e(t-1), e(t-2), \ldots e(t-N)]$ θ is the vector of the parameters; $\xi(t)$ is residual.

The estimation of parameter is:

$$\hat{\theta}_t = (\Phi_t^T\Phi_t)^{-1}\Phi_t^T Y_t \tag{2}$$

where, $\Phi_t^T = [\psi_1, \psi_2, \ldots \psi_t]$ $Y_t^T = [e(1), e(2), \ldots e(t)]$

The new data is input into auto-regression updating model at every time step, so that the estimated parameters are time-varying. The auto-regression updating model for real time flood forecasting demands tracking of the current properties of the observed discharge by means of emphasizing the new data that contains more current information. In addition, many literaturereports (Karhunen, 1998; Haykin, 1996; Eleftherious *et al.*, 1986; Poor *et al.*, 1997) the RLS algorithm is better than the Least-Square Algorithm (LSM) in convergence rate and tracking capability. Therefore, the recursive least square algorithm is considered to estimate θ.

This recursive algorithm could be expressed as:

$$\text{Set } P(t) = (\Phi_t^T\Phi_t)^{-1} \tag{3}$$

$$\hat{\theta}_t = P(t)\Phi_t^T Y_t \tag{4}$$

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \frac{P(t)\psi_{t+1}}{(1+\psi_{t+1}^T P(t)\psi_{t+1})}(e(t+1) - \psi_{t+1}^T\hat{\theta}(t)) \tag{5}$$

$$P(t+1) = P(t) - \frac{P(t)\psi_{t+1}\psi_{t+1}^T P(t)}{(1+\psi_{t+1}^T P(t)\psi_{t+1})} \tag{6}$$

For discharge correction, $e(t)$ in Eq. (1) is $e(t) = Q(t) - QC(t)$, where, $Q(t)$ and $QC(t)$ stands for observed and simulated discharge, respectively. The simulated discharge (another input of AR model) is calculated by a widely used XAJ rainfall-runoff model.

2.2 Hydrological Model

The flood forecasting model used in this paper is a semi distributed XAJ rainfall-runoff model developed in 1973 and published in 1980 (Zhao *et al.*, 1980). The model has been applied successfully over a very large area including all of the agricultural, pastorals and forested lands of China except the loess. In China, the XAJ model is used mainly for hydrological forecasting. Many large projects such as Gezhouba, Panjakou, Danjiangkou, Lubuge, Longyangxia and Ertan have used this model. Many forecasting systems such as Sanmenxia to Huayuenkou on the Yellow River, the middle reach of the Hui River and the Yangtze Gorges have used the model, sometimes with adjustment in real time. The main feature is the concept of runoff formation on repletion of storage, which means that runoff is not produced until the soil moisture content of the aeration zone reaches field capacity, and thereafter runoff equals the rainfall excess without further loss. The 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 main basin outlet. The model parameters are calibrated for a long rainfall-runoff time series. The time series can be divided into two parts, calibration and validation. The parameters are optimized in calibration series, the calibrated parameters are verified in validation. There are some criterions to evaluate the performance of parameters, if it can reach the measurement, the calibrated parameters can be determined.

Notwithstanding the performance of the XAJ model is successful in many watersheds in China, the real-time updating is necessary in real-time forecasting. The reasons can be inferred in Fig. 1. It can be seen from Fig. 1 (Li, 2008) that the simulation

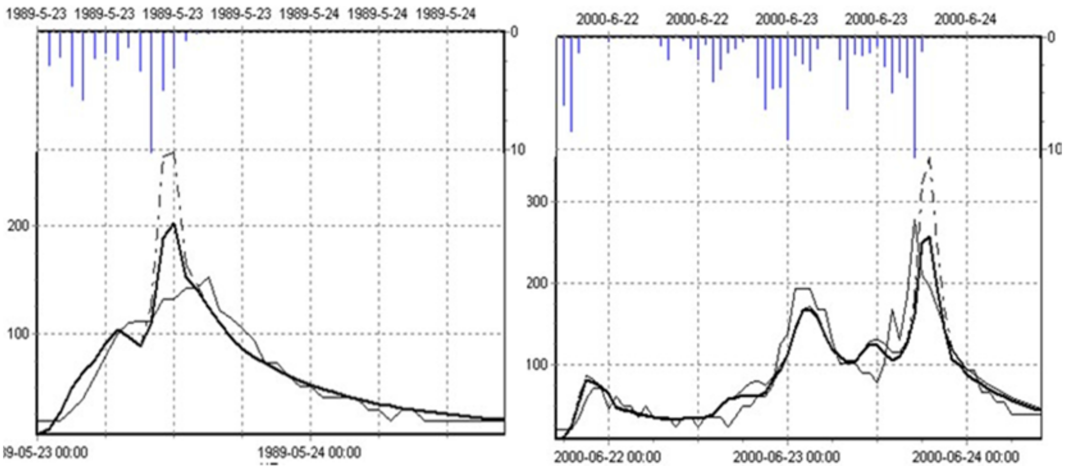

Fig. 1. Updating Effect about the XAJ Model

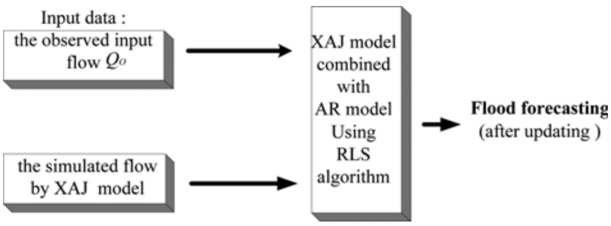

Fig. 2. Flow Chart of the Traditional Error Updating System

performance with error updating is better than that without error updating, the benefit is exist, but not obvious because the negative effect of coarse errors. Therefore, the improvement about error updating and how to decrease the negative effect of coarse errors is important.

More details about the XAJ model can be found in the model manual (Zhao *et al*., 1980; 1992).

The outline of the traditional error-updating in real-time forecasting system is shown in Fig. 2.

2.3 Problem Interpretation

The observed input flow Q_o for the reservoir is not measured immediately, but calculated by the observed water stage, out flow and stage-storage relation curve. Based on the water balance, the equation is given as:

$$\frac{Q_1+Q_2}{2}\Delta t-\frac{q_1+q_2}{2}\Delta t = V_2-V_1 \tag{7}$$

where, Q = Input flow of reservoir
q = Output flow of reservoir
$\bar{q}$ = Interval scale
V = Storage of the reach;

Subscripts 1 and 2 represent the starting and ending times of interval, respectively.

Then it becomes:

$$\bar{Q} = \frac{V_2-V_1}{a\times\Delta t}+\bar{q} \tag{8}$$

where, $\bar{Q}$ = Average of Q
$\bar{q}$ = Average of q
a = 3600 (by converting Δt from hour to sec)

Set $\Delta H = H_2-H_1$; H = The water lever of reservoir and $\bar{A}$ = the average area of reservoir section between H_1 and H_2, then Eq. (8) could be rewritten as:

$$\bar{Q} = \frac{\Delta H\times\bar{Q}}{a\times\Delta t}+\bar{q} \tag{9}$$

where, ζ is error existed in ΔH. Then the error of Q could be expressed as:

$$\Delta Q = \frac{b\times\zeta[m]\times\bar{A}[km^2]}{c\times\Delta t[h]} \tag{10}$$

where, b = 1000, c = 3.6

Table 1. Error Produced in Inflow (ΔQ) with Error Existed in Stage (ζ) and Reservoir Areas ($\bar{A}$)

	Reservoir Area $\bar{A}$ (km^2)					
Error existed in stage ζ(m)	10	20	50	100	200	500
0.01	27.8	55.6	139	278	556	1390
0.02	55.6	111	278	556	1110	2780
0.05	139	278	694	1390	2780	6940
0.08	222	444	1110	2220	4440	11100
0.10	278	556	1390	2780	5560	13900

In most forecasting system, we always set Δt =1 h, then ΔQ has a linear relationship to error ζ, coefficient = 1000 × A, that is error amplification ratio; it could be easily know from the Table 1 (Qu and Bao, 2003).

In general, during a flood accompanied by wind, the measures of reservoir stage often are affected by the wind, it is very common that 0.01or 0.02 m error existed in stage, but the coarse error of Q amplified by the tiny error of H and $\overline{A}$ could not be neglect. The coarse errors that are outliers have an unknown distribution with a much bigger variance, and appear to be inconsistent with the remainder of data set (Barnet and Lewis, 1994; Han and Kamber, 2001) but are relatively large in magnitude.

The calibration procedure of the standard AR model in updating the simulation errors of a rainfall-runoff model is firstly done by simulating the errors of the rainfall-runoff (substantive) model:

$$e_i = Q_i - QC_i \tag{11}$$

where, e_i denotes the simulation error of the XAJ model, Q_i is the observed discharge, and QC_i is the simulated discharge.

If the mean value of the simulation error series of the calibration period, e, is not equal to zero, then that mean should be subtracted from the simulated errors to produce a corresponding zero-mean time series, ε_i:

$$\varepsilon_i = e_i - \overline{e} \tag{12}$$

Thus, the AR updating model at the current time step i is:

$$\hat{\varepsilon}_i = a_1\varepsilon_{i-1} + a_2\varepsilon_{i-2} + \ldots + a_p\varepsilon_{i-p} \tag{13}$$

where, $a_1, a_2, \ldots, a_p$ are the coefficients used in the AR model, $\hat{\varepsilon}_i$ is the estimate of ε_i, and p is the number of AR coefficients. Recalling that the autocorrelation function of the ε_i time series satisfies an analogous form of linear difference equation to that of Eq. (13), one can obtain the Yule-Walker estimates of the parameters by replacing the theoretical autocorrelations by their respective estimates (obtained from the ε_i time series) (Box and Jenkins, 1976). Clearly, the AR error-forecast updating procedure is successful only when the residual $\varepsilon(t)$ is the random process of white noise. However, for real-time flood forecasting in reservoir watershed, the observed input flow Q_o for the reservoir is not measured immediately, but calculated by the observed water stage, out flow and stage-storage relation curve. The process produce the coarse error, causing $\varepsilon(t)$ displays abnormal distribution, which not only caused the fluctuation of the input flow hydrograph also weaken the ability of error updating for real-time flood forecasting. This paper introduce robust procedure to improve the observed input flow Q_o, to prevent these coarse error obtained in Q_o into the error updating system.

3. Error Updating System with Proposed Robust Procedure

The great characteristic of the system is the proposed robust process developed to modify the observed input flow by down-weighting the influence of outliers (Huber, 1981; Rey, 1977). A primary goal of robust procedure is to safeguard against coarse errors and produce a new modified input flow Q_o' instead of Q_o as the input of the error-updating system. The procedure is described in detailed:

Theoretically, discharge hydrograph should be a smooth and continuous curve; in this study, the observed input flow was firstly smoothed by the curve fitting method; the curve fitting means that the observed discharge hydrograph should be fitted by a conic, a smoothing observed discharge which across the saw tooth can be obtained. It is just called Q_{gh}.

In order to down-weight the influence of coarse error, the robust procedure is introduced to modify the observed input flow. In the robust procedure the assigned weight is a function of the residuals, and a loss function is selected to assign more weight to the bulk of small residuals while down-weighting small portion of coarse error. As the information contained in the input flow of reservoir is too limited to eliminate any discharge value, this paper selects the loss function expressed as:

$$w(t) = \begin{cases} 1 & |\xi(t)| \le k\sigma \\ k\sigma/\xi(t) & |\xi(t)| < k\sigma \end{cases} \tag{14}$$

$$\xi(t) = Q_o(t) - Q_{gh}(t) \tag{15}$$

$$\sigma = \sqrt{\sum_{i=1}^{m} w_i\xi_i^2/(m-1)} \tag{16}$$

where, k in Eq. (14) is constant, Good choice for K in robust theory are in the range between 1 and 2, say k = 1.5; for each event or each reservoir, k is the same value. The m in Eq. (16) is the number of data series, i means each data point, ω_i means the weight for each point, commonly, the weight for each data point is equal, so $\omega_i = 1$. It can be easy to known that the σ is a constant for each event. The effect of using (14) is to assign less weight to a small portion of large residuals so that the outliers will not greatly influence the final estimation, while giving unity weight to the bulk of small moderate residual.

3) The modified input flow can be expressed as:

$$Q_o'(t) = w(t) \times Q_o(t) + (1 - w(t)) \times Q_{gh}(t) \tag{17}$$

where, $w(t)$= Weighting factor
$\xi(t)$= Residual

Q_o and Q_o are the observed input discharge before and after modifying; Q_{gh} is the discharge after making smooth by curve fitting method. It should be noted that Q_o' is called ‘modified input flow’ in follow section.

Based on the above-mentioned robust procedure, the coarse error updating system with proposed robust process is seen in Fig. 3 and summarized as:

Step 1: the observed input flow is transformed by the proposed robust procedure to get to modified input flow. In this process, the coarse errors are down-weighted by the robust loss function.

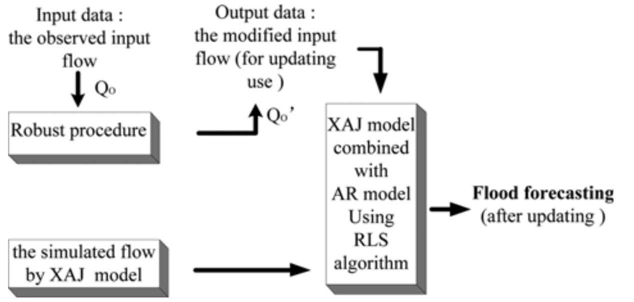

Fig. 3. Flow Chart of Error Updating System with Proposed Robust Process

Step 2: the modified input flow and the simulated flow are of the input data of AR model, the error updating process is conducted by RLS algorithm.

Step 3: after updating, the results of real-time flood forecasting is obtained.

4. Experimental Results

4.1 Study Area and Evaluation Criteria

Ten reservoirs with different features in humid area of China are selected to test the performance of the updating system reasonably and effectively. The characteristics of each reservoir basins are presented in Table 2.

In order to assess the accuracy of modeling results, two statistical indices are selected to judge the updating system performance. The detailed information about these indices is expressed as:

Relative runoff depth error:

$$e \times \frac{R_c - R_o}{R_o} \times 100\% \tag{18}$$

Robust root mean square error V:

$$V = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} w_i (Q_o(i) - Q_c(i))^2} \tag{19}$$

where, w_i is equal to the one in Eq. (16); R_C and R_O are the forecast and observed runoff depth; $Q_o(i)$ stands for the modified input flow Q_o'. $Q_c(i)$ represents the simulated flow with updating.

Relative runoff depth error (e) measures the bias of model performance. The optimal value is 0.0, which means that the model has an unbiased flow simulation. Positive values indicate a tendency of overestimation and negative values indicate a tendency of underestimation. The format of V is similar to that of *RMSE* which measure the difference between the forecast and observed flows. However, *RMSE* contain all information about the observed input flow including the outlier, which will have a negative effect on the authenticity of the updating results. V is developed to measure the dispersion degree between the forecast flows and observed one without outlier, called valid observed flows. There are two reasons for using V as the criterion: 1) It is used to evaluate the efficiency about the error updating system.2) it can measure the improvement of the error updating efficiency when introducing the robust theory. To some extent, it can reflect the efficiency of proposed real-time updating for resisting outlier.

4.2 Results

In order to confirm the advantage and evaluation of the proposed method, the performance of the error updating flood forecasting system with robust procedure is compared with that

Table 3. Comparison of the Performance for Updating System With and Without Robust Procedure involving Ten Reservoirs

Reservoir	e_{sx} (%)	e_{kc} (%)	V_{sx}	V_{kc}	E_V (%)
Longjinshan	2.16	2.05	85.658	70.614	17.56
Qinshitan	2.92	2.82	48.199	39.565	17.91
Dongxi	1.54	1.51	21.346	16.330	23.50
Dakai	2.64	2.44	15.943	11.941	25.10
Duihekou	0.98	1.01	8.623	6.255	27.46
Lishimen	1.57	1.48	19.523	15.205	22.12
Lushui	1.83	1.61	112.914	73.838	34.61
Naban	1.47	1.42	71.998	61.309	14.85
Dongzhen	0.76	0.79	48.511	40.188	17.16
Nanjiang	1.92	1.84	19.889	18.022	9.39
Average	1.78	1.70	45.260	35.327	20.97

Table 2. Characteristics of each Reservoir Basins

Reservoir	Average annual precipitation (mm)	Watershed area (km^2)	Water area (10^{5}m^2)	Calibration periods	Validation periods
Longjinshan	1723	285	30.5	1964-1979	1980-2002
Qinshitan	1800	474	150.6	1987-1993	1994-2000
Dongxi	1000	554	31.7	1988-1992	1994-1999
Dakai	1600	427	144.9	1981-1989	1991-2000
Duihekou	1100~1150	148.7	36.3	1984-1990	1992-1999
Lishimen	1895.4	296	34.7	1981-1990	1992-2000
Lushui	1550	3400	351.2	1979-1980	1981-1987
Naban	1715	490	148.6	1987-1992	1993-2001
Dongzhen	1200	321	108.8	1960-1985	1988-2000
Nanjiang	1200~2200	210	37.7	1986-1989	1990-2001

without robust procedure. The performance of the error updating flood forecasting system for ten reservoirs is showed in Table 3. The criterion Ev is expressed as:

$$Ev = (V_{sx} - V_{kc})/V_{sx} \qquad (20)$$

The subscript *sx* and *kc* stand for updating without and with robust procedure. E_v, which stands the decrease between V with robust procedure and without, is set to imply the efficiency of proposed the robust procedure. The value is much greater, the efficiency of proposed the robust procedure about resisting the coarse error is much higher and vice versa.

The Statistical results of ten reservoirs showed in Table 3 suggested that the proposed method was suited to resist the coarse error of input flow and improved the effect of correction in real-time forecasting. V_{kc} Was decrease obviously, the mean decrease rate of V ($\overline{E}_V$) was close to 20%. The maximal $\overline{E}_V$ was 34.61%, even for a single flood event, E_v = 60% (flood code is 31850603 in lushui). The relative error between the computed and observed runoff was decreased by the real-time forecasting correction. However, there were no obvious differences between the relative error of runoff with robust correction and that without .in other words, the proposed robust correction method kept the water balance. Furthermore, in order to illustrate the efficiency of the proposed updating system for resisting outliers and the influence of the outliers on updating results more detailed, the performance of updating system in Dongxi reservoir is shown and discussed. The comparison performance for updating system with and without robust procedure in Dongxi reservoir is presented in Table 4. The average E_v is 23.5%, which means the updating system with robust procedure is better than

Table 4. Comparison Performance Results in Dongxi Reservoir

Flood code	Start time	End Time	e_{sx} (%)	e_{kc} (%)	V_{sx}	V_{kc}	E_V (%)
31880520	1988-05-20 22:00	1988-05-23 10:00	4.15	4.02	38.753	32.899	15.11
31880618	1988-06-18 14:00	1988-06-23 19:00	1.27	1.41	29.459	28.146	4.46
31880904	1988-09-04 12:00	1988-09-06 21:00	5.52	5.51	17.293	10.729	37.96
31890522	1989-05-22 9:00	1989-05-25 12:00	3.29	3.56	16.772	15.889	5.26
31890526	1989-5-26 21:00	1989-05-30 16:00	0.77	0.35	17.959	13.765	23.35
31890721	1989-07-22 06:00	1989-07-24 21:00	1.64	0.24	39.556	27.142	31.38
31900611	1990-06-12 00:00	1990-06-15 11:00	1.53	1.53	18.736	17.817	4.90
31920516	1992-05-16 18:00	1992-05-19 18:00	-0.20	-0.02	36.720	28.459	22.50
31920614	1992-06-14 20:00	1992-06-18 23:00	0.38	0.34	11.279	9.145	18.92
31920622	1992-06-22 20:00	1992-06-27 21:00	-0.08	0.01	21.881	20.231	7.54
31920703	1992-07-03 15:00	1992-07-08 09:00	-1.55	-1.79	65.308	30.592	53.16
31930505	1993-05-05 16:00	1993-05-10 20:00	-0.74	-0.65	19.170	18.807	1.89
31930531	1993-05-31 11:00	1993-06-03 20:00	-0.64	-0.99	9.644	7.799	19.13
31930604	1993-06-04 09:00	1993-06-06 22:00	3.14	3.15	10.750	7.356	31.57
31930613	1993-06-13 14:00	1993-06-26 11:00	-0.11	-0.17	18.764	11.781	37.21
31930630	1993-06-30 13:00	1993-07-04 00:00	0.51	0.47	20.701	13.204	36.22
31930728	1993-07-28 10:00	1993-07-30 22:00	-4.52	-5.06	10.987	7.652	30.35
31940613	1994-06-13 12:00	1994-06-19 14:00	-1.25	-1.33	21.946	16.383	25.35
31950429	1995-04-29 19:00	1995-05-02 14:00	0.38	0.28	14.752	12.742	13.63
31950527	1995-05-28 08:00	1995-05-30 17:00	-0.34	-0.53	14.867	11.666	21.53
31950602	1995-06-03 04:00	1995-06-06 09:00	-0.08	-0.05	11.386	8.685	23.72
31950619	1995-06-19 09:00	1995-06-21 08:00	-0.42	-0.90	43.163	34.362	20.39
31950626	1995-06-26 21:00	1995-07-05 22:00	-0.35	-0.16	22.535	16.462	26.95
31950813	1995-08-13 09:00	1995-08-16 16:00	0.73	0.82	9.258	6.590	28.82
31960524	1996-05-25 23:00	1996-05-27 06:00	1.20	0.70	19.764	17.403	11.95
31970622	1997-06-24 13:00	1997-06-27 23:00	1.46	1.30	28.734	27.321	4.92
31970707	1997-07-07 22:00	1997-07-12 01:00	0.59	0.59	26.406	19.243	27.13
31970818	1997-08-18 20:00	1997-08-21 16:00	-2.18	-2.27	9.718	9.206	5.27
31980113	1998-01-13 14:00	1998-01-16 14:00	0.92	1.21	14.839	10.230	31.06
31980306	1998-03-07 23:00	1998-03-10 21:00	3.06	3.02	14.778	12.541	15.14
31980612	1998-06-12 09:00	1998-06-30 04:00	-0.83	-0.82	42.168	25.417	39.72
31980904	1998-09-04 15:00	1998-09-06 20:00	1.15	0.95	6.210	5.243	15.57
31990416	1999-04-16 22:00	1999-04-19 18:00	3.25	3.21	14.729	14.445	1.93
31990516	1999-05-16 09:00	1999-05-19 11:00	3.76	3.57	12.473	10.109	18.95
31990830	1999-08-30 09:00	1999-09-02 08:00	-1.80	-1.88	15.646	12.092	22.72
Average			1.54	1.51	21.346	16.330	23.50

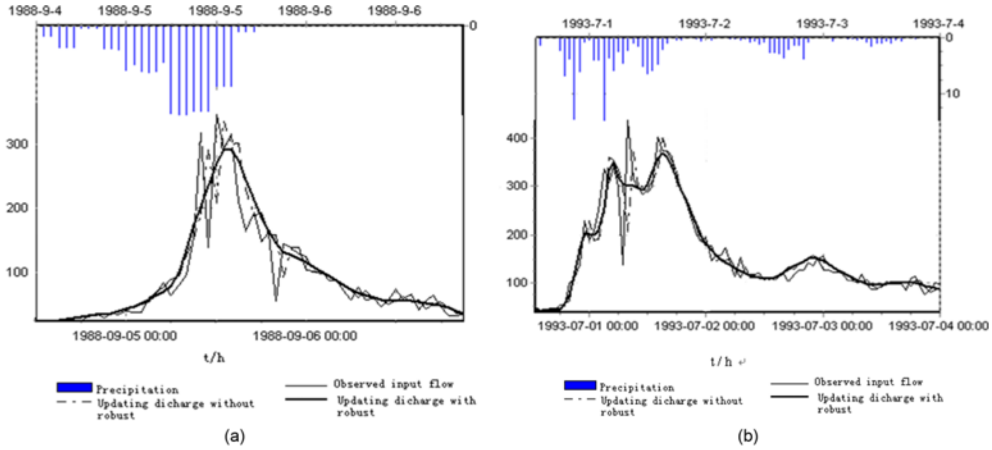

Fig. 4. (a) Hydrograph of Flood Event 31880904 in Dongxi (EV = 37.96%), (b) Hydrograph of Flood Event 31930630 in Dongxi (*EV* = 36.22%)

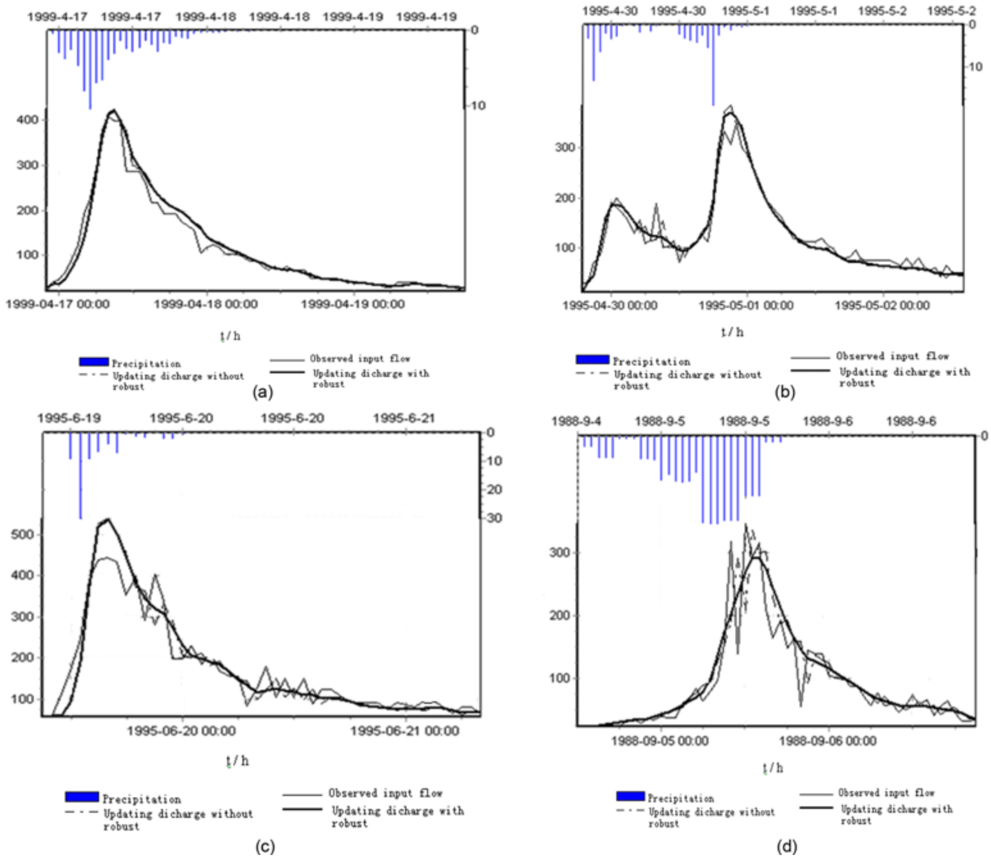

Fig. 5. (a) Hydrograph of Flood Event 31990416 (*EV* = 1.93%), (b) Hydrograph of Flood Event 31950429 (*EV* = 13.63%), (c) Hydrograph of Flood Event 31950619 (*EV* = 20.39%), (d) Hydrograph of Flood Event 31880904 (EV = 37.96%)

without that. Flood number 31920703 had a maximum *EV* of 53.16%. The difference of *e* is not obvious, which means the proposed new method keep the water balance. Summarizing, the updating system with robust procedure show the effective error-correction ability, even the coarse error exist.

In Fig. 4, the hydrographs of two flood events with relative large E_V are illustrated. It can be easily seen from the hydrographs of the two flood events that the observed input flow with the serration waves phenomenon because of coarse errors. After real-time updating, the forecast hydrograph is smooth and continuous. However, the graph of forecast discharge without robust procedure still show jagged and big jump similar to that ofobserved input flow. It indicates that the real-time updating system without robust procedure fail to overcome the influence of coarse errors from observed input flow. The forecast hydrograph with robust procedure show good outlier-resisting ability. The smoothing hydrograph through the jagged observed input flow felicitously. For a hydrological simulation research, the hydrograph obtained here is reasonable and be considered as more close to the actual trend of observed flow.

Then four flood events with different degrees of fluctuations are revealed in Fig. 5. With the increase in the degree of fluctuation, the range of E_V is from 1.93% to 37.96%. For 31990416, the hydrograph of observed input flow is relative smooth; the updating result with robust procedure is almost as same as that without. However, the difference is most obvious in 31880904 because of the highest fluctuationsin the observed input flow. The results indicate that the proposed robust procedure does well in resisting coarse errors in real-time updating. Thus, the more the coarse error, the more obvious is the error-resisting effect. As a whole, it was apparent that the results obtained by the smoothing robust correction method was reliable and effectively. The better correction effect indicated that the proposed method had ability to reduce the influence of the coarse error on real-time updating, especially for reducing the errors for peak stage.

5. Conclusions

Observed discharge is used as the main calibration quality, thus their errors may have significant influence on model performance. In most cases water levels are observed and rating curves are used to transform them to discharges, this is an important source of errors comprising the real-time forecasting. To prevent the negative influence of these errors and improve the effect of real-time forecasting, this paper proposed a robust procedure and obtained the following conclusion:

1. In this paper, the main source of the coarse error and the effect of coarse error on the observed discharge are investigated. It could be shown that a tiny measurement error of water level could cause huge error of observed discharge due to the amplification of water area of reservoir.
2. To prevent abnormal factors (such as coarse error) from entering the flood forecasting updating system, and ensure the stability of the system and the accuracy of updating, this paper modified the observed discharge with robust procedure before updating the errors of real-time forecasting; Seen from the results of ten reservoirs that this proposed process performed well in resisting the coarse error and improving the effect of correction, especially in case of observed input flow with acutely fluctuated.

References

Bao, W. M., Ji, H. X., and Hu, Q. M. (2004). “Robust estimation theory and its application to hydrology.” *Advances in Water Science*, Vol. 14, No. 4, pp. 133-137 (in Chinese).

Bao, W. M., Qu, S. M., and Li, Q. S. (2005). “Study on estimation methods of rainfall gauge errors in remote system.” *Journal of Hydraulic Engineering*, Vol. 4, No. 32, pp. 30-35.

Bao, W. M., Zhang, X. Q., and Yu, Z. B. (2011). “Real-time equivalent conversion correction on river stage forecasting with Manning’s formula.” *Journal of Hydrologic Engineering*, Vol. 16, No. 1, pp. 1-9.

Bao, W. M., Qu, S. M., Huang, X. Q., Hu, Q. M., Zhao, C., and Ji, H. X. (2003). “Analysis of robust weighting functions for hydrological systems.” *J. T singhua Univ*, Vol. 43, No. 8, pp. 1127-1129.

Barnett, V., and Lewis, T. (1994). *Outliers in statistical data*, 3rd Ed., John Wiley, Chichester, UK.

Box, G. E. P. and Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*, Holden-Day, Oakland, California, USA.

Broersen, P. M. T. and Weerts, A. H. (2005). “Automatic error correction of rainfall-runoff models in flood forecasting systems.” *Instrumentation and Measurement Technology Coference*, May 17-19, Otawa, Canada.

Eleftheriou, E. and Faleoner, D. (1986). “Traeking properties and steady-state performance of RLS adaptive filter algorithm.” *IEEE Trans. ASS*, Vol. 34, No. 5, pp. 1097-1109.

Goswani, M, O’Connor, K. M., Bhattarai, K. P., and Shamsedlin, A. Y. (2005). “Assessing the performance of eight real-time updating models and procedures for the Brosna River.” *Hydrology and Earth System Sciences,* Vol. 9, No. 4, pp. 394-411.

Han, J. W. and Kamber, M. (2001). *Data mining: Concepts and Techniques*, Morgan Kaufmann Publishers, USA.

Haykin, S. (1996). *Adaptive filter theory*, 3rd Ed., Prentive Hall, NewYork.

Hu, S. Y. (1987). “Problems with outlier test methods in flood frequency analysis.” *J. Hydrol.*, pp. 375-383.

Huber, P. J. (1981). *Robust statistics*, John Wiley, New York.

Karhunen, J. and Pajunen, P. Oja, E. (1998). “The nonlinear PCA criterion in blind source separation: Relation with other approaches.” *Neurcomputing*, Vol. 22, No. 1, pp. 5-20.

Li, R. R. and Wu, G. X. (2008). “Research on the robust correction method of a reservoir’s input flow.” *China Rural Water and Hydropower*, Vol. 11, pp. 12-14.

Poor, H. and Wang, X. D. (1997). “Code-aided interference suppression for DS-CDMA communications, Part 11: Parallel blind adaptive implementations.” *IEEETrans. Comm1*, Vol. 45, No. 9, pp. 1112-1122.

Qu, S. M. and Bao, W. M. (2003). “Comprehensive correction of real2time flood forecasting[J].” *Advances In Water Science*, Vol. 14, No. 2, 167-172.

Rey, W. J. (1977). *Robust statistical methods*, Springer, New York.

Toth, E., Montanari, A., and Brath, A. (1999). “Real-time flood forecasting via combined use of conceptual and stochastic models.” *Physics and Chemistry of the Earth Part B,* Vol. 24, No. 7, pp. 793-798.

Zhao, C., Bao, W. M., and Hong, H. S. (2009). “The robust detection of outlying rainfall observation.” *Hydroinformatics in Hydrology*, Hydrogeology and Water Resources, pp. 154-162.

Zhao, C., Hong, H. S., Bao, W. M., and Zhang, L. P. (2008). “Robust recursive estimation of auto-regressive updating model parameters for real-time flood forecasting.” *Journal of Hydrology,* Vol. 349, No. 5, pp. 376-382.

Zhao, R. J. (1992). “The Xinanjiang model applied in China.” *Journal of Hydrology*, Vol. 135, No. 16, pp. 371-381.

Zhao, R. J., Zhuang, Y. L., Fang, L. R., Liu, X. R., and Zhang, Q. S. (1980). “The Xinanjiang model.” *Hydrologica Forecasting Proceedings Oxford Symposium IAHS,* Vol. 129, No. 7, pp. 351-356.