

Real-time correction method of Muskingum model based on Kalman filter

Li qian, Yu huanfei, Hu linlin, Ge hangjian, and Zheng hongri

Zhejiang Institute of Hydraulics and Estuary, Zhejiang Provincial Key Laboratory of Hydraulic Disaster Prevention and Mitigation, Hangzhou, Zhejiang 310000, China

Abstract. In flood forecasting, general flood forecasting models or empirical forecasts reflect the average optimal value or relationship curve under the previous data. However, in the operation forecast, the forecast plan value often deviates from the actual situation. This paper takes Muskingum model as an example, and uses the Kalman filter algorithm to correct the forecast results. The algorithm structure and principles were described detailed, and the numerical simulation test was set to verify the efficiency of the Kalman filter algorithm. The correct results with corrected method were compared. The results indicated that the efficiency of the updating system using Kalman filter algorithm was improved. Conclusively, the proposed method could be widely applied in real-time flood forecast updating.

1 Introduction

Watershed hydrological system is a very complex natural system. When using the hydrological model for flood forecasting, due to various reasons, the actual situation is that each forecast has more or less errors (Bogner et.al 2008). Traditional hydrological forecasting methods are difficult to make up for some errors in the actual situation. Real-time forecasting is not appropriate for such considerations that were not considered in the original model, cannot be considered, or even considered, and there is a certain amount of flood estimation (Madsen, 2005). The factors that cause certain errors in the forecast, such as the structure, parameters, state variables, or input values of the model, are considered. In addition to a reasonable and effective model of river basin hydrological forecasting for flood forecasting, a set of reasonable and effective real-time correction techniques is also needed.

The research of traditional real-time correction technology mainly focuses on the improvement of correction methods and correction contents, such as Kalman filter technology (Babovic 2001; Komma, 2008; Trushnamayee 2019) etc. The improvement of research technology has shown obvious results in industrial automatic control and national defense cutting-edge science. The reason is that the former has a large amount of information that can be used, and simple technology cannot be fully utilized (Kidanidis, 1980). When the structure of the correction technology is complicated and the amount of information used is increased, the correction effect is improved. The amount of information on flood forecast errors is only sufficient to provide simple correction technology. Or even the amount of

information required for simple correction methods is not enough. For complex correction techniques, the amount of available real-time information has not been increased, and the correction effect naturally cannot be improved. In the real-time flood forecasting system, there are many reasons for the error, and the mechanism that affects the error is very complicated. Although the model calculates the real-time error series, although it contains all the error information, the amount of information that can be distinguished is too small to be able to reach the model parameters, Input errors, etc. Therefore, the expansion of real-time correction information utilization and improvement of utilization technology is the key to real-time flood forecast correction technology.

2 Methodology

2.1 Algorithm of Kalman Filtering

In the absence of external disturbances, the future state of the dynamic system can be determined based on the current state using a motion equation that describes the dynamic change of the system. Unfortunately, for an actual physical system, it is inevitable that there is some external interference, or people's description of the dynamic changes of the system is not very accurate enough. Therefore, the behavior of any actual physical system can be considered as consisting of two parts: one is predicted from a known equation of motion, and the other is a random component whose mean can be regarded as zero. Such a system can be regarded as a Markov sequence. If the dynamic system is assumed to be linear, its system equations and measurement equations are:

$$\begin{cases} \mathbf{X}_k = \Phi_k \mathbf{X}_{k-1} + \mathbf{G}_k \mathbf{U}_k + \Gamma_k \mathbf{W}_k \\ \mathbf{Y}_k = \mathbf{H}_k \mathbf{X}_k + \mathbf{V}_k \end{cases} \quad (1)$$

Among them, Φ_k represents the linear state equation at time k under the condition of time k-1; \mathbf{H}_k is the linear observation equation at time k; X and Y are state variables and measured variables; dynamic noise \mathbf{W}_k and measurement noise \mathbf{V}_k are A white noise sequence with zero mean, and the two are uncorrelated, that is, for all i and j:

$$\begin{cases} E(\mathbf{W}_i) = 0 \\ E(\mathbf{V}_i) = 0 \\ Cov(\mathbf{W}_i, \mathbf{W}_j) = E[\mathbf{W}_i, \mathbf{W}_j] = \mathbf{Q}_i \delta_{ij} \\ Cov(\mathbf{V}_i, \mathbf{V}_j) = E[\mathbf{V}_i, \mathbf{V}_j] = \mathbf{R}_i \delta_{ij} \\ Cov(\mathbf{W}_i, \mathbf{V}_j) = 0 \end{cases} \quad (2)$$

Where δ_{ij} is the Keronecker function, ie let the initial state \mathbf{X}_0 are a normal random variable, and its statistical characteristics are:

$$\begin{cases} E(\mathbf{X}_0) = \boldsymbol{\mu}_0 \\ \mathbf{Var}(\mathbf{X}_0) = \mathbf{P}_0 \\ Cov(\mathbf{X}_0, \mathbf{W}_i) = 0 \\ Cov(\mathbf{X}_0, \mathbf{V}_i) = 0 \end{cases} \quad (3)$$

The recurrence formula of kalman filtering is:

Filtering:

$$\hat{\mathbf{X}}_k = \hat{\mathbf{X}}_{k/k} = \hat{\mathbf{X}}_{k/k-1} + \mathbf{K}_k (\mathbf{Y}_k - \mathbf{H}_k \hat{\mathbf{X}}_{k/k-1}) \quad (4)$$

Forecasting:

$$\hat{\mathbf{X}}_{k/k-1} = \Phi_k \hat{\mathbf{X}}_k + \mathbf{G}_k \mathbf{U}_k \quad (5)$$

Gain matrix:

$$\mathbf{K}_k = \mathbf{P}_{k/k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k/k-1} \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (6)$$

State vector error covariance:

$$\mathbf{P}_{k/k-1} = \Phi_k \mathbf{P}_{k-1} \Phi_k^T + \Gamma_k \mathbf{Q}_k \Gamma_k^T \quad (7)$$

Filtering error covariance:

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k/k-1} \quad (8)$$

In addition, the initial filtering value of the recursive algorithm generally includes the first and second moments of the initial state value:

$$\begin{cases} E(\mathbf{X}_0) = \boldsymbol{\mu}_0 \\ \mathbf{Var}(\mathbf{X}_0) = \mathbf{P}_0 \end{cases} \quad (9)$$

2.2 Kalman Filtering for Muskingum Model (KFMM)

The prerequisite for the application of Kalman filtering is that the system must be linear, and river confluence can be viewed as a time-varying linear system in hydrology. The Muskingum model is the only model that can rely on hydrological methods to derive the parameter matrix. The mathematical form conforming to the linear system

equations provides the basis for the application of Kalman filtering. Muskingum's matrix solution is as follows:

A long reach is used as the research system, which is divided into n sub river sections. For the linear flow calculation, the calculation parameters K_i and x_i of each sub river section is equal. For the i-th sub-reach, the Muskingum equation is:

$$\begin{aligned} \mathbf{W}_i^{t+1} - \mathbf{W}_i^t &= \frac{\Delta t}{2} (\mathbf{I}_i^{t+1} + \mathbf{I}_i^t + \mathbf{q}_i^{t+1} + \mathbf{q}_i^t) - \frac{\Delta t}{2} (\mathbf{Q}_i^{t+1} + \mathbf{Q}_i^t) \\ \mathbf{W}_i &= \mathbf{K}_i [\mathbf{x}_i (\mathbf{I}_i + \mathbf{q}_i) + (1 - \mathbf{x}_i) \mathbf{Q}_i] \end{aligned} \quad (10)$$

Where: t is the time sequence number; \mathbf{I}_i and \mathbf{Q}_i is the inflow and outflow of the i-th sub-reach; \mathbf{q}_i are the interval or branch inflow of the i-th sub-reach; \mathbf{W}_i are the tank reserves of the sub-reach; K_i and x_i are the calculation parameters.

Solving Equ. (10), get

$$\mathbf{a}_i \mathbf{Q}_i^{t+1} + \mathbf{b}_i \mathbf{I}_i^{t+1} = \mathbf{c}_i \mathbf{Q}_i^t + \mathbf{d}_i \mathbf{I}_i^t + \mathbf{d}_i \mathbf{q}_i^t - \mathbf{b}_i \mathbf{q}_i^{t+1} \quad (11)$$

$$\begin{cases} \mathbf{a}_i = \mathbf{K}_i^{t+1} (1 - \mathbf{x}_i^{t+1}) + \frac{1}{2} \Delta t \\ \mathbf{b}_i = \mathbf{K}_i^{t+1} \mathbf{x}_i^{t+1} - \frac{1}{2} \Delta t \\ \mathbf{c}_i = \mathbf{K}_i^t (1 - \mathbf{x}_i^t) - \frac{1}{2} \Delta t \\ \mathbf{d}_i = \mathbf{K}_i^t \mathbf{x}_i^t + \frac{1}{2} \Delta t \end{cases}$$

Where

If the section flow on the forecasted river section \mathbf{I}_i

is merged into \mathbf{q}_i , and 4 sub river sections are taken as an example, the simplified Muskingum vector matrix equation can be obtained by derivation as:

$$\begin{bmatrix} \mathbf{Q}_1^{t+1} \\ \mathbf{Q}_2^{t+1} \\ \mathbf{Q}_3^{t+1} \\ \mathbf{Q}_4^{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_2 & & & \\ \mathbf{C}_0 \mathbf{C}_3 & \mathbf{C}_2 & & \\ \mathbf{C}_0^2 \mathbf{C}_3 & \mathbf{C}_0 \mathbf{C}_3 & \mathbf{C}_2 & \\ (\mathbf{C}_0)^2 \mathbf{C}_3 & \mathbf{C}_0 \mathbf{C}_3 & \mathbf{C}_3 & \mathbf{C}_2 \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1^t \\ \mathbf{Q}_2^t \\ \mathbf{Q}_3^t \\ \mathbf{Q}_4^t \end{bmatrix} + \begin{bmatrix} 1 & & & \\ \mathbf{C}_0 & 1 & & \\ (\mathbf{C}_0)^2 & \mathbf{C}_0 & 1 & \\ (\mathbf{C}_0)^3 & (\mathbf{C}_0)^2 & \mathbf{C}_0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{C}_0 \mathbf{q}_1^{t+1} + \mathbf{C}_1 \mathbf{q}_1^t \\ \mathbf{C}_0 \mathbf{q}_2^{t+1} + \mathbf{C}_1 \mathbf{q}_2^t \\ \mathbf{C}_0 \mathbf{q}_3^{t+1} + \mathbf{C}_1 \mathbf{q}_3^t \\ \mathbf{C}_0 \mathbf{q}_4^{t+1} + \mathbf{C}_1 \mathbf{q}_4^t \end{bmatrix}$$

$$\begin{cases} \mathbf{C}_0 = -\frac{\mathbf{b}}{\mathbf{a}} \\ \mathbf{C}_1 = \frac{\mathbf{d}}{\mathbf{a}} \\ \mathbf{C}_2 = \frac{\mathbf{c}}{\mathbf{a}} \\ \mathbf{C}_3 = \mathbf{C}_0 \mathbf{C}_3 + \mathbf{C}_1 \end{cases} \quad (12)$$

Where

It can be found that the expression of the Muskingum matrix equation is a linear dynamic system:

$$\mathbf{Q}_{t+1} = \boldsymbol{\phi} \mathbf{Q}_t + \mathbf{B} \mathbf{U}_{t+1} \quad (13)$$

Where:

$$\varphi_{i,j} = \begin{cases} 0 & i < j \\ C_{2,i} & i = j \\ C_{3,i} & i = j + 1 \\ C_{0,i}\varphi_{i-1,j} & i > j + 1 \end{cases};$$

$$b_{i,j} = \begin{cases} 0 & i < j \\ 1/a_j & i = j \\ C_{0,i}b_{i-1,j} & i > j \end{cases}, U_i = d_i q_i^t - b_i q_i^{t+1}$$

Based on the Muskingum matrix method, the state equations and measurement equations of the Kalman filter are constructed, and the default initial value X_0 and state vector filter error covariance matrix P_0 is given. The initial values of Q and R can be obtained from the statistical variance of the historical data samples of the actual measurement stations.

3 Synthetic Study and discussion

3.1 Study basin and data

The basin considered in following example is the Tingxia, located in the southeast of China, with a drainage area of 176 km². The climate is relatively humid and belongs to the semi-humid region. From the analysis of observed data, 70% of the mean annual rainfall falls between March and June, while evaporation is greatest between May and September. In this study, 13 years of historical data, 1988-2000, including hourly precipitation, pan evaporation and discharge, were used for parameter calibration while the data between 1999 and 2000 were used for the parameter validation. Hydrological data are used for calibration and validation of the hydrological model's parameters, including daily and hourly rainfall and runoff and daily evaporation rate. The runoff data are calculated based on the change in the water level of

the reservoirs. The rainfall data are obtained from rain gauges near the dams. The daily evaporation data are obtained by using daily evaporation pan data from evaporation station near the dams.

3.2 Numerical Results and discussion

In order to confirm the advantage and evaluation of the selected method (KFMM), the performance of the error correction flood forecasting system with KFMM is compared with that uncorrected. 26 flood events during 1988-2000 are used in numerical study. The Nash - Sutcliffe efficiency (NSE), the relative error of peak flow (ΔQ_m), and the relative error of runoff depth (ΔR) were used as evaluation criteria.

The relative error of runoff depth :

$$\Delta R(\%) = ((R_{cal} - R_{obs}) / R_{obs}) \times 100\%$$

The relative error of peak flow :

$$\Delta Q_m(\%) = ((Q_{mcal} - Q_{mobs}) / Q_{mobs}) \times 100\%$$

Nash - Sutcliffe Efficiency (NSE):

$$NSE = 1 - \frac{\sum_{i=1}^n (Q_{t-cal} - Q_{t-obs})^2}{\sum_{i=1}^n (Q_{t-obs} - \bar{Q}_{obs})^2}$$

Where R_{cal} is the calculated runoff depth and R_{obs} the measured runoff depth; Q_{mcal} to calculate the flood peak flow, Q_{mobs} is the measured flood peak flow; Q_{t-cal} is the calculated flow at time t, Q_{t-obs} is the measured flow at time t, and \bar{Q}_{obs} is the measured average flow. The forecast of flood peak and flood volume is based on the allowable error of 20% of the measured peak and flood volume.

The detailed results obtained by application of the KFMM, compared with uncorrected is displayed in Table 1.

Table 1 Comparison of the Performance for Updating System With and Without KFMM

Floodcode	Uncorrected			Corrected with Kalman Filtering		
	$\Delta R_u(\%)$	$\Delta Q_{m_u}(\%)$	NSE_u	$\Delta R_c(\%)$	$\Delta Q_{m_c}(\%)$	NSE_c
	/%					
31010623	-17.06	-17.46	0.40	-12.37	-26.27	0.64
31000709	1.76	2.15	0.66	1.06	3.78	0.75
31970707	-0.21	-3.05	0.9	-0.21	-4.35	0.93
31950702	2.30	6.17	0.79	1.47	2.61	0.81
31950428	7.94	1.69	0.76	5.60	3.30	0.84
31930703	-12.85	-9.30	0.94	-9.41	-6.70	0.96
31900623	5.61	-7.67	0.95	4.10	-7.04	0.96
31900614	-48.93	-0.48	0.67	-36.26	0.29	0.78
31890701	-7.48	1.40	0.87	-5.40	1.97	0.91
31890521	-16.75	15.90	0.91	-12.35	10.43	0.95

31890412	-4.18	3.12	0.94	-3.12	2.20	0.96
31880617	12.35	3.93	0.87	8.92	3.01	0.92
31000913	0.61	-0.96	0.82	0.02	-4.04	0.9
31000829	-7.13	8.85	0.86	-4.44	-4.23	0.92
31970816	-2.14	0.10	0.92	-1.35	4.85	0.92
31940821	14.63	1.03	0.68	9.25	1.89	0.8
31920922	-4.78	-11.91	0.61	-3.03	-5.18	0.69
31920830	-0.70	-8.95	0.89	-0.37	-1.27	0.89
31900908	1.23	-17.06	0.96	0.77	-15.44	0.97
31900904	-2.18	-5.54	0.91	-1.17	-0.81	0.93
31900830	0.49	8.08	0.91	0.25	7.05	0.94
31890912	-18.98	10.40	0.93	-11.86	0.99	0.95
31890831	2.01	18.91	0.66	1.53	7.05	0.84
31890818	-17.46	7.86	0.51	-10.74	-10.23	0.79
31880807	36.99	-14.45	0.71	22.83	-7.30	0.7
31880729	0.52	-15.81	0.5	0.27	-4.22	0.57

Notes: ΔR_u and ΔR_c mean the relative error of runoff depth without and with correction; ΔQ_{m_u} and ΔQ_{m_c} indicate the relative error of peak flow without and with correction.

Results analysis and discussion:

1. On the whole, the forecast of hydrological element state quantity correction is more accurate than the forecast without correction. Relative errors (ΔR and ΔQ_m) have been reduced. In particular, the certainty coefficient *NSE*, which reflects the consistency of the overall process of the sub-flood, has been improved after each flood.

2. However, the peak relative error of some floods increased after correction. The author made a detailed analysis of the Kalman filtering principle and the Muskingum matrix solution to draw the following conclusions: (1) Kalman filtering is mainly corrected by (Equ.4) and Prediction (Equ.5) consists of two major steps. The correction is to use the measured section state value at this moment to correct the outflow value of the upper section, and predict the exit section state value at the next moment based on the corrected outflow value; (2) The state equation of the Kalman filter constructed by the Muskingum matrix solution has input terms and the inflow of each interval; (3) Kalman filtering can only modify the state quantity that can be expressed in a matrix, that is, the section Outflow, it is not possible to modify the interval inflow obtained through the sloping field convergence calculation, and the existence of inflow errors will lead to unsatisfactory Kalman filtering prediction.

4 Summary and Conclusion

Real-time correction of flood forecast is an important link to improve forecast accuracy. The real-time flood forecasting correction models also have their own advantages and disadvantages. How to combine rich practical experience with modern control theory to develop a real-time correction model suitable for the characteristics of flood forecasting in this area is an important issue for forecasters and also improves the

development direction of flood forecast accuracy.

Kalman filtering technology, because of its scientific design ideas and design institutions, is widely used in hydrology. The correction effect of the Kalman filter technology has a certain effect in general, especially the improvement of the *NSE*. However, due to the existence of interval inflow errors, the correction effect is not particularly significant. Therefore, further research is needed on the applicability of Kalman filters in real-time correction of hydrological models and how to correct interval inflows that cannot be expressed by a matrix.

Acknowledgements

This study was supported by Zhejiang Provincial Natural Science Foundation of China (Grant No. LQ20E090002) and Zhejiang Provincial Research Institute Support Fund (Grant No. 2018F10027, 2018).

Reference

1. Bogner k, Kalas m. Error-correction methods and evaluation of an ensemble based hydrological forecasting system for the upper Danube catchment. *Atmospheric science letters*, 2008, 9: 95-102.
2. Babovic, V., Jensen, H. R., and Klinting, A., 2001. Neural networks as routine for error updating of numerical models. *Journal of Hydraulic Engineering*, 3(181), 181–193.
3. Kidanidis, P. K., Bras, R.L., 1980. Real-Time Forecasting with a Conceptual Hydrologic Model. *Water Resource Research*, (1)-16(6)1025-1033, 1980, (2)-16(6)1034-1044.
4. Komma j, Blöschl g, Reszler c. Soil moisture updating by ensemble kalman filtering in real-time flood forecasting. *Journal of hydrology*, 2008, 357: 228-242.

5. Madsen h, Skotner c. Adaptive state updating in real-time river flow forecasting-a combined filtering and error forecasting procedure. *Journal of Hydrology*, 2005: 308: 302–312.
6. Trushnamayee, N., Bhabagrahi, S., Chandranath, C., 2019. Enhancing real-time streamflow forecasts with wavelet-neural network based error-updating schemes and ECMWF meteorological predictions in Variable Infiltration Capacity model. *Journal of Hydrology*, 574 (000), 890-910.