

基于卡尔曼滤波技术的河道汇流实时校正

王船海, 白耀玲

(河海大学水文水资源与水利工程科学国家重点实验室, 江苏 南京 210098)

摘要 基于卡尔曼滤波的河道汇流实时校正算法, 采用非线性马斯京根矩阵解法, 将河道汇流表达为一个时变线性系统, 满足了卡尔曼滤波状态空间表达的要求. 在进行滤波时, 对系统参数矩阵的实时修正减小了河道汇流的非线性对线性滤波的影响. 通过对每个子流域的“远程”和“本地”卡尔曼滤波校正, 使全流域得到了全面合理的校正. 提出一种新的模型噪声方差阵和量测噪声方差阵实时统计方法, 该统计方法可有效地防止滤波发散现象, 提高滤波预报的精度. 通过在三峡区间流域上的应用, 证明了基于卡尔曼滤波的河道汇流实时校正算法的有效性.

关键词 河道汇流; 实时校正; 卡尔曼滤波; 实时统计方法; 滤波发散

中图分类号 P333 **文献标识码** A **文章编号** 1000-1980(2007)02-0181-05

水文实时预报的误差主要来自水文观测数据的误差和流域水文模型的误差. 为提高预报的精度, 有必要对河道汇流进行实时校正. 卡尔曼滤波法是一个有效的实时校正方法, 是水文预报的重要工具. 本文采用新安江三水源模型^[1]计算坡面产汇流, 采用马斯京根法进行河道洪水演进, 并在此基础上提出了一个基于卡尔曼滤波的河道汇流实时校正算法. 在该算法中, 采用马斯京根矩阵方程^[2]将河道汇流系统表达为一个线性时变系统, 而河道流量作为状态变量, 采用正规卡尔曼滤波进行校正. 为了弱化河道汇流的非线性对滤波的影响, 本文采用非线性马斯京根模型^[3]以校正状态方程的参数矩阵.

近年来对卡尔曼滤波在水文预报中的应用研究较多, 但是确定模型噪声方差矩阵和量测噪声方差矩阵是个难点, 因为这 2 种矩阵在水文系统中的数学规律尚不完全明确, 没有通用的处理方法^[4]. 为了防止滤波发散, 在校正中采用了多种误差补偿技术^[5-6]以及自适应方法, 但模型和系统的不匹配导致滤波不稳定的情况仍然存在^[3, 5]. 本文提出了一种新的实时统计方法以确定模型噪声方差矩阵和量测噪声方差矩阵, 大大降低了滤波发散的可能性.

1 河道汇流实时校正算法

1.1 一种确定噪声的实时统计方法

1.1.1 估计噪声的难点

应用卡尔曼滤波的难点之一是如何估计模型噪声方差矩阵 Q_k 和量测噪声方差矩阵 R_k . 由于噪声序列真值是未知的, Q_k 和 R_k 不能准确统计, 而增益矩阵 K_k 的估计完全依赖于 Q_k 和 R_k . 所以一般的滤波递推模型用于实时洪水预报容易引起预报值发散(或者偏离真值).

在数学上, 自适应滤波模型可以在线估计出 Q_k 和 R_k , 其中比较典型的是 Sage-Husa 算法. 但自适应滤波模型只适用于稳态系统, 要求 Q_k 和 R_k 是平稳随机过程^[7]. 在 Sage-Husa 算法中, 从 Q_k 及 R_k 的递推公式可以发现随着计算时间的增长, Q_k , R_k 的估计进入稳态, 参数递推成为等值延伸. 实际上, 由于河道流量是非平稳的随机过程, 模型噪声和量测噪声一般也是非平稳的随机过程.

在其他研究中, 采用自适应滤波模型、衰减记忆滤波来克服预报值发散是最常见的 2 种方式. 自适应滤波模型存在本文所述的问题, 而衰减记忆滤波方式通过引入一个衰减因子改变 $P_k|_{k-1}$ 的计算公式, 来对 $P_k|_k$ 进行补偿调节. 根据数值试验确定衰减因子大小. 但是由于滤波递推算式中各变量的变化对洪水实时预报的影响均很大, 其算式本身遵循一定的递推规律. 因此, 衰减因子只能起到微调的作用, 其值接近于单位阵 I , 偏大偏

小都会造成滤波器的不稳定^[8]。因此本文提出了估计噪声协方差矩阵的方法,有效地防止了滤波发散。

1.1.2 本文提出的估计噪声协方差矩阵方法

首先由滤波公式可以发现增益矩阵 K_k 为预报值和实测值的权重:

$$X_k = X_k|_{k-1} + K_k(Z_k - H_kX_k|_{k-1})$$

式中: X_k ——状态向量的滤波值; $X_k|_{k-1}$ ——状态向量的预测值; Z_k ——观测向量; H_k ——观测矩阵。 K_k 与信噪比 (Q_{k-1} 与 R_k 的能量比) 的大小成比例。故要得到合理的滤波值, 取决于 Q_{k-1} 与 R_k 的相对关系, 因此不一定要精确求出 Q_k, R_k , 只要得到近似估计, 而该估计能大体反映量测值与模型计算值哪个更值得“信任”即可。

对于 R , 其反映的是量测值与真值的偏差程度。对于山区性河流, 水位流量多呈单一线, 此时, 将资料整编得到的水位流量单一线中的流量看作真值。在量测系统误差为零的前提下, 可以假设各水位级下量测误差序列是平稳序列。据该单一线统计出不同水位级下的量测噪声方差矩阵的元素值 r , 然后写出对角阵的形式, 即可得到量测噪声方差矩阵。

$$r = \frac{1}{n} \sum_{t=1}^n (Q_{\text{实}t} - Q_{\text{线}})(Q_{\text{实}t} - Q_{\text{线}})$$

式中: $Q_{\text{线}}$ ——不同水位级下逐级通过线性计算得到的流量; $Q_{\text{实}t}$ —— t 时刻的实测流量。

流量 Q 反映的是模型计算值与真值的偏差程度。虽然模型误差是一个非平稳序列, 但是考虑到水文系统是一个时间渐变系统^[9]及模型的误差无法避免, 相邻时刻间的误差统计特性具有一定程度的相似性, 故可以把实测值看成真值。在实时预报中, 动态选择一个较小的样本, 统计计算出模型噪声方差矩阵的元素值 q , 写成对角阵的形式, 即可得到模型噪声方差矩阵。

$$q = \frac{1}{n} \sum_{t=1}^n (Q_{\text{计算}} - Q_{\text{实}t})(Q_{\text{计算}} - Q_{\text{实}t})$$

式中 $Q_{\text{计算}}$ 为模型计算得到的流量。

1.2 算法步骤

卡尔曼滤波算法适合于大的流域和水情测报系统。若流域内有较多个水文站, 卡尔曼滤波算法可以对系统的状态变量进行最优估计, 对全流域进行全面、合理的校正, 既能达到最小方差, 又不损失预见期, 是比较理想的方法。

基于卡尔曼滤波的河道汇流实时校正算法的步骤如下:

a. 利用可视化数字高程流域水系模型系统 DEDNM^[10](一种数字地形分析工具) 自动生成水系及子流域和河网结构拓扑关系。在此基础上, 用该系统进行干支流分级, 设置子流域级别。

b. 遍历各子流域, 如果该子流域有实测水文站(流量站), 在每个水文站设置一个校正要素。每个校正要素包含了实测和计算的对比信息, 以及该实测点量测噪声、计算点模型噪声的信息, 实测资料的数据类型等。

c. 生成校正子流域组。校正子流域组是校正的最小单元。分析各子流域的拓扑关系, 将全流域划分为一些不同级别的校正子流域组, 并分析各组间的拓扑关系。校正子流域组是为校正需要而设立的, 是一组入流连续的子流域, 且具有同一级别。校正子流域组可以动态生成, 基本方法是从某个“头”子流域向流入的子流域不断搜索, 直至发现其流入的子流域“变级”为止, 历经的这些子流域就是一个校正子流域组。

d. 如果该校正子流域组内含的子流域具有校正要素, 就为该校正子流域组内建卡尔曼滤波器。该卡尔曼滤波器的状态向量维数为该组内子流域河段数总和, 量测向量维数为该组内子流域内置的校正要素总和。根据组内每个河段 t 时刻和 $t+1$ 时刻的汇流参数 $K_i^{t+1}, x_i^{t+1}, K_i^t, x_i^t$, 计算得到状态转移矩阵^[2]。而量测矩阵的建立方式决定于实时校正要素的类型。实测资料为水位时, 量测矩阵为流量到水位的转换矩阵; 实测资料为流量时, 量测矩阵为单位阵。

赋予默认的初值 x_0 、状态向量的滤波误差协方差 P_0 。如果系统稳定, x_0 和 P_0 的值只会对前几个时段的计算结果有影响, 算法默认 x_0 取 0, R 是 P 的上界^[8], P_0 值取一个较大的值^[5]。 Q, R 的初值可以根据各实测站历史资料样本统计方差得到。

e. 在实际预报中, 当“新息”出现时, 由低一级别的校正子流域组(支流)向高一级别的校正子流域组(主干)进行实时校正。如一个流域(图 1 图中箭头表示站点在流域的位置), 由子流域 A, B, ..., G、水文站 a, b,

... e 组成, A 为三级子流域, B, C, D 为二级子流域, E, F, G 为主干, 即一级子流域. 根据级别及分析拓扑关系, 其构成 4 个校正子流域组: A , $B-C, D, E-F-G$. A 是三级子流域组, $B-C, D$ 为二级子流域组, $E-F-G$ 为一级子流域组. 该流域的校正顺序如下: (a) 校正 A , A 产生“本地”校正量, 该校正量传入 $B-C$. (b) 由于旁侧入流更新, $B-C$ 重新进行流量演算, 2 次演算的差值就是 $B-C$ 的远程校正量. $B-C$ 再进行校正, 产生“本地”校正量. $B-C$ 的校正量为“远程”校正量与“本地”校正量之和. 此校正量传入 $E-F-G$. (c) 校正 D , D 产生“本地”校正量, 该校正量传入 $E-F-G$. (d) 与 $B-C$ 的校正方式一样, $E-F-G$ 产生的校正量为“远程”校正量与“本地”校正量之和, 因再无流入的子流域组, 校正完成.

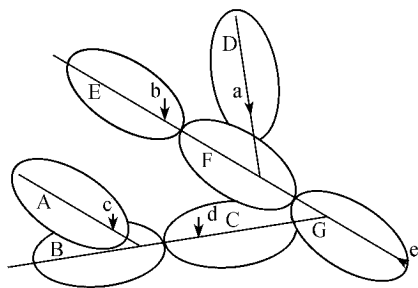


图 1 示例流域

Fig.1 Example of watershed

校正子流域组内任何一个校正要素通过卡尔曼滤波器都会对整组子流域的校正发生作用, 该校正量可称为“本地”卡尔曼滤波校正量. 校正某个校正子流域组的状态量后, 其流入高一级别校正子流域组(主干)的旁侧入流需要立即更新, 方法是将计算拉回到上一时刻, 重新计算高一级别子流域组在当前时刻的流量. 低一级别子流域经校正的入流使得该子流域需要第二次校正, 故这一校正量可称为“远程”卡尔曼滤波校正量. 每一次校正都让该校正子流域组的状态量更接近真值. 必须指出, 第二次校正时状态转移矩阵应根据经校正的入流量重新计算, 因为由某个校正子流域组经滤波器校正后得到当前子流域的入流后, $K, \hat{x}$ 势必要重新得到校正. 若不校正, 在马斯京根非线性差分迭代中, 由于 $\hat{x}, \bar{K}$ 无法取准, 相当于目标函数偏离, 有时会使流量过程线呈锯齿状.

基于卡尔曼滤波的河道汇流实时校正算法的优点是, 对于某一校正子流域组, 其自身将得到二次校正, 每一次校正都让其状态量更接近真值, 有望取得很好的校正效果. 而在并入数字流域系统后, 又具有比较好的通用性, 具体表现为当子流域划分, 河段数目产生变化, 或增、减某个实测站点时, 都将动态生成不同维数的卡尔曼滤波器.

2 实例验证

通过非线性显著的三峡区间的应用, 验证本文提出的基于卡尔曼滤波的河道汇流实时校正算法.

2.1 三峡区间水文情况简介

三峡区间一般指长江流域中寸滩—武隆—三峡坝址之间的区域. 该区间长江干流河长 614 km, 区间面积 52726 km². 河道分段情况如下: 寸滩—清溪场, 干流河段长 129 km, 区间面积 16262 km²; 清溪场—万县, 干流河段长 192 km, 区间面积 9026 km²; 万县—奉节, 干流河段长 127 km, 区间面积 12830 km²; 奉节—三峡坝址, 干流长 166 km, 区间面积 8650 km². 图 2 是数字流域水系模型^[8]生成的三峡区间水系和子流域分布图. 现有资料时段为 1981~1987 年. 资料并不十分完整, 只有 1981~1983 年寸滩和武隆有上游来水过程资料. 选择若干场次洪统计, 可得清溪场、万县、奉节、宜昌的预见期大约为 9 h, 27 h, 39 h, 50 h. 根据现有资料, 有 13 个目标断面在实时预报中会有实测“新息”出现, 可以为其逐一建立实时校正要素.

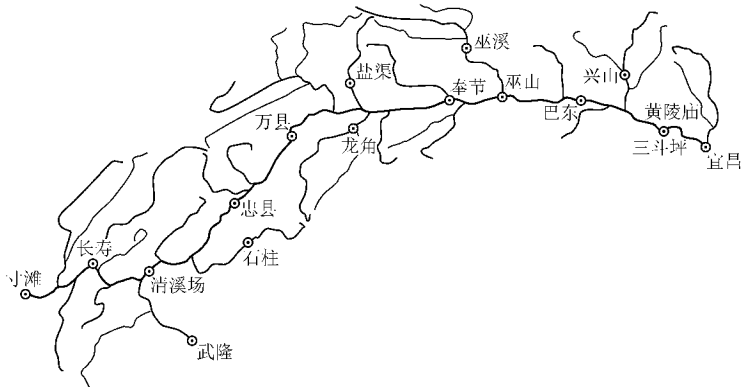


图 2 三峡区间水系和子流域分布

Fig.2 Drainage network and sub-basins in Three-Gorges region

2.2 验证结果

1983年7~9月三峡区间经历了3个较大的洪峰.采用本文提出的算法进行滤波预报.表1~3分别是各目标断面预见期不校正的模拟预报(采用线性马斯京根法进行洪水演进)和滤波预报(采用非线性马斯京根模型校正汇流参数和系统参数矩阵)的结果精度对比.图3是各目标断面预见期滤波预报与实测过程线的对比.

表1 洪峰1不校正的模拟预报和滤波预报的结果精度对比

Table 1 Comparison of accuracy of non-corrected result of simulative forecasting with that of filtering forecasting for flood peak 1						
站名	预见期/h	实测流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	不校正的模拟预报 (采用线性马斯京根法)		滤波预报 采用非线性 马斯京根模型)	
			流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	相对 误差/%	流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	相对 误差/%
清溪场	10	62000	56979.37	-8.10	56748.40	-8.47
万县	30	56800	54939.27	-3.28	52539.96	-7.50
奉节	40	52100	52441.98	0.66	50620.49	-2.84
宜昌	50	52400	51999.51	-0.76	50382.69	-3.85

表2 洪峰2不校正的模拟预报和滤波预报的结果精度对比

Table 2 Comparison of accuracy of non-corrected result of simulative forecasting with that of filtering forecasting for flood peak 2						
站名	预见期/h	实测流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	不校正的模拟预报 (采用线性马斯京根法)		滤波预报 采用非线性 马斯京根模型)	
			流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	相对 误差/%	流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	相对 误差/%
清溪场	10	59300	59149.46	-0.25	57678.56	-2.73
万县	30	56600	58784.99	3.86	55681.69	-1.62
奉节	40	54200	60341.21	11.33	56654.81	4.53
宜昌	50	53500	60896.35	13.82	57087.30	6.71

表3 洪峰3不校正的模拟预报和滤波预报的结果精度对比

Table 3 Comparison of accuracy of non-corrected result of simulative forecasting with that of filtering forecasting for flood peak 3						
站名	预见期/h	实测流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	不校正的模拟预报 (采用线性马斯京根法)		滤波预报 采用非线性 马斯京根模型)	
			流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	相对 误差/%	流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	相对 误差/%
清溪场	10	46700	47942.35	2.66	46332.45	-0.79
万县	30	47300	48822.17	3.22	45454.78	-3.90
奉节	40	45000	49579.55	10.18	46302.45	2.89
宜昌	50	46800	50254.29	7.38	49195.15	5.12

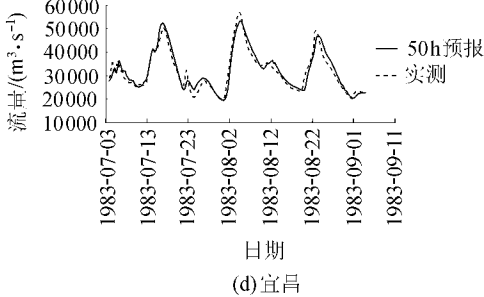
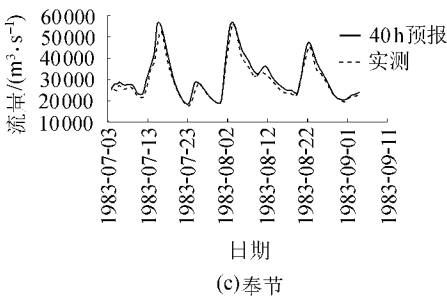
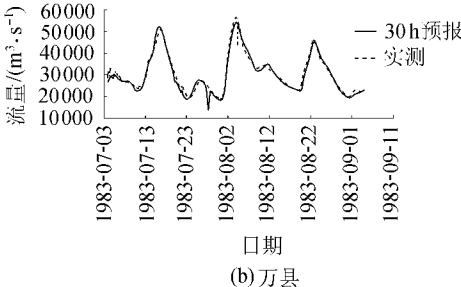
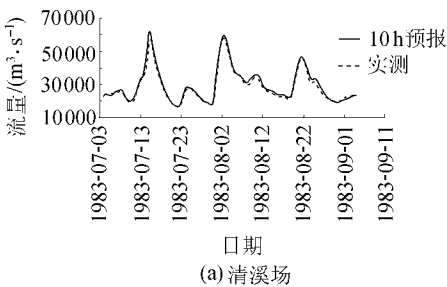


图3 各目标断面预见期滤波预报与实测对比

Fig.3 Comparison of discharge curves of each objective section between filtering forecasting and field measurement for different forecast periods

2.3 成果分析

在给预报模型设置实时校正子要素后,就可以在预报中实现基于卡尔曼滤波的水文要素状态量校正预报.

根据表1~3,可以发现:

a. 从总体上看,水文要素状态量实时校正的预报,较不做校正的模拟预报精度要高.第2,3个洪峰,在

各目标断面预见期 4 个目标断面滤波预报较模拟预报精度都有明显提高 校正效果明显.

b. 当参数率定不准确或是遇到区间有水库放水时 校正会产生“负效应”即随着预报时段数的增加 校正预报的结果会比不校正预报精度有一定程度下降. 分析以上预报结果 发现洪峰 1 流量相对误差采用的效果较模拟预报差. 究其原因 由于在流域内有一些小水库的放水情况无法得知 而根据 1981 ~ 1983 年的实测资料率定得到的模型参数涵盖了水库放水对河道汇流的影响 因此很可能使得率定的参数失真 现有模型参数与校正后的状态量不匹配 需要重新调整模型参数. 所以 1983 年 7 ~ 9 月率定成果的预报精度反而更高. 此外 全流域的水文状态向量是一个整体 每个分量的校正是互相制约影响的 当上游水量以校正的方式增减时 势必要影响到下游 当参数率定不准确或是遇到区间有水库放水时 这样的影响有时候是消极的 会“破坏”原本比较准确的不校正预报结果.

3 结 语

本文提出的基于卡尔曼滤波的河道汇流实时校正算法在流域水文预报上的应用表明 该算法能提高实时预报的精度 同时有效地防止了滤波发散 达到了滤波的稳定以及系统和模型的匹配. 在未来卡尔曼滤波法在水文实时预报的应用研究中 需要进一步研究如何更好地确定模型噪声方差矩阵和量测噪声方差矩阵 以及非线性卡尔曼滤波在水文预报中的应用. 同时 需要深入研究流域水文模型产流参数的校正方法以提高实时水文预报的精度. 流域内有实测资料的站较少时 可以利用历史资料 采用人工神经网络和卡尔曼滤波结合起来的方法进行实时校正.

参考文献：

[1] 赵人俊. 流域水文模拟——新安江模型与陕北模型[M]. 北京: 水利电力出版社, 1984: 50-73.
[2] 朱华. 马斯京根法的矩阵方程求解法[J]. 水文, 1987(4): 13-15.
[3] 陈森林. 马斯京根法非线性演算模型的差分牛顿迭代解法[J]. 水文, 1995(1): 30-33.
[4] 葛守西. 现代洪水预报技术[M]. 北京: 水利水电出版社, 2002: 117-120.
[5] 陈小红, 陈俊合. 嘉陵江流域三江汇流洪水联机实时预报研究[J]. 水文, 1995(4): 31-34.
[6] 王燕生, 王舜燕. 适应性卡尔曼滤波在水文预报中的应用[J]. 水文, 1985(1): 56-58.
[7] 王俊德. 水文统计[M]. 北京: 水利电力出版社, 1993: 100-105.
[8] 宋文尧, 张牙. 卡尔曼滤波[M]. 北京: 科学出版社, 1991: 67-70.
[9] 王船海, 郭丽君, 芮孝芳, 等. 三峡区间入库洪水实时预报系统研究[J]. 水科学进展, 2003, 14(6): 677-681.
[10] 丁晶, 邓育仁. 随机水文学[M]. 成都: 成都科技大学出版社, 1998: 90-112.

Real-time correction of flow concentration of river channels based on Kalman filter technique

WANG Chuan-hai , BAI Yao-ling

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

Abstract :The real-time correction algorithm based on Kalman filter technique was proposed to improve the accuracy of real-time hydrological forecasting under the circumstances that the nonlinearity of flow concentration of river channels is evident. In Muskingum matrix algorithm , the flow concentration was expressed as a linear system varying with time , which satisfied the requirement of spatial expression of Kalman filter. The real-time correction of the parameter matrix of the system decreased the influences of the nonlinearity of flow concentration on the linear filter. The advantage of the algorithm is that each sub-basin can be corrected twice by obtaining of the remote and local correction values , and then the whole basin can be rationally updated. Moreover , a real-time statistical method was put forward to estimate the system noise covariance matrix and measurement noise covariance matrix , which effectively prevented the filter divergence and improved the accuracy of forecasting. The effectiveness of the algorithm was verified by its application to some river channels in Three-Gorges region.

Key words :flow concentration of river channel ; real-time correction ; Kalman filter ; real-time statistical method ; filter divergence