International Seminar on Hydrology of Extremes

1-3 December, 1988
Roorkee (India)

MULTI-DAY STREAMFLOW FORECASTING USING THE KALMAN FILTER

Murray Fitch Edward McBean

Department of Civil Engineering, University of Waterloo
Waterloo, Ontario, N2L 3G1, Canada

SYNOPSIS

The performance of a river flow forecasting model employing a Kalman Filter algorithm is evaluated for increasing forecast times. The expected decrease in forecast accuracy is quantified and a decrease in forecast precision is noted for increased lead times.

INTRODUCTION

The prediction of future river flows is an important component of the operation of many water resources systems, including the opportunity for flood damage mitigation or assistance in the operations planning of reservoirs. In response, a variety of mathematical approaches have been developed. Techniques based on existing and expected meteorological conditions, past behavior of the river basin, the physical condition of the soil, and the historical correlation structure between flows from different time periods, have all been used with varying degrees of success.

An important element of the forecasting problem in Canada is that the forecasts must determine the magnitude of river flow arising from both the dissipation in the spring of the winter's snowpack and from the incidence of precipitation on the melting snowpack. As well, since many of these rivers are situated in isolated locations and often drain large areas, forecasts of spring runoff can be hampered by a lack of readily available data. The availability of data can be lacking both in terms of timely data (i.e. the data needed at time t is not accessible to the forecaster until time t+1 or later), and in terms of spatial variability (i.e. the locations of data collection are remote from the forecaster). The sparsity of the data networks can result in considerable uncertainty in the data which are available.

The lack of an extensive data network in Canada must be reflected in any forecasting procedure implemented. This implies that a detailed physical representation of the snowpack dissipation component of the flow generation process will not be feasible. The uncertainty associated with the data which are available indicates that there is a need for a technique which can reflect errors in the data employed as input to the model.

In this context, the state-space approach for developing flow forecasts consists of a dynamic model which describes the state of the system over time, subject to modelling errors, and a measurement equation which provides noisy estimates of some of the states of the system. Although there are many techniques available to update the state of a system using information gained from measurements taken of some of the states of the system, the most powerful of these methods is probably the Kalman filter. Previous work in this context includes Todini et al (1976), Todini and O'Connell (1972), Wood and Szollosi-Nagy (1978), and Burn and McBean (1985 and 1986).

While the research published-to-date has demonstrated the potential for utilizing Kalman filtering techniques for forecasting with actual data, an important need exists to quantify the accuracy of the forecasts with alternative forecast periods and gain insights into the assignment of parameters for application to real problems. The focus of this paper is to examine some of these considerations in application to the Sturgeon River Basin.

KALMAN FILTER AND FORECAST RESULTS

The Kalman filter employed herein is the mutually interactive state/parameter estimation (MISP) technique, developed by Todini (1978) and modified in Burn and McBean (1985). The technique involves using two filters: one in the state-space and one in the parameter-space. The two filters are used recursively each time step, with information from each filter being used in the optimal estimation performed by the other filter.

The equations used in the MISP filter are documented in Table 1 and interested readers are referred to Todini (1978) or Burn and McBean (1985) for further details. The case study region for which the forecasts were developed is the Sturgeon River basin located in the Canadian shield region of northern Ontario. The Sturgeon River drains an area of approximately 7000 km^2. The data employed for the prediction of river flow using the MISP algorithm and the Martinec snowmelt model includes:

(i) use of Theissen polygons to generate areal precipitation data;
(ii) daily maximum and minimum temperatures;
(iii) snow course survey data available at 2 week intervals; and,
(iv) a stream gauge at the location of forecasting.

The model's forecasting ability was tested in application to five different spring periods. The state vector used to forecast river flow was

$$X_t = [q_t, P_t, T_t, q_{t-1}, P_{t-1}, P_{t-2}, P_{t-3}, P_{t-4}, T_{t-1}, T_{t-2}]^T \qquad (1)$$

where q_t = the flow at time t, P_t = the precipitation (as rainfall) at time t, and T_t = the value of the temperature index at time t. The vector X_t is then used to forecast the flow at time t+1 through the equation

$$X_{t+1} = \phi_t X_t + \Gamma_t W_t \qquad (2)$$

where ϕ_t = transition matrix containing the model parameters

Γ_t = a scaling matrix

W_t = system noise matrix

Table 1 Changes to the MISP Algorithm

(a) $\hat{\vartheta}_t = \frac{t-1}{t}\hat{\vartheta}_{t-1} + \frac{1}{t}(Z_t - H_t\hat{X}_{t/t-1})$	Updating of the measurement error vector.
$\nu_t = Z_t - H_t\hat{X}_{t/t-1} - \hat{\vartheta}_t$	Filter innovation is calculated.
(b) $\hat{R}_t = \frac{t-1}{t}\hat{R}_{t-1} + \frac{1}{t}(\nu_t\nu_t^T - H_tP_{t/t-1}H_t^T)$	Measurement noise matrix is updated.
$K_t = P_{t/t-1}H_t^T(H_tP_{t/t-1}H_t^T + \hat{R}_t)^{-1}$	Kalman filter gain is calculated.
$\hat{X}_{t/t} = \hat{X}_{t/t-1} + K_t\,\nu_t$	State vector is updated.
$P_{t/t} = (I - K_tH_t)P_{t/t-1}$	Error covariance matrix is updated.
(c) $\nu_t^* = H_t\,K_t\,\nu_t$	Parameter filter innovation is calculated.
$R^* = H_t\,K_t\,C_0\,K_t^T\,H_t^T \;:\; C_0 = (R + H_tP_{t/t-1}H_t^T)$	Parameter measurement noise is calculated.
$K_t^* = P_{t/t-1}^*\,H_t^{*T}(H_t^*\,P_{t/t-1}^*\,H_t^{*T} + R_t^*)^{-1}$	Kalman filter gain for parameter filter is calculated.
(d) $\hat{\theta}_{t/t} = \hat{\theta}_{t/t-1} + K_t^*\,\nu_t^*$	Parameter vector is updated.
$P_{t/t}^* = (I - K_t^*\,H_t^*)P_{t/t-1}^*$	Paramter error covariance matrix is updated.
$\hat{\theta}_{t+1/t} = \hat{\theta}_{t/t} + \Gamma_{t+1}^*\,\bar{W}^*$	Parameter vector is extrapolated.
$P_{t+1/t}^* = P_{t/t}^* + \Gamma_{t+1}^*\,Q^*\,\Gamma_{t+1}^{*T}$	Extrapolated error covariance matrix for the parameter filter is calculated.
$\hat{\bar{W}}_{t+1/t} = \hat{\bar{W}}_{t/t-1} + \frac{1}{t}(\Gamma_t^T\,\Gamma_t)^{-1}\Gamma_t^T\,K_t\,\nu_t$	Updating of the system error vector.
(b) $\hat{Q}_{t+1/t} = \frac{t-1}{t}\hat{Q}_{t-1/t} + \frac{1}{t}(\Gamma_t^T\Gamma_t)^{-1}\Gamma_t^T(K_t\nu_t\nu_t^TK_t^T + P_{t/t} - \hat{\phi}_{t-1}P_{t-1/t-1}\phi_{t-1}^T)\Gamma_t(\Gamma_t^T\Gamma_t)^{-1}$	System noise matrix is updated.
$\hat{X}_{t+1/t} = \hat{\phi}_t\hat{X}_{t/t} + \Gamma_{t+1}\,\hat{\bar{W}}_{t+1/t}$	State vector is forecast.
$P_{t+1/t} = \hat{\phi}_tP_{t/t}\hat{\phi}_t^T + \Gamma_{t+1}\hat{Q}_{t+1/t}\Gamma_{t+1}^T$	Extrapolated error covariance matrix is calculated.

Notes:

(a) Prior value (t-1) was set to zero.

(b) Variance (diagonal) elements were constrained to be non-negative.

(c) Parameter filter innovation was divided in half.

(d) Parameters were only updated when there was an input associated with the parameter.

For the two-day ahead forecast (t.2), the measurements at t+1 are unknown and were taken as those at time t. Research work is underway to improve upon this assumption.

Examples of the forecasting results obtained are provided in Figures 1 through 4, for 1968 one-day and one- to three-day forecasts for 1970, respectively. Quantitative results on peak forecasts for the four years of forecasts with respect to timing and magnitude, are provided in Table 2. As expected, the results for the one-day ahead forecasting are better than the two-day, and, in turn, for the three-day forecasts.

Table 2 Sturgeon River Model Tests

Year	RMS Error			Peak Flow Error (forecast/measured)		
	* 1	2	3	1	2	3
1954	1.16	1.25	1.60	1.03	1.23	1.05
1964	1.00	1.20	1.52	0.81	0.57	0.39
1968	0.78	1.00	1.26	0.72	0.61	0.51
1970	0.83	1.55	2.39	0.96	0.96	0.68

* Refers to one-, two-, and three-day ahead forecasts

As further quantification of the relative merit of the forecasting models, the root-mean-squared error was calculated by first squaring the difference between the measured and forecasted results. The squared values were then summed over all the forecasts for a given year and divided by the number of values forecasted. The results for the four years used in the case study are included in Table 2.

COMMENTS ON THE KALMAN FILTER MODEL

Findings on the use of model in the research include:

(i) the model parameters in the MISP algorithm can vary in time (as hydrologic conditions change);

(ii) the model is particularly useful when inputs (e.g. temperature and precipitation) are corrupted by errors;

(iii) there is a 'learning period' where the model forecasts are sensitive to initial parameter estimates;

(iv) there are occasions of filter divergence and instability that warrant further research;

(v) the model runs on an IBM-PC with a math co-processor in 5 to 6 minutes for 120 days of forecasting.

Figure 1

One-day ahead forecast

1968

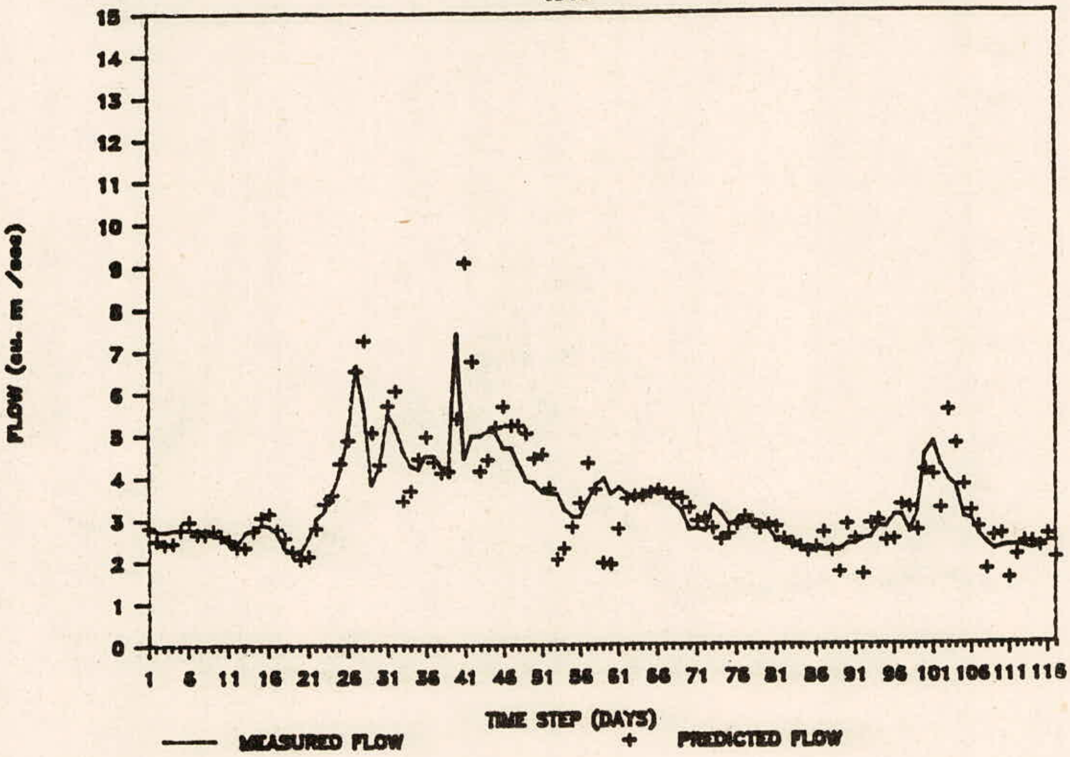


Figure 2

One-day ahead forecast

1970

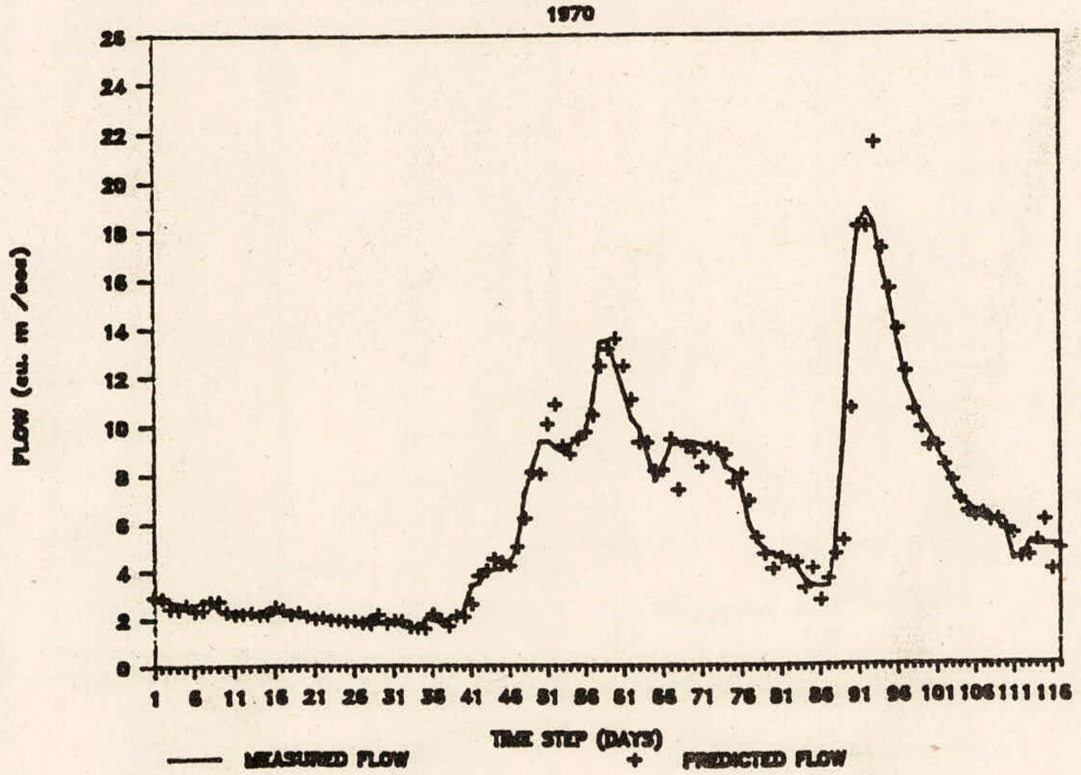


Figure 3

Two-day ahead forecast

1970

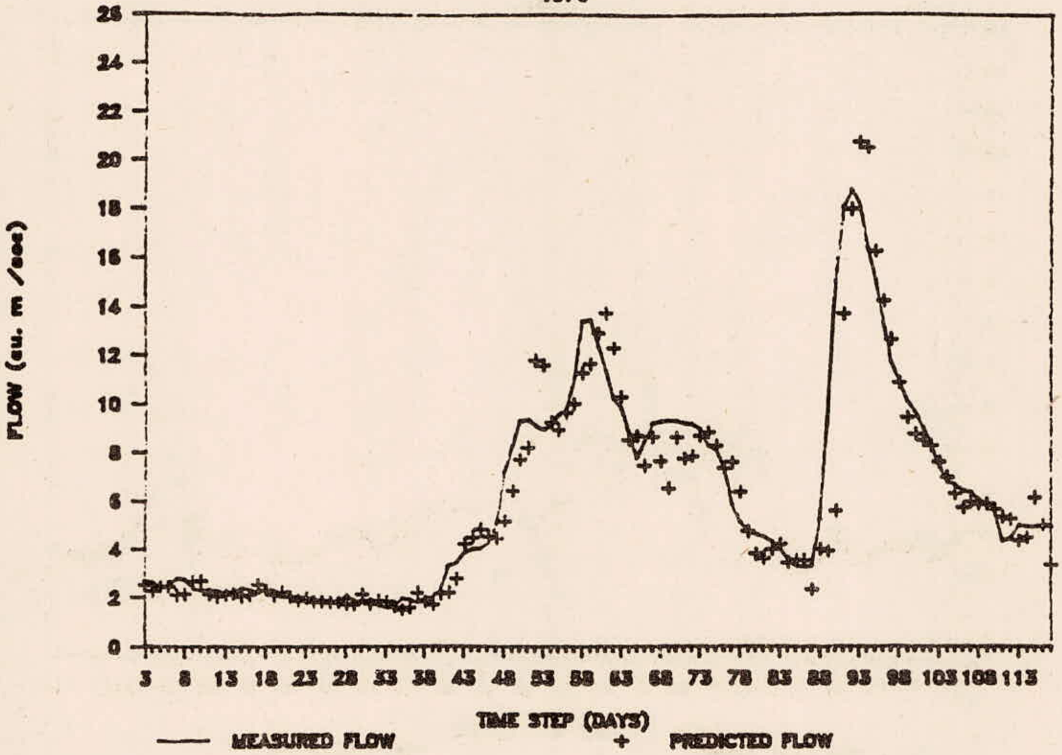
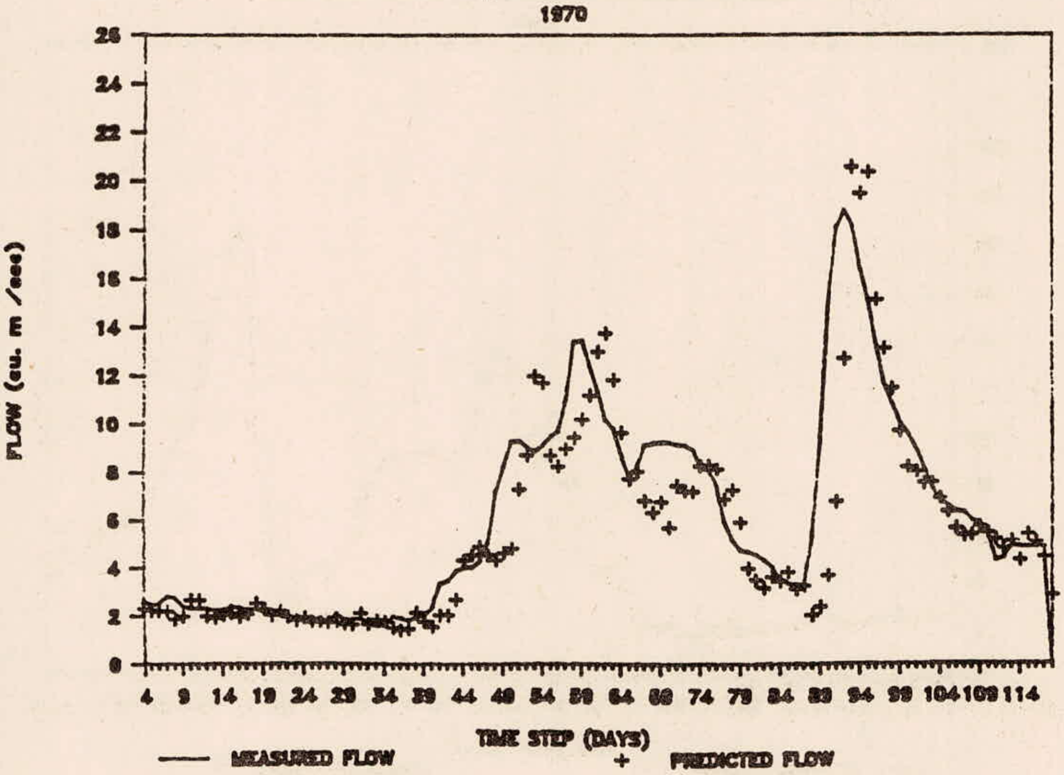


Figure 4

Three-day ahead forecast

1970



CONCLUSIONS

The proposed model provides good one-day forecasts, and fairly good two-day forecasts. The three-day forecasts show a deteriorating accuracy of forecast. The degree to which this decrease in forecast accuracy occurs is a function at least in part, of the Sturgeon basin. The model performance improves as the filter "learns" - good performance at early times depends on good initial parameter estimates.

BIBLIOGRAPHY

1. Aboitiz, M., Labadie, J.W., and Heermann, D.F., "Stochastic Soil Moisture Estimation and Forecasting for Irrigated Fields", Water Resources Research, Vol. 22, No. 2, 180-190, 1986.

2. Burn, D.H., and McBean, E.A., "River Flow Forecasting Model for Sturgeon River", Journal of Hydraulic Engineering, ASCE, Vol. III, No. 2, 316-333, 1985.

3. Burn, D.H., and McBean, E.A., "Accuracy Dieoff of Flow Forecasts Using the MISP Algorithm", presented at the International Association of Hydrological Sciences meeting, VITUKI, Budapest, Hungary, July 2-10, 1986.

4. Todini, E., "Mutually Interactive State-Parameter (MISP) Estimation", Application of Kalman Filter to Hydrology, Hydraulics and Water Resources, C.L. Chiu, Ed., Proceedings of AGU Chapman Conference, University of Pittsburgh, 1978.

5. Todini, E., and O'Connell, P.E., "Basic Methodology: Kalman Filter Estimation Problems", Proceedings Extended Workshop in Real-Time Hydrological Forecasting and Control, 1977.

6. Todini, E., Szollosi-Nagy, A., and Wood, E.F., "Adaptive State Parameter Estimation Algorithms for Real-Time Hydrological Forecasting: A Case Study", presented at the IIASA/WMO Workshop on the Recent Developments in Real-Time Forecasting/Control of Water Resources Systems, October 1976.

7. Wood, E.F., and Szollosi-Nagy, A., "An Adaptive Algorithm for Analyzing Short-term Structural and Parameter Changes in Hydrologic Prediction Models", Water Resources Research, Vol. 14, No. 4, 1978.