International Seminar on Hydrology of Extremes

1-3 December, 1988
Roorkee (India)

A BIVARIATE MODEL FOR REAL TIME FLOOD FORECASTING

P. F. Krstanovic
Louisiana Geological Survey
P.O. Box G, University Station
Baton Rouge, LA 70893, U.S.A.

V. P. Singh
Department of Civil Engineering
Louisiana State University
Baton Rouge, LA 70803, U.S.A.

SYNOPSIS

A bivariate model for real-time flood forecasting is developed from Burg's multivariate channel analysis. The model may be applied to any bivariate hydrologic processes, provided that there exists cross-correlation between processes. The model was tested for short-term flood forecasting for various rainfall-governed events from five climatologically different watersheds. The model is particularly suited for flood forecasting in the data-scarce regions, since it does not require prior parameter estimation.

1.0 INTRODUCTION

A number of multivariate forecasting models have been developed in the last quarter of the century. A good discussion of some of the models is given in proceedings of two recent meetings: Workshop on Recent Developments in Real Time Forecasting/Control of Water Resources Systems [11], and International Workshop on Operational Applications of Mathematical Models (Surface Water) in Developing Countries [5], and also by Salas et al. [8].

The most common forecasting models are time series models including AR, ARMA, ARIMA, ARMAX, CARMA, TFN, state-space. These models are easy to use, but only for certain hydrologic processes and under certain conditions. For complex processes, the parameter estimation becomes more difficult. Rainfall forecasting models (temporal and space-time models) lack the forecasting accuracy and cannot yet be used for long-term forecasting. Their success depends on other technologies such as radar, remote sensing, etc. Both long historical record and technical equipment are needed to provide useful inputs for flood forecasting. The Bayesian forecasting models employ Bayesian theory for forecasting, but have the same problems as the time series models (excessive number of parameters). They are also limited only to short-term daily forecasting, since they exclude higher order dependencies.

The entropy forecasting models are still in the early developmental stage. Currently, there are only three types of models. The first is general, and not specifically developed for hydrological forecasting [9]. The second is based on the entropy-minimax approach and has been applied to long-term annual forecasting of drought using seven stations in northern California [3]. The third employs maximum entropy spectral analysis (MESA), which was tested on Spring Creek in Louisiana [7].

This study develops a bivariate recursive model for real-time flood forecasting for streamflow governed by one or more stochastic mechanisms. The model is tested for short-term forecasting of the rainfall-governed streamflow.

2.0 MATHEMATICAL DEVELOPMENT OF THE MODEL

We study the bivariate hydrologic sequence of rainfall series and its runoff response. The dependence of runoff on rainfall is expressed through cross-correlation function (ccf). From known values of rainfall R(t) and runoff Q(t), we may compute cross-correlation $\rho_{12}(k)$, extend it if necessary, and forecast the next flood value(s). This forecasting is effective for small sampling time intervals (STI) and long data base usually satisfied on large watersheds and for long duration events when numerous sampling intervals are available before the occurrence of the upper rising part of the flood hydrograph. In that case, the ccf need not be extended. However, for very small watersheds and for short duration flash flood events, the data base is small. In that case, an extension of the ccf is essential before the forecasting stage.

2.1 Algorithm

Let T_1 denote the length of the known rainfall and runoff sequence and T the length of the complete record. Then, we compute the ccf matrix and determine the coefficients using the Burg algorithm for the bivariate hydrologic sequence [2]. General formulas for the coefficients are obtained after extending the ccf matrix by one row or using the equation:

$$\begin{bmatrix} \rho(0) & \cdots & \rho(-N) \\ \rho(1) & \cdots & \rho(-N+1) \\ \vdots & & \vdots \\ \rho(N-1) & \cdots & \rho(-1) \\ \rho(N) & \cdots & \rho(0) \end{bmatrix} \begin{bmatrix} I \\ F_1 \\ \vdots \\ F_{N-1} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ B_{N-1} \\ \vdots \\ B_1 \\ I \end{bmatrix} [c_N] = \begin{bmatrix} E_{N-1} \\ 0 \\ \vdots \\ 0 \\ \Delta_N \end{bmatrix} + \begin{bmatrix} \Delta_N^* \\ 0 \\ \vdots \\ 0 \\ E_{N-1}^* \end{bmatrix} [c_N] \qquad (1)$$

where $\rho(k)$ (k = 0, ..., N) are 2x2 ccf matrices, I is 2x2 identity matrix and the superscript * denotes the matrix transpose. All other elements of eq. (1) are 2x2 matrices. F_j and B_j (j = 1, ..., N) are extension matrices essential for forecasting. The solutions of eq. (1) are expressed as matrix equations:

$$\Delta_N = \sum_{n=0}^{N-1} \rho(N-n)\ F_n \tag{2.1}$$

$$\Delta_N^* = \sum_{n=0}^{N-1} \rho(N-n)\ B_n \tag{2.2}$$

$$c_N = -\ (E_{N-1}^*)^{-1}\ \Delta_N \tag{2.3}$$

$$E_N^* = E_{N-1}\ (I - c_N^*\ c_N) \tag{2.4}$$

$$F_N = \begin{bmatrix} F_{N-1} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ B_{N-1}^* \end{bmatrix} c_N \tag{2.5}$$

2.2 Forecasting

In each forecasting step, the ccf matrix is recomputed and extended by one row and column. Then the system of eqs. (2.1) - (2.5) is solved and F_N, B_N matrices are determined. Forecasting equations were developed by Krstanovic [6] who used F_n (n = 0, ..., N) matrices as weights for each rainfall-runoff pair X_t = (R_t, Q_t) such that

$$\sum_{n=0}^{N} F_n\ X_{T-n} = 0 \tag{3}$$

From eq. (3), 1-step ahead forecasting is

$$[R_{T+1}, Q_{T+1}] = -\sum_{n=1}^{N} F_n\ X_{T+1-n} \tag{4.1}$$

and k-step ahead forecasting is

$$[R_{T+k}, Q_{T+k}] = -\sum_{n=1}^{N} F_n\ X_{T+k-n} \tag{4.2}$$

Since the solution domain involves the bivariate stationary sequence of Gaussian nature, both rainfall and runoff sequences must be Gaussian. From the nature of the multivariate normal distribution [6], all ccf matrices must be positive definite. Furthermore, all E_n and ccf matrices belong to the class of non-negative definite matrices. Under these conditions, eq. (1) always has solution, and the extension of ccf matrices satisfies maximum entropy conditions [2].

3.0 VERIFICATION OF THE MODEL

Necessary steps for real-time forecasting are: (a) Determine the nature of the hydrologic event: long event, associated with a long record, or short flash flood. (b) Transform and normalize the data (optionally). (c) For longer STI's and short duration, high intensity rainfall, create shorter STI's assuming some rainfall distribution within them, interpolate rainfall and runoff values, and then proceed with multistep ahead forecasting with $k > STI$. (d) Evaluate the performance of 1-step ahead and multistep ahead forecasting by bivariate model.

The model was tested on five different hydrological records, as shown in Table 1. These records belong to three different climatological groups: (a) watersheds from southeastern USA (Goodwin Creek, Hillsborough River), (b) watershed from the Indian monsoon climate (Krishna Wuna), and (c) watersheds from central Italy (Tevere Torgiano and Tevere St. Lucia). The available STI's ranged from 2 minutes (smallest watershed - Goodwin Creek) to 6 minutes and 1 hour (medium and large watersheds). All STI's satisfied the forecasting rule specified by Gosain [4], STI in the range 1/3 - 1/2 watershed lag time or 1/6 - 1/3 watershed concentration time. Thus, we can issue reliable forecasts from 1/2 hour (Goodwin Creek watershed) to maximum several hours (other watersheds). Runoff was measured at the watershed mouth, except for the Goodwin Creek watershed. Rainfall was recorded simultaneously by the rainfall stations located throughout the watershed, and then weighted by the Thiessen method (all watersheds).

Runoff events on the Goodwin Creek watershed were examined for different seasons (spring and winter) and for different STI's. For smaller STI's and forecast lead times (FLT's), relative errors in predicting time to peak were greater than for larger STI's and longer FLT's. In forecasting the flood peak, the magnitude of STI's was not proportional with forecast errors. A typical example of forecasting the multiple flood event for STI = 10 min and FLT = 30 min is given in Figure 1a. This spring event yielded three flood hydrographs, all with sharp rising limbs with no intense rainfall before the first one. Thus, forecasting of the first flood hydrograph was similar to ephemeral flood forecasting. The Hillsborough River in central Florida is perennial, but mostly governed by rainfall. The presented forecast corresponds to a case with very short STI (= 2 hours) and multistep ahead FLT (= 1 day) as shown in Figure 2. The forecasts of the bivariate model were then better than simple forecast with STI = FLT = 1 day.

For watersheds in central Italy, we found that for longer STI's and FLT's, the relative forecast errors rose correspondingly. This was expected due to large watershed areas. With STI = 6 minutes, the maximum reliable forecast was for FLT = 1 hour. More reliable forecasts were not possible since significant

Table 1. The watersheds used in testing of the bivariate forecasting model.

No.	Watershed	Streamgage Location	Drainage Area (km^2)	Sampling Time Interval (STI)	Record	Source
1	Goodwin Creek watershed, Mississippi, USA	No. 4 at Goodvin Creek watershed, Panola County Mississippi	22	2 minutes	Continuous rainfall-runoff observations (1981-1983)	USDA-ARS Oxford, Mississippi
2	Krishna Wuna watershed, India	Bridge No. 501/807	824	1 hour	Events	Indian Institute of Technology, Delhi, India
3	Watershed Tevere, subwatershed Torgiano, Tiber basin, Italy	P.N. Torgiano	4147	6 minutes	Events	National Research Council, Institute for Hydrological Protection Research in Central Italy, Perugia, Italy
4	Watershed Tevere, subwatershed St. Lucia, Tiber basin, Italy	St. Lucia	934	6 minutes	Events	
5	Hillsborough River, Florida	Zephyrhills, Florida	---	1 day	Continuous record	Cooper and Wood (1982)

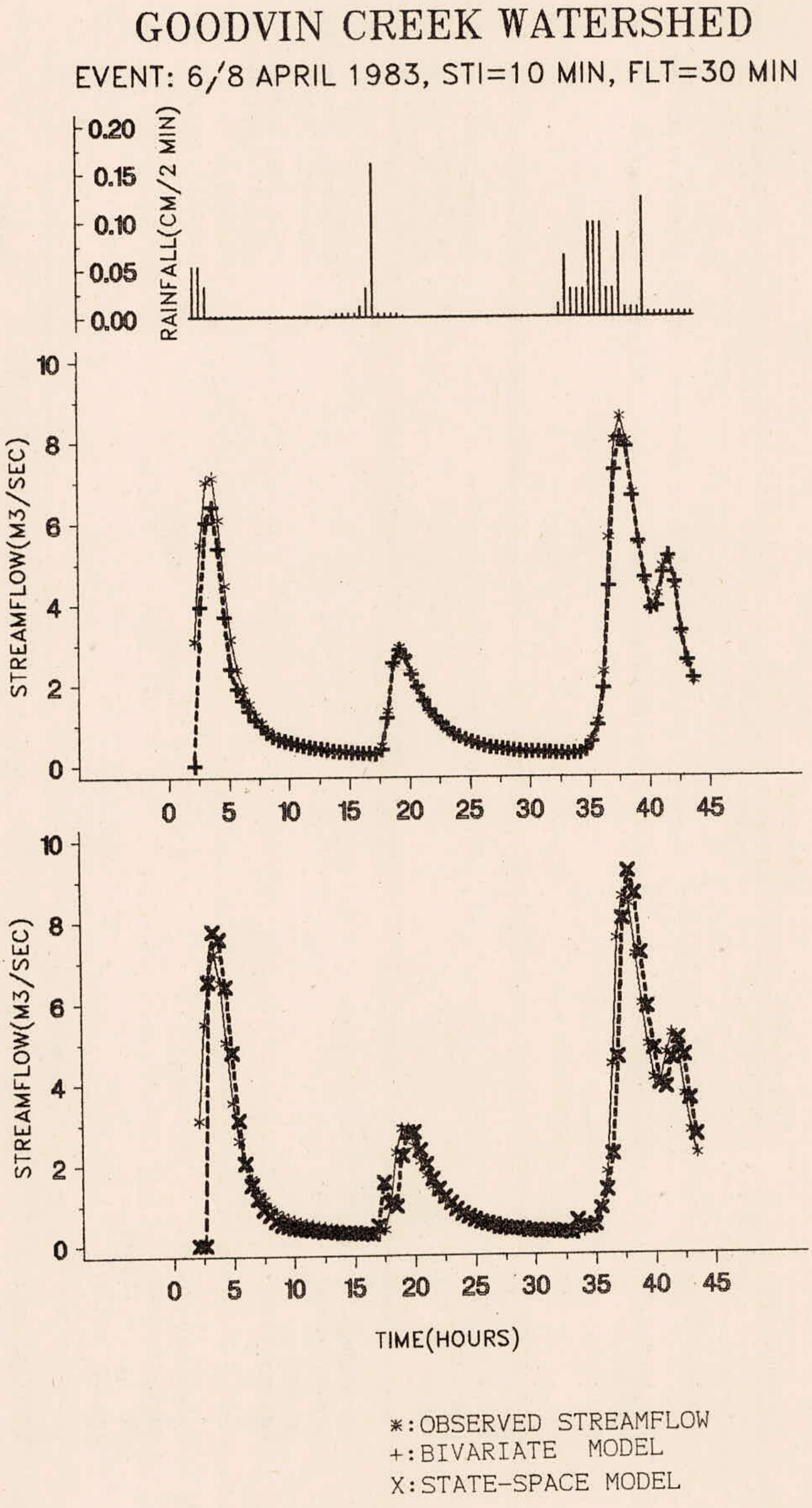

Figure 1. Goodwin Creek watershed forecasts: a. Bivariate model forecasts, and b. state-space model forecasts.

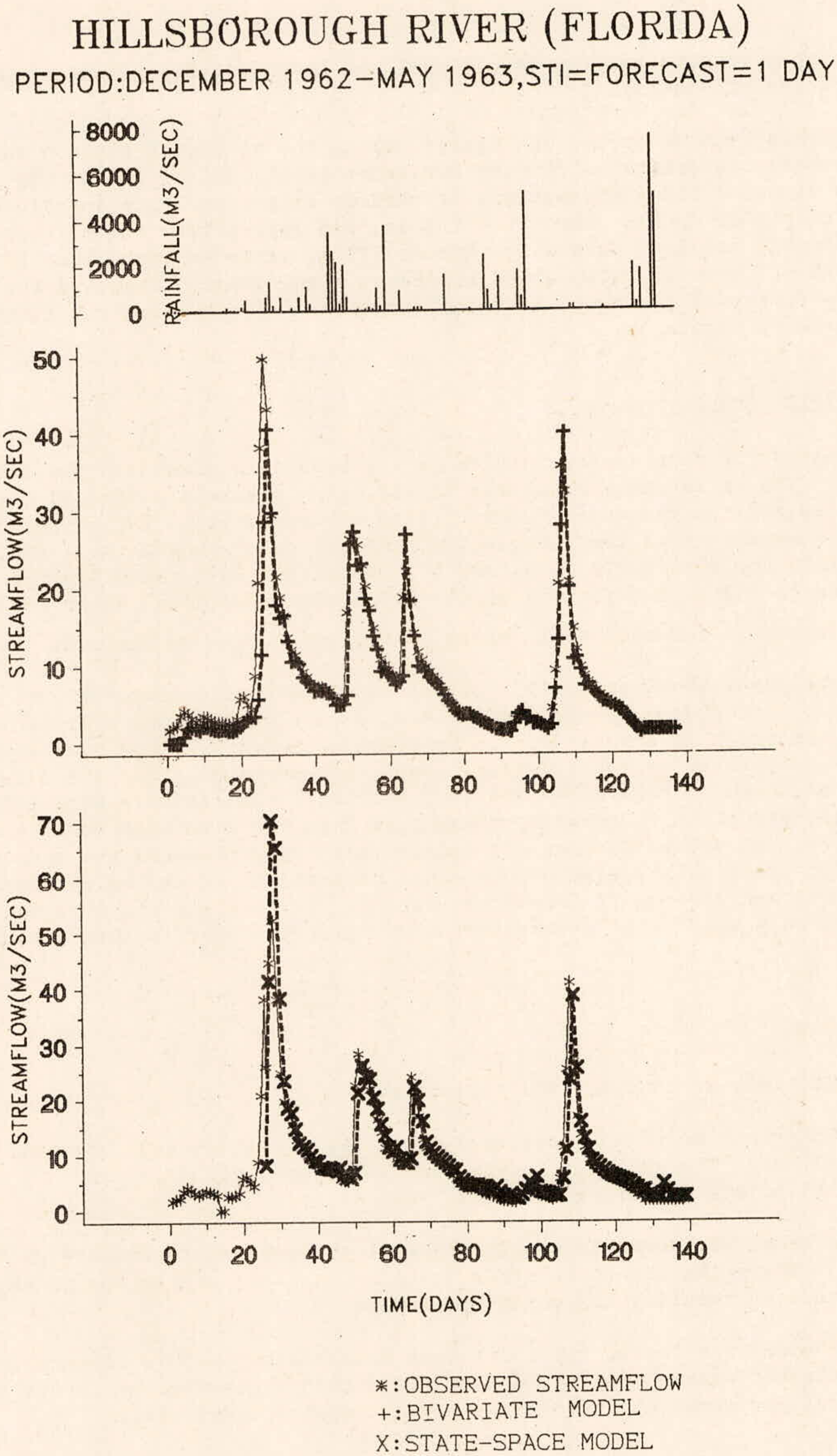

Figure 2. Hillsborough River forecasts: a. Bivariate model forecasts, and b. state-space model forecasts.

rainfall stopped long before the runoff peak was reached. This limited the dimension of the ccf matrix.

The Krishna Wuna watershed was tested during the southwest monsoon season (June-September), associated with very extensive precipitation (80 to 120 cm). This season produced flood hydrographs of various shapes and long durations (usually a couple of days). The STI = 1 hour, and appropriate FLT = 1 hour. The record apparently required data with shorter STI's, since we found that STI = 6 min and FLT = 1 hour (10 step ahead forecast) significantly improved the results in the early forecasting stage. A typical example of forecasting the flash flood event is given in Figure 3a.

4.0 COMPARATIVE EVALUATION

We compared the forecasting results of the bivariate model and the state-space model. The state-space model was fitted using Akaike's canonical correlation technique and the recursive Kalman filter algorithm [1]. Two criteria were employed for comparison of the forecasting models: (a) graphical or visual fit of the forecast to the real-world data, and (b) numerical [10]: mean squared error of the forecasts (MSE), coefficient of variation of the residual error (v_1), ratio of relative error to the mean (v_2), ratio of absolute error to the mean (v_3).

The state-space model was unable to initiate forecast during the early forecasting stage. For sudden flash floods, it also gave delayed and overpredicted flood peaks, as shown in Figure 3b. In forecasting multiple flood hydrographs it gave results comparable to the bivariate model, except usually for the first hydrograph, as shown in Figures 1b and 2b. Numerical coefficients were comparable for all cases, except in forecasting flood peak when the bivariate model was usually better. We emphasize that two models were compared under the same conditions, with no prior calibration of the model parameters. Thus, no information about previous rainfall-runoff events was needed. To initiate the forecasting algorithm, we only need first several measured rainfall-runoff values during the test event.

5.0 CONCLUSIONS

From this study we conclude the following:

(a) A bivariate model was developed from Burg's multivariate channel analysis. The explicit forecasting equations, valid for any multivariate hydrologic process, were derived.

(b) The bivariate model is an extension of the univariate streamflow model [6], where the autocorrelation function (acf) of streamflow is replaced by ccf of rainfall and runoff processes.

(c) The model was tested for short term forecasting on five climatologically different watersheds for various flood events governed by rainfall. The model performed satisfactorily for all events considered.

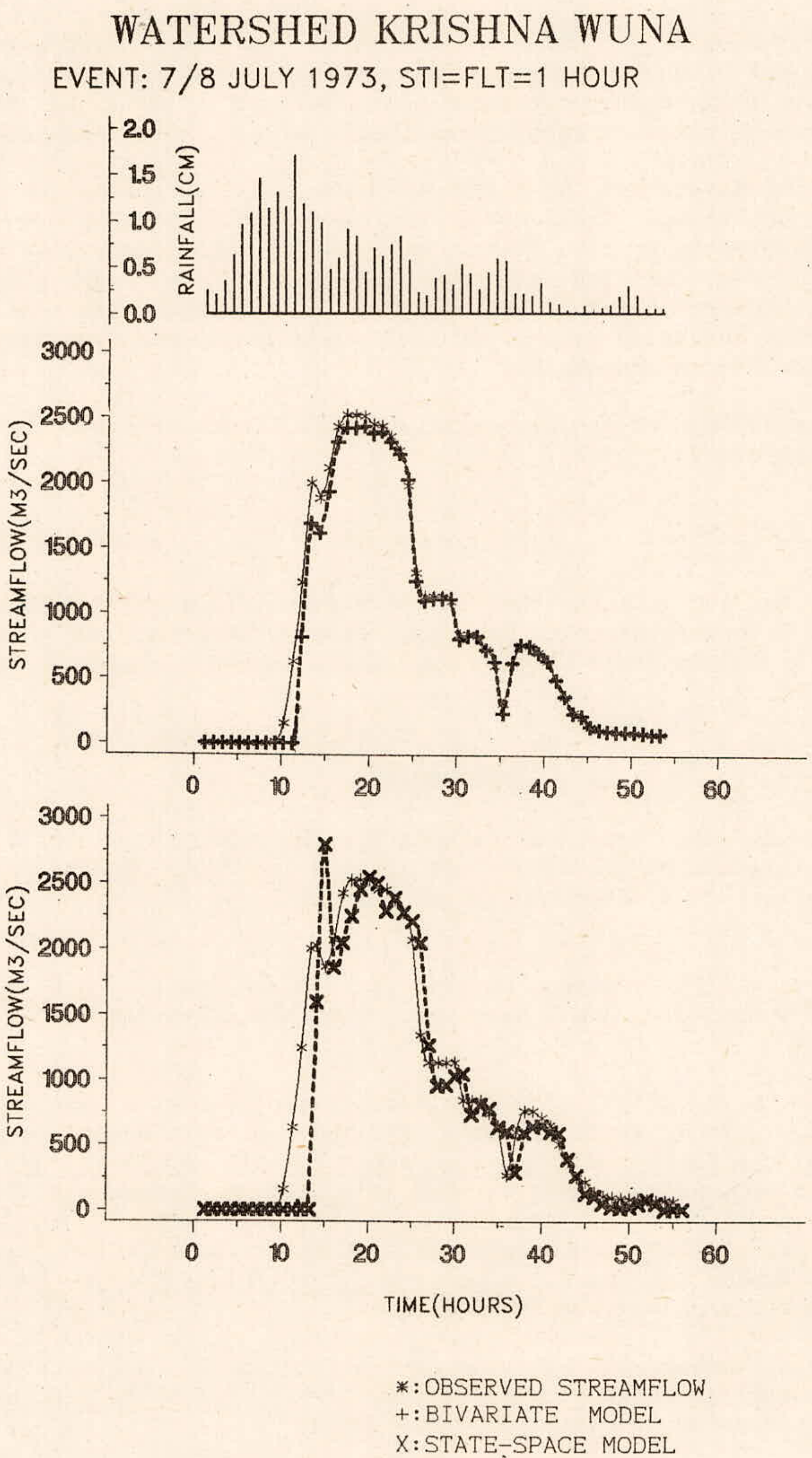

Figure 3. Krishna Wuna watershed forecasts: a. Bivariate model forecasts, and b. state-space model forecasts.

(d) For forecasting, no prior calibration of the model parameters is needed. Thus, the bivariate model may be particularly suited for flood forecasting in data-scarce regions where it would be difficult to calibrate parameters of other sophisticated models (i.e., state-space model).

(e) The bivariate model and the state-space model were compared under the same conditions. These models were comparable for single events on large watersheds (i.e., Tevere watersheds). For sudden, very intense flood flows, the bivariate model accommodated the rising part of the flood hydrograph much better than the state-space model. It was superior both with respect to flood warning (issuing forecasts on time) and the hydrograph shapes.

(f) A disadvantage of the bivariate model is slight underprediction of the flood hydrograph peaks.

ACKNOWLEDGEMENTS

This study was supported by the U.S. Department of Interior through the Louisiana Water Resources Research Institute and by the Department of Civil Engineering at Louisiana State University. Their support is gratefully acknowledged.

REFERENCES

1. Akaike, H., (1976), 'Canonical correlation analysis of time series and the use of an information criterion,' in Advances and Case Studies in System Identification, by R. Mehra and D. G. Lainiotis, pp. 27-96, Academic Press, New York.

2. Burg, J. P., (1975), 'Maximum entropy spectral analysis,' Ph.D. Thesis, Stanford University, 123 p., Palo Alto, California, University Microfilms, 75-25, 499.

3. Christensen, R. A., (1981), 'An exploratory application of entropy minimax to weather prediction: estimating the likelihood of multi-year droughts in California,' in Entropy Minimax Sourcebook Vol. IV: Applications, by R. A. Christensen, pp. 495-544, Entropy Limited, Lincoln, Massachusetts.

4. Gosain, A. K., (1984), 'Intercomparison of real-time highflow forecasting models for Yamina catchment,' unpublished Ph.D. Dissertation, Indian Institute of Technology, Delhi, India.

5. International Workshop on Operational Applications of Mathematical Models (surface water) in Developing Countries, (1985), 'IWOAM Proceedings,' Vol. I, Indian Institute of Technology, Delhi, India.

6. Krstanovic, P. F., (1988), 'Application of entropy theory to multivariate hydrologic analysis,' Ph.D. Dissertation, Vols. I and II, p. 558, Louisiana State University, Baton Rouge, Louisiana, U.S.A.

7. Krstanovic, P. F. and V. P. Singh, (1987), 'A multivariate stochastic flood analysis using entropy.' in Hydrologic frequency modeling by V. P. Singh, 515-539, D. Reidel Publishing Co., Dordrecht, Holland.

8. Salas, J. D., G. Q. Tabios III and P. Bartolini, (1985), 'Approaches to multivariate modeling of water resources time series,' Water Resources Bulletin, Vol. 21, No. 4, pp. 683-708.

9. Souza, R. C., (1978), 'A Bayesian entropy approach to forecasting,' Ph.D. Thesis, University of Warwick, Coventry, U.K.

10. World Meteorological Organization (WMO), (1975), 'Intercomparison on conceptual models used in operational hydrologic forecasting,' Operational Hydrology Report No. 7, Secretariat of WMO, Geneve, Switzerland.

11. Wood, E. F., editor, (1980), 'Real-time forecasting/control of water resource systems,' Pergamon Press, Oxford, U.K.