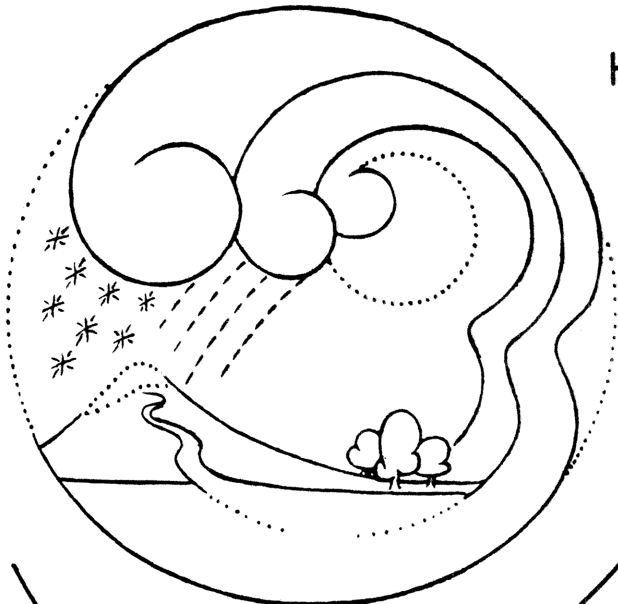
HYDRO TECHNICAL NOTE - 1



NOVEMBER 1983

HYDROLOGIC

RESEARCH

LABORATORY

STOCHASTIC DECOMPOSITION SCHEMES
FOR THE REAL-TIME FORECASTING OF
RIVER FLOWS IN A LARGE RIVER SYSTEM

KONSTANTINE P. GEORGAKAKOS

OFFICE
OF
HYDROLOGY

NATIONAL
WEATHER
SERVICE

NOAA

HYDROLOGIC RESEARCH LABORATORY

OFFICE OF HYDROLOGY

NATIONAL WEATHER SERVICE, NOAA

8060 13TH STREET

SILVER SPRING, MD 20910

HYDROLOGIC RESEARCH LABORATORY
NATIONAL WEATHER SERVICE, NOAA
HRL TECHNICAL NOTE

STOCHASTIC DECOMPOSITION SCHEMES FOR THE REAL-TIME FORECASTING OF RIVER FLOWS IN A LARGE RIVER SYSTEM

by

KONSTANTINE P. GEORGAKAKOS

National Research Council
National Oceanic and Atmospheric Administration
Research Associate

CONTENTS

FIGURES

TABLE

STOCHASTIC DECOMPOSITION SCHEMES FOR THE REAL-TIME FORECASTING OF RIVER FLOWS IN A LARGE RIVER SYSTEM

Konstantine P. Georgakakos
National Research Council
National Oceanic and Atmospheric Administation
Research Associate

1. INTRODUCTION

In recent years, modern estimation theory techniques have been successfully utilized in the real-time forecasting of river flows for headwater basins (no upstream inflows). Second moment estimators of the Kalman type (Gelb, 1974) have been used together with conceptual models that simulate catchment processes (Kitanidis and Bras, 1980a, 1980b) and precipitation-catchment processes (Georgakakos and Bras, 1982).

Straightforward application of the above mentioned stochastic models to large river systems with several tributary basins creates serious computational problems. In such a case, one would formulate a composite state vector whose components are the states of the precipitation, soil, and channel processes in each tributary basin. (See Appendix C in Georgakakos and Bras, 1982, for the relevant formulation.) However, application of estimation theory techniques that treat the system as a whole can be infeasible when the number of tributary basins grows large. This is primarily due to the number of computations related to the propagation and updating of the system covariance matrix, which provides a measure of the error in the state estimates in real time. If each tributary basin requires a state vector of order n, N basins require a composite state vector of order nxN. The covariance matrix is then a (nxN) by (nxN) square matrix, while a filtering procedure of the Kalman type requires computations that increase as $(nxN)^3$. With n of the order of 10, a value of N equal to 10 requires computations of the order of 10^6, which may be impracticable especially when microprocessors are used or when very short forecast lead times are required.

Typically, five to ten tributary basins have been aggregated to form a forecast group during preliminary testing of the National Weather Service River Forecast System (NWSRFS) at the Tulsa and Minneapolis River Forecast Centers (Smith, 1983). Work on the Potomac River has segmented that river into 24 tributary basins (Smith, Sheer, and Schaake, 1982).

It is the purpose of this note to indicate possible solutions to the presented problem that are computationally efficient and preserve the useful properties of the modern estimation theory techniques.

2. STOCHASTIC DECOMPOSITION

It is rather clear that in the case of large river networks there is a trade-off between the computational efficiency of the algorithms used and the accuracy of the forecasts. Therefore, to gain efficiency one has to simplify the models used. Since the cause of the computational problem is the set of computations related to the covariance matrix, it is only natural to concentrate on ways to simplify those.

There appear to be three main classes of simplification procedures. (See Georgakakos, Restrepo-Posada, and Bras, 1980, for a detailed review.)

1) Avoiding the covariance computations in real time as is the case with prespecified filter gains.

2) Filtering only part of the state vector. (In general, the states comprising the filtered part of the state vector are different for different time steps.)

3) Partitioning the state covariance matrix by omitting some of the state cross-correlations, thus reducing the original high-dimensional problem to several ones of low dimensionality.

In all the above simplification procedures, the state mean propagation is done in the same way as for the composite system. It is the set of covariance computations that is simplified.

Due to the uncertainty of real world data and the varying conditions for each river basin, one cannot assess a priori which strategy is best. In fact, in cases where one can afford the cost of comparison, one should compare each simplification procedure to the exact solution in terms of accuracy. Then choose the one that, while meeting the cost restrictions, has the highest accuracy. However, some general comments can be made a priori and several problem (research) areas can be identified for each class of simplification.

2.1 A Priori Computed Filter Gain

A priori specification of the filter gain matrix offers considerable real time operational computational savings (perhaps the most compared to the other two classes). However, this line of approach has several problems associated with it. The main one concerns how to choose the function of time that will represent the filter gain. Whatever method is chosen, one cannot avoid running the complex system with the filter in order to determine the gain as a function of time. This can be quite expensive, if possible at all.

A serious problem, characteristic of this class of simplifications in river flow forecasting, is the fact that the observations of the input variables and of the output variables of the system have errors that are time varying and dependent on the magnitude of the observation. Thus, discharge measurements tend to be more accurate in cases of low base-flow activity. This, and the fact that the gain is a function of those errors, produce erratic time-traces of filter gain which are difficult to approximate by smooth functions of time. Figures 1-10 present time-traces of the gains for 6-hr time steps corresponding to the ten states of the Georgakakos and Bras (1982) model for the mean areal precipitation output (thin dashed line) and for the outflow discharge (thick solid line). The gain time-traces are for the month of July, 1959, with 6-hr data from the Bird Creek

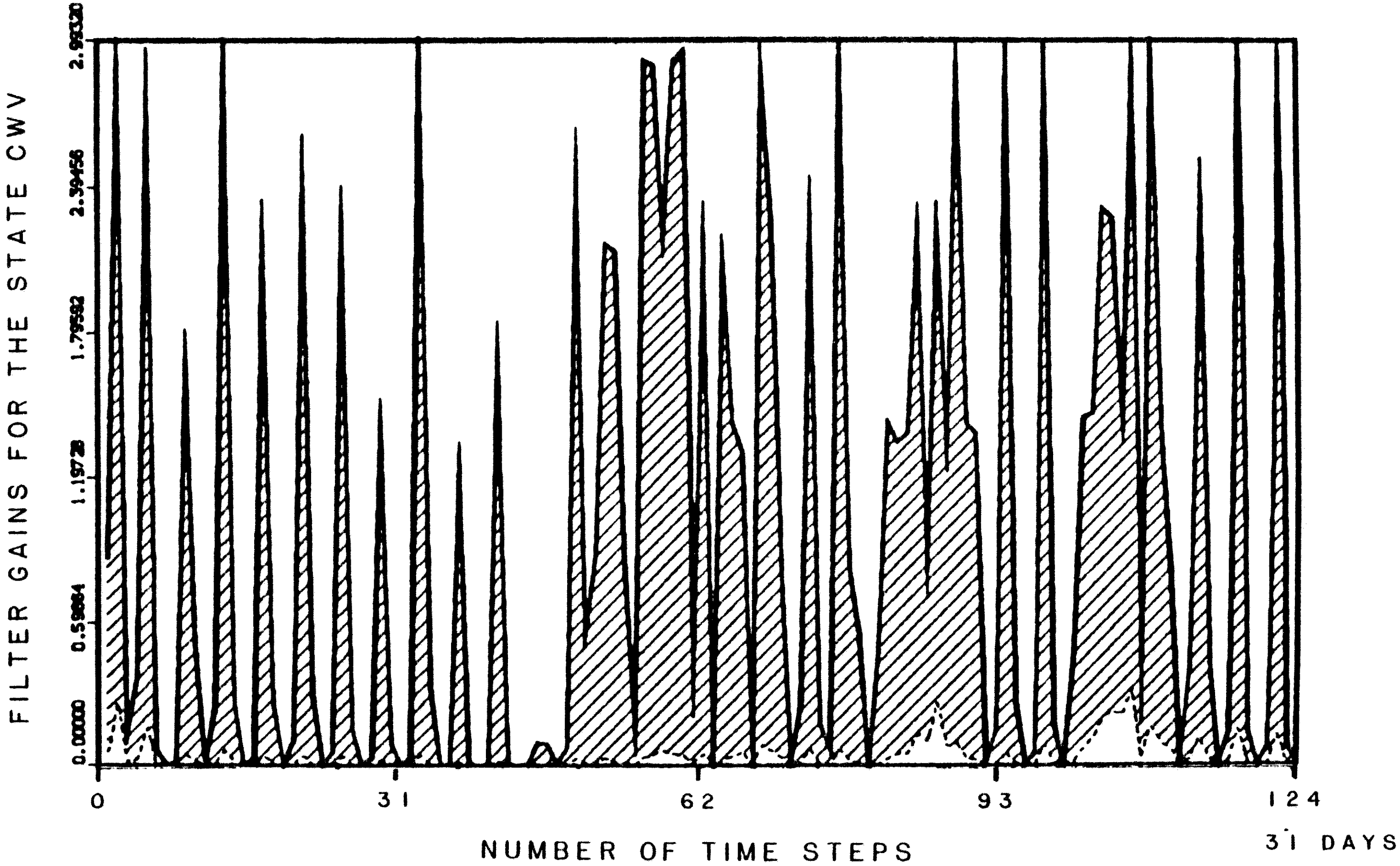


Figure 1. Gains for the precipitation model state for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

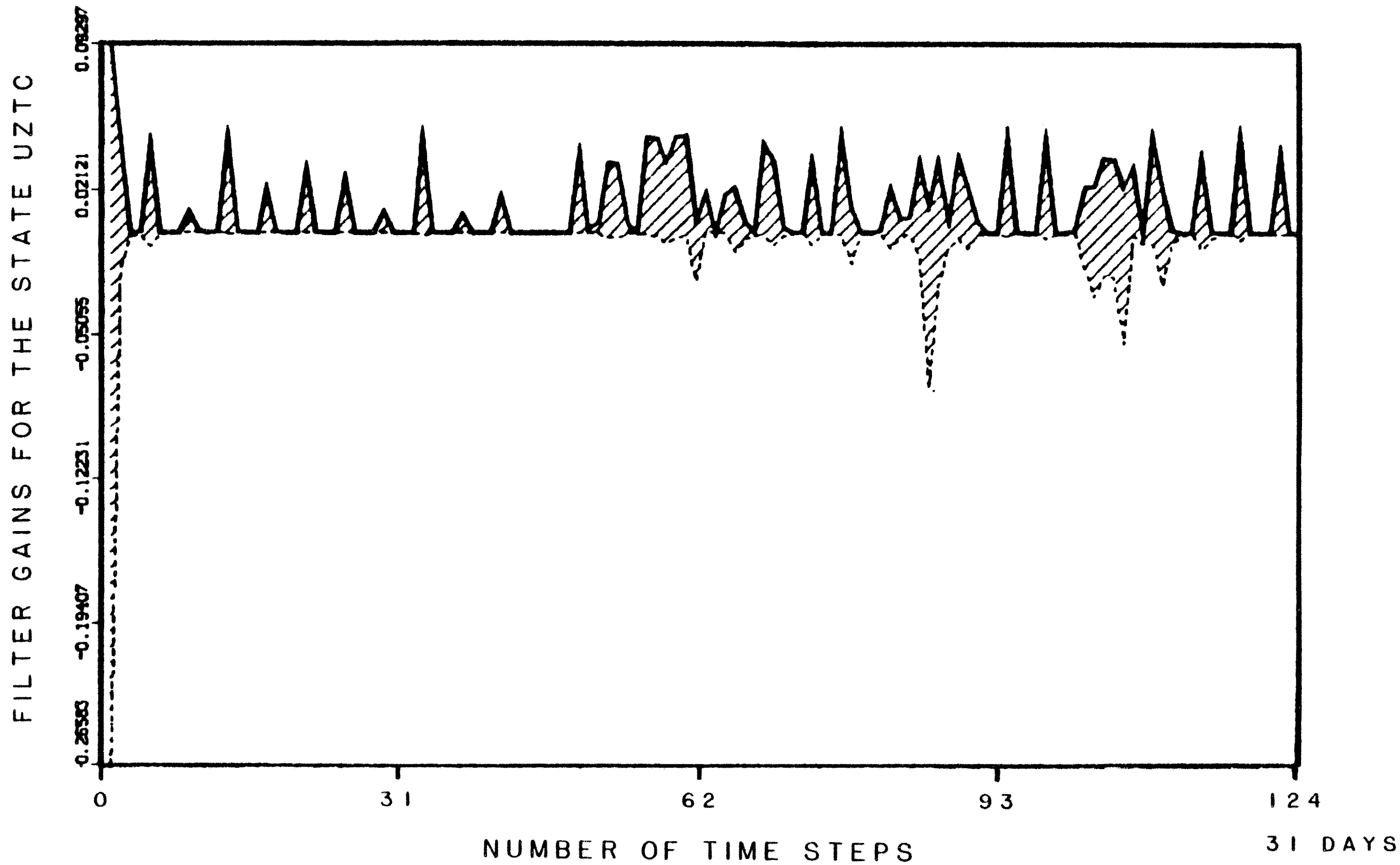


Figure 2. Gains for the upper zone tension water element of the NWSRFS for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

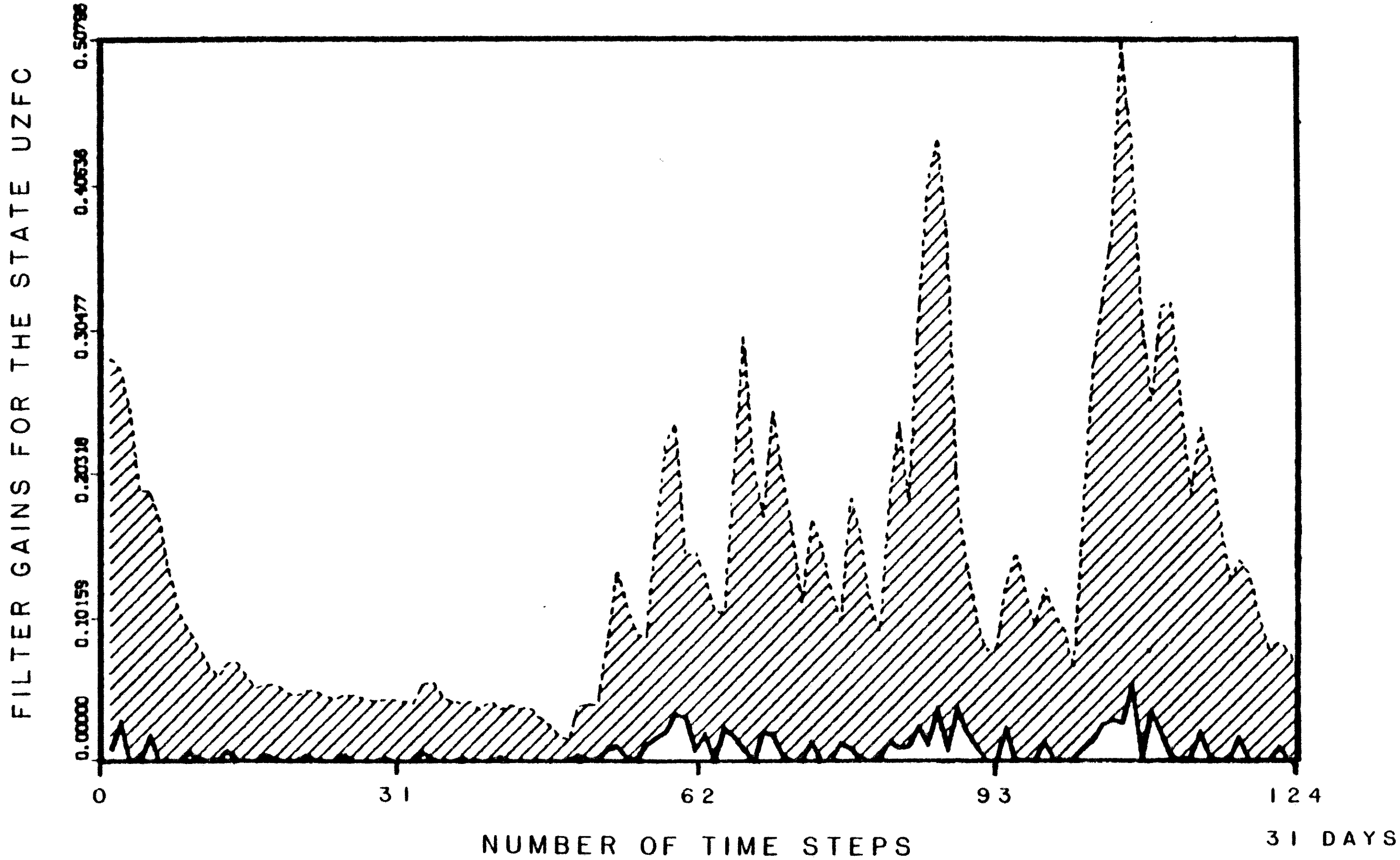


Figure 3. Gains for the upper zone free water element of the NWSRFS for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

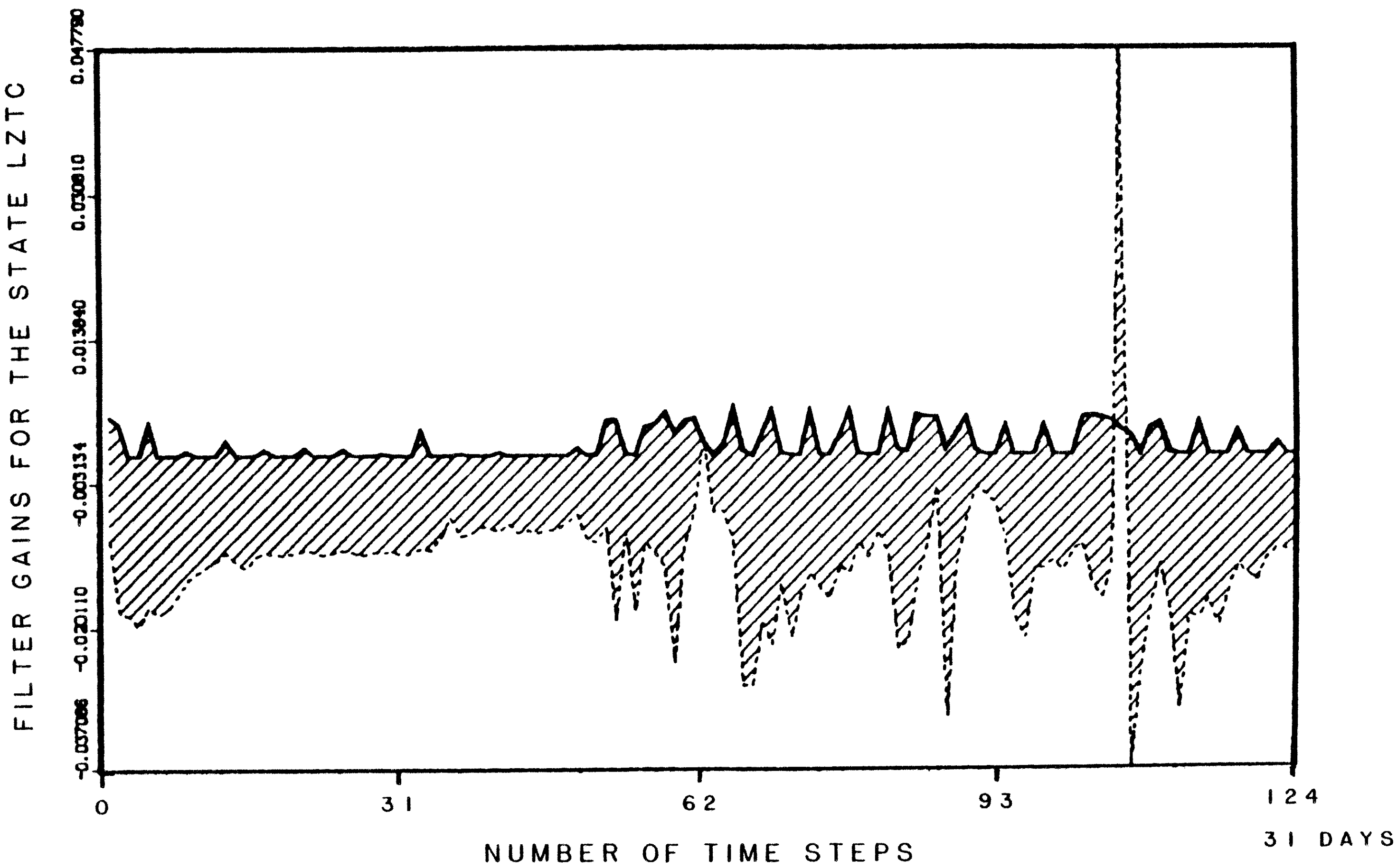


Figure 4. Gains for the lower zone tension water element of the NWSRFS for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

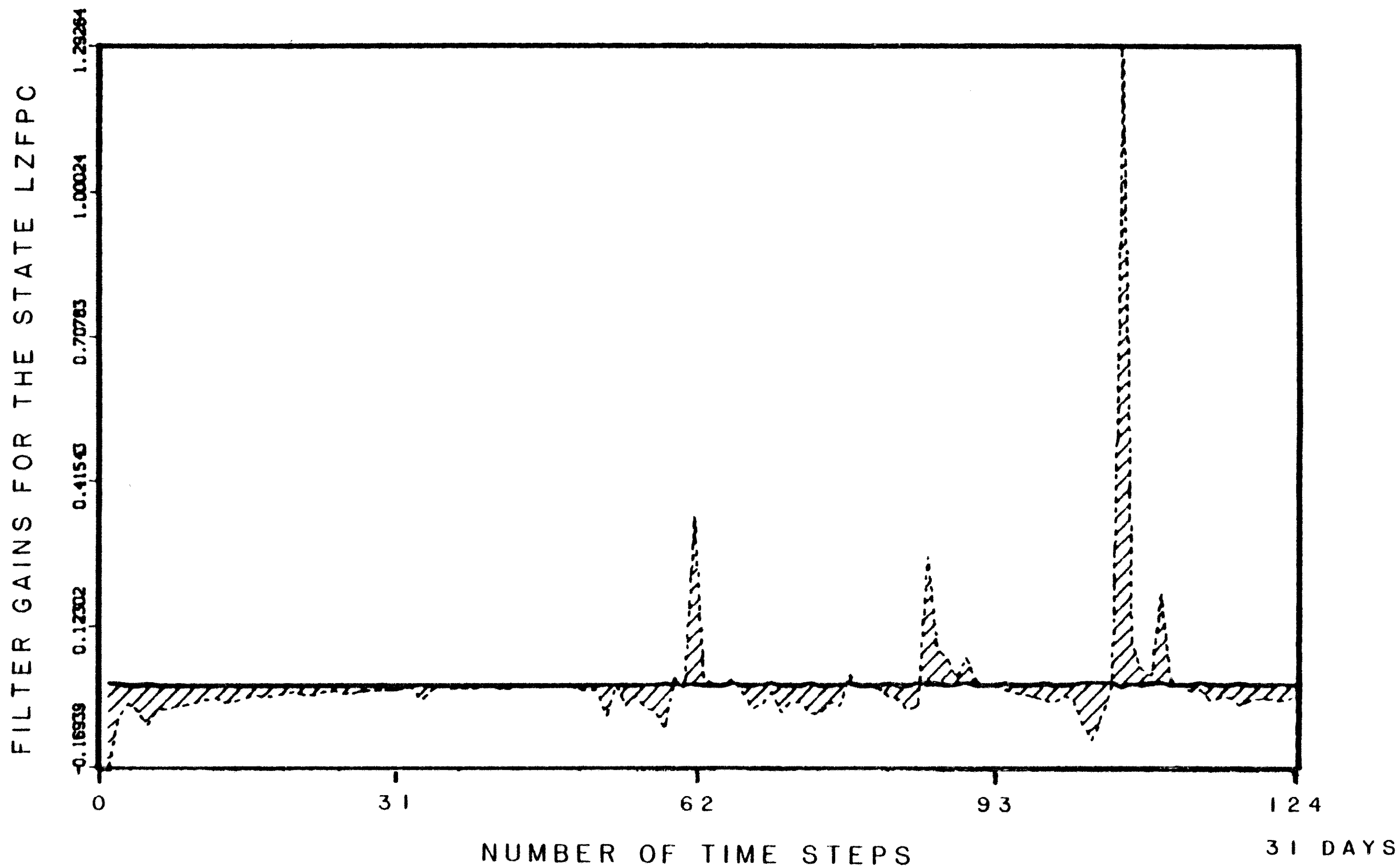


Figure 5. Gains for the lower zone free primary water element of the NWSRFS for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

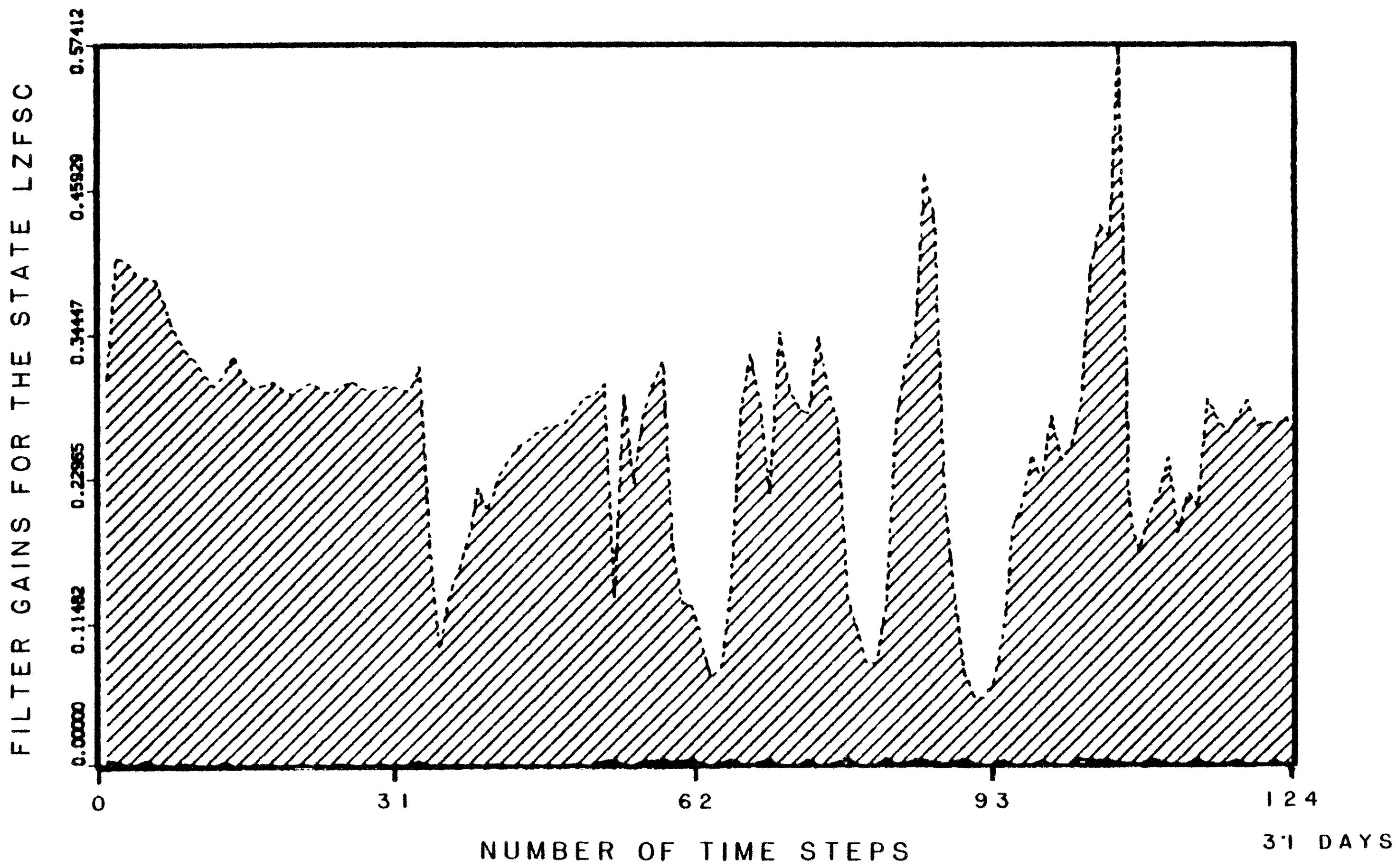


Figure 6. Gains for the lower zone free supplementary water element of the NWSRFS for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

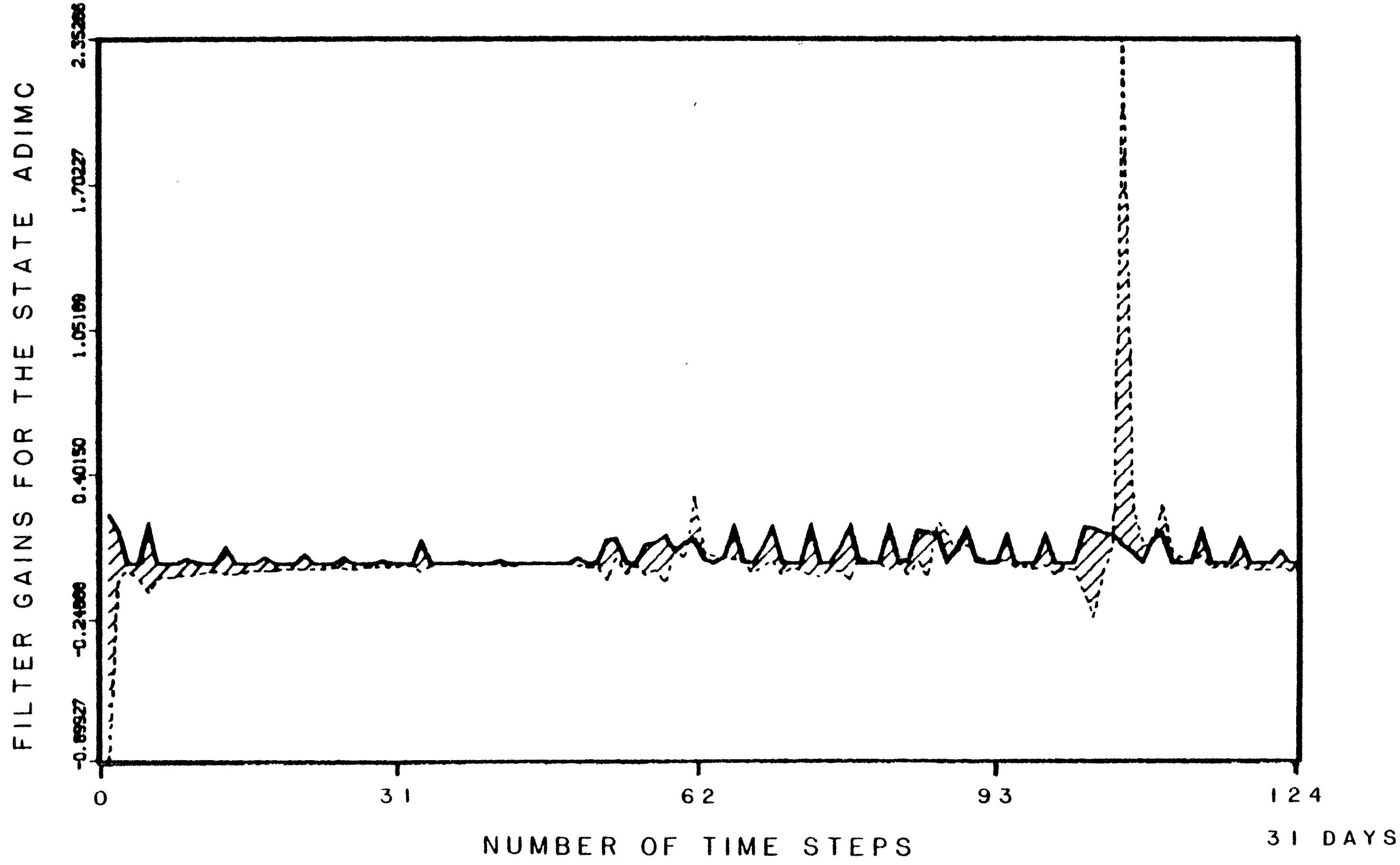


Figure 7. Gains for the additional impervious water element of the NWSRFS for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

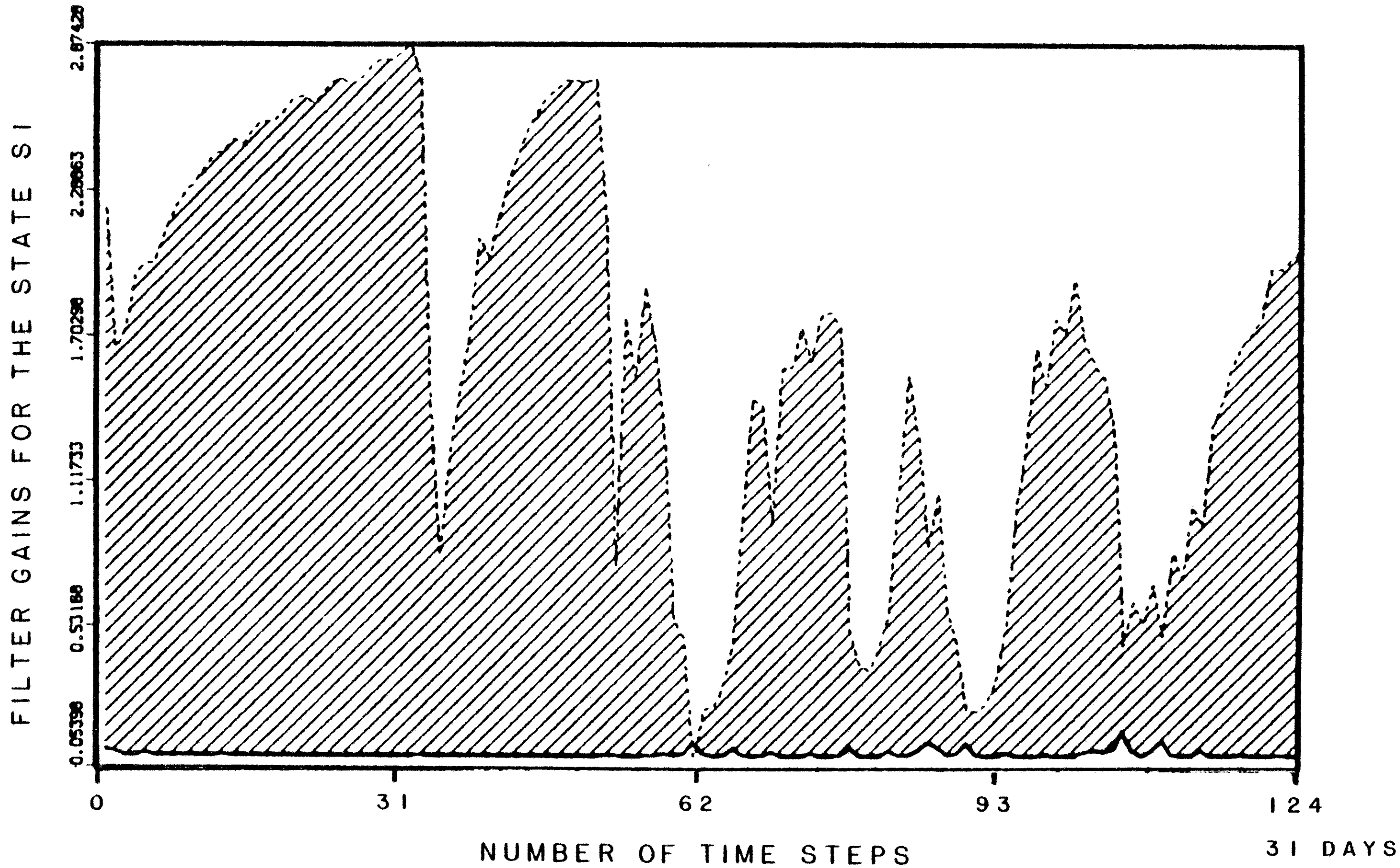


Figure 8. Gains for the first (upstream) channel state for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

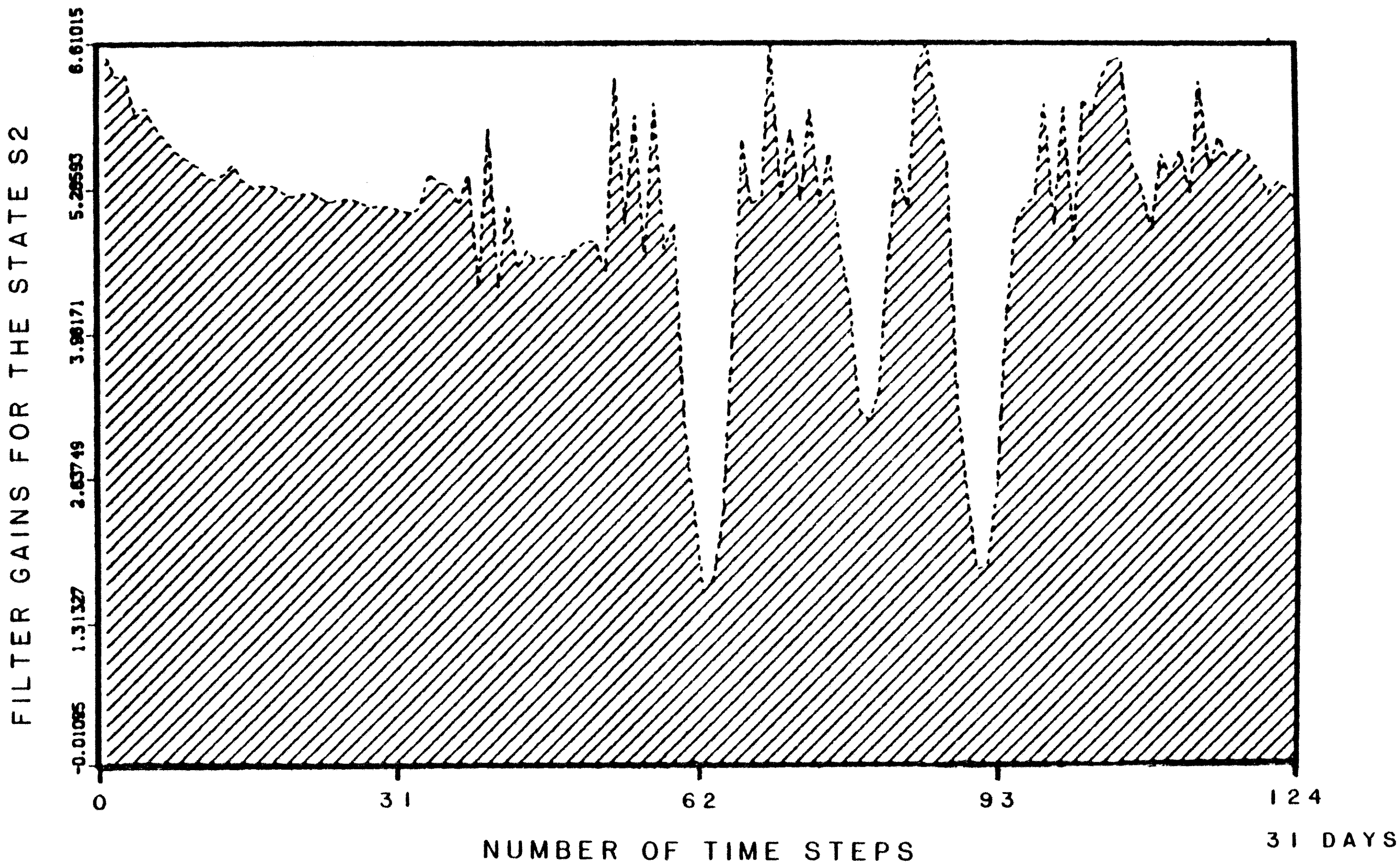


Figure 9. Gains for the second channel state for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

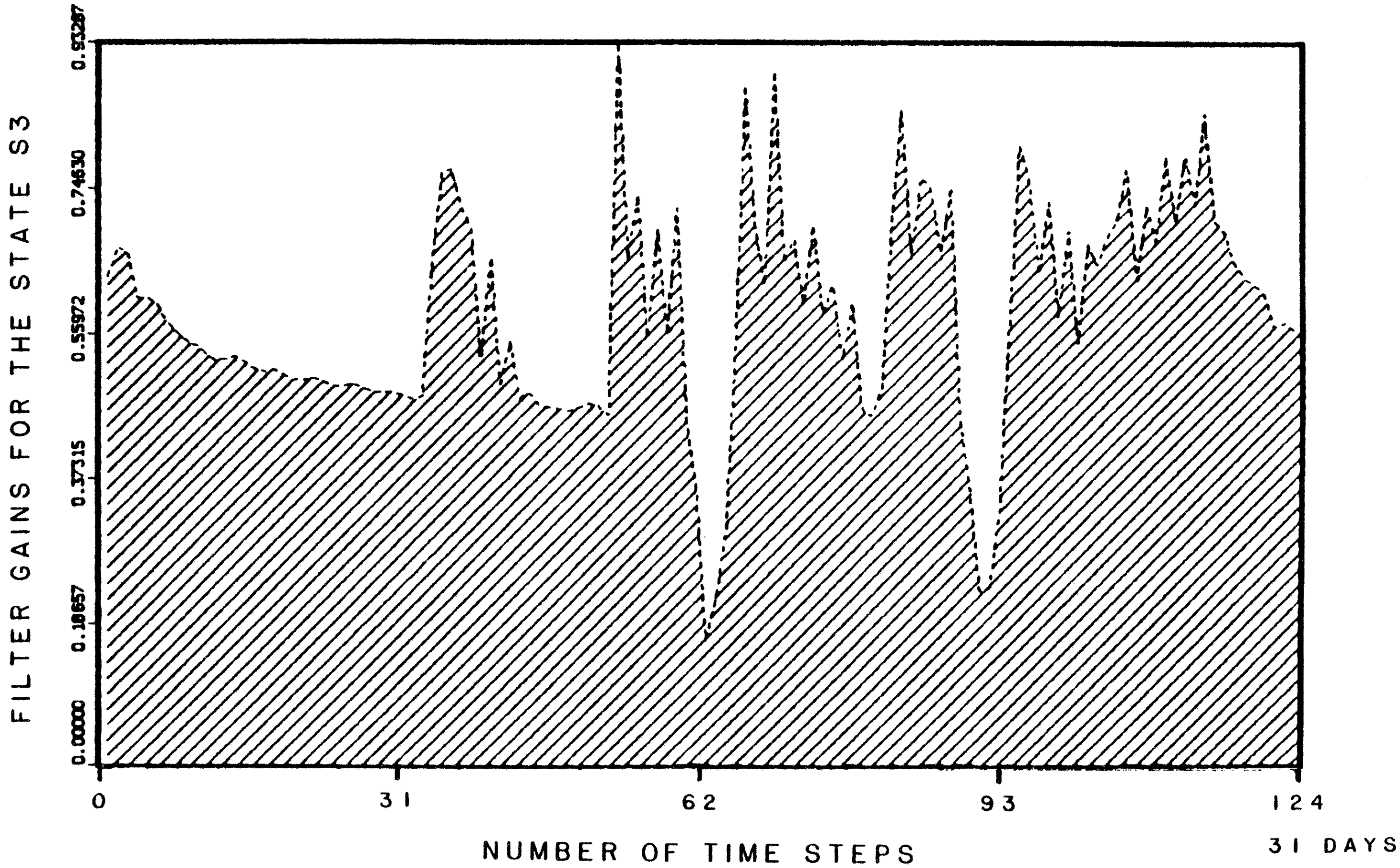


Figure 10. Gains for the third channel state for six-hour time steps. The thick solid line shows mean areal precipitation output, and the thin dashed line shows outflow discharge.

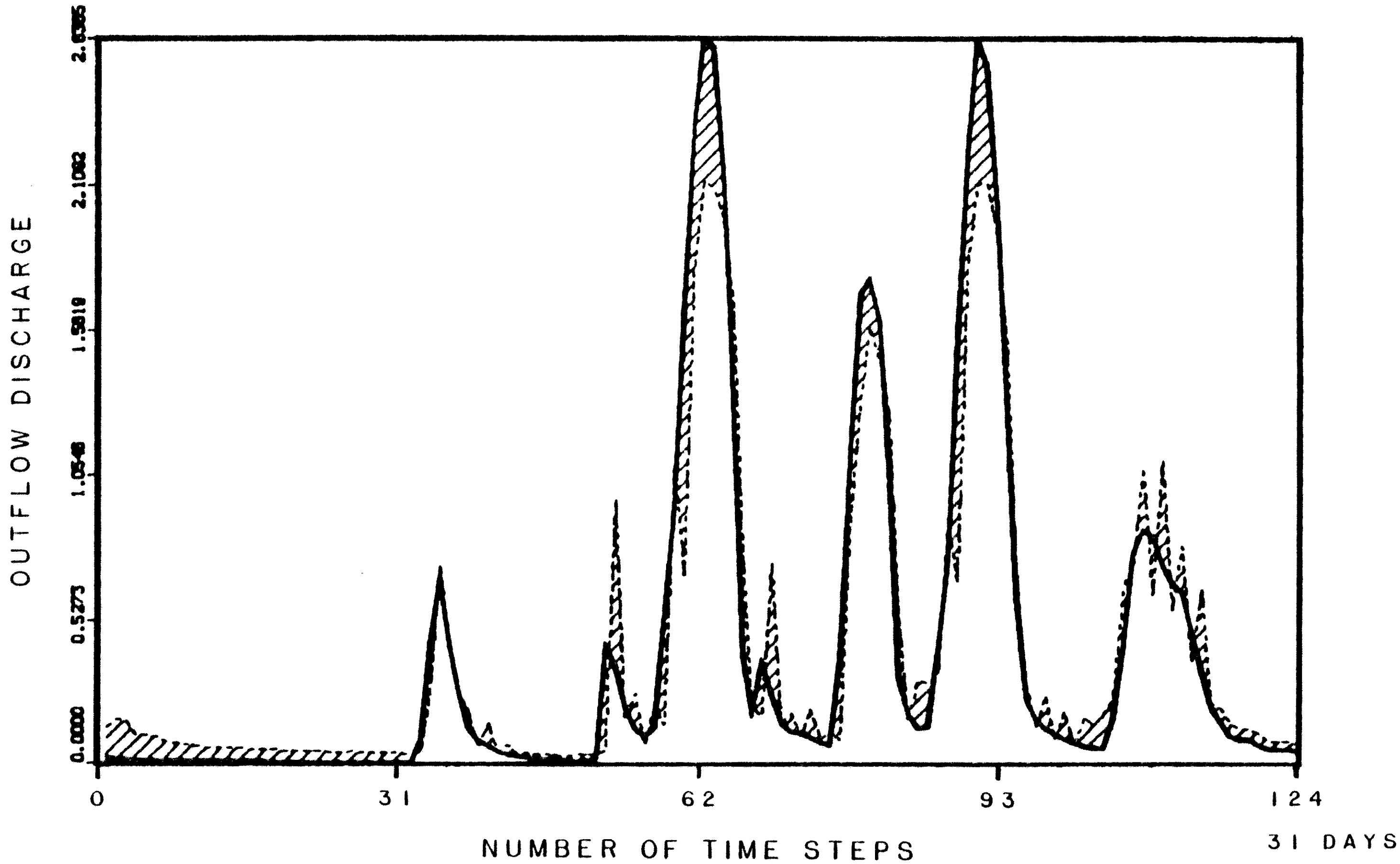


Figure 11. Six-hour forecasts of the outflow discharge in mm/6 hrs (dashed line), with corresponding observed values (solid line).

basin (2344 square kilometers). The model predictions of the discharge outflow for July, 1959, in mm/6 hrs are represented by a thin dashed line in Figure 11. The corresponding observations are represented by a thick solid line. Characteristic of Figures 1-10 is the fact that the highly irregular gain time-traces are radically different for different states.

2.2 Filtering Part of the State Vector

Sims (1974) developed an algorithm for estimating a portion of the state of a linear dynamical system. Based on this idea, one could envision a filter that processes only part of the state vector in real time (the temporarily active part) running together with a decision algorithm which specifies the part of the state vector to be filtered at each time step.

The challenging problem here is to develop the decision algorithm based on past and current data. For example, one can base the decision on high or low past forecasted rainfall activity in each tributary basin; on the relative saturation of the tension water elements; or on the response time of the different basins. Runs of the complete model which will produce the gain time-traces (as in Figures 1-10) can give an indication as to what states are filtered for which types of hydrometeorological conditions. Along the same lines, the sensitivity analysis suggested in Gelb (1974) will prove useful in establishing the criteria for the decision algorithm.

Due to the operational nature of the real time forecasting procedures, the methodology that accompanies this type of simplification should be able to accommodate the full spectrum of possible hydrometeorological situations. Since the decision algorithm is likely to depend on the data available, comparisons with the full complex model running together with the filter for all types of hydrometeorological regimes are indispensible in this case too.

2.3 Partitioning of the State Covariance Matrix

In contrast to the above two classes, this class of simplification can give, a priori, good simplified models that do not need a great number of runs to be calibrated. The idea is to omit the elements of the covariance matrix that are likely to be insignificant due to the weak coupling of the corresponding states.

Since the precipitation process is the one that fills the soil elements of the different tributary basins, it is natural to suppose that the correlation among the soil elements of different basins is due to the spatial correlation of the precipitation input.

Once this is accepted, the model of Georgakakos and Bras (1982) can be used to obtain decomposition algorithms, since it integrates:

- A precipitation process state.
- Soil states as they are represented in the soil moisture accounting scheme of the National Weather Service River Forecast System (model description in Peck, 1976).
- Channel states of the nonlinear channel routing model of Georgakakos and Bras (1980).

The model is physically based and uses, as input, estimates of surface mean areal temperature, mean areal potential evapotranspiration, pressure, and dew-point temperature for each basin. It gives forecasts of mean areal precipitation over a basin and of the outflow discharge. The model forecasts are compared to corresponding observations in a filtering framework in order to update the state estimates in real time. The data-base used is the typical data-base available operationally for time intervals down to a few hours.

With this approach, the correlation between the precipitation model states in different basins is preserved (during the predict-update cycle in the filtering procedure); the soil and channel states of different basins in the covariance matrix are completely decoupled; the influence of the precipitation spatial correlation structure is preserved; and, at the same time, each tributary basin is processed separately.

Based on this idea, given that the precipitation model has one state, the number of computations for N basins increases as n^3+N^3, instead of $(nxN)^3$ for the full model.

The meteorological data can be obtained by linear interpolation from the nearest meteorological stations or the nearest forecast nodes of the Limited-Area Fine Mesh model which runs operationally (NOAA-NMC, 1981). This is due to the high linear correlation of temperature, pressure and dew-point. The data required for each tributary basin are:

1) Forecasts of mean areal temperature, pressure, and dew-point temperature.

2) Observations of mean areal precipitation.

3) Forecasts of mean areal potential evapotranspiration.

4) Observations of outflow discharge.

In cases where areal precipitation or outflow discharge data are not available for some basins in the river system, these basins must be combined with the nearest ones that have data. This part of the river system is thus considered as a whole.

The stochastic decomposition algorithm (presented in Table 1) assumes that the outflow discharge observations in all tributary basins are of the same quality, therefore there is nothing to be gained by updating the model states corresponding to a certain basin by using discharge observations at the outlet of another basin. This implies a diagonal river-system discharge sub-matrix of the observations error covariance matrix. In addition, it assumes that there is little to be gained by updating the soil-channel part of the model from the mean areal precipitation observations.

Table 1.--A stochastic decomposition algorithm

STEP1: Starting from the upstream tributary basins, forecast the state mean and covariance for each basin using the Extended Kalman Filter (Gelb, 1974) prediction equations and the precipitation-soil-channel model of Georgakakos and Bras (1982).

For basins with upstream inflow, the inflow is treated as an uncorrelated (in time) random variable of mean and covariance predicted from the filter when it was run for the upstream basin whose outflow is the inflow under consideration.

STEP2: Form an additional state vector from only the precipitation states corresponding to all tributary basins. Since the precipitation model of Georgakakos and Bras is linear in its state, forecast only the covariance matrix of this new state vector using the filter covariance prediction equations.

STEP3: For the precipitation states mean predicted in STEP1 and the precipitation states covariance predicted in STEP2, use the filter update equations for the precipitation models state vector of STEP2. Thus, obtain a new estimate of the precipitation model state mean and covariance across the observation vector consisting of the observations of mean areal precipitation for all basins.

STEP4: Substitute the mean and variance of each precipitation state of STEP3 for the forecasted mean and variance of each precipitation state in the state mean vector and state covariance matrix of STEP1. In addition, change the cross-covariance elements of the state covariance matrix that correspond to the precipitation state, in such a way that the correlation forecasted in STEP1 and the variance of the precipitation state updated in STEP3 are preserved. The above changes are done for each basin.

STEP5: Obtain new estimates of the mean and covariance of the state vector of the precipitation-soil-channel model for each basin, using the filter equations to update across the basin outflow discharge observation. This completes the predict-update cycle for one forecast period. Starting from STEP1, repeat the above sequence of steps for the next forecast period.

Research areas within the framework given are:

1) The determination of the filter parameters for each tributary basin and for the stack of precipitation models of STEP3 in the algorithm. In particular, the initial state covariance matrix, the system noise covariance parameter matrix, and the observations error covariance matrix.

2) The verification of the algorithm assumptions by testing them in real world applications.

Note that the first research area is also an important one for the full stochastic model of the river system. Note also that there are no parameters in the algorithm that need to be estimated through comparison of the decomposed model forecasts with the forecasts of the composite stochastic model. Therefore, for cases in which one cannot afford to run the full model even for comparison purposes, the proposed algorithm can be used.

3. SUMMARY-CONCLUSIONS

This technical note has examined the problem of the efficient use of modern estimation theory techniques for the real time forecasting of river flows in large river systems of several tributary basins. Various stochastic decomposition techniques were discussed in terms of computational efficiency and suitability for operational use. A stochastic decomposition algorithm based on the partition of the state covariance matrix in real time was proposed as the least dependent, for the calibration of its parameters, on expensive comparison runs with the full stochastic model.

ACKNOWLEDGEMENTS

The comments and suggestions of Michael D. Hudlow, Eric A. Anderson, George F. Smith, and Carlos Puente are gratefully acknowledged.

REFERENCES

Gelb, A., ed., 1974: Applied Optimal Estimation. The MIT Press, Cambridge, Mass., 374 pages.

Georgakakos, K.P. and Bras, R.L., 1980: A Statistical Linearization Approach to Real-Time Nonlinear Flood Routing. Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics Report No. 235, Dept. of Civil Engineering, MIT, 218 pages.

Georgakakos, K.P. and Bras, R.L., 1982: A Precipitation Model And Its Use In Real-Time River Flow Forecasting. Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics Report No. 286, Dept. of Civil Engineering, MIT, 302 pages.

Georgakakos, K.P., Restrepo-Posada, P.J., and Bras, R.L., 1980: On-Line River Discharge Forecasting Using Filtering and Estimation Theory. Progress Report on Contract No. 7-35112 for Aug. 13,1979 - Jan. 13, 1980, Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, Dept. of Civil Engineering, MIT, 129 pages.

Kitanidis, P.K. and Bras, R.L., 1980a: Real-Time Forecasting With A Conceptual Hydrologic Model: 1. Analysis of Uncertainty. Water Resources Research, 16(6), pp. 1025-1033.

Kitanidis, P.K. and Bras, R.L., 1980b: Real-Time Forecasting With A Conceptual Hydrologic Model: 2. Applications and Results. Water Resources Research, 16(6), pp. 1034-1044.

NOAA-NWS, 1981: The LFM Model 1980. NOAA Technical Memorandum NWS NMC-66, U.S. Dept. of Commerce, Washington, D.C., 20 pp.

Peck, E.L., 1976: Catchment Modeling and Initial Parameter Estimation for the National Weather Service River Forecast System. NOAA Technical Memorandum NWS HYDRO-31, U.S. Dept. of Commerce, Washington, D.C., 24 pages.

Sims, C.S., 1974: An Algorithm for Estimating a Portion of the State Vector. IEEE Trans. Automatic Control, AC-19(4), pp. 391-393.

Smith, G.F., 1983: Personal Communication. Hydrologic Research Laboratory, National Weather Service, NOAA, Silver Spring, Md.

Smith, J.A., Sheer, D.P., and Schaake, J.C., Jr., 1982: The Use of Hydrometeorological Data in Drought Management: Potomac River Basin Case Study. Presented at the AWRA International Symposium on Hydrometeorology, Denver, Colorado, 20 pp.