OPTIMAL FILTERING TECHNIQUES FOR ANALYTICAL STREAMFLOW FORECASTING

Vinay Kantamneni, Dan Simon, and Stuart Schwartz
Department of Electrical Engineering
Cleveland State University

Abstract

This paper describes the development of a streamflow forecasting model based on the the Sacramento Soil Moisture Accounting Model and applies optimal filtering techniques to sequentially update watershed-scale soil moisture state values, to improve streamflow predictions. In general hydrology is the study of the waters of the earth, especially with relation to the effects of precipitation and evaporation upon the occurrence and character of water in streams, lakes and water on or below the land surface. The Sacramento model is a hydrologic simulation model developed by the National Weaterer service, and used throughout the nation for operational streamflow forecasting. Here optimal filtering techniques are used in order to predict hydrological variables. An H-infinity filter is used to update daily estimates of the water content in model states representing watershed-scale soil moisture storage. Updated soil moisture storages are then used to predict daily streamflow. The output from the estimator is then compared with the model output without state updating.

1. INTRODUCTION

The Sacramento model [1] is divided into six states and two inputs and one output. Streamflow computed by the Sacramento Catchment [2] Model is the result of processing observed precipitation through a conceptual representation of the uppermost soil mantle identification as the upper zone, and the deeper portion of the soil mantle or lower layer. The model utilizes a set of water storages of determinable capacities, to approximate dynamic soil moisture conditions and the interactions of water in the landscape, which controls the production of streamflow. The soil moisture storages (states) fill and drain with nonlinear dynamics in response to rainfall, verticall percolation, evaporation and lateral drainage.

The streamflow forecasting model has two inputs: precipitation and potential evapotranspiration. The six states of the system are upper zone tension water, upper zone free water, lower zone tension water, lower zone free water primary, lower zone free water secondary, and additional impervious area. The output of the system is streamflow. The inputs of the system are derived from meteorological measurements, from which streamflow is calculated. The results that are obtained from calculation are then compared with the known values of the flow. In order to predict the streamflow, a known set of data of inputs and the outputs from previous records are taken and then the model is initially calibrated to reproduce the observed flow. The results that are obtained from calculation have errors in it. The main errors are the measurement errors in input and the output, and model errors. In order to make the system more accurate the H-infinity estimator is used to update estimates of the model states and these next states are used to calculate the streamflow. The next section gives an overview of the hydrology model and how it is framed in terms of

E-mail:vinay_2@yahoo.com, d.j.simon@csuohio.edu and s.schwartz1@csuohio.edu. This work is supported in part by CSU Center for Environmental Science Technology and Policy

mathematical equations. Section 3 discusses the estimator design. Section 4 presents simulation results and Section 5 contains concluding remarks.

2. HYDROLOGY MODEL

2.1. Principle of water flow

When rainfall occurs over a catchment, the water, which falls over the rocks, hard soil or ponds, lakes, and streams, produces direct runoff. Some water gets into the pervious areas where the water infiltrates into the soil.

The soil is divided into 2 zones.

1. Upper zone
2. Lower zone

The upper zone can be categorized into 3 soil moisture contents.

1. Upper zone tension water.
2. Upper zone free water
3. Hygroscopic water

The lower zone is again divided into 3 soil moisture contents.

1. Lower zone tension water
2. Lower zone free water
   a. Supplementary free water storage
   b. Primary free water storage

When water gets into the upper zone the water is absorbed by the upper zone soil if it is water deficient. Then when the upper zone tension water requirements are met the water will be free to flow within the surface, and is called upper zone free water. When the upper zone water requirements are met the water will be available for the lower zone. Water percolates from upper zone to the lower zone where the water is first absorbed by the lower zone for its tension water requirements. Then this water is allowed to free flow, which in turn increases the lower zone primary and secondary free water storages.

Drainage from free water contents in the lower zone constitutes base flow in that catchment. Base flow takes its maximum value when both of the lower zone free water storages are filled to capacity. When all the soil moisture requirements are met, the excess rainfall that falls constitutes surface runoff and direct runoff.

The model is applied to a 228 square mile river basin in southern Ohio, USA. In this climate, the seasonal precipitation maximum occurs in the spring season. At the start of the spring season the water in the catchment is above the threshold level because of the ice that melts over the soil and because of the percolation that takes place at the end of the winter season. After a prolonged dry season, the water table begins to fall below the required level. Then when there is rainfall this water starts filling up the deficient water table. This is high at the end of the summer season.

The Sacramento Catchment Model is a “Generalized Streamflow Forecasting Model” which was developed by personnel at the California-Nevada River Forecast Center. It was developed by Burnash et al. in 1973 [1] to forecast the river flow in the California-Nevada region where previous approaches were deficient. Streamflow computed by the Sacramento Catchment Model is the result of processing precipitation through an algorithm representing the uppermost soil mantle identification as the upper zone, and the deeper portion of the soil mantle or lower layer.

2.2. Mathematical calculations involved

The Sacramento Model can be represented using derivative equations. This model can be modified as a system having 2 inputs and 6 states, with each state begin non-negative.

States:

x_1 = upper zone tension water content

x_2 = upper zone free water content

x_3 = lower zone tension water content

x_4 = lower zone primary free water content

x_5 = lower zone secondary free water content

x_6 = additional impervious storage

Inputs:

U_p = mean aerial precipitation

U_e = mean aerial evapotranspiration demand

The state equations are given as follows.

$$\frac{dx_1}{dt} = [1 - \left(\frac{x_1}{x^o{}_1}\right)^{m1}]u_p - u_e\left(\frac{x_1}{x^o{}_1}\right)$$

$$\frac{dx_2}{dt} = \left(\frac{x_1}{x^o{}_1}\right)^{m1} u_p[1-\left(\frac{x_2}{x^o{}_2}\right)^{m2}] - d_u x_2 - C_1(1+\varepsilon y^{\theta})\left(\frac{x_2}{x^o{}_2}\right)$$

$$\frac{dx_3}{dt} = C_1(1+\varepsilon y^{\theta})\left(\frac{x_2}{x^o{}_2}\right)(1-p_f)[1-\left(\frac{x_3}{x^o{}_3}\right)^{m3}] - u_e(1-\frac{x_1}{x^o{}_1})\left(\frac{x_3}{x^o{}_1+x^o{}_3}\right)$$

$$\frac{dx_4}{dt} = -d'_l x_4 + C_1(1+\varepsilon y^{\theta})\left(\frac{x_2}{x^o{}_2}\right)[1-(1-p_f)[1-\left(\frac{x_2}{x^o{}_2}\right)^{m3}]][(C_2\frac{x_5}{x^o{}_5}-1)\left(\frac{x_2}{x^o{}_2}\right)+1]$$

$$\frac{dx_5}{dt} = -d''_l x_5 + C_1(1+\varepsilon y^{\theta})\left(\frac{x_2}{x^o{}_2}\right)[1-(1-p_f)[1-\left(\frac{x_3}{x^o{}_4}\right)^{m3}]][(1-C_2\frac{x_5}{x^o{}_5})\left(\frac{x_4}{x^o{}_4}\right)]$$

$$\frac{dx_6}{dt} = [1-\left(\frac{x_6}{x^o{}_3}\right)^{2}[1-\left(\frac{x_2}{x^o{}_2}\right)^{m2}]\left(\frac{x_1}{x^o{}_1}\right)^{m1} u_p - u_e(1-\frac{x_1}{x^o{}_1})\left(\frac{x_6}{x^o{}_3+x^o{}_1}\right)$$

The output u_c from the soil moisture accounting model, referred to as channel inflow per unit time, is given by:

$$u_c = (d_u x_2 + \frac{d''_l x_4 + d_l'' x_5}{1+\mu})(1-\beta_1-\beta_2) + u_p\beta_2 + \left(\frac{x_6}{x^o{}_3}\right)^2 u_p\left(\frac{x_1}{x^o{}_1}\right)^{m1}\beta_1 +$$

$$u_p\left(\frac{x_1}{x^o{}_1}\right)^{m1}\left(\frac{x_2}{x^o{}_2}\right)^{m2}(1-\beta_1-\beta_2) + [1-\left(\frac{x_6}{x^o{}_3}\right)^2]\left(\frac{x_2}{x^o{}_2}\right)^{m2}\left(\frac{x_1}{x^o{}_1}\right)^{m1} u_p\beta_1$$

where

$x^o{}_1$ = upper zone tension water capacity

$x^o{}_2$ =upper zone free water capacity

x^o_3 =lower zone tension water capacity
x^o_4 =lower zone primary free water capacity
x^o_5 =lower zone secondary free water capacity
d_u =upper zone instantaneous drainage coeff
d'_1= lower zone primary instantaneous
d''_1 =lower zone secondary instantaneous drainage
ϵ =parameter in percolation function
θ =exponent in percolation function
μ =fraction of base flow not appearing in river flow
β_1 =fraction of basin that becomes impervious when tension water requirements are met
β_2 =fraction of basin presently impervious
m_l =exponent of upper zone tension water nonlinear reservoir
m_2 =exponent of upper zone free water nonlinear reservoir
m_3 =exponent of lower zone tension water nonlinear reservoir

3. ESTIMATOR DESIGN

The linear system is used to implement the estimator. The estimator is basically used to estimate the 6 states of the system and then these estimated states are used to predict the streamflow. The model processes, and can accumulate measurement errors that the estimator needs to filter. This error is basically due to representations of aggregate model inputs (e.g. mean areal precipitation) from noisy point measurements (e.g. depth of rainfall at a precipitation gauge). Therefore the measurement errors are included in the system, which are provided to the first 3 states of the system and to the output of the system.

$$\Delta \dot{x} = A\Delta x + B\Delta U + B_w \Delta W$$

$$\Delta y = C\Delta x + D\Delta U + D_w \Delta W$$

Here ΔW is a 3 x 1 error vector with errors in U and some noise induced in the system.

Here the H-infinity [3] state estimator is used. H- infinity estimator is called the minmax filter as it tries to minimize the worst case gain from the noise to the estimation error. In order to implement H-infinity estimation there are some basic conditions to be satisfied.

$$D_m B_w^T = 0$$

$$D_m D_m^T = I$$

where D_m will be introduced shortly. The H-infinity estimator estimates the linear combination of the state given as the measured output of the plant. An optimal H-infinity estimator generates estimates that minimizes the worst-case gain between the input and estimation error

$$J = \sup_{w \neq 0} \frac{\| y - \hat{y} \|_2}{\| w \|_2}$$

J is the objective function of the H-infinity estimator.

In order to implement the H-infinity estimator the system is modified as shown below

$$\Delta\dot{x} = A\Delta x + B\Delta U + B_w\Delta W$$
$$\Delta m = C_m\Delta x + D_m\Delta U + D_w\Delta W$$
$$\Delta y = C_y\Delta x$$

where Δm is the change in the measured signal from nominal. The H-infinity estimator equations are given as follows.

$$\Delta\dot{x} = A\Delta x + B\Delta U + G(\Delta m - C_m\Delta x - D_m\Delta u)$$
$$G = QC_m^{\ T}$$
$$\dot{Q} = QA^T + AQ + B_w B_w^{\ T} - Q(C_m^{\ T}C_m - \gamma^{-2}C_y^{\ T}C_y)Q$$
$$\Delta x = \Delta x + \Delta\dot{x}\,dt$$
$$x = xnom + \Delta x$$

where G is the H infinity estimator gain matrix and Q is the solution to the Riccati equation. Here as we are going to estimate the states, $C_y = I$, γ is set to 80 and the system is simulated.

4. EXPERIMENTAL RESULTS

The results that are obtained from the simulation of H-infinity estimator are quite encouraging. They are quantified using the Nash Sutcliffe statistic [4] and correlation coefficient. State updating with the H-infinity filter improved the output of the system (compared to straightforward modeling) in most of the occasions. The Nash Coefficient is increased from .68 to .75 and correlation coefficient is increased from .82 to .86. The following figure is an example of daily streamflow simulation comparing observed flow to simulated flow both with and without state updating, using the H-infinity estimator.

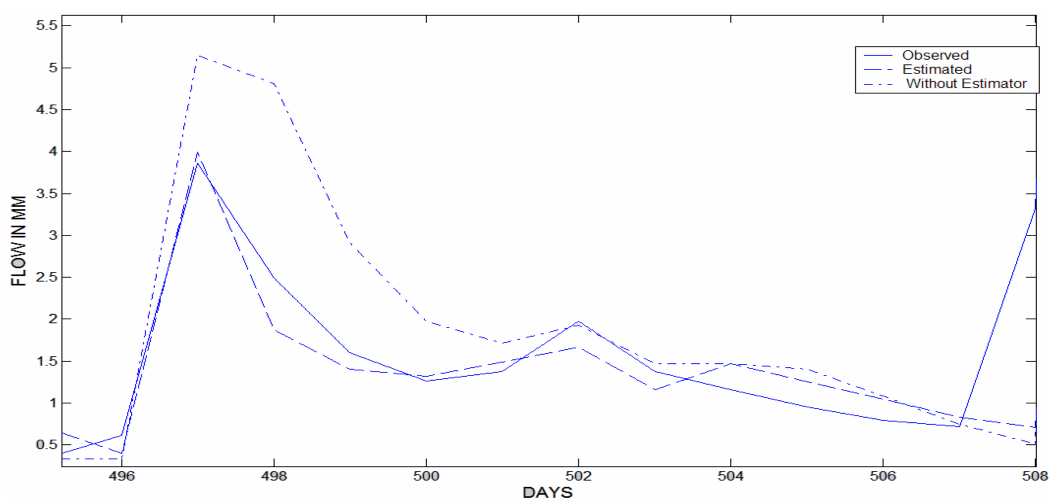

Fig. 1 **Streamflow (mm) vs. Time (days)**

Upon comparing the flow in Fig. 1 in a certain period of the simulated record, the H-infinity estimator worked more effectively than the nonlinear model tracking the observed flow. During this period there was a small amount of precipitation which accounted for the low flow. But the nonlinear model showed some excess streamflow which was not present. This shows that the estimator estimated the states effectively.

6. CONCLUSIONS

The H-infinity filter improved the estimation but it couldn't perfectly estimate the states. This is a combination of unresolved measurement errors, and model error. The errors related to evapotranspiration and precipitation are handled effectively when these are high but less effectively when these are low. In order to make the estimate more accurate there is some tuning that can be done in the system by modeling more noise that effects a state which has an undesirable effect on the system.

Here this model can be further improved using a combination of H-infinity and Kalman filter or a neural network. The future work will be to implement the above mentioned techniques to further improve the behavior of the system.

REFERENCES

[1] Deg Hyo Bae and Konstantine P. Georgakakos "Hydrologic Modeling for Flow Forecasting and Climate Studies In Large Drainage Basins", University of Iowa, IIHR Report No. 360 1992.

[2] Veissman, W., Jr., et al, *Introduction to Hydrology*" 2nd Edition, 1972.

[3] J. Burl, *Linear Optimal Control*, Addison Wesley, 1999.

[4] Nash J.E. and Sutcliffe J.V., River flow forecasting through conceptual models, Part I – A discussion of principles, *Journal of Hydrology*, 10, 282-290, 1970.