Mausam, (1985), 36, 4, 463-468

556·16 : 556.53

Flow discharge prediction of river *Tigris* using self tuning predictors*

A. K. SINHA†

Indian Institute of Technology, New Delhi

and

SADIQ JASSIM

University of Basrah, Basrah, Iraq

(*Received 2 April 1981*)

सार — इस पत्र/लेख में सैल्फ ट्यूनिंग प्रैडिक्टर कहलाने वाले पुनरावर्तन स्वतः अनुकूली प्रागुक्ति एल्गोरिथ्मों के विकास तथा ईराक में बगदाद की टिगरिस नदी के प्रवाह विसर्जन की प्रागुक्ति में इनके अनुप्रयोग को प्रस्तुत किया गया है। इसमें कुछ सामान्य मूल्यांकन तकनीकों का उपयोग किया गया है। चार प्रकार के प्रागुक्ति कारकों/प्रैडिक्टरों यथा, (1) न्यूनतम वर्ग प्रैडिक्टर, (2) न्यूनतम प्राचल प्रेडिक्टर, (3) प्रसम्भाव्य सन्निकटन का उपयोग करने वाला प्रडिक्टर तथा एक दो चरणों वाला प्रैडिक्टर, का विकास किया गया है। उपलब्ध आंकड़ों का उपयोग करके टिगरिस नदी के औसत दैनिक, मासिक तथा वार्षिक विसर्जनों के प्रागुक्ति अनुमान प्राप्त कर लिये गए है। प्रागुक्ति की प्रत्येक किस्म में बहुत से निदर्शों को आजमाया गया है। तुलना के लिये विभिन्न प्रागुक्ति परिणाम ग्राफ के रूप में तथा सारणी रूप में प्रस्तुत किए गए हैं।

ABSTRACT. The paper presents development of recursive self-adaptive prediction algorithms called the self-tuning predictors using some common estimation techniques and their application to prediction of flow discharge of river *Tigris* at Baghdad, Iraq. Four kinds of predictors, *namely*, the least square predictor, the minimum variance predictor, predictor using stochastic approximation and a two stage predictor have been developed. Using available data for the river *Tigris*, prediction results have been obtained for average daily discharge, average monthly discharge and average yearly discharge. In each typt of prediction, a number of models have been tried. The various prediction results are presented graphically and in tabular forms for comparison.

1. Introduction

The application of modern recursive estimation techniques, such as least squares and Kalman filtering (Graupe 1976) is well known in the areas of communication and control. In communication, the techniques are used generally to filter out signals from noisy informations. In control, these are used normally for state and parameter estimation of various processes to be controlled. In the recent years researchers have also used these techniques to solve the problem of modelling and prediction in other areas, such as educational system (Sinha and Singh 1973), weather process (Sinha and Sharma 1975) and river flows (Rao and Kashyap 1974).

In the present paper, the authors have tried to use a unified approach to the problem of recursive prediction in such processes by developing four self adaptive prediction algorithms called the self tuning predictors (Wittenmark 1974) based on some common estimation techniques like the least square, minimum variance, stochastic approximation (Graupe 1976) and two stage estimation (Prasad, Sinha and Mahalanabis 1977) algorithm. These have been used to process real flow discharge data of river *Tigris* at Baghdad (Iraq) to obtain three kinds of predictions, *namely*, the average daily discharge prediction one day in advance the average monthly discharge prediction one month in advance and the average yearly discharge prediction one year in advance. The models developed are stochastic difference equations with or without deterministic input terms and also with sinosoidally varying terms in some cases to take into account the periodic nature of the river flow data. The approximate model orders are obtained by ordinary auto-correlation and cross correlation tests (Sinha and Sharma 1975) and the models are validated by obtaining the error of prediction.

2. Predictor algorithms

In this section we consider the problem of prediction of a process variable, such as river flow discharge and develop various kinds of prediction algorithms.

*The paper was presented in the Symposium "Indo-French School on recent advances in Computer techniques in Meteorology, Biomechanics and Applied Systems" held at I.I.T., New Delhi, February 1980.

†Demised in A.I. *Kanishka* air-crash on 23 June 1985.

It is assumed that the river flow process is represented by the following discrete time linear stochastic difference equation model :

$$y(k) = \sum_{i=1}^{n} a_i\, y(k-i) + \sum_{j=1}^{m} b_j\, u(k-j) + e(k) \qquad (1)$$

where, $k=0, 1, 2\ldots.$ represents a sampling period of either a day, a month or an year depending upon the type of data, $y(k)$ is the average flow discharge during kth period, $u(k)$ is the measured average rainfall in the entire catchment area during kth period, $e(k)$ is a random sequence representing the modelling error and disturbance inputs and a_i and b_j are the unknown coefficients of the model.

The one step ahead prediction of $y(k)$ is defined as the conditional expectation $\hat{y}\{k/(k-1)\}$ given by

$$\hat{y}\,[k/(k-1)] = E\left\{y(k)/y(0), y(1) \ldots y(k-1)\right\}$$

If $e(k)$ is assumed zero mean, using Eqn. (1) this is given by :

$$\hat{y}\,[k/(k-1)] = \sum_{i=1}^{n} \hat{a}_i\, y(k-1) + \sum_{j=1}^{m} \hat{b}_j\, u(k-j) \qquad (2)$$

where $\hat{a}_i$ and $\hat{b}_j$ are some estimates of the unknown coefficients. There are various estimation methods existing in the literature (*e.g., see* Sage and Melsa 1971) which can be used to find out the estimates $\hat{a}_i$ and $\hat{b}_j$ on the basis of observed data. Eqn. (2) will then be called on 'Offline predictor'.

A 'self tuning predictor' is defined as a set of recursive algorithms which makes use of the observations online to yield the predictor parameter estimates and the predicted output simultaneously. A block diagram of a self-tuning predictor is shown in Fig. 1 and the algorithms are described below:

It is simple to show that Eqn. (1) can be rewritten in the following form :

$$y(k) = h(k)\, \theta(k) + e(k) \qquad (3)$$

where,

$$h(k) = [y(k-1)\, y(k-2) \ldots.\, y(k-n)\, u(k-1) \ldots u(k-m)]$$

and $\theta(k) = [a_1\ a_2 \ldots a_n \quad b_1\ b_2 \ldots b_m]^T$

Taking the conditional expectation of both side of Eqn. (3). and assuming $e(k)$ to be zero mean, the predicion equation is given by :

$$\hat{y}[k/(k-1)] = h(k)\, \hat{\theta}[k/(k-1)] \qquad (4)$$

where, $\hat{\theta}\{k/(k-1)\}$ represents the predictor parameter estimate.

For constant parameters it can be seen that (since $\theta(k) = \theta(k-1)$)

$$\hat{\theta}\{k/(k-1)\} = \hat{\theta}\,(k-1)/(k-1) \qquad (5)$$

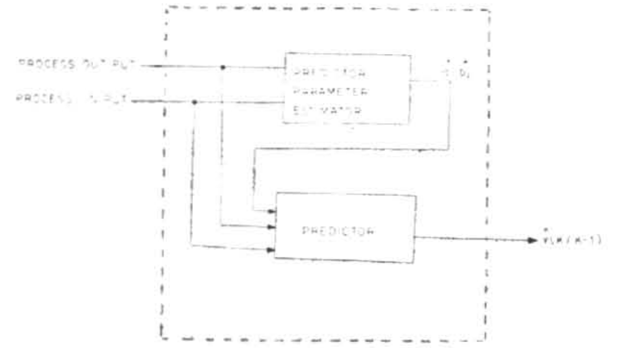

Fig. 1. A block diagram of the self tuning predictor

Let the error of prediction be defined as

$$\tilde{y}\,[k/(k-1)] = y(k) - \hat{y}\,[k/(k-1)] \qquad (6)$$

The parameter updating algorithm is then assumed to be of the following recursive form :

$$\hat{\theta}(k/k) = \hat{\theta}\,(k-1)/(k-1) + K_\theta(k)\, \tilde{y}\,\{k/(k-1)\} \qquad (7)$$

where $K\theta(k)$ is a gain term. Selection criteria for this gain term defines the types of self tuning predictors described below :

(*i*) *The Least square predictor* (*L.S.P.*) — Selection of $K_\theta(k)$ is based on minimization of J_L where

$$J_L = E\,[\,\tilde{y}\,\{k/(k-1)\}]^2 \qquad (8)$$

The gain expression for the above least square criterion is given by :

$$K_\theta(k) = A(k-1)\ h^T(k)\,[h(k)\ A(k-1)\ h^T(k) + 1]^{-1} \qquad (9)$$

where, $\quad A(k-1) = [H^T\,(k-1)\, H(k-1)]^{-1}$

The matrix $H(k-1)$ is a matrix formed by initial $(k-1)$ observation vector $[h(1),\ h(2) \ldots.\, h(k\text{-}1)]$. The recursive formula for updating the matric $A(k-1)$ is given by :

$$A\,(k) = A\,(k-1) - K_\theta\,(k)\, h\,(k)\, A\,(k-1) \qquad (10)$$

The set of Eqns. (4), (7), (9) and (10) from the algorithms for least square self tuning predictor.

(*ii*) *The minimum variance predictor* (*M.V.P.*) — The minimization criterion is the variance of parameter estimates given by

$$J_M = E\,[\theta(k) - \hat{\theta}\{k/(k-1)\}]\ [\theta(k) - \hat{\theta}\{k/(k-1)\}]^T \qquad (11)$$

It is assumed that the noise sequence $e(k)$ is zero mean white Gaussian with a constant variance R. Then treating the following relation

$$\theta(k) = \theta(k-1) \qquad (12)$$

as the state equation and Eqn. (3) as the observation equation, $K_\theta(k)$ is given by the Kalman filter gain expression (Sage and Melsa 1971).

$$K_\theta(k) = P(k-1)/(k-1)\, h^T(k)\,[h(k)\,P(k-1)/(k-1)\, h^T(k) + R]^{-1} \quad (13)$$

where $P(k/k)$ is the filter error covariance matrix given by

$$P(k/k) = P(k-1)/(k-1) - K_\theta(k)\, h(k)\, P\{(k-1)/(k-1)\} \quad (14)$$

Eqns. (4), (7), (13) and (14) constitute the algorithms for minimum variance self tuning predictor. It can be seen that both L.S.P. and M.V.P. are similar except that the knowledge of variance R is utilised in the latter case.

(iii) Predictor using stochastic approximation (P.S.A.)

In the stochastic approximation method (Saridis 1974) the gain $K_\theta(k)$ is selected on the basis of convergence of the parameter estimates. Several expressions are are described by Saridis (1974) and others. A common expression for gain is given by :

$$K_\theta(k) = a/(b+k) \quad (15)$$

where a and b are positive constants which can be chosen by trial for fastest convergence. Eqns. (4), (7) and (13) constitute the algorithms for self tuning predictor using stochastic approximation. This is most economical computationally compared to other algorithms.

(iv) Two stage predictor (T.S.P.)

In model (1), the river flow data has been taken as the values for actual flow. These data ate generally erroneous and contain errors of measurement. In such cases, the following model can be suggested :

$$x(k) = \sum_{j=1}^{n} a_1\, x(k-i) + \sum_{j=1}^{m} b_j\, u(k-j) \quad (16)$$

where $x(k)$ represent the true value of flow discharge. The recorded data is represented by $y(k)$ and is related to $x(k)$ by

$$y(k) = x(k) + v(k) \quad (17)$$

where $v(k)$ is the measurement noise. For convenience, this noise is assumed to be zero mean while Gaussian having a constant variance R_v. Substituting from Eqn. (17), Eqn. (16) can be rewritten in the following form :

$$y(k) = C(x)\, \theta(k) + v(k) \quad (18)$$

Where $C(x) = [x(k-1)\ \ x(k-2) \ldots x(k-n)\ \ u(k-1) \ldots u(k-m)]$

The one step ahead prediction algorithm for model (18) is then given by :

$$\hat{y}\{k/(k-1)\} = C(\hat{x})\, \hat{\theta}\{k/(k-1)\} \quad (19)$$

The above predictor is now function of not only the parameter estimates $\hat{\theta}$ but also of the estimates $\hat{x}$. The above problem can be solved by applying a two stage estimation method [Prasad *et al.* (1977)] briefly explained below :

Stage 1 : Parameter estimation and prediction

With apriori initial assumed estimates of x, minimum variance estimates of θ is obtained using Eqns. (7), (13) and (14) and prediction is made through Eqn. (19).

Stage 2 : Updating of estimate of x

Following the procedure of Prasad *et al.* (1977) the estimates of state can be updated either through a Kalman filter or an stochastic approximation algorithm after transforming Eqns. (16) and (17) to a suitable state variable form. This estimate can be feedback to stage 1 for updating the parameter estimate and then making the prediction for the next period.

3. Application and results

In this section we briefly present some of the results of application of the predictor algorithms described in the previous section to flow discharge prediction of river *Tigris* at Baghdad. Three kinds of data, *namely*, the average daily discharge, the monthly average discharge and yearly average discharge have been used for daily, monthly and yearly discharge predictions. The various data are processed on the HP-2000 computer at the University of Basrah, Iraq.

Daily discharge prediction — The data refers to one water year period from October 1956 to September 1957. From the seasonal behaviour of the data, two separate models have been selected for different periods in the year, Model D-1 valid for the months February to July and Model D-2 valid for the months August to January. On the basis of autocorrelation tests (Sinha and Sharma 1975), the approximate model orders are decided. The two models are given as

Model D-1

$$y_1(k) = a_{11}\, y_1(k-1) + a_{12}\, y_1(k-2) + a_{13}\, y_1(k-3) + a_{14}\, y_1(k-4) + e_1(k) \quad (20)$$

Model D-2

$$y_2(k) = a_{21}\, y_2(k-1) + a_{22}\, y_2(k-2) + a_{23}\, y_2(k-3) + + e_2(k) \quad (21)$$

To provide the initial estimates of the coefficients for the recursive estimation algorithm, the first 40 data is processed nonrecursively using nonrecursive least square algorithm (Graupe 1976) to yield approximate estimates of the coefficients and also the matrix $A(k-1)$ (*see* Eqn. 9) or the matrix $P[(k-1)/(k-1)]$ (*see* Eqn. 13). Three predictor algorithms, *namely*, the L.S.P., the M.V.P and T.S.P. have been used to obtain one day ahead prediction of flow discharge. In case of M.V.P. and T.S.P., the variance of noise $e_1(k)$ and $e_2(k)$ are assumed to be 1000 and 10 respectively on the basis

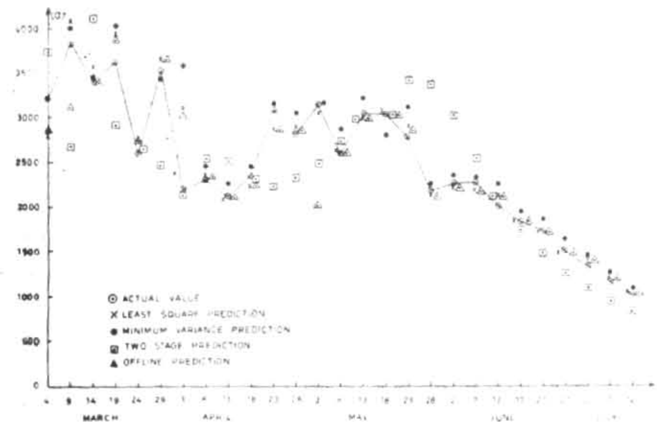

Fig. 2. Graphs of the actual daily discharge and predicted values using model D-1

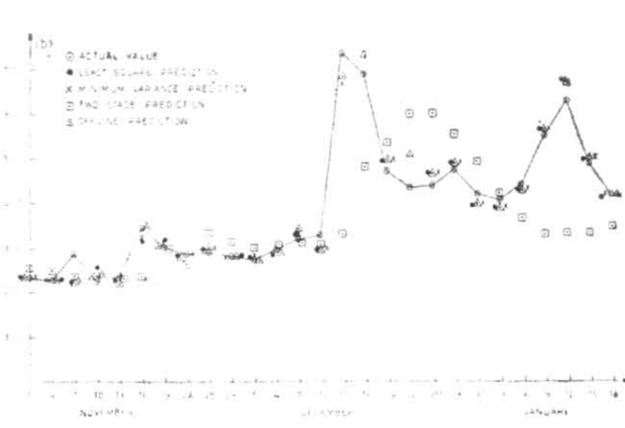

Fig. 3. Graphs of the actual daily discharge and predicted values using model D-2

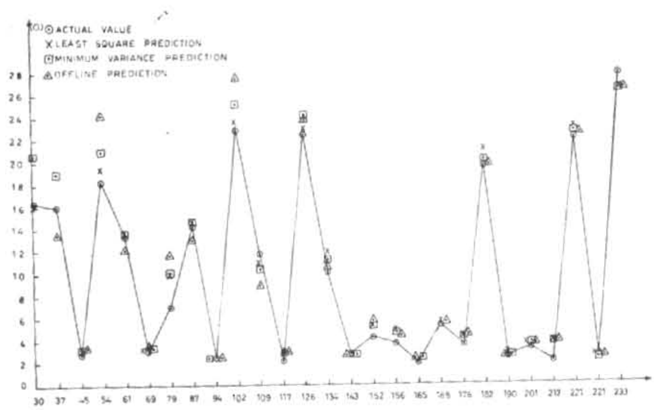

Fig. 4. Graphs of the actual monthly discharge and predicted values using model M-1

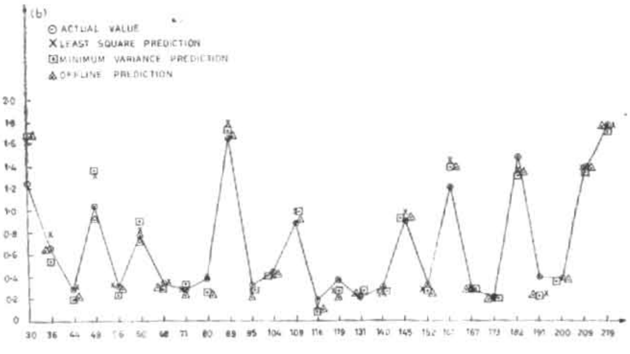

Fig. 5. Graphs of the actual monthly discharge and predicted values using model M-2

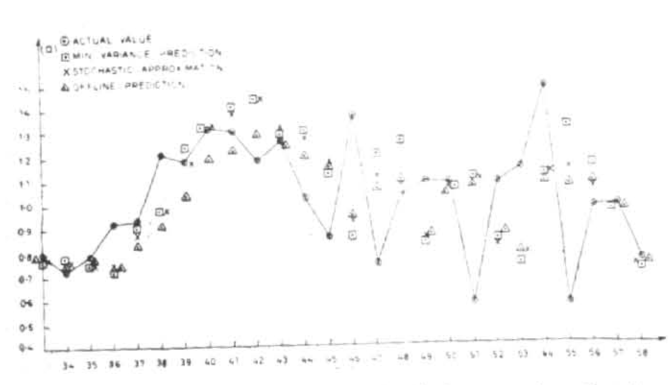

Fig. 6. Graphs of the actual yearly discharge and predicted values for model Y-1

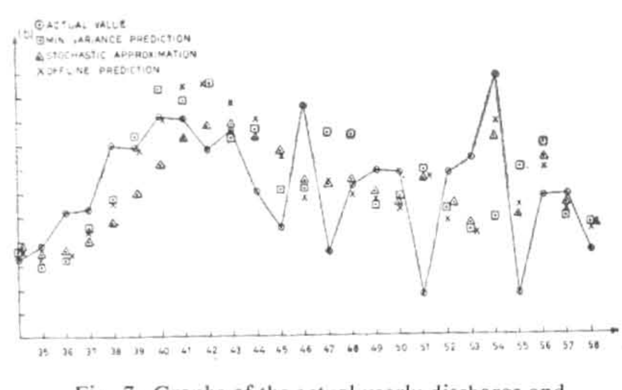

Fig. 7. Graphs of the actual yearly discharge and predicted values using model Y-2

of about 5 per cent measurement error. For the purpose of comparison, offline predictions have also been made using final estimated values of parameter as follows :

$\hat{a}_{11}=1.023$ $\hat{a}_{12}=0.083$

$\hat{a}_{13}=0.317$ $\hat{a}_{14}=0.204$

$\hat{a}_{21}=1.74$ $\hat{a}_{22}=-0.99$

$\hat{a}_{23}=0.25$

Figs. (2) and (3) show the results of predictions using model D-1 and D-2 for various predictor algorithms.

Monthly discharge prediction — The monthly discharge data refers to the period January 1938 to December 1958, a total of 240 values. Data for average monthly rainfall in the catchmen area have also been used in one of the models as deterministic inputs. The main characteristic of the data is that it is periodic in nature with a periodicity of 12 months. Following two models of the type considered by Rao and Kashyap (1974) have been assumed :

Model M-1 :

$$y_1(k) = a_{11} y_1(k-1) + a_{12} y_1(k-12) + b_{11} u(k) + e_1(k) \qquad (22)$$

Model M-2

$$y_2(k) = a_{20} + a_{21} y_2(k-1) + a_{22} \sin[w(k-1)] + a_{23} \cos[w(k-1)] \qquad (23)$$

where $u(k)$ represents the average monthly rainfall and the frequency $w = \frac{2\pi}{12}$

Only L.S.P. and M.V.P. algorithms have been used to obtain one-month ahead predictions. Offline predictions have also been made with following estimated values of parameters

$\hat{a}_{11} = 0.51$, $\hat{a}_{12} = 0.41$, $\hat{b}_{11} = 10.8$,
$\hat{a}_{20} = 0.46$, $\hat{a}_{21} = 0.53$, $\hat{a}_{22} = 0.55$,
$\hat{a}_{23} = 0.27$

Yearly discharge prediction — For the yearly discharge, only 28 years data for the period 1931 to 1958 are available. With this small number of data, the autocorrelation test does not give a clear idea of the model order, hence two models, one a second order and another third order, have been tried. The models assumed are

Model Y-1 :

$$y_1(k) = a_{11} y_1(k-1) + a_{12} y_1(k-2) + e_1(k) \qquad (24)$$

Table 1

The commulative root mean square error (C.R.M.S.E.) of prediction

Kind of prediction	Model	L.S.P.	M.V.P.	P.S.A.	T.S.P.	Offline
Daily	D-1	373.9	639.1	—	645.4	252.5
	D-2	39.8	38.1	—	137.9	30.3
Monthly	M-1	581	577	—	—	552
	M-2	544	545	—	—	521
Yearly	Y-1	—	361	341	—	325
	Y-2	—	384	336	—	317

Model Y-2 :

$$y_2(k) = a_{21} y_2(k-1) + a_{22} y_2(k-2) + a_{23} y_2(k-3) + e_2(k) \qquad (25)$$

For better convergence with small number of data, the predictor algorithm using stochastic approximation has been used with gain $K_v(k)=a/(b+k)$.

Various values of a and b have been tried. A good convergence is obtained for $a=b=1$.

For the sake of comparison, prediction results using M.V.P. and Offline method have also been obtained using the parameter estimates as

$\hat{a}_{11} = 0.43$, $\hat{a}_{12} = 0.55$ $\hat{a}_{21} = 0.26$,
$\hat{a}_{22} = 0.41$, $\hat{a}_{23} = 0.28$

The results are shown in Figs. 6 and 7.

4. Comparison and conclusions

In each type of prediction with various models, the prediction results are compared by the graphs (Figs. 2-7) and also by evaluating the cummulative root mean square error (C.R.M.S.E.) defined as

$$\text{C.R.M.S.E.} = \sqrt{\frac{1}{N} \sum_{k=1}^{N} [\, y(k) - \hat{y}\{k/(k-1)\}\,]^2}$$

where N is the total number of data used.

Values of the C.R.M.S.E. for different prediction results are given in Table 1 for the purpose of comparison. It can be seen from this table that the offline prediction gives the least error in all the cases. This is expected because the prediction results are obtained only after processing all the data to obtain parameter estimates. In the daily discharge prediction, the two stage method has not given better results, meaning thereby that the recorded data does not contain much error. This method, therefore, has been discarded for further prediction.

The above results are however not the best as there are many other considerations regarding the variance of noise, variations of parameters and such other things. The aim of the above work is to illustrate the application of the recursive self adaptive prediction algorithms to a hydrological process of river flow dischare which has been achieved to a certain extent.

Acknowledgement

The authors wish to express their sincere thanks to Dr. Kais Baker Mirza, Head of the Electrical Engineering Department, University of Basrah for his keen interest in the work and for allowing to use the facilities of the department.

References

Graupe, D., 1976, Identification of Systems (New York : Krieger).

Prasad, R. M., Sinha, A. K. and Mahalanabis, A. K., 1977, *International J. Systems Sci*, **8**, pp. 1365-1374.

Rao, R. A. and Kashyap, R.L., 1974, *I.E.E.E. Trans. Aut. Contr.*, **AC-19**, pp. 874-881.

Sage, A. P. and Melsa, J. L., 1971, *Estimation Theory with Application to Communication and Control* (McGraw Hill).

Saridis, G. N, 1974, *I.E.E.E. Tran. Aut. Contr.*, **AC-19**, pp. 798-809.

Sinha, A. K. and Sharma, J. P., 1975, *International J. Systems Sci.*, **6**, pp. 681-688.

Sinha, A. K. and Singh, M. P., 1973, *International J. Systems Sci.*, **4**, pp. 501-509.

Wittenmark, B., *IEEE Trans. Aut. Contr.*, **AC-19**, pp. 848-851.