

MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

12

NOSC TD 556

NOSC TD 556

Technical Document 556

**ADAPTIVE STATE VARIABLE ESTIMATION  
USING ROBUST SMOOTHING**

**F. D. Groutage  
Naval Ocean Systems Center**

**R. G. Jacquot  
University of Wyoming**

**D. E. Smith  
Computer Sciences Corp.**

**6 December 1982**

**Prepared for  
Naval Air Systems Command**

AD A197433

Approved for public release; distribution unlimited

DTIC FILE COPY

**DTIC  
ELECTE**  
APR 29 1983

**B**

**NOSC**

**NAVAL OCEAN SYSTEMS CENTER  
San Diego, California 92152**

**83 04 29 023**

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NOSC Technical Document 556 (TD 556)	2. GOVT ACCESSION NO. AD-A27433	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle)  ADAPTIVE STATE VARIABLE ESTIMATION USING ROBUST SMOOTHING	5. TYPE OF REPORT & PERIOD COVERED	
	6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) F.D. Groutage (Naval Ocean Systems Center) R.G. Jacquot (University of Wyoming) D.E. Smith (Computer Sciences Corp.)	8. CONTRACT OR GRANT NUMBER(s)	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Ocean Systems Center San Diego CA 92152	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS  63303N	
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Air Systems Command Washington DC 20361	12. REPORT DATE 6 December 1982	
	13. NUMBER OF PAGES 26	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report)  Unclassified	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Adaptive states Variable estimations Robust smoothing Electromagnetic Radiating Source Elimination (ERASE)		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  ↓ The development of a conventional Kalman filter is based on a full knowledge of system <i>a priori</i> information. A problem of concern is that associated with determining the estimates of the state variables of a system from observation data when a full knowledge of some <i>a priori</i> system information is unknown. The information includes a knowledge of noise statistics, system forcing functions, and descriptions of system dynamics. This paper addresses only one of the important aspects of the above problem: state variable estimation in the absence of knowledge about deterministic system forcing functions. A robust estimation concept for weighting certain (Continued on reverse side)		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 68 IS OBSOLETE  
S/N 0102-LF-014-6601UNCLASSIFIED  
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Continued.

elements of the Kalman gain and covariance matrices is presented. Robust statistical procedures are used to smooth estimates of the state variables once the estimates are determined by an adaptive Kalman filter. The weights for the elements of the Kalman gain and covariance matrices are functions of the sample means and sample variances of the innovations sequence. A primary application of the techniques presented in this paper is that of determining the estimates of position, velocity, and acceleration of a maneuvering body in three-dimensional space from observed data collected by a remote sensor tracking the maneuvering body.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	



S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

## INTRODUCTION

The idea of estimating the state of a system based on observed data has been the focal point of much research. The concept of least squares and curve fitting was introduced by Gauss<sup>1</sup> in the early 1800s. In more recent times, Norbert Wiener<sup>2</sup> was asked to solve an important specific World War II problem: "How can optimum properties for servomechanisms be formulated and specified?" Wiener formulated a general solution to this problem based on rigorous probabilistic approaches. Among his works, "The Wiener Filter" is widely acclaimed today as a cornerstone of modern estimation theory. Over the past decade and a half, modern sequential estimation theory has been coming of age. Kalman<sup>3</sup> and Kalman and Bucy<sup>4</sup> brought modern estimation theory into the limelight with their historical publications, "A New Approach to Linear Filtering and Prediction Theory" and "New Results in Linear Filtering and Prediction Theory." Scientists, engineers, statisticians, mathematicians, and technical personnel faced with the problem of estimating the behavior or monitoring the state variables of a physical system find themselves relying heavily on modern estimation theory. The problem is not an easy one in that the choice of the estimation technique, in many cases, is problem-dependent. There are many criteria available to specify the methodology for estimating a parameter or variable. Ho and Lee<sup>5</sup> discuss three choices: a) most probable estimate; b) conditional mean estimate; and c) minimax estimate. Jacquot<sup>6</sup> presents a

- 
1. Gauss, K. F., *Theory of Motion of the Heavenly Bodies*, Dover, 1963.
  2. Wiener, N., *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*, MIT Press, 1949.
  3. Kalman, R. E., *A New Approach to Linear Filtering and Prediction Theory*, Trans ASME, (Journal of Basic Engineering) ser D, vol 82, pp 35-45, March 1960.
  4. Kalman, R. E., and R. S. Bucy, *New Results in Linear Filtering and Prediction Theory*, Trans ASME, (Journal of Basic Engineering), ser D, vol 83, pp 95-108, March 1961.
  5. Ho, V. C., and R. C. K. Lee, *A Bayesian Approach to Problems in Stochastic Estimation and Control*, IEEE Trans Automatic Control, vol AC-9, pp 333-339, October 1969.
  6. Jacquot, R. G., *Modern Digital Control*, Marcel Dekker, Inc, 1981.

detailed derivation of the classical discrete-time vector Kalman filter where the system of interest is governed by the stochastic matrix-vector difference equation

$$x(k + 1) = Ax(k) + w(k) + Bu(k) \quad (1)$$

with the measurement process defined by

$$z(k) = Hx(k) + v(k). \quad (2)$$

In equation (1),  $x(k)$  is an  $n \times 1$  vector;  $A$  is an  $n \times n$  state transition matrix;  $B$  is an  $n \times j$  distribution matrix; and  $w(k)$  and  $u(k)$  are  $n \times 1$  and  $j \times 1$  vectors, respectively. The stochastic sequence  $w(k)$  is defined as the state or process noise. In equation (2), the measurement equation,  $z(k)$  is an  $r \times 1$  vector of observations,  $H$  is an  $r \times n$  measurement matrix, and  $v(k)$  is the  $r \times 1$  vector of measurement noise.

The assumptions will be made that the noises  $w(k)$  and  $v(k)$  are stationary, Gaussian random sequences. The problem is to obtain the best estimate of  $x(k)$  such that the conditional mean square error

$$J = 1/2 E[\tilde{x}^T(k) \tilde{x}(k) | Z^k] \quad (3)$$

is minimized. The estimation error  $\tilde{x}(k)$  is given as

$$\tilde{x}(k) = x(k) - \hat{x}(k). \quad (4)$$

The notation  $Z^k \triangleq \{z(1), z(2), \dots, z(k)\}$  is used to represent the past observation or measurement set. The vector  $\hat{x}(k)$  is the estimate of the true vector  $x(k)$ . The conditional mean square error of (3) is conditional on the past observations or measurements. The estimation process is formalized in the following algorithm:

$$\bar{x}(k) = A\hat{x}(k - 1) + q + Bu(k - 1) \quad (5)$$

$$\bar{P}(k) = A\hat{P}(k-1)A^T + Q. \quad (6)$$

Observation Residual:

$$y(k) = z(k) - H\bar{x}(k). \quad (7)$$

Kalman Gain:

$$K(k) = \bar{P}(k) H^T [H\bar{P}(k)H^T + R]^{-1}. \quad (8)$$

Estimation Equations:

Measurement update-

$$\hat{x}(k) = \bar{x}(k) + K(k)[y(k) - r - HBu(k-1) - Hq] \quad (9)$$

Covariance update-

$$\hat{P}(k) = \bar{P}(k) - K(k)H\bar{P}(k). \quad (10)$$

A sequence which is of extreme importance in the field of estimation is termed the innovations sequence and is defined as the observation residual given by equation (7).

This development of the estimation problem is the classical problem for known statistics; ie, R and Q are known quantities; also, the noise sequences have the properties

$$E\{v(k)\} = r \quad (11)$$

and

$$E\{w(k)\} = q. \quad (12)$$

The classical problem also assumes that if the system is driven by a deterministic forcing function, it is completely known for all time.

The noise covariance matrix  $Q$  is defined by

$$Q = E \left\{ [w(k) - q] [w(k) - q]^T \right\} \quad (13)$$

where  $w(k)$  is the random forcing sequence of expression (1) and  $q$  as defined in relation (12), where the mean has been assumed not to be a function of time or  $k$ . The observation noise covariance matrix  $R$  is given as

$$R = E \left\{ [v(k) - r] [v(k) - r]^T \right\} \quad (14)$$

where  $v(k)$  is the measurement noise sequence and  $r$  is the expected value of  $v(k)$ .

Figure 1 presents a block diagram of a system model, measurement system, and discrete Kalman filter. The system model is a discrete representation of a continuous system being observed at discrete times by a measurement system. The assumptions are that  $u(k)$  and  $w(k)$  are constant over the sampling interval; ie,

$$w(t) = w(k); kT \leq t \leq (k+1)T \quad (15)$$

and

$$u(t) = u(k); kT \leq t \leq (k+1)T \quad (16)$$

where  $T$  is the sampling interval.

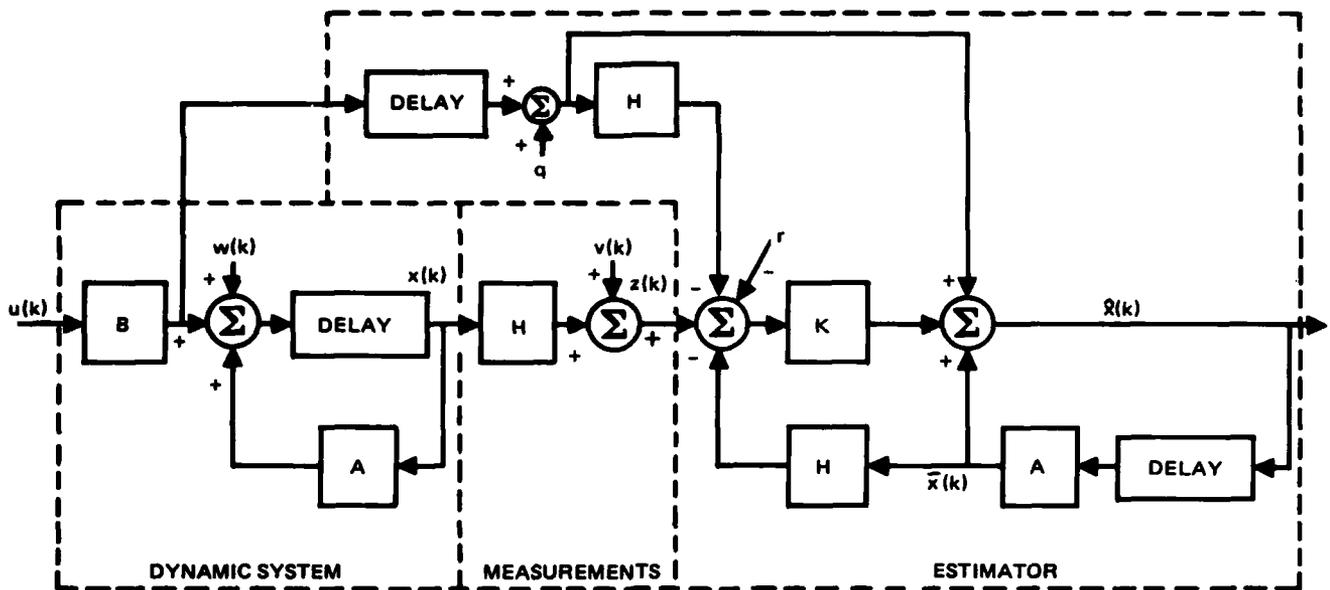


Figure 1. System model and conventional discrete Kalman filter.

A further assumption is that knowledge of the forcing function exists; thus the Kalman filter, as illustrated in figure 1, reflects this knowledge. Consider now the case where a system is being forced by forcing functions  $u(k)$  and  $w(k)$  and the noise corrupted output is observed by a sensor. The deterministic forcing function,  $u(k)$ , is fast varying with respect to time. It is desired to formulate estimates of the value of the state variables in a timely manner by using a Kalman filter; however, for this case, knowledge of the system forcing function,  $u(k)$ , is unknown to the filter. A problem of equal importance is the case where the statistics of the process noise,  $w(k)$ , are unknown. An excellent treatment of estimation in the presence of unknown noise statistics is presented in Myers and Tapley<sup>7</sup>. Empirical estimators are developed in reference [7] which estimate the noise statistics.

For the case at hand, state estimation without knowledge of the deterministic forcing functions, several modifications are incorporated into the estimation process. These include adaptive weighting of the elements of the

7. Myers, K. A. and B. D. Tapley, Adaptive Sequential Estimation with Unknown Noise Statistics, IEEE Trans Automatic Control, vol AC-21, pp 520-523, August 1976.

conventional Kalman gain and covariance matrices, as well as robust statistical smoothing of the estimates made by the adaptive Kalman filter using the modified gain and covariance matrices. Figure 2 illustrates the estimation process with the modifications incorporated. The assumptions are that the noises  $w(k)$  and  $v(k)$  are stationary, Gaussian, random sequences. The modified gain matrix  $\Lambda[K(k)]$  is defined in the next section as well as the robust statistical smoothing procedure.

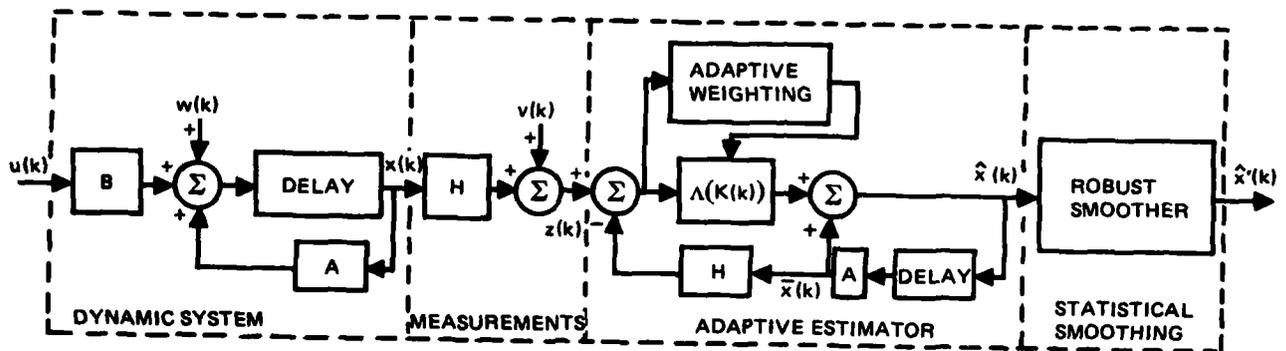


Figure 2. Adaptive robust estimation.

Two concepts are investigated in this paper; 1) adaptive weight functions for the Kalman filter gain and error covariance matrices, and 2) robust smoothing of the estimated state variables.

Robust smoothing can be applied to both the observed and nonobserved state variables. The concept of robust statistical smoothing of nonobserved variables is addressed in Groutage<sup>10</sup>, where it is applied to the maneuvering target tracking problem.

Basically, two robust statistical procedures are used for the robust smoothing of the estimated state variables:

- 1) robust measure of spread, and

2) robust measure of location.

Details on these robust statistics are found in Bickel<sup>8</sup> and Huber<sup>9</sup>. The concepts of using adaptive weights for Kalman filter gain and covariance matrices are presented in [10].

#### ROBUST ESTIMATION OF OBSERVED STATE VARIABLES BY USING ADAPTIVE WEIGHTS FOR GAIN AND ERROR COVARIANCE MATRICES

The concepts presented here are empirical in nature; ie, they are based on observations and experimental data. An intuitive perception of these concepts can be obtained by examining the Kalman filtering algorithm for a scalar parameter. Consider the scalar system with process noise  $w(k)$

$$x(k + 1) = a x(k) + w(k) \quad (17)$$

and the associated measurements

$$z(k) = x(k) + v(k). \quad (18)$$

The covariance associated with the process noise is

$$E[w^2(k)] = Q \quad (19)$$

and the measurement noise covariance is

$$E[v^2(k)] = R. \quad (20)$$

---

8. Bickel, P. J., One-Step Huber Estimates in the Linear Model, Journal of the American Statistical Association, vol 70, no 350, pp 428-433, June 1975.

9. Huber, P. J., Robust Estimation of a Location Parameter, The Annals of Mathematical Statistics, vol 35, no 1, pp 73-101, March 1964.

10. Groutage, F. D., Adaptive Robust Sequential Estimation with Application to Tracking a Maneuvering Target, PhD Dissertation, Department of Electrical Engineering, University of Wyoming, May 1982.

The noises  $w(k)$  and  $v(k)$  are zero mean, Gaussian random sequences. The Kalman filter equations for this scalar case, which formulate estimates of the single state variable  $x(k)$ , are:

Propagation Equations

$$\bar{x}(k) = a \hat{x}(k - 1) \quad (21)$$

$$\bar{P}(k) = a^2 \hat{P}(k - 1) + Q \quad (22)$$

Gain Relationship

$$c(k) = \frac{\bar{P}(k)}{\bar{P}(k) + R} \quad (23)$$

Estimation Equations

$$\hat{x}(k) = \bar{x}(k) + c(k)[z(k) - \bar{x}(k)] \quad (24)$$

$$\hat{P}(k) = [1 - c(k)]\bar{P}(k). \quad (25)$$

Equation (26) for estimating the state variable  $x(k)$  contains interesting information concerning the estimation process. Note that the scalar gain,  $c(k)$ , is bounded from above and below as

$$0 \leq c(k) \leq 1. \quad (26)$$

For the case when  $c(k) = 0$ , equation (24) indicates that total faith is placed in the estimation process. In fact, the measurements are ignored and the previous estimate is the updated estimate. Now consider the case where  $c(k) = 1$ . For this limit, equation (24) indicates that there will be no faith in the estimation process. In fact, the current measurement for the updated estimate of the state variable is used. With these concepts in mind, the idea of a pseudogain,  $\alpha(k)$ , is investigated. Let

$$\alpha(k) \triangleq 1 + e^{-\beta(k)} c(k) - e^{-\nu\beta(k)} \quad (27)$$

where  $v$  and  $\beta(k)$  are parameters to be established and  $v$  can be a constant. The observation residual,  $y(k)$ , is defined as the difference between the measurement and the propagated state estimate

$$y(k) = z(k) - \bar{x}(k). \quad (28)$$

Now consider the recursive sample mean of the observation residual sequence and the recursive sample variance of the residual sequence. Let the sample space be  $N_\ell$  and let  $\bar{y}(k)$  designate the recursive sample mean of the observation residual sequence; ie,

$$\bar{y}(k) = \frac{1}{N_\ell} \sum_{j=k-N_\ell+1}^k y(j). \quad (29)$$

Let  $\hat{\sigma}_y^2(k)$  designate the recursive sample variance (see appendix A) of the observation residual sequence; ie,

$$\begin{aligned} \hat{\sigma}_y^2(k) = & \hat{\sigma}_y^2(k-1) + \frac{1}{N_\ell - 1} \{ [y(k) - \bar{y}(k)]^2 - [y(k-N_\ell) - \bar{y}(k)]^2 \\ & + \frac{1}{N_\ell} [y(k) - y(k-N_\ell)]^2 \}. \end{aligned} \quad (30)$$

If the parameter  $\beta(k)$  of equation (27) is chosen in the following manner

$$\beta(k) = \gamma \hat{\sigma}_y^2(k) \quad (31)$$

then  $\alpha(k)$  is a function of the dispersion of the residuals  $y(k)$ . The value of the constant,  $\gamma$  (a weighting factor), must be determined. Equation (27) is now interpreted. Note that in the limit as  $\beta(k) \rightarrow 0$ , the pseudogain approaches the Kalman filter gain or

$$\lim_{\beta(k) \rightarrow 0} \alpha(k) = c(k) \quad (32)$$

and for the limit as  $\beta(k) \rightarrow \infty$

$$\lim_{\beta(k) \rightarrow \infty} \alpha(k) = 1. \quad (33)$$

Thus for a small dispersion of the residuals, the gain is the optimum Kalman gain  $c(k)$ . For a large dispersion of the residuals, no faith is given to the estimation process.

From practical experiments, it was found that a robust weighting function for the propagated error covariance  $\bar{P}(k)$  was also required. The modified error covariance was defined as

$$\theta'(k) = \bar{P}(k) + f'[1 - \sigma(k)] \quad (34)$$

where

$$\bar{P}'(k) = \sigma(k)\bar{P}(k). \quad (35)$$

For the scalar case, where  $\sigma(k)$  is defined as  $e^{-\nu_1 \hat{\sigma}_y^2(k)}$ ,

$$\theta'(k) = e^{-\nu_1 \hat{\sigma}_y^2(k)} \bar{P}(k) + e^{\alpha_1 \hat{\sigma}_y^2(k)} [1 - e^{-\nu_1 \hat{\sigma}_y^2(k)}]. \quad (36)$$

The quantity,  $e^{\alpha_1 \hat{\sigma}_y^2(k)}$  is limited to some a priori upper bound, T. This is illustrated in figure 3. Note that for a small dispersion of the residuals [ $\hat{\sigma}_y^2(k)$  approaches zero], the modified error covariance  $\theta'(k)$  of expression (36) approaches the Kalman propagated error covariance  $\bar{P}(k)$  or

$$\lim_{\hat{\sigma}_y^2(k) \rightarrow 0} \theta'(k) = \bar{P}(k) \quad (37)$$



Figure 3. Influence curve of sample variance for observation residual.

For the other limit, that is, when the dispersion of the residuals is large [ $\hat{\sigma}_y^2(k)$  becomes large], the term  $\theta'(k)$  approaches the a priori upper bound, or

$$\lim_{\hat{\sigma}_y^2(k) \rightarrow \infty} \theta'(k) = T \quad (38)$$

Note that the gain,  $c(k)$ , is also bounded when the dispersion of the residuals is large; ie,

$$\lim_{\hat{\sigma}_y^2(k) \rightarrow \infty} c(k) = \frac{T}{T + R} \quad (39)$$

These concepts for the scalar case, robust weighting of the Kalman gain and error covariance matrices, are expanded to the vector case. Note that for the scalar case, the gain and covariance matrices are also scalars.

## ADAPTIVE GAIN MATRIX WEIGHTING

To be useful in an adaptive filter, expression (27) must be expanded to the vector dynamic case. This is accomplished with the  $\Lambda[K(k)]$  matrix (for linear measurements or pseudolinear measurements) similar to the  $\alpha(k)$  function for the scalar case of equation (27). When the measurements are linear,  $\Lambda[K(k)]$  will replace the Kalman gain matrix  $K(k)$  in the estimation algorithm.

Consider a matrix  $\Lambda[K(k)]$  defined as

$$\Lambda[K(k)] = [I' + EK(k) - F] \quad (40)$$

where  $I'$ ,  $E$  and  $F$  are  $n \times r$ ,  $n \times n$ , and  $n \times r$  matrices, respectively. Recall that  $x(k)$  and  $z(k)$  are  $n \times 1$  and  $r \times 1$  vectors, respectively. The individual matrices of equation (42) can best be explained through example. Consider the linear system of figure 4. This system is driven by a deterministic forcing function,  $u(t)$ , and process noise,  $w(t)$ . The output is observed discretely with a sensor system. The measurements,  $z(k)$ , made by the sensor are corrupted by a noise sequence  $v(k)$ . The measurement equation is thus

$$z(k) = x_1(k) + v(k). \quad (41)$$

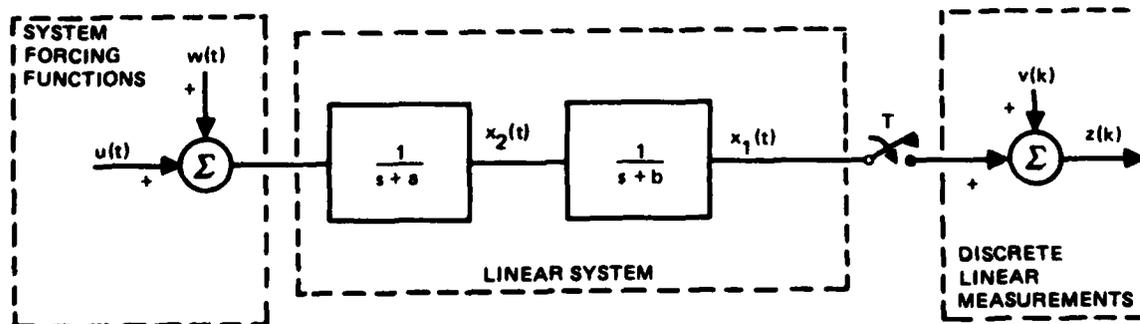


Figure 4. Linear system example.

In matrix form,

$$z(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + v(k). \quad (42)$$

For this example, the I', E, and F matrices are defined as follows:

$$I' = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (43)$$

$$E = \begin{bmatrix} e^{-\beta(k)} & 0 \\ 0 & 1 \end{bmatrix} \quad (44)$$

$$F = \begin{bmatrix} e^{-\nu\beta(k)} \\ 0 \end{bmatrix} \quad (45)$$

where  $\beta(k)$  is as defined in relation (31). It should also be recalled that for this case the K(k) matrix is a 2 X 1 vector.

#### ADAPTIVE ERROR COVARIANCE MATRIX WEIGHTING

The equivalent matrix formulation of expression (36) is

$$\theta'(k) = \bar{P}'(k) + F'(I - \Sigma). \quad (46)$$

Again, the individual matrices of (46) can best be explained through application to the linear problem of figure 4. The propagated error covariance matrix,  $\bar{P}(k)$ , is partitioned into three separate matrices. The following notation is introduced

$$\bar{P}(k) = L(k) + D(k) + V(k) \quad (47)$$

where L(k) is a lower triangular matrix with zeros on the diagonal, D(k) is a diagonal matrix, and V(k) is an upper triangular matrix with zeros on the diagonal. The  $\bar{P}'(k)$  matrix is defined as

$$\bar{P}'(k) = V(k) + \Sigma D(k) + L(k) \quad (48)$$

where the  $\Sigma$  matrix is a diagonal matrix. The individual elements of the  $\Sigma$  matrix are of the form  $e^{-v_i \hat{\sigma}_{y_i}^2(k)}$  where  $\hat{\sigma}_{y_i}^2$  is a sample variance of the  $i$ th observation residual sequence. For the case when all state variables are measured, then all of the diagonal elements of  $\Sigma$  would be of the form

$e^{-v_i \hat{\sigma}_{y_i}^2(k)}$ . The  $i$ th diagonal element,  $\Sigma_{ii}$ , is  $e^{-v_i \hat{\sigma}_{y_i}^2(k)}$ . For the problem at hand, where only discrete measurements of the output variable,  $x_1(k)$ , are available

$$\Sigma = \begin{bmatrix} e^{-v_1 \hat{\sigma}_{y_1}^2(k)} & 0 \\ 0 & 1 \end{bmatrix} \quad (49)$$

and

$$F' = \begin{bmatrix} e^{\alpha_1 \hat{\sigma}_{y_1}^2(k)} & 0 \\ 0 & 1 \end{bmatrix} \quad (50)$$

For a case where measurements of both state variables,  $x_1(k)$  and  $x_2(k)$ , were available the  $\Sigma$  and  $F'$  matrices would be as follows:

$$\Sigma = \begin{bmatrix} e^{-v_1 \hat{\sigma}_{y_1}^2(k)} & 0 \\ 0 & e^{-v_2 \hat{\sigma}_{y_2}^2(k)} \end{bmatrix} \quad (51)$$

$$F' = \begin{bmatrix} e^{\alpha_1 \hat{\sigma}_{y_1}^2(k)} & 0 \\ 0 & e^{\alpha_2 \hat{\sigma}_{y_2}^2(k)} \end{bmatrix}. \quad (52)$$

The quantities  $e^{\alpha_1 \hat{\sigma}_{y_1}^2(k)}$  and  $e^{\alpha_2 \hat{\sigma}_{y_2}^2(k)}$  would be limited similarly as shown in figure 3 to upper bounds of  $T_1$  and  $T_2$ , respectively.

Both the adaptive Kalman gain and error covariance weighting procedures were incorporated into the algorithm of equations (5) through (10). It was determined through experimentation that these robust adaptive procedures were most effective if they were activated only after the sample variance of the innovations sequence reached a predetermined threshold level. As a rule of thumb, the threshold was taken to be one and one-half the anticipated value of the standard deviation of the observation noise.

#### ROBUST SMOOTHING

The estimates of the state variables made by the Kalman filter with the modified gain and covariance matrices contain periodic outliers. This is a result of the sampling procedure and the way in which the sample statistics of the innovations sequence are utilized to formulate weights for the elements of the gain and covariance matrices. To alleviate the outliers in the state estimates, a robust statistical smoothing procedure was incorporated into the estimation procedure. The robust smoother uses a regression procedure in the following manner. Consider  $n$  samples of estimates of the  $i$ th state variable  $\hat{x}_i(k)$  where the samples are defined as the set

$$\lambda = \{\hat{x}_i(k-n-1), \hat{x}_i(k-n), \hat{x}_i(k-n+1), \dots, \hat{x}_i(k)\}. \quad (53)$$

It is desired to find a weighted-least-squares solution for the straight-line regression fit through the  $n$  samples of the estimates of the  $i$ th state variable,  $\hat{x}_i$ , over the discrete interval spanned from discrete time  $k-n-1$  to

time  $k$ . The specific weighted-least-squares solution for the straight-line regression case (ie,  $\bar{Y} = \beta_0 + \beta_1 \bar{X} + \bar{E}$ ) is given by the formulas

$$\hat{\beta}_1 = \frac{\sum_{j=1}^n [(w_j)(j - \bar{x}')(\hat{x}_i(k - n - j) - \bar{Y}')] }{\sum_{j=1}^n [(w_j)(j - \bar{x}')^2]} \quad (54)$$

and

$$\hat{\beta}_0 = \bar{Y}' - \hat{\beta}_1 \bar{x}' . \quad (55)$$

When  $n$ , the sample size, is an odd integer,  $\bar{x}'$  is defined as

$$\bar{x}' = n - \frac{n+1}{2} + 1 \quad (56)$$

and  $\bar{Y}$  is defined as

$$\bar{Y} = \frac{\sum_{j=1}^n [w_j \hat{x}_i(k - n - j)]}{\sum_{j=1}^n w_j} . \quad (57)$$

The weighting term,  $w_j$ , which appears in equations (54) and (57), is called the biweight (See Mosteller and Tukey<sup>11</sup>), which is an abbreviation for bi-square weight. Observations (meaning a sample of a random variable) are weighted according to the relationship

$$w(e_i) = \left. \begin{array}{l} (1 - e_i^2)^2 \\ 0 \end{array} \right\} \begin{array}{l} |e_i| \leq 1 \\ \text{elsewhere} \end{array} \quad (58)$$

11. Mosteller, F., and J. W. Tukey, Data Analysis and Regression, Addison-Wesley Publishing Company, 1977.

where

$$e_i = \frac{t_i - \hat{t}}{cs} \quad (59)$$

and  $t_i$  is the  $i$ th observation with  $\hat{t}$  the estimate of location based on  $n$  observations,

$$\hat{t} = \frac{\sum_{i=1}^N [W(e_i)t_i]}{\sum_{i=1}^N W(e_i)} \quad (60)$$

A robust measure of scale is defined in [12] as

$$s = \frac{(\text{Interquartile Distance})}{2(0.6745)} \quad (61)$$

where the interquartile range is defined as the third quartile minus the first quartile and thus gives the length of the interval in which the middle 50 percent of the data fall. For samples that arise from Gaussian distributions,  $s$  is an estimate of  $\sigma$ , the standard deviation. The value of the constant  $c$  is arbitrary. To have a feel for the range of the value of  $c$ , note that with  $c = 4$

$$cs = 4\sigma. \quad (62)$$

Discrete values of the weighted-least-squares solution along the regression line are obtained from the relationship

$$\hat{Y}(\bar{k}) = \beta_0 + \beta_1 \bar{k} \quad \text{for } \bar{k} = 1, 2, \dots, n \quad (63)$$

---

12. Launer, R. L., and G. N. Wilkinson, *Robustness in Statistics*, Academic Press, New York, 1979.

where  $\beta_0$  and  $\beta_1$  are defined by equations (55) and (54). The vertical distances from regression line of (63) to individual data points at the  $n$  discrete times are called residuals and defined by the relationship

$$r(j) = \hat{x}_1(k - j - 1) - \tilde{y}(j) \quad \text{for } j = 1, 2, \dots, n. \quad (64)$$

The above expressions of (63) and (64) are used in formulating the robust smoothing procedure illustrated in figure 2. This robust procedure is implemented by using the weighted-least-squares solution of (63) to project  $n - 1$  past values of the estimates (as formulated by the adaptive filter) of the  $i$ th state variable up to the present discrete time,  $t = k$ . The  $n - 1$  past values of the estimates of the  $i$ th state variable; ie,

$$\{\hat{x}_1(k - n - 1), \hat{x}_1(k - n), \dots, \hat{x}_1(k - 1)\} \quad (65)$$

are projected to discrete time  $t = k$  and define  $n$  values of the random variable,  $\check{x}_j(k)$ ; ie,

$$\check{x}_j(k) = \tilde{y}(k) + r(j) \quad \text{for } j = 1, 2, \dots, n. \quad (66)$$

The newly formed random variable,  $\check{x}_j(k)$ , is smoothed by using the relationship

$$\hat{x}_1'(k) = \frac{\sum_{j=1}^n w_j \check{x}_j(k)}{\sum_{j=1}^n w_j} \quad (67)$$

where  $\hat{x}_1'(k)$  is the smoothed value of the estimated value of the  $i$ th state variable as generated by the modified gain and covariance Kalman filter. A new estimate, at discrete time  $k+1$ , of the  $i$ th state variable is generated,  $\hat{x}_1'(k+1)$ , which is subsequently smoothed by means of the above process; however, the sample space now spans the discrete time interval from time  $k-n$  to time  $k+1$ . The sample set  $\lambda$  of equation (53) is now defined as

$$\lambda = \{\hat{x}_1(k - n), \hat{x}_1(k - n + 1), \dots, \hat{x}_1(k + 1)\} \quad (68)$$

A new weighted-least-squares solution for the straight-line regression fit through the  $n$  samples is found and the process repeats as outlined above.

Note that the solution of the nonlinear relationships of (54), (57), and (67) are obtained from an iterative procedure. Equations (54), (57), and (67) are nonlinear as a result of the bisquare weight function,  $W_j$ , given by the relationship of equation (60).

#### SIMULATION RESULTS

The system of figure 4 (with  $a = 2.0$  and  $b = 3.0$ ) was simulated on a digital computer; a simulated sensor monitored the output,  $x_1(t)$ , where the output was measured discretely in time and corrupted by sensor noise,  $v(k)$ . The measurement noise,  $v(k)$ , was zero mean with a variance of 25. The system was driven by a deterministic forcing function  $u(t)$ . No process noise,  $w(t)$ , was added to the system forcing function. The deterministic forcing function was a pulse with a duration of 22 seconds and a magnitude of 500 units, as illustrated in figure 5. Also shown in figure 5 are records of the values, as functions of time, of the state variables  $x_2(t)$  and  $x_1(t)$ . The output  $x_1(t)$  was sampled at a rate of five times per second.

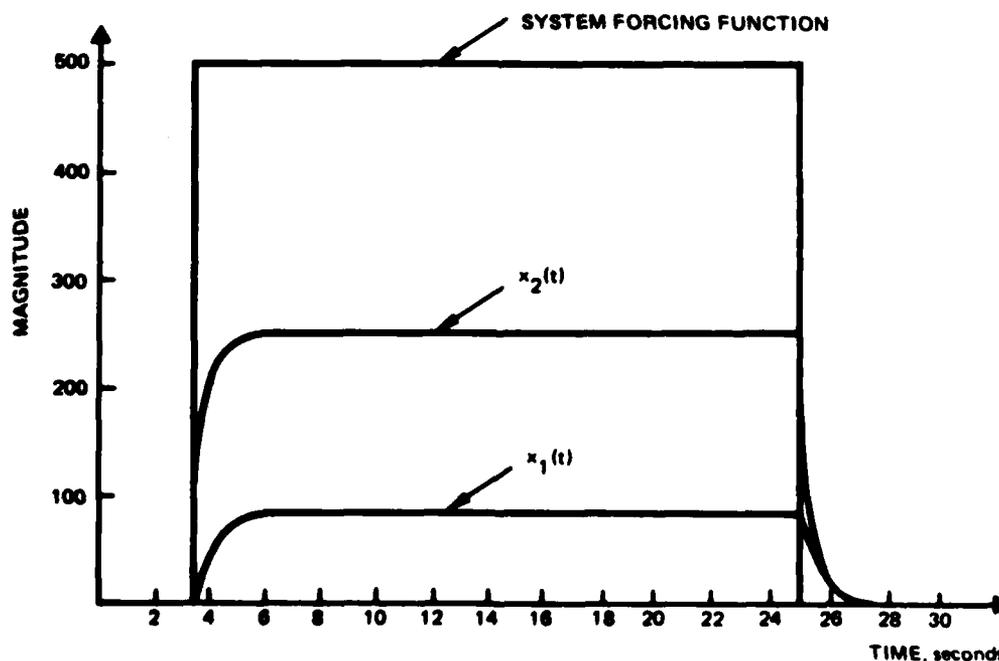


Figure 5. System forcing function and values of state variables.

A conventional Kalman filter, without any a priori knowledge of the forcing function,  $u(t)$ , or the time at which the forcing function was initiated, was used to process the measurement data  $z(k)$ . The conventional Kalman filter did not detect the influence of the deterministic forcing function on the state variables, as illustrated in figure 6. Since a priori data dictated that there was no process noise, the elements of the Kalman filter gain matrix associated with the observed variables approach zero; thus the estimation process has severed itself from the measurement process and ignores new data brought forth by additional measurements.

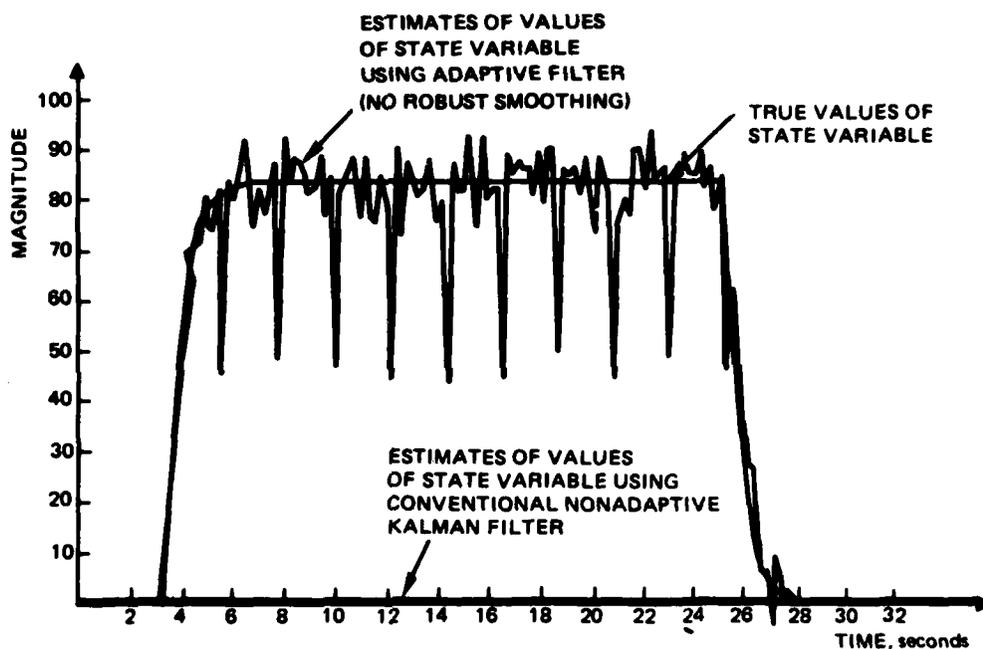


Figure 6. Estimation using adaptive Kalman filter without robust smoothing compared to estimation using nonadaptive conventional Kalman filter.

When the elements of the Kalman filter gain and covariance matrices are weighted by the adaptive procedure outlined above (sample statistics of the innovations sequence are used to adapt the respective weights), the filter no longer divorces itself from the measurement process. Additional data brought forth by the measurement process are used to update the estimates of the state variables. This is illustrated in figure 6. However, since the adaptive procedure uses sample statistics, the estimates contain periodic outliers.

The filter will run for a period of time, then monitor the innovations sequence to update the adaptive weights. It is this monitoring of the innovations sequence to obtain new information which causes the periodic outlier to appear in the estimates. The subsequent processing of the adaptive estimates by a robust smoother reduces the level of mean square error and the periodic outliers.

The smoothed estimates of the measured output state variable,  $x_1(k)$ , are shown in figure 7. Figure 8 presents an overlay of the records of figures 6 and 7 over a 5-second expanded time interval.

### CONCLUSIONS

State variable estimation in the presence of unknown a priori system information (noise statistics, forcing functions, and system dynamics) is not an easy problem. There are no clear-cut solutions. This paper addresses only one of the above problems (no information about the system deterministic forcing functions). The concepts presented relative to this particular problem address the limited class of linear system dynamics with associated linear measurements. Nonlinear system dynamics with associated linear measurements, however, are not precluded.

Estimates of the state variables using the adaptive process for the system during the periods when the system is not being forced are relatively close to those of the conventional Kalman filter for congruent periods, but there is some degeneration because the estimator is no longer optimal. During the periods when the system is being forced, a vast improvement, as compared with those estimates of the conventional Kalman filter, is realized with the adaptive gain, covariance weight, and associated robust smoothing procedure. The estimates derived with the adaptive procedure during the periods of system forcing do, however, contain a considerable level of mean-square error. This seems to be a prevailing shortfall of adaptive estimation procedures. The tradeoff is knowing more about the values of the state variables (less mean error) against more mean square error in the respective estimates.

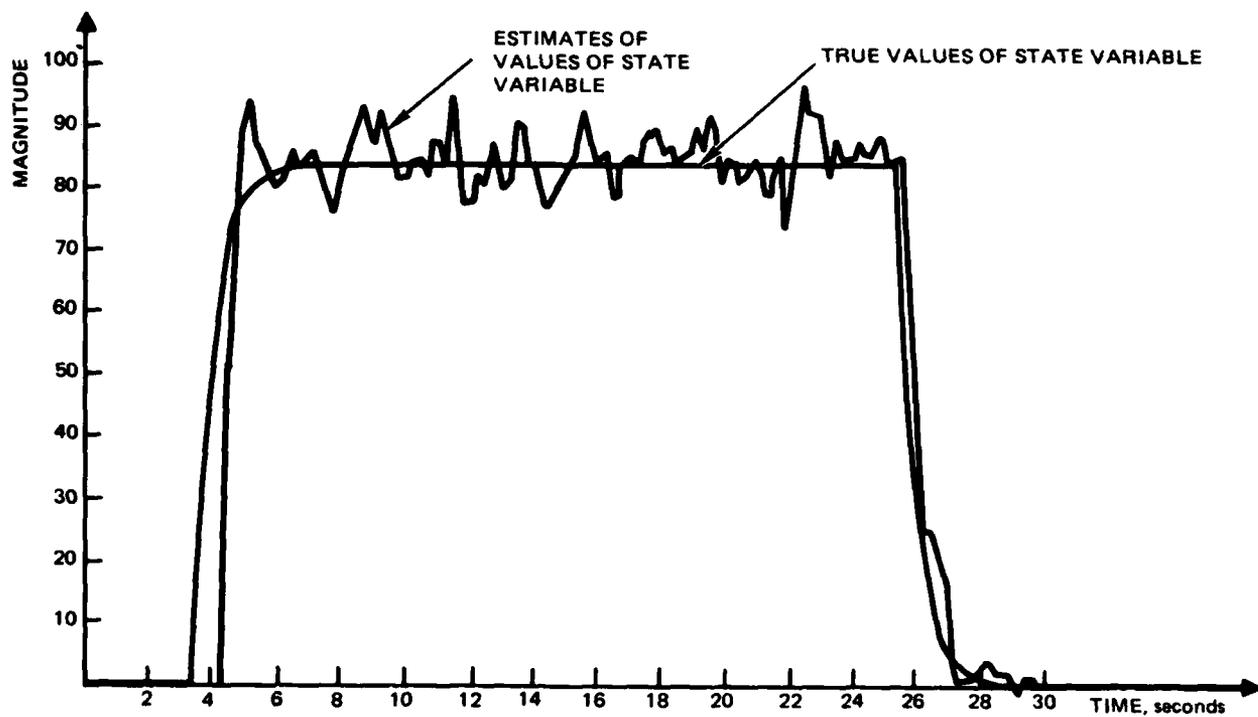


Figure 7. Estimation using adaptive Kalman filter with robust smoothing.

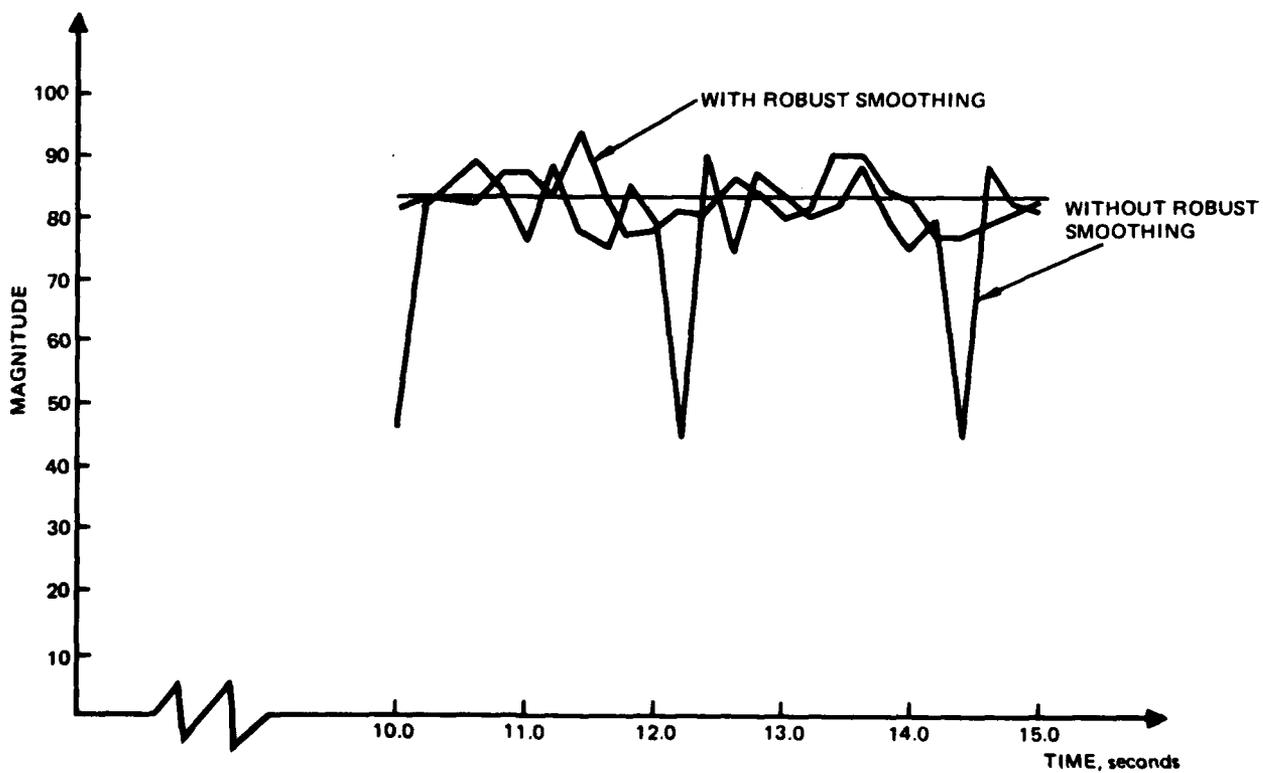


Figure 8. Comparison of filtering techniques with and without robust smoothing.

## REFERENCES

1. Gauss, K. F., *Theory of Motion of the Heavenly Bodies*, Dover, 1963.
2. Wiener, N., *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*, MIT Press, 1949.
3. Kalman, R. E., A New Approach to Linear Filtering and Prediction Theory, Trans ASME, (Journal of Basic Engineering) ser D, vol 822, pp 35-45, March 1960.
4. Kalman, R. E., and R. S. Bucy, New Results in Linear Filtering and Prediction Theory, Trans ASME, (Journal of Basic Engineering), ser D, vol 83, pp 95-108, March 1961.
5. Ho, V. C., and R. C. K. Lee, A Bayesian Approach to Problems in Stochastic Estimation and Control, IEEE Trans Automatic Control, vol AC-9, pp 333-339, October 1969.
6. Jacquot, R. G., *Modern Digital Control*, Marcel Dekker, Inc, 1981.
7. Myers, K. A. and B. D. Tapley, Adaptive Sequential Estimation with Unknown Noise Statistics, IEEE Trans Automatic Control, vol AC-21, pp 520-523, August 1976.
8. Bickel, P. J., One-Step Huber Estimates in the Linear Model, *Journal of American Statistical Association*, vol 70, no 350, pp 428-433, June 1975.
9. Huber, P. J., Robust Estimation of a Location Parameter, *The Annals of Mathematical Statistics*, vol 35, no 1, pp 73-101, March 1964.
10. Groutage, F. D., Adaptive Robust Sequential Estimation with Application to Tracking a Maneuvering Target, PhD Dissertation, Department of Electrical Engineering, University of Wyoming, May 1982.
11. Mosteller, F. and J. W. Tukey, *Data Analysis and Regression*, Addison-Wesley Publishing Company, 1977.
12. Launer, R. L., and G. N. Wilkinson, *Robustness in Statistics*, Academic Press, New York, 1979.

## APPENDIX A

### DERIVATION OF RECURSIVE ESTIMATORS FOR SAMPLE STATISTICS

Presented in this appendix, the derivations of the recursive estimators for the sample mean and sample variance based on  $N_t$  observations.

#### RECURSIVE SAMPLE MEAN

The expression for the recursive estimator for the sample mean at time  $t_{k-1}$  is

$$\hat{n}(k-1) = \frac{1}{N_t} \sum_{j=k-N_t}^{k-1} n(j). \quad (\text{A.1})$$

The expression for the recursive estimator for the sample mean at time  $t_k$  is

$$\hat{n}(k) = \frac{1}{N_t} \sum_{j=k-N_t+1}^k n(j). \quad (\text{A.2})$$

Equation (A.1) is subtracted from (A.2) to give

$$\hat{n}(k) - \hat{n}(k-1) = \frac{1}{N_t} \left[ \sum_{j=k-N_t+1}^k n(j) - \sum_{j=k-N_t}^{k-1} n(j) \right]. \quad (\text{A.3})$$

When the terms under the summation of (A.3) are expanded out and appropriate cancellations take place, the recursive estimator for the sample mean at time  $t_k$  is given as

$$\hat{n}(k) = \hat{n}(k-1) + \frac{1}{N_t} [n(k) - n(k-N_t)] \quad (\text{A.4})$$

RECURSIVE SAMPLE VARIANCE

The expression for the recursive estimator for the sample variance at time  $t_k$  is

$$\hat{\sigma}_n^2(k) = \frac{1}{N_k - 1} \sum_{j=k-N_k+1}^k [n(j) - \hat{n}(k)]^2. \quad (\text{A.5})$$

Equation (A.5) is rewritten as

$$\begin{aligned} \hat{\sigma}_n^2(k) &= \frac{1}{N_k - 1} \{ [n(k) - \hat{n}(k)]^2 - [n(k - N_k) - \hat{n}(k)]^2 \\ &+ \sum_{j=k-N_k}^{k-1} [n(j) - \hat{n}(k)]^2 \}. \end{aligned} \quad (\text{A.6})$$

When the expression for the sample mean, (A.4), is substituted into the summation term of the above expression, (A.6), it can be rewritten as

$$\begin{aligned} \hat{\sigma}_n^2(k) &= \frac{1}{N_k - 1} \{ [n(k) - \hat{n}(k)]^2 - [n(k - N_k) - \hat{n}(k)]^2 \\ &+ \sum_{j=k-N_k}^{k-1} \left[ (n(j) - \hat{n}(k - 1)) - \frac{1}{N_k} (n(k) - n(k - N_k)) \right]^2 \}. \end{aligned} \quad (\text{A.7})$$

Expanding the terms under the summation gives

$$\begin{aligned}
 \hat{\sigma}_n^2(k) &= \frac{1}{N_k-1} \{ [n(k) - \hat{n}(k)]^2 - [n(k - N_k) - \hat{n}(k)]^2 \\
 &+ \sum_{j=k-N_k}^{k-1} \{ (n(j) - \hat{n}(k-1))^2 - \frac{2}{N_k} (n(j) - \hat{n}(k-1)) \\
 &\quad (n(k) - \hat{n}(k-N_k)) + \frac{1}{N_k^2} (n(k) - \hat{n}(k-N_k))^2 \} \}. \tag{A.8}
 \end{aligned}$$

The cross terms under the summation of the above expression of (A.8) can be shown to be zero, thus,

$$\begin{aligned}
 \hat{\sigma}_n^2(k) &= \frac{1}{N_k-1} \{ [n(k) - \hat{n}(k)]^2 - [n(k - N_k) - \hat{n}(k)]^2 \\
 &+ \sum_{j=k-N_k}^{k-1} \{ (n(j) - \hat{n}(k-1))^2 + \frac{1}{N_k^2} (n(k) - \hat{n}(k-N_k))^2 \} \}. \tag{A.9}
 \end{aligned}$$

Note that by definition

$$\hat{\sigma}_n^2(k-1) = \sum_{j=k-N_k}^{k-1} [n(j) - \hat{n}(k-1)]^2; \tag{A.10}$$

thus, the recursive estimator for the sample variance is

$$\begin{aligned}
 \hat{\sigma}_n^2(k) &= \hat{\sigma}_n^2(k-1) + \frac{1}{N_k-1} \{ [n(k) - \hat{n}(k)]^2 - [n(k - N_k) \\
 &\quad - \hat{n}(k)]^2 + \frac{1}{N_k} (n(k) - \hat{n}(k-N_k))^2 \}. \tag{A.11}
 \end{aligned}$$