NASA CONTRACTOR
REPORT

NASA CR-129020

MODERN CONTROL CONCEPTS
IN HYDROLOGY

By Nguyen Duong, Gearold R. Johnson,
and C. Byron Winn

Colorado State University
Fort Collins, Colorado   80521

February 1974

Prepared for

NASA-GEORGE C. MARSHALL SPACE FLIGHT CENTER
Marshall Space Flight Center, Alabama   35812

1. REPORT NO. NASA CR-129020	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.	
4. TITLE AND SUBTITLE MODERN CONTROL CONCEPTS IN HYDROLOGY		5. REPORT DATE February 1974	
		6. PERFORMING ORGANIZATION CODE	
7. AUTHOR(S) Nguyen Duong, Gearold R. Johnson, and C. Byron Winn		8. PERFORMING ORGANIZATION REPORT #	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Colorado State University Fort Collins, Colorado 80521		10. WORK UNIT NO.	
		11. CONTRACT OR GRANT NO. NAS 8-28655	
12. SPONSORING AGENCY NAME AND ADDRESS NASA Washington, DC 20546		13. TYPE OF REPORT & PERIOD COVERED CONTRACTOR REPORT	
		14. SPONSORING AGENCY CODE	

15. SUPPLEMENTARY NOTES

Prepared under the technical monitorship of the Aerospace Environment Division, Aero-Astrodynamics Laboratory, NASA-Marshall Space Flight Center.

16. ABSTRACT

Two approaches to an identification problem in hydrology are presented based upon concepts from modern Control and Estimation Theory. The first approach treats the identification of unknown parameters in a hydrologic system subject to noisy inputs as an adaptive linear stochastic control problem; the second approach alters the model equation to account for the random part in the inputs, and then uses a nonlinear estimation scheme to estimate the unknown parameters. Both approaches use state-space concepts. The identification schemes are sequential and adaptive and can handle either time invariant or time dependent parameters. They are used to identify parameters in the Prasad model of rainfall-runoff. The results obtained are encouraging and conform with results from two previous studies; the first using numerical integration of the model equation along with a trial-and-error procedure, and the second, by using a quasi-linearization technique. The proposed approaches offer a systematic way of analyzing the rainfall-runoff process when the input data are imbedded in noise.

17. KEY WORDS hydrology models, control and estimation theory Prasad model parameters linear stochastic non-linear estimation time invariant or time dependent parameters	18. DISTRIBUTION STATEMENT Unclassified-Unlimited Ludie G. Richard Acting Director Aero-Astrodynamics Laboratory		
19. SECURITY CLASSIF. (of this report) Unclassified	20. SECURITY CLASSIF. (of this page) Unclassified	21. NO. OF PAGES 108	22. PRICE 8.50 NTIS

FOREWORD

The present study is part of the program "Feasibility

of Aircraft Surveys for Stream Discharge Prediction" con-

ducted by Colorado State University, Fort Collins, Colorado,

under George C. Marshall Space Flight Center Contract

NAS8-28655. The authors would like to express their apprecia-

tion for the NASA support. Mr. Joseph Sloan, Aerospace Environ-

ment Division, Marshall Space Flight Center, was the contract

monitor.

iii

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Mathematics is a universal tool in the physical sciences, and much of the insight gained in other fields, especially in systems engineering, is directly applicable to hydrology (Dawdy, 1969). Since Modern Control and Estimation Theory have been applied successfully to aerospace engineering problems (i.e., satellite tracking, orbit determination, space navigation, etc...) in the last two decades, and there exist many similarities between satellite tracking problems and the identification of unknown parameters of hydrologic processes (i.e., the models are not known precisely; the system under study is stochastic and highly non-linear; also, there is noise in the observations), an attempt to use this new approach for the study of hydrologic systems is worthwhile to be investigated.

Characteristics of Hydrologic Systems

A hydrologic system may be defined as an interconnection of physical elements which are related to water in its natural state. The essential feature of a hydrologic system lies in its role in generating outputs (i.e., runoff,...) from inputs (i.e., rainfall, snowmelt, temperature,...), or in interrelating inputs and outputs. The stochastic nature of the inputs and outputs of hydrologic systems has been discussed by Yevjevich (1971).

Hydrologic processes are complex time-varying distributed phenomena, which are controlled by an unknown number of climatic and physiographic factors. The later descriptors tend to be static

1

or change slowly in relation to the time scale of hydrologic
fluctuations. Observations of results in the laboratory and in the
field also indicate that many of the component processes in hydrol-
ogy are nonlinear due to the following reasons (Amorocho and Orlob,
1961):

(1) the time variability of watersheds due to the natural
processes of weathering, erosion, climatic changes, etc...

(2) the uncertainty with respect to the space and time
distribution of the inputs and outputs of hydrologic systems, and
with respect to the states and characteristics of the interior
elements of the system in time; and,

(3) the inherent nonlinearity of the processes of mass and
energy transfer that constitute the hydrologic cycle.
Thus, for systems engineers, hydrologic processes can be considered
as nonlinear dynamic distributed-parameter systems with partially
known or unknown structures, operated in a continuously changing
environment. The inputs and outputs of these systems are measurable
but the data obtained are imbedded in noises with partially known
or unknown characteristics.

Modeling Problems in Hydrology

Precise mathematical models developed for the study of
hydrologic systems should be nonlinear, dynamic, distributed
parameter models. However, at the present time, because of lack of
data on parameter distribution, the assumption of space invariance is
unavoidable. The subdivision of large watersheds into environmental

zones, where environmental conditions which affect the behavior of hydrologic systems can be assumed as uniform, and the use of lumped-parameter model for each zone, are then required to improve the modeling situation. By routing the flow spacewise through all the lumped-parameter models representing the environmental zones the total simulation of the entire watershed represents a distributed-parameter system.

Lumped-parameter models of hydrologic systems can be divided into deterministic and stochastic models. The deterministic approach often is called parametric modeling. The choice of the model is determined by the type of problem to be solved. Parametric models require input data with considerable detail in time, therefore, they model transient responses well and are most widely used for short-term simulation or for actual prediction for water management purposes (Dawdy, 1969). Stochastic models have the advantage of taking into account the chance dependent nature of hydrologic events. Stochastic synthesis models are concerned with the simulation of the relationship between input and output data (cross-correlation models) and between successive values of each data series (serial correlation models). In stochastic simulation models, statistical measures of hydrologic variables are used to generate future events to which probability levels are attached. But in this case long term records, which in many instances are not available, are needed to estimate the parameters of the stochastic model in order to obtain a proper representation of their stochastic nature. Stochastic simulation models usually are used for planning purposes to develop many

"equally likely" long-term traces of monthly streamflow or similar smoothly varying responses (Fiering, 1967).

Efficient management of water runoff from a watershed requires that hydrologic systems be described by dynamic models with sufficient accuracy. Since many controlling factors of hydrologic systems such as weather conditions, soil moisture variations, are known very little, or unknown, deterministic models do not at present offer satisfactory results. However, in this case, with certain reasonable assumptions, one can use random variables to approximate the stochastic nature of the system, and, then, can analyze it if one can track these "random" variables with time.

For dynamic systems which are well characterized by finite-order ordinary differential equations (differential systems) with additive noise terms, when the analysis in the time-domain is to be preferred, the use of the so-called state-space approach will offer a great deal of convenience conceptually, notationally, and, sometimes, analytically. The use of state-space concepts and modern control and estimation techniques for the analysis of lumped-parameter response models of hydrologic systems, when input and output data are corrupted by additive noises, will be discussed further in subsequent chapters.

Identification Problems in Hydrology

Zadeh (1962) defines the problem as

"Identification is the determination, on the basis of input and

output of a system within a class of systems (models), to which the system under test is equivalent (in some sense)."

Thus, techniques for system identification must be based on data. For deterministic models, errors in data are reflected in the identification results and in errors in predicted outputs from an incorrect model of the system. An empirical study of the response of a simulation model of the rainfall-runoff process to input and output errors has been made by Dawdy, Lichty, and Bergmann (1972). For their model, they found that the errors of streamflow estimates are approximately linearly related to errors of rainfall input data.

Errors in rainfall-runoff data can result from the following sources (Rao and Delleur, 1971): (a) errors in reading stage hydrographs, (b) errors inherent in the rating table, (c) errors associated with the method used in base-flow separation, (d) errors resulting from an inadequate network of precipitation stations and (e) errors in the method of determining the rainfall excess. While the errors associated with reading the stage hydrograph can be estimated, at present no accurate assessment can be made of the other errors involved. Since precipitation changes rapidly in time and space over a watershed and a network of precipitation stations may not be adequate, point rainfall measurements are unlikely to be representative of the actual rainfall on the watershed. On the contrary, the runoff is eventually collected at a single point, the mouth of the watershed, and the discharge data are usually of very good quality. It could, therefore, be reasonably assumed that the

runoff is reasonably noise-free compared to the rainfall. Thus, techniques that can use the observed outputs to compensate the disturbance in the inputs (i.e., control the inputs) and at the same time adaptively learn the characteristics of the unknown noise in order to get better system identification are needed in the study of the rainfall-runoff process. These techniques can easily be derived from modern control and estimation theory, based on state-space concepts.

For large scale hydrologic systems it is not often possible to specify the a priori structure or functional form of the model (Kisiel, 1969). Structural information may be none, partial or complete. In the complete-information case, the system identification task reduces to a parameter estimation problem. The study reported in this thesis deals only with the applications of the state-variable approach in modern control and estimation theory to the identification of unknown parameters of nonlinear lumped-parameter response models of hydrologic systems subject to noisy input-output data.

Advantages of the State-Variable Approach

Modern control theory is based on the state-space concept. The idea of state as a basic concept in the representation of systems was first introduced in 1936 by A. M. Turing (Tou, 1964). Later, the concept was employed by C. E. Shannon in his basic work on information theory. The application of the state-space concept in the control field was initiated in the forties by the Russian scientists M. A. Aizerman, A. A. Fel'dbaum, A. M. Letov, A. I. Lur'e

and others. In the United States the introduction of the concept of state and related techniques into the optimum design of linear as well as nonlinear systems is due primarily to R. Bellman. The basic work of R.E. Kalman in estimation and control theory, and the extension of his work by others, played an important role in the advancement of modern control theory.

The state of a dynamic system is defined (Ogata, 1970) as the smallest set of variables, called state-variables, such that the knowledge of these variables at $t = t_o$ together with the input for $t \geq t_o$ completely determines the behavior of the system for any time $t \geq t_o$. The state-space representation of a system is not unique. In the design of optimum control systems it is extremely desirable that all the state variables be accessible for measurement and observation. For a linear system in the linear filtering problem, Athans (1967) showed that the choice of the state-variables is not crucial since one can obtain the estimates of another set of state variables using a simple linear nonsingular transformation. The same linear transformation also links the error covariance matrices in the two models.

The advantages of the state-space concept over the conventional transfer function approach can be listed as follows (Ogata, 1970):

(1) The state-variable formulation is natural and convenient for computer solutions.

(2) The state-variable approach allows a unified representation of digital systems with various types of sampling schemes.

(3)   The state-variable method allows a unified representation of single variable and multi-variable systems.

(4)   The state-variable method can be applied to certain types of nonlinear and time-varying systems.

Control Systems Terminology

A system to be controlled, called a plant, has a set of outputs represented by the vector Y and a set of inputs represented by the vector U.   A priori information about the plant may also be available, which usually is the desired output and represented by the vector R.

Definition 1: An open-loop control system is one in which the control action is independent of the output.   In this case, U is obtained as an operation on R, and the operator is known as an open-loop controller.   The block diagram of an open-loop control system is illustrated in Figure 1-1.

R
REFERENCE INPUT → | open-loop controller | → U CONTROLLED INPUT → | PLANT | → Y CONTROLLED OUTPUT
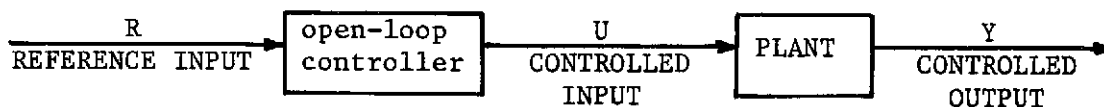
Figure 1-1:   Block diagram of an open-loop control system.

Two outstanding features of open-loop control systems are:

(1)   Their ability to perform accurately is determined by their calibration.   To calibrate means to establish or re-establish the input-output relation to obtain a desired system accuracy.

(2)   They are not generally troubled with problems of instability.

Definition 2: A closed-loop control system is one in which the
control action is somehow dependent on the output. Closed-loop
control systems are more commonly called feedback control systems.
In this case, U is an operation on R and Y, the operator is called
a feedback controller. The block diagram of a feedback control
system is illustrated in Figure 1-2.
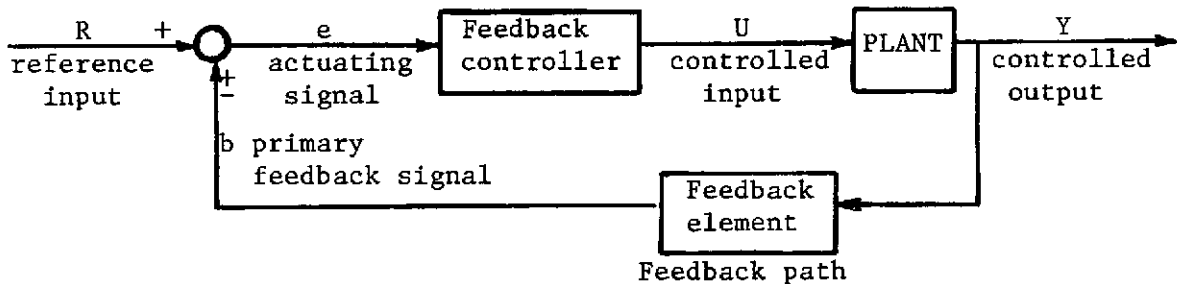
Figure 1-2: Block diagram of a feedback control system.

When the summing point is a subtracter, i.e., e = R - b, one
has negative feedback. When it is an adder, i.e., e = R + b, one
has positive feedback.

The most important features the presence of feedback imparts
to a system are the following:

(1) Increased accuracy. For example, the ability to faithfully
reproduce the input.

(2) Reduced sensitivity of the ratio of output to input to
variations in system characteristics.

(3) Reduced effects of nonlinearities and distortion.

(4) Increased bandwidth. The banwidth of a system is that
range of frequencies (of the input) over which the system will
respond satisfactorily.

(5) Stabilized effect to an unstable system, i.e., without
the addition of feedback, the unavoidable uncertainties in initial
conditions and the inaccuracies in the model that would be used for
determining an open-loop control would render such a system useless.

If everything about the environment and process is known
a priori, the design of the control law is straightforward and can
be accomplished by means of proven techniques. On the other hand,
if the environment or process is poorly defined, more advanced and
sometimes less-proven techniques must be used to design the law.
In the latter situation, control specialists have devised adaptive
control systems and learning control systems.

Definition 3: An adaptive control system is one which is provided
with: (1) a means of continuously monitoring its own performance
relative to desired performance conditions, and (2) a means of
modifying a part of its control law, by closed-loop action, so as
to approach these conditions (Cooper and Gibson, 1960).

A comprehensive survey of adaptive control systems was
presented by Aseltine et al. (1958). The definition of adaptivity
and characteristics of adaptive control systems have also been
treated by Zadeh (1963), Braun (1959), Donalson and Kishi (1965),
Eveleigh (1967) and Davies (1970).

Definition 4: A learning control system is an improved adaptive
system which can memorize the optimal control function once
established through adaptation and can immediately execute optimal
control without adaptive search when a once experienced situation
takes place again (Tamura, 1971).

A learning control system, therefore, should have memory facilities to store pairs of experienced situations and the results of adaptation. Moreover, it should have the capability to relate a certain control function with the present situation. A representative learning control system can be illustrated as in Figure 1-3.
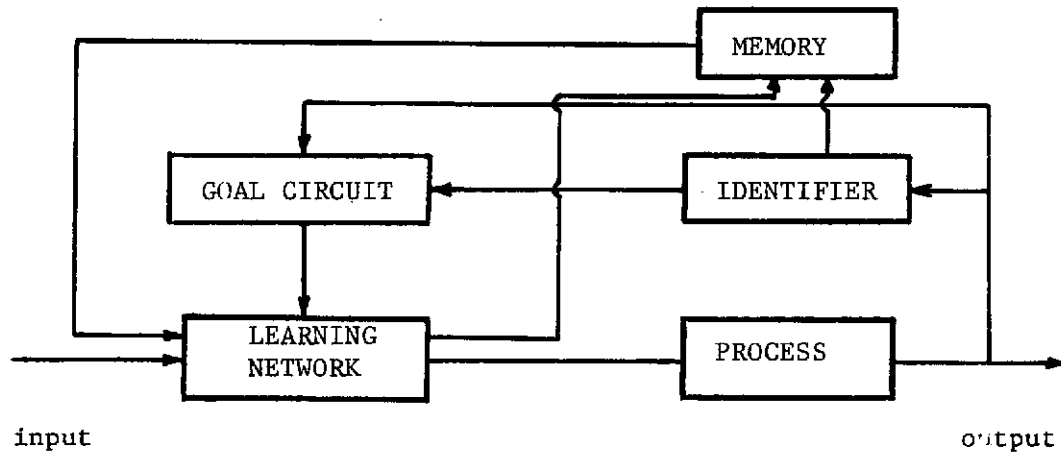
Figure 1-3: A representative learning control system.

The state-of-the-art of learning control theory and applications have been given by Gibson (1963), Fu et al. (1963), Mosteller (1963), Sklansky (1966), and Mendel and Zapalac (1968). In their studies, various learning control models have been discussed and an extensive bibliography on the subject was presented.

Objectives

The main objective of the study reported in this report is the introduction of two approaches, namely adaptive control and sequential non-linear estimation, for the identification of the state and unknown parameters of a nonlinear hydrologic system response model.

A state-space implementation of these techniques was used
for the analysis of the Prasad equation for the rainfall-
runoff process. The algorithms developed are a Kalman filtering
scheme with adaptive estimation of the error-covariance matrices
and secondly, an interated extended Kalman filter. Two rather
general computer programs were developed during the investiga-
tion and are discussed in detail.

Report Outline

Following is the outline of the contents of each chapter
in this report. Chapter 2 is devoted to a review of parameter
identification techniques used in hydrology in the past. These
techniques range from optimum search, recursive least-squares,
to a few more sophisticated optimization methods using sensi-
tivity analysis and quasi-linearization to identify unknown
time-invariant parameters in a nonlinear model. Chapter 3
presents the formulation of the identification problem, first,
as a regulator problem in stochastic control theory, and then,
as an optimum sequential estimation problem in a noisy
situation; both methods use state-space concepts. The results
of this chapter are two adaptive identification algorithms which
can track unknown parameters in a nonlinear model with noisy
observations. Finally, the implementation of the two proposed
approaches to the study of the rainfall-runoff processes in a

selected watershed and a discussion of the results are presented

in Chapter 4. Chapter 5 presents a summary of the study and

some recommendations for further investigation. The derivation

of the necessary equations mentioned in various chapters and

a description of the computer programs that were written to

implement the two approaches are presented in the attached

appendices.

CHAPTER 2

REVIEW OF PARAMETER IDENTIFICATION TECHNIQUES IN HYDROLOGY

When the structure of the model of a hydrologic system is
known, the system identification problem becomes the problem of
estimation of unknown parameters in the model. The following devel-
opment summarizes the various approaches proposed in the past to
solve the parameter identification problem in hydrology.

Optimum Search Techniques

Search techniques are very useful for engineers and hydrologists
to solve optimization problems when it is impossible or impractical
to solve them directly by analytical optimization techniques. The
only requirements are that a value of the function can be determined
for any given set of variables and that, when a global extremum is
sought, the function has no unbounded or multiple peaks (well-behaved
functions). When constraint equations on the variables are
associated with a given problem, the objective or cost function may
be augmented by penalty functions such that the extremum of the
augmented but otherwise unconstrained functions converge to the
contrained extremum of the cost function in the limit, and the usual
search techniques may be applied with little modifications. This
very useful penalty function concept was first introduced by Courant
(1943) and later modified by Carroll (1961) and by Goldstein and
Kripke (1964). The penalty argument has the defect that it may
yield fictitious solutions when the problem is ill-posed. More

discussion on this subject has been given by Beltrami (1970). A great number of search techniques have already been proposed in the past. One can easily find them in various textbooks or reports on optimization theory and control engineering (Davidson, 1959; Norris, 1960; Wilde, 1964; Leon, 1966; Pierre, 1969), or in various technical journals dealing with computational methods in optimization (Brooks, 1959; Rosenbrock, 1960; Powell, 1964; Shah et al., 1964; Fletcher, 1965; Young, 1965).

Search techniques can be grouped into two broad categories: deterministic search and random search. Techniques belonging to the latter category are superior in solving optimization problems of complex nonlinear hydrologic systems, such as rainfall-runoff processes, where discontinuities of the first derivatives and noise in the system can cause deterministic algorithms to become inefficient or to fail. Practical algorithms for random search have been proposed or discussed by Brooks (1958), Hooke (1958), Gurin and Rastrigin (1965), Gurin (1966), Schumer and Steiglitz (1968), Zakharev (1969) and Hill (1969). Good survey papers on random search methods were also given by Karnopp (1963) and White (1972). For general comparison purposes, search techniques can be divided into two classes: those techniques which utilize derivatives of the performance measure, and those techniques which do not. In general, the best sequential search techniques are more efficient than the best nonsequential ones, and the best sequential methods which utilize the gradient are more efficient than those which do not (Pierre, 1969). Search techniques, especially gradient methods, were used very often by

hydrologists for fitting parameters in a conceptual model of catchment hydrology. Ibbitt (1970) has tested eight deterministic optimum search methods and one random search method for application to hydrologic models. Among deterministic optimum search methods, he found that Rosenbrock's method was the best for the following reasons: (a) it was the most efficient among the methods tested; (b) it had an extremely flexible constraint technique; (c) its demands for computer storage were small; and (d) it could be applied to any type of objective function.

Least-Squares Procedure

Identification techniques based on least-squares procedure are applicable to both linear and nonlinear systems. The method of least-squares was initiated by Karl Friedrich Gauss in 1795 but detailed description of this method was not published by Gauss until 1809, in his book Theoria Motus Corporum Coelestium. Some basic ideas of Gauss about the method of least-squares are:

(1) minimum number of observations are required for the determination of the unknown parameters;

(2) model equations must be exact descriptions of real systems;

(3) observation errors are unknown; and

(4) the estimates of the unknown parameters must satisfy the observations in the most accurate manner possible.

The best estimates of the unknown parameters are defined as the set of values that minimizes the sum of the squares of the observation residuals. Based on the least-squares concept, a recursive least-squares approximation algorithm can be formulated as follows (Duong, 1970).

Given a time-invariant nonlinear hydrologic model

$$y_c = f(X,P) \qquad (2\text{-}1)$$

where y_c is the computed output (scalar) from the model;

X is the set of state variables of the system; and

P is the set of unknown parameters to be identified;

let y_o be the observed output from the system. The output residual

at sampled time t_i is defined to be

$$\Delta y_i = (y_o)_i - (y_c)_i \qquad (2\text{-}2)$$

The best estimates of all unknown parameters of the model

are computed at the same time by proceeding as follows:

After guessing initial values, P^o, for the unknown parameters,

one writes Δy_i in terms of the corrections Δp_j, related to the

parameters p_j, as

$$\Delta y_i = \left[\sum_{j=1}^{n} (\frac{\partial f}{\partial p_j}) \Delta p_j \right]_i \qquad (2\text{-}3)$$

where n is the total number of unknown parameters to be identified.

In the least-squares regression method, the best corrections Δp_j^*

to the a priori known values of the parameters will minimize the

function

$$J = (\Delta Y - A\Delta P^*)^T (\Delta Y - A\Delta P^*) \qquad (2\text{-}4)$$

where the superscript T denotes the transpose of a matrix, ΔY

is the output residual vector whose elements are the Δy_i, and A

is the matrix of the partial derivatives of dimension N x n (N is

the number of observations, $N \geq n$), with elements

$$a_{ij} = \left[\left(\frac{\partial f}{\partial p_j} \right)_{p_j = p_j^*} \right]_i \quad ; \quad i=1,2,\ldots,N \; ; \; j=1,2,\ldots,n \quad (2\text{-}5)$$

Minimization of the function (2-4) relative to ΔP^* leads to the unique solution, if $A^T A$ is non-singular,

$$\Delta P^* = (A^T A)^{-1} A^T \Delta Y \qquad (2\text{-}6)$$

and the best estimates p_j^* of the unknown parameters p_j are given by

$$P^* = P^O + \Delta P^* \qquad (2\text{-}7)$$

The p_j^* are then substituted back into equation (2-1) and the $(y_c)_i$ are computed again. New residuals are formed from equations (2-2) and new corrections to the p_j are computed from equation (2-6), which provides the next approximations to the parameters. One important assumption of least-squares regression technique is that the unknown parameters must be independent. If this assumption is not satisfied, i.e., the "independent" variables are not truly independent, then the correction Δp_j will not be uniquely associated with p_j and convergence of the method cannot be insured.

To account for the difference in accuracy that might exist between various measured outputs and possible relations between them, the values p_j^* (best estimates of the p_j) to minimize the function

$$J^* = (\Delta Y - A\Delta P^*)^T \; W \; (\Delta Y - A\Delta P^*) \qquad (2\text{-}8)$$

are often sought, where W is an NxN weighting matrix and it is assumed to be symmetric and known. In general, W is chosen to be the inverse of the covariance matrix of the errors. If $A^T W A$ is non-singular, the unique solution for the optimization problem

will then be given by

$$\Delta P^* = (A^T W \ A)^{-1} A^T W \Delta Y \tag{2-9}$$

The method of least-squares has been used extensively in hydrology during the last two decades for fitting parameters of parametric models of hydrologic systems. The recursive least-squares approach and the method of <u>differential correction</u> introduced to hydrologists by Snyder (1962) and later refined by Decoursey and Snyder (1969) to optimize hydrologic parameters are part of a more general theory on <u>sensitivity analysis</u> originated by Bode (1945). The mathematical background of the theory of sensitivity analysis of dynamic systems was treated by Tomovic (1962) and can be found in many textbooks in control engineering (i.e., Perkins, 1972) and therefore is omitted here. Sensitivity analysis has been used by Vemuri et al. (1967, 1969) in the analysis of ground-water systems.

Quasi-linearization

In this section, we consider the possibility of solving a non-linear hydrologic system problem by first transforming it into a related linear system problem whose solution is then modified to obtain the desired solution. Only one such indirect computational method for solving system identification problems is treated here. This technique is known as quasi-linearization and has been used by hydrologists in the identification of a non-linear hydrologic system response (Labadie, 1968) and of unconfined aquifer parameters (Yeh and Tauxe, 1971).

The quasi-linearization technique, which is often referred to as a generalized Newton-Raphson technique, was originally presented by Bellman and Kalaba (1965). Sage and Burt (1965), Sage and Smith (1966), Sage and Melsa (1971) and Graupe (1972) have examined the application of discrete and continuous quasi-lineari-zation to system identification problems. Following is the summary of the development of the continuous quasi-linearization technique.

Consider a non-linear hydrologic system described by a vector differential equation of the form

$$\dot{X} = F(X,P,U) \qquad (2-10)$$

where X is an n-dimensional state-vector;

P is an r-dimensional unknown parameter-vector to be identified; and

U is an m-dimensional input-vector.

The n+r boundary conditions of the equation (2-10) are assumed to be linear and known, and the elements of P are stationary, i.e.,

$$\dot{P} = 0 \qquad (2-11)$$

The two equations (2-10) and (2-11) can be combined to have the form

$$\dot{Z}(t) = G\left[Z(t),U(t)\right] \qquad (2-12)$$

with boundary conditions:

$$H(t)Z(t) = B(t) \qquad (2-13)$$

where Z is the new (augmented) state vector defined by

$$Z(t) = \left[x_1(t),\ldots,x_n(t);p_1,\ldots,p_r \right]^T \qquad (2\text{-}14)$$

Expanding $G\left[Z(t),U(t)\right]$ in a Taylor series about the ith estimate of the state vector, $Z^i(t)$, the $(i+1)^{th}$ estimate of the state is then given by

$$\dot{Z}^{i+1}(t) = G\left[Z^i(t),t\right] + \frac{\partial G\left[Z^i(t),t\right]}{\partial Z^i(t)} \left[Z^{i+1}(t) - Z^i(t) \right]$$

$$+ \text{ higher order terms.} \qquad (2\text{-}15)$$

Assuming that the initial estimate is close to the predicted value, and dropping the higher order terms in equation (2-15), one obtains

$$\dot{Z}^{i+1}(t) = \frac{\partial G\left[Z^i(t),t\right]}{\partial Z^i(t)} Z^{i+1}(t) + \left\{ G\left[Z^i(t),t\right] - \right.$$

$$\left. \frac{\partial G\left[Z^i(t),t\right]}{\partial Z^i(t)} Z^i(t) \right\}$$

This equation has the form:

$$\dot{Z}^{i+1}(t) = A^i(t)Z^{i+1}(t) + V^i(t) \qquad (2\text{-}16)$$

where

$$A^i(t) = \frac{\partial G\left[Z^i(t),t\right]}{\partial Z^i(t)} \qquad (2\text{-}17)$$

$$V^i(t) = G\left[Z^i(t),t\right] - \frac{\partial G\left[Z^i(t),t\right]}{\partial Z^i(t)} Z^i(t) \qquad (2\text{-}18)$$

The general solution of equation (2-16) has the form

$$z^{i+1}(t) = \phi^{i+1}(t,t_o)z^{i+1}(t_o) + Q^{i+1}(t) \qquad (2\text{-}19)$$

where $\phi^{i+1}(t,t_o)$ is the fundamental solution of equation (2-16), given by

$$\dot{\phi}^{i+1}(t,t_o) = A^i(t)\,\phi^{i+1}(t,t_o) \qquad (2\text{-}20)$$

with $\qquad \phi^{i+1}(t_o,t_o) = I \qquad (2\text{-}21)$

and $Q^{i+1}(t)$ is the particular solution of equation (2-16), satisfying

$$\dot{Q}^{i+1}(t) = A^i(t)Q^{i+1}(t) + V^i(t) \qquad (2\text{-}22)$$

with $\qquad Q^{i+1}(t_o) = 0 \qquad (2\text{-}23)$

Substituting the general solution (2-19) into the boundary conditions (2-13), one obtains

$$H(t)\left[\phi^{i+1}(t,t_o)z^{i+1}(t_o) + Q^{i+1}(t)\right] = B(t) \qquad (2\text{-}24)$$

or $\qquad H(t)\,\phi^{i+1}(t,t_o)z^{i+1}(t_o) = B(t) - H(t)Q^{i+1}(t) \qquad (2\text{-}25)$

which is of the form

$$\tilde{A}\,z^{i+1}(t_o) = \tilde{B} \qquad (2\text{-}26)$$

where $\qquad \tilde{A} = H(t)\phi^{i+1}(t,t_o) \qquad (2\text{-}27)$

$$\tilde{B} = B(t) - H(t)Q^{i+1}(t) \qquad (2\text{-}28)$$

Thus, the solution for $Z^{i+1}(t)$ has been reduced to a set of linear initial condition problems given by Eqns. (2-26) which are easily solved.

The quasi-linearization technique for system identification often requires a good initial estimate of the states in order to converge. The computational effort in identification by this technique is considerable, and the approach is limited mainly to cases where only some states (not necessarily the same states) are accessible at different times (Graupe, 1972).

Summary

In this Chapter, a review of various techniques for the identification of unknown parameters of lumped-parameter models of hydrologic systems is presented. These techniques range from optimum search methods, least-squares procedure to more sophisticated optimization techniques using sensitivity analysis and quasi-linearization to identify unknown parameters in a non-linear model of a hydrologic process. Most of these techniques are suited for the analysis of linear or non-linear deterministic time-invariant systems; some of them can be used when the inputs are imbedded in noise; but none of them are good for the analysis of non-linear time-varying systems with noisy observations. For this case, adaptive learning control techniques and nonlinear filtering approaches, using state-space concepts, might be more suitable. Two such techniques will be presented in the next Chapter.

CHAPTER 3

STATE-SPACE APPROACH FOR THE IDENTIFICATION OF NONLINEAR

HYDROLOGIC SYSTEMS FROM NOISY OBSERVATIONS

As already mentioned in Chapter 2, techniques used in the
past for the determination of the instantaneous unit hydrograph and
the identification of unknown parameters in a conceptual model of
a hydrologic process are not adequate. The main reasons for this
come from the fact that the input and output hydrologic data are
imbedded in noise, the hydrologic processes are more or less non-
linear, and the changing of the environmental conditions with time
may affect the model output. In this chapter, a new approach using
state-space concept is investigated, and techniques for optimal
adaptive identification of unknown parameters and of the control
inputs are presented.

Problem Formulation

A general nonlinear lumped-parameter model for a hydrologic
system can be represented by the following vector differential
equation

$$\dot{X}(t) = f\left[X(t), U(t), t\right] + w(t) \qquad (3-1)$$

where $X(t)$ is the actual (nx1) state-vector of the system, $U(t)$ is
the noisy forcing input (i.e., rainfall), $f(\cdot)$ is the nonlinear
mapping from E^n into E^n, and $w(t)$ is an (nx1) noise term which is
assumed to be a zero-mean white noise process with covariance

matrix

$$E\left\{w(t)w(\tau)^T\right\} = W\,\delta(t-\tau) \tag{3-2}$$

where $\delta(\cdot)$ is the Dirac-delta function and superscript T indicates transpose.

In most cases one cannot measure the state directly and precisely, but one can measure a vector Y which represents the output (i.e., runoff) of the system, and is related to the state by the following equation

$$Y(t) = h\left[X(t)\right] + v(t) \tag{3-3}$$

where $h(\cdot)$ is the nonlinear mapping from E^n into E^1, and $v(t)$ is the observation noise which is again assumed to be a zero-mean white noise process with covariance matrix

$$E\left\{v(t)v(\tau)^T\right\} = V\,\delta(t-\tau) \tag{3-4}$$

The identification problem can now be stated as follows:

Given the system model (3-1), the observation equation (3-3), the noise characteristics defined by Eqns. (3-2) and (3-4), and the set of noisy input-output pairs $\left\{U_i\right\}$ and $\left\{Y_i\right\}$; from the initial conditions of the state and the estimation-error covariance matrix, find the best estimate of the state of the process under some optimality criterion (i.e., minimization of the mean-square-error of the estimate).

Two approaches for solving this problem are presented. In the first approach the given problem is treated as a regulator problem in stochastic control theory. In the second approach a non-linear

filter will be used to estimate the state and unknown parameters of the system under noisy observations.

A Linear Stochastic Control Problem

In order to apply linear stochastic control theory to the identification problem formulated above, one must linearize the given nonlinear system equations and the observation model to get the linearized model of the system.

Let the nominal values of the state and the input be denoted by $X^*(t)$ and $U^*(t)$ respectively and the deviations of the actual state and the input from nominal values be represented by

$$
\begin{aligned}
\delta x(t) &= X(t) - X^*(t) \\
\delta u(t) &= U(t) - U^*(t)
\end{aligned}
\tag{3-5}
$$

The linearized equations of the system about the nominal values are then given by

$$
\dot{\delta x}(t) = F(t)\ \delta x(t) + G(t)\ \delta u(t) + e(t) \tag{3-6}
$$

where the elements of the matrices $F(t)$ and $G(t)$ are partial derivatives of f with respect to $X(t)$ and to $U(t)$ evaluated about the nominal values,

$$
F(t) = \left[\frac{\partial f(X,U,t)}{\partial X} \right]_{X=X^*,U=U^*} \tag{3-7}
$$

$$
G(t) = \left[\frac{\partial f(X,U,t)}{\partial U} \right]_{X=X^*,U=U^*}
$$

and $e(t)$ denotes random disturbances which are used to model such effects as unknown dynamics and truncation errors. It is assumed

that e(t) is a zero-mean white noise process with covariance matrix

$$E\left\{ e(t)e(\tau)^T \right\} = N\delta(t - \tau) \qquad (3-8)$$

From now on, the vectors $\delta x(t)$ and $\delta u(t)$ will be referred to as the state and the control input, respectively, for the linear perturbation model.

For calculations with digital computers, Eqn.(3-6) has to be converted into a difference equation. To that end one needs to have a state transition matrix $\Phi(t_{n+1}, t_n)$ which satisfies the homogeneous part of the differential equation (3-6); i.e., $\Phi(t_{n+1}, t_n)$ is defined by

$$\dot{\Phi}(t_{n+1}, t_n) = F(t_n)\ \Phi(t_{n+1}, t_n) \qquad (3-9)$$

with the initial condition

$$\Phi(t_n, t_n) = I \quad \text{for all } t_n \qquad (3-10)$$

where I denotes the unit matrix having the same dimension as the matrix F.

The difference equation of the system, derived from the differential equation (3-6) for the time interval (t_n, t_{n+1}), will have the following form

$$\delta x(t_{n+1}) = \Phi(t_{n+1}, t_n)\delta x(t_n) + \int_{t_n}^{t_{n+1}} \Phi(t_{n+1}, \tau)G(\tau)\delta u(t_n)d\tau$$

$$+ \int_{t_n}^{t_{n+1}} \Phi(t_{n+1}, \tau)e(\tau)d\tau$$

or, in abbreviated notations,

$$\delta x_{n+1} = \Phi_{n+1,n} \; \delta x_n \; + \; \Gamma_{n+1,n} \; \delta u_n \; + \; w_{n+1} \tag{3-11}$$

The covariance matrix of the process noise w_n then becomes

$$E \{w_n w_m^T\} = \left[\int_{t_n}^{t_{n+1}} \Phi(t_{n+1},\tau) \; N \; \Phi^T(t_{n+1},\tau) \; d\tau \right] \delta_{nm}$$

$$= Q_{n+1} \; \delta_{nm} \tag{3-12}$$

where δ_{nm} is the Kronecker delta.

Similarly, linearizing the observation equation (3-3) about the nominal states and converting the result into a difference equation, one obtains:

$$\delta y_n = H_n \; \delta x_n \; + \; v_n \tag{3-13}$$

where δy_n represents the deviation of the measured output from the nominal value, i.e.,

$$\delta y_n = Y_n - Y_n^* \tag{3-14}$$

the elements of the mapping matrix H_n are partial derivatives of h with respect to X evaluated at the nominal values:

$$H_n = \left[\frac{\partial h_n(X)}{\partial X} \right]_{X=X^*} \tag{3-15}$$

H_n is often called the observation matrix; and v_n is the observation noise which is again assumed to be a zero-mean white noise process with covariance matrix

$$E \{v_n v_m^T\} = R_n \; \delta_{nm} \tag{3-16}$$

In general, the Q_n and R_n are considered to be positive definite and assumed to be given in advance. For complex systems, Q_n might be very hard to get. In this case, technique for adaptive estimation of Q_n is very useful and will be discussed later in subsequent sections.

The control vector δu in this optimal stochastic control problem is to be selected so that the performance index

$$E \{J_N\} = E\left\{ \sum_{j=1}^{N} (\delta x_j^T A_j \ \delta x_j + \delta u_{j-1}^T B_{j-1} \ \delta u_{j-1}) \right\} \qquad (3\text{-}17)$$

is minimized; where N is the total number of measurements made during the identification period, and the matrices A_j and B_j are arbitrary non-negative definite, symmetric weighting matrices. An approximate choice of these matrices must be made to obtain good results for the given problem. A choice that often turns out to be quite reasonable is (Bryson and Ho, 1969)

$$A_j^{-1} = n(t_f - t_o) \times \text{maximum acceptable value of diag } \delta x(t) \delta x^T(t);$$

$$B_j^{-1} = m(t_f - t_o) \times \text{maximum acceptable value of diag } \delta u(t) \delta u^T(t);$$

where n and m are dimensions of the state-vector and the control input vector, respectively.

The Optimum Controller

In the above formulation, the equations (3-11),(3-12),(3-13) and (3-16) represent a linear time-varying system with Gaussian statistics; one can separate the solution of this optimum stochastic control problem into a deterministic optimum control problem and an optimum

estimation problem (Sorenson, 1968). The proof of the Separation Principle is given in Appendix A. The optimum stochastic control law is described by

$$\delta \hat{u}_{N-k-1} = - \Lambda_{N-k} \Phi_{N-k,N-k-1} \delta \hat{x}_{N-k-1} \tag{3-18}$$

where $\Lambda_{N \cdot k}$ is obtained from the solution of the deterministic control problem, and $\delta \hat{x}_{N-k-1}$ is the optimum estimate of the state δx_{N-k-1} obtained by solving the optimum estimation problem. The Λ_{N-k} are given by the following recursive formulas:

$$\Lambda_{N-k} = \left[\Gamma^T_{N-k,N-k-1} L'_{N-k} \Gamma_{N-k,N-k-1} + B_{N-k-1} \right]^{-1} \Gamma^T_{N-k,N-k-1} L'_{N-k} \tag{3-19}$$

$$L'_{N-k} = \Phi^T_{N-k+1,N-k} L_{N-k+1} \Phi_{N-k+1,N-k} + A_{N-k} \tag{3-20}$$

$$L_{N-k} = L'_{N-k} - L'_{N-k} \Gamma_{N-k,N-k-1} \Lambda_{N-k} \tag{3-21}$$

$$k = 0,1,2,\ldots,N-1.$$

To start the iteration process, one can assume that $L_{N+1} = 0$ and then proceed the calculations backwards in time to the initial control time t_o. The L_{N-k} and Λ_{N-k} are often called the control cost matrix and the control gain matrix, respectively.

Identification of State-Variables

The optimum feedback control law defined by Eqn. (3-18) depends on the optimum estimate of the state at each stage of the process under study. This optimum estimate can be obtained through

the well-known Kalman filter (Kalman, 1960) which is the best

linear minimum variance estimator since the estimate $\delta\hat{x}_n$, defined

as the conditional expectation of δx_n given the available data

set Z_n at time t_n, i.e., $\delta\hat{x}_n = E\{\delta x_n \mid Z_n\}$, is chosen to minimize

the mean-square-error

$$E\{(\delta\hat{x}_n - \delta x_n)^T(\delta\hat{x}_n - \delta x_n)\} = \text{trace } E\{(\delta\hat{x}_n - \delta x_n)(\delta\hat{x}_n - \delta x_n)^T\}$$

$$(3-22)$$

The derivation of the standard Kalman filter is omitted here

since one can easily find it in any textbook on estimation theory

(Sorenson, 1966; Sage and Melsa, 1971). Following is a summary of

the filter algorithms for discrete cases where the state equation

and observation model are given in the form of Eqns. (3-11) and

(3-13), with noise covariance matrices defined by Eqns. (3-12) and

(3-16).

one-stage
prediction
$$\delta x'_{n+1/n} = \Phi_{n+1,n}\delta\hat{x}_n + \Gamma_{n+1,n}\delta\hat{u}_n \tag{3-23}$$

a priori
variance
$$P'_{n+1/n} = \Phi_{n+1,n}P_n\Phi^T_{n+1,n} + Q_{n+1} \tag{3-24}$$

Kalman
gain
$$K_{n+1} = P'_{n+1/n}H^T_{n+1}\left[H_{n+1}P'_{n+1/n}H^T_{n+1} + R_{n+1}\right]^{-1} \tag{3-25}$$

Filter
algorithm
$$\delta\hat{x}_{n+1} = \delta x'_{n+1/n} + K_{n+1}\left[\delta y_{n+1} - H_{n+1}\delta x'_{n+1/n}\right] \tag{3-26}$$

a posteriori
variance
$$P_{n+1} = (I - K_{n+1}H_{n+1})P'_{n+1/n} \tag{3-27}$$

The initial conditions δx_o and P_o must be given, also the exact values of Q_n and R_n must be known. After each observation, the optimum estimate of the state X is then given by

$$X = X^* + \delta \hat{x} \qquad (3-28)$$

The most important properties of a Kalman filter can be summarized as follows:

(1) The filter estimates are all variables of the state vector in the least-mean-square-error sense.

(2) The estimation is based upon statistical data of all error sources and is completely carried out in the time domain.

(3) The filter formulae satisfy minimum variance criteria for all problem parameters.

(4) The formulae implemented are recursive. This means that the optimum estimate at the present time can be computed from the previous estimate and the current observation without recourse to earlier estimates or observations.

(5) The recursive formulae are well suited to digital computers.

The use of the Kalman filter to estimate the unknown parameters in a lumped-parameter model of a hydrologic system is discussed in the next section, and the application of this approach to the analysis of rainfall-runoff is presented in Chapter 4.

Parameter Identification

Consider an unknown parameter $\bar{a}$ which is slowly varying in time. One could model its behavior satisfactorily by the random walk model

$$a_{n+1} = \bar{a}_n + (w_{\bar{a}})_n \qquad (3-29)$$

where $(W_{\bar{a}})_n$ is a zero mean white noise sequence with covariance matrix $(Q_{\bar{a}})_n$. Equation (3-29) combined with Eqn. (3-11) gives an augmented state equation of the form

$$\begin{bmatrix} \delta x \\ \bar{a} \end{bmatrix}_{n+1} = \begin{bmatrix} \Phi & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \delta x \\ \bar{a} \end{bmatrix}_n + \begin{bmatrix} \Gamma \\ 0 \end{bmatrix} \delta u_n + \begin{bmatrix} w \\ w_{\bar{a}} \end{bmatrix}_n \qquad (3-30)$$

The observation model is also modified as

$$\delta y_n = \begin{bmatrix} H_n & \vdots & 0 \end{bmatrix} \begin{bmatrix} \delta x \\ \bar{a} \end{bmatrix}_n + v_n \qquad (3-31)$$

One could then apply the Kalman filtering algorithms mentioned above to estimate the unknown parameter along with values of the previous state variables of the system, using the augmented state equation and observation model defined by Eqns. (3-30) and (3-31).

If the secular variation in a parameter a_n is rapid it can be decomposed into the product of a known, non-singular and rapidly time-varying transformation matrix, T_n, and an unknown but fixed or only slowly varying parameter $\bar{a}_n$. Thus,

$$a_n = T_n \bar{a}_n$$

and from Eqn. (3-29):

$$a_n = \Phi_a a_{n-1} + \Gamma_a (w_{\bar{a}})_{n-1} \qquad (3-32)$$

where $\Phi_a = T_n T_{n-1}^{-1}$ and $\Gamma_a = T_n$. One then uses Eqn. (3-32) to form an augmented state equation and new observation model, and proceeds as mentioned earlier.

Estimation of Error-Covariance Matrices

In most practical situations, complete knowledge of the
hydrologic process and measurement noise statistics is hard to get.
The use of wrong a priori statistics in the design of a Kalman filter
can lead to large estimation errors or even to a divergence of
errors. The covariance matrix becomes unrealistically small and
optimistic; the filter gain thus becomes small, and subsequent
measurements are ignored. The state and its estimate then diverge,
due to model errors in the filter. Analyses of error divergence
in the Kalman filter have been performed by Heffes (1966), Schlee et
al. (1967), Nishimura (1967), Price (1968) and Fitzgerald (1971).

Several approaches have been proposed for preventing filter
divergence. Horlick and Sward (1965) investigated the technique of
filter reset to keep the diagonal elements of the state covariance
matrix above a specified value. Peschon and Larson (1965) proposed
to include a random variable at the model input to account for any
unrealistic assumptions made about the system model. Schmidt (1967)
suggested two methods. In one method an estimate of the state is
computed as a linear combination of the estimate given all prior
data with the estimate given no prior data. Past information is
degraded. The other method assumes a priori lower bounds on certain
projections of the covariance matrix. Schmidt et al. (1968) also
designed the modified Kalman filter equations by including an addi-
tive gain matrix to the conventional gain matrix of the Kalman
filter to prevent divergence of the filter resulting from unmodeled
and computational errors. Recently, Tarn and Zaborszky (1970)

used a scalar weighting factor s for the observation error-variance
and came up with a non-diverging filter which reduces to the
standard Kalman filter for s = 1.

The most interesting approach to prevent filter divergence
is probably to cover model errors with noise and adaptively
estimate the noise levels. The purpose of an adaptive filter is,
then, to reduce or bound the noise by adapting the Kalman filter
to the real data. A number of approaches to adaptive filtering has
been proposed in the past few years. Good survey papers in this
area were presented by Sage and Husa (1969), Weiss (1970) and
Mehra (1972). Following are some adaptive schemes for estimating
the observation error covariance matrix R and the process error
covariance matrix Q.

1/Estimation of R

One approach to reduce the effect of earlier measurement as
new measurements are included is to replace the finite time
averaging operation by exponentially weighted past averages. This
technique, applied to the estimation of R_n, is expressed as

$$\hat{R}_n = (1-W_n)\hat{R}_{n-1} + W_n R'_n \qquad (3-33)$$

where R'_n is the predicted value of R_n at the nth measurement; i.e.,

$$R'_n = (\delta y_n - H_n \delta x'_{n/n-1})(\delta y_n - H_n \delta x'_{n/n-1})^T \qquad (3-34)$$

This method has been applied successfully to satellite tracking
problems. R is assumed to be constant during the sample period

$t_o \leq t \leq t_N$. For these problems, Tapley and Born (1971) proposed

the following expression for the weighting coefficient W_n:

$$W_n = (n-1)(n-2) \ldots (n-k)/n^{k+1} \qquad (3-35)$$

in which k is an integer. W_n is zero for all $n \leq k$, and as $n \to \infty$,

W_n approaches $1/n$. Hence, for $n \leq k$ the value of R is not changed

from the a priori value. The choice of k will depend upon how well

R_o is known, i.e., the more accurate R_o the larger k may be.

In the case of a rainfall-runoff process, there are typically

few data points obtained for one storm. By using the weighting co-

efficient defined by Eqn. (3-35) one might not get the correct value

for R. At the end of the observation period, the estimate $\hat{R}_n$ might

still depend very much upon its initial guess. For this case, it

is suitable to choose the weighting coefficient W_n as

$$W_n = \frac{\alpha}{1-(1-\alpha)^n} \qquad (3-36)$$

where α is a positive constant, normally less than 0.1 (Young, 1965).

This choice of the weighting coefficient also assures that new data

continues to have some effect on the estimate of R as long as the

observation of the process is still going on.

2/Estimation of Q

Let $r_{n+1/n}$ denote the predicted measurement residual, i.e.,

$$r_{n+1/n} \triangleq \delta y_{n+1} - H_{n+1} \, \delta x'_{n+1/n} \qquad (3-37)$$

one then has

$$E\{r_{n+1/n}r^T_{n+1/n}\} = H_{n+1} P'_{n+1/n} H^T_{n+1} + R_{n+1}$$

$$= H_{n+1}\Phi_{n+1,n}P_n\Phi^T_{n+1,n}H^T_{n+1} + H_{n+1}Q_nH^T_{n+1} + R_{n+1}$$

$$(3-38)$$

Assume Q_n has the form

$$Q_n = q\, S_n S^T_n$$

where q is a scalar; then, for the case of single output, Eqn. (3-38) can be written as

$$E\{r^2_{n+1/n}\} = H_{n+1}\Phi_{n+1,n}P_n\Phi^T_{n+1,n}H^T_{n+1} + q\, H_{n+1}S_nS^T_nH^T_{n+1} + R_{n+1} \quad .$$

$$(3-39)$$

To determine q, Jazwinski (1971) suggested that one finds the q value which satisfies the requirement

$$\max_{q \geq 0} p\{r_{n+1/n}\} \qquad (3-40)$$

or, when the probability density p is normal and has zero-mean,

$$r^2_{n+1/n} = E\{r^2_{n+1/n}\} \quad . \qquad (3-41)$$

Equation (3-41) represents the consistency requirement for the estimates of q. From Eqns. (3-39) and (3-41), the following adaptive scheme is derived:

$$\tilde{q}_{n+1} = \begin{cases} \dfrac{r^2_{n+1/n} - H_{n+1}\Phi_{n+1,n}P_n\Phi^T_{n+1,n}H^T_{n+1} - R_{n+1}}{H_{n+1}S_nS^T_nH^T_{n+1}} & , \text{ if positive} \\ \\ 0 & , \quad \text{otherwise} \end{cases}$$

$$(3-41)$$

Since the estimate of q is based on only one residual, the
estimator (3-41) will respond to (large) measurement noise samples
as well as large residuals caused by model errors. One approach to
remedy this defect is to apply the exponential smoothing technique
described in the previous section, i.e.,

$$\hat{q}_{n+1} = (1 - W_{n+1}) \hat{q}_n + W_{n+1} \tilde{q}_{n+1} \qquad (3-42)$$

An Adaptive Control Algorithm

From the material in the above sections, the following
algorithm can be set up to estimate the values of the state variables
and unknown parameters of a time-varying system in a noisy situation.

(1) Set initial conditions:

$$\delta x'_{1/0} = \delta \hat{x}_o = \mu_o \text{ (usually set equal to zero)}$$

$$P'_{1/0} = P_o = V_o$$

(2) Compute (off-line) the optimum control law by Eqns. (3-19),
(3-20) and (3-21) from nominal values of the states.

(3) Set $n = 1$. Start the sequential estimation process.

(4) Adjust the observation error covariance matrix R_n by
Eqns. (3-33) and (3-36).

(5) Compute the filter gain K_n by Eqn. (3-25).

(6) Process the observation δy_n by Eqn. (3-26) to get the best
estimate $\delta \hat{x}_n$ of the state.

(7) Replace $\delta x'_{n/n-1}$ by $\delta \hat{x}_n$ and return to step (4) to re-estimate
R_n and $\delta \hat{x}_n$.

(8)   Continue step (7) until one gets stable values for R_n and
$\delta\hat{x}_n$. Hence,

$$\hat{X}_n = X_n^* + \delta\hat{x}_n$$

(9)   Compute the a posteriori error variance P_n of the estimates
by Eqn. (3-27).

(10) Adjust the model-error covariance matrix Q_n by Eqns. (3-41)
and (3-42).

(11) Compute the predicted state $\delta x'_{n+1/n}$ by Eqn. (3-23).

(12) Compute the predicted error variance $P'_{n+1/n}$ of the estimates
by Eqn. (3-24).

(13) Set $n = n+1$ and return to step (4).

Three remarks can now be made about the above algorithm.

First, during computation one must somehow design a technique
to maintain the positive definiteness and symmetry of the matrix P_{n+1}
to avoid filter divergence. One way to maintain the desired property
of the matrix P_{n+1} is to replace Eqn. (3-27) by an equivalent form

$$P_{n+1} = \begin{bmatrix} I - K_{n+1} \ H_{n+1} \end{bmatrix} P'_{n+1/n} \begin{bmatrix} I - K_{n+1} \ H_{n+1} \end{bmatrix}^T + K_{n+1} \ R_{n+1} \ K_{n+1}^T \quad . (3-43)$$

This equation is the sum of two symmetric positive definite matrices,
and, when these are added, the sum will also be positive definite;
therefore it is better conditioned for numerical computations than
the previously mentioned expression (3-27).

Secondly, one could improve the above algorithm by correcting
the nominal values after each new estimate of the state was obtained,
i.e., one takes the new nominal value $\tilde{X}^*$ as

$$\tilde{X}_n^* = X_n^* + \delta\hat{x}_n = \hat{X}_n \qquad (3\text{-}44)$$

the new deviation of the state at stage n becomes

$$\tilde{\delta\hat{x}}_n = 0 \qquad (3\text{-}45)$$

This process is often called <u>trajectory rectification</u>. As a consequence of relinearization, large initial estimation errors are not allowed to propagate through time, and, therefore, the linearity assumptions are less likely to be violated.

Lastly, in order for the above algorithm to be called adaptive control, at each iteration step the new estimate of the state must somehow be used to alter the control law. However, the recursive equations used to compute the control history run backward in time. To update such a control law, the gains over the entire control interval would have to be recomputed each time a new estimate of the state was made available. The computational requirements of such a technique are enormous. A sub-optimum control law can be obtained as follows: first, the control law is computed off-line from nominal value of the state and the values of the matrix L are stored. Then, as each new estimate of the state is obtained, the control is updated by using new values of the matrices Φ and Γ. The adaptive control algorithm can now be summarized as illustrated in Figure 3-1.

When applying the above algorithm to the determination of unknown parameters in rainfall-runoff processes, the disturbance from rainfall data will be compensated by the optimum value of the control term, $\Gamma_{n+1,n} \, \delta\hat{u}_n$, computed at each observation from the
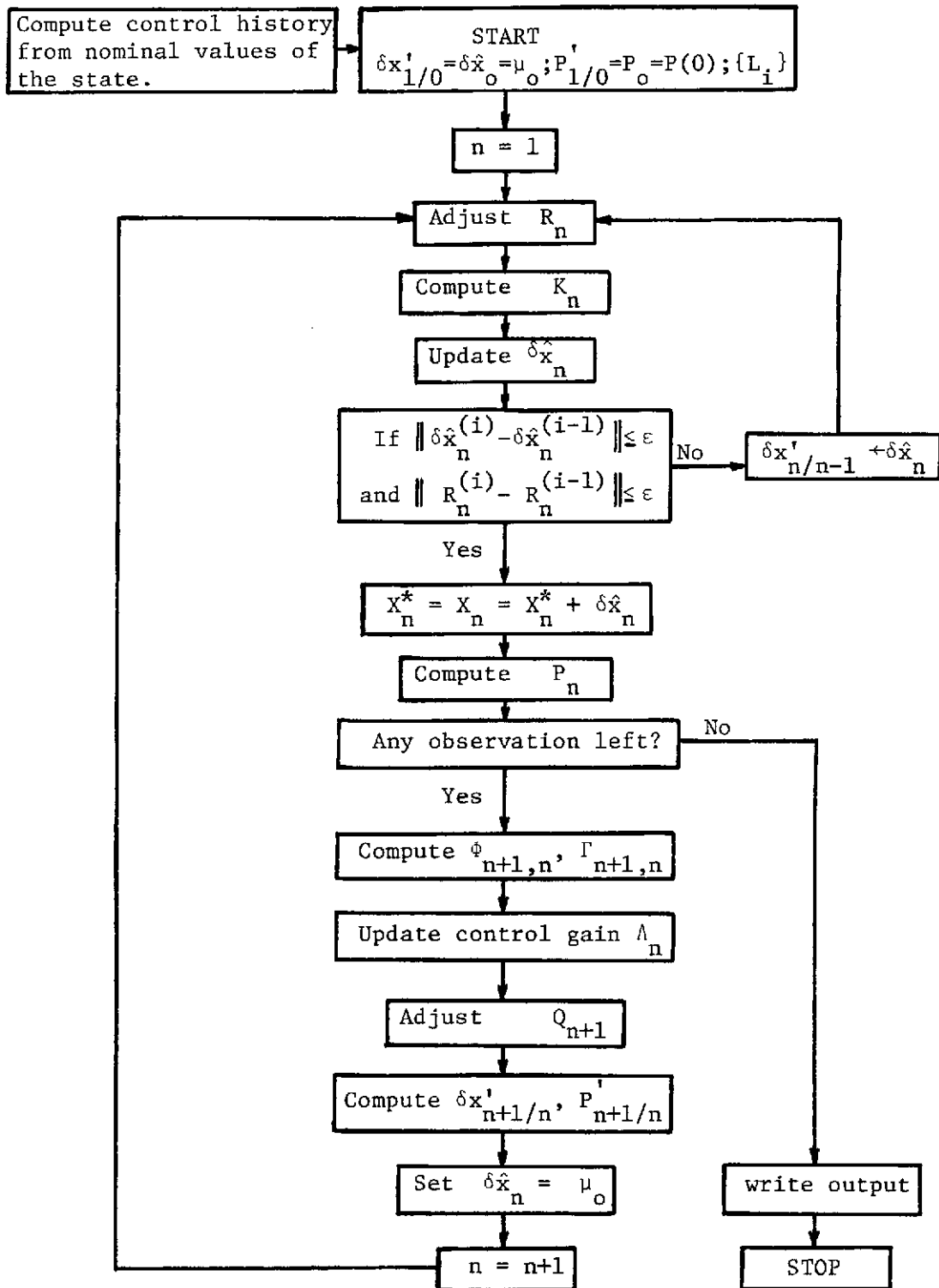
Figure 3-1: Block diagram of Adaptive Control Algorithm.

estimate of the state of the system and from Eqns. (3-18), (3-19), (3-20) and (3-21). The adjustment for noise in rainfall data by this adaptive control approach is effective if the system is not highly non-linear, since the estimation of the best value for the state and the derivation of the optimum control law are all based on the validity of the linearized model. For this reason, another method for the identification of unknown parameters in nonlinear lumped-parameter models of hydrologic systems, using nonlinear estimation techniques, is presented in the following section. The comparison of results of the two approaches, when applying to the identification of rainfall-runoff processes, is discussed in Chapter 4.

A Nonlinear Filtering Problem

A lumped-parameter model for the rainfall-runoff system is represented by Eqns. (3-1),(3-2),(3-3) and (3-4). Since rainfall data are noisy, one can approximate the correct rainfall input U(t) in the model (3-1) by

$$U(t) = U^*(t) + w_{U^*}(t) \qquad (3-46)$$

where $U^*(t)$ represents the actual rainfall data, $w_{U^*}(t)$ is the rainfall measurement noise which is assumed to be a zero-mean white noise process with covariance matrix

$$E \{w_{U^*}(t) \ w_{U^*}(\tau)^T\} = W_{U^*}\delta(t - \tau)$$

The process equation is then written as

$$\dot{X}(t) = f \left[X(t),U^*(t),t\right] + \tilde{w}(t) \qquad (3-47)$$

where $\tilde{w}(t)$ is the combined noise term in the model which includes errors due to unknown dynamics and noise in input data. The form of $\tilde{w}(t)$ depends on the form of $w(t)$ and the relation of $U^*(t)$ in the model equation. Assume that $\tilde{w}(t)$ is also a zero-mean white noise process, having new covariance matrix defined by

$$E \ \{\tilde{w}(t) \ \tilde{w}(\tau)^T\} = \ \tilde{W} \ (t - \tau) \qquad (3\text{-}48)$$

Since $U^*(t)$ is a given scalar at time t, Eqn. (3-47) can be written simply as

$$\dot{X}(t) \ = \ f \left[X(t),t\right] \ + \ \tilde{w}(t) \qquad (3\text{-}49)$$

The identification problem of the rainfall-runoff system then becomes the estimation of the state of a nonlinear system defined by the Eqns. (3-49),(3-48),(3-3) and (3-4), given the initial conditions $\hat{X}(0)$ and $E\{X(0)X(0)^T\} = P(0)$.

Optimal estimation in the nonlinear case involves the solution of an infinite-dimensional process, as shown by Kushner (1967). Since the computational aspects of the truly optimum nonlinear filter are prohibitive, several approaches to sub-optimal filtering have been proposed in the past few years (Friedland and Bernstein, 1966; Schwartz and Stear, 1968; Athans et al.; 1968; Sage and Melsa, 1971). These algorithms can be roughly subdivided into the so-called first-order filters and higher-order filters with increasing complexity and computational requirements. Because in hydrology the estimate of the state of a system is usually not required to be highly accurate, only the extended (first-order) Kalman filter is considered here.

This filter is simple but effective and has been used very often in similar problems in the aerospace field.

The Extended Kalman Filter

This filter is the result of the relinearization procedure mentioned in the previous section. If, initially one linearizes the model equation (3-49) about $\hat{X}(t_o)$, then

$$\delta \hat{x}(t_o) = 0 \quad .$$

The predicted deviation, given by

$$\delta x'(t_1/t_o) = \Phi(t_1, t_o) \, \delta \hat{x}(t_o)$$

is therefore equal to zero.

Since one subsequently linearizes about $\hat{X}(t_1)$,

$$\delta \hat{x}(t_1) = 0$$

it follows that $\quad \delta x'(t_2/t_1) = 0 \quad .$

Thus, in general,

$$\delta x'(t/t_n) = 0 \qquad t_n \le t \le t_{n+1} \qquad \text{for all } n \quad ; \qquad (3\text{-}50)$$

that is, the best estimate of the state between observations is the nominal value of the state. Accordingly, one has

$$\dot{X}'(t/t_n) = f\left[X'(t/t_n), t\right] \qquad . \qquad (3\text{-}51)$$

Since $\qquad \delta \hat{x}_{n+1} = \hat{X}_{n+1} - X'_{n+1/n}$

in view of relinearization, and using Eqn. (3-50), the correction
to the estimate at an observation (Eqn. (3-26)) leads to

$$\hat{X}_{n+1} = X'_{n+1} + K_{n+1}\left[Y_{n+1} - h(X'_{n+1/n}, t_{n+1})\right] \qquad (3-52)$$

Thus, the extended Kalman filter can be summarized in the following
operations:

A priori
estimate
$$X'_{n+1/n} = \hat{X}_n + \int_{t_n}^{t_{n+1}} f\left[X(t, t_n), t\right] dt \qquad (3-53)$$

$$P'_{n+1/n} = \Phi_{n+1,n}(\hat{X}_n) P_n \Phi^T_{n+1,n}(\hat{X}_n) + \tilde{Q}_{n+1} \qquad (3-54)$$

A posteriori
estimate
$$\hat{X}_{n+1} = X'_{n+1/n} + K_{n+1}(X'_{n+1/n})\left[Y_{n+1} - h_{n+1}(X'_{n+1/n})\right]$$
$$(3-55)$$

$$P_{n+1} = \left[I - K_{n+1}(X'_{n+1/n}) H_{n+1}(X'_{n+1/n})\right] \cdot P'_{n+1/n}$$

$$\cdot \left[I - K_{n+1}(X'_{n+1/n}) H_{n+1}(X'_{n+1/n})\right]^T$$

$$+ K_{n+1}(X'_{n+1/n}) R_{n+1} K^T_{n+1}(X'_{n+1/n}) \qquad (3-56)$$

Kalman
gain
$$K_{n+1}(X'_{n+1/n}) = P'_{n+1/n} H^T_{n+1}(X'_{n+1/n})$$

$$\cdot \left[H_{n+1}(X'_{n+1/n}) P'_{n+1/n} H^T_{n+1}(X'_{n+1/n})\right.$$

$$\left. + R_{n+1}\right]^{-1} \qquad (3-57)$$

The matrices Φ and H are those of the linearized system defined by

$$\delta x_{n+1} = \Phi_{n+1,n}(\hat{X}_n) \delta x_n + \tilde{w}_n \qquad (3-58)$$

$$\delta y_n = H_n(\hat{X}_n)\, \delta x_n + v_n \qquad\qquad (3\text{-}59)$$

The Iterated Extended Kalman Filter

To improve the performance of the extended Kalman filter, one can use the technique derived by Denham and Pines (1966) to reduce the effect of measurement function (h) nonlinearity which occurs very often in hydrology when output data from a hydrologic system are imbedded in noise. This technique is a local iteration algorithm based on the relinearization about the new estimate.

Consider the estimator, Eqn. (3-55), in the extended Kalman filter. It was obtained by evaluating the correction to the estimate

$$\delta\hat{x}_{n+1} = \delta x'_{n+1/n} + K_{n+1}(X^*_{n+1})\left[\delta y_{n+1} - H_{n+1}(X^*_{n+1})\,\delta x'_{n+1/n}\right] \qquad (3\text{-}60)$$

about $X^*_{n+1} = X'_{n+1/n}$. Then, processing the observation Y_{n+1} via Eqn. (3-55) one gets $\hat{X}_{n+1}$. Assuming that $\hat{X}_{n+1}$ is closer to the true state than $X'_{n+1/n}$, one then would expect to get a better result by relinearizing the system equation about $\hat{X}_{n+1}$ and recomputing the estimate. Thus, the iterated extended Kalman filter consists of Eqns. (3-53) – (3-57) with Eqn. (3-55) replaced by

$$\hat{x}^{(i+1)}_{n+1} = X'_{n+1/n} + K_{n+1}(\hat{x}^{(i)}_{n+1})\left[Y_{n+1} - h_{n+1}(\hat{x}^{(i)}_{n+1})\right.$$
$$\left. - H_{n+1}(\hat{x}^{(i)}_{n+1})\{X'_{n+1/n} - \hat{x}^{(i)}_{n+1}\}\right] \qquad i = 1,2,\ldots,\ell \quad (3\text{-}61)$$

where $\hat{x}^{(1)}_{n+1} = X'_{n+1/n}$. This local iteration terminates when there is no significant difference between consecutive iterations. The covariance matrix in Eqn. (3-56) is then computed based on the last estimate

of X_{n+1}.

Combining the sequential estimation of error-covariance matrices mentioned in the previous section with the above iteration for the improvement of the estimate of the state, the following algorithm is formed as illustrated in Figure 3-2.

The local iteration process mentioned in this algorithm is designed for measurement nonlinearities and does not improve the previous nominal value chosen for the state on the interval $|t_n, t_{n+1})$. To include the nominal value in the iteration loop, one needs to smooth back the new estimate at t_{n+1} to t_n to get an improved nominal value for prediction to t_{n+1}. The linearized smoother is given by Jazwinsky (1970) as

$$\delta \bar{x}_{n/n+1} = \delta \hat{x}_n + S_n(\xi_i) \left[\delta \hat{x}_{n+1} - \delta x'_{n+1/n} \right]$$

with
$$S_n(\xi_i) = P_n \, \phi^T_{n+1,n}(\xi_i) \left[P'_{n+1/n} \right]^{-1} \qquad (3\text{-}62)$$

or, using the smoothed estimate as ξ_{i+1},

$$\xi_{i+1} = X_n + S_n(\xi_i) \left[X^{(i)}_{n+1} - X'_{n+1/n} \right] \qquad (3\text{-}63)$$

The iteration starts with $\xi_1 = \hat{X}_n$ and $\hat{X}^{(1)}_{n+1} = X'_{n+1/n}$.

This iterated linear filter-smoother, as named by Jazwinsky (1970), was apparently first derived by Wishner et al. (1968) in a different way, and was called by these authors a "single state iteration filter". Although the performance of this algorithm was shown (Wishner et al., 1968) to be better than the iterated extended
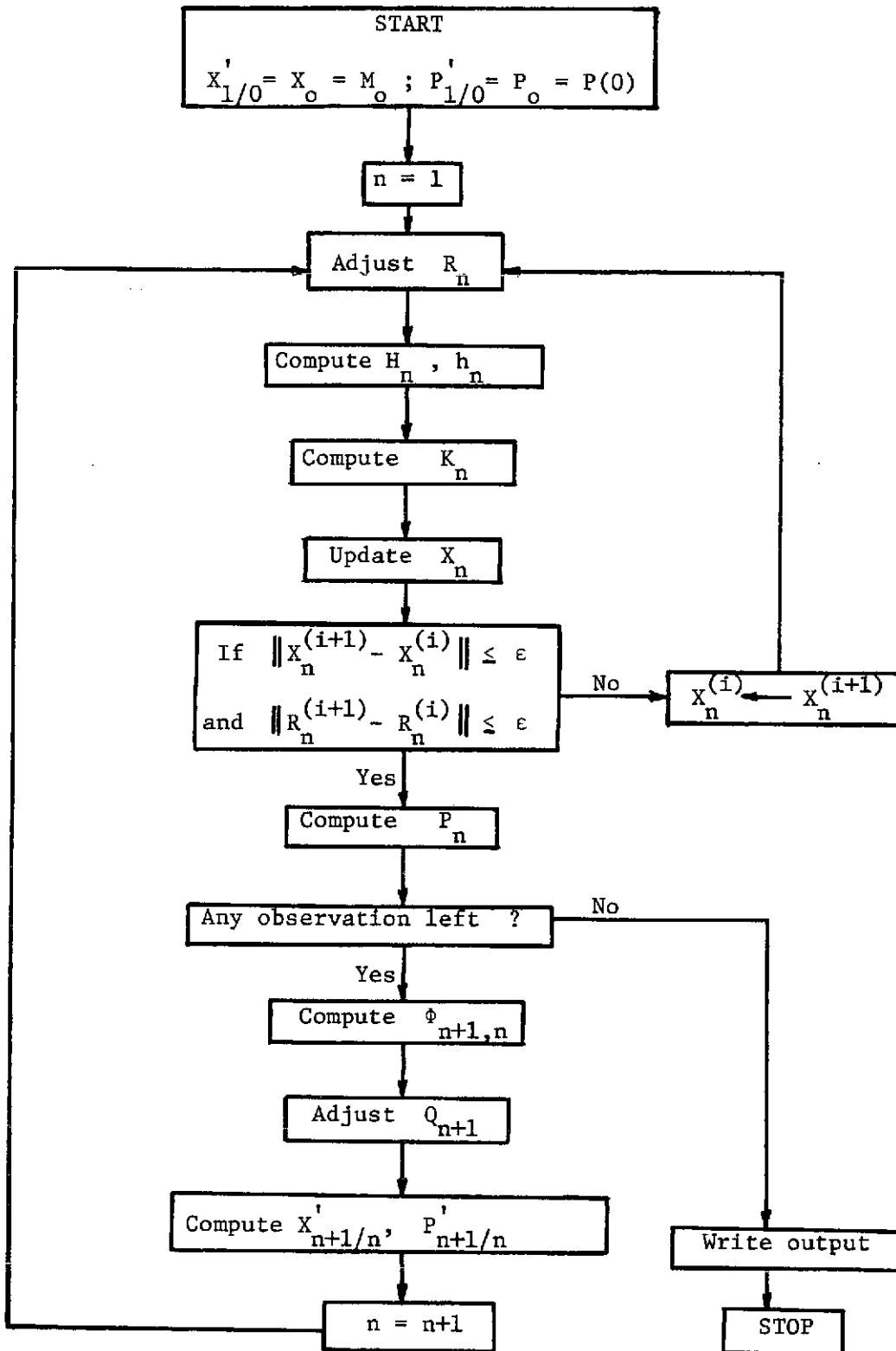
Figure 3-2: Block diagram of the Iterated Extended Kalman Filter.

Kalman filter, the amount of computer time required is also two or three times greater. Hence this technique is suitable for those cases when one has only a small set of observation data and wishes to come up with acceptable estimates for the state of the system. For large data sets, the other algorithms can, hopefully, give the desired values for the estimates after processing all the data with a reasonably small computer time.

Summary

In this Chapter, two approaches are presented for the estimation of the state and unknown parameters in a nonlinear lumped-parameter model of a hydrologic system. The optimum linear stochastic control approach is suitable for those cases when the error-covariance matrix of the input disturbances is unknown. The nonlinear estimation approach is simpler and more powerful, but requires knowledge about input noise characteristics. Subsequent chapters are devoted to the implementation of these techniques to the study of a particular model of the rainfall-runoff system.

CHAPTER 4

IMPLEMENTATION AND RESULTS

Nonlinear Lumped-Parameter Models for Rainfall-Runoff

The nonlinearity of the rainfall-runoff relationship has been
of concern only in the last decade; however, the concept of non-
linearity and its methods of analysis are still very limited (Chow,
1967). Following are two nonlinear lumped-parameter models that have
been used quite often by hydrologists in the past, namely the
Kulandaiswamy model and the Prasad model.

The Kulandaiswamy Model

Direct runoff may be considered as the result of the transform-
ation of rainfall excess by a basin system. The physical process of
this transformation is very complex, depending mainly upon the storage
effects in the basin. Kulandaiswamy (1964) derived the following
general expression for the storage

$$S = \sum_{n=0}^{N} a_n(Q,R) \frac{d^n Q}{dt^n} + \sum_{m=0}^{M} b_m(Q,R) \frac{d^m R}{dt^m} \qquad (4-1)$$

where S is the storage, t is the time, N and M are integers, and
$a_n(Q,R)$ and $b_m(Q,R)$ are parametric functions of the direct runoff Q
and the excess rainfall R. To apply Eqn. (4-1) to the study of the
rainfall-runoff process in a particular watershed, the values of N
and M must be determined. Both Q(t) and R(t) are available in the
form of curves and differentiation has to be done by numerical

approximation techniques. Taking into consideration the nature of

the curves representing Q(t) and R(t) and the magnitude of error

likely to be introduced by numerical differentiation, the values of

N = 1 and M = 0 have been adopted in Kulandaiswamy's study. Eqn.

(4-1) reduces to

$$S = a_o(Q,R) \ Q + a_1(Q,R) \ \frac{dQ}{dt} + b_o(Q,R) \ R \qquad . \qquad (4-2)$$

Plots of a_o, a_1 and b_o versus Q_p, the peak discharge, for Willscreek

Basin are illustrated in Figure 4-1. Kulandaiswamy found that a_1 and

b_o vary from storm to storm, but do not show any well defined trend

in the variations; hence, he took these two parameters as constants

(Kulandaiswamy and Subramanian, 1967). The storage equation can now

be written as

$$S = a_o(Q) \ Q + a_1 \ \frac{dQ}{dt} + b_o \ R \qquad . \qquad (4-3)$$

With the continuity equation

$$\frac{dS}{dt} = R(t) - Q(t) \qquad , \qquad (4-4)$$

the rainfall-runoff process can be represented by the following

differential equation

$$a_1 \ \frac{d^2Q}{dt^2} + A(Q) \ \frac{dQ}{dt} + Q = R - b_o \ \frac{dR}{dt} \qquad (4-5)$$

where
$$A(Q) = a_o + Q \ \frac{da_o}{dQ} \qquad .$$

A plot of Q versus A(Q) was made for various basins and two types of

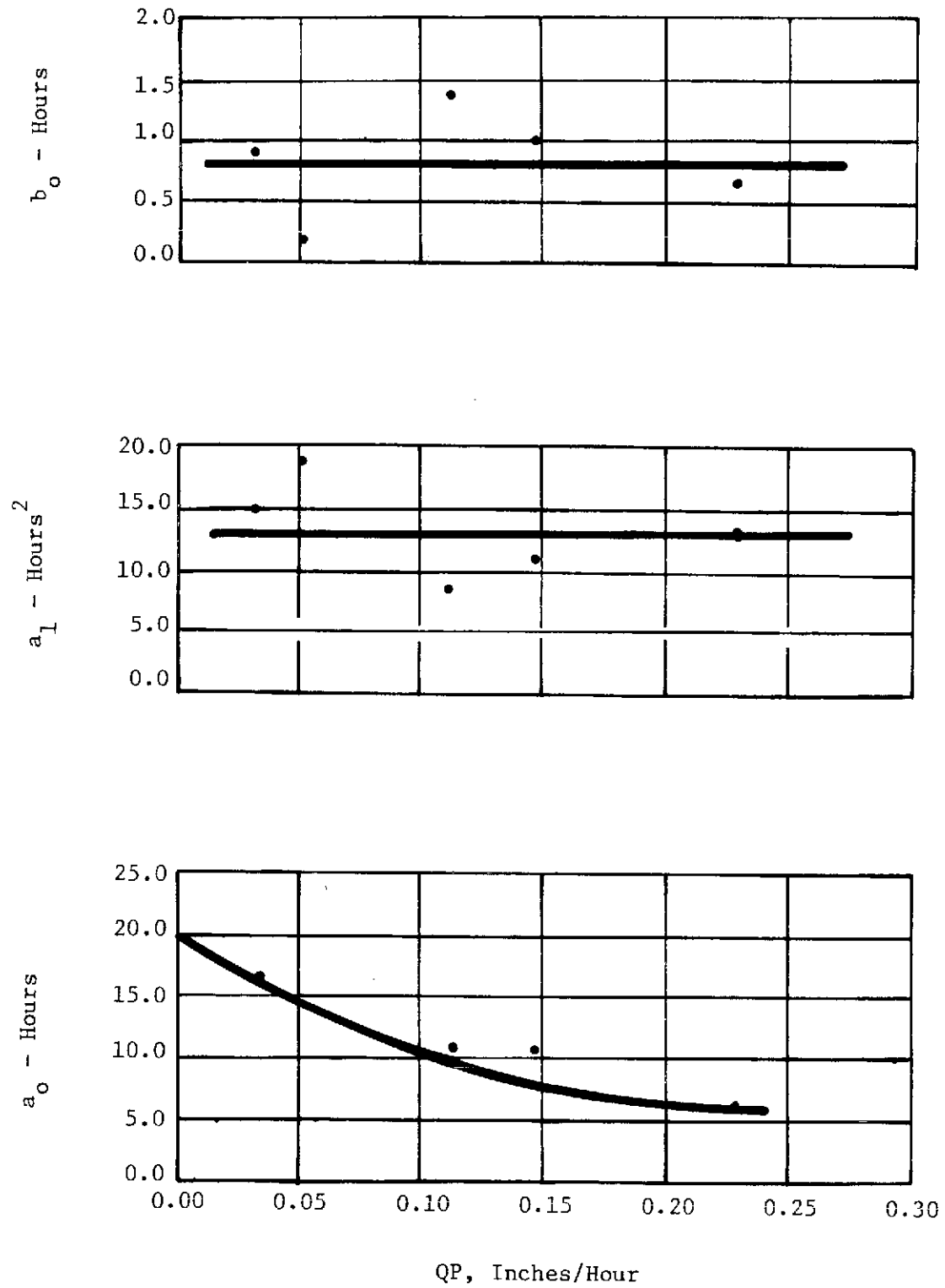regions could be differentiated. The system equations for these

Figure 4-1: Plots of a_o, a_1 and b_o vs. QP for Willscreek Basin.

regions are

(1) Non-linear region:

$$a_1 \frac{d^2 Q}{dt^2} + (c_1 + mQ) \frac{dQ}{dt} + Q = R - b_o \frac{dR}{dt} \quad ; \quad (4\text{-}6)$$

(2) Linear region:

$$a_1 \frac{d^2 Q}{dt^2} + c_2 \frac{dQ}{dt} + Q = R - b_o \frac{dR}{dt} \quad . \quad (4\text{-}7)$$

The general nonlinear storage equation (4-1) proposed by Kulandaiswamy has been accepted by many hydrologists in the simulation of the rainfall-runoff process by lumped-parameter response models, but the approach used in the determination of the model parameters has also been criticized (Eagleson, 1967). Kulandaiswamy used characteristics of the surface runoff hydrograph at peak discharge ($dQ/dt = 0$), on the falling limb ($R = 0$), and on the rising limb up to the end of rainfall excess to get various plots of a_o, a_1 and b_o vs. Q_p and Q vs. $A(Q)$; then from these plots the values of a_1, c_1, m, b_o and c_2 were determined. The evaluation of a_o from a single discharge (the peak discharge) and a_1, b_o from a portion of the surface runoff hydrograph should be replaced by some other means that can evaluate the model coefficients over the full range of observed discharges.

The Prasad Model

A simplification of the above model by retaining only two terms of the general nonlinear storage equation was proposed by Prasad (1967); in this case the storage equation is

$$S = K_1 Q^N + K_2 \frac{dQ}{dt} \qquad (4\text{-}8)$$

in which K_2 may be a complicated function of several variables affecting the wedge-storage as well as the storage-discharge relationship. In his study, Prasad assumed that K_1, K_2 and N are constant for a particular hydrograph. Using the continuity equation (4-4), one gets the following differential equation for the rainfall-runoff process

$$K_2 \frac{d^2Q}{dt^2} + K_1 N \, Q^{N-1} \frac{dQ}{dt} + Q = R \qquad . \qquad (4\text{-}9)$$

Comparing the Prasad model for nonlinear storage (Eqn. (4-8)) with the Kulandaiswamy model defined by Eqn. (4-3), one can recognize that $a_o(Q)$ and a_1 have been taken as $K_1 Q^{N-1}$ and K_2, respectively, and $b_o = 0$.

In the Prasad model, the time-invariant coefficients K_1, K_2 and N were evaluated by a trial-and-error method which is computationally inefficient and requires the knowledge of the initial conditions with sufficient accuracy. These coefficients were later computed by Labadie (1968) using quasi-linearization technique which has two main inherent weaknesses: (1) Initial approximations must be within, or at least close to, the convex region surrounding the optimal solution, or convergence is not attained. (2) If convergence does not result for a particular set of initial approximations, it is not possible to determine systematically a better set of initial approximations from these results.

All the above approaches for estimating the model coefficients are suitable only for deterministic models and not suitable for the

analysis of real input–output data where the values to be used in the model are imbedded in partially known or unknown noise. The two methods proposed in Chapter 3 are very useful in solving parameter identification problems in this case.

Since the Prasad model for rainfall-runoff is typically non-linear and the data set made available to the author was related to Prasad's work, only the Prasad model will be used in the investigation of the proposed identification schemes' performances.

Reformulation of the Prasad Model in State-Space

Eqn. (4-9) can be written as

$$\frac{d^2Q}{dt^2} = -(\frac{1}{K_2})K_1 N Q^{N-1}\frac{dQ}{dt} - (\frac{1}{K_2})Q + (\frac{1}{K_2})R \quad . \qquad (4-10)$$

Using the transformation

$$
\begin{aligned}
X_1 &= Q \\
X_2 &= \dot{Q} \\
X_3 &= K_1 \\
X_4 &= 1/K_2 \\
X_5 &= N
\end{aligned}
\qquad (4-11)
$$

and the assumption that the model coefficients are time-invariant, Eqn. (4-10) can be written in the following form

$$
\begin{bmatrix}
\dot{X}_1 \\
\dot{X}_2 \\
\dot{X}_3 \\
\dot{X}_4 \\
\dot{X}_5
\end{bmatrix}
=
\begin{bmatrix}
X_2 \\
-X_3 X_4 X_5 X_1^{X_5-1} X_2 + X_4(R - X_1) \\
0 \\
0 \\
0
\end{bmatrix}
, \quad (4-12)
$$

or, in abbreviated notation,

$$\dot{X}(t) = f\left[X(t),R(t)\right] \qquad . \qquad (4\text{-}13)$$

Eqn. (4-13) is the model equation in state-space. Let Y(t) denote the measured runoff which is embedded in noise, one then has

$$Y(t) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \end{bmatrix} + v(t) \qquad , \qquad (4\text{-}14)$$

or, in abbreviated notation,

$$Y(t) = h\left[X(t)\right] + v(t) \qquad (4\text{-}15)$$

where v(t) represents the noise term.

Eqns. (4-12), (4-13), (4-14) and (4-15) are the basis for further development.

Computation of the System State-Transition Matrix

The crucial problem in applying the proposed estimation schemes to continuous systems with discrete measurements is the evaluation of the state-transition matrix.

It is known that ϕ_{t,t_o} can be obtained from the coefficient matrix F(t) by the differential equation

$$\dot{\phi}_{t,t_o} = F(t)\, \phi_{t,t_o} \qquad (4\text{-}16)$$

with the initial condition

$$\Phi_{t_o,t_o} = I \qquad (4\text{-}17)$$

In the general case an explicit closed form solution of Eqn. (4-16) is not possible, but an infinite series can be derived which converges uniformly in t for every matrix $F(t)$. This solution is given as follows, by the use of the Peano-Baker method of integration (Pipes, 1963)

$$\Phi_{t,t_o} = I + \int_{t_o}^{t} F(\tau_1)d\tau_1 + \int_{t_o}^{t} F(\tau_1)\int_{t_o}^{\tau_1} F(\tau_2)d\tau_2 d\tau_1 +$$

$$\int_{t_o}^{t} F(\tau_1)\int_{t_o}^{\tau_1} F(\tau_2)\int_{t_o}^{\tau_2} F(\tau_3)d\tau_3 d\tau_2 d\tau_1 + \ldots \qquad (4\text{-}18)$$

If the matrix $F(t)$ satisfies the commutativity condition, i.e.,

$$F(t_1)\ F(t_2) = F(t_2)\ F(t_1) \quad \text{for all } t_1 \text{ and } t_2 \quad,$$

then the state-transition matrix is given by (DeRusso, Roy and Close, 1967)

$$\Phi_{t,t_o} = \exp\int_{t_o}^{t} F(\tau)d\tau$$

$$= \left\{ I + \int_{t_o}^{t} F(\tau)d\tau + \frac{1}{2!}[\int_{t_o}^{t} F(\tau)d\tau\ ]^2 \right.$$

$$\left. + \frac{1}{3!}[\int_{t_o}^{t} F(\tau)d\tau\ ]^3 + \ldots\ldots \right\} \qquad (4\text{-}19)$$

Expansion of each term and rearrangement yield:

$$\Phi_{t,t_o} = \left\{ I + \Delta t.F_o + \frac{\Delta t^2}{2!}(\dot{F}_o + F_o^2) + \frac{\Delta t^3}{3!}[\ddot{F}_o + \frac{3}{2}(F_o\dot{F}_o + \dot{F}_o F_o) + F_o^3] \right.$$

$$\left. + \ldots\ldots \right\} \qquad (4\text{-}20)$$

where $F_o = F(t_o)$ and $\Delta t = t - t_o$.

The computation of Eqn. (4-20) contains only a sufficient number of terms so that additional terms are negligible by comparison with the partial sum to that point. However, there are some difficulties associated with attempting to use a truncated form of this expansion. These stem from the impracticality of obtaining the first- and higher-order derivatives of F(t), if second- or higher-order accuracy is required. In these cases, it is better to subdivide the interval $(t - t_o)$ and to consider F as a constant matrix during these partial intervals. The partial transition matrix can now be computed by

$$\Phi_{t_i, t_{i-1}} = I + F \, \Delta t_N + F^2 \frac{\Delta t_N^2}{2!} + \dots = e^{F \Delta t_N} \quad (4\text{-}21)$$

where $\Delta t_N = (\Delta t)/N$ and N is the number of sub-intervals. The state-transition matrix Φ_{t, t_o} over the whole interval $(t - t_o)$ is simply the product of the partial matrices $\Phi_{t_i, t_{i-1}}$

$$\Phi_{t, t_o} = \Phi_{t=t_p, t_{p-1}} \cdot \Phi_{t_{p-1}, t_{p-2}} \dots \Phi_{t_2, t_1} \cdot \Phi_{t_1, t_o} \quad (4\text{-}22)$$

$$t = t_p > t_{p-1} > t_{p-2} > \dots > t_1 > t_o \quad .$$

For fast time-varying systems, it is better to use small sub-intervals and consider only the linear term in the expansion (4-21) than use bigger sub-intervals and include higher-order terms in the computation of Φ_{t, t_o} (Unger and Ott, 1970).

The formulation of the estimation schemes for estimating system parameters with noisy input-output data have been implemented for the study of the rainfall-runoff process. The data to be used were from the storm of April 10, 1953, on South Fork Vermilion River

Basin above Catlin, Illinois. The following results are obtained

from computer programs written in EXTENDED FORTRAN for use on the

CDC 6400 digital computer at Colorado State University.

Adaptive Control Approach

The matrices $F(t)$ and $G(t)$ in the linearized expression of

Eqn. (4-12) have the following forms

$$F(t) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ E_1 & -X_3X_4X_5X_1^{X_5-1} & -X_2X_4X_5X_1^{X_5-1} & E_2 & E_3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (4-23)$$

where

$$E_1 = -X_2X_3X_4X_5(X_5-1)X_1^{X_5-2} - X_4$$

$$E_2 = R-X_1-X_2X_3X_5X_1^{X_5-1}$$

$$E_3 = -X_2X_3X_4X_1^{X_5-1}-X_2X_3X_4X_5X_1^{X_5-1} \operatorname{Log}(X_1) \ ;$$

$$G(t) = \begin{bmatrix} 0 & X_4 & 0 & 0 & 0 \end{bmatrix}^T \qquad (4-24)$$

The error term $e(t)$, used to model the truncation error, is

chosen to be a zero-mean white noise process with covariance matrix

$$Q = \begin{bmatrix} 0 & & & & \\ & .01 & & & \\ & & 0 & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix}$$

The observation matrix derived from Eqn. (4-14) has the following

form

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (4\text{-}25)$$

the observation noise $v(t)$ is also assumed to be a zero-mean white noise process with variance

$$R = .0001 \qquad .$$

R is taken to be much smaller than Q because the observed outputs of the rainfall-runoff system are relatively noise-free compared to the inputs; also, there are errors in the model equations due to the incomplete knowledge of the nature of the system.

The following initial conditions are assumed for starting the adaptive control algorithm:

$$x_1^*(0) = x_2^*(0) = 0.$$

$$x_3^*(0) = 10. \quad , \quad x_4^*(0) = .1 \quad , \quad x_5^*(0) = 1.$$

$$\delta\hat{x}_i(0) = 0. \quad , \quad i = 1, 2, 3, 4, 5$$

$$P_o = \begin{bmatrix} 10 & & & & \\ & 10 & & & \\ & & 10 & & \\ & & & 10 & \\ & & & & 10 \end{bmatrix}$$

The initial estimation-error variances are chosen somewhat arbitrarily; the important thing is that they must be large enough so that the filter will forget the initial values as more data arrive.

For this linear stochastic control problem, the state-transition matrix must be computed carefully to avoid introducing further errors into the model equations; therefore, in the study, second-order terms are also taken into the computation of Φ. Since the state variables

of the rainfall-runoff system vary relatively slowly with time, a
sub-interval of the order of a few minutes is sufficient for the system
to be considered as time-invariant during this period of time. Thus,
the interval between two consecutive observations (1 hour) is subdivided
into only 10 sub-intervals to avoid excessive computational requirements,
and the value of Φ is computed from the approximation formulae (4-20)
and (4-21).

Based on nominal values of the state, the control gains are
computed first by the Subroutine ADAPT. A sample listing of these
gains is given in Table 4-1. Later, values of the control gains are
updated as soon as new values of the estimates are obtained. The
values of the time-invariant parameters in the model converge relatively
quickly to their optimal estimates after only 10 iterations, as shown
in Fig. 4-2. The optimal estimates are:

$$K_1 = 19.99 \qquad , \qquad \frac{1}{K_2} = 0.16 \qquad , \qquad N = 1.18 \quad .$$

Thus, using the same coefficients as those used by Labadie (1968), one
has
$$A_1 = \frac{K_1 N}{K_2} = 3.77 \qquad , \qquad A_2 = \frac{1}{K_2} = 0.16 \qquad , \qquad N = 1.18 \quad .$$

These values are not very much different from those obtained by Prasad,
using numerical integration of the nonlinear equation along with a
trial and error procedure; Prasad obtained the values 3.79, 0.076 and
1.27 for A_1, A_2 and N, respectively. Labadie, using the quasi-
linearization technique, obtained 4.473, 0.0943 and 1.27, respectively.
The differences are mainly due to noise terms introduced into the model
equations to make them more realistic and conform with the nature of the
rainfall-runoff data.

.539454116E-01	.668759662E-01	-.408799974E-05	.819002790E-04	-.252710147E-04
.534591990E-01	.668278519E-01	-.225537746E-05	.411390483E-04	.163041946E-03
.529202102E-01	.667745150E-01	.667714257E-07	-.577832919E-05	-.138525250E-04
.523227319E-01	.667153902E-01	.195283176E-05	-.424874495E-04	-.145489795E-03
.516604359E-01	.666498512E-01	.196881815E-05	-.390819943E-04	-.140989337E-03
.509263138E-01	.665772045E-01	.180376887E-05	-.358057368E-04	-.129816408E-03
.501126047E-01	.664966820E-01	.164561338E-05	-.326663210E-04	-.119019061E-03
.492107153E-01	.664074335E-01	.149424528E-05	-.296616348E-04	-.108598278E-03
.482111331E-01	.663085175E-01	.134958663E-05	-.267901305E-04	-.985568599E-04
.471033289E-01	.661988922E-01	.121158868E-05	-.240508404E-04	-.888995269E-04
.458756508E-01	.660774046E-01	.108023272E-05	-.214433929E-04	-.796330341E-04
.445152074E-01	.659427788E-01	.955531026E-06	-.189680315E-04	-.707662904E-04
.430077393E-01	.657936038E-01	.837527969E-06	-.166256367E-04	-.623104887E-04
.413374776E-01	.656283193E-01	.726301129E-06	-.144177481E-04	-.542792452E-04
.394869901E-01	.654452000E-01	.621962609E-06	-.123465907E-04	-.466887482E-04
.374370114E-01	.652423397E-01	.524660405E-06	-.104151020E-04	-.395579165E-04
.351662582E-01	.650176321E-01	.434579901E-06	-.862696154E-05	-.329085672E-04
.326512274E-01	.647687515E-01	.351945452E-06	-.698662243E-05	-.267655931E-04
.298659757E-01	.644931306E-01	.277022066E-06	-.549934470E-05	-.211571493E-04
.267818812E-01	.641879369E-01	.210117154E-06	-.417123013E-05	-.161148474E-04
.233673834E-01	.638500475E-01	.151582362E-06	-.300925850E-05	-.116739587E-04
.195877036E-01	.634760206E-01	.101815444E-06	-.202132480E-05	-.787362217E-05
.154045420E-01	.630620663E-01	.612621616E-07	-.121627698E-05	-.475705816E-05
.107757534E-01	.626040140E-01	.304181887E-07	-.603953703E-06	-.237178397E-05
.565500788E-02	.620972797E-01	.983088953E-08	-.195219667E-06	-.769822890E-06
0.	.615384615E-01	0.	0.	0.

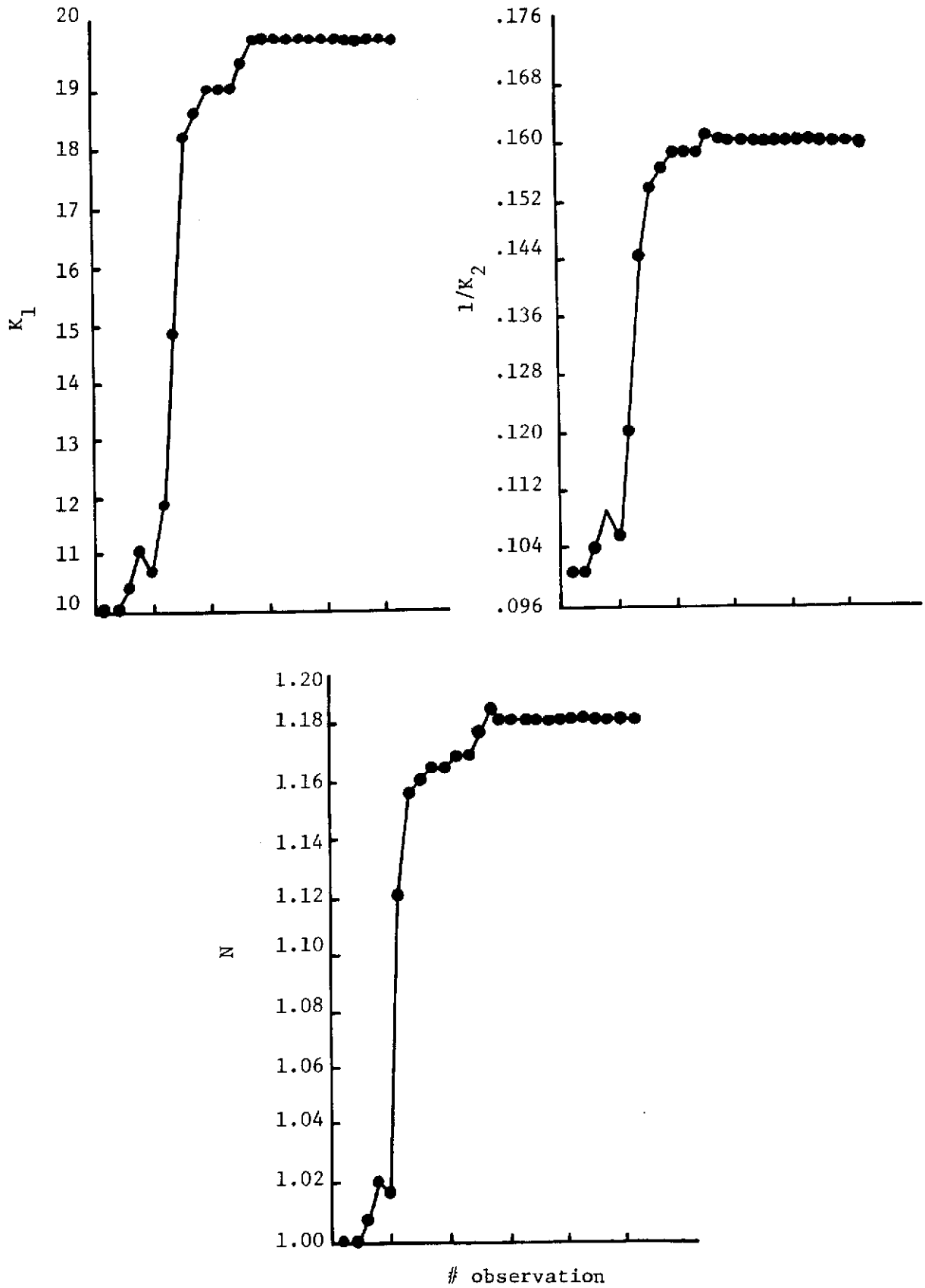Table 4-1: Sample values of control gains.

Figure 4-2: Estimation of time-invariant parameters by
linear adaptive control approach.

Values of the estimated surface runoff compared with observed
runoff are given in Figure 4-3 for control and without control of
the input disturbance. A sample listing of the results for these
two cases is given in Table 4-2, in which the peak values of the
estimated and observed hydrographs are indicated by small arrows.
From these results, the following remarks can be made:

(1) Adaptive control of the inputs is important for the analysis
of the rainfall-runoff process. Without controlling the rainfall data,
the system identification results would be bad, and, therefore, the
estimation of the surface runoff from these noisy inputs would be
unacceptable.

(2) Controlling the inputs requires additional computer time.
For the particular data set under study, with the same initial conditions
mentioned previously, only 6 secs. are needed to add to 12 secs. which
is the computer time used by the Kalman filtering scheme, including the
adaptive estimation of the error-covariance matrices R and Q.

(3) Even with input-control procedure, the approximation of a
nonlinear system by linearized equations cannot offer good results,
unless the system under study is not highly nonlinear.

The effect of adaptive estimation of the model-error covariance
matrix can be seen from Figure 4-4. In this test case, the Q matrix
can be seen as

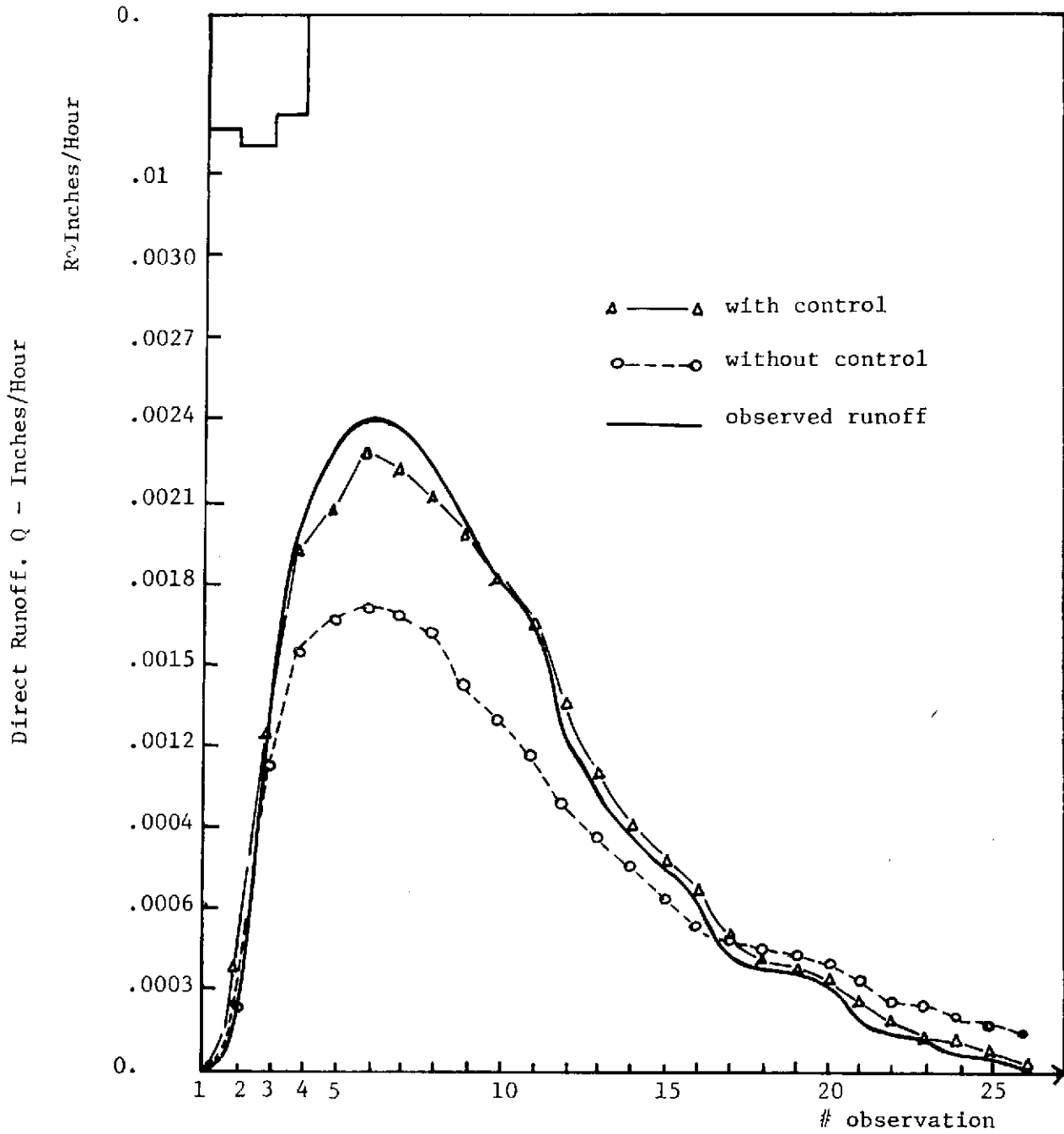$$Q = \begin{bmatrix} 0 & & & & \\ & 100 & & & \\ & & 0 & & \\ & & & 0 & \\ & & & & 0 \end{bmatrix} .$$

Figure 4-3: Estimate of direct runoff with and without control.

ESTIMATED	ESTIMATED	MEASURED
.991807573E-07	.991807573E-07	-0.
.421776526E-03	.252270195E-03	.150000000E-03
.134001409E-02	.112551851E-02	.147000000E-02
.197257266E-02	.158049358E-02	.205000000E-02
.206403744E-02	.169617079E-02	.230000000E-02
→ .218885424E-02	→ .171930534E-02	→ .240000000E-02
.216242101E-02	.168333312E-02	.238000000E-02
.210813662E-02	.159430782E-02	.225000000E-02
.195767699E-02	.144564441E-02	.200000000E-02
.177955574E-02	.131368394E-02	.178000000E-02
.160873416E-02	.120310747E-02	.160000000E-02
.137013063E-02	.102278394E-02	.128000000E-02
.113689880E-02	.879052731E-03	.103000000E-02
.941007460E-03	.771580819E-03	.850000000E-03
.777018873E-03	.679555547E-03	.700000000E-03
.652996927E-03	.613792339E-03	.600000000E-03
.521249447E-03	.522779863E-03	.450000000E-03
.426869681E-03	.473724189E-03	.380000000E-03
.374135431E-03	.449590865E-03	.360000000E-03
.310153269E-03	.396322809E-03	.280000000E-03
.240757733E-03	.343625931E-03	.200000000E-03
.184397605E-03	.306895386E-03	.150000000E-03
.133759896E-03	.270461470E-03	.100000000E-03
.100924958E-03	.249638757E-03	.800000000E-04
.659088663E-04	.219309945E-03	.400000000E-04
.298520185E-04	.189475272E-03	-0.
(I)	(II)	

Table 4-2: Sample results of estimated surface runoff by linear adaptive control approach:
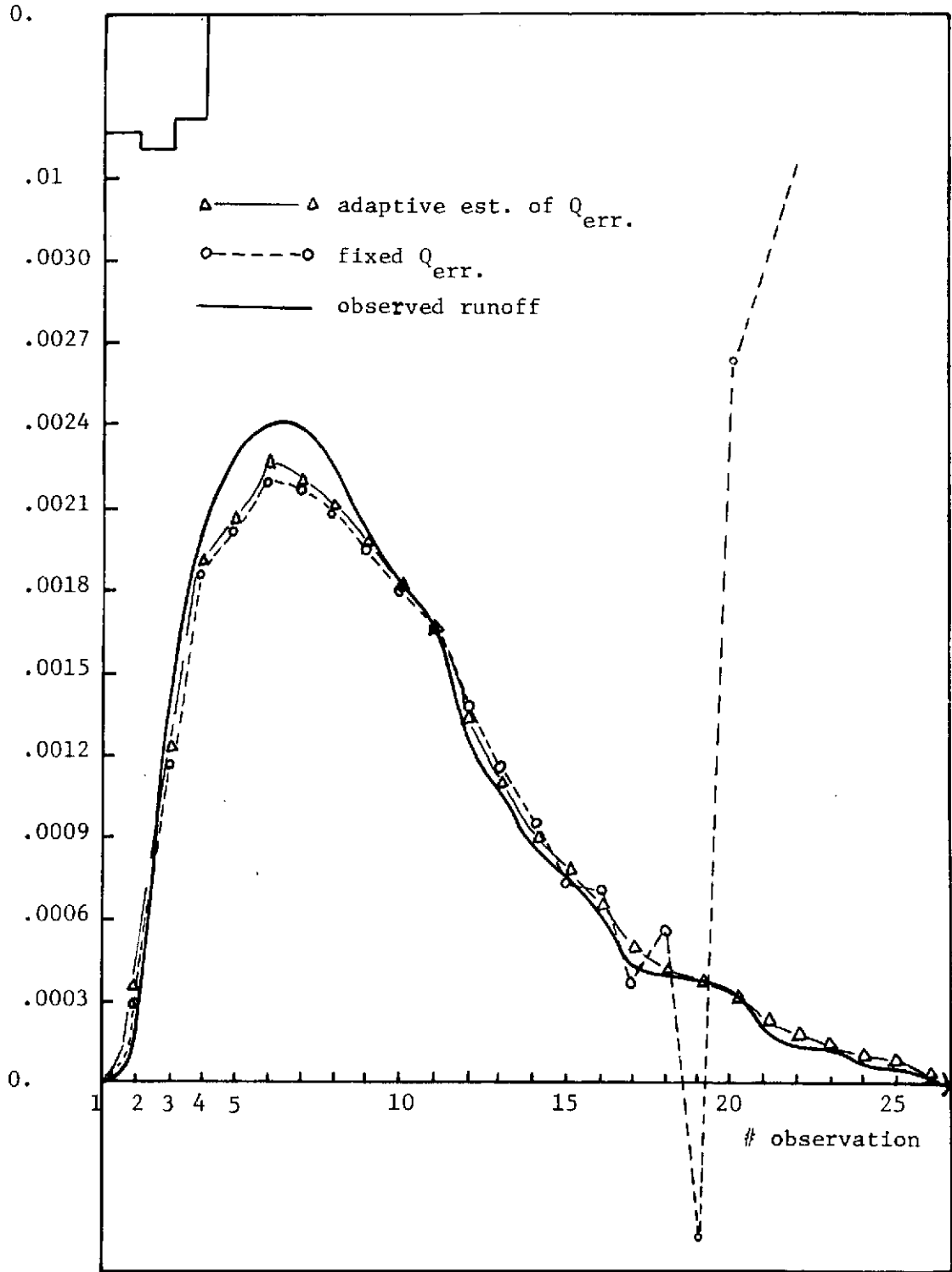
(I) with control
(II) without control

Figure 4-4: Effect of adaptive estimation of model-error covariance matrix.

Without adaptive estimation of the Q matrix at each stage, the filter started to diverge at the 20th observation.

Finally, performances of the adaptive control algorithm with and without rectification of the nominal state at each stage are given in Figure 4-5 for comparison. A description of subroutine ADAPT and the related working subroutines are given in Appendix B.

Nonlinear Estimation Approach

Since rainfall data are noisy, the continuity equation is written as

$$\frac{dS}{dt} = R^* - Q + w(t) \tag{4-26}$$

where R^* denotes the actual noisy input data and $w(t)$ represents the input noise which is assumed, as usual, to be a zero-mean white noise process with covariance matrix

$$E\left(w(t)w(\tau)^T \right) = W \delta(t - \tau) \quad . \tag{4-27}$$

Combining Eqn. (4-26) with the nonlinear storage equation (4-8), one obtains

$$\frac{d^2Q}{dt^2} = -(\frac{1}{K_2})K_1 N Q^{N-1}\frac{dQ}{dt} + (\frac{1}{K_2})(R^* - Q) + (\frac{1}{K_2})w \quad . \tag{4-28}$$

Let $X_1 = Q$, $X_2 = \dot{Q}$, $X_3 = K_1$, $X_4 = \frac{1}{K_2}$, $X_5 = N$,

one then gets the following state equation

$$
\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \\ \dot{X}_4 \\ \dot{X}_5 \end{bmatrix} = \begin{bmatrix} X_2 \\ -X_4 X_3 X_5 X_1^{X_5-1} X_2 + X_4(R^* - X_1) \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ X_4 \\ 0 \\ 0 \\ 0 \end{bmatrix} w \quad ,
$$

$$(4-29)$$

or, in abbreviated notations,

$$
\dot{X} = f\left(X(t), R^*\right) + g\left(X(t), w(t)\right) . \tag{4-30}
$$

The observation equation is the same as in the previous section.

The matrix of partial derivatives is

$$
F(t) = \begin{bmatrix} 0 & X_2 X_5-1 & 0 & 0 & 0 \\ E_1 & -X_3 X_4 X_5 X_1^{X_5-1} & -X_2 X_4 X_5 X_1^{X_5-1} & E_2 & E_3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4-31}
$$

where

$$
E_1 = - X_2 X_3 X_4 X_1 (X_5 - 1) X_1^{X_5-2} - X_4
$$

$$
E_2 = R^* - X_1 - X_2 X_3 X_5 X_1^{X_5-1}
$$

$$
E_3 = - X_2 X_3 X_4 X_1^{X_5-1} (1 + X_5 \text{Log}(X_1))
$$

Subroutine ITERA was developed using the iterated extended Kalman filter to estimate the state and unknown parameters in the system.
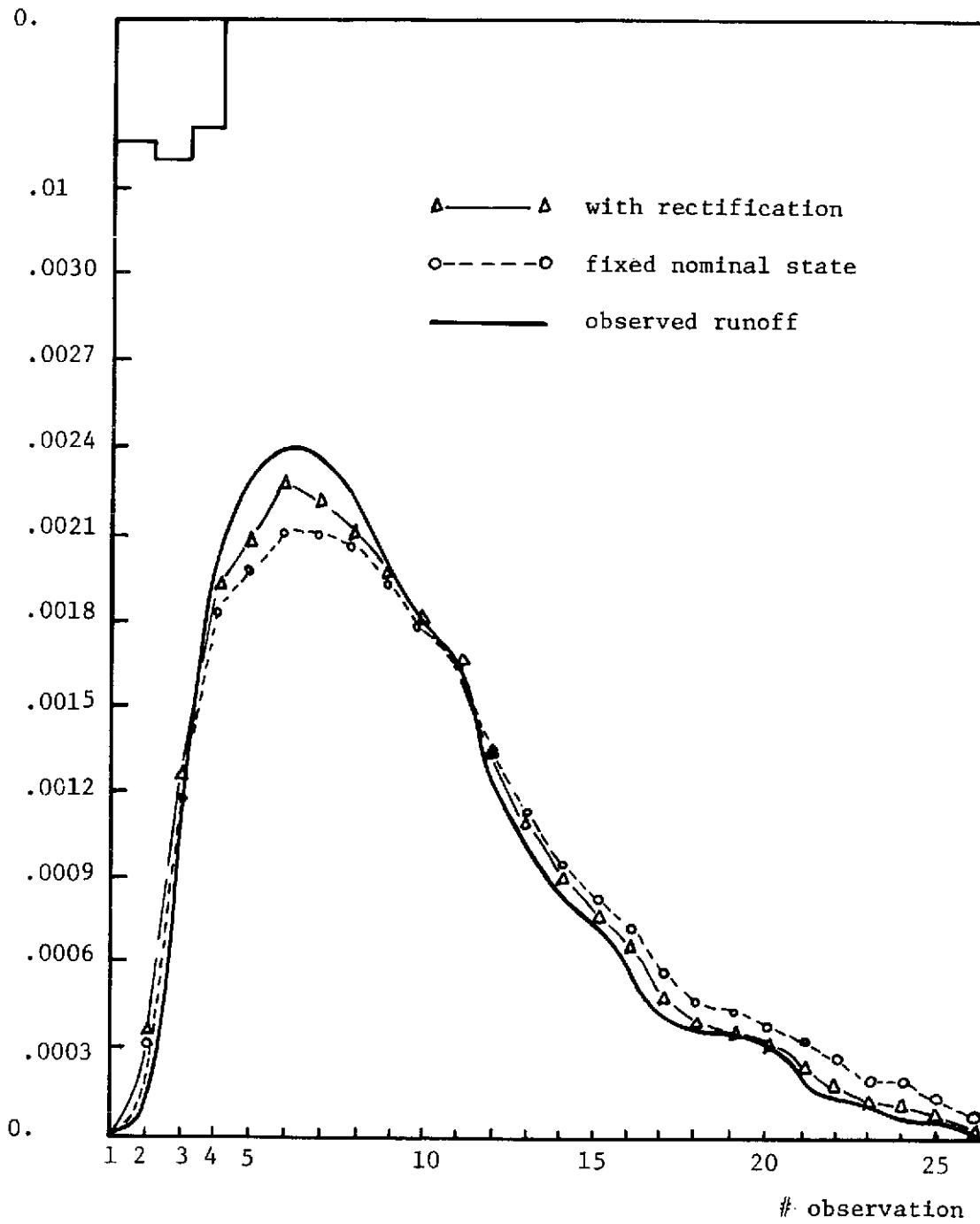
Figure 4-5:   Effect of rectification of nominal state.

Assume that the error-covariance matrices have the following values

$$R = .001 \quad , \quad W = .01$$

and the given initial conditions for $X_i(0)$ and $P(0)$ are the same as in the previous example. The optimum values of the model parameters converge a little faster than in the case of using the adaptive control algorithm; this is shown in Figure 4-6. These optimum estimates are

$$K_1 = 19.998 \quad , \quad \frac{1}{K_2} = 0.162 \quad , \quad N = 1.182.$$

One could expect that these results are better than those obtained previously, since in this case a nonlinear filter has been used and therefore model-error has been reduced. This fact is verified by a better fit of the estimated outflow with the measured outflow for this case as shown on Figure 4-7. A sample listing of the values of these estimates is given in Table 4-3. Since the estimates of model coefficients converge to stable values, one may conclude that these coefficients are constant or can be approximated by time-invariant parameters for a particular hydrograph.

The variation of the estimation-error variances of the state with time is shown in Figure 4-8. The effect of adaptive estimation of the model-error covariance matrix is also tested in this case. Using the same set of initial conditions as used previously, the divergence of the filter was less rapid than in the adaptive control approach. The result is shown in Figure 4-7.

A description of subroutine ITERA and related working subroutines are also presented in Appendix B.
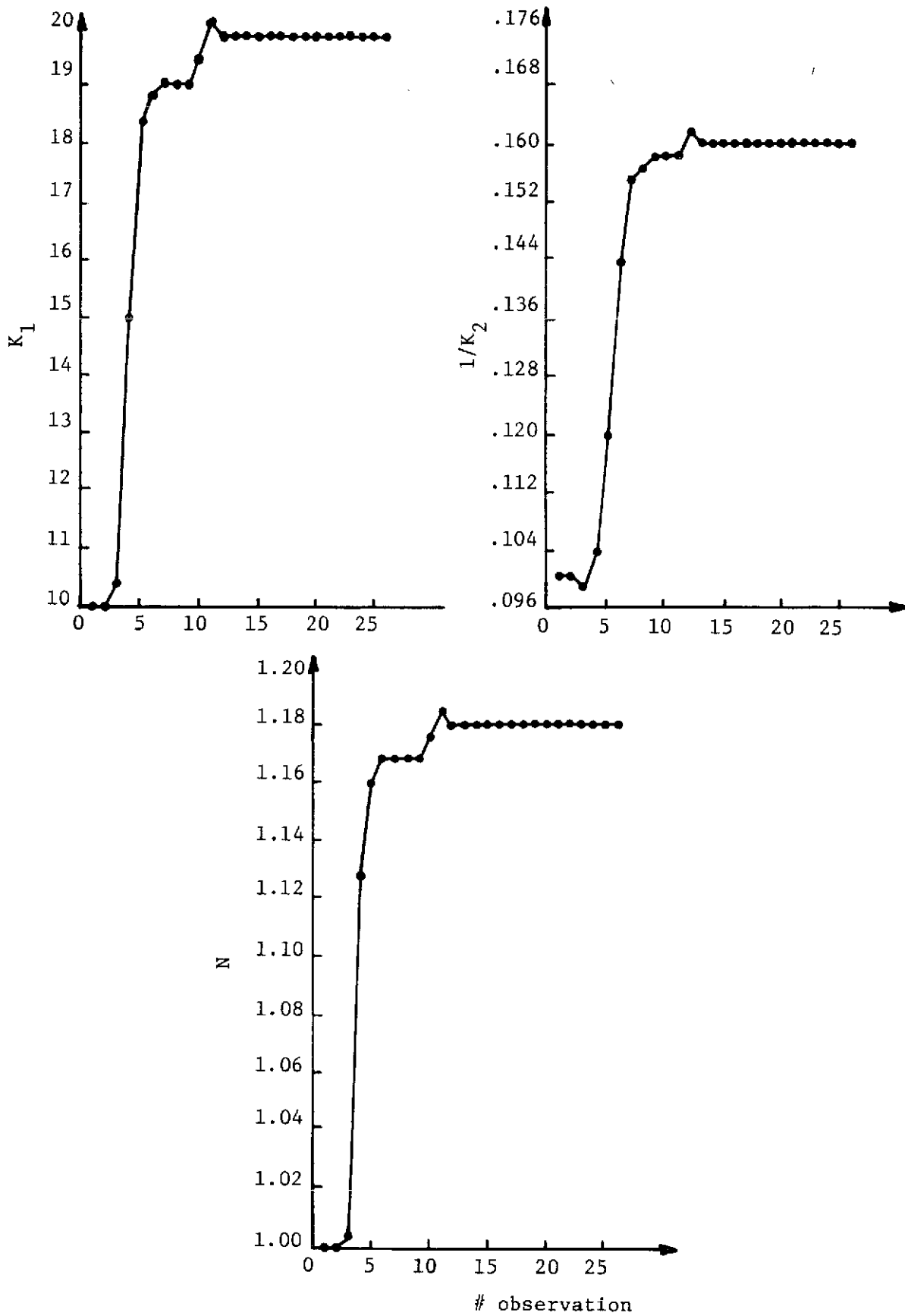
Figure 4-6:   Estimation of time-invariant parameters by
iterated extended Kalman filter.

ESTIMATED	MEASURED
.119619071E-15	-0.
.150062718E-03	.150000000E-03
.146647386E-02	.147000000E-02
.205084022E-02	.205000000E-02
.221341347E-02	.230000000E-02
→ .240452681E-02	→ .240000000E-02
.236577599E-02	.238000000E-02
.224040644E-02	.225000000E-02
.200300775E-02	.200000000E-02
.180874269E-02	.178000000E-02
.161181444E-02	.160000000E-02
.127554449E-02	.128000000E-02
.103091575E-02	.103000000E-02
.849778807E-03	.850000000E-03
.700038029E-03	.700000000E-03
.599990928E-03	.600000000E-03
.450001643E-03	.450000000E-03
.379999577E-03	.380000000E-03
.359997687E-03	.360000000E-03
.280013859E-03	.280000000E-03
.199952195E-03	.200000000E-03
.150323660E-03	.150000000E-03
.994354885E-04	.100000000E-03
.818362730E-04	.800000000E-04
.364464120E-04	.400000000E-04
.431316959E-04	-0.

Table 4-3: Sample results of estimated surface runoff by nonlinear filter.
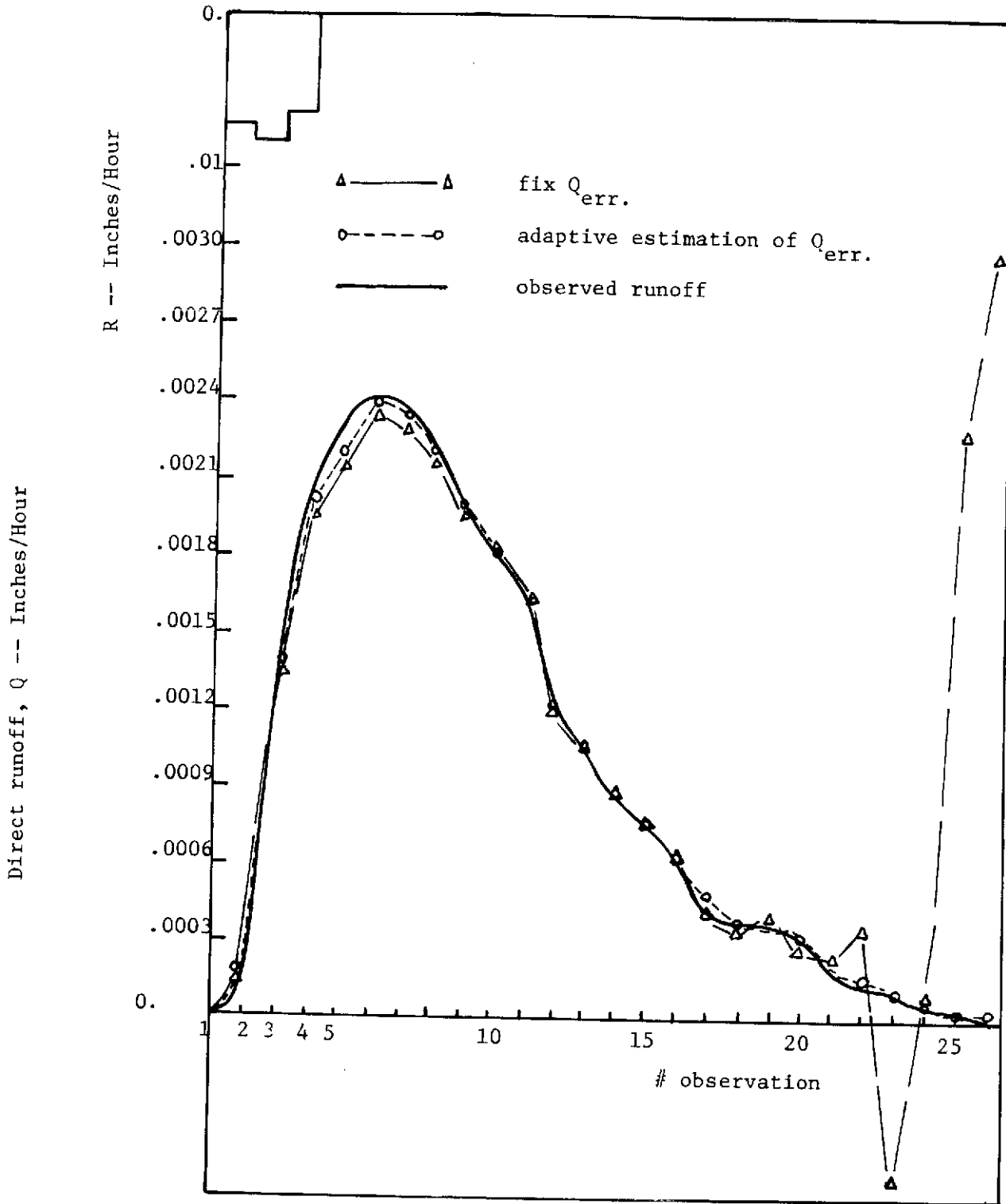
Figure 4-7: Estimate of direct runoff with and without adaptive estimation of model-error covariance matrix.
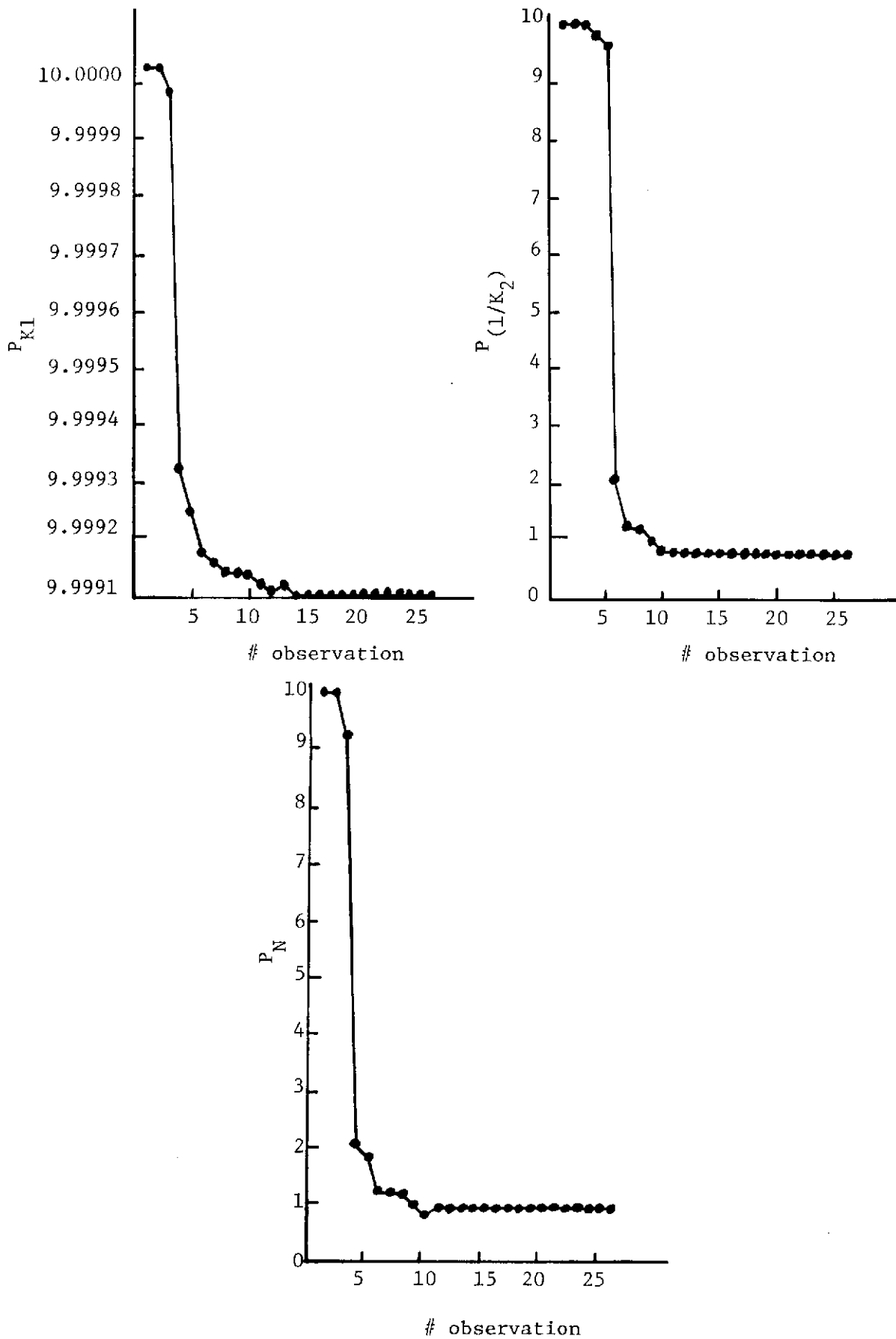
Figure 4-8: Variations of the estimation-error variances of the estimates.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

Summary of the work described in previous chapters

After reviewing various parameter identification techniques
used in hydrology in the past, two new approaches for the analysis
of the rainfall-runoff process are proposed, namely the adaptive
control of input disturbances and the nonlinear estimation of
unknown parameters in a noisy environment. Both techniques are
then applied to the identification of time-invariant parameters
in the Prasad model of rainfall-runoff. The results obtained
are encouraging and conform with previous studies made by Prasad
(1967) and Labadie (1968).

Advantages of the proposed approaches

The identification techniques presented in this report
offer the following advantages:

1) The formulation of the parameter identification problem
is natural and gives more insight to researchers about the
process under study.

2) Both techniques offer a better and systematic way to
analyze the rainfall-runoff process.

3) The identification schemes are sequential and adaptive
and therefore the proposed approaches can handle large data
set which may be imbedded in noise with unknown characteristics.

4)   Both techniques can handle any type of parameters, either time-invariant or time dependent.

5)   They can handle multiple-input - multiple-output systems.

6)   The computational requirements are relatively small, since the algorithms are simple and the computer times needed to process a set of 26 observations are just 14 sec. and 18 sec. for the nonlinear estimation approach and the adaptive control approach, respectively.

LIMITATIONS:

Any approach must have certain operational limitations which stem from the formulation of the problem and assumptions used to derive necessary equations for further considerations. Therefore, the following limitations on the use of the proposed techniques can be listed:

1)   They are restricted to jointly Gaussian variables with white noise processes or with a certain class of colored noise, i.e., noise $v(t)$ which is not white, but can be expressed as

$$v(t) \;=\; A(t) \; v(t) \;+\; \xi(t)$$

where $A(t)$ is known and $\xi(t)$ is white.

2)   The system under study must be controllable and observable. This problem of stochastic controllability and observability has been discussed in detail by Aoki (1967) and Sorenson (1968).

3)   When using linearized equations for the adaptive control
scheme, the initial guess for the value of the state cannot be far
away from the true value, if filter divergence is to be avoided.

Suggestions for Future Research

Possible extensions of the study reported in this report are
suggested as follows:

1)   Extend this work to the study of distributed-parameter response
models of hydrologic systems. In this case, the systems will be
represented by partial differential equations and the estimation
algorithms developed in Chapter 3 will need only slight modifications
to include the spatial variations of the states.

2)   Investigate the use of the proposed approaches for short-term
streamflow prediction in small watersheds, using only measured runoff
at the mouth of the basins. In this case, the adaptive control scheme
will be used to estimate the unknown input from measured runoff and
good short-term streamflow prediction can be obtained by propagating
the state and the estimation-error covariance matrix forward in time.

3)   Investigate the possibility of combining the identification
of lumped-parameter models of various environmental zones of a large
mountaineous watershed in one task, using the proposed identification
techniques and total measured runoff. In this case, routing models
for two adjacent zones must be incorporated into the system equations
and the use of the state-variable approach can give a simple matrix
differential equation representing the response model of the whole
watershed under study. Thus, using total measured runoff, one can

estimate all unknown parameters in various environmental zones and in the routing models at the same time.

4) Apply these techniques to the analysis of other hydrologic processes which can be represented by lumped-parameter response models. For this task, no further modifications of the proposed approaches are required because they are derived for use in the general case of time-invariant or time-dependent, linear or nonlinear lumped-parameter hydrologic response models.

REFERENCES

1. Amorocho, J. and G. T. Orlob. 1961. Nonlinear Analysis of Hydrologic
        Systems. Water Resources Center, Univ. of California,
        Contribution 40, Berkeley, California.

2. Aoki, Masanao. 1967. Optimization of Stochastic Systems - Topics
        in Discrete-time Systems. Academic Press, New York.

3. Aseltine, J. A., Mancini, A. R. and C. W. Sarture. 1958. A
        Survey of Adaptive Control Systems. I. R. E. Trans. on
        Automatic Control, PGAC-6:102-108.

4. Athans, M. 1967. The Relationship of Alternate State-Space Represen-
        tations in Linear Filtering Problems. I. E. E. Trans. on
        Automatic Control, Correspondence: 775-776.

5. Athans, M., Wishner, R. P. and A. Bertolini. 1968. Suboptimal
        State Estimation for Continuous time Nonlinear System from
        Discrete Noisy Measurements. I. E. E. E. Trans. on Automatic
        Control, AC-13: 504-514.

6. Barrera, A. B. and F. E. Perkins. An extension of the role of
        linear systems analysis in Hydrograph Theory. Report No.106,
        M. I. T. Dept. of Civil Engineering, Hydrodynamics Laboratory.
        1967.

7. Bellman, R. 1963. Mathematical Optimization Techniques. Chapter
        17. Univ. of California Press.

8. Bellman, R. and R. E. Kalaba. 1965. Quasi-linearization and Non-
        linear Boundary Value Problems. Elsevier, Amsterdam.

9. Beltrami, E. J. 1970. An Algorithmic Approach to Nonlinear Analysis
        and Optimization. Academic Press. New York and London.

10. Blank, D. and J. W. Delleur. 1968. A Program for Estimating
        Runoff from Indiana Watersheds - Part I: Linear System
        Analysis in Surface Hydrology and its Application to Indiana
        Watersheds. Tech. Report No. 4. Purdue University, Water
        Resources Research Center.

11. Bode, H. W. 1945. Network Analysis and Feedback Amplifier Design.
        Van Nostrand, Princeton, N. J.

12. Braun, L., Jr. 1959. On Adaptive Control Systems. Doctoral
        Dissertation, Polytechnic Institute of Brooklyn.

13. Brooks, S. H. 1958. A Discussion of Random Methods for Seeking
        Maxima. Operations Research, 6, 244-251.

14. Brooks, S. H. 1959. A comparison of Maximum-Seeking Methods. *Operations Research*, 7, 430-457.

15. Bryson, A. E., Jr. and Yu-Chi Ho. 1969. Applied Optimal Control. Blaisdell Publishing Co., Waltham, Massachusetts.

16. Chow, V. T. 1967. General Report, Technical Session No. 1. *Inter. Hydrology Symposium*. Fort Collins, Colo. Sept. 6-8.

17. Cooper, G. R. and J. E. Gibson. 1960. A Survey of Philosophy and State-of-the-Art of Adaptive Systems. *Tech. Report. No. 1*, School of Electrical Engineering, Purdue University.

18. Courant, R. 1943. Variational methods for the solution of problems of equilibrium and vibrations. *Bull. Amer. Math. Soc.* 49: 1-23.

19. Davidon, W. C. 1959. Variable Metric Method for Minimization. A. E. C. *Research and Development Report*, ANL-5990, 21 pp.

20. Davies, W. D. T. 1970. *System Identification for Self-adaptive Control*. Wiley-Interscience.

21. Dawdy, D. R. 1969. Mathematic Modeling in Hydrology. In: *The Progress of Hydrology*. Proc. of the First Inter. Seminar for Hydrology Professors. Univ. of Illinois, Urbana.Illinois.

22. Decoursey, D. G. and W. M. Snyder. 1969. Computer-oriented method of optimizing hydrologic model parameters. *J. of Hydrology*, 9, 34-56.

23. Delleur, J. W. 1971. Developpements Recents en Matiere d'Analyse des Chroniques d'Evenements Hydrologiques: Liaison en Chaine, Analyse Spectrale. (Recent Developments in the Analysis of Hydrologic Time Series - Serial Correlations and Spectral Analysis.) *La Houille Blanche*, No. 2:147-158 (in French with English summary).

24. Denham, W. F. and S. Pines. 1966. Sequential Estimation when Measurement Function Nonlinearity is comparable to error. *A. I. A. A. Journal*, 4, 1071-1076.

25. Diskin, M. H. 1964. A Basic Study of the Linearity of the Rainfall-Runoff Process in Watersheds. *Ph.D. Thesis*. University of Illinois, Civil Engineering.

26. Dooge, J. C. I. 1959. A General Theory of the Unit Hydrograph. *J. of Geophysical Research*, Vol. 64.

27. Dooge, J. C. I. Analysis of Linear systems by means of Laguerre Functions. *J. SIAM (Control)* Ser. A, Vol. 12(3). 1965.

28. Duong, N. 1970. Space-based orbit determination. M. S. thesis.
        Colorado State University, Fort Collins, Colorado.

29. Eagleson, P. S., Mejia, R. R. and F. March. 1966. Computation
        of Optimum Realizable Unit Hydrographs. Water Resources
        Research. Vol. 2 (4): 755-764.

30. Eagleson, P. S. 1967. General Report. Technical Session No. 2.
        Inter. Hydrology Symposium. Fort Collins, Colo. Sept. 6-8.

31. Fitzgerald, R. J. 1971. Divergence of the Kalman Filter. I.E.E.E.
        Trans. on Automatic Control, Vol AC-16(6): 736-747.

32. Fletcher, R. 1965. Function Minimization without evaluating
        derivatives - A Review. The Computer Journal, 8, 33-41.

33. Friedland, B. and I. Bernstein. 1966. Estimation of the State of
        a Nonlinear process in the presence of non-gaussian noise
        and disturbances. J. Franklin Inst. 281: 455-480.

34. Fu, K. S., Gibson, J. E., Hill, J. D., Luisi, J. A., Raible, R. H.
        and M. D. Waltz. 1963. Philosophy and State-of-the-art
        of Learning Control Systems. Report No. TR-EE 63-7.
        Purdue University.

35. Gibson, J. E. 1963. Adaptive Learning Systems. Available from
        ASTIA, Alexandria, Virginia, as AD 292796.

36. Graupe, D. 1972. Identification of Systems. Van Nostrand
        Reinhold Co. New York.

37. Gurin, L. S. and L. A. Rastrigin. 1965. Convergence of the
        Random Search method in the presence of noise. Automation.
        and Remote Control, 26, 1505-1511.

38. Gurin, L. S. Random search in the presence of noise. Engineering
        Cybernetics. 3. 1966.

39. Heffes, H. 1966. The effect of erroneous models of the Kalman
        filter response. I. E. E. E. Trans. on Automatic Control
        (short paper) Vol AC-11, 541-543.

40. Hill, J. D. 1969. A Search Technique for Multi-modal Surfaces.
        I. E. E. E. Trans. on Systems Science and Cybernetics,
        Vol. SSC-5 (1): 2-8.

41. Hooke, R. and T. A. Jeeves. 1958. Comment on Brook's discussion
        of random methods. Operations Research. Vol. 6(6).

42. Horlick, A. and D. V. Sward. 1965. Computer Precision Study II.
        I. B. M. Report 65-504-30. I. B. M. Federal Systems Div.

43. Ibbitt, R. P. 1970. Systematic parameter fitting for conceptual models of catchment hydrology. Ph.D. Thesis, Imperian College, University of London.

44. Jazwinski, A. 1971. Stochastic Processes and Filtering Theory. New York: Academic Press, 305-318.

45. Kalman, R. E. 1960. A New Approach to Linear Filtering and Prediction Problems. Trans. ASME, Vol. 82D, p. 35.

46. Karnopp, D. C. 1963. Random Search Techniques for Optimization Problems. Automatica, Vol. 1: 11-121.

47. Kavvas, Mustafa L. Derivation of Stable Non-oscillating Unit Hydrographs. M. S. Thesis, Dept. of Civil Engineering, Colorado State University, Fort Collins, Colo. 1972.

48. Kisiel, C. C. 1969. Time Series Analysis of Hydrologic Data. Advances in Hydroscience, V. T. Chow, editor, Academic Press.

49. Kulandaiswamy, V. C. 1964. A basic study of the rainfall excess Surface Runoff Relationship in a basin system. Ph.D. Thesis, University of Illinois.

50. Kulandaiswamy, V. C. and C. V. Subramanian. 1967. A nonlinear approach to runoff studies. Inter. Hydrology Symposium. Fort Collins, Colo. Sept. 6-8.

51. Kushner, H. J. 1967. Dynamical equations for optimum nonlinear filtering. J. Differential Eqns. 3: 179-190.

52. Labadie, J. 1968. Optimal identification of nonlinear hydrologic system response models by quasi-linearization. M. S. Thesis, Univ. of California at Los Angeles.

53. Leon, A. 1966. A comparison among eight known optimizing procedures. Pages 23-46 of Recent Advances in Optimization Techniques. T. P. Vogl and A. Lavi (eds) Wiley, N.Y.

54. Linsley, R. K. and Crawford, N. H. 1966. Digital Simulation in Hydrology: Stanford Watershed Model IV. Tech. Report 39. Stanford University, Department of Civil Engineering.

55. Mehra, R. K. 1972. Approaches to Adaptive Filtering. I. E. E. E. Trans. on Automatic Control (short papers): 693-698.

56. Mendel, J. M. and Zapalac, J. J. 1968. The application of techniques of artificial intelligence to control system design. In: Advances in Control Systems: Theory and Applications. C. T. Leondes, ed. Vol. 6, pp. 1-94. Academic Press, New York.

57. Mosteller, H. W. 1963. Learning Control Systems. Electronic Lab., General Electric, Syracuse, New York.

58. Nishimura, T. 1967. Error bounds of continuous Kalman filters and the application to orbit determination problems. I. E. E. E. Trans. on Automatic Control, Vol AC-12: 268-275.

59. Norris, R. C. 1960. A Method for Locating the Maximum of a Function of a single variable in the presence of noise. Report 22G-0035, Lincoln Laboratory, M. I. T., Cambridge, Mass. 33 pp.

60. O'Donnell, T. 1960. Instantaneous Unit Hydrograph Derivation by Harmonic Analysis. Inter. Assoc. Scientific Hydrology Pub. 51: 546-557.

61. Ogata, Katsuhiko. 1970. Modern Control Engineering. Prentice-Hall, Inc. Englewood Cliffs, N. J.

62. Peschon, J. and R. E. Larson. 1965. Analysis of an Intercept System. SRI Project 5188, Stanford Research Institute.

63. Pierre, D. A. 1969. Optimization Theory with Applications. John Wiley & Sons, Inc. New York.

64. Pipes, L. A. 1963. Matrix methods for Engineering. Prentice-Hall Inc. Englewood Cliffs, N. J. pp 90-92.

65. Powell, M. J. D. 1964. An Efficient Method for Finding the Minimum of a Function of Several Variables without calculating derivatives. The Computer Journal, 7, 155-162.

66. Prasad, R. 1967. Nonlinear Simulation of a Regional Hydrologic System. Ph. D. Thesis, University of Illinois.

67. Price, C. F. 1968. An analysis of the divergence problem in the Kalman filter. I. E. E. E. Trans. on Automatic Control (short papers) Vol. AC-13: 699-702.

68. Rao, R. A. and J. W. Delleur. 1971. The Instantaneous Unit Hydrograph: its Calculation by the Transform Method and Noise Control by Digital Filtering. Tech. Report No. 20, Purdue University, Water Resources Research Center.

69. Rosenbrock, H. H. 1960. An Automatic Method for Finding the Greatest or the Least Value of a Function. The Computer Journal, 3, 175-184.

70. Sage, A. P. and R. W. Burt. 1965. Optimum design and error analysis of digital integrators for discrete system simulation. A. F. I. P. S. Proc. FJCC 27: 903-914. Las Vegas, Nevada.

71. Sage, A. P. and S. L. Smith. 1966. Real-time digital simulation
     for systems control. Proc. I. E. E. E. 54: 1802-1812.

72. Sage, A. P. and G. W. Husa. 1969. Algorithms for sequential
     adaptive estimation of prior statistics. Proc. 8th I. E. E. E.
     Symposium Adaptive Processes, Pennsylvania State University.

73. Sage, A. P. and Melsa, J. L. 1971. Estimation Theory with appli-
     cation to Communications and Control. McGraw-Hill, New York.

74. Schlee, F. H., Standish, C. J. and N. F. Toda. 1967. Divergence
     in the Kalman Filter. A. I. A. A. Journal, Vol. 5 (6):
     1114-1120.

75. Schmidt, S. F. 1967. Estimation of state with acceptable accuracy
     constraints. Analytical Mechanics Associates, Inc. Interim
     Report 67-4.

76. Schmidt, S. F., Weinberg, J. D. and J. S. Lukesh. 1968. Case
     study of Kalman Filtering in the C-5 Aircraft Navigation
     System. Case Study Workshop, 1968 JACC. The University of
     Michigan, 59-109.

77. Schumer, M. A. and K. Steiglitz. 1968. Adaptive step size random
     search. I. E. E. E. Trans. on Automatic Control, Vol. AC-13(3).

78. Schwartz, L. and E. B. Stear. 1968. A computational comparison of
     several nonlinear filters. I. E. E. E. Trans. on Automatic
     Control Vol. AC-13: 83-86.

79. Shah, B. V., Buehler, R. J. and O. Kempthorne. 1964. Some
     algorithms for minimizing a function of several variables.
     J. of the Soc. for Ind. and Appl. Math. 12: 74-92.

80. Sherman, L. K. 1932. Stream Flow from Rainfall by Unit-Graph
     Method. Engineering News-Record 108: 501-505.

81. Singh, Krishan Piara. 1962. A Non-linear Approach to the
     Instantaneous Unit Hydrograph Theory. Ph. D. Thesis,
     University of Illinois, Dept. of Civil Engineering.

82. Sklansky, J. 1966. Learning Systems for Automatic Control.
     I. E. E. E. Trans. on Automatic Control, Vol. AC-11(1): 6-19.

83. Snyder, W. M. 1962. Some possibilities for multivariate analysis
     in hydrologic studies. J. Geophys. Research, 67: 721-729.

84. Sorenson, H. W. 1968. Controllability and Observability of
     Linear, Stochastic, Time-Discrete Control Systems. In:
     Advances in Control Systems, Theory and Applications, Vol.6:
     95-158.

85. Tamura, Hiroshi. 1971. Self-learning method for time-optimal Control. In: Pattern Recognition and Machine Learning. K. S. Fu, ed.

86. Tapley, B. D. and G. H. Born. 1971. Sequential Estimation of the State and Observation-error Covariance Matrix. A. I. A. A. Journal, Vol. 9(2): 212-217.

87. Tarn, T. J. and J. Zaborszky. 1970. A Practical, Nondiverging Filter. A. I. A. A. Journal, Vol. 8(6): 1127-1133.

88. Tomovic, R. 1962. Sensitivity Analysis of Dynamic Systems. McGraw-Hill Book Co., New York.

89. Tou, J. T. 1964. Modern Control Theory. McGraw-Hill Book Co., New York.

90. Unger, F. G. and N. Ott. 1970. Optimal Use of redundant information in an inertial navigation system. In: Theory and Applications of Kalman Filtering. AGARDograph 139. NATO.

91. Vermuri, V. and J. A. Dracup. 1967. Analysis of nonlinearities in ground-water hydrology: A hybrid computer approach. Water Resources Res. 3(4): 1047-1058.

92. Vermuri, V., Dracup, J. A., and R. C. Erdmann. 1969. Sensitivity Analysis Method of System Identification and its Potential in Hydrologic Research. Water Resources Res. 5(2): 341-349.

93. Weiss, I. M. 1970. A Survey of discrete Kalman-Bucy filtering with unknown noise covariances. Presented at the A. I. A. A. Guidance, Control, and Flight Mechanics Conf. Santa Barbara, Calif. 70-p 995.

94. White, R. C., Jr. 1971. A Survey of Random Methods for parameter optimization. Simulation: 197-204.

95. Wilde, D. J. 1964. Optimum Seeking Methods. Prentice-Hall, Englewood Cliffs, N. J.

96. Wishner, R. P., Tabaczynski, J. A. and M. Athans. 1968. A Comparison of Three Nonlinear Filters. Presented at the IFAC Symposia on Multivariable Control Systems, Dusseldorf, Germany; also in Automatica, Vol. 5: 487-496. Pergamon Press, 1969.

97. Yamaoka, I. and M. Fujita. 1969. Evaluation of Simulation Models for River Runoff through Nyquist Plots. Proc. XIIIth Congress IAHR, Kyoto, Vol. 1: 171-180.

98. Yeh, W. W-G., and G. W. Tauxe. 1971. A Proposed Technique for Identification of Unconfined Aquifer Parameters. _J. of Hydrology 12_: 117-128.

99. Yevjevich, Vujica. 1971. The Structure of Inputs and Outputs of Hydrologic Systems. _Proc. of the First Bilateral United States-Japan Seminar in Hydrology_, Honolulu, Jan. 1971.

100. Young, P. C. 1965. The determination of the Parameters of a Dynamic Process. _The Radio and Electronic Engineer_, 29: 345-362.

101. Zadeh, L. A. 1962. From circuit theory to system theory. _Proc. Institute of Radio Engineers_, Vol. 50: 856-865.

102. Zadeh, L. A. 1963. On the definition of Adaptivity. _Proc. of the I. E. E. E._ 51: 469-470.

103. Zakharev, V. V. 1969. A Random Search Method. _Engineering Cybernetics_, No. 2: 26-30.

APPENDIX A

PROOF OF THE SEPARATION PRINCIPLE FOR

THE LINEAR OPTIMAL STOCHASTIC CONTROL PROBLEM

Problem:

Given the model

$$\delta x_n = \Phi_{n,n-1} \delta x_{n-1} + \Gamma_{n,n-1} \delta u_{n-1} + w_{n-1} \tag{A-1}$$

$$\delta y_n = H_n \delta x_n + v_n \tag{A-2}$$

where

$$E\left\{\delta x_o\right\} = 0 \qquad , \qquad E\left\{\delta x_o \, \delta x_o^T\right\} = P_o \qquad ,$$

$$E\left\{w_n\right\} = E\left\{v_n\right\} = 0 \quad \text{for all } n \qquad ,$$

$$E\left\{w_n w_m^T\right\} = Q_n \, \delta_{nm} \quad , \quad E\left\{v_n v_m^T\right\} = R_n \, \delta_{nm} \qquad .$$

The various random variables are also assumed to be mutually uncorrelated, so

$$E\left\{\delta x_o w_n^T\right\} = E\left\{\delta x_o v_n^T\right\} = 0 \quad \text{for all } n \qquad ,$$

and $\qquad E\left\{w_n v_m^T\right\} = 0 \quad \text{for all } n, m.$

Choose the N control vectors $\delta u_i \quad (i=0,1,2,\ldots,N-1)$ as functions of the measurements $\delta y_o, \delta y_1, \ldots, \delta y_i$ so that the performance index

$$E\left\{J_N\right\} = E\left\{\sum_{j=1}^{N} (\delta x_j^T A_j \delta x_j + \delta u_{j-1}^T B_{j-1} \delta u_{j-1})\right\} \tag{A-3}$$

is minimized.

Separation Principle:

For the model described by Eqns. (A-1), (A-2) and (A-3), the optimal stochastic control law is described by

$$\delta u^{*}_{N-k-1} = - \Lambda_{N-k} \, \Phi_{N-k,N-k-1} \, \delta \hat{x}_{N-k-1} \qquad (A-4)$$

where Λ_{N-k} is the control gain obtained by solving the deterministic control problem, δx_{N-k-1} represents the optimal linear estimate of the state obtained from the set of measurements δY_{N-k-1}. In obtaining the estimate, δU_{N-k-2} is treated as a deterministic quantity.

Using the definition of expectation, Eqn. (A-3) can be written as

$$
\begin{aligned}
E\left\{J_N\right\} &= \int \sum_{j=1}^{N} (\delta x_j^T A_j \delta x_j + \delta u_{j-1}^T B_{j-1} \delta u_{j-1}) p(\delta X_N, \delta Y_{N-1}) d(\delta X_N, \delta Y_{N-1}) \\
&= \int \sum_{j=1}^{N-1} (\delta x_j^T A_j \delta x_j + \delta u_{j-1}^T B_{j-1} \delta u_{j-1}) p(\delta X_N, \delta Y_{N-1}) d(\delta X_N, \delta Y_{N-1}) \\
&\quad + (\delta x_N^T A_N \delta x_N + \delta u_{N-1}^T B_{N-1} \delta u_{N-1}) p(\delta X_N, \delta Y_{N-1}) d(\delta X_N, \delta Y_{N-1})
\end{aligned}
$$

$$(A-5)$$

where $\int$ denotes multiple integrals and

$$d(\delta X_N, \delta Y_{N-1}) = d(\delta x_o) d(\delta x_1) \ldots d(\delta x_N) d(\delta y_o) d(\delta y_1) \ldots d(\delta y_{N-1})$$

But by hypothesis, the control δu_i is determined from the measurements δY_i, so the δx_N and δy_{N-1} can be integrated out of the first term of Eqn. (A-5)

$$E\left\{J_N\right\} = \int \sum_{j=1}^{N-1} (\delta x_j^T A_j \delta x_j + \delta u_{j-1}^T B_{j-1} \delta u_{j-1}) p(\delta X_{N-1}, \delta Y_{N-2}) d(\delta X_{N-1}, \delta Y_{N-2})$$

$$+ \ (\delta x_N^T A_N \delta x_N + \delta u_{N-1}^T B_{N-1} \delta u_{N-1}) p(\delta X_N, \delta Y_{N-1}) d(\delta X_N, \delta Y_{N-1}) \qquad (A-6)$$

Consider the last control interval, only the last term of Eqn. (A-6) is involved. Using the property of joint probability density function, one has

$$p(\delta X_N, \delta Y_{N-1}) \ = \ p(\delta x_N | \delta X_{N-1}, \delta Y_{N-1}) p(\delta X_{N-1}, \delta Y_{N-1}) \qquad (A-7)$$

and

$$p(\delta X_N, \delta Y_{N-1}) \ = \ \int p(\delta x_N | \delta X_{N-1}, w_{N-1}, \delta Y_{N-1}) p(w_{N-1} | \delta X_{N-1}, \delta Y_{N-1})$$

$$\cdot p(\delta X_{N-1}, \delta Y_{N-1}) dw_{N-1} \qquad (A-8)$$

But the set of measurements δY_{N-1} determines δu_{N-1}, and δx_{N-1}, w_{N-1} and δu_{N-1} completely define δx_N by Eqn. (A-1). Thus

$$p(\delta x_N | \delta x_{N-1}, w_{N-1}, \delta u_{N-1}) \ = \ \delta(\delta x_N - \Phi_{N,N-1} \delta x_{N-1} - \Gamma_{N,N-1} \delta u_{N-1} - w_{N-1})$$

$$(A-9)$$

where $\delta(.)$ is the Dirac-delta function.

Since w_i is uncorrelated with δx_i and δy_i, one has

$$p(w_{N-1} | \delta X_{N-1}, \delta Y_{N-1}) \ = \ p(w_{N-1}) \qquad (A-10)$$

Using Eqns. (A-8), (A-9) and (A-10), $E\left\{J_1\right\}$ can now be written as

$$E\left\{J_1\right\} \ = \ \int (\delta x_N^T A_N \delta x_N + \delta u_{N-1}^T B_{N-1} \delta u_{N-1})$$

$$\cdot \ \delta(\delta x_N - \Phi_{N,N-1} \delta x_{N-1} - \Gamma_{N,N-1} \delta u_{N-1} - w_{N-1})$$

$$\cdot \ p(w_{N-1}) p(\delta X_{N-1}, \delta Y_{N-1}) d(w_{N-1}, \delta X_N, \delta Y_{N-1}) \qquad (A-11)$$

Integrate Eqn. (A-11) with respect to δx_N and w_{N-1}, then

$$E\left\{J_1\right\} = \int\left[\delta x_{N-1}^T \Phi_{N,N-1}^T A_N \Phi_{N,N-1}\delta x_{N-1} + 2\delta x_{N-1}^T \Phi_{N,N-1}^T A_N \Gamma_{N,N-1}\delta u_{N-1}\right.$$

$$+ \delta u_{N-1}^T (B_{N-1} + \Gamma_{N,N-1}^T A_N \Gamma_{N,N-1})\delta u_{N-1}$$

$$\left.+ \operatorname{trace}(A_N Q_{N-1})\right] p(\delta X_{N-1}, \delta Y_{N-1})d(\delta X_{N-1}, \delta Y_{N-1}) \qquad \text{(A-12)}$$

only δx_{N-1} occurs explicitly in (A-12), so integration will eliminate δX_{N-2}. The control is to be computed as a function of the δY_{N-1}, so write

$$p(\delta x_{N-1}, \delta Y_{N-1}) = p(\delta x_{N-1}|\delta Y_{N-1})p(\delta Y_{N-1})$$

and choose δu_{N-1} to minimize

$$\int\left[\delta u_{N-1}^T (B_{N-1} + \Gamma_{N,N-1}^T A_N \Gamma_{N,N-1})\delta u_{N-1} + 2\delta x_{N-1}^T \Phi_{N,N-1}^T A_N \Gamma_{N,N-1}\delta u_{N-1}\right.$$

$$\cdot p(\delta x_{N-1}|\delta Y_{N-1})dx_{N-1}$$

$$= \delta u_{N-1}^T (B_{N-1} + \Gamma_{N,N-1}^T A_N \Gamma_{N,N-1})\delta u_{N-1} +$$

$$+ 2 E\left\{\delta x_{N-1}^T \delta Y_{N-1}\right\} \Phi_{N,N-1}^T A_N \Gamma_{N,N-1}\delta u_{N-1} \qquad \text{(A-13)}$$

The optimal control δu_{N-1}^* is

$$\delta u_{N-1}^* = -(B_{N-1} + \Gamma_{N,N-1}^T A_N \Gamma_{N,N-1})^{-1}\Gamma_{N,N-1}^T A_N \Phi_{N,N-1}\delta \hat{x}_{N-1} \qquad \text{(A-14)}$$

where $\delta \hat{x}_{N-1} = E\left\{\delta x_{N-1}|\delta Y_{N-1}\right\}$ is the minimum variance estimate of δx_{N-1}.

Let
$$\Lambda_N \triangleq (B_{N-1} + \Gamma_{N,N-1}^T A_N \Gamma_{N,N-1})^{-1}\Gamma_{N,N-1}^T A_N \qquad \text{(A-15)}$$

Eqn. (A-14) then becomes

$$\delta u_{N-1}^* = -\Lambda_N \Phi_{N,N-1}\delta \hat{x}_{N-1} \qquad \text{(A-16)}$$

Thus, the separation principle is true for the last stage.

The control (A-16) allows (A-11) to be rewritten as

$$E\left\{J_1^*\right\} = \int\left[\left[\delta x_{N-1}^T \Phi_{N,N-1}^T A_N \Phi_{N,N-1}\delta x_{N-1} - 2\delta x_{N-1}^T \Phi_{N,N-1}^T A_N \Gamma_{N,N-1}\Lambda_N \Phi_{N,N-1}\delta\hat{x}_{N-1}\right.\right.$$

$$\left.+ \delta\hat{x}_{N-1}^T \Phi_{N,N-1}^T A_N \Gamma_{N,N-1}\Lambda_N \Phi_{N,N-1}\delta\hat{x}_{N-1}\right] p(\delta X_{N-1},\delta Y_{N-1})$$

$$\cdot d(\delta X_{N-1},\delta Y_{N-1}) + \text{trace}(A_N Q_{N-1}) \qquad\qquad (A-17)$$

Let $\delta\tilde{x}_{N-1} \triangleq \delta\hat{x}_{N-1} - \delta x_{N-1}$ represent the error in the estimate, Eqn. (A-17) can then be rewritten in simpler form as

$$E\left\{J_1^*\right\} = \int\left[\left[\delta x_{N-1}^T \Phi_{N,N-1}^T A_N \Phi_{N,N-1}\delta x_{N-1} + \delta\tilde{x}_{N-1}^T \Phi_{N,N-1}^T A_N \Gamma_{N,N-1}\Lambda_N \Phi_{N,N-1}\delta\tilde{x}_{N-1}\right]\right.$$

$$\cdot p(\delta X_{N-1},\delta Y_{N-1})d(\delta X_{N-1},\delta Y_{N-1}) + \text{trace}(A_N Q_{N-1}) \qquad (A-18)$$

Using again the property of joint probability density function, one gets

$$E\left\{J_1^*\right\} = \int\left[\left[\delta x_{N-1}^T \Phi_{N,N-1}^T A_N \Phi_{N,N-1}\delta x_{N-1} + \delta\tilde{x}_{N,N-1}^T \Phi_{N,N-1}^T A_N \Gamma_{N,N-1}\Lambda_N\right.\right.$$

$$\left.\cdot \Phi_{N,N-1}\delta\tilde{x}_{N-1}\right]\cdot\delta(\delta y_{N-1} - H_{N-1}\delta x_{N-1} - v_{N-1})p(v_{N-1})$$

$$\cdot p(\delta X_{N-1},\delta Y_{N-2})d(v_{N-1},\delta X_{N-1},\delta Y_{N-1}) \qquad\qquad (A-19)$$

The first term of the integrand does not depend upon v_{N-1} and the error $\delta\tilde{x}_{N-1}$ is not affected by the control δu_{N-2}, so

$$E\left\{J_1^*\right\} = \int \delta x_{N-1}^T \Phi_{N,N-1}^T A_N \Phi_{N,N-1} \delta x_{N-1} p(\delta X_{N-1}, \delta Y_{N-2}) d(\delta X_{N-1}, \delta Y_{N-2})$$

$$+ \text{trace}(\Phi_{N,N-1}^T A_N \Gamma_{N,N-1} \Lambda_N \Phi_{N,N-1} P_{N-1} + A_N Q_{N-1}) \qquad (A-20)$$

where $\qquad E\left\{\delta x_{N-1} \delta x_{N-1}^T\right\} \triangleq P_{N-1}$

Now, consider the last two control intervals

$$E\left\{J_2\right\} = \int (\delta x_{N-1}^T A_{N-1} \delta x_{N-1} + \delta u_{N-2}^T B_{N-2} \delta u_{N-2}) p(\delta X_{N-1}, \delta Y_{N-2}) d(\delta X_{N-1}, \delta Y_{N-2})$$

$$+ \int (\delta x_N^T A_N \delta x_N + \delta u_{N-1}^T B_{N-1} \delta u_{N-1}) p(\delta X_N, \delta Y_{N-1}) d(\delta X_N, \delta Y_{N-1}) \qquad (A-21)$$

The principle of optimality allows one to use expression (A-16) for δu_{N-1}. Thus, with Eqn. (A-20), the above equation becomes

$$E\left\{J_2\right\} = \int \left[\delta x_{N-1}^T (A_{N-1} + \Phi_{N,N-1}^T A_N \Phi_{N,N-1}) \delta x_{N-1} + \delta u_{N-2}^T B_{N-2} \delta u_{N-2} \right.$$

$$\cdot p(\delta X_{N-1}, \delta Y_{N-2}) d(\delta X_{N-1}, \delta Y_{N-2})$$

$$+ \text{trace}(\Phi_{N,N-1}^T A_N \Gamma_{N,N-1} \Lambda_N \Phi_{N,N-1} P_{N-1} + A_N Q_{N-1}) \qquad (A-22)$$

Proceeding in a manner analogous to that employed in deriving Eqns. (A-11) - (A-13), one obtains

$$\delta u_{N-2}^* = - \Lambda_{N-1} \Phi_{N-1,N-2} \delta \hat{x}_{N-2} \qquad (A-23)$$

where Λ_{N-1} is determined by

$$\Lambda_{N-1} = (\Gamma_{N-1,N-2}^T L_{N-1}' \Gamma_{N-1,N-2} + B_{N-2})^{-1} \Gamma_{N-1,N-2}^T L_{N-1}' \qquad (A-24)$$

$$L_{N-1}' = \Phi_{N,N-1}^T L_{N-1} \Phi_{N,N-1} + A_{N-1} \qquad (A-25)$$

$$L_{N-1} = L'_{N-1} - L'_{N-1} \Gamma_{N-1, N-2} \Lambda_{N-1} \qquad (A-26)$$

The cost associated with δu^*_{N-2} is found to be

$$E\left\{J^*_2\right\} = \int \delta x^T_{N-2} \Phi^T_{N-1,N-2} L_{N-1} \Phi_{N-1,N-2} \delta x_{N-2} p(\delta X_{N-2}, \delta Y_{N-3}) d(\delta X_{N-2}, \delta Y_{N-3})$$

$$+ \text{ trace } \Phi^T_{N,N-1} L'_N \Gamma_{N,N-1} \Lambda_N \Phi_{N,N-1} P_{N-1}$$

$$+ \Phi^T_{N-1,N-2} L'_{N-1} \Gamma_{N-1,N-2} \Lambda_{N-1} \Phi_{N-1,N-2} P_{N-2}$$

$$+ L'_N Q_{N-1} + L'_{N-1} Q_{N-2} \qquad (A-27)$$

The proof of the general result is done inductively. For any k one can assume that

$$\delta u^*_{N-k-1} = -\Lambda_{N-k} \Phi_{N-k,N-k-1} \delta x_{N-k-1} \qquad (A-28)$$

where
$$\Lambda_{N-k} = (\Gamma^T_{N-k,N-k-1} L'_{N-k} \Gamma_{N-k,N-k-1} + B_{N-k-1})^{-1}$$

$$\cdot \Gamma^T_{N-k,N-k-1} L'_{N-k} \qquad (A-29)$$

$$L'_{N-k} = \Phi^T_{N-k+1,N-k} L_{N-k+1} \Phi_{N-k+1,N-k} + A_{N-k} \qquad (A-30)$$

$$L_{N-k} = L'_{N-k} - L'_{N-k} \Gamma_{N-k,N-k-1} \Lambda_{N-k} \qquad (A-31)$$

The cost at stage k is

$$E\left\{J^*_k\right\} = \int \delta x^T_{N-k} \Phi^T_{N-k+1,N-k} L_{N-k+1} \Phi_{N-k+1,N-k} \delta x_{N-k} p(\delta X_{N-k}, \delta Y_{N-k-1})$$

$$\cdot d(\delta X_{N-k}, \delta Y_{N-k-1}) + \text{ trace } \left[\sum_{j=1}^{k} \Phi^T_{N-j+1,N-j} L'_{N-j+1} \Gamma_{N-j+1,N-j} \right.$$

$$\left. \cdot \Lambda_{N-j+1} \Phi_{N-j+1,N-j} P_{N-j} + \sum_{j=1}^{k} L'_{N-j+1} Q_{N-j} \right] \qquad (A-32)$$

The proof for the (k+1)st stage is accomplished in the same way so that details shall be omitted. Eqns. (A-28) – (A-32) define the optimal stochastic control policy. This completes the proof of the Separation Principle.

APPENDIX B

PROGRAM DESCRIPTION

<u>SUBROUTINE ADAPT</u> (MX, MEAS, DELT, NSTEPS, AJ, BJ, OX, R, Q, CP, CONTR, Z)

** This subroutine estimates the state-variables of a linear system

noisy observations using linear stochastic control approach. The

optimal estimates are given by a Kalman filter with adaptive estimation

of the noise covariance matrices **

<u>Working parameters</u>:

MX = number of state-variables,

MEAS = number of observations,

DELT = time-period for sub-intervals,

NSTEPS = number of sub-intervals,

AJ(.,.) = weighting matrix for the state-variables,

BJ = weighting factor for the control input,

OX(.) = set of initial estimates of the state-variables,

R = observation-error variance,

Q(.,.) = model-error covariance matrix,

CP(.,.) = estimation-error covariance matrix,

CONTR(.) = set of inputs,

Z(.) = set of observations.

<u>Subroutines required</u>: STATE, HAD, ADJUSTR, GAIN, STAEST, ERVAR, FPHI,

PREDICT, CGAIN, GAMMA, UPCON.

<u>SUBROUTINE STATE</u> (OX, DELT, CTR)

** This subroutine computes the next value of the state of the

system **

<u>Working parameters</u>: OX, DELT

CTR = the control input.

<u>Subroutines required</u>: None.

SUBROUTINE HAD (MX, OXP, H)

** This subroutine computes the observation matrix **

Working parameters: MX

OXP(.) = set of predicted state-variables,

H(.) = the observation matrix.

Subroutines required: None

SUBROUTINE ADJUSTR (MX, R, WG, DZ, H, XP, DEL)

** This subroutine adjusts the value of the observation-error variance **

Working parameters: MX, R, H

WG = observation weight,

DZ = measurement residual,

XP(.) = set of predicted state-variables,

DEL = difference between computed and measured residuals.

Subroutines required: None.

SUBROUTINE GAIN (MX, PP, H, R, GK)

** This subroutine computes the Kalman gains **

Working parameters: MX, H, R

PP(.,.) = predicted estimation-error covariance matrix,

GK(.) = vector of Kalman gains.

Subroutines required: None.

SUBROUTINE STAEST (MX, GK, DEL, XP, X)

** This subroutine gives the best estimates of the state-variables. **

Working parameters: MX, GK, DEL, XP

X(.) = set of best estimates of the state-variables.

Subroutines required: None.

SUBROUTINE ERVAR (MX, PP, GK, H, R, CP)

** This subroutine updates the estimation-error covariance matrix **

Working parameters: MX, PP, GK, H, R

   CP(.,.) = estimation-error covariance matrix.

Subroutine required: None.

SUBROUTINE FPHI (MX, OX, CONTR, DELT, NSTEPS, PHI)

** This subroutine computes the system state-transition matrix. **

Working parameters: MX, OX, CONTR, DELT, NSTEPS

   PHI(.,.) = the system state-transition matrix.

Subroutines required: STATE.

SUBROUTINE PREDICT (MX, PHI, GAM, CONGK, CP, Q, X, XP, PP)

** This subroutine computes predicted values of the state-variables
and estimation-error covariance matrix **

Working parameters: MX, PHI, GAM, CONGK, CP, Q, X, XP

   PP(.,.) = predicted estimation-error covariance matrix.

Subroutines required: None.

SUBROUTINE CGAIN (MX, MEAS, AJ, BJ, GAMS, PHIS, CGK, OL)

** This subroutine computes the control gains **

Working parameters: MX, MEAS, AJ, BJ, GAMS, PHIS, CGK

   OL(.,.,.) = set of control cost matrices.

Subroutines required: None

SUBROUTINE GAMMA (MX, OX, PHI, GAM)

** This subroutine is used to compute the model-error covariance
matrix Q. **

Working parameters: MX, OX, PHI

   GAM(.) = control-input matrix.

Subroutines required: None.

SUBROUTINE UPCON (MX, PHI, GAM, OL, AJ, BJ, CONGK, KK)

**   This subroutine is used to update the control law **

Working parameters: MX, PHI, GAM, OL, AJ, BJ, KK

    CONGR(.) = set of control gains.

Subroutines required: None.

SUBROUTINE ITERA (MX, MEAS, DELT, NSTEPS, X, R, QEL, CP, CONTR, Z)

**   This subroutine estimates the state-variables of a nonlinear system

through the iterated extended Kalman filter with adaptive estimation of

the noise covariance matrices **

Working parameters: MX, MEAS, DELT, NSTEPS, R, CP

    X(.) = state-variable vector

    QEL = variance element of the model-error covariance matrix.

Subroutines required: ISTATE, IHAD, IADJUSTR, GAIN, STAEST, ERVAR,

                                 IFPHI, IPREDICT, IGAMMA.

SUBROUTINE ISTATE (X, DELT, CONTR)

**   This subroutine computes the next value of the state from model

equations **

Working parameters: X, DELT

    CONTR = the control input.

Subroutines required: None.

SUBROUTINE IHAD (MX, XP, H, HSM)

**   This subroutine computes the observation matrix and its partial

derivatives **

Working parameters: MX, H

    XP(.) = set of predicted state-variables,

    HSM = h(X) = observation-vector.

Subroutines required: None.

SUBROUTINE IADJUSTR (MX, R, WG, Z, HSM, H, OXP, SX, X, DEL)

** This subroutine adjusts the value of the observation-error

variance **

Working parameters: MX, R, WG, Z, HSM, H, OXP, X, DEL SX(.) = set of

stored predicted-values of the state-variables.

Subroutines required: None.

SUBROUTINE IFPHI (MX, X, CONTR, DELT, NSTEPS, PHI)

** This subroutine computes the system-state-transition matrix **

Working parameters: MX, X, CONTR, DELT, NSTEPS, PHI.

Subroutines required: ISTATE.

SUBROUTINE IPREDICT (MX, PHI, GAM, CP, QEL, X, NSTEPS, DELT, CONTR,

      XP, PP)

** This subroutine computes predicted values of the state-variables

and estimation-error covariance matrix **

Working parameters: MX, PHI, GAM, CP, QEL, X, NSTEPS, DELT, CONTR,

      XP, PP.

Subroutines required: None.

SUBROUTINE IGAMMA (MX, X, GAM)

** This subroutine is used to compute the model-error covariance matrix

Q. **

Working parameters: MX, X, GAM.

Subroutines required: None.

SUBROUTINE ADJUSTQ (MX, PHI, H, CP, R, SN, DEL, WG, Q)

** This subroutine adjusts the value of the model-error covariance

matrix Q **

<u>Working parameters</u>: MX, PHI, H, CP, R, SN, DEL, WG, Q

<u>Subroutines required</u>: None.