R80-28 OSP 87627

MIT

A STATISTICAL LINEARIZATION APPROACH TO REAL TIME NONLINEAR FLOOD ROUTING

by
Konstantine P. Georgakakos
and
Rafael L. Bras

DEPARTMENT OF CIVIL ENGINEERING

RALPH M. PARSONS LABORATORY
FOR
WATER RESOURCES AND HYDRODYNAMICS

Report No. 256

The work upon which this publication is based was sponsored by the Hydrologic Research Laboratory, National Weather Service, U.S. Department of Commerce

SCHOOL OF ENGINEERING
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Cambridge, Massachusetts 02139

June 1980

R80-28

OSP 87627

A STATISTICAL LINEARIZATION APPROACH TO REAL TIME NONLINEAR FLOOD ROUTING

by

Konstantine P. Georgakakos

and

Rafael L. Bras

Ralph M. Parsons Laboratory

for

Water Resources and Hydrodynamics

Massachusetts Institute of Technology

June 1980

Report No. 256

The work upon which this publication is based was sponsored by the Hydrologic Research Laboratory, National Weather Service, U.S. Department of Commerce

A STATISTICAL LINEARIZATION APPROACH TO REAL TIME NONLINEAR FLOOD ROUTING

ABSTRACT

This work examines the flood routing problem. A nonlinear router based on a series of reservoirs served as the model and modern estimation theory techniques were used to improve its performance in real time river discharge forecasting. The model was statistically linearized to become compatible with a Gaussian minimum variance estimator. The Taylor-Gauss methodology was proposed for the analytical determination of the expected value of nonlinear functions when the independent variable is approximately normally distributed. Having bypassed the heavy computational requirements for the numerical calculation of the statistical linearization gains, a recursive estimator for the states and parameters was designed. The outcome of the synthesis procedure described, the stochastic flood routing model, was used in a real world application to forecast six-hour discharge values at the Bird Creek drainage basin in Oklahoma, U.S.A. Procedures for the determination of initial parameter values are described. An adaptive run with suboptimal error statistics showed fast convergence in the direction of minimum average squared error.

ACKNOWLEDGEMENTS

This report essentially constitutes the Masters degree thesis of Mr. Konstantine Georgakakos. The work was sponsored by the U.S. Department of Commerce, National Weather Service, Contract No. NA79SAC00650. The authors are grateful for the continuing cooperation of the staff of the Hydrologic Research Laboratory, N.W.S. Administrative support was provided by M.I.T. through the Office of Sponsored Programs, OSP 87627. All the computer work was performed at the M.I.T. Information Processing Center on a Honeywell 6180, Multics Operating System.

Particular words of appreciation go to Prof. E. M. Laurenson of Monash University, Australia, and to Prof. A. Koussis of the University of Florida at Gainesville for many discussions on the physical aspects of flood routing.

Ms. Anne Clee deserves recognition for her patient, efficient and fast typing of this material with its innumerable equations. Ms. Carolyn Comer typed Chapters 6 and 7.

Messrs. Pedro Restrepo-Posada, David Ostendorf and Telis Mantoglou participated in many useful discussions.

TABLE OF CONTENTS

Page No.

LIST OF FIGURES

LIST OF TABLES

LIST OF PRINCIPAL SYMBOLS

Symbol	Description	Equation No.
a	coefficient in the outflow function of a conceptual reservoir	2.12
a_i	coefficient in the outflow function of the i^{th} reservoir in a cascade of conceptual reservoirs or of the i^{th} channel reach	2.27
A	cross sectional area of flow	2.7
b	channel width	2.1
b_i	channel width in the i^{th} reach	2.21
c	wave celerity	2.14
$\bar{c}$	spatial average value of wave celerity in a channel reach	2.15
C	Chezy's constant coefficient	2.5
C_i	Chezy's constant coefficient in the i^{th} channel reach	2.22
D	difussion coefficient	2.9
$e(t)$	error of statistical linearization	4.12
E_1	absolute percent error	3.79
E_2	absolute percent error	3.80
f	infiltration rate	2.1
$\underline{\bar{F}}(t)$	system matrix	5.44
g	gravitational acceleration	2.2
$\underline{\bar{H}}$	Hessian matrix	4.39

Symbol	Description	Equation No.
$\underline{\bar{H}}(t_k)$	measurement gradient matrix at time t_k	5.45
$\underline{\bar{H}}_a(t_k)$	measurement gradient matrix of parameters	5.137
$\underline{\bar{H}}_x(t_k)$	measurement gradient matrix of states	5.137
i	rainfall intensity	2.1
I_i	inflow discharge at the i^{th} grid point or conceptual reservoir	2.15
$\underline{k}_a(t_k)$	parameter estimator gain	5.140
$\underline{k}_x(t_k)$	state estimator gain	5.137
$\underline{\bar{K}}(t_k)$	Gaussian minimum variance estimator gain at time t_k	5.65
L_i	Length of the i^{th} channel reach	2.21
m	exponent of the outflow function of a conceptual reservoir	2.27
N_{b_i}	statistical linearization bias gain	4.32
N_{a_i}	statistical linearization residual gain	4.32
N_{x_i}	statistical linearization residual gain	4.32
p_i	proportion of the total channel inflow that serves as an input to the i^{th} reach or conceptual reservoir	2.32
p_1	percent error	3.81
p_2	percent error	3.82

Symbol	Description	Equation No.
p_3	percent error	3.83
$p_x(x; \underline{a})$	probability density function of the random variable X evaluated at x, with parameter vector $\underline{a}$	3.5
$\bar{\underline{P}}(t/t_k)$	error covariance matrix corresponding to the estimate of the state at time t given observations up to and including time t_k	5.51
$\bar{\underline{P}}_y(t_{k+1}/t_k)$	error covariance matrix of prediction corresponding to the observation process	5.59
$\bar{\underline{P}}_{xy}(t_{k+1}/t_k)$	error covariance matrix of prediction corresponding to the observation process	5.61
$\bar{\underline{P}}_x(t)$	estimation error covariance matrix of states	5.129
$\bar{\underline{P}}_{xa}(t)$	estimation error cross-covariance matrix between states and parameters	5.129
$\bar{\underline{P}}_a(t)$	estimation error covariance matrix of parameters	5.129
q_L	continuous lateral channel inflow	2.1
Q	discharge of water	2.12
$\bar{\underline{Q}}(t)$	white noise intensity matrix at time t	5.51
$\bar{\underline{Q}}_a(t)$	white noise intensity matrix of parameters	5.135
Q_i	outflow discharge from the i^{th} conceptual reservoir	2.20
Q_m	maximum input discharge	B.1

Symbol	Description	Equation No.
Q_{p_i}	maximum outflow rate from the i^{th} conceptual reservoir	2.36
$\bar{Q}_x(t)$	white noise intensity matrix of states	5.130
r_{a_i}	normally distributed residual parameter process	4.30
$r_u(t)$	normally distributed residual input process	5.79
r_{x_i}	normally distributed residual state process	4.29
$\hat{\underline{r}}_a(t_{k+1}/t_{k+1})$	minimum variance estimate of parameter vector of time t_{k+1}	5.144
$\hat{\underline{r}}_x(t_{k+1}/t_{k+1})$	minimum variance estimate of state vector at time t_{k+1}	5.143
R	hydraulic radius	2.5
$\bar{\underline{R}}(t_k)$	covariance matrix of measurement noise	5.49
$R1_{N'}$	expected value of the N'^{th} term in a Taylor's series expansion of a nonlinear function about its mean	3.37
$R2_{N'}$	expected value of the N'^{th} term in a Taylor's series expansion of a nonlinear function about its mean	3.38
S_f	friction slope	2.4
S_i	volume of water in storage in the i^{th} conceptual reservoir	2.20
$S1_{N'}$	ratio of the N' vs. N'-1 terms in the Taylor's series expansion of Equation 3.34	3.39

Symbol	Description	Equation No.
$S2_{N'}$	ratio of the N' vs. N'-1 terms in the Taylor's series expansion of Equation 3.36	3.40
S_o	slope of the channel bed	2.2
t	time variable	2.1
t_{p_i}	time to the hydrograph peak	2.38
t_{r_i}	time duration of the hydrograph falling limb	2.38
T_1	time to peak of a triangular hydrograph input	B.1
T_2	time duration of the falling limb of a triangular hydrograph input	B.2
u	velocity of flow in the longitudinal direction of the channel	2.1
$u(t)$	spatially lumped channel inflow at time t	2.31
$\underline{v}(t_k)$	white noise vector sequence in the observations process	5.47
V_i	coefficient of variation	3.35
V_{x_i}	coefficient of variation	3.61
V_{a_i}	coefficient of variation	3.62
$\underline{w}(t)$	white noise vector process driving the system at time t	5.46
$W(t,\tau)$	response at time t due to a unit impulse input at time τ	4.1

Symbol	Description	Equation No.
x	position variable in the longitudinal direction of the channel	2.1
$\underline{x}(t)$	state vector	5.44
$\hat{\underline{x}}(t/t_k)$	minimum variance estimate at time t of the normally distributed random process $\underline{x}(t)$ given observations up to and including time t_k	5.49
y	water depth in channel	2.1
$\underline{y}(t_k)$	vector of observations at time t_k	5.47
$z(t_k)$	discrete time observations process at time t_k	5.112
$z_r(t_k)$	discrete time observations process of the residuals at time t_k	5.83
$\Gamma(\mu)$	Gamma function evaluated at μ	3.11
$\delta(t)$	Dirac delta function (Figure 3.3)	
Δt_i	time difference between the peak discharges of the input and output hydrographs associated with the i^{th} reservoir in a cascade of conceptual reservoirs	2.36
δ_{kj}	Kronecker's delta	5.49
θ	parameter in the discretization of the kinematic wave equations	2.15
λ_{i_k}	the i^{th} eigenvalue of a system of linear equations for times in the interval $[t_k, t_{k+1}]$	5.9
$\mu_{a_i}(t)$	the mean function of the random parameter $a_i(t)$	5.93

Symbol	Description	Equation No.
$\mu_u(t)$	the mean function of the random input	5.79
$\mu_{x_i}(t)$	the mean function of the random state $x_i(t)$	5.94
$\nu(t_k)$	innovations sequence of the minimum variance estimator	5.125
Π	product operator	3.34
$\rho_{x_i a_i}$	correlation coefficient between the random variables (processes) x_i and a_i	3.58
$\sum$	summation operation	3.2
$\sigma_{a_i}(t)$	standard deviation of the parameter process	3.54
$\sigma_{x_i}(t)$	standard deviation of the state process	3.53
$\sigma^2_{x_i a_i}(t)$	cross-covariance between the state and parameter processes	3.53
$\phi_{i,i_k}(t,t_o)$	transition function corresponding to the i^{th} variable in a system of linear equations for the time interval $[t_k, t_{k+1}]$	5.33
$\underline{\Phi}_r(t_{k+1},t_k)$	one-step transition matrix of a system of linear equations in the vector variable $\underline{r}(t)$	5.106

Chapter 1

INTRODUCTION

1.1 Scope of Study

Optimal design and operation of flood control reservoirs require the implementation of efficient flood routing schemes. For instance, flood forecasting systems together with stage-discharge relationships are used to predict the depth of flooding, the main factor governing direct damage to structures in the flood plain. Subsequently, depth-damage cost relationships are utilized for the design of flood control works, which will reduce the damage cost associated with a given flood (Bhavnagri and Bugliarello, 1965).

The flood routing model is also the basic tool in the determination of the damage cost-frequency curves associated with an existing reservoir. The historic floods with known seasonal or annual probabilities of occurrence are routed through the reservoir for various levels of initial storage and release schedules. The computed releases and overflows approximate (since the optimal release schedule is not known a priori) downstream flows. The computed flows are, then, used in damage cost-discharge (or depth) relationships for the construction of damage cost-frequency curves (Hughes, 1971).

In many cases, determination of the expected value of the flood stage and damage costs is not adequate, and a measure of the error involved in this estimate is needed. Coefficient of variations of the flood damage cost greater than one have been reported (Langbein, 1958).

This requires the use of flood routing schemes that provide an estimate of the discharge (mean value) as well as a measure of the uncertainty in this estimate. Literature over the past years have used Kalman filtering (Kalman, 1960) to give an efficient solution to the problem. Unfortunately, the algorithm for the optimal estimator requires linear system dynamics and linear observation equation. Thus, if the system in nonlinear, then some sort of linearization is required.

The purpose of this work is to design a stochastic flood routing model, compatible with large conceptual soil moisture accounting schemes, to be used in the real time forecasting of river flows. The model should be capable of utilizing, in real time, the information available in the observations of the river discharges at a single location. A real world application will serve to verify the design.

1.2 Program of Study

A review of existing models is presented and a selection made based on the goals of this work. Since the chosen model is nonlinear, the results of optimal estimation theory are not readily applicable, and linearization is required. Statistical linearization is used due to its superior properties over ordinary linearization using a truncated Taylor series expansion. This technique has the disadvantage that the resultant linearization gains are functions of expected values of the nonlinear functions involved. The problem is almost prohibitive in the case of multi-dimensional functions. A methodology is developed for the approximate, analytical, determination of the expected value of vector, multi-dimensional, nonlinear functions. The requirements are that the function is differentiable to some order and that its arguments have probability distributions that can be characterized by a finite number of parameters.

A Gaussian minimum variance estimator of parameters and states is designed for the statistically linearized model. The issue of stability for the state estimator is examined. Tests of the stochastic model in the 2344 km^2 drainage basin of Bird Creek, Oklahoma, U.S.A., are performed.

In Chapter 2, a review of existing flood routing models is offered. The models are grouped in dynamic, conceptual and regression type. A conceptual model based on a series of nonlinear (in general) reservoirs is chosen and its response characteristics are examined.

A methodology to determine analytically the expected value of a nonlinear function is presented in Chapter 3. Results of comparisons with numerical integration schemes are also discussed for the case of a two-dimensional function of the type $a \cdot x^m$. The arguments a, x of the function are assumed to be normally distributed.

The statistical linearization technique is presented in Chapter 4. The theory for stationary processes is outlined first and subsequently, results for the nonstationary case are used to linearize the nonlinear flood routing model. A Gaussian distribution is assumed for the states and parameters of the model. The gains of linearization are computed using the results of Chapter 3.

The minimum variance estimator based on the assumption of Gaussian distribution for the initial states and the driving noise terms is discussed in Chapter 5. The canonical form of the system equations is developed. Use of the canonical form is made to examine the state estimator stability properties. The formulation for simultaneous state and parameter estimation is also given.

The stochastic flood routing model designed is used in Chapter 6 to predict six-hour discharge values at the Bird Creek drainage basin in Oklahoma, U.S.A. Implementation considerations as well as evaluation of the filter parameters are discussed.

Chapter 2

FLOOD ROUTING

2.1 Introduction

Flood routing is defined by Viessman, et al. (1972), as the procedure used to predict the temporal and spatial variations of a flood wave, as it traverses a river reach or reservoir.

During the past years, several descriptions of the phenomenon, in terms of mathematical models, have been formulated. Flood routing models differ in: the representation of the process physics, the ways of estimating model parameters, and the input data requirements.

In the sections to follow, a review of the commonly used flood routing models is offered. The criteria for selection among them are discussed in terms of the scope of this study. Subsequently, the important features of the chosen model (based on the predetermined criteria) are examined. Characteristic applications of modern estimation theory to flood routing problems are also presented.

2.2 Flood Routing Models

It is customary, in order to facilitate the discussion on the subject, to group the different routing models under categories. In the past, division of the models has been based on the number of their parameters (Dooge, 1973) or on the aspects of the physical phenomenon (flood wave movement) that they best simulate (Weinmann and Laurenson, 1977). In recent years, advances in estimation theory permitted the use of simplified models in river flow forecasting with satisfactory results.

This fact calls for a somewhat more general categorization of models.

For the purposes of this development, the modern flood routing models will be divided in a) dynamic (or physically oriented), b) conceptual (or phenomenological) and c) regression type (or black box) models.

All models that are based on the description of the physical phenomenon using the conservation of mass and momentum equations (usually referred to as the St. Venant equations) are grouped under the heading "dynamic models." Representatives of this group of models are those utilizing the full equations for gradually varied, unsteady channel flow, those that are based on equations derived from a linearization of the complete nonlinear St. Venant equations, and the kinematic wave model.

Conceptual hydrologic models are the reservoir storage type and the so-called kinematic models. Their characteristic feature is that they retain some of the physical laws (e.g., conservation of mass) in their mathematical formulation, without being exact representations of the physical reality.

The last category includes the regression type of models. They rely heavily on an input-output description of the phenomenon, without simulating any of the physical processes involved. In these models, the flood wave-channel network system is replaced by some sort of impulse response function.

Admittedly, there is no exact representation of the physical reality given the assumptions inherent in the best available description of the flood wave propagation in channel networks; the main assumptions being one-dimensional flow and resistance laws similar to those of steady uniform flow in prismatic channels. Furthermore, exact analytical solution to the partial differential equations of motion and continuity for a given set of initial and boundary conditions is not available. Consequently, a major abstraction is introduced in most cases due to the discretization schemes used for the numerical integration of these equations. Therefore, one is tempted to include the dynamic models under the second category. However, the distinction is made since it is believed that physical models are superior to the other ones in that they are capable of good performance in ungaged drainage basins, utilizing hydro-geologic data.

Concerning the so-called conceptual models, the common procedure to evaluate their parameters is to use input-output data and as a consequence little or no physical meaning can be attributed to the obtained values. In this way, they resemble very much the models of the third category. Nevertheless, conceptual models are expected to have a better performance than the "black box" models, in forecasting under conditions dissimilar to those of their calibration period (Kitanidis and Bras, 1978).

In the following, a rather qualitative discussion of the characteristic models of each category is presented.

2.2.1 Dynamic Models

The laws of conservation of mass and momentum for gradually varied, unsteady flow can be written as (Eagleson, 1970),

$$\frac{\partial y}{\partial t} + u \cdot \frac{\partial y}{\partial x} + y \cdot \frac{\partial u}{\partial x} = i - f + \frac{2\, q_L}{b} \tag{2.1}$$

$$\frac{\partial u}{\partial t} + u \cdot \frac{\partial u}{\partial x} + g \cdot \frac{\partial y}{\partial x} = [i - f + \frac{2\, q_L}{b}] \cdot \frac{u}{y} - (1 + \frac{2y}{b}) \cdot \frac{\tau_o}{\rho \cdot y} + g \cdot S_o \tag{2.2}$$

where y is the flow depth measured perpendicular to the channel bottom; u is the flow velocity, assumed parallel to the channel bed (one-dimensional problem); x is the direction of flow; t is the time; i is the rainfall intensity; f is the infiltration rate; q_L is the lateral inflow to the channel; b is the width of the channel, assumed rectangular; g is the gravitational acceleration; ρ is the fluid mass density; S_o is the slope of the channel bed and τ_o is the boundary shear stress.

Under the assumptions of negligible rates of rainfall (i) and infiltration (f), and small contribution of the lateral inflow in the momentum equation (see corresponding discussion in Eagleson, 1970), one can rewrite the governing equations (2.1) and (2.2) as

$$\frac{\partial y}{\partial t} + u \cdot \frac{\partial y}{\partial x} + y \cdot \frac{\partial u}{\partial x} - \frac{2q_L}{b} = 0 \tag{2.3}$$

$$S_f = S_o - \frac{\partial y}{\partial x} - \frac{1}{g} \cdot \frac{\partial u}{\partial t} - \frac{u}{g} \cdot \frac{\partial u}{\partial x} \tag{2.4}$$

where S_f is the friction slope (Henderson, 1966), being equal to the expression: $(1 + \frac{2y}{b}) \cdot \frac{\tau_o}{\rho \cdot g \cdot y}$.

If the assumption is made that the friction slope can be determined as in steady, uniform flow, then,

$$S_f = u^2 \cdot \frac{1}{C^2 \cdot R} \tag{2.5}$$

where C is a constant coefficient and R is the hydraulic radius.

Substitution of Eq. (2.5) in the momentum Eq. (2.4) yields:

$$u = C \cdot R^{1/2} \cdot \left(S_o - \frac{\partial y}{\partial x} - \frac{1}{g} \cdot \frac{\partial u}{\partial t} - \frac{u}{g} \cdot \frac{\partial u}{\partial x}\right)^{1/2} \tag{2.6}$$

or, in terms of discharge Q (= u•A):

$$Q = A \cdot C \cdot R^{1/2} \cdot \left(S_o - \frac{\partial y}{\partial x} - \frac{1}{g} \cdot \frac{\partial u}{\partial t} - \frac{u}{g} \cdot \frac{\partial u}{\partial x}\right)^{1/2} \tag{2.7}$$

where A is the cross-sectional area of flow. It should be noted that Eqs. (2.6) and (2.7) imply that the velocity and discharge are not single valued functions of the depth. Thus, it is common to refer to them (in particular, Eq. (2.7)) as looped rating curves.

Equations (2.3) and (2.6) have been used for the simulation of the flood wave propagation in rivers. Major assumptions inherent in the derivation of these equations are a) one-dimensional flow, b) small bottom slope S_o, c) hydrostatic pressure distribution along the depth of flow, d) the resistance laws for steady, uniform, turbulent flow are considered applicable. The last two terms in the momentum equation (2.4) are the local and convective acceleration terms, respectively, while $\frac{\partial y}{\partial x}$ represents the pressure gradient.

Equations (2.3) and (2.6) are coupled, nonlinear, first order partial differential equations of the hyperbolic type. They require one initial and two boundary conditions for their solution. For subcritical flow, the boundary conditions are specified at both ends of the river reach under consideration, while for supercritical flow (not usually found in natural rivers), both boundary conditions are specified upstream (since disturbances propagate only downstream). Discharge or depth hydrographs and stage vs. discharge (rating) curves are utilized as boundary conditions.

The objective in solving these equations is to determine the water depths and velocities at all points in the reach, for all times. Unfortunately, there is no known analytical solution. Numerical schemes are required to discretize the space and time axes and convert the differential equations to difference ones. Different numerical schemes have been suggested.

Barnes (1967) made a comparison of three numerical schemes in terms of their ability to give stable solutions and to adapt to sub- and supercritical flow. The results of the numerical experiments were compared with the results obtained from an experimental facility that was operated under known initial and boundary conditions. The method of characteristics gave superior results as compared with a scheme utilizing a triangular mesh of solution points in the position-time space and with another utilizing a rectangular grid and second order differences. The initial conditions were taken as those of a steady uniform flow, while a discharge hydrograph and a rating curve were the boundary conditions.

An implicit characteristic method using a characteristic network, an explicit direct method and an implicit direct method were compared by Amein and Fang (1969). Their tests were made for a) flow through an idealized channel and b) flow through a natural channel. These investigators suggested the use of an implicit scheme that would take advantage of the sparse nature of the simultaneous algebraic equations that need to be solved at each iteration step.

Due to the computational burden associated with the numerical solution of the "complete" equations (2.3) and (2.4), as well as due to the quantity and quality of the input data required for their solution, several approximations have been proposed, within the class of dynamic models.

Dooge and Harley (1967) presented a linearized version of Eqs. (2.3) and (2.6) about a reference discharge Q_o, as follows:

$$(g \cdot y_o - u_o^2) \cdot \frac{\partial^2 Q}{\partial x^2} - 2 \cdot u_o \cdot \frac{\partial^2 Q}{\partial x\, \partial t} - \frac{\partial^2 Q}{\partial t^2} = 3 \cdot g \cdot S_o \cdot \frac{\partial Q}{\partial x} + \frac{2 \cdot g \cdot S_o}{u_o} \cdot \frac{\partial Q}{\partial t} \tag{2.8}$$

where Q represents discharge; y_o and u_o are the reference depth and velocity, respectively, corresponding to the reference discharge. Note that the lateral inflow has been assumed insignificant, and it does not enter Eq. (2.8). Since the perturbation analysis was based on small deviations from the reference trajectories, these investigators performed sensitivity analysis with respect to the values of Q_o. They reported that a) an increase in Q_o results in a decrease in the lag-time for the channel reach under study, b) the second moment of the outflow hydrograph was not sensitive to changes in the value of Q_o, c) for floods of long duration, changes in Q_o have a small effect on the shape of the hydrograph. Based on these observations, they subsequently suggested that the operation of the linearized channel on the inflow appears to consist of a translation dependent on the reference discharge Q_o, and an attenuation effect practically independent of the reference discharge. That is, the system has been replaced by two subsystems, one causing a nonlinear translation of the inflow and the other causing a linear attenuation to the translated inflow. This linearized version of the complete equations was mainly used to evaluate simpler linear models (Dooge, 1973).

Simplification of Eq. (2.8) leads to the so-called diffusion analogy models whose describing equation is of the type:

$$\frac{\partial Q}{\partial t} + c \cdot \frac{\partial Q}{\partial x} = D \cdot \frac{\partial^2 Q}{\partial x^2} \tag{2.9}$$

where c is a coefficient related to the translation of the flood wave; D is a diffusion coefficient that introduces attenuation of the flood wave. If lateral inflow is significant, it is added in the right hand term in the form of $c \cdot q_L$. Price (1973) proposed analytical expressions for the determination of the coefficients D and c in natural channels.

In the diffusion analogy models, the acceleration terms of the momentum equation (2.6) have been neglected, while the pressure gradient term $\frac{\partial y}{\partial x}$ has been retained. This suggests that this type of model is well-suited for very mild slopes of the channel bed where pressure terms become significant. The diffusion analogy models preserve the double valued relationship between discharge and stage.

Perhaps the dynamic model used most often is the kinematic wave model (Lighthill and Whitham, 1955). In this case, the momentum equation is reduced to the following:

$$S_f = S_o \tag{2.10}$$

Consequently, there is a one to one correspondence between stage and discharge. In this case, the describing equations can be determined to be (Eagleson, 1970):

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = q_L \tag{2.11}$$

$$Q = a \cdot A^m \tag{2.12}$$

where a and m are constant coefficients, or in a combined form (Lighthill and Whitham, 1955):

$$\frac{1}{c} \cdot \frac{\partial Q}{\partial t} + \frac{\partial Q}{\partial x} = q_L \tag{2.13}$$

where c is the celerity of the kinematic wave and can be determined as:

$$c = \left.\frac{dQ}{dA}\right|_{x=x_c} \tag{2.14}$$

where x_c is any value of the position coordinate.

Solution of the system of Equations (2.11) and (2.12) is usually based on the method of characteristics, since this simplified model possesses only one system of forward characteristics, in contrast to the hyperbolic equations (2.3) and (2.4) that possess two systems of characteristics.

The kinematic wave model is a good representation of the physical process of the flood wave movement when inflow, free surface slope and inertia terms are negligible in comparison with those of bottom slope and friction (Eagleson, 1970).

Henderson (1966) suggests that the kinematic wave behavior is very close to the one observed for natural floods, in steep rivers whose slopes are of the order of 10 feet per mile or more.

A kinematic wave does not subside or disperse, but it will change shape (it steepens) due to the dependence of the velocity on the depth. If the steepening process ceases, the result will be a steady state formation called the kinematic shock (or monoclinal wave) (Henderson, 1966). This kind of behavior is confirmed by observed natural floods, but the waves do not steepen as much and do exhibit attenuation (Weinmann, 1977).

The term kinematic is used to indicate that the generated waves initiate in the continuity equation. Dynamic waves result from the use of the complete equations (2.3) and (2.6).

Both kinematic and dynamic waves are present in natural floods. The bulk of the flow behaves like a kinematic wave, while dynamic waves precede and follow it. It can be shown (Eagleson, 1970) that for natural floods ($F < 2$), the dynamic wave preceding the kinematic one will attenuate rapidly, making little contribution to the progress of the flood wave.

Due to the ability of the kinematic wave model to represent the flood wave movement in natural floods (except perhaps at the junction of the main river with tributaries carrying significant flow), and its analytical tractability, especially for overland flow type of problems, it has been widely used.

Eagleson (1967) used this model to derive the contribution to peak catchment discharge from an impulse of rainfall excess at a certain location in the catchment. Bras and Perkins (1975) utilized the kinematic wave equations to simulate flow in a conceptual basin represented as a network of overland, stream, pipe and gutter segments, in order to determine the effects of urbanization in the runoff hydrographs of selected areas.

Bras and Rodríguez-Iturbe (1975) used a finite difference representation of the kinematic wave equations in network design problems.

Recently, Chan and Bras (1978) derived the distribution of the water volume above a given threshold discharge for overland flow based on the kinematic wave theory.

2.2.2 Conceptual Models

Due to existing uncertainty in the input data, system parameters, and initial-boundary conditions, conceptual models for flood routing have been developed. Almost all of them use the conservation of mass law and to a varying extent try to approximate the conservation of momentum law.

One group of these models, called kinematic (Weinmann and Laurenson, 1977), are generalizations of the kinematic wave model, in that they account for attenuation effects through the numerical methods used for the discretization of the kinematic wave equations.

The most commonly used representative of these models is the Muskingum River model (Viessman, et al., 1972). If the kinematic wave equation (2.13) is discretized, based on the scheme of solution points of Figure 2.1, then

$$\frac{1}{\bar{c}} \cdot \frac{\theta \cdot (I_2 - I_1) + (1-\theta)(Q_2 - Q_1)}{\Delta t} + \frac{Q_1 + Q_2 - I_1 - I_2}{2\Delta x} = 0 \quad (2.15)$$

I_1, I_2: input discharges

where the lateral inflow has been omitted for simplicity and $\bar{c}$ represents an average value of c in the reach. Setting $\frac{\Delta x}{c}$ equal to a constant K and solving for Q_2 results in the commonly used form of the Muskingum model. Cunge (1969) showed that if $\bar{c}$ evaluated as a representative value of the kinematic wave celerity in the reach with length Δx, and θ is set equal to $\frac{1}{2} \cdot (1 - \frac{Q}{b \cdot S_o \cdot \bar{c} \cdot \Delta x})$, then the Muskingum scheme is a second order approximation of the Equation (2.9), which corresponds to diffusion analogy models.

A second order approximation to Equation (2.9) was obtained from the kinematic wave equation (2.13) by discretizing in space only and solving the resultant differential equation for the unknown discharge. The solution can be written in the form:

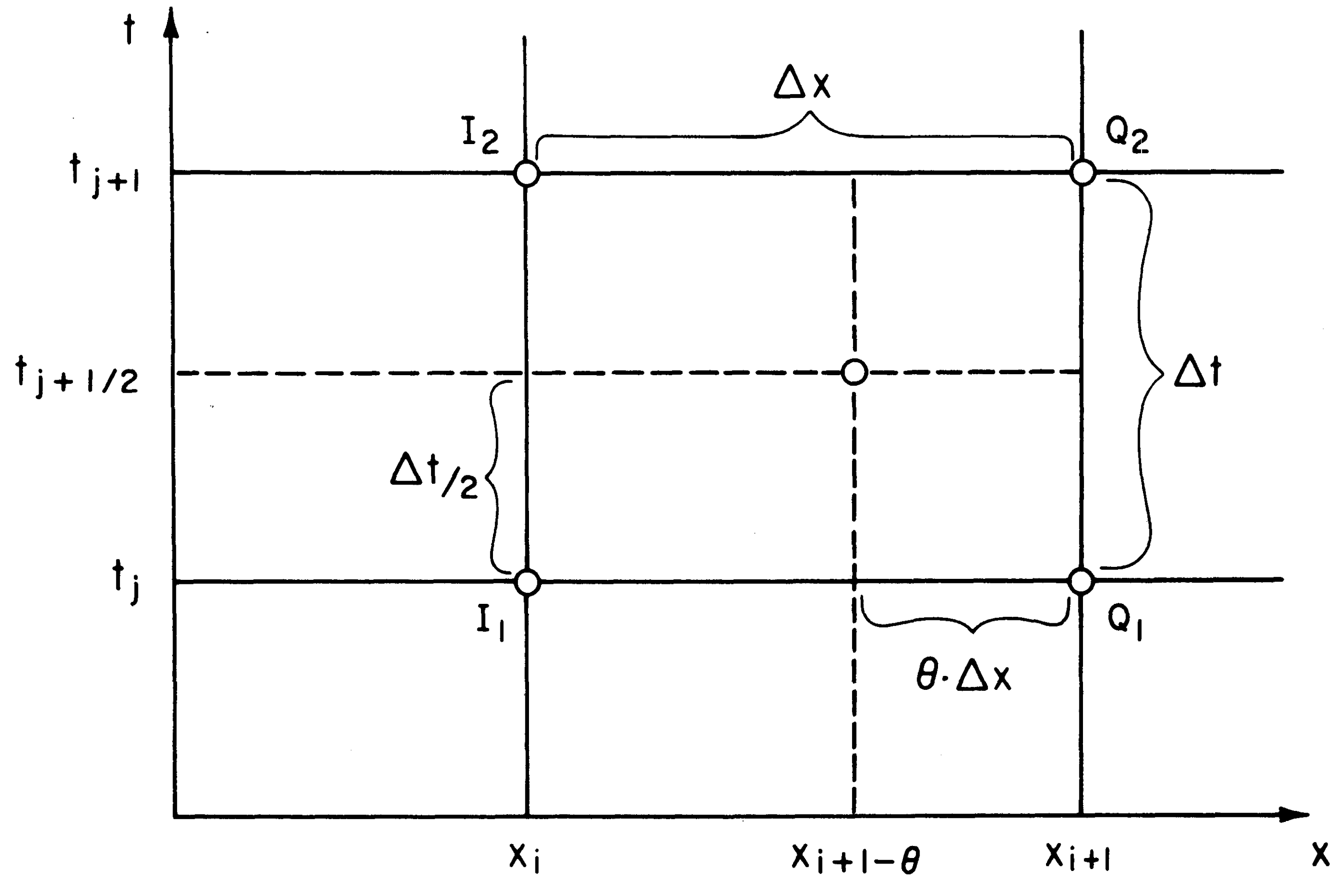

Figure 2.1

FINITE DIFFERENCE GRID FOR KINEMATIC MODELS
(Adopted from Weinmann and Laurenson, 1977)

$$Q_2 = k_1 \cdot I_1 + k_2 \cdot I_2 + k_3 \cdot Q_1 \qquad (2.16)$$

where

$$k_1 = \frac{\Delta x}{\bar{c} \cdot \Delta t} \cdot (1 - k_3) - k_3 \qquad (2.17)$$

$$k_2 = 1 - \frac{x}{\bar{c} \cdot \Delta t} (1 - k_3) \qquad (2.18)$$

$$k_3 = e^{-\frac{\bar{c} \cdot \Delta t}{\Delta x \cdot (1-\theta)}} \qquad (2.19)$$

It can be shown (Weinmann and Laurenson, 1977) that the model described by Eqs. (2.16) through (2.19) can be reduced to the model proposed by Nash (1959), to the successive routing method of Kalinin and Milyukov (Dooge, 1973) and to the kinematic wave model (e.g., set $\theta = 0.5$ and $\bar{c}$ is evaluated from a steady flow rating curve). This model is attributed to A. Koussis. The same investigator generalized this model (Koussis, 1976) by introducing the effects of pressure for very mild bed slopes, through the "Jones formula" (Henderson, 1966), based on the kinematic wave assumption of no subsidence.

In general, it can be stated that the group of kinematic models presented above are introducing attenuation effects for values of the parameter θ in the interval $0 \leq \theta < 0.5$.

For $\theta = 0.5$, a pure translation results (kinematic wave), while for $\theta = 0$, a pure attenuation is in effect.

The parameter $\bar{c}$ is a measure of the translation effects.

A second class of models in the conceptual category consists of those of the reservoir type. These models introduce attenuation due to the storage properties of the channel network or reservoirs involved. Different configurations for the reservoirs have been proposed (Nash, 1959; Overton, 1967; Mein, et al., 1974). Perhaps that of Mein, Laurenson and McMahon (1974) is the most general one. In this case, the continuity equation has been retained and a nonlinear relationship between storage and discharge is assumed. The channel network is divided in subreaches and the model is used with each one of them. The different reservoirs are connected through the use of the conservation of mass law at each junction of subreaches. Detailed discussion about this type of model is reserved for later sections.

The parameters of the reservoir type of models are estimated from hydro-geologic data (Mein, et al., 1974) or from observed input-output data (which is the most common approach when these data are available).

Weinmann (1977) compared complete dynamic models, diffusion analogy models, kinematic wave models and kinematic models in regular and highly irregular channels. The criteria for comparison were:

a. the ability of the models to closely approximate the physical phenomenon,
b. the ability to produce stable and accurate results for a wide range of conditions when used with numerical schemes,
c. the flexibility of the associated parameter estimation procedures to adjust to different forms of input information and to make use of all available data,

d. the flexibility of the model structure to permit a matching of the model complexity and outputs (considering accuracy and format), to the quality of the available input data and to the purpose of the routing results.

He concluded that none of the models included in the study had all the desirable features of an "ideal" model. However, he suggested that the model proposed by Koussis (1976) shows great potential of being developed into a practical model with general applicability.

2.2.3 Black Box Models

The models of this category part completely from any physical analogy to the flood routing phenomenon. They describe the discharge (or depth) at a given position and time as a (usually linear) function of the discharges at (n) previous time steps in all (m) positions as well as a function of the inflows at (ℓ) previous time steps in (p) inflow locations. The major problem in those models is to estimate the parameters involved. For example, if the model is linear, then estimation of the regression coefficients and the integers n, m, ℓ is sought. This is done exclusively using inflow-outflow observed data and linear regression techniques (Dooge, 1973).

Concern regarding the predictive capability of black box type models for periods dissimilar to the calibration period suggests their use in short-term prediction. Preferably, "black box" models should be used with observation processors to incorporate additional streamflow and input measurements in the estimated value of their parameters (e.g., as in a Kalman filter).

2.3 Use of Modern Estimation Theory in Flood Routing Studies

Kalman and Bucy (1961) solved the linear filtering problem that consists of obtaining the linear, minimum variance, recursive estimator of the state variables of a system (described by a set of first order differential equations) at time t, given observations on the output of the system up to and including time t. The major assumption of the developed theory are a) linear system dynamics equation and linear equation relating the observed quantities with the system states, and b) the second moment statistics of the additive disturbance processes must be known (see also Chapter 5).

Several extensions to the basic theory were proposed to relax the assumptions involved (Jazwinski, 1970). Mainly, for the purpose of parameter estimation, the so-called extended Kalman filter was developed. This is based on a linearization, through a Taylor series expansion, of the nonlinear multi-dimensional functions involved and an augmentation of the state vector of the system by the vector of parameters. Then, the basic theory was applied to the resultant higher order, linear system.

Based on the fact that the so-called Kalman filter is ideally suited for use in the hydrologic forecasting problem (equivalent to prediction in the sense of Kalman, et al., 1961), several investigators have applied this technique within the hydrologic context (e.g., Hino, 1973; Todini and Bouillot, 1975; Bras and Rodríguez-Iturbe, 1975;

Kitanidis and Bras, 1978).

In this section, only some of the characteristic applications of this estimator to flood routing studies will be presented.

Hino (1973) used the Kalman filter to recursively estimate the parameters of unit hydrographs (models of the regression category).

Bras and Rodríguez-Iturbe (1975) used a state-space formulation of the kinematic wave model, coupled with a multivariate non-stationary rainfall generator to assess the accuracy inherent in the estimation of the rainfall input by a raingage network.

Logan, Lennox and Unny (1978) utilized a nonlinear storage type model with time varying parameters together with a Kalman filter for state and parameter estimation, with satisfactory results.

Wood (1978) used a simplified model of channel routing (black box type) to forecast river flows at different points in a river network with a short lead time of forecast (2 hours).

Li, Duong and Simons (1978) used a state space formulation of the stage-discharge relationship with an iterated extended Kalman filter (Jazwinski, 1970) to study the parameters that affect the form of the looped rating curve at a location along the river.

Kitanidis and Bras (1978) used a simple linear reservoir as the routing model of the National Weather Service River Forecast System (NWS HYDRO-31), together with an extended Kalman filter for state and parameter (coefficient of the linear reservoir model) estimation. Their formulation included adaptive estimation of model errors and input error identification techniques. A comparison of a black box type of model with the conceptual storage type hydrologic model favored the second type, especially for long forecast lead times.

All the above presented studies convert the describing equations of the flood wave movement into a system of first order, ordinary, differential equations and then apply the results of modern estimation theory. Since a basic prerequisite for the use of optimal estimation is the linearity in the model and measurement equations, it is a standard procedure to linearize the system equations by keeping the leading terms in a Taylor series expansion of the nonlinear functions about some reference point. Kitanidis and Bras (1978) were the first to use the technique of statistical linearization in hydrologic studies. Although superior to ordinary linearization (Gelb, 1974), this technique has not been adequately explored.

Another characteristic of the studies presented is that there has been little work in using conceptual models for real time flood routing, even though whenever used, they produced satisfactory results (Kitanidis and Bras, 1978).

To the best of the authors' knowledge, there has not been an attempt to work directly with the partial differential equations of motion (Equations (2.3) and (2.6)), even though estimation techniques for distributed systems are available (Seinfeld, et al., 1971). Clearly, an approach like this would be ideal in cases where accuracy is the only objective and at the same time adequate hydrologic data are available for good observability properties of the resultant filters.

2.4 Model Selection

In mathematical terms, the problem of model selection can be formulated as a mathematical programming problem, whose objective function is a weighted combination of accuracy, computational economy and availability of input data functions. One can go even further and convert the problem to a stochastic programming one by considering the uncertainty inherent in the values of the weights of the objective function. The solution to the generalized model selection problem can then be obtained through successive evaluations of the objective function with respect to all the available alternatives.

Even if the weights of the objective function were given, time constraints do not permit the above described venture. In addition, such a procedure has not been used in cases with objectives similar to those of this study. Therefore, the selection of the flood routing model to be used in this work will be based on qualitative arguments.

It is the purpose of this work to formulate a stochastic flood routing model to be used in real time streamflow forecasting, together with large conceptual hydrologic models. The available data consist of an inflow discharge hydrograph, which is usually the output of a soil moisture accounting scheme, and an observed outflow discharge hydrograph. The flood routing model is intended to be used with real time estimation schemes that will compensate for errors in the input hydrographs, model formulation and estimated parameter values. On-line parameter estimation schemes will be designed to tune the model for the particular drainage basin under study.

It is consistent with the above considerations to prefer a conceptual model over a dynamic one. On the other hand, when used in real time with recursive estimators, conceptual models perform better than regression type models for relatively large forecast lead times (Kitanidis and Bras, 1978). Consequently, a conceptual flood routing model seems well suited for the purposes of this work.

A general model based on a series of reservoirs similar to the one proposed by Mein, et al. (1974) is adopted.

2.5 Flood Routing Using a Series of Nonlinear Reservoirs

Based on the relative homogeneity of the hydro-geomorphologic properties of a catchment (e.g., roughness of the channel bottom, slope of the channel, distribution of channel inflow), one can divide the total length of the main channel in n reaches. Each one of the reaches can be thought of as a reservoir that stores water and releases it according to some law.

Denote by $S_i(t)$ the volume of water in storage at the i^{th} reach, $I_i(t)$ the lumped inflow rate in the i^{th} reach, and by $Q_i(t)$ the corresponding outflowing discharge.

The conservation of mass law applied to the i^{th} conceptual reservoir gives,

$$\frac{dS_i(t)}{dt} = I_i(t) - Q_i(t) \quad ; \qquad i = 1, 2, \ldots, n \qquad (2.20)$$

for all times t.

To complete the model, it is necessary to define the release law of the reservoirs, which would provide a relationship between $S_i(t)$ and $Q_i(t)$ for all i and t.

There is analytical and field evidence (Mein, et al., 1974) that supports a power law between the outflowing discharge at time t and the volume of water in storage at the same time t, for each reservoir. For example, consider a wide rectangular channel and let $y_i(t)$ represent the depth at the i^{th} reach at time t, on the average. If L_i is the length of the reach and b_i is the average width of the channel, then, the total volume in storage at time t for the i^{th} reach is,

$$S_i(t) = b_i \cdot y_i(t) \cdot L_i \qquad (2.21)$$

Use of Chezy's formula for the resistance law in uniform flow and of a kinematic relationship that equates the friction slope with the slope of the bottom of the channel results in,

$$Q_i(t) = C_i \cdot A_i(t) \cdot R_i^k(t) \cdot S_{o_i}^{1/2} \qquad (2.22)$$

where C_i is a constant; $A_i(t)$ is the cross-sectional area of flow at time t; $R_i(t)$ is the hydraulic radius at time t; k is a parameter and S_{o_i} is the bedslope for the i^{th} reach.

Using

$$R_i(t) \approx y_i(t) \qquad (2.23)$$

in Eq. (2.22) and substituting,

$$A_i(t) = b_i \cdot y_i(t) \tag{2.24}$$

yields,

$$Q_i(t) = C_i \cdot b_i \cdot y_i(t) \cdot y_i^k(t) \cdot S_{o_i}^{1/2} \tag{2.25}$$

Substitution of $y_i(t)$ from Eq. (2.21) gives,

$$Q_i(t) = C_i \cdot \frac{S_i(t)}{L_i} \cdot \left(\frac{S_i(t)}{b_i \cdot L_i}\right)^k \cdot S_{o_i}^{1/2} \tag{2.26}$$

or

$$Q_i(t) = a_i \cdot S_i^m(t) \tag{2.27}$$

where in this case,

$$a_i = \frac{C_i \cdot S_{o_i}^{1/2}}{b_i^k \cdot L_i^{k+1}} \tag{2.28}$$

and

$$m = k + 1 \tag{2.29}$$

If a $Q_i(t)$ vs. $S_i(t)$ relationship of the type in Eq. (2.27) is accepted, then Eq. (2.20) can be rewritten as,

$$\frac{dS_i(t)}{dt} = I_i(t) - a_i \cdot S_i^m(t) \tag{2.30}$$

The input distribution can be taken as,

$$I_1(t) = p_1 \cdot u(t) \tag{2.31}$$

$$I_i(t) = p_i \cdot u(t) + a_{i-1} \cdot S_{i-1}^m(t) \ ; \quad i = 2, \ldots, m \tag{2.32}$$

where u(t) is the total channel inflow and p_i, i = 1, 2, ..., n, is the proportion of u(t) that serves as an input to the i^{th} reservoir. Under these conditions, the equations of the system of reservoirs can be expressed as

$$\frac{dS_i(t)}{dt} = p_i \cdot u(t) + a_{i-1} \cdot S_{i-1}^m(t) - a_i \cdot S_i^m(t) \quad ; \quad i = 1, 2, \ldots, n \tag{2.33}$$

with

$$a_o \triangleq 0 \tag{2.34}$$

A schematic representation of the system is given in Figure 2.2.

The final output of the system is the discharge $Q_n(t)$ given by

$$Q_n(t) = a_n \cdot S_n^m(t) \tag{2.35}$$

The parameters a_i, i = 1, 2, ..., n, and m of the model can be determined either from survey data or/and from input-output discharge records. If detailed survey data are available, for example, average values of b_i, L_i, S_{o_i}, C_i, k for all reaches, and the assumptions used to arrive in Eq. (2.30) are a good approximation to reality, then initial estimates of the parameters can be obtained from Eqs. (2.28) and (2.29). If inflow and outflow discharge data are available, then an estimation scheme should be used to produce the best fit to the records, according to some criterion, utilizing the values obtained

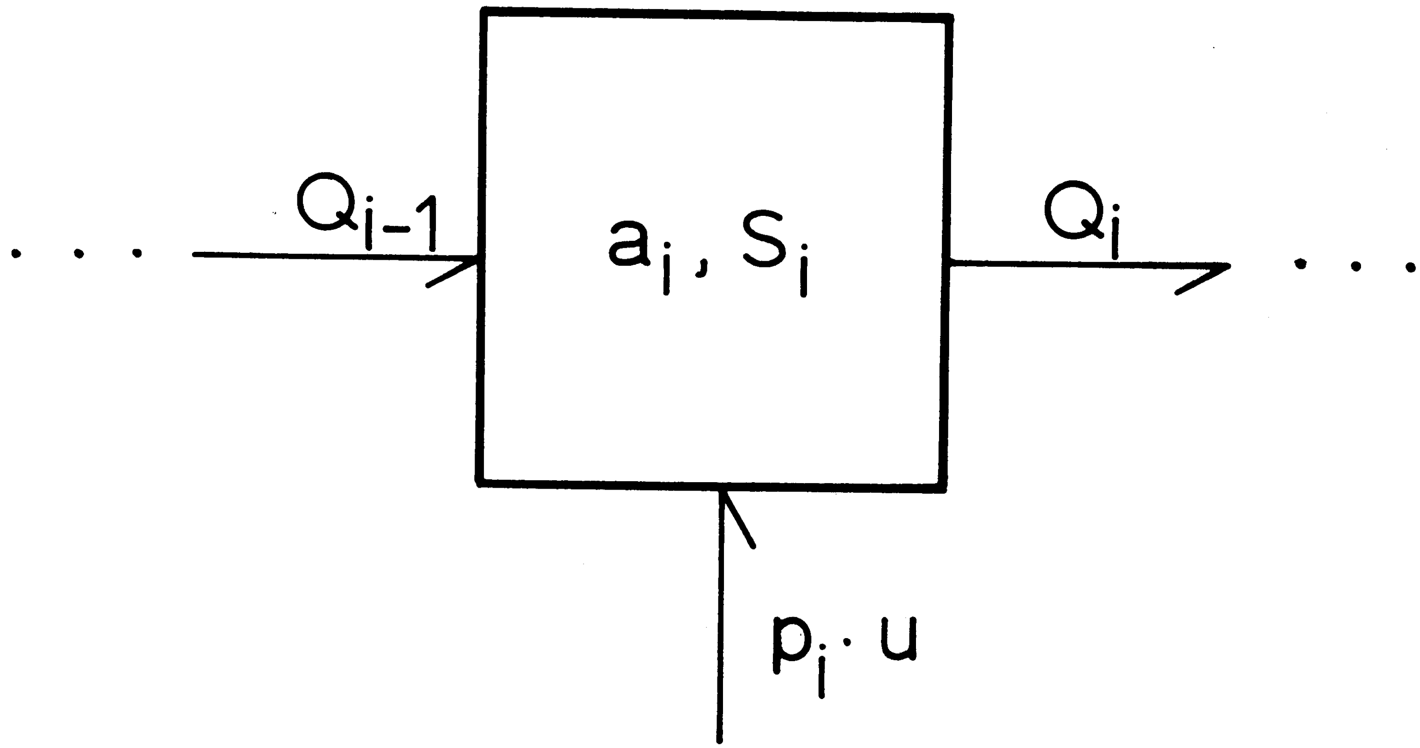

Figure 2.2

SCHEMATIC REPRESENTATION OF THE i^{th} NONLINEAR RESERVOIR IN A CASCADE OF RESERVOIRS from Eqs. (2.28) and (2.29) as initial values.

Parameter estimation based on input-output information is recommended since b_i, L_i, S_{o_i}, C_i, k are distributed quantities in space and possibly varying in time, rather than deterministic quantities clearly identifiable for each lumped reach.

In the absence of survey data, the estimation of parameters must be solely based on the input and output discharge records. There is a variety of parameter estimation schemes. In particular, advances in filtering theory permit the on-line evaluation of the parameters. That is, each observation is processed as it becomes available. Chapter 5 will deal with the problem of the on-line parameter estimation. It is a well-known fact, though, that the main problem of on-line parameter estimation schemes is to obtain convergence of the estimates. In other words, the estimation schemes produce stable parameter values if the initial estimates used were good. The most common way to obtain good initial estimates is to conduct simulation studies for a certain period when historical input-output data are available. Parameters are selected based on a criterion of performance (usually, a time averaged squared error criterion). Selected values are then used as initial estimates for the on-line parameter estimation schemes. In the following, certain properties of the nonlinear reservoir model are presented that will facilitate the determination of good initial parameter values.

It is obvious from Eq. (2.33) that if the system is not operated in real time, one can consider each reservoir individually (with input the sum of the hydrograph produced by the previous reservoir and the channel input). Consequently, reasonable parameters may result from studies of a single reservoir response to a given input. It will be convenient to characterize the output hydrograph from a single reservoir with the same parameters as the input hydrograph. The objective is to relate output and input characteristics as a function of model parameters. The inverse relationship could then be used to identify a model that will reproduce (up to some limited characteristics) an observed output hydrograph. It should be noted that the invariant in the transformation of the input, when routed through a reservoir, is its total volume of water, provided that the initial volume in storage is zero. The identification problem could be a trial and error procedure to determine the values of the parameters of the model (a_i for all reservoirs and the number of reservoirs) from given values of the parameters of the observed input and the output hydrographs.

In this work, the input hydrograph is simplified to a triangular shaped one. The characterizing parameters in this case are the peak discharge Q_{p_i}, the time to peak t_{p_i} and the duration of the falling limb t_{r_i}. The subscript i denotes that the parameters correspond to the output of the i^{th} reservoir (the input to the next one). One can convert any hydrograph to triangular form by preserving Q_{p_i} and the volumes of water under the rising and the falling limb.

When the reservoir is assumed linear, the analysis is simple since one can determine the output parameters analytically. This is pursued in Appendix B. The results can be presented in a recursive form as follows:

$$\frac{Q_{p_i}}{Q_{p_{i-1}}} = 1 - \left(\frac{t_{p_{i-1}}}{t_{r_{i-1}}}\right) \cdot \left(\frac{\Delta t_i}{t_{p_{i-1}}}\right) \tag{2.36}$$

$$\frac{Q_{p_i}}{Q_{p_{i-1}}} = g_1\left(a_i \cdot t_{p_{i-1}}, \frac{t_{p_{i-1}}}{t_{r_{i-1}}}\right) \tag{2.37}$$

$$\frac{t_{p_i}}{t_{r_i}} = g_2\left(a_i \cdot t_{p_{i-1}}, \frac{t_{p_{i-1}}}{t_{r_{i-1}}}\right) \tag{2.38}$$

$$a_i \cdot t_{p_i} = g_3\left(a_i \cdot t_{p_{i-1}}, \frac{t_{p_{i-1}}}{t_{r_{i-1}}}\right) \tag{2.39}$$

where Δt_i is the translation in time of the peak discharge due to the routing through the reservoir and g_1, g_2, g_3 are nonlinear functions. (Note: Q_{p_o}, t_{p_o} and t_{r_o} would be the characteristic parameters of the input to the cascade of reservoirs.) Expressions for the nonlinear functions g_1, g_2, g_3 are given in Eqs. (B.49), (B.51) and (B.52) of Appendix B.

Given values for $\frac{t_{p_{i-1}}}{t_{r_{i-1}}}$, and $a_i \cdot t_{p_{i-1}}$, one can determine $\frac{Q_{p_i}}{Q_{p_{i-1}}}$, $\frac{\Delta t_i}{t_{p_{i-1}}}$, $\frac{t_{p_i}}{t_{r_i}}$ and $a_i \cdot t_{p_i}$. Inversely, if Δt_i and $t_{p_{i-1}}$ are given, then, one can use these equations to determine the value of a_i (e.g., using iterations on Eq. (2.37) or using graphical methods). A typical identification procedure would be the following. Q_{p_o}, t_{p_o}, t_{r_o} are given. The observed output parameters Q_{p_n}, t_{p_n}, t_{r_n} and Δt (translation time of the peak) are also given. The problem is to determine the sequence of a_i,

$i = 1, 2, \ldots, n$ and the value of n. One can immediately check if one reservoir is adequate, from Eq. (2.36). If n can be one, then the value of a_i is determined from the set of Equations (2.36) through (2.39) as previously indicated. If one reservoir is not adequate, it seems unavoidable to follow a trial and error procedure. Δt_i would be specified for each added reservoir and the corresponding values of the parameters would be obtained from the available equations. Δt_i needs to be specified such that the sum over i will give the total observed translation time.

Although tedious, the procedure described above eliminates the need for costly simulation studies for the linear case. Simulation appears to be a necessity when dealing with nonlinear reservoirs. Exploratory studies for various values of Q_{p_o}, t_{p_o}, t_{r_o} (given in Table 2.1), indicated that Eq. (2.36) holds true with great precision for values of the exponent m in the interval [0.8, 1.6]. Figures 2.3 and 2.4 present plots of the parameter a_i vs. $\frac{Q_{p_i}}{Q_{p_{i-1}}}$ and $\frac{\Delta t_i}{t_{p_{i-1}}}$ for m = 0.8 and m = 1.6, respectively. The parameters of the input used in these figures were $Q_m = 850\ m^3/sec$, $t_{p_i} = t_{r_i} = 9.5$ hours. It requires a significant change in the value of a_i to attain a relatively small change in $\frac{Q_{p_i}}{Q_{p_{i-1}}}$ or $\frac{\Delta t_1}{t_{p_{i-1}}}$. This implies that small errors in the parameter a_i have a negligible effect on the output characteristics. These figures also bring about a phenomenon observed in identification studies, called the problem of multiple optima. Based on Figures 2.3 and 2.4, one can find combinations of a_i and m that produce identical results in terms of

Table 2.1

INPUT PARAMETERS USED IN SIMULATION STUDIES

Q_{p_o} (m^3/sec)	t_{p_o} (hours)	t_{r_o} (hours)	Volume (m^3)
100	18	18	$6.48x10^6$
100	6	6	$2.16x10^6$
500	12	12	$21.6x10^6$
500	6	18	$21.6x10^6$
1000	6	6	$21.6x10^6$
600	6	4	$10.8x10^6$
600	5	5	$10.8x10^6$
600	6	14	$21.6x10^6$
850	9.5	9.5	$29.07x10^6$

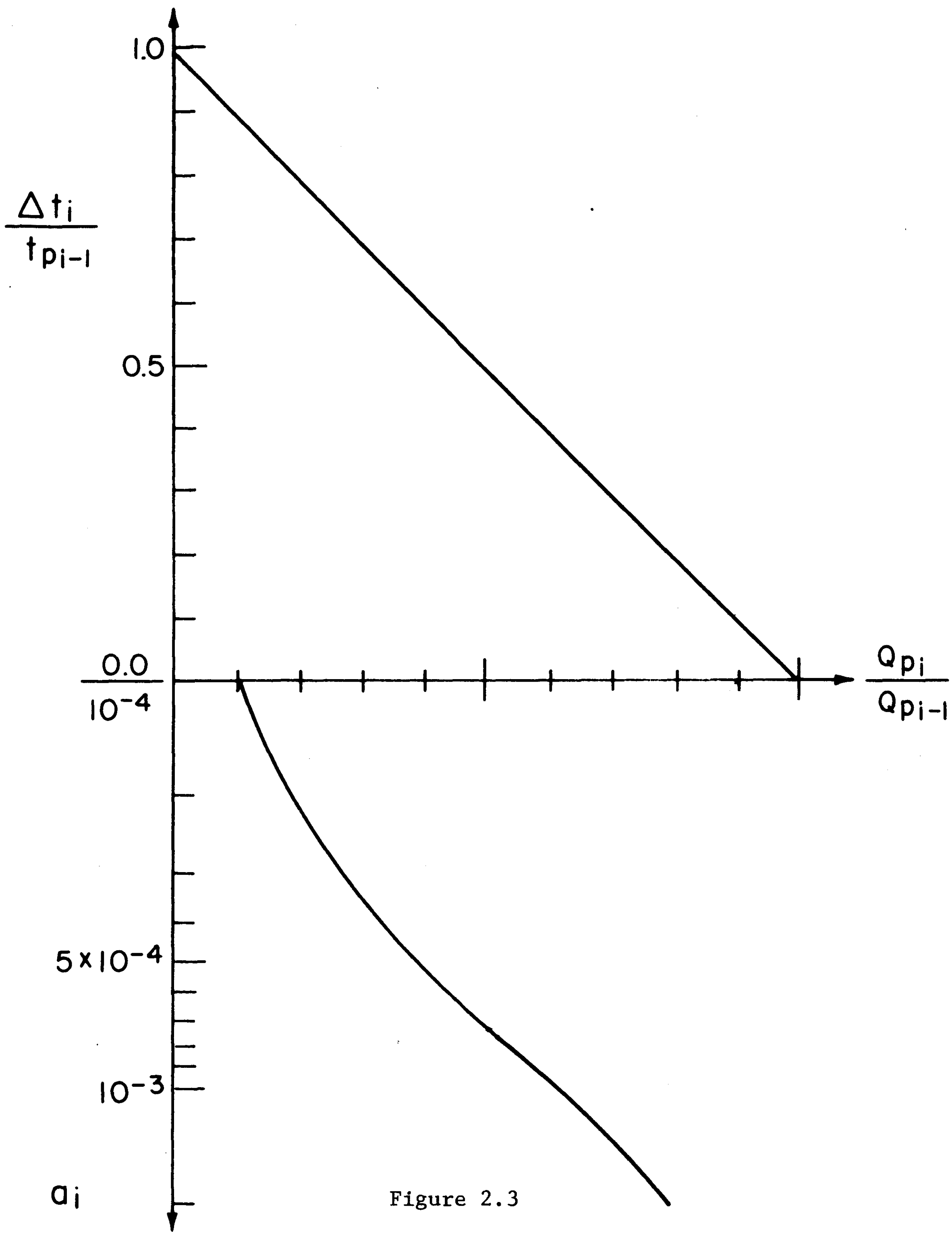

Figure 2.3

NORMALIZED MEASURES OF TRANSLATION AND ATTENUATION OF A TRIANGULAR HYDROGRAPH INPUT TO A NONLINEAR RESERVOIR AS A FUNCTION OF PARAMETER a_1. m = 0.8

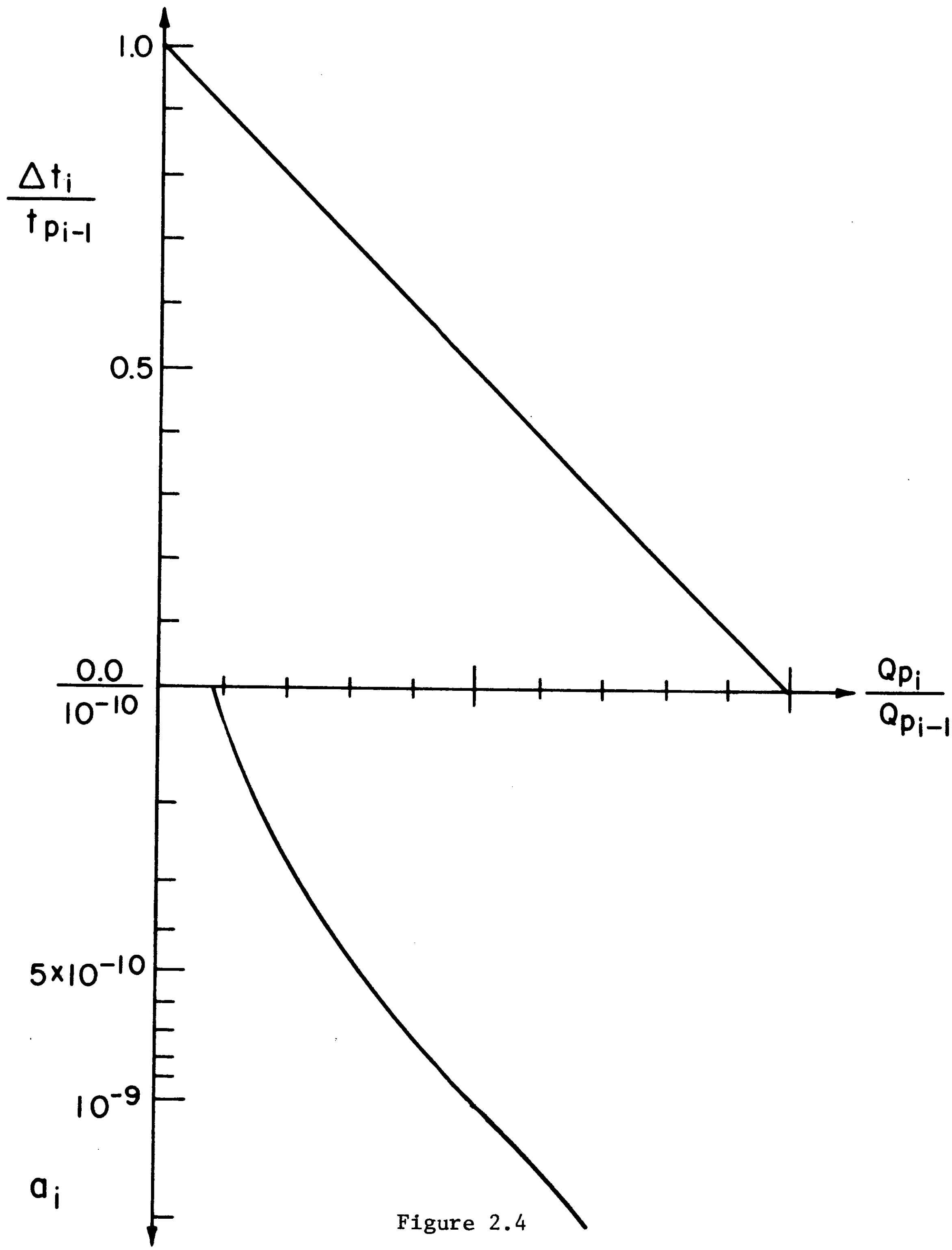

Figure 2.4

NORMALIZED MEASURES OF TRANSLATION AND ATTENUATION OF A TRIANGULAR HYDROGRAPH INPUT TO A NONLINEAR RESERVOIR AS A FUNCTION OF PARAMETER a_1. m = 1.6

$$\frac{Q_{p_i}}{Q_{p_{i-1}}} \text{ and } \frac{\Delta t_i}{t_{p_{i-1}}} .$$

Hence, it would seem that, if the objective function used in the identification study is to minimize, in some sense, deviations from $\frac{Q_{p_i}}{Q_{p_{i-1}}}$ and $\frac{\Delta t_i}{t_{p_{i-1}}}$ and a single reservoir was adequate, there would be more than one pair of a_i and m that would give minimum objective value. This is only due to the fact that the problem was not well posed. In other words, all the information available from the observed hydrograph was not used, while all the characteristic parameters of the input hydrograph were utilized to obtain the plots in Figures 2.3 and 2.4. The ambiguity is removed by adding, as an objective, the preservation of the observed $\frac{t_{p_i}}{t_{r_i}}$. Figure 2.5 shows the sensitivity of the parameter $\frac{t_{p_i}}{t_{r_i}}$ to variations in m, for the same set of input parameters. It can be stated that for a given input hydrograph and given level of attenuation,

$$\left(\frac{Q_{p_{i-1}} - Q_{p_i}}{Q_{p_{i-1}}}\right)$$

large values of m give rise to fast rising hydrographs with long duration of the falling limb.

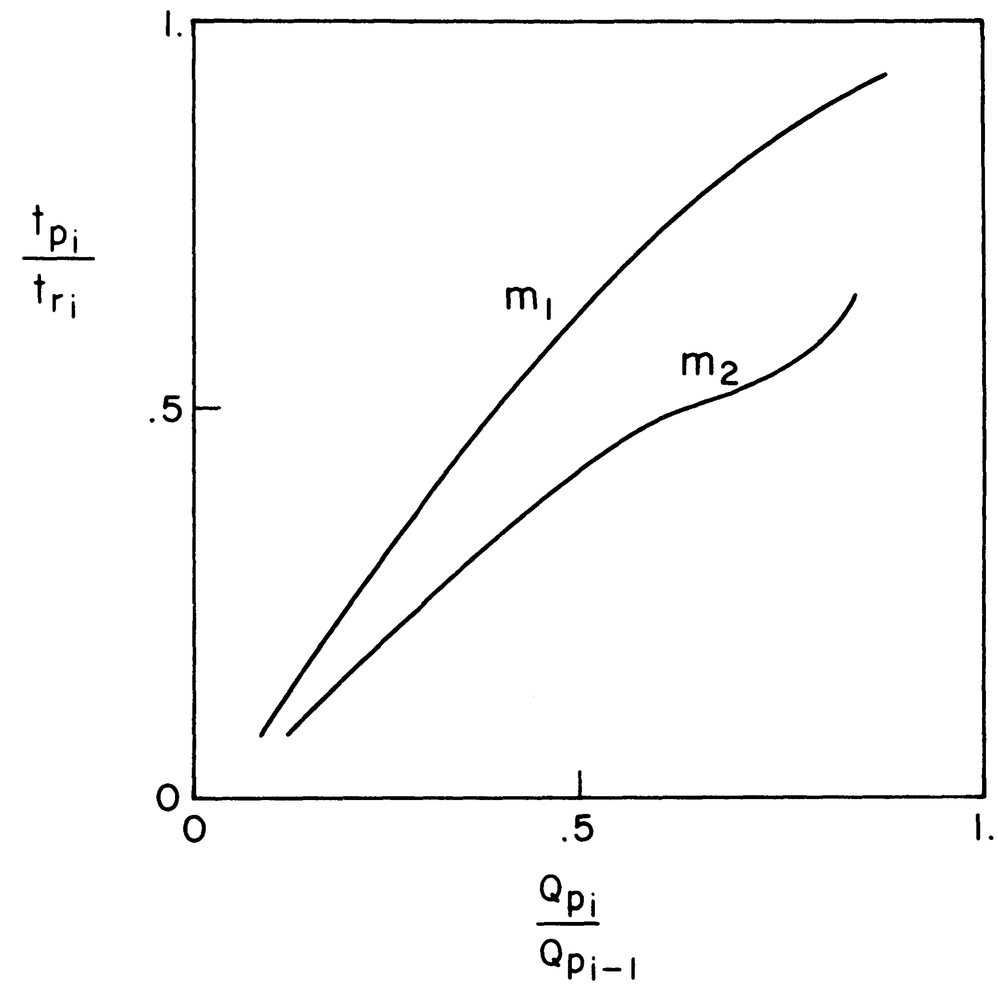

Figure 2.5

LOCUS OF FEASIBLE VALUES OF $\frac{t_{p_i}}{t_{r_i}}$ AND $\frac{Q_{p_i}}{Q_{p_{i-1}}}$ FOR A NONLINEAR RESERVOIR OF EXPONENT m_i, $i = 1,2$. $m_1 = 0.8$ and $m_2 = 1.6$

Chapter 3

ANALYTICAL DETERMINATION OF THE EXPECTED VALUE OF A NONLINEAR FUNCTION

3.1 Introduction

This chapter will derive analytical approximations to the expected value of nonlinear functions. The proposed methodology can be applied to any scalar or vector valued functions of a scalar or vector random variable. The assumptions of this development are

a. the nonlinear function under study possesses derivatives of some order, and

b. the joint probability density of the independent variables is characterized by a finite number of parameters.

The idea is to expand the elements of the vector valued function in a Taylor series about the point whose coordinates are the mean values of the independent variables. Then, use the expectation operator with each term in the series expansion. Since it was assumed that the joint probability density of the independent variables is characterized by a finite number of parameters, it follows, in principle, that the joint central moments of any order of these variables will be functions of these parameters only. Thus, the desired expected value will be expressed as a series in the parameters of the given joint distribution.

The resultant series will not be convergent, in general, since the expectation operator will utilize values of the independent variables that lie outside the radius of convergence of the initial Taylor series expansion. Hence, the approximation series will have minimum residual error for a finite number of terms. Studies of the magnitude of the individual terms in the series or comparison with numerical integration schemes can be used to determine the number of terms that provide the best approximation to the expected value of the nonlinear function under study.

This methodology will be useful in cases where analytical expressions are sought to assess the importance of the parameters of the input joint distribution to the expected value of the nonlinear function, or in cases where the expected value of the nonlinear function has to be calculated repeatedly (e.g., in a sequential estimator formulation), since the analytically derived solution is more efficient and economical than numerical integration schemes.

In the following, the case of a scalar function of one variable will be treated first. Generalization to higher dimensional functions is given subsequently. At the end, examples of application of the proposed methodology within the flood routing problem are presented and a comparison with numerical integration schemes is made.

3.2 Expected Value of a Nonlinear Function of Random Variables or Vectors

Consider the function f(x) of the random variable x, whose probability density function is denoted by $p_x(x; \underline{a})$. $\underline{a}$ represents the vector (of finite dimension) with elements $a_1, a_2, \ldots, a_\ell$ the parameters of $p_x(x; \underline{a})$. At this point, it should be noted that **x** is a continuous random variable.

A Taylor series expansion of f(x) about E{x} yields:

$$f(x) = f(E\{x\}) + \left.\frac{df(x)}{dx}\right|_{E\{x\}} \frac{(x - E\{x\})}{1!} + \left.\frac{d^2f(x)}{dx^2}\right|_{E\{x\}} \frac{(x - E\{x\})^2}{2!} + \ldots \quad (3.1)$$

If f(x) is approximated by the sum of the first N+1 terms in ther series of Eq. (3.1), this expression can be rewritten as:

$$f(x) = \sum_{n=0}^{N} \left.\frac{d^n f(x)}{dx^n}\right|_{E\{x\}} \frac{(x - E\{x\})^n}{n!} \quad (3.2)$$

where $n! = 1 \cdot 2 \ldots (n-1) \cdot n$. In Equation (3.2), $\sum_{n=0}^{N} (\cdot)$ represents the sum of the first N+1 terms starting with the 0^{th} order term, and $\left.\frac{d^n f(x)}{dx^n}\right|_{E\{x\}}$ is the n^{th} order derivative of the function f(x) evaluated at the point E{x}.

The expected value of f(x) can be approximated by means of Eq. (3.2),

$$E\{f(x)\} = E\left\{\sum_{n=0}^{N} \left.\frac{d^n f(x)}{dx^n}\right|_{E\{x\}} \cdot \frac{(x - E\{x\})^n}{n!}\right\} \quad (3.3)$$

Interchange of the expectation operator E(•) and the summation operator Σ(•) (since they both are linear operators) results in:

$$E\{f(x)\} = \sum_{n=0}^{N} \frac{1}{n!} \cdot \left.\frac{d^n f(x)}{dx^n}\right|_{E\{x\}} \cdot E\{(x - E\{x\})^n\} \quad (3.4)$$

Clearly, Equation (3.4) gives the expected value of the nonlinear function f(x) in a series involving the central moments of the random variable x, up to the N^{th} order, as well as the first initial moment E{x}. In principle, these moments can be expressed as functions of the parameters of the distribution of x. Expressions for the n^{th} central and the first moment about the origin of the most common probabilistic models are presented below. The details of their derivation are given in Appendix A.

Exponential

$$\text{p.d.f.:} \quad p_X(x;\ \lambda) = \lambda \cdot e^{-\lambda x} \quad ; \quad \lambda > 0;\ x \geq 0 \tag{3.5}$$

$$E\{x\} = 1/\lambda \tag{3.6}$$

$$E\{(x - E(x))^n\} = \frac{n!}{\lambda^n} \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!} \tag{3.7}$$

Rectangular

$$\text{p.d.f.:} \quad p_X(x;\ \mu,\ \lambda) = \begin{cases} \frac{1}{\lambda}\ , & \mu - \frac{\lambda}{2} \leq x \leq \mu + \frac{\lambda}{2} \\ 0\ , & x < \mu - \frac{\lambda}{2}\ ,\ x > \mu + \frac{\lambda}{2} \end{cases} \qquad \lambda > 0 \tag{3.8}$$

$$E\{x\} = \mu \tag{3.9}$$

$$E\{(x-E\{x\})^n\} = \frac{n!}{\lambda} \cdot \sum_{i=0}^{n} (-1)^i \cdot \frac{\mu^i}{(n-i+1)!\ i!} \left[(\mu + \tfrac{\lambda}{2})^{n-i+1} - (\mu - \tfrac{\lambda}{2})^{n-i+1}\right] \tag{3.10}$$

Gamma

$$\text{p.d.f.:} \quad p_X(x;\ \mu) = \frac{x^{\mu-1} \cdot e^{-x}}{\Gamma(\mu)} \quad ; \quad \mu > 0\ , \quad x > 0 \tag{3.11}$$

where if μ is an integer $\Gamma(\mu) = (\mu-1)!$, while if μ is not an integer $\Gamma(\mu) = (\mu-1)(\mu-2)\ldots\delta\cdot\Gamma(\delta)$ where δ is a positive real number less than one.

$$E\{x\} = \mu \tag{3.12}$$

$$E\{(x-E\{x\})^n\} = n!\cdot\sum_{i=0}^{n-1}\frac{(-1)^i\cdot\mu^i}{(n-i)!\ i!}\cdot\prod_{j=1}^{n-i}(\mu+n-i-j) + n!(-1)^n\cdot\frac{\mu^n}{n!} \tag{3.13}$$

Log-normal

$$\text{p.d.f.:}\quad p_X(x;\ m,\ \sigma) = \frac{1}{x\cdot\sqrt{2\pi}\cdot\sigma}\cdot e^{-\frac{1}{2}\left[\frac{1}{\sigma}\cdot\ln(\frac{x}{m})\right]^2};\ x \geq 0,\ \sigma > 0 \tag{3.14}$$

$$E\{x\} = m\cdot e^{\frac{1}{2}\cdot\sigma^2}$$

$$E\{(x-E\{x\})^n\} = m^n\cdot n!\cdot\sum_{i=0}^{n}\frac{1}{(n-i)!\ i!}\cdot(-1)^i\cdot e^{\frac{1}{2}\cdot\sigma^2[(n-i)^2+i]} \tag{3.15}$$

Normal

$$\text{p.d.f.:}\quad p_X(x;\ \mu,\ \sigma) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2};\ \sigma > 0 \tag{3.16}$$

$$E\{x\} = \mu \tag{3.17}$$

$$E\{(x-E\{x\})^n\} = \begin{cases} 1\cdot3\cdot5..(n-1)\,\sigma^n, & n = 2k;\ k = 0,\ 1,\ 2,\ \ldots \\ 0 & ,\ n = 2k+1;\ k = 0,\ 1,\ 2,\ \ldots \end{cases} \tag{3.18}$$

A result similar to Equation (3.4) is available for the expectation of $f(x_1,\ \ldots,\ x_m)$: $E\{f(x_1,\ \ldots,\ x_m)\}$. Extension to vector valued functions is straightforward and can be done by applying the following development to each element of the vector valued function.

Expansion of $f(x_1, \ldots, x_m)$ about the point $(E\{x_1\}, E\{x_2\}, \ldots, E\{x_m\})$ in the m dimensional space, in an N terms Taylor series yields (Hildebrand, 1976):

$$f(\underline{x}) = \sum_{n=0}^{N-1} \frac{1}{n!} \cdot \left([x_1 - E\{x_1\}] \cdot \frac{\partial}{\partial_{x_1}} + \ldots + [x_m - E\{x_m\}] \cdot \frac{\partial}{\partial_{x_m}}\right)^n \cdot f(x_1, \ldots, x_m)\Bigg|_{\substack{x_1 = E\{x_1\} \\ \vdots \quad \vdots \\ x_m = E\{x_m\}}} \tag{3.19}$$

Given real numbers $y_1, y_2, \ldots, y_m$,

$$(y_1 + \ldots + y_m)^n = \sum_{\substack{n_1, \ldots, n_m \\ n_1 + n_2 + \ldots + n_m = n}} \frac{n!}{n_1!\, n_2! \ldots n_m!} \cdot y_1^{n_1} \cdot y_2^{n_2} \ldots \cdot y_m^{n_m} \tag{3.20}$$

Using Eq. (3.20) in Eq. (3.19) and defining,

$$r_i = x_i - E\{x_i\} \quad ; \quad i = 1, 2, \ldots, m \tag{3.21}$$

$$\underline{x} = [x_1\ x_2\ \ldots\ x_m]^T \tag{3.22}$$

$$\underline{\mu} = [E\{x_1\}\ E\{x_2\}\ \ldots\ E\{x_m\}]^T \tag{3.23}$$

(where "T" denotes the transpose of a matrix or vector quantity) results in ,

$$f(\underline{x}) = \sum_{n=0}^{N-1} \frac{1}{n!} \cdot \sum_{\substack{n_1,\ldots,n_m \\ n_1+\ldots+n_m=n}} \frac{n!}{n_1!\, n_2! \ldots n_m!} \cdot r_1^{n_1} \cdot r_2^{n_2} \ldots \cdot r_m^{n_m} \cdot \left. \frac{\partial^n \cdot f(\underline{x})}{(\partial x_1)^{n_1} \cdot (\partial x_2)^{n_2} \ldots \cdot (\partial x_m)^{n_m}} \right|_{\underline{x} = \underline{\mu}} \tag{3.24}$$

Thus, the expected value of the function $f(\underline{x})$ is given by means of Equation (3.24) as follows

$$E\{f(\underline{x})\} = \sum_{n=0}^{N-1} \sum_{\substack{n_1,\ldots,n_m \\ n_1+\ldots+n_m=n}} \frac{1}{n_1!\, n_2! \ldots n_m!} \cdot \left. \frac{\partial^n f(\underline{x})}{(\partial x_1)^{n_1} \cdot (\partial x_2)^{n_2} \ldots (\partial x_m)^{n_m}} \right|_{\underline{x} = \underline{\mu}} \cdot E\left\{ r_1^{n_1} \ldots \cdot r_m^{n_m} \right\} \tag{3.25}$$

In the derived expression, the n^{th} order central moment of the random vector $\underline{x} = [x_1, x_2, \ldots, x_m]^T$ is involved. If this moment can be expressed as a function of a finite number of parameters, characterizing the joint probability density of $\underline{x}$, then the number N* that provides the best approximation to the $E\{f(\underline{x})\}$ can be determined by comparison with numerical integration schemes.

Expressions for the n^{th} order central moment of the common multivariate normal distribution are as follows:

$$\text{p.d.f.:} \quad p_{\underline{X}}(\underline{x}; \underline{\mu}\ \overline{\underline{P}}) = \frac{1}{(2\pi)^{m/2} \cdot (\det \overline{\underline{P}})^{1/2}} \cdot e^{-\frac{1}{2} \cdot (\underline{x}-\underline{\mu})^T \cdot \overline{\underline{P}}^{-1} \cdot (\underline{x} \cdot \underline{\mu})} \tag{3.26}$$

where $\overline{\underline{P}}$ is a positive definite, mxm, symmetric matrix, "det $\overline{\underline{P}}$" denotes the determinate of the matrix $\overline{\underline{P}}$, "$\overline{\underline{P}}^{-1}$" denotes the inverse of the matrix $\overline{\underline{P}}$, $\underline{x}$ and $\underline{\mu}$ are m dimensional vectors.

$$E\{\underline{x}\} = \underline{\mu} \tag{3.27}$$

$$E\{r_{i_1} \cdot r_{i_2} \cdots r_{i_L}\} = \begin{cases} 0, & L = 2k+1,\ k = 0, 1, \ldots \\ \sum\limits_{\substack{j_1, j_2, \ldots, j_L \\ P_e\{i_1, i_2, \ldots, i_L\}}} \sigma^2_{j_1 j_2} \cdot \sigma^2_{j_3 j_4} \cdot \ldots\ \sigma^2_{j_{L-1} j_L}, & L=2k,\ k=0,1,\ldots \end{cases} \tag{3.28}$$

where $P_e\{i_1, i_2, \ldots, i_L\}$ denotes all distinct pairs of subscripts j_1, j_2, ..., j_L that are permutations of $\{i_1, i_2, \ldots, i_L\}$. $\sigma^2_{j_i j_k}$ is the (j_i, j_k) element of matrix $\overline{\underline{P}}$. Note that repetition is allowed in the subscripts i_1, i_2, ..., i_L. For example,

$$E\{(x_1 - \mu_1) \cdot (x_2 - \mu_2) \cdot (x_3 - \mu_3) \cdot (x_4 - \mu_4)\}$$

$$= \sigma^2_{12} \cdot \sigma^2_{34} + \sigma^2_{13} \cdot \sigma^2_{24} + \sigma^2_{14} \cdot \sigma^2_{23}$$

$$E\{(x_i - \mu_i)^4\} = 3 \cdot \sigma^2_{ii}$$

3.3 A Taylor-Gauss Application

Non-linear flood routing provides several examples of the methodology presented in the previous section. In particular, expectations used in the determination of the statistical linearization gains (see Chapter 4) will be approximated for the case when the independent variables involved are normally distributed with given means, variances and covariances.

3.3.1 Nonlinear Function of a Single Random Variable

Consider the function

$$f(x_i) = x_i^m \tag{3.29}$$

Assume that x_i can be modeled as the sum of a mean level, μ_i, and a Gaussian random process, r_i, with zero mean and variance equal to σ_i^2, where the dependence of x_i, μ_i, r_i, σ_i^2 on time is not shown explicitly for notational convenience. Then:

$$x_i = \mu_i + r_i \tag{3.30}$$

and

$$f(r_i) = (\mu_i + r_i)^m \tag{3.31}$$

A Taylor series expansion of the function $f(r_i)$ about the point $E\{r_i\} = 0$, keeping the N+1 leading terms, yields,

$$(\mu_i + r_i)^m = \mu_i^m + \frac{m}{1!}\,\mu_i^{(m-1)} \cdot r_i + \frac{m\cdot(m-1)}{2!} \cdot \mu_i^{(m-2)} \cdot r_i^2 + \ldots + \frac{m\cdot(m-1)\ \ldots\ (m - N + 1)}{N!} \cdot \mu_i^{(m-N)} \cdot r_i^N \tag{3.32}$$

Applying the expectation operator to both sides of Eq. (3.32) results in

$$E\{(\mu_i + r_i)^m\} = \mu_i^m + \frac{m}{1!} \cdot \mu_i^{(m-1)} \cdot E\{r_i\} + \frac{m(m-1)}{2!} \cdot \mu_i^{(m-2)} \cdot E\{r_i^2\} + \ldots + \frac{m\cdot(m-1)\ \ldots\ (m - N + 1)}{N!} \cdot \mu_i^{(m-N)} \cdot E\{r_i^N\} \tag{3.33}$$

By means of Eq. (3.18), the function $E\{(\mu_i + r_i)^m\}$ can be rewritten as:

$$E\{(\mu_i + r_i)^m\} = \mu_i^m \cdot \left\{ 1 + \sum_{k=1}^{N'} \frac{\prod_{i=0}^{2k-1} (m-i) \cdot \prod_{j=1}^{k} (2 \cdot j - 1)}{(2 \cdot k)!} \cdot V_i^{2k} \right\} \tag{3.34}$$

with

$$N' = \frac{(\text{max even integer} \leq N)}{2}$$

where the coefficient of variation, V_i, of the random variable x_i is given by,

$$V_i = \frac{\sigma_i}{\mu_i} \tag{3.35}$$

In a similar manner, the function $E\{r_i(\mu_i + r_i)^m\}$ can be approximated by,

$$E\{r_i \cdot (\mu_i + r_i)^m\} = \mu_i^{(m+1)} \cdot \sum_{k=1}^{N'} \frac{\prod_{i=0}^{2k-2} (m-i) \cdot \prod_{j=1}^{k} (2 \cdot j - 1)}{(2 \cdot k - 1)!} \cdot V_i^{2k} \tag{3.36}$$

Approximate analysis of the magnitude of the residual terms is possible, leading to a procedure to analytically determine the order N' that minimizes the magnitude of the last term in the series. Alternatively, comparison of the results derived by the methodology proposed (called a Taylor-Gauss approximation) with numerical integration results can determine the value of N' that minimizes the truncation error.

Denote by $R1_{N'}$, $R2_{N'}$, the N'^{th} term of the series in Eqs. (3.34) and (3.36), respectively. Then,

$$R1_N = \mu_i^m \cdot \prod_{i=0}^{2N'-1} (m-i) \cdot \prod_{j=1}^{N'} (2j-1) \cdot \frac{V_i^{2N'}}{(2N')!} \quad ; \quad N' \geq 1 \tag{3.37a}$$

$$R1_0 = 1 \tag{3.37b}$$

$$R2_{N'} = \mu_i^{(m+1)} \cdot \prod_{i=0}^{2N'-2} (m-i) \cdot \prod_{j=1}^{N'} (2j-1) \cdot \frac{V_i^{2N'}}{(2N'-1)!} \quad ; \quad N' \geq 1 \tag{3.38}$$

To determine the N_1^*, N_2^* that minimize the terms $R1_{N'}$, $R2_{N'}$ over the range of N', one has to investigate the ratio of two successive terms. Namely,

$$S1_{N'} = \frac{R1_{N'}}{R1_{N'-1}} \tag{3.39}$$

$$S2_{N'} = \frac{R2_{N'}}{R2_{N'-1}} \tag{3.40}$$

In regions where $|S1_{N'}|$ ($|S2_{N'}|$) is less than one, the successive terms are decreasing in magnitude, while where the absolute value is greater than one, the successive terms are increasing in magnitude. Hence, the maxima N_1^*, N_2^* are sought, that result in $|S1_{N+1}|$ and $|S2_{N+1}|$ equal to or less than one.

Direct substitution of Eqs. (3.37) and (3.38) in Eqs. (3.39) and (3.40), respectively, gives:

$$|S1_{N'}| = \frac{|(m-2N'+2) \cdot (m-2N'+1)|}{2N'} \cdot V_i^2 \quad ; \quad N' \geq 1 \tag{3.41}$$

$$|S2_{N'}| = \frac{|(m-2N'+2) \cdot (m-2N'+3)|}{2N'-2} \cdot V_i^2 \quad ; \quad N' \geq 2 \tag{3.42}$$

The value N_1^* is the solution to the following mathematical integer programming problem,

$$\max_{N'} N'$$

subject to

$$\frac{|(m-2N'+2) \cdot (m-2N'+1)|}{2N'} \cdot V_i^2 \leq 1$$

N': integer ≥ 1; m, V_i are given, while the integer programming problem to find N_2^* is

$$\max_{N'} N'$$

subject to

$$\frac{|(m-2N'+2) \cdot (m-2N'+3)}{(2N'-2)} \cdot V_i^2 \leq 1$$

N': integer ≥ 2; m, V_i are given.

Solution of the integer programming problems presented above (e.g., by enumeration), for different values of the parameters m and V_i^2, leads to Figure 3.1. There, plots of N_1^*, N_2^* vs. V_i are presented for different values of m. Sensitivity analysis showed that N_1^* and N_2^* are relatively insensitive to the value of m for m in the range (0.8, 2.0), which is the range of m reported in the literature as applicable to non-linear reservoir routing models.

The asymptotic behavior of the series in Equations (3.34) and (3.36) is depicted in Figure 3.2. In this figure, the magnitudes (absolute value) of the ratios of two successive terms in the series is plotted against the order of the term of the numerator in the ratio (N') for m = 0.8 and V_i = 0.4. Similar results were obtained (not shown) for different parameter values. Notice that the absolute value of the ratios

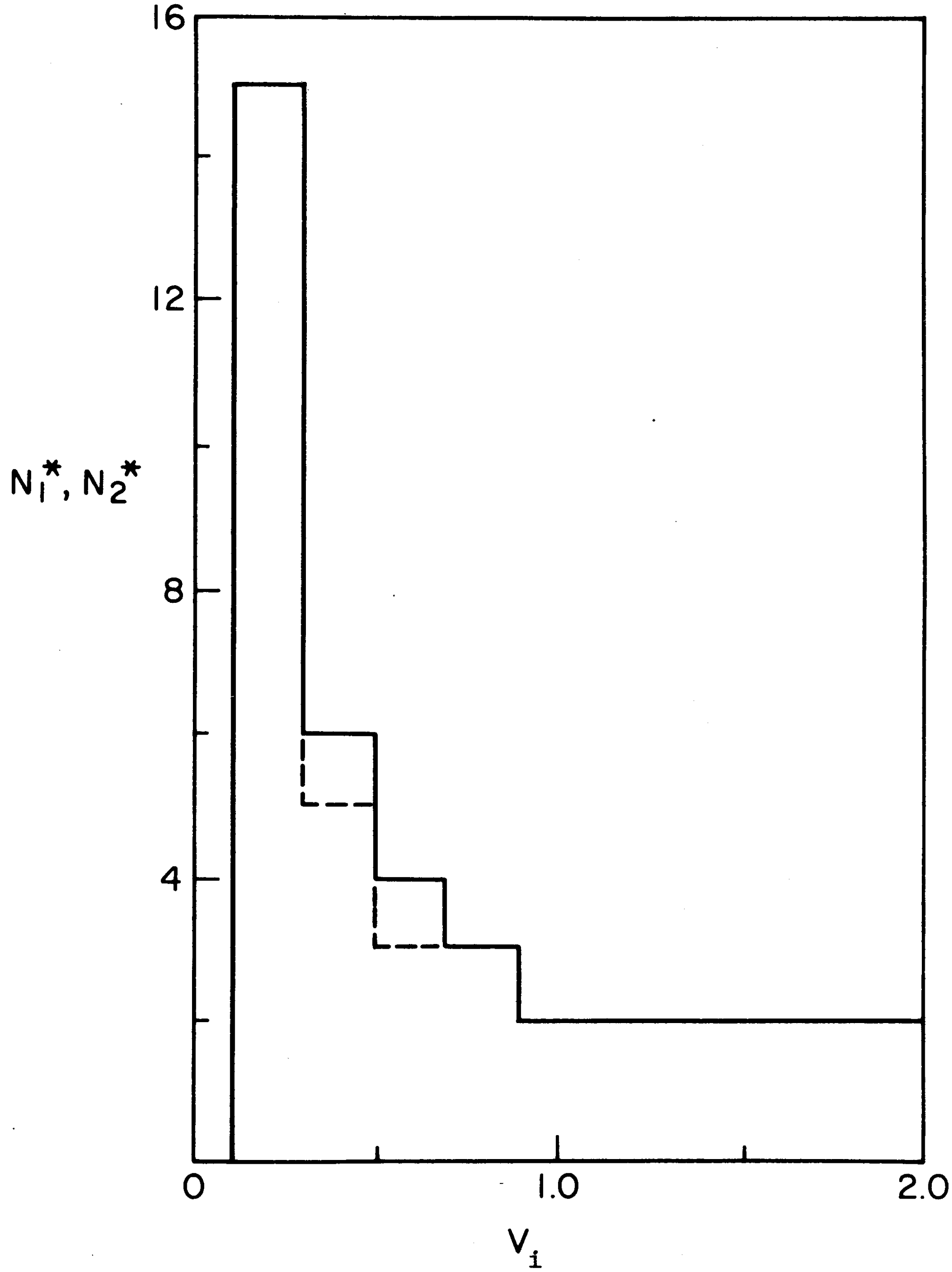

Figure 3.1

OPTIMAL INTEGERS N_1^*, N_2^* AS FUNCTIONS OF V_i

SOLID LINE IS FOR N_1^*, DASHED LINE IS FOR N_2^*

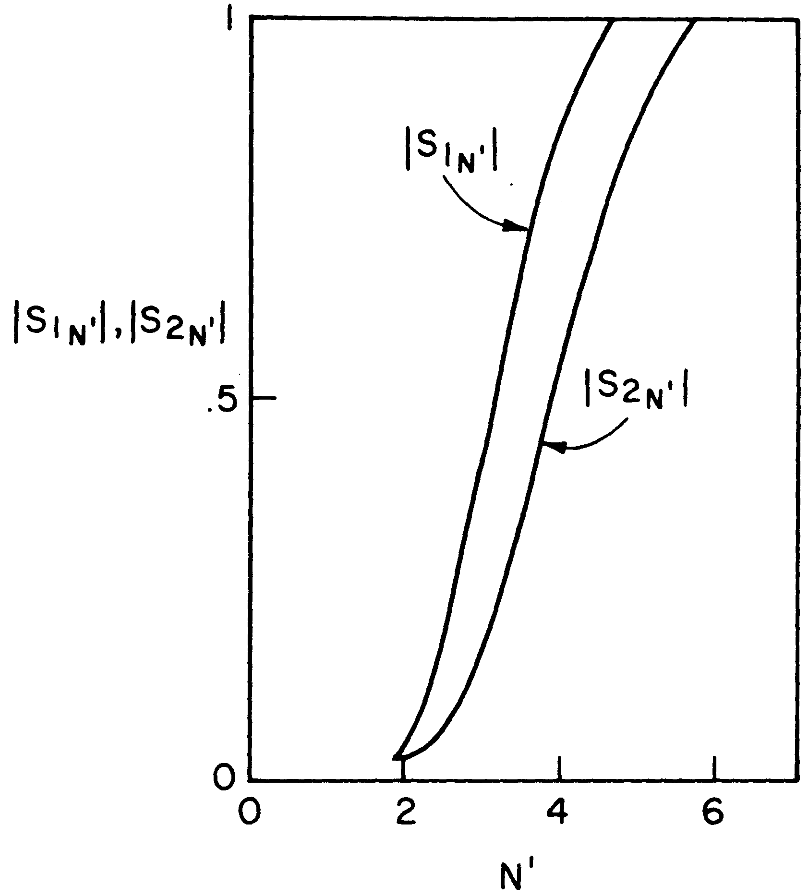

Figure 3.2

VARIATIONS OF $|S_{1_{N'}}|$, $|S_{2_{N'}}|$ DUE TO CHANGES IN THE NUMBER OF TERMS N', FOR m = 0.8 AND V_i = 0.4

decreases rapidly with N'. Thus, it is expected to obtain a very good approximation by just keeping the first two or three terms of the series in Equations (3.34) and (3.36).

Summarizing, it is possible to state the following:

a. For given m and V_i, there is a minimum value for $|R_{i_N}|$; i = 1, 2. Thus, the series approximation used in the Taylor-Gauss method generates always an error which cannot be reduced (i.e., the series is an asymptotic one).

b. N_i^*; i = 1, 2 depends largely on V_i. It decreases rapidly as V_i increases.

c. N_i^*; i = 1, 2 is relatively insensitive to the value of m, for m in the range (0.8, 2.0). Nevertheless, it decreases as m decreases.

d. A few terms in the series are enough to give a good approximation to the functions $E\{(\mu_i + r_i)^m\}$ and $E\{r_i \cdot (\mu_i + r_i)^m\}$. Keeping only the two leading terms, these functions are approximated by (i.e., keeping up to fourth order terms in the original Taylor series),

$$E\{(\mu_i + r_i)^m\} = \mu_i^m \cdot \{1 + \frac{m(m-1)}{2} \cdot V_i^2\} \tag{3.43}$$

and

$$E\{r_i \cdot (\mu_i + r_i)^m\} = \mu_i^{m+1} \cdot \{m \cdot V_i^2 + \frac{m(m-1)(m-2)}{2} \cdot V_i^4\} \tag{3.44}$$

3.3.2 Nonlinear Function of Two Random Variables

An approach similar to the above is possible to determine the expected value of a nonlinear function of two random variables which are assumed to obey a normal probability law. This situation arises very often within a sequential estimator algorithm (Chapter 5) when the simultaneous estimation of the states and parameters of a stochastic nonlinear system is sought. As will be shown in Chapter 4, statistical linearization gains depend on functions of the type to be treated below.

Consider the function $g_1(a_i, x_i)$ given by

$$g_1(a_i, x_i) = a_i \cdot x_i^m = g_1(r_{x_i}, r_{a_i}) \tag{3.45}$$

where

$$a_i = \mu_{a_i} + r_{a_i} \tag{3.46}$$

$$x_i = \mu_{x_i} + r_{x_i} \tag{3.47}$$

with μ_{a_i}, μ_{x_i} being the means of the random variables a_i, x_i and r_{a_i}, r_{x_i} the normally distributed random residuals with variances $\sigma^2_{a_i}$, $\sigma^2_{x_i}$ respectively, and with covariance $\sigma^2_{x_i a_i}$.

Also define

$$g_2(r_{x_i}, r_{a_i}) = r_{x_i} \cdot (\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m \cdot \tag{3.48}$$

$$g_3(r_{x_i}, r_{a_i}) = r_{a_i} \cdot (\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m \tag{3.49}$$

The Taylor-Gauss method will be used to approximate the functions $E\{g_1(r_{x_i}, r_{a_i})\}$, $E\{g_2(r_{x_i}, r_{a_i})\}$ and $E\{g_3(r_{x_i}, r_{a_i})\}$.

At first, the partial derivatives of the functions g_1, g_2 and g_3 are,

$$\left.\frac{\partial^{n_1+n_2} g_1(r_{x_i}, r_{a_i})}{(\partial r_{x_i})^{n_1} \cdot (\partial r_{a_i})^{n_2}}\right|_{(0,0)} = \begin{cases} m \cdot (m-1) \ldots (m-n_1+1) \cdot \mu_{x_i}^{(m-n_1)} \cdot \mu_{a_i}^{(1-n_2)} & ; n_2 \leq 1,\ n_1 > 0 \\ 0 & ; n_2 > 1 \end{cases} \quad (3.50)$$

$$\left.\frac{\partial^{n_1+n_2} g_2(r_{x_i}, r_{a_i})}{(\partial r_{x_i})^{n_1} \cdot (\partial r_{a_i})^{n_2}}\right|_{(0,0)} = \begin{cases} n_1 \cdot m \cdot (m-1) \ldots (m-n_1+2) \cdot \mu_{x_i}^{(m-n_1+1)} \cdot \mu_{a_i}^{(1-n_2)} & ; n_2 \leq 1, n_1 > 0 \\ 0 & ; n_2 > 1 \end{cases} \quad (3.51)$$

and

$$\left.\frac{\partial^{n_1+n_2} g_3(r_{x_i}, r_{a_i})}{(\partial r_{x_i})^{n_1} \cdot (\partial r_{a_i})^{n_2}}\right|_{(0,0)} = \begin{cases} n_2 \cdot m \cdot (m-1) \ldots (m-n_1+1) \cdot \mu_{x_i}^{(m-n_1)} \cdot \mu_{a_i}^{(2-n_2)} & ; n_2 \leq 2, n_1 > 0 \\ 0 & ; n_2 > 2 \end{cases} \quad (3.52)$$

The n^{th} order central moment of two jointly Gaussian random variables is given by ($n = n_1 + n_2$):

$$E\{r_{x_i}^{n_1} \cdot r_{a_i}^{n_2}\} = (n_1+n_2-1) \cdot (n_1+n_2-3) \ldots 3 \cdot 1 \cdot \sigma_{x_i}^{n_1-n_2} \cdot \sigma_{x_i a_i}^{2n_2} \quad ; \quad n_2 \leq 1 \quad (3.53)$$

$$E\{r_{x_i}^{n_1} \cdot r_{a_i}^{n_2}\} = (n_1-1) \cdot (n_1-3) \ldots 3 \cdot 1 \cdot \sigma_{x_i}^{n_1} \cdot \sigma_{a_i}^{n_2} + n_1 \cdot (n_1-1) \cdot (n_1-3) \ldots 3 \cdot 1 \cdot \sigma_{x_i}^{n_1-n_2} \cdot \sigma_{x_i a_i}^{2n_2} \ ; \ n_2 = 2 \quad (3.54)$$

Using up to fourth order terms in the Taylor series expansion of Eq. (3.25) (in view of its asymptotic behavior) and using Eqs. (3.50) through (3.54) yields,

$$E\{g_1(r_{x_i},r_{a_i})\} = \mu_{a_i}\cdot\mu_{x_i}^{m} + \frac{m\cdot(m-1)}{2}\cdot\mu_{a_i}\,\mu_{x_i}^{(m-2)}\cdot\sigma_{x_i}^{2} + m\cdot\mu_{x_i}^{(m-1)}\cdot\sigma_{x_ia_i}^{2} \tag{3.55}$$

$$E\{g_2(r_{x_i},r_{a_i})\} = m\cdot\mu_{a_i}\cdot\mu_{x_i}^{(m-1)}\cdot\sigma_{x_i}^{2} + \mu_{x_i}^{m}\cdot\sigma_{x_ia_i}^{2} + \frac{m\cdot(m-1)\cdot(m-2)}{6}\cdot\mu_{a_i}\cdot\mu_{x_i}^{(m-3)}\cdot\sigma_{x_i}^{4} + \frac{3\cdot m\cdot(m-1)}{2}\cdot\mu_{x_i}^{(m-2)}\cdot\sigma_{x_i}^{2}\cdot\sigma_{x_ia_i}^{2} \tag{3.56}$$

$$E\{g_3(r_{x_i},r_{a_i})\} = m\cdot\mu_{a_i}\cdot\mu_{x_i}^{(m-1)}\cdot\sigma_{x_ia_i}^{2} + \mu_{x_i}^{m}\cdot\sigma_{a_i}^{2} + \frac{m\cdot(m-1)\cdot(m-2)}{2}\cdot\mu_{a_i}\cdot\mu_{x_i}^{(m-3)}\cdot\sigma_{x_i}^{2}\cdot\sigma_{x_ia_i}^{2} + \frac{m\cdot(m-1)}{2}\cdot\mu_{x_i}^{(m-2)}\cdot[\sigma_{x_i}^{2}\cdot\sigma_{a_i}^{2} + 2\cdot\sigma_{x_ia_i}^{4}] \tag{3.57}$$

Normalization of Eqs. (3.55), (3.56) and (3.57) results in:

$$E\{g_1(r_{x_i},r_{a_i})\}/[\mu_{a_i}\cdot\mu_{x_i}^{m}] = 1 + \frac{1}{2}\cdot m\cdot(m-1)\cdot V_{x_i}^{2} + m\cdot\rho_{x_ia_i}\cdot V_{x_i}\cdot V_{a_i} \tag{3.58}$$

$$E\{g_2(r_{x_i},r_{a_i})\}/[\mu_{a_i}\cdot\mu_{x_i}^{(m+1)}] = m\cdot V_{x_i}^{2} + \rho_{x_ia_i}\cdot V_{x_i}\cdot V_{a_i} + \frac{m\cdot(m-1)\cdot(m-2)}{2}\cdot V_{x_i}^{4} + \frac{3\cdot m\cdot(m-1)}{2}\cdot V_{x_i}^{2}\cdot\rho_{x_ia_i}\cdot V_{x_i}\cdot V_{a_i} \tag{3.59}$$

$$E\{g_3(r_{x_i},r_{a_i})\}/[\mu_{a_i}^2 \cdot \mu_{x_i}^m] = m \cdot \rho_{x_i a_i} \cdot V_{x_i} \cdot V_{a_i} + V_{a_i}^2$$

$$+ \frac{m \cdot (m-1) \cdot (m-2)}{2} \cdot V_{x_i}^2 \cdot \rho_{x_i a_i} \cdot V_{x_i} \cdot V_{a_i} + \frac{m \cdot (m-1)}{2} \cdot [V_{x_i}^2 \cdot V_{a_i}^2$$

$$+ 2 \cdot \rho_{x_i a_i}^2 \cdot V_{x_i}^2 \cdot V_{a_i}^2] \tag{3.60}$$

where

$$V_{x_i} = \frac{\sigma_{x_i}}{\mu_{x_i}} \tag{3.61}$$

$$V_{a_i} = \frac{\sigma_{a_i}}{\mu_{a_i}} \tag{3.62}$$

$$\sigma_{x_i a_i} = \frac{\sigma_{x_i a_i}^2}{\sigma_{x_i} \cdot \sigma_{a_i}} \tag{3.63}$$

3.4 Comparison of Taylor-Gauss Method with Numerical Integration

The Gaussian joint probability density of the random variables r_{x_i}, r_{a_i} is:

$$p_{R_{x_i},R_{a_i}}(r_{x_i},r_{a_i}) = \frac{e^{-\frac{r_{x_i}^2 \cdot \sigma_{a_i}^2 - 2 \cdot r_{x_i} \cdot r_{a_i} \cdot \rho_{x_i a_i} \cdot \sigma_{x_i} \cdot \sigma_{a_i} + r_{a_i}^2 \cdot \sigma_{x_i}^2}{2 \cdot \sigma_{x_i}^2 \cdot \sigma_{a_i}^2 \cdot (1 - \rho_{x_i a_i}^2)}}}{2\pi \cdot \sigma_{x_i} \cdot \sigma_{a_i} \cdot (1 - \rho_{x_i a_i}^2)^{1/2}} \tag{3.64}$$

For any function $g_i(r_{x_i}, r_{a_i})$, the expected value $E\{g_i(r_{x_i}, r_{a_i})\}$ is given by,

$$E\{g_i(r_{x_i}, r_{a_i})\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g_1(r_{x_i}, r_{a_i}) \cdot p_{R_{x_i} R_{a_i}}(r_{x_i}, r_{a_i}) \, dr_{x_i} \cdot dr_{a_i} \tag{3.65}$$

Use of the functions $g_1(r_{x_i}, r_{a_i})$, $g_2(r_{x_i}, r_{a_i})$, $g_3(r_{x_i}, r_{a_i})$ (Eqs. (3.45), (3.48), (3.49)) in Eq. (3.65) and substitution for the joint probability density (Equation (3.64)) results in,

$$E\{g_1(r_{x_i}, r_{a_i})\} = \int_{-\varepsilon_1}^{+\infty}\int_{-\varepsilon_2}^{+\infty} \mu_{a_i}\cdot\mu_{x_i}^m\cdot(1+V_{a_i}\cdot t)\cdot(1+V_{x_i}\cdot u)^m\cdot \frac{e^{-\frac{u^2-2u\cdot t\cdot\rho_{x_ia_i}+t^2}{2\cdot(1-\rho_{x_ia_i}^2)}}}{2\pi(1-\rho_{x_ia_i}^2)^{1/2}}\, dudt\ ;$$

$$\varepsilon_1 > 0,\ \varepsilon_2 > 0 \qquad (3.66)$$

or

$$E\{g_1(r_{x_i}, r_{a_i})\}/(\mu_{a_i}\cdot\mu_{x_i}^m) = \frac{1}{2\pi(1-\rho_{x_ia_i}^2)^{1/2}}\cdot\int_{-\varepsilon_1}^{\infty}\int_{-\varepsilon_2}^{\infty}(1+V_{a_i}\cdot t)\cdot(1+V_{x_i}\cdot u)^m\cdot e^{-\frac{u^2-2\cdot u\cdot t\cdot\rho_{x_ia_i}+t^2}{2\cdot(1-\rho_{x_ia_i}^2)}}\, dudt \qquad (3.67)$$

Similarly,

$$E\{g_2(r_{x_i}, r_{a_i})\}/(\mu_{a_i}\cdot\mu_{x_i}^{(m+1)}) = \frac{V_{x_i}}{2\pi(1-\rho_{x_ia_i}^2)^{1/2}}\cdot\int_{-\varepsilon_1}^{+\infty}\int_{-\varepsilon_2}^{+\infty}u\cdot(1+V_{a_i}\cdot t)\cdot(1+V_{x_i}\cdot u)^m\cdot e^{-\frac{u^2-2\cdot u\cdot t\cdot\rho_{x_ia_i}+t^2}{2(1-\rho_{x_ia_i}^2)}}\, dudt \qquad (3.68)$$

$$E\{g_3(r_{x_i}, r_{a_i})\}/(\mu_{a_i}^2\cdot\mu_{x_i}^m) = \frac{V_{a_i}}{2\pi(1-\rho_{x_ia_i}^2)^{1/2}}\cdot\int_{-\varepsilon_1}^{+\infty}\int_{-\varepsilon_2}^{+\infty}t\cdot(1+V_{a_i}\cdot t)\cdot(1+V_{x_i}\cdot u)^m\cdot e^{-\frac{u^2-2\cdot u\cdot t\cdot\rho_{x_ia_i}+t^2}{2(1-\rho_{x_ia_i}^2)}}\, dudt \qquad (3.69)$$

In the above set of equations, the lower limits of integration, usually taken as "$-\infty$" for truly normally distributed random variables, have been set equal to $-\varepsilon_1$, $-\varepsilon_2$ where ε_1, ε_2 are suitably chosen positive parameters.

With the perspective of using the results of this chapter in the statistical linearization of the nonlinear functions involved in the differential equations describing a flood routing model, it is essential to retain the physical constraints: $x_i \geq 0$ and $a_i \geq 0$. Hence, the random variables x_i, a_i cannot be distributed normally. However, to retain the attractive properties of the normally distributed random variables, assume a truncated or mixed distribution, Gaussian like, based on the assumption of small coefficients of variation for the random variables under study.

Consider, for example, the one-dimensional problem; analogous arguments can be made for the two-dimensional problem. A random variable u is normally distributed a priori with zero mean and unit variance. The constraint $1 + V_{x_i} \cdot u \geq 0$ is imposed, thus restricting u to regions: $u \geq -\frac{1}{V_{x_i}}$; $V_{x_i} > 0$. It is desired to reshape the Gaussian probability density to account for the imposed constraint but, at the same time, to retain the properties of the a priori density (i.e., odd moments equal to zero, the a priori mean and variance) necessary for the use of the Taylor-Gauss methodology. Based on the assumption of small V_{x_i}, the compound distribution of Figure 3.3 can be used, where the area of probability less than $-\frac{1}{V_{x_i}}$ is concentrated on $(-\frac{1}{V_{x_i}})$. This type of a posteriori probability density is also consistent with the physical

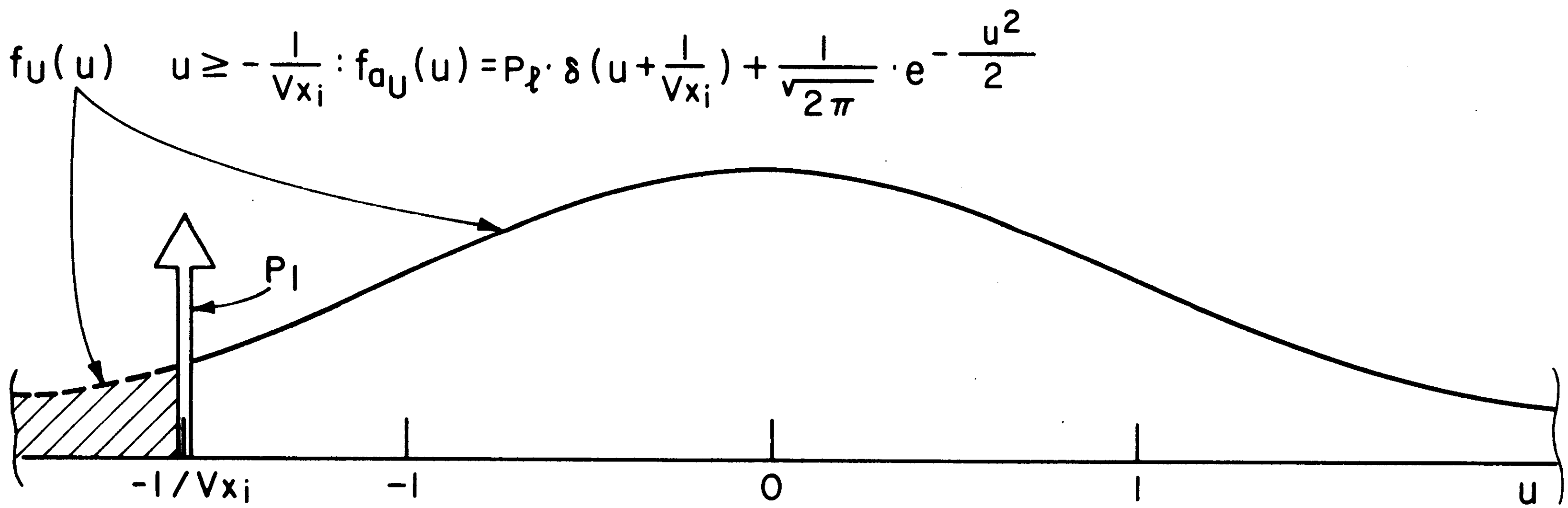

Figure 3.3

GAUSSIAN A PRIORI DENSITY $f_U(u)$ AND COMPOUND A POSTERIORI DENSITY $f_{a_U}(u)$

$\delta(\cdot)$ REPRESENTS THE DIRAC DELTA FUNCTION reality of ephemeral streams, where this situation usually arises ($x_i = 0$).

Using the notation of Figure 3.3, one can determine the a posteriori statistics of the constrained random variable as follows:

$$P_\ell = \int_{-\infty}^{-1/V_{x_i}} f_U(u)\,du = \int_{1/V_{x_i}}^{+\infty} f_U(u)\,du \tag{3.70}$$

$$E_a\{u\} = P_\ell \cdot \left(-\frac{1}{V_{x_i}}\right) + \int_{-1/V_{x_i}}^{+\infty} u \cdot f_U(u)\,du$$

consequently,

$$E_a\{u\} = \int_{1/V_{x_i}}^{+\infty} u \cdot f(u)\,du - \frac{P_\ell}{V_{x_i}} \tag{3.71}$$

with f(u) denoting the a priori probability density of the random variable u (Gaussian). By means of Eq. (3.71), it is concluded that $E_a\{u\}$ is a monotonically increasing function of V_{x_i}. Thus, e.g., the bias associated with the a posteriori mean $E_a\{u\}$ is found to be 0.083 for $V_{x_i} = 1$. Denote by σ_a^2 the a posteriori variance of u. Then,

$$\sigma_a^2 = E_a\{(u - E_a\{u\})^2\} \tag{3.72}$$

or

$$\sigma_a^2 = E_a\{u^2 + E_a^2\{u\} - 2 \cdot E_a\{u\} \cdot u\}$$

consequently,

$$\sigma_a^2 = E_a\{u^2\} - E_a^2\{u\} \tag{3.73}$$

where

$$E_a\{u^2\} = \left(-\frac{1}{V_{x_i}}\right)^2 \cdot P_\ell + \int_{-1/V_{x_i}}^{+\infty} u^2 \cdot f_U(u)\, du \tag{3.74}$$

which results in

$$E_a\{u^2\} = \frac{P_\ell}{V_{x_i}^2} + \int_{-\infty}^{+\infty} u^2 \cdot f_U(u)\, du - \int_{-\infty}^{-1/V_{x_i}} u^2 \quad f_U(u)\, du$$

or

$$E_a\{u^2\} = \frac{P_\ell}{V_{x_i}^2} + 1 - \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{V_{x_i}} \cdot e^{-\frac{1}{2V_{x_i}^2}} - P_\ell \tag{3.75}$$

rearranging terms in Eq. (3.75) yields:

$$E_a\{u^2\} = P_\ell \cdot \left[\frac{1}{V_{x_i}^2} - 1\right] + 1 - \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{V_{x_i}} \cdot e^{-\frac{1}{2V_{x_i}^2}} \tag{3.76}$$

Substitution of Equation (3.76) into Equation (3.73) yields

$$\sigma_a^2 = P_\ell \cdot \left[\frac{1}{V_{x_i}^2} - 1\right] + 1 - \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{V_{x_i}} \cdot e^{-\frac{1}{2V_{x_i}^2}}$$

$$-\left[\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2V_{x_i}^2}} - \frac{P_\ell}{V_{x_i}}\right]^2 \tag{3.77}$$

since from Eq. (3.71),

$$E_a\{u\} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2V_{x_i}^2} - \frac{P_\ell}{V_{x_i}}} \qquad (3.78)$$

If we denote by E_1 and E_2 the absolute percent errors, relative to the standard deviation, in the mean and the standard deviation, respectively, when approximating a Gaussian distribution by a mixed one, then,

$$\frac{1}{100} \times E_1 = \left|\frac{E_a\{u\} - 0}{1}\right| = |E_a\{u\}| \qquad (3.79)$$

and

$$\frac{1}{100} \times E_2 = \left|\frac{\sigma_a - 1}{1}\right| = |\sigma_a - 1| \qquad (3.80)$$

Figure 3.4 displays the error indices E_1, E_2 as functions of the coefficient of variation V_{x_i}. Note that the error in the mean is about half the error in the standard deviation and that the maximum error in the standard deviation (for $V_{x_i} = 1.0$) is about 10%. Thus, the biases due to the reshaping of the a priori Gaussian distribution to an a posteriori compound one are not large. Consequently, the a posteriori distribution $f_{a_U}(u)$ of Figure 3.3 will be used for the numerical integration of Eq. (3.65). In addition, it is evident that due to the functional form of the integrand in Eqs. (3.67) through (3.69), the discrete part of the a posteriori distribution has zero contribution in the calculation of the mean of the nonlinear functions. The right-hand side of Eqs. (3.66) through (3.69) with $\varepsilon_1 = +\frac{1}{V_{a_i}}$ and $\varepsilon_2 = +\frac{1}{V_{x_i}}$ is the contribution of

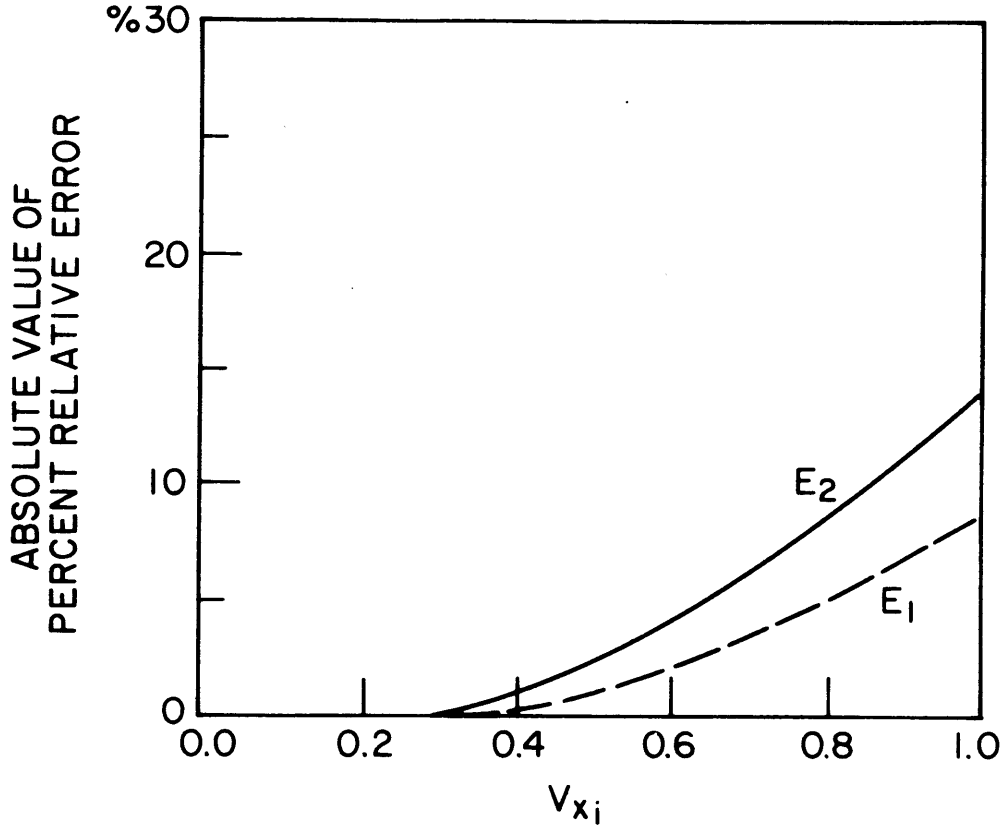

Figure 3.4

ERRORS E_1, E_2 IN THE MEAN AND THE VARIANCE DUE TO THE APPROXIMATION OF A GAUSSIAN DISTRIBUTION BY A COMPOUND ONE, AS A FUNCTION OF V_{x_i}

continuous part of the compound distribution.

For the purposes of the comparison to numerical integration, up to fourth order terms were used in the Taylor-Gauss approximation. Call I_1, I_2, I_3 the results of the numerical integration of the functions $E\{g_1(r_{x_i}, r_{a_i})\}$, $E\{g_2(r_{x_i}, r_{a_i})\}$, $E\{g_3(r_{x_i}, r_{a_i})\}$ (Equations (3.45), (3.48) and (3.49), respectively and T_1, T_2, T_3, the results of the Taylor-Gauss approximation for the same functions, then define

$$p_1 = \left(\frac{T_1 - I_1}{I_1}\right) \times 100 \tag{3.81}$$

$$p_2 = \left(\frac{T_2 - I_2}{I_2}\right) \times 100 \tag{3.82}$$

and

$$p_3 = \left(\frac{T_3 - I_3}{I_3}\right) \times 100 \tag{3.83}$$

the percentage errors of approximation. The parameters in these errors are m, V_{a_i}, V_{x_i} and $\rho_{x_i a_i}$. Tables 3.1 to 3.3 present values of p_1, p_2, p_3 for m equal to 1.2 and V_{x_i}, V_{a_i} in the range (0.1, 1.0). Similar results are given in Tables 3.4 to 3.6 for m equal to 0.8. Tables 3.1 and 3.4 correspond to $\rho_{x_i a_i}$ equal to 0.2, Tables 3.2 and 3.5 to $\rho_{x_i a_i}$ equal to 0.4 and Tables 3.3 and 3.6 to $\rho_{x_i a_i}$ equal to 0.6. If $V_{a_i} = 0$ (it follows that $\rho_{x_i a_i}$ is also equal to zero), the results correspond to the functions $E\{(\mu_{x_i} + r_{x_i})^m$ and $E\{r_{x_i} \cdot (\mu_{x_i} + r_{x_i})^m\}$ treated in Section 3.3.1. The absolute percentage errors $|p_1|, |p_2|$ for this case are displayed in Figures 3.5 (m = 1.2) and 3.6 (m = 0.8).

Based on these results, the following comments can be made for the one-dimensional case:

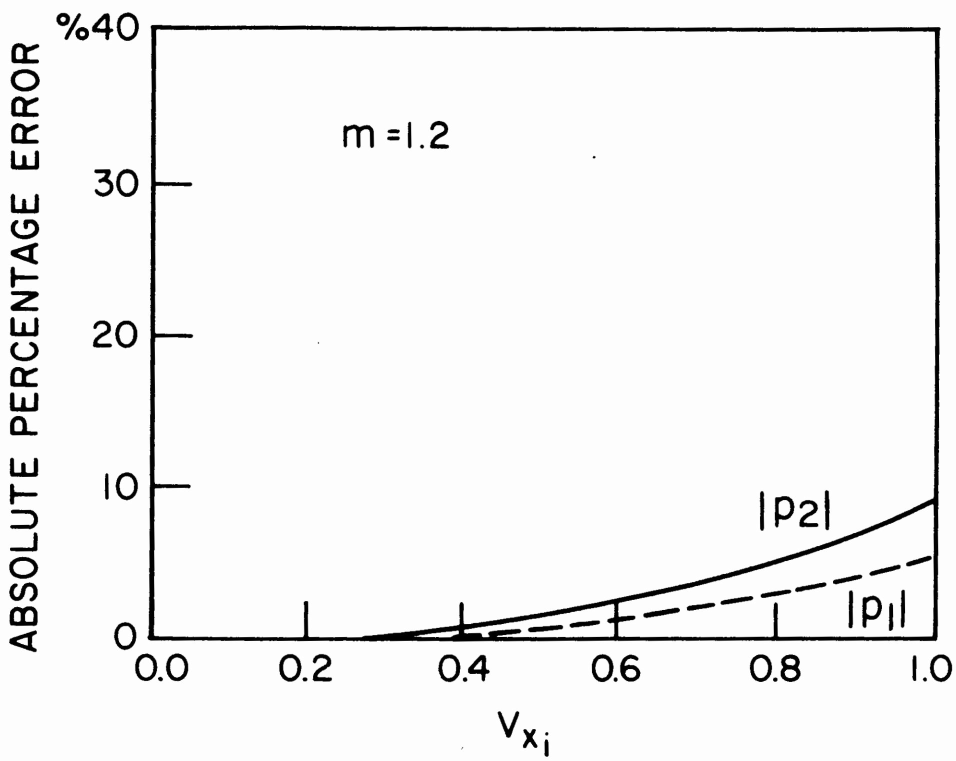

Figure 3.5

ERRORS $|p_1|$, $|p_2|$ IN THE TAYLOR-GAUSS APPROXIMATION VS. V_{x_i} FOR m = 1.2

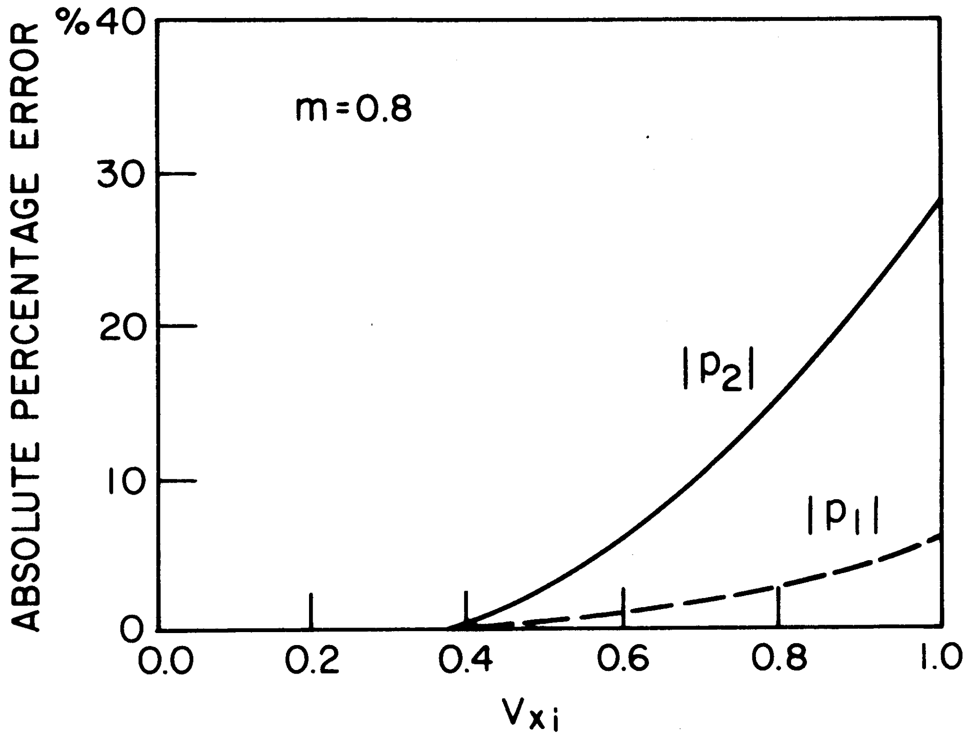

Figure 3.6

ERRORS $|p_1|$, $|p_2|$ IN THE TAYLOR-GAUSS APPROXIMATION

VS. V_{x_i} FOR $m = 0.8$

a. The absolute percentage errors $|p_1|$ and $|p_2|$ are increasing functions of the coefficient of variation, V_{x_i}.

b. These errors are greater for m = 0.8 than for m = 1.2, for all the range of values for V_{x_i}. In particular, the absolute percentage error in the approximation of the function $E\{r_{x_i} \cdot (\mu_{x_i} + r_{x_i})^m\}$, is very sensitive to m.

c. The error in the approximation of the function $E\{(\mu_{x_i} + r_{x_i})^m\}$ is always less than the error in the approximation of the function $E\{r_{x_i} \cdot (\mu_{x_i} + r_{x_i})^m\}$ for all m (m = 0.8, m = 1.2) and all V_{x_i} $(0.1 \leq V_{x_i} \leq 1.0)$. The maximum error in the approximation of $E\{r_{x_i} \cdot (\mu_{x_i} + r_{x_i})^m\}$ (for $V_{x_i} = 1.0$) is about 30%, the error corresponding to the same function for $V_{x_i} = 0.6$, being less than 10%.

d. In general, the Taylor-Gauss approximation gives satisfactory results for the one-dimensional case, being exact for values of V_{x_i} less than 0.3.

For the two-dimensional case, one can state that for constant m the error indeces $|p_1|$, $|p_2|$ and $|p_3|$ in general tend to decrease as the cross-correlation $\rho_{x_i a_i}$ increases. $|p_1|$ attains its maximum value 9.85% for m = 0.8, $\rho_{x_i a_i} = 0.2$ and $V_{x_i} = V_{a_i} = 1.0$, while the maxima 28% and 20.6% of the errors $|p_2|$ and $|p_3|$, respectively, are obtained for m = 0.8, $\rho_{x_i a_i} = 0.2$, $V_{x_i} = 1.0$, $V_{a_i} = 0.1$ and m = 0.8, $\rho_{x_i a_i} = 0.6$, $V_{x_i} = 1.0$, $V_{a_i} = 0.1$. Note that the maximum errors in the two-dimensional case are bounded by those of the one-dimensional one.

Table 3.1

PERCENTAGE ERRORS p_1, p_2 AND p_3 (%)

$m = 1.2$; $\rho_{x_i a_i} = 0.2$

V_{x_i}	V_{a_i}	p_1	p_2	p_3
0.1	0.1	-0.3×10^{-3}	0.1×10^{-2}	0.5×10^{-3}
0.1	0.5	-0.40	0.76	2.09
0.1	1.0	-7.30	6.80	17.30
0.5	0.1	-0.58	1.82	0.71
0.5	0.5	-0.67	1.08	1.09
0.5	1.0	-5.82	-1.64	11.88
1.0	0.1	-5.68	7.47	2.69
1.0	0.5	-4.42	4.87	-1.87
1.0	1.0	-7.25	1.09	5.10

Table 3.2

PERCENTAGE ERRORS p_1, p_2 AND p_3 (%)

$m = 1.2$; $\rho_{x_i a_i} = 0.4$

V_{x_i}	V_{a_i}	p_1	p_2	p_3
0.1	0.1	-0.68×10^{-4}	0.11×10^{-2}	0.68×10^{-3}
0.1	0.5	-0.36	1.13	1.85
0.1	1.0	-6.80	10.34	15.75
0.5	0.1	-0.53	1.64	1.05
0.5	0.5	-0.34	0.59	0.56
0.5	1.0	-3.75	-0.01	7.34
1.0	0.1	-5.29	6.79	4.38
1.0	0.5	-2.70	2.57	-1.23
1.0	1.0	-2.70	-1.72	0.71

Table 3.3

PERCENTAGE ERRORS p_1, p_2 AND p_3 (%)

$m = 1.2$; $\rho_{x_i a_i} = 0.6$

V_{x_i}	V_{a_i}	p_1	p_2	p_3
0.1	0.1	0.98×10^{-4}	0.89×10^{-3}	0.71×10^{-3}
0.1	0.5	-0.33	1.26	1.64
0.1	1.0	-6.35	11.59	14.40
0.5	0.1	-0.48	1.46	1.16
0.5	0.5	-0.04	0.12	0.10
0.5	1.0	-2.04	0.32	3.87
1.0	0.1	-4.92	6.17	4.91
1.0	0.5	-1.24	0.70	-1.23
1.0	1.0	0.69	-3.18	-2.35

Table 3.4

PERCENTAGE ERRORS p_1, p_2 AND p_3 (%)

$m = 0.8$; $\rho_{x_i a_i} = 0.2$

V_{x_i}	V_{a_i}	p_1	p_2	p_3
0.1	0.1	0.32×10^{-3}	-0.46×10^{-3}	0.1×10^{-2}
0.1	0.5	-0.41	1.04	2.17
0.1	1.0	-7.43	9.43	17.75
0.5	0.1	-0.01	1.68	0.79
0.5	0.5	-0.37	1.19	1.75
0.5	1.0	6.22	-0.96	13.90
1.0	0.1	-7.44	28.00	13.70
1.0	0.5	-6.47	21.35	0.79
1.0	1.0	-9.85	11.75	6.69

Table 3.5

PERCENTAGE ERRORS p_1, p_2 AND p_3 (%)

$m = 0.8$; $\rho_{x_i a_i} = 0.4$

V_{x_i}	V_{a_i}	p_1	p_2	p_3
0.1	0.1	0.1×10^{-3}	0.3×10^{-3}	0.1×10^{-2}
0.1	0.5	-0.39	1.41	2.00
0.1	1.0	-7.11	12.50	16.70
0.5	0.1	-0.02	1.54	1.05
0.5	0.5	-0.31	0.92	1.22
0.5	1.0	-4.90	1.65	9.70
1.0	0.1	-7.13	26.25	18.75
1.0	0.5	-5.00	15.15	3.49
1.0	1.0	-5.78	5.47	3.25

Table 3.6

PERCENTAGE ERRORS p_1, p_2 AND p_3 (%)

$m = 0.8$; $\rho_{x_i a_i} = 0.6$

V_{x_i}	V_{a_i}	p_1	p_2	p_3
0.1	0.1	-0.1×10^{-3}	0.8×10^{-3}	0.1×10^{-2}
0.1	0.5	-0.37	1.53	1.85
0.1	1.0	-6.79	13.55	15.75
0.5	0.1	-0.3×10^{-1}	1.41	1.15
0.5	0.5	-0.25	0.60	0.74
0.5	1.0	-3.61	2.29	6.51
1.0	0.1	-6.82	24.55	20.60
1.0	0.5	-3.73	10.21	4.26
1.0	1.0	-2.53	0.63	-0.76

Chapter 4

STATISTICAL LINEARIZATION

4.1 Introduction

It is a well-known fact that the study of nonlinear systems is a much more difficult task than the study of linear ones. The superposition principle does not hold in nonlinear systems, thus making the analysis input and initial condition specific. It is not possible to generalize from the response to a generic input form (e.g., unit impulse function), to the response to other classes of input. Even within a certain class of input, the response of the nonlinear system will depend on the characteristics (e.g., magnitude) of the individual inputs.

Since powerful results have been obtained in estimation (and control) theory for stochastic linear system models, it is often desirable to convert the stochastic nonlinear model of the physical phenomenon to an equivalent linear one, where the word "equivalent" is defined according to a prespecified criterion. Common approximations involve: the use of linear models of high dimensionality as a piece-wise linear approximation to nonlinear models; a Taylor series expansion of the nonlinear functions about reference trajectories, keeping the leading linear term in the expansion, and use of models resulting from the statistical linearization of the nonlinear functions involved.

For the purposes of this work, statistical linearization seems to be the most promising technique since:

1. it provides a model that has all the advantages of being linear,
2. it gives an unbiased estimate for the mean of the nonlinear function involved,
3. it is especially convenient when the nonlinear function to be linearized is not differentiable, thus excluding a Taylor series expansion,
4. it does not increase the dimensionality of the nonlinear model,
5. it provides linear systems with weighting functions dependent on the form and magnitude of the input, thus retaining the essential properties of the nonlinear model, and
6. its use in streamflow forecasting has given satisfactory results (Kitanidis and Bras, 1978).

This chapter deals with the statistical linearization technique. The next section outlines the theory as it has been developed for the case of stationary inputs. The last two sections present extensions of the technique to account for nonstationarity of the input as well as applications to the flood routing model based on a cascade of reservoirs.

4.2 Statistical Linearization for Stationary Processes

To establish notation, some of the basic concepts of linear theory are reviewed first.

Denote the response of a linear system (in general time varying) to the unit impulse $\delta(t_o)$, by $W(t, t_o)$, with $\delta(t_o)$ representing the Dirac delta function. Then for every input x( ), it holds

$$y(t) = \int_{-\infty}^{t} x(\tau) \cdot W(t, \tau)\, d\tau \qquad (4.1)$$

where y(t) represents the output at time t. If the system is time-invariant, Eq. (4.1) reduces to,

$$y(t) = \int_{0}^{+\infty} x(t - \tau) \cdot W(\tau)\, d\tau \qquad (4.2)$$

When x(t) can be considered to be a random process (e.g., due to the noise level inherent in x(t)), then it is readily shown that,

$$E\{y(t)\} = m_y = \int_{0}^{+\infty} E\{x(t - \tau)\} \cdot W(\tau)\, d\tau \qquad (4.3)$$

where E(.) denotes expected value.

If x(t) is a wide sense stationary process, then

$$m_y = m_x \cdot \int_{0}^{+\infty} W(\tau)\, d\tau \qquad (4.4)$$

where m_y and m_x stand for the means of the processes y(t) and x(t), respectively.

Also,

$$E\{(y(t) - m_y)(y(t+\tau) - m_y)\}$$
$$= \int_{0}^{+\infty} W(\tau_1) \cdot \int_{0}^{+\infty} W(\tau_2) \cdot E\{(x(t-\tau_1) - m_x)\cdot(x(t+\tau-\tau_2) - m_x)\} \cdot d\tau_2 \cdot d\tau_1 \qquad (4.5)$$

and

$$E\{(x(t) - m_x)(y(t+\tau) - m_y)\} =$$

$$= \int_0^{+\infty} W(\tau_1) \cdot E\{(x(t) - m_x)(x(t+\tau-\tau_1) - m_x)\} \, d\tau_1 \qquad (4.6)$$

Denoting

$$\sigma^2_{yy}(\tau) = E\{(y(t) - m_y) \cdot (y(t+\tau) - m_y)\} \qquad (4.7)$$

$$\sigma^2_{xx}(\tau) = E\{(x(t) - m_x) \cdot (x(t+\tau) - m_x)\} \qquad (4.8)$$

and

$$\sigma^2_{xy}(\tau) = E\{(x(t) - m_x) \cdot (y(t+\tau) - m_y)\} \qquad (4.9)$$

Equations (4.5) and (4.6) reduce to

$$\sigma^2_{yy}(\tau) = \int_0^{+\infty} W(\tau_1) \cdot \int_0^{+\infty} W(\tau_2) \cdot \sigma^2_{xx}(\tau + \tau_1 - \tau_2) \cdot d\tau_1 \cdot d\tau_2 \qquad (4.10)$$

and

$$\sigma^2_{xy}(\tau) = \int_0^{+\infty} W(\tau_1) \cdot \sigma^2_{xx}(\tau - \tau_1) \cdot d\tau_1 \qquad (4.11)$$

Equations (4.4), (4.10) and (4.11) give expressions for the mean and autocovariance functions of the output process as well as the input-output cross-covariance function, given the mean and autocovariance functions of the input, for a linear time-invariant system and wide sense stationary inputs.

Based on the Riemann definition of the integral, Eq. (4.2) can be approximated by a discrete summation. Since the sum of Gaussian random variables is also Gaussian, if x(t) is a Gaussian process, then y(t) should be also Gaussian. It is the unique property of the class of linear systems to permit the Gaussian input form to be retained in the output too. Furthermore, if one uses the central limit theorem for the case of uncorrelated (in time) input processes (e.g., white sequences), Eq. (4.2) indicates that the output process y(t) will tend to a Gaussian process, irrespectively of the distribution of $x(t_i)$, $\forall\, t_i$. It is also true, even though not yet theoretically proven, that, in general, non-Gaussian inputs (even non-white) tend to become more nearly Gaussian as a result of linear filtering (Gelb and Vander Velde, 1968).

All the properties mentioned hold for the class of the linear systems. When the system under study does not satisfy the superposition principle, then it is evident that one has to study every input form (and initial condition) separately. To avoid the construction of large catalogs of input-output pairs for a nonlinear system, several approximations can be proposed. Surely, the most common one is to linearize the nonlinear functions of the system at hand using a Taylor series expansion, retaining up to and including the first order term.

The major problem of this approach is that it requires differentiability of the nonlinear functions involved. Thus, if the function is discontinuous, this methodology cannot be used. In addition, it is justified for small deviations of the independent variables from the nominal trajectory, and it fails when the higher order terms in the Taylor series expansion are of significant magnitude. In particular, when the inputs to the system are random processes, such a truncation introduces biases.

Another approach is quasi-linearization. Quasi-linearization is the operation on the nonlinear function that results in linear approximations dependent on some properties of the input. In particular, when the input to the system is a random process, the operation is called statistical linearization. Gelb and Vander Velde (1968) and Graham and McRuer (1961) present comprehensive treatments of quasi-linearization. These are the major references in the development to follow.

The fundamental idea of quasi-linearization is to construct linear time-invariant operators, called describing functions, that will preserve certain response characteristics of the nonlinear functions, for some generic input forms.

As a result of linear filtering, general input forms reduce to the form of biases, sinusoids and Gaussian random processes. If the system configuration is a feedback one, the application of the quasi-linearization technique requires the presence of a linear operator before the nonlinear function, so that the signal to be "fed-back" (to the nonlinear function) is filtered to approximate generic input forms.

For the purposes of this development, only biases and Gaussian random processes will be utilized.

Having defined the generic input forms, one has to specify what input-output transfer characteristics to preserve in the quasilinear approximation. In quantitative terms, if we denote by $y_a(t)$ the resulting linear approximation of the function $y(t)$, possible choices for a criterion function to be minimized might be,

$$E\{(y(t) - y_a(t))^2\}; \quad E\{|y(t) - y_a(t)|\}; \; E\{[y(t) - E\{y(t)\}]^2\}$$

$$-E\{[y_a(t) - E\{y_a(t)\}]^2\}; E\{y(t)\} - E\{y_a(t)\} \; ; \; E\{x(t) \cdot y(t)\}$$

$$- E\{x(t) \cdot y_a(t)\} \; ; \quad E\{(y(t) - E\{y(t)\}) \cdot (y(t+\tau)$$

$$- E\{y(t+\tau)\})\} - E\{(y_a(t) - E\{y_a(t)\})(y_a(t+\tau) - E\{y_a(t+\tau)\})$$

Usually a minimum mean squared error criterion is used. The major reasons are that the derivatives of the criterion function will result in a system of linear equations, which is easily solved, and that minimization of this objective function implies optimization with respect to some of the other criteria presented above as well.

Denote by $e(t)$ the error of approximation,

$$e(t) = y_a(t) - y(t) \tag{4.12}$$

Then,

$$E\{e^2(t)\} = E\{y_a^2(t)\} + E\{y^2(t)\} - 2E\{y_a(t)\; y(t)\} \tag{4.13}$$

The objective is to minimize the mean squared error with respect to the set of functions: $w_i(t)$, $i = 1, 2, \ldots, n$ where

$$y_a(t) = \sum_{i=1}^{n} \cdot \int_{0}^{+\infty} w_i(\tau) \cdot x_i(t - \tau)\, d\tau \tag{4.14}$$

with $x_i(t)$, $i = 1, 2, \ldots, n$, denoting the generic input forms that add up to the total input in the nonlinearity. Clearly, the problem fits within the calculus of variations framework. Its solution (Graham and McRuer (1961)) for broad sense stationary inputs results in the well known in estimation theory Wiener-Hopf integral equation,

$$\sum_{j=1}^{n} \int_{o}^{+\infty} w_j(\tau_2) \cdot \sigma^2_{x_i x_j}(\tau_1 - \tau_2)\, d\tau_2 = E\{y(t) \cdot x_i(t - \tau)\};$$

$$\tau_1 \geq 0,\ i = 1, 2, \ldots, n \qquad (4.15)$$

The left-hand side of Eq. (4.15) is the cross-correlation of the i^{th} input process with the output of the approximation $y_a(t)$. Thus, the optimal set of weighting functions is the set that equates the input-output cross-correlation corresponding to the nonlinear function with the input-output cross-correlation corresponding to the linear one. This property is also true for the Wiener filter.

Several properties of the approximation are given in the following:

1. The error of approximation is uncorrelated with the input to the nonlinear function.
2. The linear approximation results in a mean-squared output always less or equal to the mean-squared output of the nonlinearity.
3. When the inputs are statistically uncorrelated, Eq. (4.15) results in,

$$\int_{0}^{+\infty} w_i(\tau_2) \cdot \sigma^2_{x_i x_i}(\tau_1 - \tau_2) \cdot d\tau_2 = E\{y(t) \cdot x_i(t - \tau)\} ;$$

$$\tau_1 \geq 0 , \; i = 1, 2, \ldots, n \quad (4.16)$$

A special case of interest in this work arises when the input processes can be decomposed in a bias term x_1 plus a zero mean Gaussian random variable $x_2(t)$,

$$x(t) = x_1 + x_2(t) \quad (4.17)$$

Since x_1 and $x_2(t)$ are uncorrelated, Eq. (4.16) can be used. Consequently,

$$\int_{0}^{+\infty} w_1(\tau_2) \cdot x_1^2 \, d\tau_2 = x_1 \cdot E\{y(t)\} \quad (4.18)$$

and

$$\int_{0}^{+\infty} w_2(\tau_2) \cdot \sigma^2_{x_2 x_2}(\tau_1 - \tau_2) \, d\tau_2 = E\{y(t) \cdot x_2(t - \tau)\} ; \; \tau_1 \geq 0 \quad (4.19)$$

Equation (4.18) is particularly meaningful when the expectation $E\{y(t)\}$ is equal to a constant. This implies that the nonlinearity $y(t)$ is dependent on time only through $x(t)$ (time invariant or static nonlinearity) and that the Gaussian process is at least wide sense stationary. Under the circumstances, Equation (4.18) can be written as,

$$N_{x_1} = \int_{0}^{+\infty} w_1(\tau_2) \cdot d\tau_2 = \frac{E\{y(t)\}}{x_1} \quad (4.20)$$

And the component $y_{a_1}(t)$ of the output approximating the nonlinearity will be (Equation (4.14)),

$$y_{a_1}(t) = x_1 \cdot N_{x_1} \tag{4.21}$$

with N_{x_1} given by (4.20).

Equation (4.21) states that, under the conditions specified, the time variation of the optimal weighting function $w_1(\tau)$ is not important.

Note that,

$$E\{y_a(t)\} = E\{N_{x_1} \cdot x_1 + \int_0^{+\infty} w_2(\tau) \cdot x_2(t - \tau)\, d\tau\} \tag{4.22}$$

Using Eqs. (4.3) and (4.20) and assuming that $x_2(t)$ is a zero mean wide sense stationary process results in,

$$E\{y_a(t)\} = N_{x_1} \cdot x_1 = E\{y(t)\} \tag{4.23}$$

Thus, the equivalent linear approximation is unbiased.

When the input process is not stationary, then one cannot, in general, minimize Eqs. (4.13) to arrive at equations similar to the Wiener-Hopf expression (4.15), since, in this case, the covariances involved are functions of two arguments.

The difficulties of the nonstationary case can be avoided using an approximation of the type,

$$y_a(t) = \sum_{i=1}^{n} N_{x_i}(t) \cdot x_i(t) \tag{4.24}$$

instead of the more general expression given in Eq. (4.14). Since the purpose of quasi-linearization is to simplify a nonlinear problem, and, in view of the fact that the expression in Eq. (4.14) is not itself a general linear approximation (a wider class of approximations would involve time-varying linear gains $w_i(t, \tau)$), the expression in Eq. (4.24) will be employed to approximate a nonlinear function when its input is nonstationary.

4.3 Statistical Linearization for Nonstationary Processes

It is of interest to minimize the mean quadratic error of approximation given in Eq. (4.13) when the approximation $y_a(t)$ is of the type in Eq. (4.24). In the following, the dependence of the time functions on t is not explicitly shown. Proceeding formally by taking the derivations of Eq. (4.13) with respect to the gains N_{x_i}, i = 1, 2, ..., n, and equating the result to zero gives,

$$\sum_{j=1}^{n} N_{x_j} \cdot E\{x_i\, x_j\} = E\{x_i\, y\} \quad ; \quad i = 1, 2, \ldots, n \qquad (4.25)$$

Solution of the set of coupled linear equations gives:

$$\underline{N}_x = \underline{\bar{P}}^{-1} \cdot \underline{\beta} \qquad (4.26)$$

provided that $\underline{\bar{P}}^{-1}$ exists.

$\underline{N}_x$ is the n-dimensional vector whose i^{th} element is N_{x_i}, $\underline{\bar{P}}^{-1}$ is an n by n positive definite matrix whose $(i, j)^{th}$ element is $E\{x_i\, x_j\}$ and $\underline{\beta}$ is the n-dimensional vector whose i^{th} element is $E\{x_i \cdot y\}$.

The condition that the determinant of $\underline{\bar{P}}$ be different from zero is exactly the condition that the set of gains in Eq. (4.25) give a minimum mean squared error, because $\overline{\underline{P}}$ is equal to the Hessian matrix of $E\{e^2\}$. If all the x_i's are zero mean processes, then $\overline{\underline{P}}$ is a covariance matrix. If x_1 is a bias input, then the (n-1) by (n-1) lower principal minor of $\overline{\underline{P}}$ is a covariance matrix. In this case, the only nonzero element of the first row and first column is the (1, 1) element, equal to x_1^2. Therefore, the condition for nonzero determinant for $\overline{\underline{P}}$ is equivalent to the (n-1) by (n-1) covariance matrix being positive definite. For a 2 by 2 covariance matrix, positive definiteness is implied by the condition that the cross-correlations between any two x_i's be less than one in absolute value.

For covariances of dimensions greater than two, this is not true and one has to examine the set of required sufficient conditions to assure existence of a solution (i.e., Equation (4.26)). To illustrate this, consider a 3 by 3 covariance matrix as follows:

$$\overline{\underline{P}} = \begin{vmatrix} \sigma_{x_1}^2 & \rho_{12}\cdot\sigma_{x_1}\cdot\sigma_{x_2} & \rho_{13}\cdot\sigma_{x_1}\cdot\sigma_{x_3} \\ \rho_{12}\cdot\sigma_{x_1}\cdot\sigma_{x_2} & \sigma_{x_2}^2 & \rho_{23}\cdot\sigma_{x_2}\cdot\sigma_{x_3} \\ \rho_{13}\cdot\sigma_{x_1}\cdot\sigma_{x_3} & \rho_{23}\cdot\sigma_{x_2}\cdot\sigma_{x_3} & \sigma_{x_3}^2 \end{vmatrix} \qquad (4.27)$$

If $\rho_{12} = \rho_{13} = \frac{1}{2}$ and $\rho_{23} = -\frac{1}{2}$, the determinant of $\overline{\underline{P}}$ is equal to zero, even though the correlations are all less than one in absolute value.

The procedure outlined in this section is used in the following for the statistical linearization of the nonlinear functions involved in the flood routing model used in this work.

4.4 Linearization of the Outflow Discharge Function of a Nonlinear Reservoir

The differential equations of motion for a flood routing model based on nonlinear reservoirs are of the type:

$$\frac{dx_i(t)}{dt} = u_i(t) + a_{i-1}(t) \cdot x_{i-1}^m(t) - a_i(t) \cdot x_i^m(t) \quad ;$$

$$i = 1, 2, \ldots, n \qquad (4.28)$$

where

- $x_i(t)$ is the i^{th} state of the system
- $u_i(t)$ is the i^{th} input to the system
- $a_i(t)$ is the (possibly) time varying parameter of the system with $a_o(t) \triangleq 0$
- m is a constant, known, exponent
- n is the number of the reservoirs used in the model or equivalently the order of the system

The state $x_i(t)$ $\forall i$ represents volume of water in storage.

Since the objective is to design algorithms to estimate $a_i(t)$ and $x_i(t)$, $\forall i$, t, in real time using the powerful results of linear filtering theory (e.g., Kalman filter), it is convenient to proceed with the statistical linearization of the function, $a_i(t) \cdot x_i^m(t)$.

The processes $a_i(t)$, $x_i(t)$ are taken as Gaussian, assuming that there is adequate linear filtering in the system (Figure 4.1) and that the inputs $u_i(t)$ are normal random processes.

The time dependence of $a_i(t)$, $x_i(t)$ will not be shown in

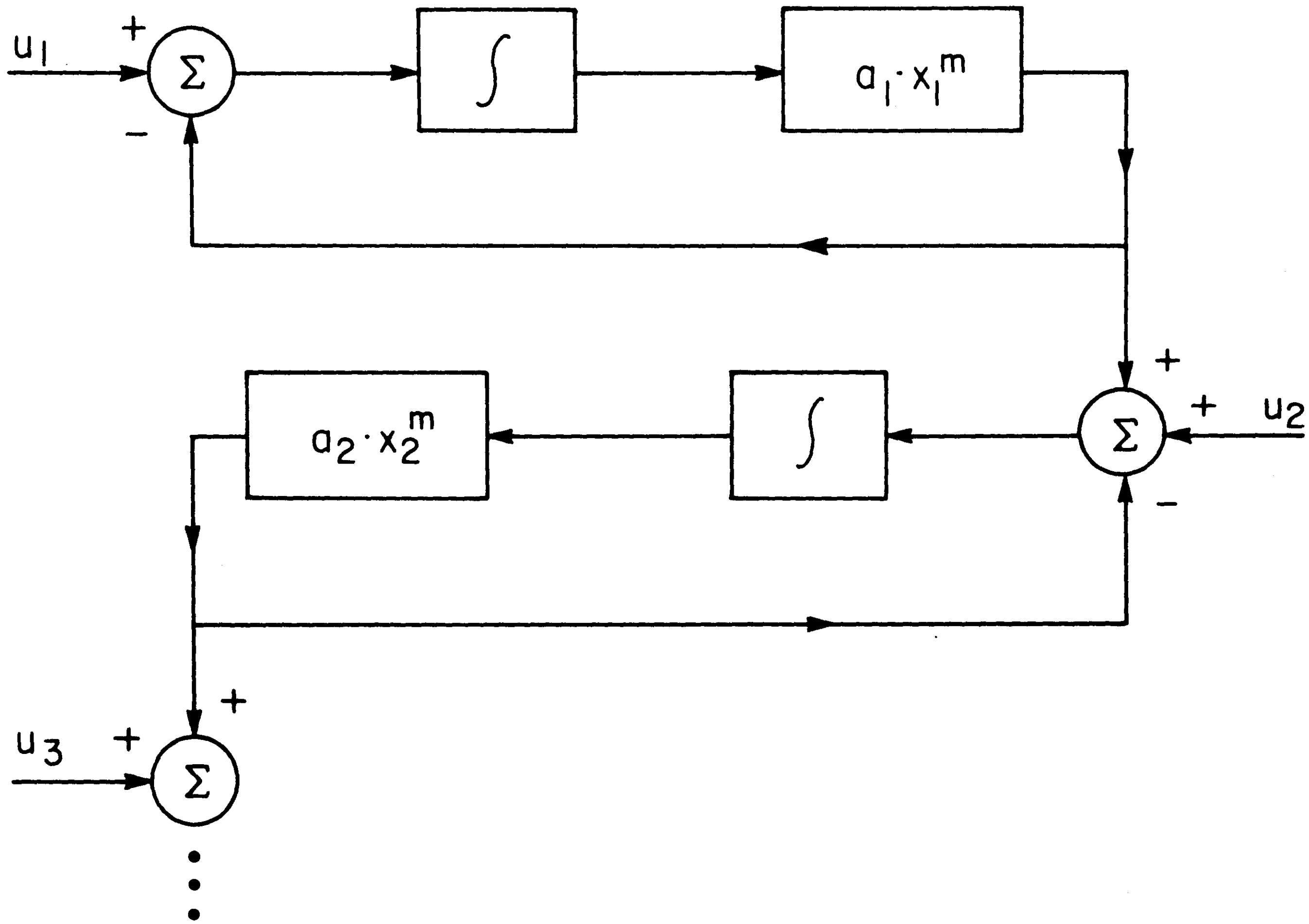

Figure 4.1

BLOCK DIAGRAM OF A CASCADE OF NONLINEAR RESERVOIRS order to facilitate notation.

Decompose a_i, x_i as:

$$a_i = \mu_{a_i} + r_{a_i} \tag{4.29}$$

$$x_i = \mu_{x_i} + r_{x_i} \tag{4.30}$$

where μ_{x_i}, μ_{a_i} are bias terms while r_{a_i}, r_{x_i} are normally distributed random variables (for a certain value of t) with zero mean, variances $\sigma^2_{a_i}$, $\sigma^2_{x_i}$ and covariance $\sigma^2_{x_i a_i}$ (= $\sigma^2_{a_i x_i}$).

Define the nonlinear function,

$$g(\mu_{x_i}, \mu_{a_i}, r_{x_i}, r_{a_i}) = (\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m \tag{4.31}$$

The goal is to find the "best" linear approximation to the function $g(\mu_{x_i}, \mu_{a_i}, r_{x_i}, r_{a_i})$. The approximation will take the form

$$g_a(\mu_{x_i}, \mu_{a_i}, r_{x_i}, r_{a_i}) = N_{b_i} \cdot (\mu_{a_i} + \mu_{x_i}) + N_{x_i} \cdot r_{x_i} + N_{a_i} \cdot r_{a_i} \tag{4.32}$$

The problem reduces in finding the set of gains N_{b_i}, N_{x_i}, N_{a_i}, such that the error of approximation,

$$e = g(\mu_{x_i}, \mu_{a_i}, r_{x_i}, r_{a_i}) - g_a(\mu_{x_i}, \mu_{a_i}, r_{x_i}, r_{a_i}) \tag{4.33}$$

has a minimum mean squared value.

The mean square error value is

$$E\{e^2\} = E\{[(\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m - N_{b_i} \cdot (\mu_{a_i} + \mu_{x_i})$$

$$- N_{x_i} \cdot r_{x_i} - N_{a_i} \cdot r_{a_i} \tag{4.34}$$

Carrying out the necessary algebraic operations, Eq. (4.34) reduces to

$$E\{e^2\} = E\{(\mu_{a_i} + r_{a_i})^2 \cdot (\mu_{x_i} + r_{x_i})^{2m}\} + N_{b_i} \cdot (\mu_{a_i} + \mu_{x_i})^2 + N_{a_i}^2 \cdot \sigma_{a_i}^2$$

$$+ N_{x_i}^2 \cdot \sigma_{x_i}^2 - 2 \cdot N_{b_i} \cdot (\mu_{a_i} + \mu_{x_i}) \cdot E\{(\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m\}$$

$$- 2 \cdot N_{a_i} \cdot E\{r_{a_i} \cdot (\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m\}$$

$$- 2 \cdot N_{x_i} \cdot E\{r_{x_i} \cdot (\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m\}$$

$$+ 2 \cdot N_{a_i} \cdot N_{x_i} \cdot \sigma_{x_i a_i}^2 \qquad (4.35)$$

Given Equation (4.35), the optimal (static) gains N_{b_i}, N_{a_i}, N_{x_i} satisfy

$$\frac{\partial}{\partial N_{b_i}} \cdot (E\{e^2\}) = 0 \qquad (4.36)$$

$$\frac{\partial}{\partial N_{a_i}} (E\{e^2\}) = 0 \qquad (4.37)$$

$$\frac{\partial}{\partial N_{x_i}} (E\{e^2\}) = 0 \qquad (4.38)$$

when the matrix

$$\underline{\bar{H}} = \{h_{ij}\} = \{\frac{\partial^2}{\partial y_i \, \partial y_j} E\{e^2\}\}; \quad i, j = 1, 2, 3 \qquad (4.39)$$

is positive definite, with y_i, y_j denoting any of the N_{b_i}, N_{a_i}, N_{x_i}.

It can be shown that,

$$\underline{\bar{H}} = \begin{vmatrix} 2\cdot(\mu_{a_i} + \mu_{x_i})^2 & 0 & 0 \\ 0 & 2\sigma^2_{x_i} & 2\sigma^2_{x_i a_i} \\ 0 & 2\sigma^2_{x_i a_i} & 2\sigma^2_{a_i} \end{vmatrix} \tag{4.40}$$

The principal minors of matrix $\underline{\bar{H}}$ are

$$D_1 = 2 \cdot (\mu_{a_i} + \mu_{x_i})^2 \tag{4.41}$$

$$D_2 = 4 \cdot (\mu_{a_i} + \mu_{x_i})^2 \cdot \sigma^2_{x_i} \tag{4.42}$$

$$D_3 = 8 \cdot (\mu_{a_i} + \mu_{x_i})^2 \cdot [\sigma^2_{x_i} \cdot \sigma^2_{a_i} - \sigma^4_{x_i a_i}] \tag{4.43}$$

It is evident that in the nontrivial cases when $\mu_{x_i} > 0$, $\mu_{a_i} > 0$, $\sigma^2_{x_i a_i} < \sigma_{x_i} \cdot \sigma_{a_i}$, the minors D_i, $i = 1, 2, 3$ are all positive. Consequently, matrix $\underline{\bar{H}}$ is a positive definite matrix and the optimal N_{b_i}, N_{a_i}, N_{x_i} are given by Eqs. (4.36), (4.37) and (4.38) implicitly.

Solution of the system of Eqs. (4.36) and (4.35) yields,

$$N_{b_i} = \frac{E\,\{(\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m\}}{(\mu_{a_i} + \mu_{x_i})} \tag{4.44}$$

$$N_{x_i} = \frac{\sigma^2_{a_i} \cdot E\{r_{x_i} \cdot (\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m\} - \sigma^2_{x_i a_i} \cdot E\{r_{a_i} \cdot (\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m\}}{\sigma^2_{x_i} \cdot \sigma^2_{a_i} - \sigma^4_{x_i a_i}} \tag{4.45}$$

$$N_{a_i} = \frac{\sigma^2_{x_i} \cdot E\{r_{a_i} \cdot (\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m\} - \sigma^2_{x_i a_i} \cdot E\{r_{x_i} \cdot (\mu_{a_i} + r_{a_i}) \cdot (\mu_{x_i} + r_{x_i})^m\}}{\sigma^2_{x_i} \cdot \sigma^2_{a_i} - \sigma^4_{x_i a_i}} \tag{4.46}$$

The above correspond to Eq. (4.26) with $\underline{\overline{P}}$ given by Eq. (4.40).

In the development presented above, the bias terms were grouped and only one gain was attributed to them. This results from the fact that the different biases are indistinguishable for the purposes of the approximation. One can verify this by using different gains for the two bias terms and then follow the procedure described above. In this case, the matrix $\underline{\overline{H}}$ is only semi-definite and the solution to the set of Eqs. (4.36) through (4.37) provides saddle points.

All the properties shown for the stationary case can be verified here also.

At this point, invoke the results of the previous chapter to derive analytical expressions for the gains using the input statistics as parameters. Substitution of the expectations involved in Eqs. (4.44), (4.45) and (4.46) by the expressions of Eqs. (3.58) through (3.63) and carrying out the algebraic operations results in

$$N_{b_i} = \frac{\mu_{a_i} \cdot \mu_{x_i}^m}{(\mu_{a_i} + \mu_{x_i})} \cdot \{1 + \frac{m \cdot (m-1)}{2} \cdot V_{x_i}^2 + m \cdot \rho_{x_i a_i} \cdot V_{x_i} \cdot V_{a_i}\} \quad (4.47)$$

$$N_{x_i} = \mu_{a_i} \cdot \mu_{x_i}^{(m-1)} \cdot \{m + \frac{m \cdot (m-1) \cdot (m-2)}{2} \cdot V_{x_i}^2 + m \cdot (m-1) \cdot \rho_{x_i a_i} \cdot V_{a_i} \cdot V_{x_i}\} \quad (4.48)$$

$$N_{a_i} = \mu_{x_i}^m \cdot \{1 + \frac{m \cdot (m-1)}{2} \cdot V_{x_i}^2\} \quad (4.49)$$

Equations (4.47) through (4.49) give the optimal gains as functions of the parameters V_{x_i}, V_{a_i}, $\rho_{x_i a_i}$.

When the parameter a_i can be adequately modeled by only a constant bias term, the expressions for the optimal gains become ($V_{a_i} = 0$; $\rho_{x_i a_i} = 0$) :

$$N_{b_i} = \frac{\mu_{a_i} \cdot \mu_{x_i}^m}{(\mu_{a_i} + \mu_{x_i})} \cdot \{1 + \frac{m \cdot (m-1)}{2} \cdot V_{x_i}^2\} \qquad (4.50)$$

$$N_{x_i} = \mu_{a_i} \cdot \mu_{x_i}^{(m-1)} \cdot \{m + \frac{m \cdot (m-1) \cdot (m-2)}{2} \cdot V_{x_i}^2\} \qquad (4.51)$$

Some of the properties of the one-dimensional case (Eqs. (4.50) and (4.51)) are described in the following:

One can easily verify that

a. for m = 1 (i.e., $g(a_i, x_i) = a_i \cdot x_i$), it follows that

$$N_{b_i} = N_{x_i} = \mu_{a_i} \qquad (4.52)$$

b. for m = 2 (i.e., $f(x) = x^2$):

$$N_{b_i} = \mu_{a_i} \cdot \mu_{x_i} \cdot (1 + V_{x_i}^2)/(\mu_{a_i} + \mu_{x_i}) \qquad (4.53)$$

$$N_{x_i} = 2 \cdot \mu_{a_i} \cdot \mu_{x_i} \qquad (4.54)$$

Equations (4.53), (4.54) give the analytical forms of the gains using the Taylor-Gauss approximation. Note that these are the exact values of the gains. Figures 4.2 and 4.3 present plots of the normalized gains (non-dimensional),

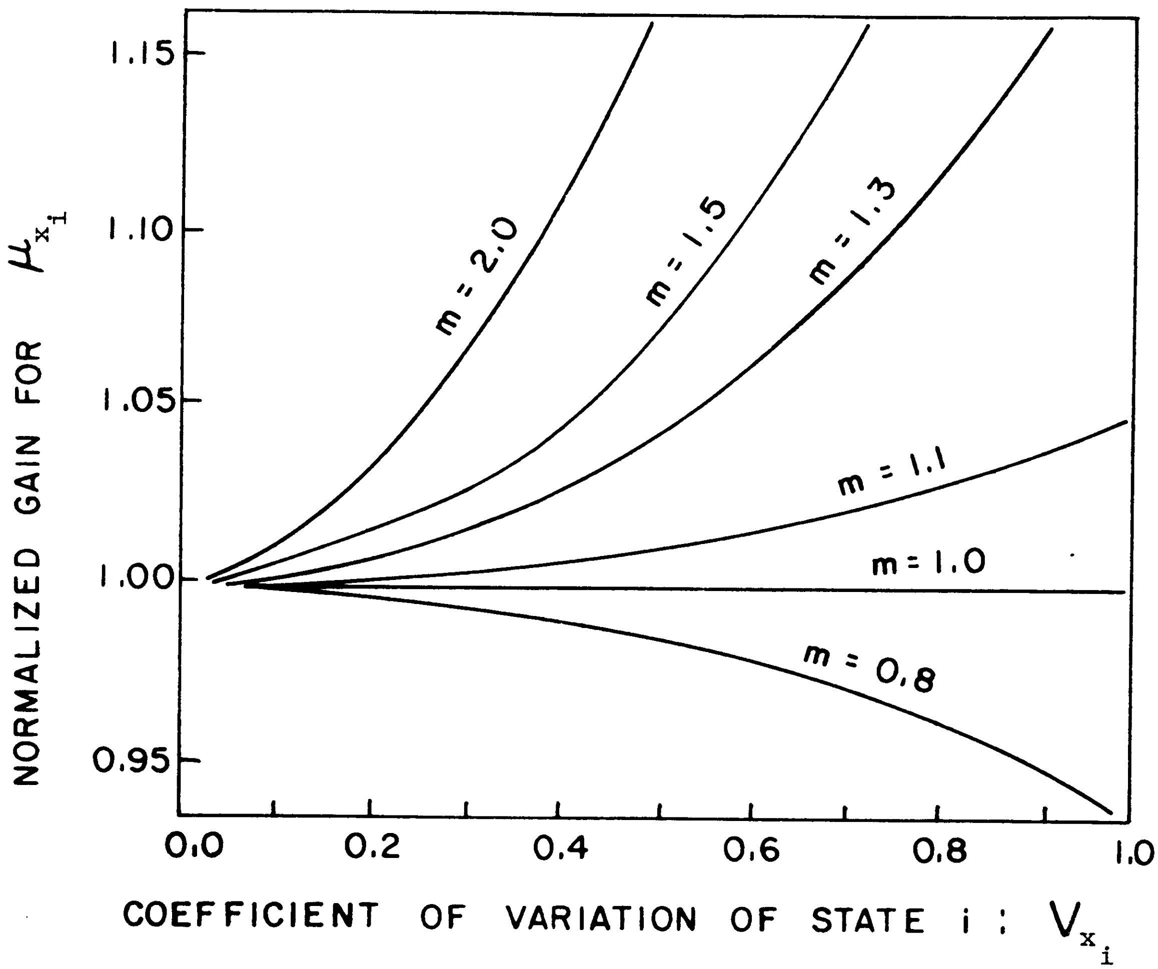

Figure 4.2

NORMALIZED GAIN FOR THE BIAS μ_{x_i} AS A FUNCTION OF THE COEFFICIENT OF VARIATION OF THE i^{th} STATE FOR DIFFERENT VALUES OF THE EXPONENT m

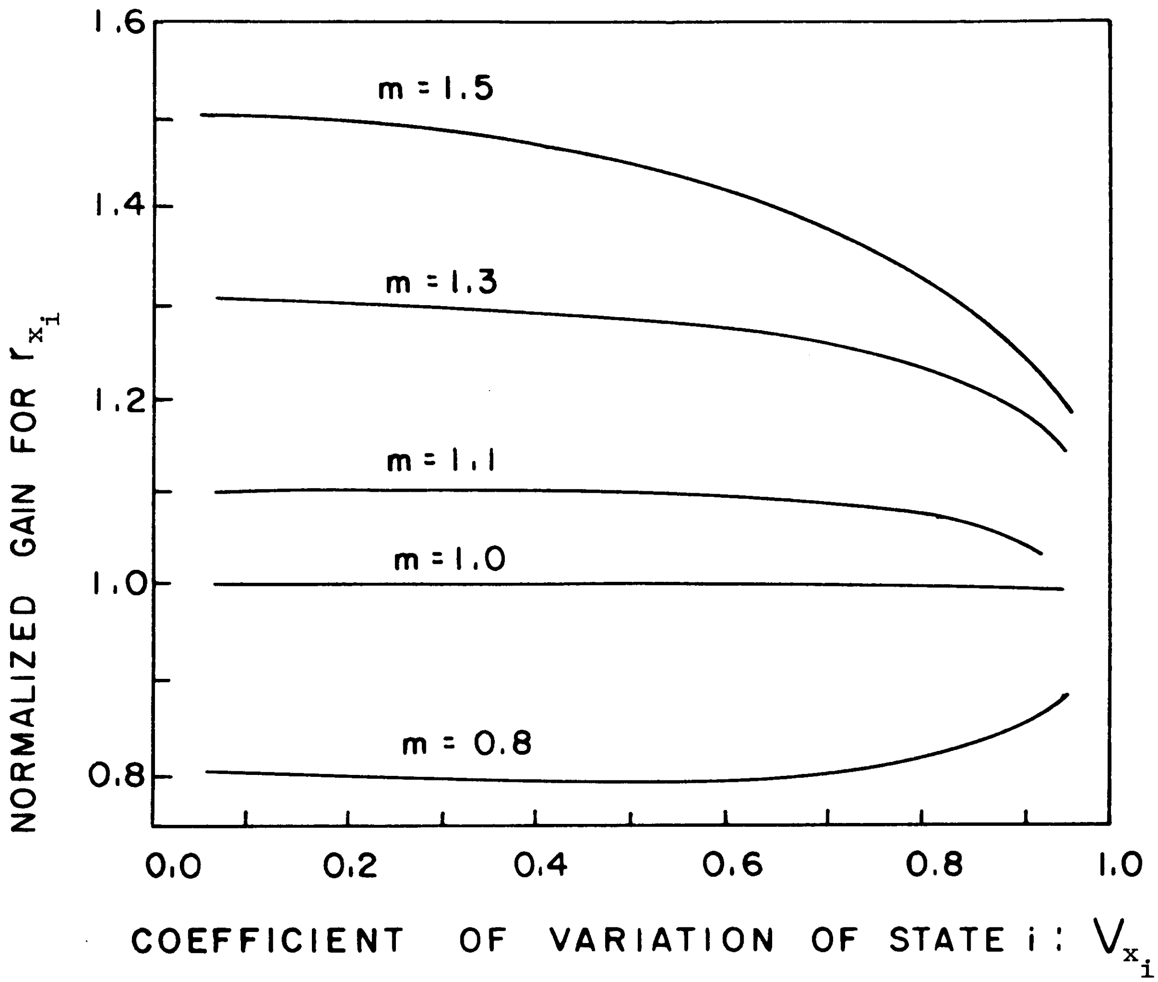

Figure 4.3

NORMALIZED GAIN FOR THE GAUSSIAN RESIDUAL r_{x_i} AS
A FUNCTION OF THE COEFFICIENT OF VARIATION OF THE i^{th} STATE
FOR DIFFERENT VALUES OF THE EXPONENT m

$$M_{b_i} = 1 + \frac{m \cdot (m-1)}{2} \cdot V_{x_i}^2 \qquad (4.55)$$

$$M_{x_i} = m + \frac{m \cdot (m-1) \cdot (m-2)}{2} \cdot V_{x_i}^2 \qquad (4.56)$$

as functions of the coefficient of variation V_{x_i}, for different values of m. Based on these figures, the following can be stated:

a. The normalized gain M_{b_i} is relatively insensitive to the value of the exponent for small values of V_i (≤ 0.30). It is quadratic in V_i.

b. The normalized gain M_{x_i} is relatively insensitive to the value of V_i for a given m. It is very sensitive to the value of m, especially for small coefficients of variation (for $V_i \leq 0.30$ is practically equal to m).

c. Consider the ordinary linearization using the linear terms in a Taylor series expansion of the function: $g(\mu_{a_i}, \mu_{x_i}, r_{x_i}) = \mu_{a_i} \cdot (\mu_{x_i} + r_{x_i})^m$ about the point $r_{x_i} = 0$. Then,

$$g_T(\mu_{a_i}, \mu_{x_i}, r_{x_i}) = \mu_{a_i} \cdot \mu_{x_i}^m + m \cdot \mu_{a_i} \cdot \mu_{x_i}^{(m-1)} \cdot r_{x_i} \qquad (4.56)$$

The normalized gains in this case are:

$$M_{b_{i_T}} = 1 \qquad (4.57)$$

$$M_{x_{i_T}} = m \qquad (4.58)$$

Direct comparison of Eqs. (4.57), (4.58) with Eqs. (4.55) and (4.56), respectively, shows that statistical linearization accounts for the effects that the variance of the random variable has on the output of the nonlinear function. If V_i can be assumed to be small (with a reasonable value of m), the results of the statistical linearization became identical to the ones of the ordinary linearization.

When the parameter a_i is equal to the sum of a bias term plus a zero mean Gaussian residual, Eqs. (4.47) through (4.49) should be used. In this case the normalized gains are

$$M_{b_i} = \frac{N_{b_i}}{\mu_{a_i} \cdot \mu_{x_i}^{m}} \cdot (\mu_{a_i} + \mu_{x_i}) \qquad (4.59)$$

$$M_{x_i} = \frac{N_{x_i}}{\mu_{a_i} \cdot \mu_{x_i}^{(m-1)}} \qquad (4.60)$$

$$M_{a_i} = \frac{N_{a_i}}{\mu_{x_i}^{m}} \qquad (4.61)$$

Figures 4.4, 4.5 and 4.6 present isometric plots (Restrepo-Posada, 1978) of M_{b_i} as a function of the coefficients of variation V_{a_i} and V_{x_i}, for correlation coefficients $\rho_{x_i a_i}$ equal to 0.3, 0.6 and 0.8 respectively.

Plots of M_{x_i} vs. V_{a_i} and V_{x_i} are displayed in figures 4.7, 4.8 and 4.9 for the same values of $\rho_{x_i a_i}$ respectively. Figure 4.10 is a plot of the normalized gain M_{a_i} as a function of V_{a_i} and V_{x_i}, which is independent of the value of $\rho_{x_i a_i}$ (Eq. (4.49)). In figures 4.4 through 4.10 the V_{a_i}, V_{x_i} axes initiate at the value 0 and terminate at the value 1. In all cases the vertical axis initiates at the value 0.7 and terminates at the value 1.5. The vertical coordinate of the point, where the surfaces intersect the vertical axis, is indicated in this group of figures. The effect of the correlation coefficient is more apparent in the case of the normalized gain M_{b_i}. It can be seen that the higher the correlation is, the more ($V_{a_i} \neq 0$, $V_{x_i} \neq 0$) the gain departs from the value it takes (i.e. one) when the nonlinear functions involved are linearized through truncation of their Taylor's series expansions after the linear terms.

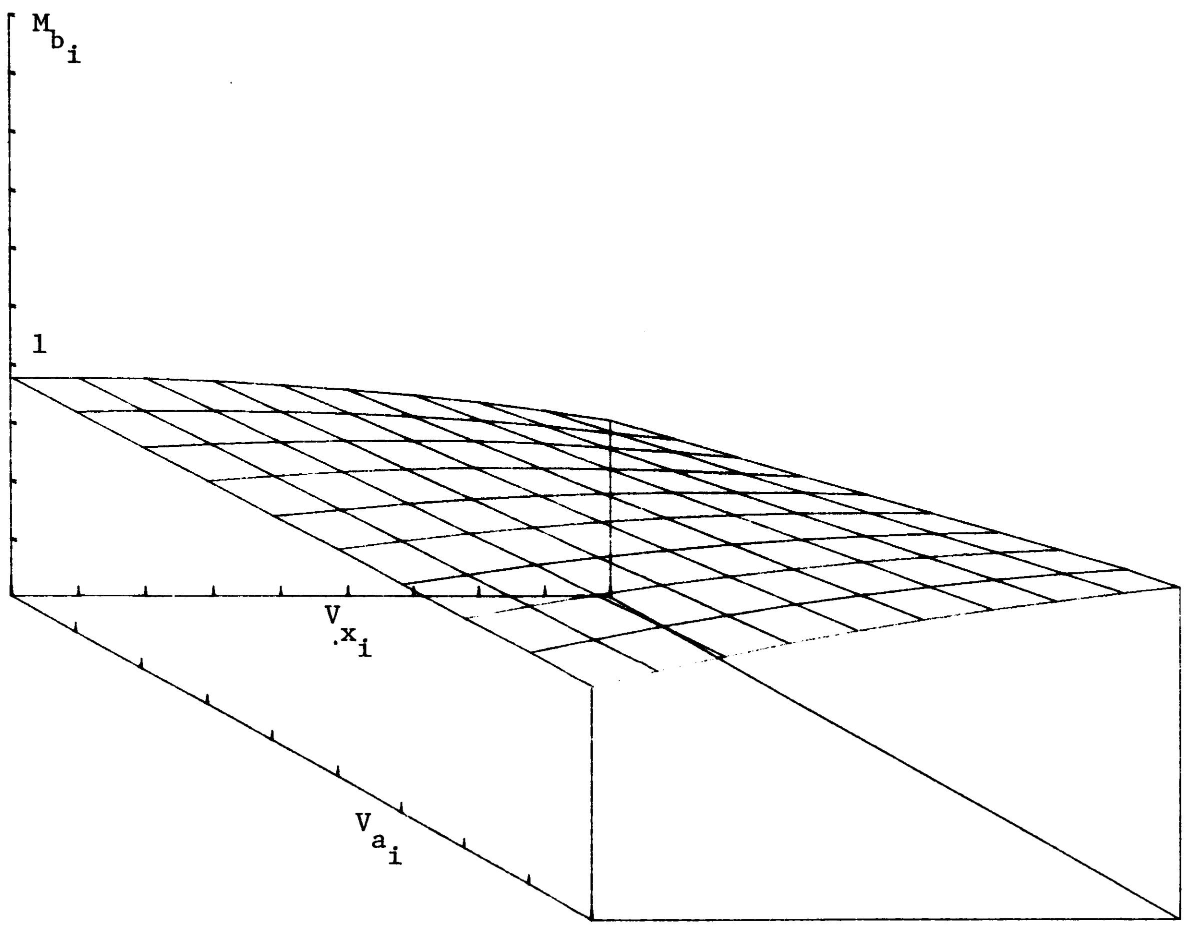

Figure 4.4

NORMALIZED BIAS GAIN AS A FUNCTION OF V_{x_i} AND V_{a_i}. $\rho_{x_i a_i} = 0.3$

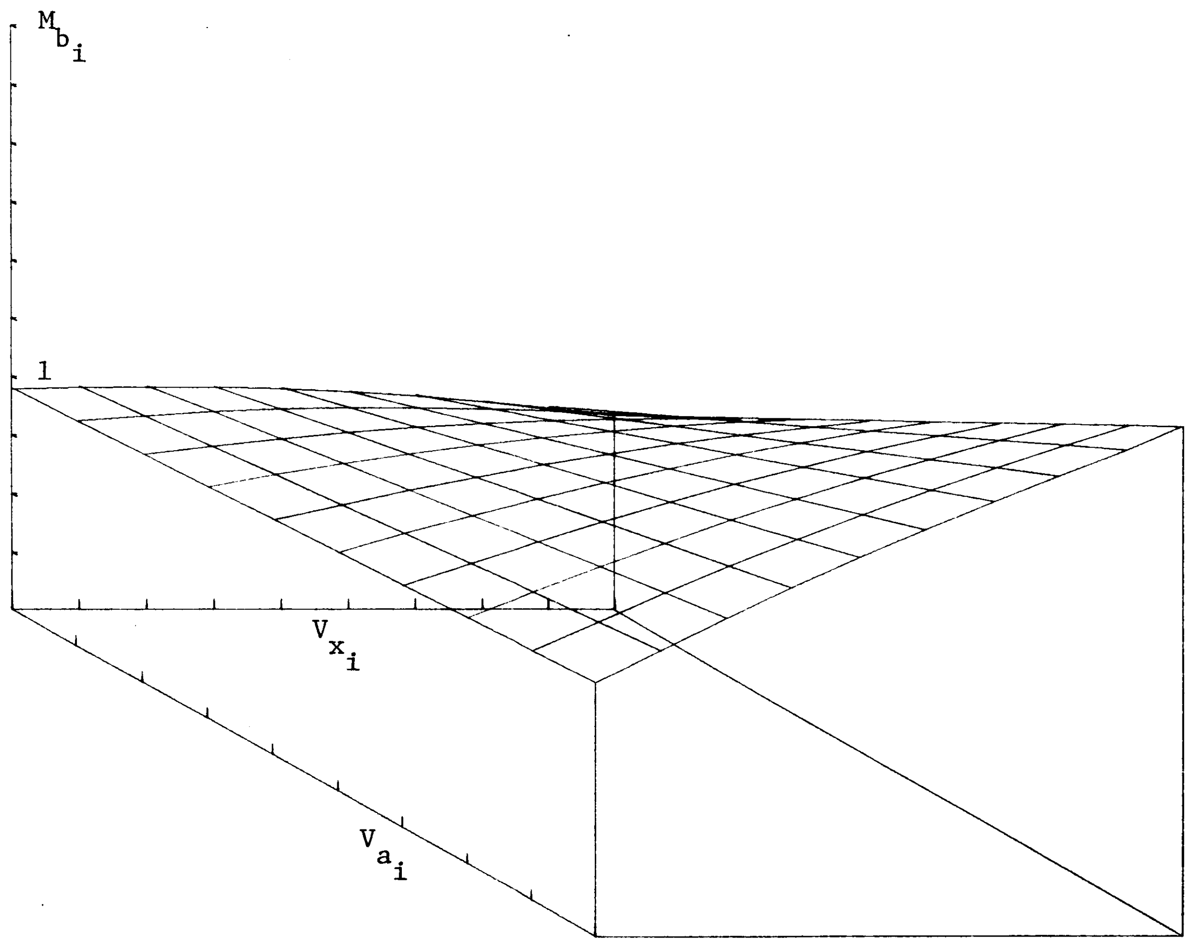

Figure 4.5

NORMALIZED BIAS GAIN AS A FUNCTION OF V_{x_i} AND V_{a_i}. $\rho_{x_i a_i} = 0.6$

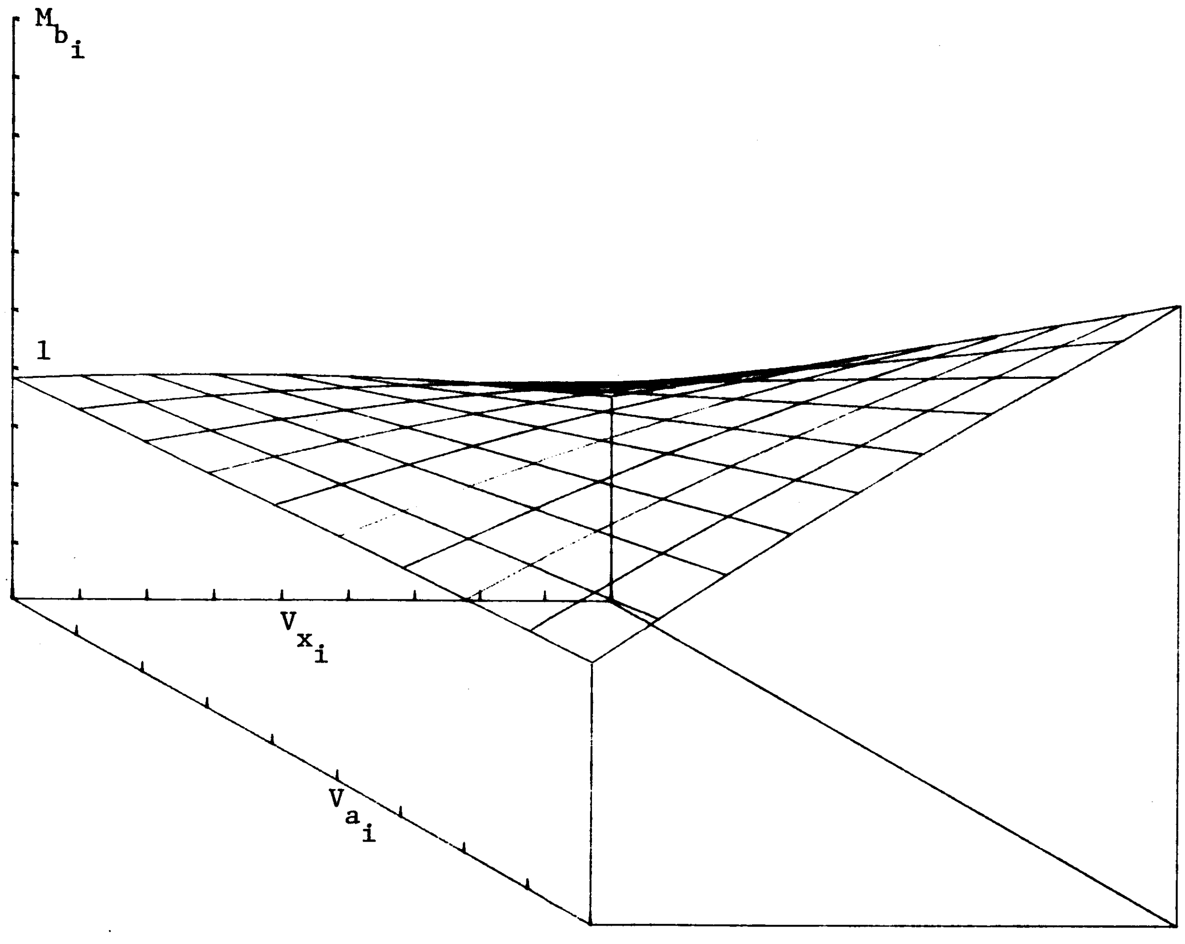

Figure 4.6

NORMALIZED BIAS GAIN AS A FUNCTION OF V_{x_i} AND V_{a_i}. $\rho_{x_i a_i} = 0.8$

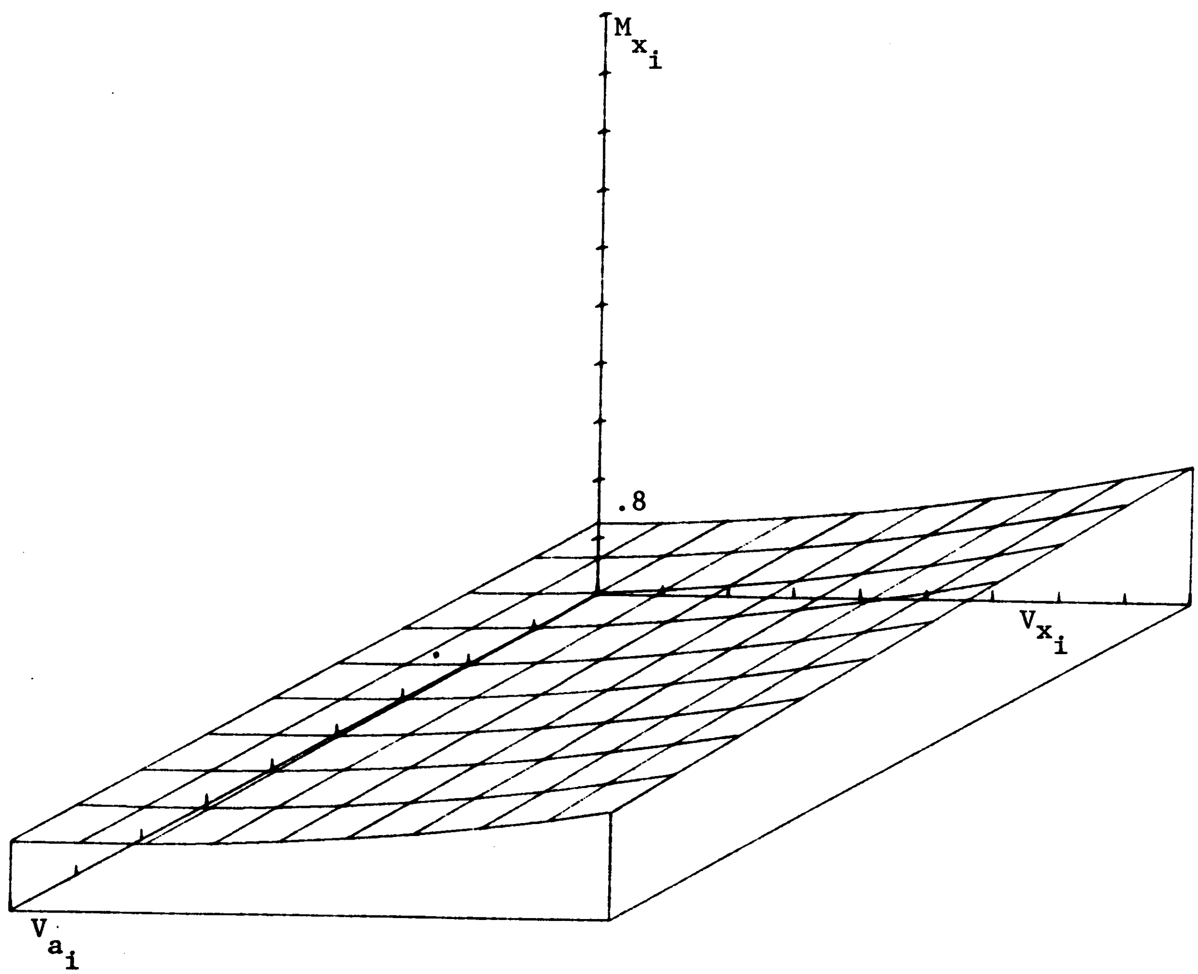

Figure 4.7

NORMALIZED RESIDUAL GAIN M_{x_i} AS A FUNCTION OF V_{x_i} AND V_{a_i}.

$\rho_{x_i a_i} = 0.3$

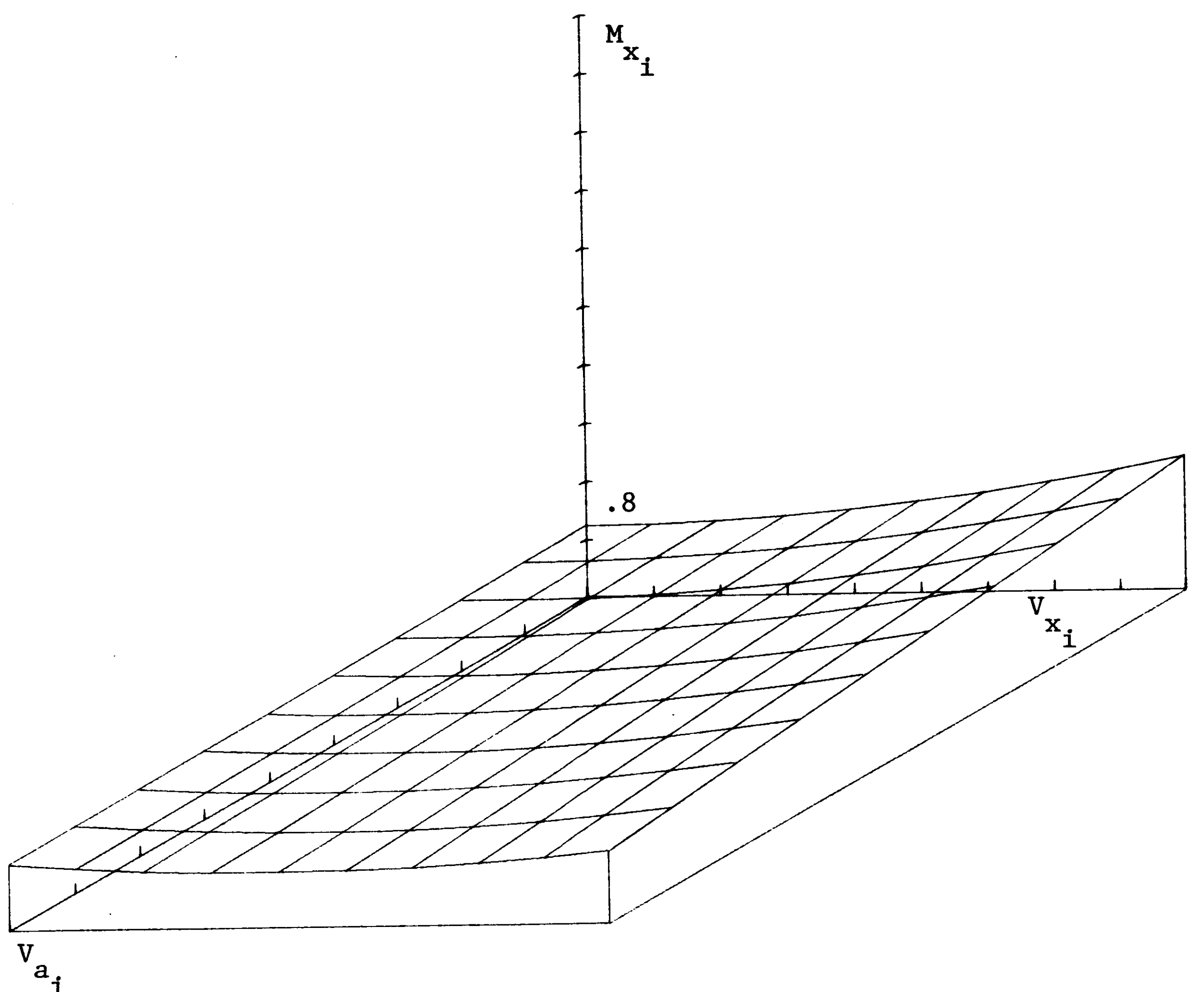

Figure 4.8

NORMALIZED RESIDUAL GAIN M_{x_i} AS A FUNCTION OF V_{x_i} AND V_{a_i}.

$\rho_{x_i a_i} = 0.6$

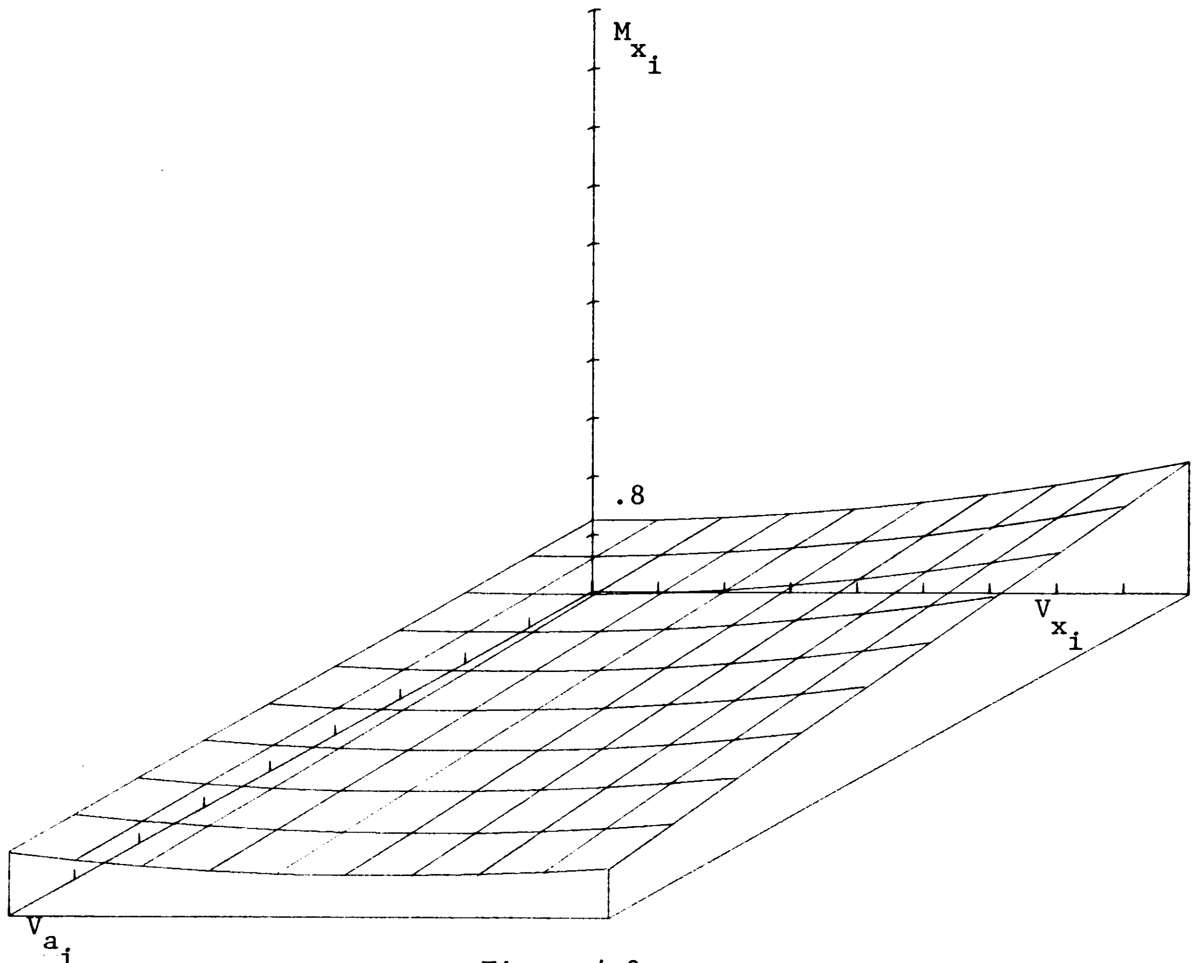

Figure 4.9

NORMALIZED RESIDUAL GAIN M_{x_i} AS A FUNCTION OF V_{x_i} AND V_{a_i}.

$\rho_{x_i a_i} = 0.8$

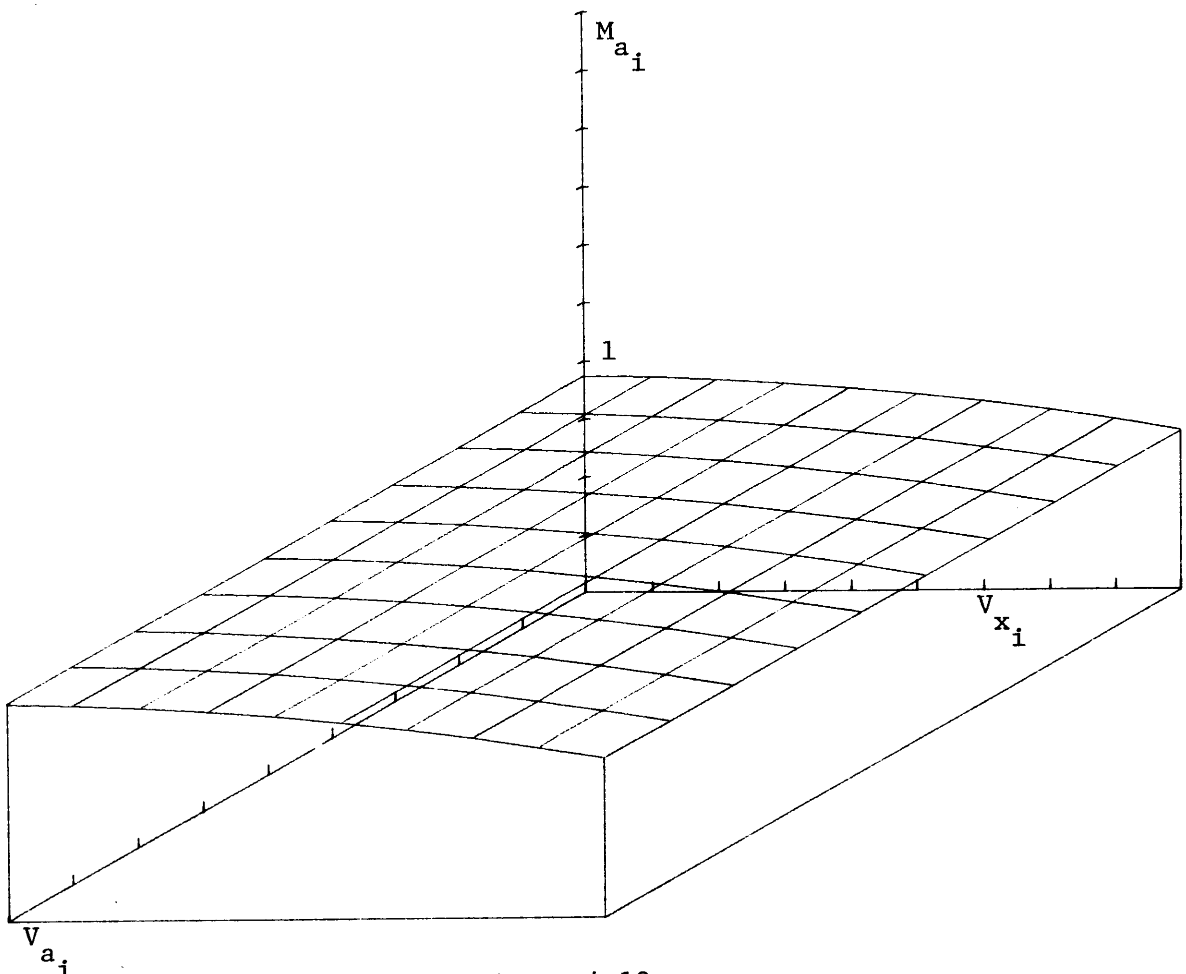

Figure 4.10

NORMALIZED RESIDUAL GAIN M_{a_i} AS A FUNCTION OF V_{x_i} AND V_{a_i}

Chapter 5

REAL TIME STATE AND PARAMETER ESTIMATION

5.1 Introduction

In recent years, the sequential, linear, minimum mean squared error estimator, called Kalman filter (Kalman and Bucy, 1961) has found widespread usage within the hydrologic forecasting framework. In particular, use of this estimator with conceptual hydrologic models has considerably improved their performance in forecasting streamflow (Kitanidis and Bras, 1978). In flood prevention studies, the Kalman filter provides a quantitative estimate of the measures of uncertainty needed.

This chapter concentrates on the study of the sequential estimator to be used with the statistically linearized flood routing model based on a series of nonlinear reservoirs.

At first, the state equations of the model are converted to canonical form, which leads to decoupled first order differential equations. This representation of the system is used for the integration of system equations as well as in the study of controllability and observability properties of the system. A Gaussian minimum variance estimator is employed to process the measurements and the necessary formulation is presented for on line state and parameter estimation. Stability of the state estimator is inferred from the observability and controllability properties of the system.

5.2 Canonical Form of Linearized State Equations

The differential equations governing the flood routing model based on n nonlinear reservoirs are of the type

$$\frac{dx_i(t)}{dt} = p_i \cdot u(t) + a_{i-1}(t) \cdot x_{i-1}^m(t) - a_i(t) \cdot x_i^m(t) \ ; \qquad i = 1, 2, \ldots, n \quad (5.1)$$

where $x_i(t)$ is the i^{th} state of the system at time t (volume of water in storage in the i^{th} reservoir); u(t) is the input to the system at time t; p_i, $a_i(t)$, $i = 0, 1, \ldots, n$, and m are the parameters of the system with $a_o(t) \triangleq 0 \ \forall t$ and $p_o \triangleq 0$.

In this section, the parameters $a_i(t)$, $i = 0, 1, 2, \ldots, n$, are assumed to be known constants.

Since nonlinear differential equations rarely have analytical solutions and, most importantly, powerful results have been obtained in estimation theory for linear systems, some sort of linearization of the system of equations described by Eq. (5.1), for all i, is in order. Suppose that the nonlinear ($m \neq 1$) functions $x_i^m(t)$, $i = 1, 2, \ldots, n$, are linearized in some way about some a priori determined reference trajectory ($\forall t$). In addition, since the input to the system is in most practical situations a piece-wise constant function (e.g., observations or predictions of the input available every six hours), it is assumed that the reference trajectory is a piece-wise constant function.

Under the above conditions, the set of Eqs. (5.1), for all i, can be written as,

$$\frac{dx_i(t)}{dt} = b_{i_k} + \beta_{i-1_k} \cdot x_{i-1}(t) - \beta_{i_k} \cdot x_i(t); \; i = 1, 2, \ldots, n$$

$$t_k \leq t \leq t_{k+1} \qquad (5.2)$$

where

$$b_{i_k} = p_i \cdot u(t_k) + a_{i-1}(t_k) \cdot x^m_{i-1_o}(t_k) - a_i(t_k) \cdot x^m_{i_o}(t_k) \; ;$$

$$i = 1, 2, \ldots, n$$

$$(5.3)$$

$$\beta_{i_k} \triangleq \beta_i(t_k), \; i = 1, 2, \ldots, n, \; k = 0, 1, 2, \ldots \qquad (5.4)$$

$\beta_i(t_k)$, $i = 1, 2, \ldots, n$, $k = 0, 1, 2, \ldots$ are the coefficients due to linearization of the functions $a_i \cdot x^m_i(t)$, $i = 1, 2, \ldots, n$, about the reference trajectory $x_{i_o}(t_k)$, $i = 1, 2, \ldots, n$, $k = 0, 1, 2, \ldots$, and, the intervals $[t_k, t_{k+1}]$ $\forall k$ are the intervals of constant input, $u(t_k)$.

Note that Eq. (5.2) describes a set of n coupled, first order, linear differential equations, for t in the interval $[t_k, t_{k+1}]$.

The system of differential equations can be converted to canonical (normal or diagonal) form. This is a particularly convenient representation since it leads to uncoupled first order differential equations for the description of the state ($x_i(t)$, $i = 1, 2, \ldots, n$) evolution in time. In addition, this type of representation is very useful in the study of system controllability and observability to be undertaken in later sections.

The constraint $t_k \leq t \leq t_{k+1}$ will be assumed (unless otherwise stated), and it will not be written explicitly for convenience in notation.

The set of equations described by Eq. (5.2) can be written in matrix form as,

$$\frac{d\underline{x}(t)}{dt} = \underline{b}_k + \overline{\underline{A}}_k \cdot \underline{x}(t) \tag{5.5}$$

where vectors are underlined and matrices under- and over-lined, and

$$\underline{x}(t) = [x_1(t) \quad x_2(t) \; . \; . \; . \; x_n(t)]^T \tag{5.6}$$

$$\underline{b}_k = [b_{1_k} \quad b_{2_k} \; . \; . \; . \; b_{n_k}]^T \tag{5.7}$$

$$\overline{\underline{A}}_k = \begin{vmatrix} -\beta_{1_k} & 0 & 0 & . & . & 0 & 0 \\ \beta_{1_k} & -\beta_{2_k} & 0 & . & . & 0 & 0 \\ 0 & \beta_{2_k} & -\beta_{3_k} & . & . & 0 & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & -\beta_{n-1_k} & 0 \\ 0 & 0 & 0 & . & . & \beta_{n-1_k} & -\beta_{n_k} \end{vmatrix} \text{(nxn)} \tag{5.8}$$

The superscript "T" denotes the transpose of a vector or matrix quantity.

The eigenvalues λ_{i_k}, i = 1, 2, ..., n, of the matrix $\overline{\underline{A}}_k$ can be found as the solution of the algebraic equation:

$$\det\{\lambda_k \cdot \underline{\overline{I}}_{nn} - \underline{\overline{A}}_k\} = 0 \qquad (5.9)$$

with det { • } denoting determinant and $\underline{\overline{I}}_{nn}$, the n^{th} order unit matrix. Since the quantity in brackets in Eq. (5.9) is a lower triangular matrix (e.g., Eq. (5.8)), its determinant is equal to: $\prod_{i=1}^{n} (\lambda_k + \beta_{i_k})$. Consequently, Eq. (5.9) is satisfied if

$$\lambda_{i_k} = -\beta_{i_k} \quad , \quad i = 1, 2, \ldots, n \qquad (5.10)$$

It is now possible to determine the set of eigenvectors $\underline{u}_{i_k}$, $i = 1, 2, \ldots, n$, associated with λ_{i_k}, $i = 1, 2, \ldots, n$. By definition,

$$\lambda_{i_k} \cdot \underline{u}_{i_k} = \underline{\overline{A}}_k \cdot \underline{u}_{i_k} \; ; \quad i = 1, 2, \ldots, n \qquad (5.11)$$

$$\underline{u}_{i_k} = [u_{1,i_k} \quad u_{2,i_k} \quad \cdot \cdot \cdot \quad u_{n,i_k}]^T \qquad (5.12)$$

Substitution of Eq. (5.12) in Eq. (5.11) yields,

$$\lambda_{i_k} \cdot \begin{vmatrix} u_{1,i_k} \\ u_{2,i_k} \\ \cdot \\ \cdot \\ \cdot \\ u_{n,i_k} \end{vmatrix} = \begin{vmatrix} \lambda_{1_k} & 0 & \cdot & \cdot & 0 & 0 \\ -\lambda_{1_k} & \lambda_{2_k} & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & -\lambda_{n-1_k} & \lambda_{n_k} \end{vmatrix} \cdot \begin{vmatrix} u_{1,i_k} \\ u_{2,i_k} \\ \cdot \\ \cdot \\ \cdot \\ u_{n,i_k} \end{vmatrix} ; \; i=1,2,\ldots,n$$

$$(5.13)$$

Due to the sparse nature of matrix $\underline{\overline{A}}_k$, one can solve the system of equations described by Eq. (5.13) to find u_{1,i_k}, ..., u_{n,i_k} for each i (i = 1, 2, ..., n). For distinct eigenvalues,

$$u_{j,i_k} = 0 \quad ; \quad j < i,\ j = 1, 2. \ldots, n-1 \tag{5.14}$$

$$u_{i,i_k} = c \quad ; \quad i = 1, 2, \ldots, n \tag{5.15}$$

$$u_{j,i_k} = \left[\prod_{\ell=i+1}^{j} \left(\frac{\lambda_{\ell-1_k}}{\lambda_{\ell_k} - \lambda_{i_k}} \right) \right] \cdot c \quad ; \quad j > i,\ j = 2, \ldots, n \tag{5.16}$$

where $\Pi(\cdot)$ represents the product operator and c is an arbitrary constant taken as unity in this work.

Next, a matrix $\underline{\bar{T}}_k$ is constructed with the eigenvectors $\underline{u}_{i_k}$, $i = 1, 2, \ldots, n$, as follows,

$$\underline{\bar{T}}_k = [\underline{u}_{1_k} \vdots\ \underline{u}_{2_k} \vdots \ \cdots \ \vdots \underline{u}_{n_k}] \quad (nxn) \tag{5.17}$$

Explicitly, $\underline{\bar{T}}_k$ takes the form,

$$\underline{\bar{T}}_k = \begin{vmatrix} 1 & 0 & \cdot & \cdot & 0 \\ \left(\frac{\lambda_{1_k}}{\lambda_{2_k} - \lambda_{1_k}}\right) & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \prod_{\ell=2}^{n} \left(\frac{\lambda_{\ell-1_k}}{\lambda_{\ell_k} - \lambda_{1_k}}\right) & \prod_{\ell=3}^{n} \left(\frac{\lambda_{\ell-1_k}}{\lambda_{\ell_k} - \lambda_{2_k}}\right) & \cdot & \cdot & 1 \end{vmatrix} \quad (nxn) \tag{5.18}$$

Since $\underline{\bar{T}}_k$ is a lower triangular matrix:

$$\det\{\underline{\bar{T}}_k\} = \prod_{\ell=1}^{n} (1) = 1 \neq 0 \tag{5.19}$$

It follows that the matrix $\underline{\overline{T}}_k^{-1}$, called the inverse of $\underline{\overline{T}}_k$, exists and it is also a lower triangular matrix, whose elements can be computed recursively by

$$t^{-1}_{i,j_k} = - \sum_{\ell=j}^{i-1} t_{i,\ell_k} \cdot t^{-1}_{\ell,j_k} \quad ; \; i > j, \; j = 1, 2, \ldots, n-1 \tag{5.20}$$

$$t^{-1}_{i,i_k} = 1 \quad ; \; i = 1, 2, \ldots, n \tag{5.21}$$

$$t^{-1}_{i,j} = 0 \quad ; \; i < j, \; j = 2, \ldots, n \tag{5.22}$$

where t_{i,j_k} represents the $(i, j)^{th}$ element of matrix $\underline{\overline{T}}_k$.

Substitution of t_{i,ℓ_k} from Eq. (5.18) reduces Eq. (5.20) (after straightforward but tedious algebraic manipulations) to,

$$t^{-1}_{i,j_k} = (-1)^{i+j} \cdot \prod_{\ell=j}^{i-1} \left(\frac{\lambda_{\ell_k}}{\lambda_{i_k} - \lambda_{\ell_k}}\right) \quad ; \; i > j, \; j = 1, 2, \ldots, n-1 \tag{5.23}$$

For example, if n = 3, then,

$$\underline{\overline{T}}_k = \begin{vmatrix} 1 & 0 & 0 \\ \left(\frac{\lambda_{1_k}}{\lambda_{2_k} - \lambda_{1_k}}\right) & 1 & 0 \\ \frac{\lambda_{2_k} \cdot \lambda_{1_k}}{(\lambda_{3_k} - \lambda_{1_k}) \cdot (\lambda_{2_k} - \lambda_{1_k})} & \left(\frac{\lambda_{2_k}}{\lambda_{3_k} - \lambda_{2_k}}\right) & 1 \end{vmatrix} \tag{5.24}$$

and

$$\underline{\bar{T}}_k^{-1} = \begin{vmatrix} 1 & 0 & 0 \\ -\left(\frac{\lambda_{1_k}}{\lambda_{2_k} - \lambda_{1_k}}\right) & 1 & 0 \\ \frac{\lambda_{2_k} \cdot \lambda_{1_k}}{(\lambda_{3_k} - \lambda_{2_k})(\lambda_{3_k} - \lambda_{1_k})} & -\left(\frac{\lambda_{2_k}}{\lambda_{3_k} - \lambda_{2_k}}\right) & 1 \end{vmatrix} \tag{5.25}$$

Define the set of canonical variables, $y_1(t)$, $y_2(t)$, ..., $y_n(t)$, by the transformation,

$$\underline{x}(t) = \underline{\bar{T}}_k \cdot \underline{y}(t) \tag{5.26}$$

with

$$\underline{y}(t) = [y_1(t) \quad y_2(t) \ . \ . \ . \ y_n(t)]^T \tag{5.27}$$

Substitution of Equation (5.26) into Equation (5.5) results in

$$\underline{\bar{T}}_k \cdot \frac{d}{dt} (\underline{y}(t)) = \underline{\bar{A}}_k \cdot \underline{\bar{T}}_k \cdot \underline{y}(t) + \underline{b}_k \tag{5.28}$$

Multiplication of both sides of Eq. (5.28) by $\underline{\bar{T}}_k^{-1}$ reveals,

$$\frac{d}{dt} \cdot (\underline{y}(t)) = \underline{\bar{T}}_k^{-1} \cdot \underline{\bar{A}}_k \cdot \underline{\bar{T}}_k \cdot \underline{y}(t) + \underline{\bar{T}}_k^{-1} \cdot \underline{b}_k \tag{5.29}$$

It can be shown (Schultz and Melsa, 1967) that,

$$\underline{\bar{T}}_k^{-1} \cdot \underline{\bar{A}}_k \cdot \underline{\bar{T}}_k = \underline{\bar{\Lambda}}_k \tag{5.30}$$

where $\underline{\bar{\Lambda}}_k$ is a diagonal matrix, whose elements are the corresponding eigenvalues of $\underline{\bar{A}}_k$. Thus, denoting by $\underline{c}_k$ the vector $\underline{\bar{T}}_k^{-1} \cdot \underline{b}_k$, the equation describing the motion of the canonical state vector $\underline{y}(t)$, of the system is

$$\frac{d}{dt}(\underline{y}(t)) = \overline{\underline{\Lambda}}_k \cdot \underline{y}(t) + \underline{c}_k \tag{5.31}$$

Each element of the vector $\underline{y}(t)$ obeys

$$\frac{d}{dt}(y_i(t)) = \lambda_{i_k} \cdot y_i(t) + c_{i_k} \quad ; \quad i = 1, 2, \ldots, n \tag{5.32}$$

Notice that the first order, linear equations derived, are decoupled. Consequently, the solution (integration) of Eq. (5.31) is in terms of a diagonal transition (or fundamental) matrix, $\overline{\underline{\Phi}}_k(t_1, t_2)$, whose $(i, i)^{th}$ element $\phi_{i,i_k}(t_1, t_2)$ satisfies,

$$\frac{d}{dt}(\phi_{i,i_k}(t, t_o)) = \lambda_{i_k} \cdot \phi_{i,i_k}(t, t_o) \tag{5.33}$$

with initial condition

$$\phi_{i,i_k}(t_o, t_o) = 1 \tag{5.34}$$

It follows that,

$$\phi_{i,i_k}(t, t_o) = e^{\lambda_{i_k}(t - t_o)} \quad ; \quad i = 1, 2, \ldots, n \tag{5.35}$$

The solution of Eq. (5.32) is then,

$$y_i(t) = \phi_{i,i_k}(t, t_o) \cdot y_i(t_o) + \int_{t_o}^{t} \phi_{i,i_k}(t, s) \cdot c_{i_k} \cdot ds \quad ; \quad i = 1, 2, \ldots, n \tag{5.36}$$

or substituting for $\phi_{i,i_k}(t, t_o)$ from Eq. (5.35) for all i results in

$$y_i(t) = y_i(t_o) \cdot e^{\lambda_{i_k}(t-t_o)} + c_{i_k} \int_{t_o}^{t} e^{\lambda_{i_k}(t-s)} \cdot ds \quad ;$$

$$i = 1, 2, \ldots, n \quad (5.37)$$

Thus,

$$y_i(t) = y_i(t_o) \cdot e^{\lambda_{i_k}(t-t_o)} - \frac{c_{i_k}}{\lambda_{i_k}} \cdot (1 - e^{\lambda_{i_k}(t-t_o)}) \quad ;$$

$$i = 1, 2, \ldots, n \qquad (5.38)$$

assuming λ_{i_k} is nonzero for all i. The above solution is in the interval $[t_k, t_{k+1}]$.

If one is only interested at the end points of the intervals $[t_k, t_{k+1}]$, k = 0, 1, 2, ... (e.g., observations of the output are available at these points), then Eq. (5.38) yields

$$y_{i_{k+1}} = y_{i_k} \cdot e^{\lambda_{i_k} \cdot \Delta t_k} - \frac{c_{i_k}}{\lambda_{i_k}} \cdot (1 - e^{\lambda_{i_k} \cdot \Delta t_k}) \quad ;$$

$$i = 1, 2, \ldots, n, \; k = 0, 1, 2, \ldots \qquad (5.39)$$

with Δt_k defined as:

$$\Delta t_k = t_{k+1} - t_k \quad ; \qquad k = 0, 1, 2, \ldots \qquad (5.40)$$

Denote by $\phi_{i,i}(t_{k+1}, t_k)$ the one step transition function. It follows that the multiple step transition function is given by

$$\phi_{i,i}(t_k, t_j) = \prod_{\ell=j}^{k-1} \phi_{i,i}(t_{\ell+1}, t_\ell) \quad ;$$

$$k > j, \; i = 1, 2, \ldots, n, \; j = 0, 1, 2, 3\ldots \qquad (5.41)$$

and

$$\phi_{i,i}(t_k, t_j) = \prod_{\ell=k}^{j-1} \phi_{i,i}(t_\ell, t_{\ell+1}) \quad ;$$

$$j > k,\ i = 1, 2, \ldots, n,\ k = 0, 1, 2, 3, \ldots \tag{5.42}$$

where $\phi_{i,i}(t_k, t_{k+1})$ is,

$$\phi_{i,i}(t_k, t_{k+1}) = e^{-\lambda_{i_k} \cdot \Delta t_k} \quad ; \quad i = 1, 2, \ldots, n,\ k = 0, 1, 2, \ldots \tag{5.43}$$

The multi-step transition matrix $\underline{\overline{\Phi}}(t_k, t_j)$ is defined as a diagonal matrix whose $(i, i)^{th}$ element is equal to the discrete transition function $\phi_{i,i}(t_k, t_j)$. These matrices will be used later in conjunction with the conditions for stochastic observability and controllability of the linearized and discretized system.

5.3 Gaussian Minimum Variance Estimator

It is a common procedure in engineering practice to work with mathematical models that are simpler, but less accurate than the best available model of a given physical process. The major reasons for this are: a) to reduce the computational burden associated with simulation, b) to simplify the analysis of the system under study. Consequently, the output of the mathematical model differs from the observed one.

Formally, one can attribute the resulting discrepancy to the following errors:

a) errors in the structure of the model,

b) errors in the values of the parameters of the model,

c) errors in the input data,

d) errors in the observation of the output.

Treatment of the above mentioned error sources within a probabilistic framework have lead to a well-defined class of stochastic models. Deterministic models have been coupled with models representing the probabilistic behavior of the above mentioned errors. For example, the time varying linear deterministic model, with dynamics,

$$\frac{d\underline{x}(t)}{dt} = \bar{\underline{F}}(t) \cdot \underline{x}(t) + \bar{\underline{B}}(t) \cdot \underline{u}(t) \tag{5.44}$$

and output equation,

$$\underline{y}(t_k) = \bar{\underline{H}}(t_k) \cdot \underline{x}(t_k) \quad ; \quad k = 1, 2, \ldots \tag{5.45}$$

can be modified to the form,

$$\frac{d\underline{x}(t)}{dt} = \bar{\underline{F}}(t) \cdot \underline{x}(t) + \bar{\underline{B}}(t) \cdot \underline{u}(t) + \bar{\underline{G}}(t) \cdot \underline{w}(t) \tag{5.46}$$

and

$$\underline{y}(t_k) = \bar{\underline{H}}(t_k) \cdot \underline{x}(t_k) + \underline{v}(t_k) \quad ; \quad k = 1, 2, \ldots \tag{5.47}$$

where $\underline{x}(t)$ is the n-dimensional state of the model at time t, $\underline{u}(t)$ is the p-dimensional input vector, $\underline{y}(t_k)$ is the m-dimensional output (or observation) vector, $\underline{w}(t)$ is a continuous time random error process of dimension r, and $\underline{v}(t_k)$ is an additive error sequence (at times t_1, t_2, ...) of dimension m. $\overline{\underline{F}}(t)$, $\overline{\underline{B}}(t)$, $\overline{\underline{G}}(t)$ are time varying coefficient matrices of dimensions (nxn), (nxp), (nxr), respectively, and $\overline{\underline{H}}(t_k)$ is the measurement coefficient matrix at time step k.

A very useful model for the continuous time error process $\underline{w}(t)$ has proven to be white noise. This process can be thought to be the formal derivative of the Wiener's process (since Wiener's process is not mean square differentiable) (Jazwinski, 1970). The unique feature that is responsible for its wide use in filtering theory is its peculiar autocorrelation function, which is a Dirac delta function. Note that this implies infinite power for this process, thus it is not physically realizable. Due to the delta autocorrelation function, when $\underline{w}(t)$ of Eq. (5.46) is modeled as a white noise process and is independent of the initial state, then the generated state vector process $\underline{x}(t)$ is a Markov process (u(t) is assumed to be a known function of time). In fact, this holds true for a much more general situation, namely, when the physical phenomenon is described by a nonlinear differential equation.

In an analogous manner, $\underline{v}(t_k)$ is usually modeled as a white sequence. That is,

$$E\{[\underline{v}(t_k) - E\{\underline{v}(t_k)\}][\underline{v}(t_j) - E\{\underline{v}(t_j)\}]^T\} = \overline{\underline{R}}(t_k) \cdot \delta_{kj} \quad (5.48)$$

where $\overline{\underline{R}}(t_k)$ is the covariance matrix of $\underline{v}(t_k)$ at time t_k and δ_{kj} is Kronecker's delta, taking the value one when k is equal to j and zero otherwise.

A fundamental problem in filtering theory is to give the best (according to some criterion of optimality) estimate of the state of a continuous system given discrete-time observations of the output up to a given instant, as well as information about the statistics of the state at some initial time and of the error processes at all times. One can decompose the problem in: a) the propagation step, where the state estimate $\underline{\hat{x}}(t_k|t_k)$ at time t_k having processed the observation at t_k, needs to be propagated in the interval of time $[t_k, t_{k+1}]$, where there are no observations of the output available, and b) the updating step, where the estimate $\underline{\hat{x}}(t_{k+1}|t_k)$ at time t_{k+1} resulting from the propagation step, needs to be corrected to the estimate $\underline{\hat{x}}(t_{k+1}|t_{k+1})$ by the incorporation of the information contained in the observation available at time t_{k+1}.

For the continuous time differential equation (5.48), it can be shown (Gelb, 1974):

$$\frac{d}{dt}\left(\underline{\hat{x}}(t|t_k)\right) = \overline{\underline{F}}(t) \cdot \underline{\hat{x}}(t|t_k) + \overline{\underline{B}}(t) \cdot \underline{u}(t) \tag{5.49}$$

with initial condition:

$$\underline{\hat{x}}(t_o|t_k) = \underline{\hat{x}}(t_k|t_k) \tag{5.50}$$

and

$$\frac{d}{dt}\left(\overline{\underline{P}}(t|t_k)\right) = \overline{\underline{F}}(t) \cdot \overline{\underline{P}}(t|t_k) + \overline{\underline{P}}(t|t_k) \cdot \overline{\underline{F}}^T(t) + \overline{\underline{G}}(t) \cdot \overline{\underline{Q}}(t) \cdot \overline{\underline{G}}^T(t) \tag{5.51}$$

with initial condition

$$\underline{\overline{P}}(t_o|t_k) = \underline{\overline{P}}(t_k|t_k) \tag{5.52}$$

$\underline{\overline{Q}}(t)$ is the covariance parameter matrix (a time varying spectral density) of the white noise process $\underline{w}(t)$ and is defined by,

$$E\{\underline{w}(t) \cdot \underline{w}^T(t)\} = \underline{\overline{Q}}(t) \cdot \delta(\tau) \tag{5.53}$$

with

$$E\{\underline{w}(t)\} = \underline{0} \tag{5.54}$$

$\delta(\tau)$ is the Dirac delta function, being zero at values of τ different from zero and infinite at τ equal to zero.

Solution of the set of Eqs. (5.49) through (5.52) gives the optimal estimate (mean process) of the state: $\hat{\underline{x}}(t_{k+1}|t_k)$ and the associated error covariance (since the optimal estimate is the mean): $P(t_{k+1}|t_k)$, given the estimate $x(t_k|t_k)$ and the error covariance $P(t_k|t_k)$. This completes the propagation step.

It is rather clear that in the processing of the observations, the conditional probability density $p(\underline{x}(t)|\underline{Y}_k)$, where $\underline{Y}_k$ is the vector of all measurements up to time t_k, plays a central role. It can be thought to be the real state of the stochastic system.

If one employs a mean quadratic estimation error criterion, then it is known that the optimal estimate is the conditional mean, for all types of probability densities. If, in addition, one assumes

Gaussian initial conditions for the state and Gaussian statistics for the driving noises $\underline{w}(t)$ and $\underline{v}(t_k)$, then the following are true:

a. During the propagation step, the state remains Gaussian with mean $\hat{\underline{x}}(t|t_k)$ and covariance $\overline{\underline{P}}(t|t_k)$ for $t_k \leq t \leq t_{k+1}$.

b. The random vectors $\underline{x}(t_{k+1})/\underline{Y}_k$ and $\underline{v}(t_{k+1})$ are jointly Gaussian due to the independence properties of $\underline{v}(t_{k+1})$. Consequently, $\underline{y}(t_{k+1})$ is a Gaussian random vector being the sum of jointly Gaussian random vectors, $\underline{y}(t_{k+1}) = \overline{\underline{H}}(t_{k+1}) \cdot \underline{x}(t_{k+1}) + \underline{v}(t_{k+1})$.

c. Following (b), $\underline{x}(t_{k+1})/\underline{Y}_k$ and $\underline{y}(t_{k+1})/\underline{Y}_k$ are jointly Gaussian random vectors since,

$$\begin{vmatrix} \underline{x}(t_{k+1}) \\ \underline{y}(t_{k+1}) \end{vmatrix} = \begin{vmatrix} \overline{\underline{I}}_{nn} & \overline{\underline{0}}_{nm} \\ \overline{\underline{H}}(t_{k+1}) & \overline{\underline{I}}_{mm} \end{vmatrix} \cdot \begin{vmatrix} \underline{x}(t_{k+1}) \\ \underline{v}(t_{k+1}) \end{vmatrix} \tag{5.55}$$

d. Since $\underline{x}(t_{k+1})/\underline{Y}_k$ and $\underline{y}(t_{k+1})/\underline{Y}_k$ are jointly Gaussian, the conditional density is also Gaussian with mean,

$$\hat{\underline{x}}(t_{k+1}|t_{k+1}) = \hat{\underline{x}}(t_{k+1}|t_k) + \overline{\underline{P}}_{xy}(t_{k+1}|t_k) \cdot \overline{\underline{P}}_y^{-1}(t_{k+1}|t_k) \cdot (\underline{y}(t_{k+1}) - \hat{\underline{y}}(t_{k+1}|t_k)) \tag{5.56}$$

and covariance:

$$\overline{\underline{P}}(t_{k+1}|t_{k+1}) = \overline{\underline{P}}(t_{k+1}|t_k) - \overline{\underline{P}}_{xy}(t_{k+1}|t_k) \cdot \overline{\underline{P}}_y^{-1}(t_{k+1}|t_k) \cdot \overline{\underline{P}}_{yx}(t_{k+1}|t_k) \tag{5.57}$$

It holds that,

$$\hat{\underline{y}}(t_{k+1}|t_k) = \overline{\underline{H}}(t_{k+1}) \cdot \hat{\underline{x}}(t_{k+1}|t_k) \tag{5.58}$$

$$\overline{\underline{P}}_y(t_{k+1}|t_k) = E\{[\overline{\underline{H}}(t_{k+1}) \cdot (\underline{x}(t_{k+1}) - \hat{\underline{x}}(t_{k+1}|t_k)) + \underline{v}(t_{k+1})] \cdot [\overline{\underline{H}}(t_{k+1}) \cdot (\underline{x}(t_{k+1}) - \hat{\underline{x}}(t_{k+1}|t_k)) + \underline{v}(t_{k+1})]^T\} \tag{5.59}$$

or

$$\overline{\underline{P}}_y(t_{k+1}\ t_k) = \overline{\underline{H}}(t_{k+1}) \cdot \overline{\underline{P}}(t_{k+1}|t_k) \cdot \overline{\underline{H}}^T(t_{k+1}) + \overline{\underline{R}}(t_{k+1}) \tag{5.60}$$

and

$$\overline{\underline{P}}_{xy}(t_{k+1}|t_k) = E\{[\underline{x}(t_{k+1}) - \hat{\underline{x}}(t_{k+1}|t_k)] \cdot [\overline{\underline{H}}(t_{k+1}) \cdot (\underline{x}(t_{k+1}) - \hat{\underline{x}}(t_{k+1}|t_k)) + \underline{v}(t_{k+1})]^T\} \tag{5.61}$$

or

$$\overline{\underline{P}}_{xy}(t_{k+1}|t_k) = \overline{\underline{P}}(t_{k+1}|t_k) \cdot \overline{\underline{H}}^T(t_{k+1}) \tag{5.62}$$

Substitution of Eqs. (5.58) through (5.62) in Eqs. (5.56) and (5.57) yields the optimal minimum variance estimator equations for Gaussian statistics,

$$\hat{\underline{x}}(t_{k+1}|t_{k+1}) = \hat{\underline{x}}(t_{k+1}|t_k) + \overline{\underline{K}}(t_{k+1}) \cdot (\underline{y}(t_{k+1}) - \overline{\underline{H}}(t_{k+1}) \cdot \hat{\underline{x}}(t_{k+1}|t_k)) \tag{5.63}$$

$$\underline{\bar{P}}(t_{k+1}|t_{k+1}) = \underline{\bar{P}}(t_{k+1}|t_k) - \underline{\bar{K}}(t_{k+1}) \cdot \underline{\bar{H}}(t_{k+1}) \cdot \underline{\bar{P}}(t_{k+1}|t_k) \tag{5.64}$$

with

$$\underline{\bar{K}}(t_{k+1}) = \underline{\bar{P}}(t_{k+1}|t_k) \cdot \underline{\bar{H}}^T(t_{k+1}) \cdot [\underline{\bar{H}}(t_{k+1}) \cdot \underline{\bar{P}}(t_{k+1}|t_k) \cdot \underline{\bar{H}}^T(t_{k+1}) + \underline{\bar{R}}(t_{k+1})]^{-1} \tag{5.65}$$

Several comments can be made at this point.

1. The optimal estimator is a linear function of the observations $\underline{y}(t_{k+1})$. Thus, the optimal estimator is a linear estimator in this case. Kalman and Bucy (1961), who essentially derived Eqs. (5.63) to (5.65), assumed a priori a linear type of estimator to arrive at the same results. In this case, one can argue that there might exist a nonlinear estimator that can do better in the least squares sense. In other words, if one makes the assumption of linearity, then the linear estimator might not be the conditional mean of the underlying distribution and $\underline{\bar{P}}(t)$ needs some other interpretation. The assumption of Gaussian statistics assures that the estimator is optimal over all linear and nonlinear estimators.

2. For the linear model treated, the covariance matrix P(t) is not a function of the observations. So one can do some covariance analysis without actual simulation. This matrix is very useful as an accuracy estimate.

3. The covariance update Eq. (5.64) can cause some problems, due to the fact that $\underline{\bar{P}}(t_{k+1}|t_{k+1})$ can become negative definite due to computational error, especially at the initial stages of the filter

operation where $\underline{\overline{P}}(t_{k+1}|t_k)$ and $\underline{\overline{K}}(t_{k+1})$ are expected to be large and $\underline{\overline{P}}(t_{k+1}|t_{k+1})$ is expected to be small. Direct algebraic calculations reveal that Eq. (5.64) is equivalent to

$$\underline{\overline{P}}(t_{k+1}|t_k) = [\underline{I}_{nn} - \underline{\overline{K}}(t_{k+1}) \cdot \underline{\overline{H}}(t_{k+1})] \cdot \underline{\overline{P}}(t_{k+1}|t_k) \cdot [\underline{I}_{nn} - \underline{\overline{K}}(t_{k+1}) \cdot \underline{\overline{H}}(t_{k+1})]^T + \underline{\overline{K}}(t_{k+1}) \cdot \underline{\overline{R}}(t_{k+1}) \cdot \underline{\overline{K}}^T(t_{k+1}) \qquad (5.66)$$

when $\underline{\overline{K}}(t_{k+1})$ is given by Eq. (5.65). In Eq. (5.66), there is no possibility for $\underline{\overline{P}}(t_{k+1}|t_k)$ to become negative definite since it is equal to the sum of two non-negative definite matrices. However, the computational burden associated with Eq. (5.66) is larger than that of Eq. (5.64).

4. During the derivation, the parameters $\underline{\overline{Q}}(t)$, $\underline{\overline{R}}(t_k)$ were assumed known for all times t, t_k. Also, the initial values of $\hat{\underline{x}}(t_k|t_k)$, $\underline{\overline{P}}(t_k|t_k)$ for $t_k = t_o$ were assumed given. If the values of these parameters used in the filter are not the true ones, then the resulting filter will not operate optimally (suboptimal filter). Sensitivity analysis is possible (Gelb (ed.), 1974) to determine the effect of the use of the wrong parameters in the filter performance.

5. Very often, one deals with scalar observations. This is commonly the case, for example, in streamflow forecasting, where observations of the river stage (or discharge) at the outlet of the catchment are available. In cases like this, the update equations can be simplified.

Suppose,

$$\underline{\overline{H}}(t_k) = [0 \quad 0 \quad . \; . \; . \quad h_n(t_k)] \quad (1xn) \tag{5.67}$$

that is the state $x_n(t)$ is observed at times t_k, $k = 1, 2, \ldots,$ and $\underline{v}(t_k)$, $\underline{\overline{R}}(t_k)$ are scalars.

Then,

$$\underline{\overline{H}}(t_{k+1}) \cdot \underline{\overline{P}}(t_{k+1}|t_k) \cdot \underline{\overline{H}}^T(t_{k+1}) + \underline{\overline{R}}(t_{k+1})$$

$$= h_n^2(t_{k+1}) \cdot p_{nn}(t_{k+1}|t_k) + R(t_{k+1}) \tag{5.68}$$

with $p_{nn}(t_{k+1}|t_k)$ denoting the $(n, n)^{th}$ element of $\underline{\overline{P}}(t_{k+1}|t_k)$. Also,

$$\underline{\overline{P}}(t_{k+1}|t_k) \cdot \underline{\overline{H}}^T(t_{k+1}|t_k) = \begin{vmatrix} p_{1n}(t_{k+1}|t_k) \\ p_{2n}(t_{k+1}|t_k) \\ \\ p_{nn}(t_{k+1}|t_k) \end{vmatrix} \cdot h_n(t_{k+1}|t_k) \tag{5.69}$$

Thus,

$$k_i(t_{k+1}) = \frac{h_n(t_{k+1}|t_k) \cdot p_{in}(t_{k+1}|t_k)}{h_n^2(t_{k+1}|t_k) \cdot p_{nn}(t_{k+1}|t_k) + R(t_{k+1})} \; ; \; i = 1, 2, \ldots, n \tag{5.70}$$

where $k_i(t_{k+1})$ is the i^{th} element of the vector (in the case of a single observation) gain $\underline{k}(t_{k+1})$ and $p_{i,j}(t_{k+1}|t_k)$ is the $(i, j)^{th}$ element of matrix $\underline{\overline{P}}(t_{k+1}|t_k)$.

It is worth noting that the correction to the i^{th} state is due to the cross correlation of this state and the observed state. Hence, if $p_{i,n}(t_{k+1}|t_k)$ is zero for some state i at some time t_{k+1}, then the gain $k_i(t_{k+1})$ is zero and from Eq. (5.63) and the fact that $\underline{\overline{K}}(t_{k+1})$ is a vector in this case, it follows that the updated estimate is the same as the propagated value at t_{k+1}.

Substitution of Eqs. (5.67) and (5.70) in Eq. (5.64) yields,

$$p_{i,j}(t_{k+1}|t_{k+1}) = p_{i,j}(t_{k+1}|t_k) - k_i(t_{k+1}) \cdot h_n(t_{k+1}) \cdot p_{n,j}(t_{k+1}|t_k);$$
$$i, j = 1, 2, \ldots, n \quad (5.71)$$

where $p_{i,j}(t_{k+1}|t_{k+1})$ is the $(i, j)^{th}$ element of matrix $\underline{\overline{P}}(t_{k+1}|t_{k+1})$.

If either $p_{i,n}(t_{k+1}|t_k)$ or $p_{n,j}(t_{k+1}|t_k)$ is zero, the updated covariance is equal to $p_{i,j}(t_{k+1}|t_k)$ (the propagated one).

6. Denote by $\underline{\nu}(t_{k+1})$ the one step ahead predicted residuals:

$$\underline{\nu}(t_{k+1}) = \underline{y}(t_{k+1}) - \underline{\overline{H}}(t_{k+1}) \cdot \underline{\hat{x}}(t_{k+1}|t_k) \quad (5.72)$$

The variance of $\underline{\nu}(t_{k+1})$ is given by Eq. (5.60). Direct calculation also shows that for the sequence of optimal gains (Eq. (5.65)), the residuals are uncorrelated in time, provided the parameters of the filter are correct (e.g., $\underline{\overline{Q}}(t)$, $\underline{\overline{R}}(t_k)$). Thus, it is often useful to examine the statistics of $\underline{\nu}(t_{k+1})$ after an application of the filter in real world, to verify if the values of $\underline{\overline{Q}}(t)$, $\underline{\overline{R}}(t_k)$ assumed are the true ones.

5.4 State Estimation

The statistically linearized flood routing model based on n nonlinear reservoirs is converted to a stochastic model of the type in Eqs. (5.46) and (5.47). At each time, the state $\underline{x}(t)$ of the system is decomposed into a mean term $\underline{\mu}_x(t)$ and a zero mean Gaussian residual process $\underline{r}_x(t)$, such that for each i (i = 1, 2, ..., n),

$$x_i(t) = \mu_{x_i}(t) + r_{x_i}(t) \tag{5.73}$$

while the parameters a_i, i = 1, 2, ..., n, are assumed constant. The mean term $\underline{\mu}_x(t)$ is treated as a deterministic quantity and is propagated according to the differential equations describing the motion of the system. The process $\underline{r}_x(t)$ is estimated by the minimum variance Gaussian estimator of the previous section. Thus, at the observation times, the a posteriori mean of $\underline{r}_x(t)$ is not necessarily zero.

Suppose the initial time is t_k. At this time, $\underline{x}(t_k)$ is decomposed in $\underline{\mu}_x(t_k)$ and $\underline{r}_x(t_k)$, where $\underline{r}_x(t_k)$ is zero mean. Both quantities are propagated in the interval (t_k, t_{k+1}) using the system equations. Thus, a priori, at time t_{k+1}, the mean of $\underline{x}(t)$ is $\underline{\mu}_x^a(t_{k+1})$ since, if the system equations are linear, $r_x(t_{k+1})$ has zero mean. At time t_{k+1}, an observation is available and is processed by the estimator to determine the a posteriori mean of $\underline{r}_x(t_{k+1})$ equal to $\hat{\underline{r}}_x(t_{k+1})$, not necessarily equal to zero. Thus, at this time step, the mean of $\underline{x}(t)$ is $\underline{\mu}_x^a(t_{k+1}) + \hat{r}_x(t_{k+1})$. As initial value $\underline{\mu}_x(t_{k+1})$ for the propagation of the mean in the time interval (t_{k+1}, t_{k+2}), the value $\underline{\mu}_x^a(t_{k+1}) + \hat{r}_x(t_{k+1})$ is used, while $\underline{r}_x(t)$ is again an a priori zero mean process. This cycle is repeated for all the observations.

The following formulation follows the above guidelines.

Denote by $\sigma_{x_i}^2(t)$ the variance of the i^{th} state at time t and by $V_{x_i}(t)$ the coefficient of variation of the same state,

$$V_{x_i}(t) = \frac{\sigma_{x_i}(t)}{\mu_{x_i}(t)} \quad ; \quad i = 1, 2, \ldots, n \tag{5.74}$$

It was found (Chapter 4) that the function

$$g(\mu_{x_i}, r_{x_i}) = a_i \cdot (\mu_{x_i}(t) + r_{x_i}(t))^m \tag{5.75}$$

can be approximated, in a least squares sense, by the function,

$$g_a(\mu_{x_i}, r_{x_i}) = N_{b_i}(t) \cdot \mu_{x_i}(t) + N_{x_i}(t) \cdot r_{x_i}(t) \tag{5.76}$$

The gains $N_{b_i}(t)$, $N_{x_i}(t)$ are given by Eqs. (4.50) and (4.51), respectively. Since in this development a_i is assumed known and constant, $\mu_{a_i}(t) = a_i$.

Substitution of Eq. (5.76) for all i, in the system of Eq. (5.1) results in

$$\frac{dx_i(t)}{dt} = p_i \cdot u(t) + N_{b_{i-1}}(t) \cdot \mu_{x_{i-1}}(t) + N_{x_{i-1}}(t) \cdot r_{x_{i-1}}(t)$$
$$- N_{b_i}(t) \cdot \mu_{x_i}(t) - N_{x_i}(t) \cdot r_{x_i}(t) \quad ; \quad i = 1, 2, \ldots, n \tag{5.77}$$

Taking the expected value of both sides of Eq. (5.77) yields

$$\frac{d}{dt}(\mu_{x_i}(t)) = p_i \cdot \mu_u(t) + N_{b_{i-1}}(t) \cdot \mu_{x_{i-1}}(t) - N_{b_i}(t) \cdot \mu_{x_i}(t) \; ;$$
$$i = 1, 2, \ldots, n \tag{5.78}$$

where $\mu_u(t)$ represents the expected value of u(t) and,

$$u(t) = \mu_u(t) + r_u(t) \tag{5.79}$$

with $r_u(t)$ a zero mean Gaussian process.

The equation describing the evolution of the residual processes $r_{x_i}(t)$, $i = 1, 2, \ldots, n$, in time, can be derived by subtracting Eq. (5.78) from Eq. (5.77),

$$\frac{d(r_{x_i}(t))}{dt} = N_{x_{i-1}}(t) \cdot r_{x_{i-1}}(t) - N_{x_i}(t) \cdot r_{x_i}(t) + p_i \cdot r_u(t) \; ;$$

$$i = 1, 2, \ldots, n \qquad (5.80)$$

The measurement equation of the system is $(k = 0, 1, 2, \ldots)$,

$$z(t_k) = a_n \cdot (\mu_{x_n}(t_k) + r_{x_n}(t_k))^m \qquad (5.81)$$

where $z(t_k)$ is the actual observation of the outflowing discharge at time t_k.

Statistical linearization of Eq. (5.81) gives (for each k)

$$z(t_k) = N_{b_n}(t_k) \cdot \mu_{x_n}(t_k) + N_{x_n}(t_k) \cdot r_{x_n}(t_k) \qquad (5.82)$$

or

$$z_r(t_k) = z(t_k) - N_{b_n}(t_k) \cdot \mu_{x_n}(t_k) = N_{x_n}(t_k) \cdot r_{x_n}(t_k) \qquad (5.83)$$

Equation (5.78) describes the propagation in time of the mean process of $x_i(t)$, $i = 1, 2, \ldots, n$.

Equation (5.80) is the propagation between observation times of the zero mean residual process $r_{x_i}(t)$, $i = 1, 2, \ldots, n$, with associated measurement equation (5.83). In matrix form,

$$\frac{d}{dt}[\underline{r}_x(t)] = \underline{\overline{F}}(t) \cdot \underline{r}_x(t) + \underline{\overline{B}} \cdot r_u(t) \tag{5.84}$$

and

$$z_r(t_k) = \underline{\overline{H}}(t_k) \cdot \underline{r}_x(t_k) \quad ; \quad k = 0, 1, 2, \ldots \tag{5.85}$$

where

$$\underline{\overline{F}}(t) = \begin{vmatrix} -N_{x_1}(t) & 0 & 0 & \cdot & \cdot & 0 \\ N_{x_1}(t) & -N_{x_2}(t) & 0 & \cdot & \cdot & 0 \\ 0 & N_{x_2}(t) & -N_{x_3}(t) & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & -N_{x_n}(t) \end{vmatrix} (n x n) \tag{5.86}$$

$$\underline{r}_x(t) = [r_{x_1}(t) \quad r_{x_2}(t) \ . \ . \ . \ r_{x_n}(t)]^T \tag{5.87}$$

$$\underline{\overline{B}} = [p_1 \quad p_2 \ . \ . \ . \ p_n]^T \tag{5.88}$$

$$\underline{\overline{H}}(t_k) = [0 \quad 0 \ . \ . \ . \ . \ N_{x_n}(t_k)] \tag{5.89}$$

Addition of driving noise terms converts Eqs. (5.84) and (5.85) to the form of Eqs. (5.46) and (5.47), respectively. Thus, the formulae of the previous section are directly applicable, provided that $r_u(t)$ is a white noise process and that at each time step the initial value of the estimate $\hat{\underline{r}}_x(t_k)$ is set to zero.

Equations (5.84) and (5.85) are linear equations with constant coefficients for the intervals (t_i, t_{i+1}) for i = 0, 1, 2. .... Using the results of Section 5.2, one can convert them into the equivalent canonical form. Addition of the discrete noise processes to the resultant equations gives the following stochastic system,

$$\underline{r}_y(t_{k+1}) = \underline{\overline{\Phi}}_r(t_{k+1}, t_k) \cdot \underline{r}_y(t_k) + \underline{\overline{\Gamma}}_k \cdot \underline{W}(t_k) \qquad (5.90)$$

$$z_r(t_{k+1}) = \underline{\overline{H}}'(t_k) \cdot \underline{r}_y(t_k) + v(t_k) \quad ; \quad k = 0, 1, 2, \ldots \qquad (5.91)$$

where $\underline{\overline{\Phi}}_r(t_{k+1}, t_k)$ is the diagonal transition matrix; $\underline{\overline{\Gamma}}_k \cdot \underline{W}(t_k)$ is a weighted discrete noise process that accounts also for the input uncertainty; $\underline{\overline{H}}'(t_k)$ is a 1xn matrix with nonzero elements (in general). $\underline{r}_y(t)$ is the canonical state related with $\underline{r}_x(t)$ by the transformation,

$$\underline{r}_x(t) = \underline{\overline{T}}_{r_k} \cdot \underline{r}_y(t) \quad ; \quad t_k \leq t \leq t_{k+1} \qquad (5.92)$$

with $\underline{\overline{T}}_{r_k}$ of the type in Eq. (5.18), provided that the eigenvalues of $\underline{\overline{F}}(t_k)$ are distinct for all k. Since Eq. (5.92) is linear, the mean and covariance matrix of $r_x(t)$ are given as $\underline{\overline{T}}_{r_k} \cdot E\{\underline{r}_y(t)\}$ and $\underline{\overline{T}}_{r_k} \cdot \underline{\overline{P}}_y(t) \cdot \underline{\overline{T}}_{r_k}^T$, respectively, where $\underline{\overline{P}}_y(t)$ is the covariance matrix of $\underline{r}_y(t)$ computed from the linear equation (5.90). The canonical form is particularly convenient to study the stability of the estimator. Deyst and Price (1968) showed that if a system is uniformly completely observable and controllable, the resultant filter is stable. That is, errors in the state estimate or in the error covariance matrix "die out".

In Appendix C, it is shown that when the elements of $\underline{\overline{H}}'(t_k)$, $\underline{\overline{\Phi}}_r(t_{k+1}, t_k)$ are well behaved (e.g., bounded and nonzero) for all times

t_k , and the covariances $R(t_k)$ and $\overline{\underline{Q}}(t_k)$ of the noise sequences have nonzero and bounded elements ($\overline{\underline{Q}}(t_k)$ is assumed diagonal and $\overline{\underline{\Gamma}}_k$ is diagonal with elements equal to unity). Then, the system of equations (5.93) and (5.94) is both uniformly completely controllable and observable.

Note that given that the filter is stable, the choice of the initial covariance matrix is not crucial in its performance. In addition, if the system is time invariant and the true covariance matrices $\overline{\underline{Q}}(t_k)$ and $R(t_k)$ are independent of time (stationary statistics), it can be stated that the filter will reach a steady state.

5.5 Simultaneous State and Parameter Estimation

It is desired to identify the value of the parameters that minimize a mean quadratic error criterion. Similar to the state estimation problem, a sequential estimator is sought, so that the observations are processed as they become available and thus, excessive computation storage requirements are avoided. Since, in most cases, the parameter estimator will operate in parallel to the state estimator, the problem is one of simultaneous state and parameter estimation (adaptive filtering).

Perhaps the most common estimator in situations like this is the extended Kalman filter, where the nonlinear multi-dimensional functions involved are linearized about a reference trajectory. Linearization through truncation of a Taylor series expansion results in a bias in the expected value of the function involved, introducing an error source that can cause divergence of the estimator from the true value or it can give biased estimates (Ljung, 1979).

As an alternative to ordinary linearization, statistical linearization has not found widespread usage in parameter estimation, even though it gives unbiased expected values for the nonlinear functions. The major reason for this is that one has to calculate the expected value of a nonlinear function in order to determine the gains of linearization (Chapter 4). The common numerical integration solution is immediately rejected due to the numerical burden associated with the evaluation of multi-dimensional integrals at each iteration step.

To overcome this problem, the approximations developed in Chapter 3 will be used in the statistical linearization of the nonlinear functions involved. The adaptive filtering problem will be converted to a state estimation one, by the introduction of differential equations to describe the dynamics of the parameters.

If the standard assumption of the states and the parameters being approximately Gaussian is used, then

$$x_i(t) = \mu_{x_i}(t) + r_{x_i}(t) \quad ; \quad i = 1, 2, \ldots, n \qquad (5.93)$$

$$a_i(t) = \mu_{a_i}(t) + r_{a_i}(t) \quad ; \quad i = 1, 2, \ldots, n \qquad (5.94)$$

with $\mu_{x_i}(t)$, $\mu_{a_i}(t)$ representing mean processes and $r_{x_i}(t)$, $r_{a_i}(t)$ zero mean normally distributed residual processes with variances $\sigma^2_{x_i}(t)$ and $\sigma^2_{a_i}(t)$, respectively. The covariance between x_i and a_i is denoted by $\sigma^2_{x_i a_i}(t)$. $x_i(t)$ and $a_i(t)$, $i = 1, 2, \ldots, n$, are the states and the parameters, respectively, of a flood routing model based on n nonlinear reservoirs, of the type discussed in Chapter 2.

Statistical linearization of the generic function $a_i(t) \cdot x_i^m(t)$, where m is a real number in the interval [0.8, 1.6] gives (Chapter 4)

$$a_i(t) \cdot x_i^m(t) = N_{b_i}(t) \cdot \mu_{a_i}(t) + N_{b_i}(t) \cdot \mu_{x_i}(t) + N_{x_i}(t) \cdot r_{x_i}(t) + N_{a_i}(t) \cdot r_{a_i}(t) \qquad (5.95)$$

where $N_{b_i}(t)$, $N_{x_i}(t)$ and $N_{a_i}(t)$ are given in Eqs. (4.47), (4.48) and (4.49), respectively, with

$$V_{x_i}(t) = \frac{\sigma_{x_i}(t)}{\mu_{x_i}(t)} \quad ; \quad i = 1, 2, \ldots, n \qquad (5.96)$$

$$V_{a_i}(t) = \frac{\sigma_{a_i}(t)}{\mu_{a_i}(t)} \quad ; \quad i = 1, 2, \ldots, n \qquad (5.97)$$

and $\rho_{x_i a_i}(t)$ represents the cross-correlation between the i^{th} state and the i^{th} parameter at time t. In the development to follow, it is assumed that the means are nonzero.

The differential equation governing the motion of the i^{th} state of the system is Eq. (5.1). Substitution of Eq. (5.95) in Eq. (5.1) yields,

$$\frac{d(\mu_{x_i}(t)+r_{x_i}(t))}{dt} = p_i \cdot \mu_u(t) + p_i \cdot r_u(t) + N_{b_{i-1}}(t) \cdot \mu_{a_{i-1}}(t) + N_{b_{i-1}}(t) \cdot \mu_{x_{i-1}}(t) + N_{a_{i-1}}(t) \cdot r_{a_{i-1}}(t) + N_{x_{i-1}}(t) \cdot r_{x_{i-1}}(t) - N_{b_i}(t) \cdot \mu_{a_i}(t) - N_{b_i}(t) \cdot \mu_{x_i}(t) - N_{x_i}(t) \cdot r_{x_i}(t) - N_{a_i}(t) \cdot r_{a_i}(t); \quad i = 1, 2, \ldots, n \qquad (5.98)$$

where u(t) is considered to be the sum of the mean process $\mu_u(t)$ and the zero mean Gaussian white noise $r_u(t)$ with variance parameter $\sigma_u^2(t)$.

The expected value of Eq. (5.98) is

$$\frac{d(\mu_{x_i}(t))}{dt} = p_i \cdot \mu_u(t) + N_{b_{i-1}}(t) \cdot \mu_{a_{i-1}}(t) + N_{b_{i-1}}(t) \cdot \mu_{x_{i-1}}(t)$$

$$- N_{b_i}(t) \cdot \mu_{a_i}(t) - N_{b_i}(t) \cdot \mu_{x_i}(t) \quad ; \quad i = 1, 2, \ldots, n \qquad (5.99)$$

Subtraction of Eq. (5.99) from Eq. (5.98) results in:

$$\frac{d(r_{x_i}(t))}{dt} = N_{x_{i-1}}(t) \cdot r_{x_{i-1}}(t) + N_{a_{i-1}}(t) \cdot r_{a_{i-1}}(t)$$

$$- N_{x_i}(t) \cdot r_{x_i}(t) - N_{a_i}(t) \cdot r_{a_i}(t) + p_i \cdot r_u(t) \; ;$$

$$i = 1, 2, \ldots, n \qquad (5.100)$$

Equations (5.99) and (5.100) hold for $N_{x_o}(t) = 0$ and $N_{a_o}(t) = 0$ for all t.

Constant parameters obey a differential equation of the type

$$\frac{d(a_i(t))}{dt} = 0 \quad ; \quad i = 1, 2, \ldots, n \qquad (5.101)$$

Consequently,

$$\frac{d(\mu_{a_i}(t))}{dy} = 0 \quad ; \quad i = 1, 2, \ldots, n \qquad (5.102)$$

$$\frac{d(r_{a_i}(t))}{dt} = 0 \quad ; \quad i = 1, 2, \ldots, n \qquad (5.103)$$

The system of equations (5.100) and (5.103) can be rewritten in a matrix notation as

$$\frac{d}{dt} (\underline{r}(t)) = \overline{\underline{N}}(t) \cdot \underline{r}(t) + \underline{p} \cdot r_u(t) \qquad (5.104)$$

where

$$\underline{r}(t) = [\underline{r}_x^T(t) \;\vdots\; \underline{r}_a^T(t)]^T \qquad (5.105)$$

$$\underline{r}_x(t) = [r_{x_1}(t) \quad r_{x_2}(t) \quad . \; . \; . \; r_{x_n}(t)]^T \qquad (5.106)$$

$$\underline{r}_a(t) = [r_{a_1}(t) \quad r_{a_2}(t) \quad . \; . \; . \; r_{a_n}(t)]^T \qquad (5.107)$$

$$\underline{p} = [p_1 \quad p_2 \; . \; . \; . \; p_n]^T \qquad (5.108)$$

$$\overline{\underline{N}}(t) = \left| \begin{array}{c:c} \overline{\underline{N}}_{r_x}(t) & \overline{\underline{N}}_{r_a}(t) \\ \hdashline \overline{\underline{0}}_{nn} & \overline{\underline{0}}_{nn} \end{array} \right| \quad (2nx2n) \qquad (5.109)$$

with $\overline{\underline{0}}_{nn}$ representing the (nxn) matrix whose elements are all zero.

$$\overline{\underline{N}}_{r_x}(t) = \begin{vmatrix} -N_{x_1}(t) & 0 & 0 & . & . & 0 \\ N_{x_1}(t) & -N_{x_2}(t) & 0 & . & . & 0 \\ 0 & N_{x_2}(t) & -N_{x_3}(t) & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & -N_{x_n}(t) \end{vmatrix} \quad (nxn) \qquad (5.110)$$

$$\underline{\overline{N}}_{r_a}(t) = \begin{vmatrix} -N_{a_1}(t) & 0 & 0 & \cdot & \cdot & 0 \\ N_{a_1}(t) & -N_{a_2}(t) & 0 & \cdot & \cdot & 0 \\ 0 & N_{a_2}(t) & -N_{a_3}(t) & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & -N_{a_n}(t) \end{vmatrix} \quad (nxn) \tag{5.111}$$

It is assumed that the observations of the output of the system at some time t_k, $k = 0, 1, 2, \ldots$, consist of a single measurement of the discharge at the outlet of the catchment of interest, at time t_k, $k = 0, 1, 2, \ldots$ It follows,

$$z(t_k) = N_{b_n}(t_k) \cdot \mu_{x_n}(t_k) + N_{b_n}(t_k) \cdot \mu_{a_n}(t_k) + N_{x_n}(t_k) \cdot r_{x_n}(t_k) + N_{a_n}(t_k) \cdot r_{a_n}(t_k) \tag{5.112}$$

where $z(t_k)$, $k = 0, 1, 2, \ldots$, represents the sequence of measurements.

Denote by $z_r(t_k)$ the function defined as

$$z_r(t_k) = z(t_k) - N_{b_n}(t_k) \cdot \mu_{x_n}(t_k) - N_{b_n}(t_k) \cdot \mu_{a_n}(t_k) \tag{5.113}$$

Then, in matrix notation,

$$z_r(t_k) = \underline{\overline{H}}(t_k) \cdot \underline{r}(t_k) \tag{5.114}$$

where

$$\underline{\overline{H}}(t_k) = [\underline{\overline{H}}_x(t_k) \; \vdots \; \underline{\overline{H}}_a(t_k)] \quad (1x2n) \tag{5.115}$$

$$\underline{\overline{H}}_x(t_k) = [0 \quad 0 \quad . . . \quad N_{x_n}(t_k)] \qquad (1xn) \qquad (5.116)$$

$$\underline{\overline{H}}_a(t_k) = [0 \quad 0 \quad . . . \quad N_{a_n}(t_k)] \qquad (1xn) \qquad (5.117)$$

Equations (5.104) and (5.114) are the system equations in matrix form. At this point, white noise, $\underline{w}'(t)$, is added to Eq. (5.104) having diagonal intensity (or variance parameter) matrix equal to $\underline{\overline{Q}}'(t)$, to account for errors in the model structure and in the value of m. Since the intensity of the white model error is not predetermined by any law and $r_u(t)$ is also modeled as a white noise process, the two are lumped into a single noise process $\underline{w}(t)$ multiplied by a weighting matrix $\underline{\overline{G}}(t)$. In the implementation of the filter equations, the variance parameter $\underline{\overline{Q}}(t)$ of this noise is taken as $\underline{p}_i \cdot \underline{p}_i^T \cdot \sigma_u^2(t) + \underline{\overline{Q}}'(t)$.

The discrete time white noise $v(t_k)$, k = 0, 1, 2, ..., is added in Eq. (5.114) to account for model and measurement schedule deficiencies. Its variance is denoted by $R(t_k)$, k = 0, 1, 2, ... Thus,

$$\frac{d}{dt}(r(t)) = \underline{\overline{N}}(t) \cdot \underline{r}(t) + \underline{\overline{G}}(t) \cdot \underline{w}(t) \qquad (5.118)$$

where $\underline{\overline{G}}(t)$ is a weighting matrix, and

$$z_r(t_k) = \underline{\overline{H}}(t_k) \cdot \underline{r}(t_k) + v(t_k) \qquad (5.119)$$

Use of the Gaussian minimum variance filter with Eqs. (5.118) and (5.119) results in the following algorithm:

State Estimate Propagation

$$\frac{d}{dt}(\hat{\underline{r}}(t|t_k)) = \underline{\overline{N}}(t) \cdot \hat{\underline{r}}(t|t_k) \qquad (5.120)$$

with initial condition

$$\underline{\hat{r}}(t_o|t_k) = \underline{\hat{r}}(t_k|t_k) \qquad (5.121)$$

Error Covariance Propagation

$$\frac{d}{dt}(\underline{\bar{P}}(t|t_k)) = \underline{\bar{N}}(t) \cdot P(t|t_k) + P(t|t_k) \cdot \underline{\bar{N}}^T(t) + \underline{\bar{G}}(t)\ \underline{\bar{Q}}(t)\ \underline{\bar{G}}^T(t) \qquad (5.122)$$

with initial condition

$$\underline{\bar{P}}(t_o|t_k) = \underline{\bar{P}}(t_k|t_k) \qquad (5.123)$$

State Estimate Update:

$$\underline{\hat{r}}(t_{k+1}|t_{k+1}) = \underline{\hat{r}}(t_{k+1}|t_k) + \underline{k}(t_{k+1}) \cdot \nu(t_{k+1}) \qquad (5.124)$$

Innovations Sequence:

$$\nu(t_{k+1}) = z_r(t_{k+1}) - \underline{\bar{H}}(t_{k+1}) \cdot \underline{\hat{r}}(t_{k+1}|t_k) \qquad (5.125)$$

Error Covariance Update:

$$\underline{\bar{P}}(t_{k+1}|t_{k+1}) = (\underline{\bar{I}}_{2n,2n} - \underline{k}(t_{k+1}) \cdot \underline{\bar{H}}(t_{k+1})) \cdot \underline{\bar{P}}(t_{k+1}|t_k) \qquad (5.126)$$

where $\underline{\bar{I}}_{2n,2n}$ is the (2nx2n) dimensional unit matrix.

Gain Matrix:

$$\underline{k}(t_{k+1}) = \underline{\bar{P}}(t_{k+1}|t_k)\ \underline{\bar{H}}^T(t_{k+1}) \cdot Q_\nu^{-1}(t_{k+1}) \qquad (5.127)$$

Variance of the Innovations:

$$Q_\nu(t_{k+1}) = \underline{\bar{H}}(t_{k+1}) \cdot \underline{\bar{P}}(t_{k+1}|t_k) \cdot H^T(t_{k+1}) + R(t_{k+1}) \qquad (5.128)$$

The error covariance matrix can be decomposed as follows:

$$\underline{\bar{P}}(t) = \begin{vmatrix} \underline{\bar{P}}_x(t) & \vdots & \underline{\bar{P}}_{xa}(t) \\ \cdots & \cdots & \cdots \\ \underline{\bar{P}}_{ax}(t) & \vdots & \underline{\bar{P}}_a(t) \end{vmatrix} \quad (2nx2n) \qquad (5.129)$$

Equation (5.129) can be substituted in the filter equations to reduce the dimensionality of the matrices involved. This results in:

Error Covariance Propagation:

$$\frac{d}{dt}(\underline{\bar{P}}_x(t|t_k)) = \underline{\bar{N}}_x(t) \cdot \underline{\bar{P}}_x(t|t_k) + \underline{\bar{N}}_a(t) \cdot \underline{\bar{P}}^T_{xa}(t|t_k) + \underline{\bar{P}}_x(t|t_k) \cdot \underline{\bar{N}}^T_x(t)$$

$$+ \underline{\bar{P}}_{xa}(t|t_k) \cdot \underline{\bar{N}}^T_a(t) + \underline{\bar{G}}_x(t) \cdot \underline{\bar{Q}}_x(t) \, \underline{\bar{G}}^T_x(t) \qquad (5.130)$$

with initial condition,

$$\underline{\bar{P}}_x(t_o|t_k) = \underline{\bar{P}}_x(t_k|t_k) \qquad (5.131)$$

where $\underline{\bar{Q}}_x(t)$ is the variance parameter of the portion of the white noise $\underline{w}(t)$ associated with $\underline{r}_x(t)$.

$$\frac{d}{dt}(\underline{\bar{P}}_{xa}(t|t_k)) = \underline{\bar{N}}_x(t) \cdot \underline{\bar{P}}_{xa}(t|t_k) + \underline{\bar{N}}_a(t) \cdot \underline{\bar{P}}_a(t|t_k) \qquad (5.132)$$

with initial condition,

$$\underline{\bar{P}}_{xa}(t_o|t_k) = \underline{\bar{P}}_{xa}(t_k|t_k) \qquad (5.133)$$

where

$$\underline{\overline{P}}_{ax}(t|t_k) = \underline{\overline{P}}^T_{xa}(t|t_k) \qquad \text{for all } t \tag{5.134}$$

and

$$\frac{d}{dt}[\underline{\overline{P}}_a(t|t_k)] = \underline{\overline{G}}_a(t) \cdot \underline{\overline{Q}}_a(t) \cdot \underline{\overline{G}}^T_a(t) \tag{5.135}$$

with initial condition,

$$\underline{\overline{P}}_a(t_o|t_k) = \underline{\overline{P}}_a(t_k|t_k) \tag{5.136}$$

where $\underline{\overline{Q}}_a(t)$ is the variance parameter of the portion of the white noise $\underline{w}(t)$ associated with $\underline{r}_a(t)$.

Error Covariance Update

$$\underline{\overline{P}}_x(t_{k+1}|t_{k+1}) = (\underline{\overline{I}}_{nn} - \underline{k}_x(t_{k+1}) \cdot \underline{\overline{H}}_x(t_{k+1})) \cdot \underline{\overline{P}}_x(t_{k+1}|t_k) - \underline{k}_x(t_{k+1}) \cdot \underline{\overline{H}}_a(t_{k+1}) \cdot \underline{\overline{P}}^T_{xa}(t_{k+1}|t_k) \tag{5.137}$$

$$\underline{\overline{P}}_{xa}(t_{k+1}|t_{k+1}) = (\underline{\overline{I}}_{nn} - \underline{k}_x(t_{k+1}) \cdot \underline{\overline{H}}_x(t_{k+1})) \cdot \underline{\overline{P}}_{xa}(t_{k+1}|t_k) - \underline{k}_x(t_{k+1}) \cdot \underline{\overline{H}}_a(t_{k+1}) \cdot \underline{\overline{P}}_a(t_{k+1}|t_k) \tag{5.138}$$

$$\underline{\overline{P}}_{ax}(t_{k+1}|t_{k+1}) = \underline{\overline{P}}^T_{xa}(t_{k+1}|t_{k+1}) \tag{5.139}$$

$$\underline{\overline{P}}_a(t_{k+1}|t_{k+1}) = - \underline{k}_a(t_{k+1}) \cdot \underline{\overline{H}}_x(t_{k+1}) \cdot \underline{\overline{P}}_{xa}(t_{k+1}|t_k) + (\underline{\overline{I}}_{nn} - \underline{k}_a(t_{k+1}) \cdot \underline{\overline{H}}_a(t_{k+1})) \cdot \underline{\overline{P}}_a(t_{k+1}|t_k) \tag{5.140}$$

Gain Vectors

$$\underline{k}_x(t_{k+1}) = [\underline{\bar{P}}_x(t_{k+1}|t_k) \cdot \underline{\bar{H}}_x^T(t_{k+1}) + \underline{\bar{P}}_{xa}(t_{k+1}|t_k) \cdot \underline{\bar{H}}_a^T(t_{k+1})] \cdot Q_\nu^{-1}(t_{k+1}) \tag{5.141}$$

$$\underline{k}_a(t_{k+1}) = [\underline{\bar{P}}_{xa}^T(t_{k+1}|t_k) \cdot \underline{\bar{H}}_x^T(t_{k+1}) + \underline{\bar{P}}_a(t_{k+1}|t_k) \cdot \underline{\bar{H}}_a^T(t_{k+1})] \cdot Q_\nu^{-1}(t_{k+1}) \tag{5.142}$$

State Estimate Update

$$\hat{\underline{r}}_x(t_{k+1}|t_{k+1}) = \hat{\underline{r}}_x(t_{k+1}|t_k) + \underline{k}_x(t_{k+1}) \cdot \nu(t_{k+1}) \tag{5.143}$$

$$\hat{\underline{r}}_a(t_{k+1}|t_{k+1}) = \hat{\underline{r}}_a(t_{k+1}|t_k) + \underline{k}_a(t_{k+1}) \cdot \nu(t_{k+1}) \tag{5.144}$$

where $\nu(t_{k+1})$ is the innovations sequence.

The assumption (for the calculation of the statistical linearization gains) that at the beginning of each propagation step (t_k, t_{k+1}), the residual processes $\underline{r}_x(t)$, $\underline{r}_a(t)$ have zero mean, leads one to add the biases $\hat{\underline{r}}_x(t_k|t_k)$, $\hat{\underline{r}}_a(t_k|t_k)$, resulting from the filter operation, to the mean values, $\underline{\mu}_x(t_k)$ and $\underline{\mu}_a(t_k)$, from the solution of Eqs. (5.99) and (5.102). Therefore, at each propagation step, the zero mean assumption is not violated. This simplifies the equations of the state estimates in the filter algorithm to the following form:

State Estimate Update

$$\hat{\underline{r}}_x(t_{k+1}|t_{k+1}) = \underline{k}_x(t_{k+1}) \cdot z_r(t_{k+1}) \tag{5.145}$$

$$\underline{\hat{r}}_a(t_{k+1}|t_{k+1}) = \underline{k}_a(t_{k+1}) \cdot z_r(t_{k+1}) \qquad (5.146)$$

where $\hat{r}(t_{k+1}|t_k)$ has been set equal to zero by means of Eq. (5.119), given that Eq. (5.121) has been modified to the following,

$$\underline{\hat{r}}(t_o|t_k) = 0 \qquad (5.147)$$

If the initial values of $\underline{\overline{P}}_a(t)$ and $\underline{\overline{P}}_{xa}(t)$ are equal to $\underline{\overline{0}}_{nn}$, then the algorithm reduces to the state estimation algorithm with the parameters being constant, equal to their initial value.

Chapter 6

CASE STUDY

6.1 Introduction

The use of the Gaussian minimum variance estimator with the statistically linearized, nonlinear flood routing model, is illustrated in this chapter. Off-line procedures to obtain parameter values are utilized, based on input-output discharge data. Given the parameters, the state estimator is used to forecast six-hour discharge values for the Bird Creek drainage basin in Oklahoma, U.S.A. Subsequently, the simultaneous state and parameter estimation scheme presented in the previous chapter is used to refine crude parameter estimates obtained from off-line estimation procedures.

Evaluation of the stochastic model performance is based on the mean squared error criterion and on the agreement between predicted and observed hydrograph characteristics (e.g. time to the peak discharge, magnitude of the peak discharge).

6.2 Available Data

The Bird Creek drainage basin, near Sperry, Oklahoma, has a drainage area of 2344 km^2 at the point of flow measurement (USGS Station No. 07177500). Six-hour discharge records were provided by the National Weather Service, Office of Hydrology, for the October 1955 to September 1962 period. The average discharge for this period is of the order of 20 m^3/sec, with maximum recorded discharge of 2535.11 m^3/sec and a minimum of 0.22 m^3/sec. Long spells with near zero discharges interrupted by sharp, high peak, hydrographs occurring, mainly, in the period from May to September are characteristic of the discharge records. Flows rising from a near zero level to 200-250 m^3/sec in 18 hours are common in the records. This kind of behavior (small base-flow contribution) suggests that the flood routing scheme will be the crucial performance regulator for any rainfall-runoff simulation model.

The soil moisture accounting scheme of the National Weather Service River Forecast System (NWSRFS) model was utilized to obtain the spatially lumped channel inflow hydrographs to be routed through the channel. Descriptions of the NSWRFS model are given in NWS HYDRO-31 (1976) and in Kitanidis and Bras (1978). Calibration was performed by the N.W.S. Hydrologic Research Laboratory. Parameters of the Bird Creek calibration are available in Georgakakos and Bras (1979).

The months of May and July of the water year 1959 were selected as test periods. Real time forecasting of six-hourly discharge values using the state estimator is done for these months. The parameter values used are based on preliminary simulation runs. Rough initial estimates are obtained utilizing the results of section 2.5. The on-line parameter estimation procedure suggested in Chapter 5 is examined during the second half of the month of May, 1959.

In each of the selected months, successive flood events produced sharp hydrographs with peak discharges separated by time periods ranging from four to six days. The maximum recorded peak discharge in May is 254 m^3/sec and in July is 286 m^3/sec. In both cases the minimum discharge is less than 0.5 m^3/sec.

6.3 State Estimation Runs

In this section the parameters of the model are held constant and the Gaussian minimum variance estimator is used to forecast the discharge with a six-hour lead time.

It is necessary to obtain values for the number n of the cascaded reservoirs, for the parameters a_i, i = 1, 2, .., n, for the exponent m and for the input distribution p_i, i = 1, 2, .., n. The test period is characterized by sharp hydrographs. Reshaping the hydrographs into triangular form, preserving the volumes under the rising and under the falling limbs, gives ratios $\frac{t_p}{t_r}$ (where t_p is the time to peak and t_r the time duration of the falling limb of the equivalent triangular hydrographs) of the order of one. Similarly, the channel inflow hydrographs, produced by the NWSRFS model have characteristic times of the same magnitude (their time ratio is about one). Section 2.5 showed that the lower the exponent m of the routing model is, the bigger the ratio $\frac{t_p}{t_r}$ of the output hydrograph will be (Section 2.5). This suggests small values for m. An m value of 0.8 is adopted for all reservoirs. The number of reservoirs is initially set to 3. To obtain a long lag time between input and output the input distribution was taken as p_1 = .95, p_2 = .05, p_3 = 0.

Rough parameter estimates result from the procedures discussed in section 2.5. Suppose that the parameter a_i, of the i^{th} reservoir is to be determined, given the quantities, $Q_{p_{i-1}}$, $t_{p_{i-1}}$, $\frac{t_{p_{i-1}}}{t_{r_{i-1}}} = k_{i-1}$. $Q_{p_{i-1}}$ represents the peak discharge of the hydrograph output from the $i\text{-}1^{\underline{th}}$ reservoir that serves as an input to the i^{th} reservoir.

From available data define a ratio, $\frac{Q_{p_i}}{Q_{p_{i-1}}}$.

From equation (B.34)

$$\Delta t_i = t_{p_{i-1}} \cdot \left(\frac{1 - \frac{Q_{p_i}}{Q_{p_{i-1}}}}{k_{i-1}} \right) \tag{6.1}$$

where Δt_i is the time difference of the input and output hydrograph peaks.

The value of a_i that gives the desired ratio $\frac{Q_{p_i}}{Q_{p_{i-1}}}$ can now be determined from graphs similar to the one in Figure 2.3, that correspond to given values of $Q_{p_{i-1}}$, $t_{p_{i-1}}$, $\frac{t_{p_{i-1}}}{t_{r_{i-1}}}$.

The ratio $\frac{t_{p_i}}{t_{r_i}}$ to be used for the next reservoir can be obtained from plots similar to the one in Figure 2.5 for the given characteristics of the input hydrograph and the ratio $\frac{Q_{p_i}}{Q_{p_{i-1}}}$.

The time to peak t_{p_i} of the output hydrograph to be used in the determination of Δt_{i+1} from Equation (6.1) is obtained by equating the input and output hydrograph volumes,

$$t_{p_i} + t_{r_i} = \frac{Q_{p_{i-1}}}{Q_{p_i}} \cdot (t_{p_{i-1}} + t_{r_{i-1}}) \qquad (6.2)$$

Given the ratio $\frac{t_{p_i}}{t_{r_i}}$, Equation (6.2) can be solved for the value of t_{p_i}.

The previously outlined procedure is then repeated to find parameter a_{i+1} of the $i + 1^{\underline{th}}$ reservoir.

In this section, parameter values are determined based on the flood of July 15, 1959. The input peak discharge is $Q_{p_o} = 850\ m^3/sec$. The output peak discharge is $Q_{p_n} \cong 280\ m^3/sec$. Also, $t_{p_o} = t_{r_o} = 9.5$ hours. The total time difference between input and output peak discharges, Δt_T, is 24 hours.

(a) For the first reservoir, set $\frac{Q_{p_1}}{Q_{p_o}} = 0.6$. Then

$Q_{p_1} = 510\ m^3/sec$, and $\Delta t_1 = 3.8$ hours $(9.5 \times \frac{0.4}{1.0})$.

From Figure 2.3 one obtains $a_1 \approx 10^{-3}$ and from Figure 2.5, $\frac{t_{p_1}}{t_{r_1}} = 0.45$, $t_{p_1} + t_{r_1} = \frac{1}{0.6} \times 19 = 31.7$ hours,

so $t_{p_1} \cong 10$ hours.

(b) For the second reservoir, set $\frac{Q_{p_2}}{Q_{p_1}} = 0.6$. Then

$$Q_{p_2} = 306 \text{ m}^3/\text{sec} \;,\; \Delta t_2 = 10 \times \frac{0.4}{0.45} \simeq 9 \text{ hours},$$

$$a_2 \approx 10^{-3} \;,\; \frac{t_{p_2}}{t_{r_2}} = 0.45.$$

Also $t_{p_2} + t_{r_2} = \frac{1}{0.6} \times 31.7 \cong 53$ hours

so $t_{p_2} \cong 16.5$ hours

(c) For the third reservoir, set $\frac{Q_{p_3}}{Q_{p_2}} = 0.7$ then

$$Q_{p_3} = 214 \text{ m}^3/\text{sec}, \; \Delta t_3 = 16.5 \times \frac{0.3}{0.45} \cong 11 \text{ hours}, \; a_3 \cong 10^{-3}.$$

Thus, the output peak discharge after a series of three reservoirs is 214 m^3/sec and the lag time between input and output peak discharge is $\Delta t_T' = \Delta t_1 + \Delta t_2 + \Delta t_3 \cong 24$ hours, when $m = 0.8$ and $a_1 = a_2 = a_3 = 10^{-3}$. Clearly, the values of a_1, a_2, a_3 obtained are rough estimates since the plots in Figures 2.3 and 2.5 were made for an input discharge of 850 m^3/sec and $t_{p_o} = t_{r_o} = 9.5$ hours. Consequently, the use of these plots for different characteristics of the input hydrograph (e.g. second and third reservoirs) is not theoretically justifiable. Obviously, there has been some trial and error procedure to arrive at the results presented above.

Emphasis has been given to preservation of the correct time lag between input and output peaks.

Realizing that the values $a_1 = a_2 = a_3 = 10^{-3}$ are only crude estimates of the parameters, simulation studies were used to improve them. In the simulation runs, a quadratic error criterion was used to guide parameter selection. The 3-reservoir model equations were integrated using a fourth order Runge integration scheme, for the period from the 913th time step up to the 960th time step in May, 1959 which is the flood period in this month. Each parameter took two values: 10^{-3}, 5×10^{-4} and the average quadratic error was evaluated at the eight vertices of the cube in Figure 6.1. The coordinates of the vertices are given in Table 6.1. The corresponding average squared error is given in Table 6.2. It can be seen that the error is less on the plane determined by the vertices 3, 5, and 7.

For state estimation, the parameters corresponding to vertex 7 were used. That is, $a_1 = 5 \times 10^{-4}$, $a_2 = a_3 = 10^{-3}$. Figures 6.2 and 6.3 show the observed (solid lines) and off-line forecasted (dashed lines) flows for May and July, 1959, respectively. The numbers on the lower part of the figures correspond to the number of the six-hour period in the water year 1959. Thus, 1093 is the first six hour period in the first day of July, 1959.

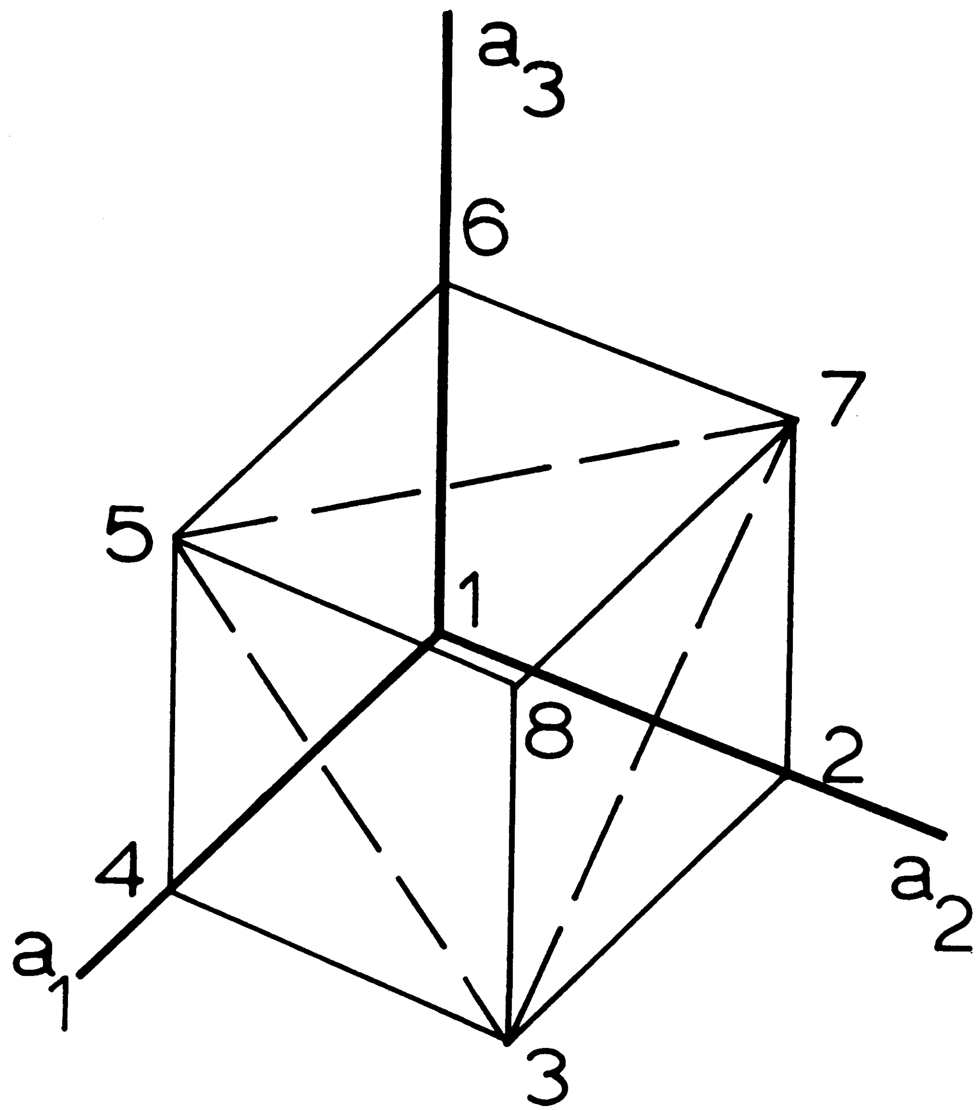

Figure 6.1

LOCUS OF PARAMETERS OF SIMULATION RUNS IN PARAMETER SPACE

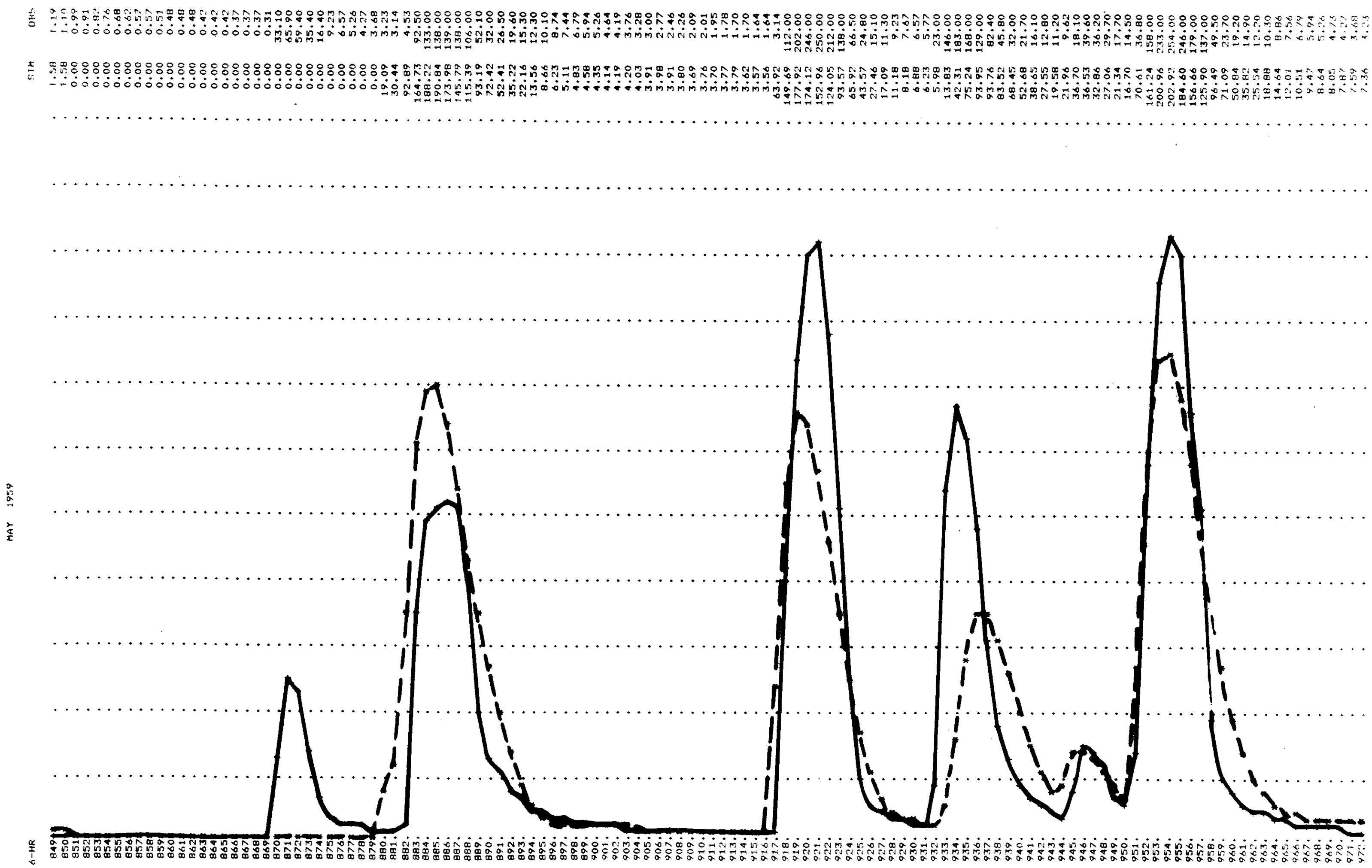

Figure 6.2

OFF-LINE RUN. SIX-HOUR LEAD FORECASTS FOR MAY, 1959. DASHED LINE CORRESPONDS TO THE FORECASTED VALUES. SOLID LINE CORRESPONDS TO THE OBSERVED VALUES - m^3/sec

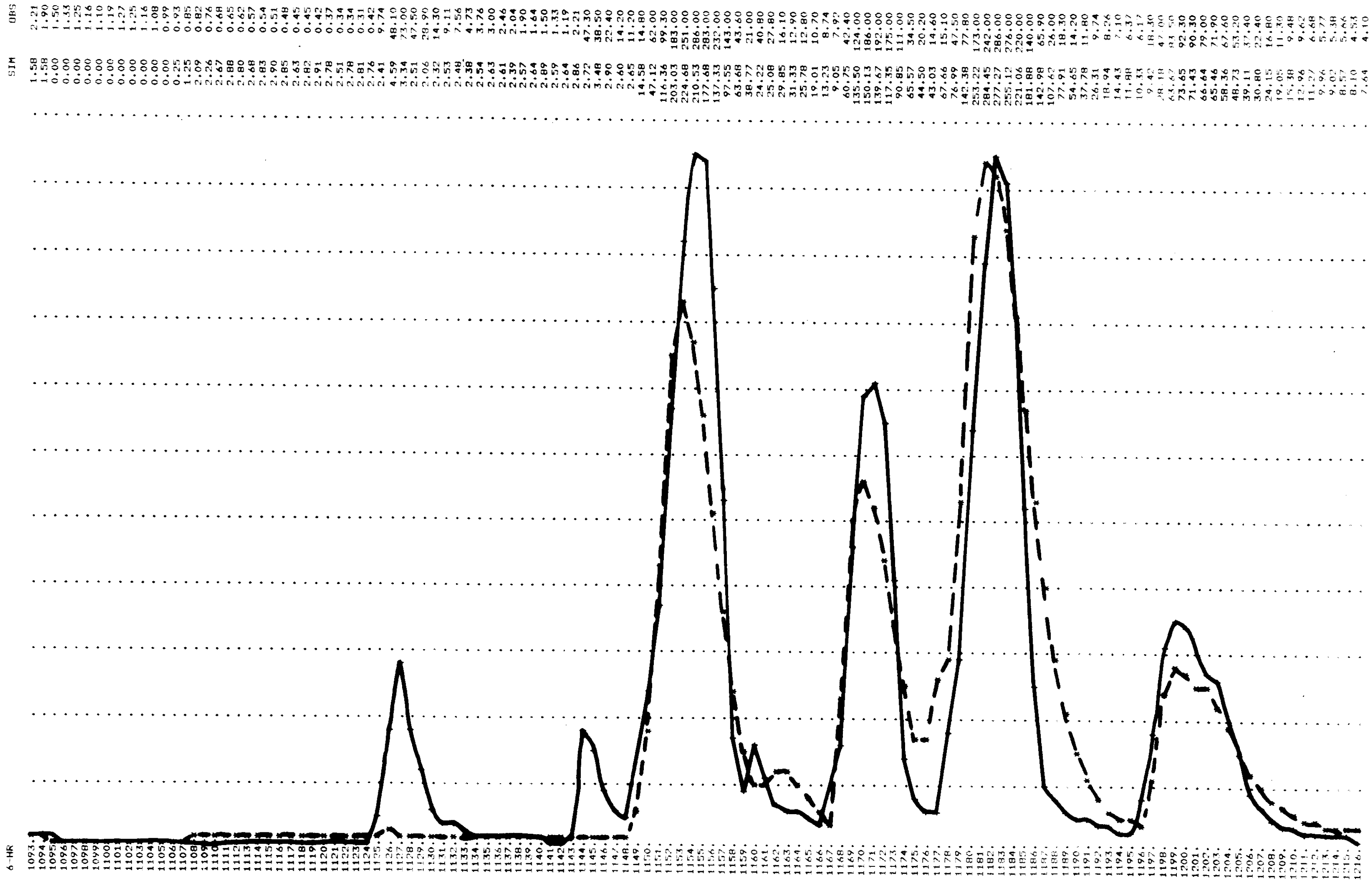

Figure 6.3

OFF-LINE RUN. SIX-HOUR LEAD FORECASTS FOR JULY, 1959. DASHED LINE CORRESPONDS TO THE FORECASTED VALUES. SOLID LINE CORRESPONDS TO THE OBSERVED VALUES - m^3/sec

Table 6.1

PARAMETER SPACE COORDINATES USED FOR SIMULATION RUNS

Vertex Number	$a_1 (\times 10^4)$	$a_2 (\times 10^4)$	$a_3 (\times 10^4)$
1	5	5	5
2	5	10	5
3	10	10	5
4	10	5	5
5	10	5	10
6	5	5	10
7	5	10	10
8	10	10	10

Table 6.2

AVERAGE QUADRATIC ERROR FOR SIMULATION RUNS

Vertex Number	Error (m^6/sec^2)
1	4402
2	2995
3	1967
4	3086
5	1974
6	2995
7	1993
8	2343

Characteristics of these figures are:

(a) The model tends to underestimate the peak discharges.

(b) The forecasted hydrographs tend to peak earlier than observed hydrographs (by about 6 hours).

(c) The hydrographs of the 872, 1128 and 1145 time steps are completely missed.

(d) There is an apparent eighteen-hour delay in the predicted hydrograph (with respect to the observed one) in time step 935.

Comments (a) and (b) can be attributed to the suboptimal set of the model parameters used as well as to errors in the input to the model. These type of errors are expected to be corrected by the on-line state estimator.

(c) and (d) errors were observed in all the simulation runs made and are attributed to large errors in the time and magnitude of the input hydrographs, obtained from the the soil moisture accounting scheme of the NWSRFS model running off-line. These errors would require the use of very high input noise intensity as well as high model error spectral densities locally, so that an on-line filter can disregard completely the input data and rely basically on the observations. On-line abrupt input error identification schemes are presented in Kitanidis and Bras (1978) and can be used in cases of unanticipated gross input errors. This is not pursued here.

It is expected, however, to observe local filter divergence in the regions of errors (c) and (d) during on-line parameter estimation and forecasting.

The square root of the average squared error, σ_p, and the lag-one correlation coefficient of the prediction errors, ρ_1, for the off-line run are as follows:

	$\sigma_p (m^3/sec)$	ρ_1
May :	33.1	0.786
July :	27.8	0.789

The high values of the correlation coefficient ρ_1 reveal that there is information in the observations that is not utilized by the output of the model (to be expected since this is an off-line run). Ideally, the values of ρ_1 should be close to zero for optimal performance.

Filter parameters for on-line state estimation and forecasting were obtained based on preliminary runs. The model error spectral density matrix was taken diagonal and so was the initial state covariance matrix. Values for the diagonal elements were obtained from a small scale (not exhaustive) sensitivity analysis and are presented in Table 6.3. The input noise intensity was taken proportional to the square of the instantaneous input. The sensitivity analysis gave a value of 6.25 for the coefficient V_I of proportionality.

Table 6.3

STATE ESTIMATOR STATISTICS

Parameter	Value (m^6/sec^2)
$Q_{x_{11}}$	2.75×10^{12}
$Q_{x_{22}}$	0.165×10^{12}
$Q_{x_{33}}$	0.165×10^{12}
$P_{x_{11}}(t_o)$	5.5×10^{9}
$P_{x_{22}}(t_o)$	0.055×10^{9}
$P_{x_{33}}(t_o)$	0.055×10^{9}

The observation error was assigned a variance proportional to the square of the observation value at each time step. The coefficient of proportionality was 0.01, that corresponds to very good measurement quality.

Figures 6.4 and 6.5 present plots of the six-hours ahead forecasted (dashed line) vs. the observed (solid line) discharges for the months of May and July, 1959, using the online state estimation procedure. The corresponding error indices σ_p and ρ_1 are given next.

	$\sigma_p (m^3/sec)$	ρ_1
May :	24	0.379
July :	16	0.489

With respect to the off-line runs of Figures 6.2 and 6.3 there is a decrease of about 30% and 40% in the square root of the average squared errors for the month of May and July, respectively. The prediction error lag are correlation coefficients of the on-line run are about half those of the off-line for the same months.

As expected the filter diverged locally in the regions of errors (c) and (d). However, it considerably improved the timing and the magnitude of the peak discharge for all cases.

Although local divergence contributer to the high correlation coefficients it is expected that adaptive estimation schemes that identify the input and model error noise intensities or more detailed sensitivity analysis will give error statistics that will reduce the correlation level of the one-step ahead predicted residuals.

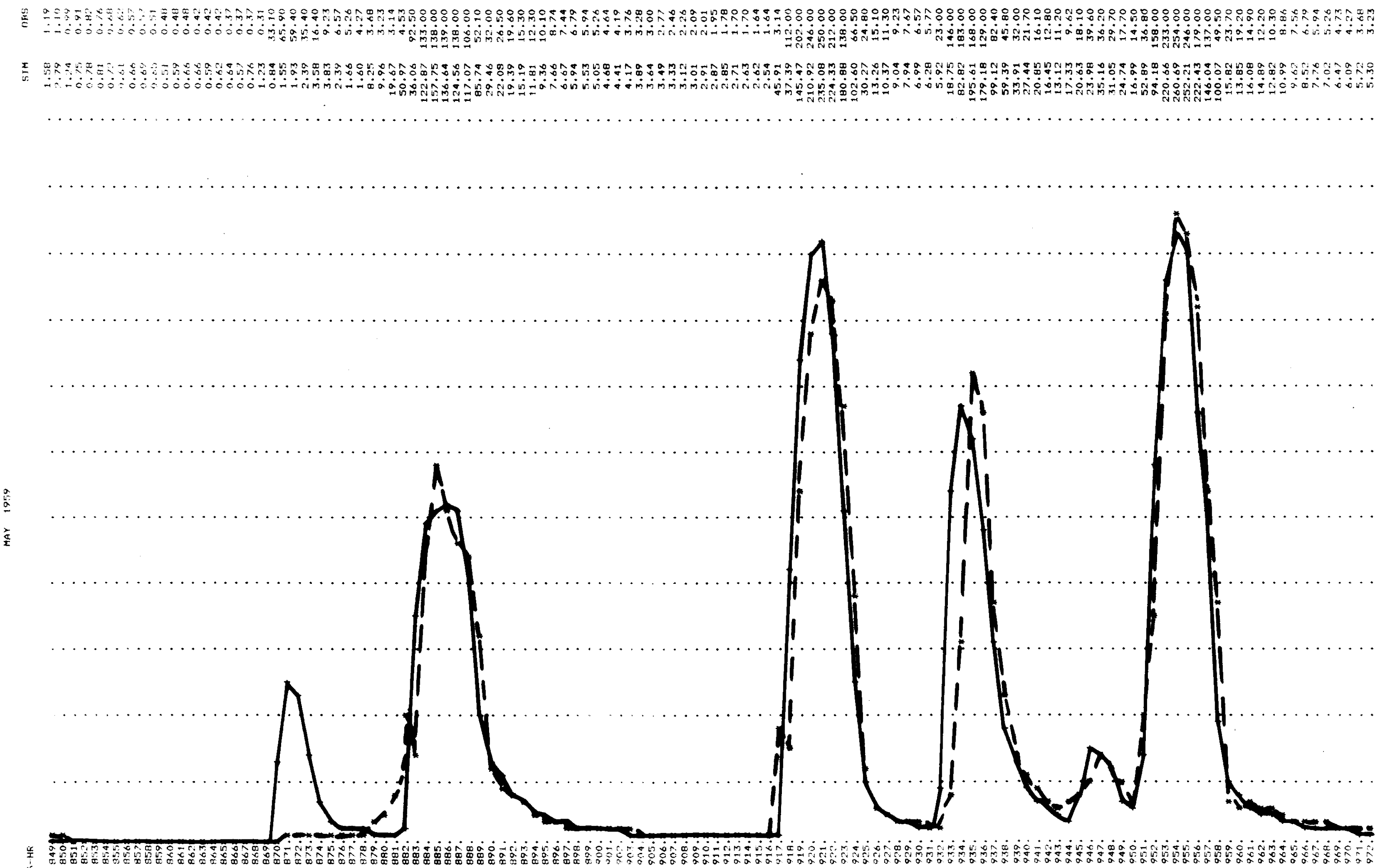

Figure 6.4

STATE ESTIMATOR SIX-HOUR LEAD FORECASTS FOR MAY, 1959. DASHED LINE CORRESPONDS TO THE FORECASTED VALUES. SOLID LINE CORRESPONDS TO THE OBSERVED VALUES - m^3/sec

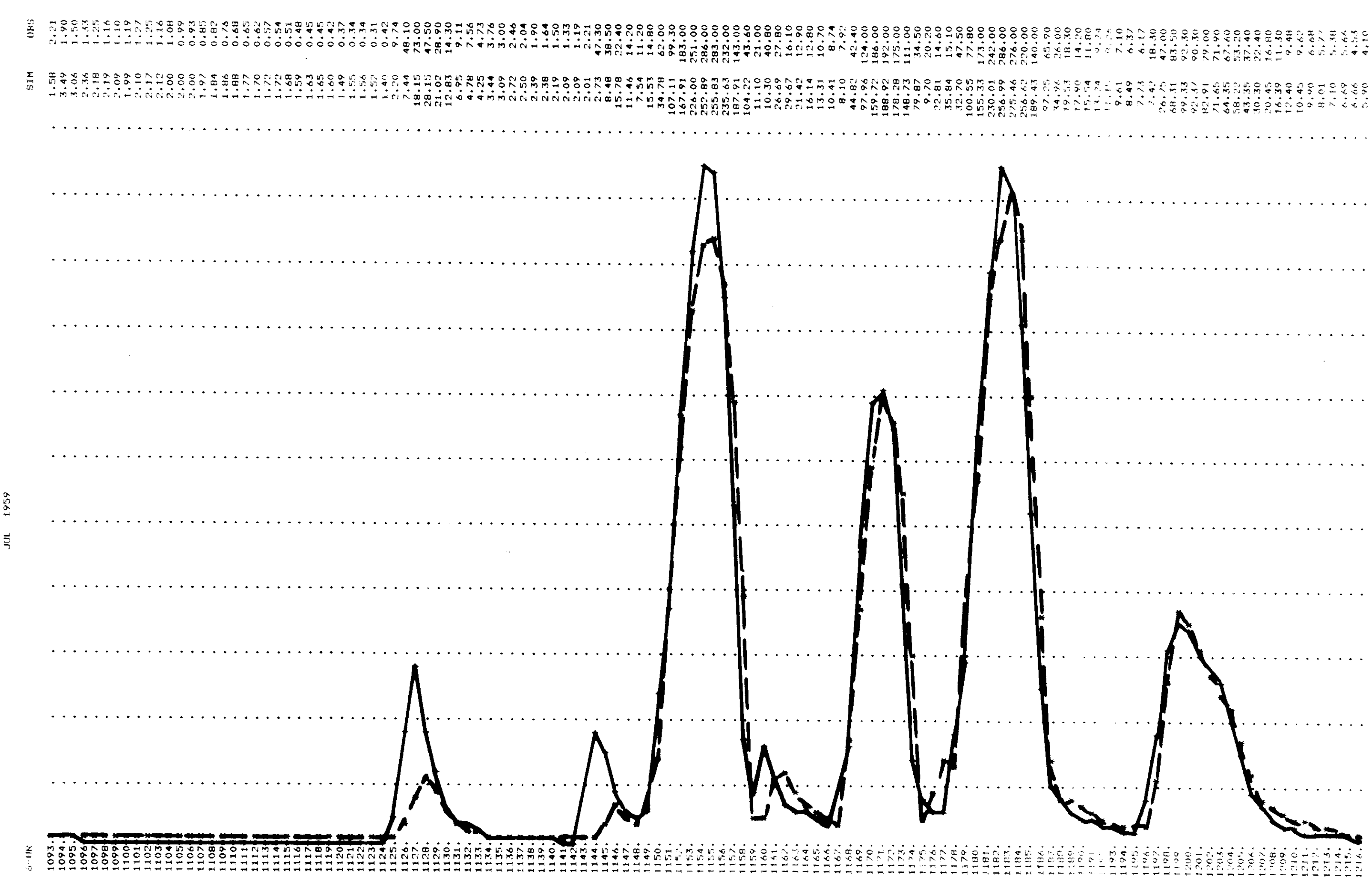

Figure 6.5

STATE ESTIMATOR SIX-HOUR LEAD FORECASTS FOR JULY, 1959. DASHED LINE CORRESPONDS TO THE FORECASTED VALUES. SOLID LINE CORRESPONDS TO THE OBSERVED VALUES - m^3/sec

6.4 Simultaneous State and Parameter Estimation Runs

A major problem in the use of simultaneous state and parameter estimation through state augmentation, is the determination of the filter statistical parameters. In particular, the model error spectral density matrix $\bar{Q}$, the input noise intensity $\sigma_u^2(t)$ and the observation error variance $R(t_k)$. $R(t_k)$ can often be estimated relatively accurately, given information about the observation mechanism and the quality of the observation data. The relative magnitude of the spectral densities $\bar{Q}_x$, corresponding to the states, and $\bar{Q}_a$ corresponding to the parameters, dictates the modelling of the parameters as random constants or as random walk (Gelb, 1974) processes. The choice between the two types of models is very important for the convergence properties of the resultant filters. It is obvious that the smaller $\bar{Q}_a$ is the faster the parameters will converge to some value. If one believes that the model structure is the true one, then it would seem appropriate to use estimates based on the whole history of the observations rather than on the few last time steps of the filter operation. It might not be appropriate, in other cases, to use a constant parameter model ($\bar{Q}_a \cong 0$), letting the filter gains go to zero and, consequently, design a filter not capable of operating when there is a future change in system behavior.

It is not a trivial task to test an identification scheme for convergence to the true parameter estimates when there is a state estimator operating in parallel. Depending on the statistical parameters of the filter, different in general, parameter values will minimize the average quadratic error criterion of performance.

It is not the purpose of this section to present detailed analysis of the simultaneous state and parameter estimation problem. The intent is to examine if the adaptive algorithm of the previous chapter converges and can be used as an analysis tool in identification studies.

The statistical parameters $\bar{\underline{P}}_x(t_o)$ and $\bar{\underline{Q}}_x$ were the ones used in the previous section. The input noise intensity was considerably reduced so that the parameters would be blamed for discrepancies between predicted and observed discharges, given that a relatively low model error spectral density matrix $\bar{\underline{Q}}_a$ was adopted to increase convergence. It was taken as the 5% of the squared value of the input at each time. $\bar{\underline{Q}}_a$ was taken diagonal and so did the initial covariance matrix $\bar{\underline{P}}_a(t_o)$. The cross covariance matrix $\bar{\underline{P}}_{xa}(t_o)$ was set equal to zero. The diagnoal elements of the matrices $\bar{\underline{Q}}_a$ and $\bar{\underline{P}}_a(t_o)$ are given in Table 6.4. For uniformity in the magnitude of the numbers corresponding to states and parameters, desired to avoid numerically singular cross-covariance matrices, the calculations were performed in units of mm/m^2/time-step for the discharge and mm/m^2 for the volume of water. The duration of one time-step was six hours. Table 6.4 reflects these units. For convenience, Table 6.5 presents the state estimator statistics in the same units.

The parameter values corresponding to the vertex 4 (Figure 6.1) were chosen as initial parameter values. Iterations using the state and parameter estimation scheme were made, where the initial parameters of each

Table 6.4

PARAMETER ESTIMATOR STATISTICS

Parameter	Value
$Q_{a_{11}}$	3×10^{-5}
$Q_{a_{22}}$	2×10^{-5}
$Q_{a_{33}}$	5×10^{-5}
$P_{a_{11}}(t_o)$	10^{-3}
$P_{a_{22}}(t_o)$	10^{-2}
$P_{a_{33}}(t_o)$	10^{-2}

Table 6.5

STATE ESTIMATOR STATISTICS USED IN ADAPTIVE RUN

Parameter	Value $(mm^2/m^4/(\text{time-step})^2)$
$Q_{x_{11}}$	.5
$Q_{x_{22}}$	.03
$Q_{x_{33}}$	.03
$P_{x_{11}}(t_o)$	10^{-3}
$P_{x_{22}}(t_o)$	10^{-5}
$P_{x_{33}}(t_o)$	10^{-5}

Table 6.6

PARAMETER ESTIMATES AND ASSOCIATED PERFORMANCE INDICES σ_p^2 and ρ_1

Run Description	$a_1(\times 10^4)$	$a_2(\times 10^4)$	$a_3(\times 10^4)$	σ_p^2	ρ_1
Off-line	10	5	5	3086	0.799
State Estimator	10	5	5	2068	0.674
Adaptive					
1st Iteration	10.09	7.10	7.66	1412	0.602
2nd Iteration	10.09	8.18	8.70	1132	0.533
3rd Iteration	9.92	8.70	9.14	1083	0.507
4th Iteration	9.83	8.87	9.31	1072	0.497
5th Iteration	9.57	9.05	9.40	1071	0.495

iteration were equal to the final estimates of the previous one. The period used for all iterations extended from the 913 time-step up to the 960 time-step in the water year 1959. Table 6.6 gives the final estimates of each iteration step with the associated average quadratic error of prediction σ_p^2 and the associated lag-one prediction error correlation coefficient, ρ_1. Results for the off-line run and the state estimator (with equivalent filter statistics) are also included. In this table (for comparison with the parameter estimates obtained from the simulation runs in the previous section), the discharge units are m^3/sec and the volume units are m^3.

It is obvious from Table 6.6 that the parameter estimation scheme has succeeded in reducing the error indices σ_p^2 and ρ_1 associated with state estimation. Convergence of the parameter values to the values 9.5×10^{-4}, 9.05×10^{-4}, 9.40×10^{-4} is apparent. Subsequent runs did not give significant improvement in σ_p^2, ρ_1.

In terms of the parameter space, the estimates moved from vertex 4 toward vertex 8 on the side defined by points 3, 4, 5, 8 of the cube. Then moved toward the plane 3, 5, 7 to arrive at their final values. Examination of the values of the errors in Table 6.2 reveals that the adaptive algorithm moved the parameters in the right direction.

The first iteration for parameter estimation reduced σ_p^2 to about 70% of the value resulting from exclusive state estimation. The second iteration reduced it to about 55% of the same value, and the third to about 48%.

The last two iterations offered minor improvement. This suggests that the first two iterations account for most of the variance reduction, implying fast convergence to optimal parameters.

The high prediction error correlation observed for the last iteration suggests that the model error spectral densities and the input noise intensity values used were suboptimal.

Note that the initial parameter values used as inputs to the adaptive scheme of Chapter 5, had an average squared error larger than that of the parameter values obtained from the off-line estimation procedures based on the results of section 2.5 ($a_1 = a_2 = a_3 = 10 \times 10^{-4}$). This suggests that in practice one can determine crude initial values from the off-line procedure and then directly use the on-line adaptive estimation scheme to refine those, without a second stage of refinement through simulation.

Chapter 7

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This work has examined the flood routing problem. A nonlinear router based on a series of reservoirs served as the model and modern estimation theory techniques were used to improve its performance in real time river discharge forecasting. The model was statistically linearized to become compatible with a Gaussian minimum variance estimator. The Taylor-Gauss methodology was proposed for the analytical determination of the expected value of nonlinear functions when the independent variable is approximately normally distributed. Having bypassed the heavy computational requirements for the numerical calculation of the statistical linearization gains, a recursive estimator for the states and the parameters was designed. The outcome of the synthesis procedure described, the stochastic flood routing model, was used in a real world application to forecast six-hour discharge values at the Bird Creek drainage basin.

Judging from the results of the case study, it is concluded that the statistical linearization technique, in conjunction with the Taylor-Gauss method, is a powerful tool of analysis for the flood forecasting problem. The conceptual flood routing model used can reproduce most of the characteristics of the output hydrographs (e.g. hydrograph shape, time to the peak discharge), while it does not require high quality input data (e.g. survey data), for the determination of initial values for its parameters.

Future research should concentrate on comparisons among the estimator design used in this work (based on statistical linearization), and other available nonlinear estimators (e.g. extended Kalman filter, second order filter, etc.), for both the state and the state-parameter estimation problems. The use of adaptive filter parameters identification schemes should be investigated in order to reduce the correlation of the predicted residuals and to establish a firm basis for comparison among the different models. Detailed sensitivity studies for the filter parameters and, in particular, for the input variance parameter and the model error spectral density matrix are suggested to see how changes in their values affect the speed of convergence of the parameter estimates to their optimal values. It will also be valuable to establish the magnitude of the region of initial values for the parameters for which the parameter estimation scheme gives convergent estimates. Finally, the use of the model for different flood events in different catchments, would increase its credibility.

REFERENCES

Amein, M. and Fang, C.S. (1969), "Stream Flow Routing (With Applications to North Carolina Rivers)," Water Resources Research Institute of the Univ. of North Carolina, Report No. 17, 106 pages.

Barnes, A.H. (1967), "Comparison of Computed and Observed Flood Routing in a Circular Cross-Section," Proc. International Hydrology Symposium, Vol. 1, Colorado State University, Fort Collins, Colorado, pp. 121-127.

Bhavnagri, V.S. and Bugliarello, G. (1965), "Mathematical Representation of an Urban Flood Plain," ASCE, J. of Hydraulics Division, HY2, March, pp. 149-173.

Bras, R.L. and Perkins, F.E. (1975), "Effects of Urbanization on Catchment Response," ASCE, J. of Hydraulics Division, HY3, March, pp. 451-466.

Bras, R.L. and Rodríquez-Iturbe, I. (1975), "Rainfall-Runoff as Spatial Stochastic Processes: Data Collection and Synthesis," Ralph M. Parsons Lab. for Water Resources and Hydrodynamics, MIT, TR #196, 382 pages.

Chan, S.O. and Bras, R.L. (1978), "Derived Distribution of Water Volume Above A Given Threshold Discharge," Ralph M. Parsons Lab. for Water Resources and Hydrodynamics, MIT, TR#234, 140 pages.

Cunge, J.A. (1969), "On the Subject of a Flood Propagation Method (Muskingum Method)," J. of Hydr. Res., IAHR , Vol. 7, No. 2, pp. 205-230.

Deyst, J.J. and Price, C.F. (1968), "Conditions for Asymptotic Stability of the Discrete, Minimum Variance, Linear Estimator," IEEE, Trans. Automatic Control, Vol. AC-13, Vol. 6, pp. 702-705.

Dooge, J.C.I. and Harley, B.M. (1967), "Linear Routing in Uniform Channels," Proc. International Hydrology Symposium, Vol. 1, Colorado State University, Fort Collins, Colorado, pp. 57-63.

Dooge, J.C.I. (1973), "Linear Theory of Hydrologic Systems," U.S. Dept. of Agriculture, Tech. Bull. No. 1468, pp. 50-53, 148-183 and 232-260.

Eagleson, P.S. (1967), "A Distributed Linear Model for Peak Catchment Discharge," Proc. International Hydrology Symposium, Vol. 1, Colorado State University, Fort Collins, Colorado, pp. 1-8.

Eagleson, P.S. (1970), "Dynamic Hydrology," McGraw-Hill Book Co., pp. 325-364.

Gelb, A. and Vander Velde, W.E. (1968), "Multiple-Input Describing Functions and Nonlinear System Design," McGraw-Hill Book Co., New York, pp. 1-40 and 365-437.

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

Georgakakos, K.P. and Bras, R.L. (1979), "On-Line River Discharge Forecasting Using Filtering and Estimation Theory," Progress Report, Sept. 13, 1978 - Feb. 13, 1979, Ralph M. Parsons Lab. for Water Resources and Hydrodynamics, Dept. of Civil Engineering, MIT, Contract No. 7-35112.

Gradshteyn, I.S. and Ryzhik, I.M. (1965), "Table of Integrals, Series and Products," Academic Press, New York, pp. 211-343.

Graham, D. and McRuer, D. (1961), "Analysis of Nonlinear Control Systems," Dover Publications, Inc., New York, pp. 230-244.

Henderson, F.M. (1966), "Open Channel Flow," McMillan Publishing Co., Inc., New York, pp. 125-164 and 285-398.

Hildebrand, F.B. (1976), "Advanced Calculus for Applications," Prentice-Hall, Inc. Englewood Cliffs, New Jersey, pp. 76-80 and 354-356.

Hino, M. (1973), "On Line Prediction of Hydrologic Systems," Proc. XVth Conf. IAHR, Istanbul, pp. 121-129.

Hughes, W.C. (1971), "Flood Control Release Optimization Using Methods From Calculus," ASCE, J. of Hydraulics Division, HY5, May, pp. 691-704.

Jazwinski, A.H. (1970),"Stochastic Processes and Filtering Theory," Academic Press, pp. 140-366.

Kalman, R.E. (1960), "A New Approach to Linear Filtering and Prediction Problems," ASME, J. of Basic Engineering, Vol. 82D, pp. 35-45.

Kalman, R.E. and Bucy, S.R. (1961), "New Results in Linear Filtering and Prediction Theory," ASME, J. of Basic Engineering, Vol. 83 D, pp. 95-108.

Kitanidis, P.K. and Bras, R.L. (1978), "Real Time Forecasting of River Flows," Ralph M. Parsons Lab. for Water Resources and Hydrodynamics, Dept. of Civil Engineering, MIT, TR #235, 324 pages.

Koussis, A. (1976), "An Approximate Dynamic Flood Routing Method," presented at the International Symposium on Unsteady Open-Channel Flow, Newcastle-Upon-Tyne, IAHR/BHRA, April.

Langbein, W.B. (1958), "Queuing Theory and Water Storage," ASCE, J. of the Hydraulics Division, Vol. 84, HY5, Proc. Paper 1811.

Li, R.M., Duong, N. and Simons, D.B. (1978), "Application of the Kalman Filter for Prediction of Stage-Discharge Relationships in Rivers," in Applications of Kalman Filter to Hydrology, Hydraulics and Water Resources, ed. Chiu, C.L., University of Pittsburgh, pp. 459-471.

Lighthill, M.J. and Whitham, G.B. (1955), "On Kinematic Waves I. Flood Movement in Long Rivers," Proc. Roy. Soc. A., Vol. 229, pp. 281-316.

Ljung, L. (1979), "Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems," IEEE, Trans. Automatic Control, Vol. AC-24, No. 1, pp. 36-50.

Logan, W.C., Lennox, W.N. and Unny, T.E. (1978), "A Time Varying Non-Linear Hydrologic Response Model for Flood Hydrography Estimation in a Noisy Environment", in Applications of Kalman Filter to Hydrology, Hydraulics and Water Resources, ed. Chiu, C.L., University of Pittsburgh, pp. 279-294.

Mein, R.G., Laurenson, E.M. adn McMahon, T.A. (1974), "Simple Nonlinear Model for Flood Estimation," ASCE, J. of Hydraulics Division, HY11, November, pp. 1507-1518.

Nash, J.E. (1959), "A Note on the Muskingum Flood-Routing Method," J. of Geophys. Res., Vol. 64, pp. 1053-1056.

NWS HYDRO-31 (1976), "Catchment Modeling and Initial Parameter Estimation for the National Weather Service River Forecast System," by Peck, E.L., 24 pages.

Overton, D.E. (1967), "Analytical Simulation of Watershed Hydrographs from Rainfall," Proc. International Hydrology Symposium, Vol. 1, Colorado State University, Fort Collins, Colorado, pp. 9-17.

Price, R.K. (1973), "Flood Routing Methods for British Rivers," Hydr. Res. Station, Wallingford, INT 111, 102 pages.

Restrepo-Posada, P. (1978), "ISPLT," FORTRAN IV Subroutine for Three Dimensional, Isometric plots, Ralph M. Parsons Lab. for Water Resources and Hydrodynamics, Dept. of Civil Engineering, MIT, unpublished.

Schultz, D.G. and Melsa, J.L. (1967), "State Functions and Linear Control Systems," McGraw-Hill Book Co., New York, pp. 1-153.

Seinfeld, J.H., Gavalas, G.R., and Hwang, M. (1971), "Nonlinear Filtering in Distributed Parameter Systems," ASME, J. of Dynamic Systems, Measurement, and Control, Vol. 93G, No. 3, pp. 157-163.

Todini, E. and Bouillot, D. (1975), "A Rainfall-Runoff Kalman Filter Model," in System Simulation in Water Resources, ed. G.C. Vansteenkiste, North Holland Publishing Co., Amsterdam.

Viessman, W. Jr., Knapp, J.W., Lewis, G.L., Harbaugh, T.E. (1972), "Introduction to Hydrology," Harper and Row Publishers, New York, pp. 231-283.

Weinmann, P.E. (1977), "Comparison of Flood Routing Methods for Natural Rivers," M. Eng. Sc. Thesis, Monash University, Australia, 182 pages.

Weinmann, P.E. and Laurenson, E.M. (1977), "Modern Methods of Flood Routing," presented in the Hydrology Symposium at Brisbane, Australia, June 28-30.

Wood, E.F. (1978), "An Application of Kalman Filtering to River Flow Forecasting," in Applications of Kalman Filter to Hydrology, Hydraulics and Water Resources, ed. Chiu, C.L., University of Pittsburgh, pp. 385-408.

Appendix A

CENTRAL MOMENTS OF COMMON PROBABILISTIC MODELS

The purpose of this Appendix is to derive analytical expressions for the n^{th} central moment of a scalar random variable, given the underlying probability distribution law.

Denote by X a scalar random variable with probability density function (p.d.f.) $p_X(x; \beta_1, \beta_2, \ldots, \beta_m)$. x is a population sample and $\beta_1, \beta_2, \ldots, \beta_m$ are the m parameters of the p.d.f. of X. It is desired to find analytical expressions for the expectation: $E\{(x - E\{x\})^n\}$.

The binomial form expansion of $E\{(x - E\{x\})^n\}$ yields,

$$E\{(x - E\{x\})^n\} = E\{\binom{n}{0}\cdot x^n + (-1)^1\cdot\binom{n}{1}\cdot x^{n-1}\cdot E\{x\} + \ldots + (-1)^{n-1}\cdot\binom{n}{n-1}\cdot x\cdot[E\{x\}]^{n-1} + (-1)^{n-1}\cdot\binom{n}{n}\cdot[E\{x\}]^n\} \quad (A.1)$$

$$E\{(x - E\{x\})^n\} = E\left\{\sum_{i=0}^{n} (-1)^i \cdot \binom{n}{i} \cdot x^{n-i} \cdot \mu^i\right\} \quad (A.2)$$

where

$$\mu = E\{x\} \quad (A.3)$$

Due to the fact that the expectation operator can be interchanged with the summation operator, Eq. (A.2) results in:

$$E\{(x - E\{x\})^n\} = \sum_{i=0}^{n} (-1)^i \binom{n}{i} \cdot \mu^i \cdot E\{x^{n-i}\} \quad (A.4)$$

Thus, the problem becomes to determine the expectation $E\{x^{n-i}\}$ as a function of the parameters of the probability density function of X for all (n-i).

In the following, the expectation $E\{x^{\ell}\}$ is determined for integer ℓ when X is obeying an exponential, rectangular, gamma, log-normal and normal (univariate) probability distribution law.

The p.d.f. for the exponential distribution is given as:

$$p_X(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & ; \quad \lambda > 0 \ , \quad x > 0 \\ 0 & ; \quad x \leq 0 \end{cases} \tag{A.5}$$

where λ is the parameter of the p.d.f.

By definition

$$E\{x^{\ell}\} = \int_{0}^{+\infty} x^{\ell} \cdot p_X(x; \lambda) \cdot dx \tag{A.6}$$

Substitution of Eq. (A.5) in Eq. (A.6) results in

$$E\{x^{\ell}\} = \int_{0}^{+\infty} x^{\ell} \cdot \lambda \cdot e^{-\lambda x} \cdot dx \tag{A.7}$$

The integral in Eq. (A.7) is equal to (Gradshteyn and Ryzhik, 1965, integral number: 3.351 (3))

$$\lambda \int_{0}^{+\infty} x^{\ell} \cdot e^{-\lambda x} \, dx = \ell! \cdot \lambda^{-\ell} \tag{A.8}$$

Consequently,

$$E\{x^{\ell}\} = \ell! \cdot \lambda^{-\ell} \tag{A.9}$$

The probability density function for a <u>rectangular</u> distribution is:

$$p_X(x;\ \beta,\ \lambda) = \begin{cases} \frac{1}{\lambda}\ , & \beta - \frac{\lambda}{2} \leq x \leq \beta + \frac{\lambda}{2} \\ 0, & \beta + \frac{\lambda}{2} < x \text{ or } x < \beta - \frac{\lambda}{2} \end{cases} \quad ;\ \lambda > 0 \qquad \text{(A.10)}$$

where β, λ are the parameters of the distribution.

In this case:

$$E\{x^{\ell}\} = \int_{\beta-\frac{\lambda}{2}}^{\beta+\frac{\lambda}{2}} \frac{1}{\lambda} \cdot x^{\ell} \cdot dx \qquad \text{(A.11)}$$

Direct evaluation of the integral in Eq. (A.11) yields

$$E\{x^{\ell}\} = \frac{1}{\lambda \cdot (\ell+1)} \cdot \left[(\beta + \frac{\lambda}{2})^{\ell+1} - (\beta - \frac{\lambda}{2})^{\ell+1}\right] \qquad \text{(A.12)}$$

A <u>Gamma</u> distributed random variable has p.d.f. as follows,

$$p_X(x;\ \lambda) = \begin{cases} \frac{x^{\lambda-1} \cdot e^{-x}}{\Gamma(\lambda)} & ;\quad \lambda > 0\ ,\ x > 0 \\ 0 & ;\quad x \leq 0 \end{cases} \qquad \text{(A.13)}$$

where $\Gamma(\lambda)$ is the Gamma function.

It follows

$$E\{x^{\ell}\} = \int_{0}^{+\infty} x^{\ell} \cdot \frac{x^{\lambda-1} \cdot e^{-x}}{\Gamma(\lambda)} \cdot dx \qquad \text{(A.14)}$$

or

$$E\{x^{\ell}\} = \frac{1}{\Gamma(\lambda)} \cdot \int_{0}^{+\infty} x^{\ell+\lambda-1} \cdot e^{-x} \cdot dx \qquad \text{(A.15)}$$

The integral in Eq. (A.15) is a Gamma function (Gradshteyn and Ryzhik, 1965, integral number: 3.381 (4)) with argument equal to $\ell+\lambda$. Thus,

$$E\{x^{\ell}\} = \frac{\Gamma(\ell + \lambda)}{\Gamma(\lambda)} \qquad \text{(A.16)}$$

Since (Hildebrand, 1976),

$$\Gamma(\beta) = (\beta-1) \cdot (\beta-2) \ldots \cdot (\delta+1) \cdot \delta \cdot \Gamma(\delta) \qquad \text{(A.17)}$$

for any $\beta > 0$ and $0 < \delta < 1$, and in addition, ℓ is an integer, it follows

$$E\{x^{\ell}\} = \prod_{j=1}^{\ell} (\lambda + \ell - j) \qquad \text{(A.18)}$$

The lognormal probability density function is:

$$p_X(x;\, \lambda,\, \beta) = \begin{cases} \dfrac{1}{x\sqrt{2\pi}\cdot\beta} \cdot e^{-\frac{1}{2} \cdot [\frac{1}{\beta} \cdot \ell n(\frac{x}{\lambda})]^2} & ; \quad \beta > 0,\ x \geq 0 \\ 0 & ; \quad x \leq 0 \end{cases} \qquad \text{(A.19)}$$

The expectation $E\{x^{\ell}\}$ can be determined by

$$E\{x^{\ell}\} = \int_0^{\infty} x^{\ell} \cdot \frac{1}{x\cdot\sqrt{2\pi}\cdot\beta} \cdot e^{-\frac{1}{2}[\frac{1}{\beta} \cdot \ell n(\frac{x}{\lambda})]^2} \cdot dx \qquad \text{(A.20)}$$

Using the transformation $y = \ell nx$, with $dy = \frac{1}{x} \cdot dx$ and $x = e^y$, in Eq. (A.20) results in

$$E\{x^{\ell}\} = \frac{1}{\sqrt{2\pi}\cdot\beta} \int_{-\infty}^{+\infty} e^{\ell\cdot y - \frac{1}{2\beta^2} [y^2 - 2\cdot\ell n\lambda\cdot y + (\ell n\lambda)^2]} \cdot dy \qquad \text{(A.21)}$$

or

$$E\{x^{\ell}\} = \frac{e^{-\frac{1}{2}(\frac{\ell n\lambda}{\beta})^2}}{\sqrt{2\pi}\cdot\beta} \cdot I(\beta,\ \lambda) \qquad (A.22)$$

where $I(\beta,\ \lambda)$ represents the integral

$$I(\beta,\ \lambda) = \int_{-\infty}^{+\infty} e^{-\frac{1}{2\beta^2}\cdot y^2 + (\frac{\ell n\lambda}{\beta^2} + \ell)\cdot y} \cdot dy \qquad (A.23)$$

Gradshteyn and Ryzhik (1965; integral number: 3.323 (2)) give ($\beta > 0$):

$$I(\beta,\ \lambda) = \sqrt{2\pi} \cdot \beta \cdot e^{(\frac{\ell n\lambda}{\beta^2} + \ell)^2 \cdot \frac{\beta^2}{2}} \qquad (A.24)$$

Noting that

$$(\frac{\ell n\lambda}{\beta^2} + \ell)^2 \cdot \frac{\beta^2}{2} = \frac{1}{2} \cdot (\frac{\ell n\lambda}{\beta} + \ell\cdot\beta)^2 \qquad (A.25)$$

Eq. (A.24) yields

$$I(\beta,\lambda) = \sqrt{2\pi} \cdot \beta \cdot e^{\frac{1}{2}[(\frac{\ell n\lambda}{\beta})^2 + \ell^2\cdot\beta^2 + 2\ell\cdot\ell n\lambda]} \qquad (A.26)$$

Substitution of this expression of $I(\beta,\ \lambda)$ in Eq. (A.22) results in the following expressions for the expectation $E\{x^{\ell}\}$:

$$E\{x^{\ell}\} = \lambda^{\ell} \cdot e^{\frac{1}{2}\cdot\ell^2\cdot\beta^2} \qquad (A.27)$$

Substitution of Eqs. (A.9), (A.12), (A.18), and (A.27) in Eq. (A.4) with $\ell = n - i$ results in the expressions given by Eqs. (3.7), (3.10), (3.13) and (3.15).

For the normal distribution, it is easier to directly work with the integral:

$$E\{(x - E\{x\})^n\} = \int_{-\infty}^{+\infty} (x - E\{x\})^n \cdot p_X(x;\ \beta,\ \lambda) \cdot dx \qquad (A.28)$$

where the p.d.f. $p_X(x;\ \beta,\ \lambda)$ is given by

$$p_X(x;\ \beta,\ \lambda) = \frac{1}{\sqrt{2\pi}\cdot\beta} \cdot e^{-\frac{1}{2}\left(\frac{x-\lambda}{\beta}\right)^2} \quad ; \qquad \beta > 0 \qquad (A.29)$$

At first it is shown

$$E\{x\} = \lambda \qquad (A.30)$$

By definition,

$$E\{x\} = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi}\cdot\beta} \cdot e^{-\frac{1}{2}\left(\frac{x-\lambda}{\beta}\right)^2} \cdot dx \qquad (A.31)$$

or

$$E\{x\} = \frac{e^{-\frac{1}{2}\cdot\left(\frac{\lambda}{\beta}\right)^2}}{\sqrt{2\pi}\cdot\beta} \cdot I(\beta,\ \lambda) \qquad (A.32)$$

where the integral $I(\beta,\ \lambda)$ is defined as

$$I(\beta,\ \lambda) = \int_{-\infty}^{+\infty} x \cdot e^{-\frac{1}{2\beta^2} \cdot x^2 + 2\cdot\left(\frac{\lambda}{2\beta^2}\right)\cdot x} \cdot dx \qquad (A.33)$$

Analytical expression for this integral is given in Gradshteyn and Ryzhik (1965; integral number: 3.462 (6)) as

$$I(\beta,\ \lambda) = \lambda \cdot \sqrt{2\pi}\cdot\beta \cdot e^{\frac{1}{2}\left(\frac{\lambda}{\beta}\right)^2} \qquad (A.34)$$

Substitution in Eq. (A.32) proves the validity of Eq. (A.30).

Under the circumstances, Eq. (A.28) can be rewritten as

$$E\{(x-\lambda)^n\} = \int_{-\infty}^{+\infty} (x-\lambda)^n \cdot e^{-\frac{1}{2}\left(\frac{x-\lambda}{\beta}\right)^2} \cdot dx \tag{A.35}$$

Using the transformation $y = \frac{x-\lambda}{\beta}$ in Eq. (A.35) results in

$$E\{(x-\lambda)^n\} = \int_{-\infty}^{+\infty} \beta^{n+1} \cdot y^n \cdot e^{-\frac{1}{2}y^2} \cdot dy \tag{A.36}$$

or

$$E\{(x-\lambda)^n\} = \beta^{n+1} \cdot \int_{-\infty}^{+\infty} y^n \cdot e^{-\frac{1}{2}y^2} \cdot dy \tag{A.37}$$

Decompose the integral of Eq. (A.37) in two as follows

$$E\{(x-\lambda)^n\} = \beta^{n+1} \cdot \left[\int_0^{+\infty} y^n \cdot e^{-\frac{1}{2}y^2} \cdot dy + \int_{-\infty}^{0} y^n \cdot e^{-\frac{1}{2}y^2} \cdot dy\right] \tag{A.38}$$

a. If n is an odd integer, then the transformation $z = -y$ converts the last integral of Eq. (A.38) to the form

$$\int_{-\infty}^{0} y^n \cdot e^{-\frac{1}{2}y^2} \cdot dy = -\int_{0}^{+\infty} z^n \cdot e^{-\frac{1}{2}\cdot z^2} \cdot dz \tag{A.39}$$

Consequently, in this case

$$E\{(x-\lambda)^n\} = 0 \quad ; \quad n \text{ is an odd integer} \tag{A.40}$$

b. If n is an even integer, then the transformation $z = -y$ converts the last integral of Eq. (A.38) to the form

$$\int_{-\infty}^{0} y^n \cdot e^{-\frac{1}{2} y^2} \cdot dy = \int_{0}^{+\infty} z^n \cdot e^{-\frac{1}{2} z^2} \cdot dz \qquad (A.41)$$

Substitution in Eq. (A.38) yields

$$E\{(x-\lambda)^n\} = 2 \cdot \beta^{(n+1)} \cdot \int_{0}^{+\infty} y^n \cdot e^{-\frac{1}{2} y^2} \cdot dy; \quad n \text{ is an even integer} \qquad (A.42)$$

Gradshteyn and Ryzhik (1965; integral number: 3.461 (2)) show:

$$\int_{0}^{+\infty} y^n \cdot e^{-\frac{1}{2} y^2} \cdot dy = \frac{(n-1)!!}{2} \cdot \sqrt{2\pi} \ ; \ n \text{ is an even integer} \qquad (A.43)$$

where

$$(n-1)!! = 1 \cdot 3 \cdot 5 \ .. \ (n-1) \quad ; \quad n \text{ is an even integer} \qquad (A.44)$$

Substitution of Eq. (A.43) in Eq. (A.42) results in:

$$E\{(x - \lambda)^n\} = \sqrt{2\pi} \cdot \beta^{n+1} \cdot (n-1)!! \ ; \quad n \text{ is an even integer} \qquad (A.45)$$

The combination of Eq. (A.40) and (A.45) gives Eq. (3.18).

Appendix B

PARAMETRIC STUDY OF A LINEAR RESERVOIR

The input hydrograph is assumed to have a triangular shape. Denote by Q_m its maximum discharge and by T_1 the time to the peak discharge. The time duration of the falling limb is denoted by T_2. Under these definitions, the input hydrograph equations are:

$$u(t) = t \cdot \frac{Q_m}{T_1} \quad ; \qquad 0 \le t \le T_1 \qquad \text{(B.1)}$$

$$u(t) = Q_m - (t - T_1) \cdot \frac{Q_m}{T_2} \quad ; \qquad T_1 \le t \le T_1 + T_2 \qquad \text{(B.2)}$$

where u(t) is the input rate at time t and it has been assumed that the initial time is equal to zero.

To facilitate notation, define the constants C_1 and C_2 as:

$$C_1 = \frac{Q_m}{T_1} \qquad \text{(B.3)}$$

$$C_2 = \frac{Q_m}{T_2} \qquad \text{(B.4)}$$

The equation describing the motion of water through a linear reservoir is

$$\frac{dS(t)}{dt} = u(t) - a \cdot S(t) \qquad \text{(B.5)}$$

where S(t) is the water in storage at time t and a is a constant parameter.

The output discharge is given by

$$Q(t) = a \cdot S(t) \tag{B.6}$$

Equation (B.5) can be solved analytically to yield (Hildebrand, 1976),

$$S(t) = \int_o^t e^{-a(t-\tau)} \cdot u(\tau) \cdot d\tau \tag{B.7}$$

Substitution of the expressions for u(t) from Eqs. (B.1) and (B.2) in Eq. (B.7) results in

$$S(t) = \int_o^t e^{-a(t-\tau)} \cdot \tau \cdot C_1 \cdot d\tau \quad ; \quad 0 \leq t \leq T_1 \tag{B.8}$$

and

$$S(t) = \int_o^{T_1} e^{-a(t-\tau)} \cdot \tau \cdot C_1 \cdot d\tau + \int_{T_1}^{t} e^{-a(t-\tau)} \cdot Q_m \cdot d\tau - \int_{T_1}^{t} e^{-a(t-\tau)} \cdot (\tau - T_1) \cdot C_2 \cdot d\tau \; ; \; T_1 \leq t \leq T_1 + T_2 \tag{B.9}$$

Integration by parts of Eq. (B.8) yields,

$$S(t) = C_1 \cdot e^{-at} \cdot \frac{1}{a} \cdot [\tau \cdot e^{a\tau}\Big|_o^t - \int_o^t e^{a\tau} \cdot d\tau] \tag{B.10}$$

or

$$S(t) = \frac{C_1 \cdot e^{-at}}{a} \cdot [t \cdot e^{at} - \frac{1}{a} \cdot e^{at} + \frac{1}{a}] \tag{B.11}$$

Rearranging terms in Eq. (B.11) gives

$$S(t) = \frac{C_1}{a^2} \cdot [at - 1 + e^{-at}] \quad ; \quad 0 \leq t \leq T_1 \tag{B.12}$$

Equation (B.9) can be written as

$$S(t) = I_1(t) + I_2(t) - I_3(t) \quad ; \quad T_1 \leq t \leq T_1 + T_2 \tag{B.13}$$

where the integrals $I_1(t)$, $I_2(t)$ and $I_3(t)$ are given by

$$I_1(t) = \int_o^{T_1} e^{-a(t-\tau)} \cdot \tau \cdot C_1 \cdot d\tau \tag{B.14}$$

$$I_2(t) = \int_{T_1}^{t} e^{-a(t-\tau)} \cdot Q_m \cdot d\tau \tag{B.15}$$

and

$$I_3(t) = \int_{T_1}^{t} e^{-a(t-\tau)} \cdot (\tau - T_1) \cdot C_2 \cdot d\tau \tag{B.16}$$

Integration by parts of Eq. (B.14) gives

$$I_1(t) = \frac{C_1 \cdot e^{-a(t-T_1)}}{a^2} \cdot [aT_1 - 1 + e^{-aT_1}] \tag{B.17}$$

Similarly, Eq. (B.15) gives,

$$I_2(t) = \frac{Q_m}{a} \cdot [1 - e^{-a(t-T_1)}] \tag{B.18}$$

The integral in Eq. (B.16) can be evaluated by a change of variables. Let,

$$u = \tau - T_1 \tag{B.19}$$

then,

$$I_3(t) = e^{-a(t-T_1)} \cdot \int_o^{t-T_1} e^{au} \cdot u \cdot C_2 \cdot du \tag{B.20}$$

Integration yields,

$$I_3(t) = e^{-a(t-T_1)} \cdot \frac{C_2}{a} \cdot [u \cdot e^{au}\Big|_o^{t-T_1} - \int_o^{t-T_1} e^{au} \cdot du] \quad (B.21)$$

or

$$I_3(t) = e^{-a(t-T_1)} \cdot \frac{C_2}{a} \cdot [(t-T_1) \cdot e^{a(t-T_1)} - \frac{1}{a} \cdot e^{a(t-T_1)} + \frac{1}{a}] \quad (B.22)$$

Substitution of Eqs. (B.17), (B.18) and (B.22) in Eq. (B.13) results in

$$S(t) = \frac{C_1 \cdot e^{-a(t-T_1)}}{a^2} \cdot [aT_1 - 1 + e^{-aT_1}] + \frac{Q_m}{a} \cdot [1 - e^{-a(t-T_1)}]$$
$$- \frac{C_2}{a^2} \cdot [a \cdot (t-T_1) - 1 + e^{-a(t-T_1)}] \quad ; \quad T_1 \leq t \leq T_1 + T_2 \quad (B.23)$$

Due to the translation in time of the input when it passes through the linear reservoir, it is expected that, if t_p denotes the time to the peak of the outflow, it holds,

$$t_p \geq T_1 \quad (B.24)$$

Also, since the outflow discharge is in a one-to-one correspondence with the volume of water stored by means of Eq. (B.6), then the maximum outflowing discharge will occur when S(t) is maximum. The time t_p when the maximum of S(t) occurs can be obtained by differentiating Eq. (B.23) with respect to time and setting the derivative equal to zero. This gives,

$$- \frac{C_1 e^{-a(t_p - T_1)}}{a} \cdot [aT_1 - 1 + e^{-aT_1}] + Q_m \cdot e^{-a(t_p - T_1)}$$

$$+ \frac{C_2}{a} \cdot e^{-a(t_p - T_1)} - \frac{C_2}{a} = 0 \quad \text{(B.25)}$$

Solving for $(t_p - T_1)$,

$$t_p - T_1 = - \frac{\ell nA}{a} \quad \text{(B.26)}$$

where

$$A = \frac{C_2}{C_2 + a\ Q_m - C_1 \cdot (aT_1 - 1 + e^{-aT_1})} \quad \text{(B.27)}$$

or

$$A = \frac{\frac{T_1}{T_2}}{\frac{T_1}{T_2} + 1 - e^{-aT_1}} \quad \text{(B.28)}$$

Use of Eq. (B.25) in Eq. (B.23), for $t = t_p$, results in a peak storage volume expression,

$$S_m = - \frac{C_2}{a^2} + \frac{Q_m}{a} - \frac{C_2}{a^2} \cdot [a(t_p - T_1) - 1] \quad \text{(B.29)}$$

or

$$S_m = \frac{Q_m}{a} \cdot \left(1 - \frac{t_p - T_1}{T_2}\right) \quad \text{(B.30)}$$

The peak outflow discharge Q_p is given by

$$Q_p = Q_m \cdot \left(1 - \frac{t_p - T_1}{T_2}\right) \quad \text{(B.31)}$$

Denote by R_1 the normalized attenuation given by

$$R_1 = \frac{Q_m - Q_p}{Q_m} \tag{B.32}$$

and by R_2 the normalized translation time given by,

$$R_2 = \frac{t_p - T_1}{T_1} \tag{B.33}$$

Using these definitions in Eq. (B.31) results in

$$\frac{R_1}{R_2} = \frac{T_1}{T_2} = K_T \tag{B.34}$$

This suggests that the ratio of the normalized attentuation vs. the normalized translation time is only dependent on the ratio of the characteristic times T_1, T_2 of the triangular input hydrograph. Thus, one linear reservoir cannot provide arbitrary normalized discharge attenuation and translation times by simply varying its coefficient a.

The results of a linear reservoir response to a triangular input hydrograph can be used to study the behavior of a cascade of reservoirs with parameters a_i. To do so, the output of each reservoir in the sequence must be approximated by a triangular hydrograph preserving the relevant characteristic features (time to peak, peak discharge, recession time). The approximated triangular hydrograph will serve as an input to the next reservoir. One way to reshape a given hydrograph to triangular form is to assure that the volume of water discharged up to the time of the peak discharge t_p is preserved in the transformation.

Similarly, for times greater than t_p.

Denote by V_1 the volume under the output hydrograph up to the time t_p. Then,

$$V_1 = \int_o^{T_1} a \cdot \frac{C_1}{a^2} \cdot (at - 1 + e^{-at}) \cdot dt$$
$$+ \int_{T_1}^{t_p} a \cdot \frac{C_1 \cdot e^{-a(t-T_1)}}{a^2} \cdot (aT_1 - 1 + e^{-aT_1}) \cdot dt$$
$$+ \int_{T_1}^{t_p} a \cdot \frac{Q_m}{a} \cdot (1 - e^{-a(t-T_1)}) \cdot dt$$
$$- \int_{T_1}^{t_p} a \cdot \frac{C_2}{a^2} \cdot [a \cdot (t - T_1) - 1 + e^{-a(t-T_1)}] \cdot dt \tag{B.35}$$

Carrying out the necessary integrations in Eq. (B.35) yields

$$V_1 = \frac{C_1}{a^2} \cdot (\frac{a^2 T_1^2}{2} - aT_1 - e^{-aT_1} + 1) - \frac{C_1}{a^2} \cdot (aT_1 - 1 + e^{-aT_1})$$
$$\cdot (e^{-a(t_p-T_1)} - 1) + Q_m \cdot (t_p - T_1) + \frac{Q_m}{a} \cdot (e^{-a(t_p-T_1)} - 1)$$
$$- C_2 \cdot \frac{(t_p-T_1)^2}{2} + \frac{C_2}{a} \cdot (t_p - T_1) + \frac{C_2}{a^2} \cdot (e^{-a(t_p-T_1)} - 1) \tag{B.36}$$

By means of Eq. (B.25), the above expression simplifies to,

$$V_1 = C_1 \cdot \frac{T_1^2}{2} + Q_m \cdot (t_p-T_1) - \frac{Q_m}{a} - C_2 \cdot \frac{(t_p-T_1)^2}{2} + \frac{C_2}{a} \cdot (t_p-T_1) \tag{B.37}$$

Substitution of C_1, C_2 from Eqs. (B.3) and (B.4) results in:

$$V_1 = Q_m \cdot \frac{T_1}{2} + Q_m \cdot (t_p - T_1) - \frac{Q_m}{a} - Q_m \cdot \frac{(t_p - T_1)^2}{2T_2} + \frac{Q_m}{a} \cdot \frac{(t_p - T_1)}{T_2} \qquad \text{(B.38)}$$

Expressing Q_m in terms of Q_p [Equation (B.31)] yields,

$$V_1 = -\frac{Q_p}{a} + Q_p \cdot \frac{(t_p - T_1)}{2} + Q_m \cdot \frac{t_p}{2} \qquad \text{(B.39)}$$

The total volume V_T is given by

$$V_T = \frac{1}{2} \cdot (T_1 + T_2) \cdot Q_m \qquad \text{(B.40)}$$

Denote by V_2 the volume discharged after time t_p. Then,

$$V_2 = V_T - V_1 \qquad \text{(B.41)}$$

or

$$V_2 = Q_p \cdot \frac{T_2}{2} + \frac{Q_p}{a} - Q_p \cdot \frac{(t_p - T_1)}{2} \qquad \text{(B.42)}$$

Denote by t_p' an equivalent time to the peak of the output hydrograph and by t_r' an equivalent falling limb time of the same hydrograph. Then,

$$\frac{1}{2} \cdot Q_p \cdot t_p' = V_1 \qquad \text{(B.43)}$$

and

$$\frac{1}{2} \cdot Q_p \cdot t_r' = V_2 \qquad \text{(B.44)}$$

It follows

$$t_p' = \frac{2 \cdot V_1}{Q_p} \qquad \text{(B.45)}$$

$$t_r' = \frac{2 \cdot V_2}{Q_p} \tag{B.46}$$

Substitution of the expressions for V_1 and V_2 from Eqs. (B.39) and (B.42), in Eqs. (B.45) and (B.46), respectively, gives,

$$t_p' = -\frac{2}{a} + t_p - T_1 + \frac{Q_m}{Q_p} \cdot t_p \tag{B.47}$$

and

$$t_r' = T_2 + \frac{2}{a} - (t_p - T_1) \tag{B.48}$$

The combination of Eqs. (B.31), (B.34) and (B.26) yields,

$$\frac{Q_p}{Q_m} = 1 + \frac{\ell nA}{aT_1} \cdot K_T \tag{B.49}$$

Equation (B.47) gives,

$$at_p' = -2 - \ell nA + (aT_1 - \ell nA) \cdot \frac{1}{1 + K_T \frac{\ell nA}{aT_1}} \tag{B.50}$$

Similarly, Eq. (B.48) gives

$$at_r' = aT_1 \cdot \frac{1}{K_T} + 2 + \ell nA \tag{B.51}$$

Then,

$$\frac{t_p'}{t_r'} = \frac{(aT_1 - \ell nA) \cdot \frac{1}{1 + K_T \cdot \frac{\ell nA}{aT_1}} - 2 - \ell nA}{aT_1 \cdot \frac{1}{K_T} + 2 + \ell nA} \tag{B.52}$$

The normalized equations (B.49), (B.50), (B.51) and (B.52) can be used to predict the characteristics of the output of a cascade of linear reservoirs, given a set of reservoir parameters for any given number of reservoirs.

Appendix C

UNIFORM COMPLETE OBSERVABILITY AND CONTROLLABILITY OF A SYSTEM IN CANONICAL FORM

The n^{th} order system,

$$\underline{r}_y(t_{k+1}) = \underline{\overline{\Phi}}_r(t_{k+1}, t_k) \cdot \underline{r}_y(t_k) + \underline{\overline{\Gamma}}_k \cdot w(t_k) \qquad (5.106)$$

with scalar observations,

$$z_r(t_k) = \underline{\overline{H}}_r(t_k) \cdot \underline{r}_y(t_k) + v(t_k); \quad k = 0, 1, 2, \ldots \qquad (5.107)$$

is said to be uniformly completely observable (Gelb, 1974) if there exist some integer $N > 0$ and positive intergers α_1, α_2 such that,

$$\alpha_1 \cdot \underline{\overline{I}}_{nn} \leq \sum_{i=k-N}^{k} \underline{\overline{\Phi}}_r^T(t_i, t_k) \cdot \underline{\overline{H}}_r^T(t_i) \cdot R^{-1}(t_i) \cdot \underline{\overline{H}}_r(t_i) \cdot \underline{\overline{\Phi}}_r(t_i, t_k)$$

$$\leq \alpha_2 \cdot \underline{\overline{I}}_{nn} \; ; \; \alpha_2 < \infty \qquad (C.1)$$

for all $k \geq N$.

Denote by $\underline{\overline{\Omega}}_{1_N}$ the matrix:

$$\underline{\overline{\Omega}}_{1_N} = \sum_{i=k-N}^{k} \underline{\overline{\Phi}}_r^{\;T}(t_i, t_k) \cdot \underline{\overline{H}}_r^{\;T}(t_i) \cdot R^{-1}(t_i) \cdot \underline{\overline{H}}_r(t_i) \cdot \underline{\overline{\Phi}}_r(t_i, t_k) \qquad (C.2)$$

and let,

$$\underline{\overline{H}}_r(t_i) = [h_{r_1}(t_i)h_{r_2}(t_i) \dots h_{r_n}(t_i)] \quad (1\text{xn}) \tag{C.3}$$

Assume that the system of Equations (5.106) and (5.107) is in canonical form. Then, the matrix $\underline{\overline{\Phi}}_r(t_j,t_k)$ is diagonal with $(i,i)^{th}$ element, $\phi_{r_{i,i}}(t_j,t_k)$. Under this condition the $(\ell,j)^{th}$ element $\omega_{1_{\ell,j}}$ of the matrix $\underline{\overline{\Omega}}_{1_N}$ can be written as,

$$\omega_{1_{\ell,j}} = \sum_{i=k-N}^{k} \phi_{r_{\ell,\ell}}(t_i,t_k) \cdot h_{r_\ell}(t_i) \cdot h_{r_j}(t_i) \cdot \phi_{r_{j,j}}(t_i,t_k) \cdot \frac{1}{R(t_i)}$$

$$; \ \ell,j = 1,2,\dots,n \tag{C.4}$$

or

$$\omega_{1_{\ell,j}} = \sum_{i=k-N}^{k} S_{\ell,i} \cdot S_{j,i} \quad ; \ \ell,j = 1,2,\dots,n \tag{C.5}$$

where

$$S_{\ell,i} = \frac{\phi_{r_{\ell,\ell}}(t_i,t_k) \cdot h_{r_\ell}(t_i)}{\sqrt{R(t_i)}} \quad ; \ \ell = 1,2,\dots,n \tag{C.6}$$

Let $\underline{\psi}$ be any non-zero real vector of dimension n. The matrix $\underline{\overline{\Omega}}_{1_N}$ is positive definite if

$$Y = \underline{\psi}^T \cdot \underline{\overline{\Omega}}_{1_N} \cdot \underline{\psi} > 0 \qquad \text{for all } \underline{\psi}.$$

Denote by ψ_i the i^{th} element of $\underline{\psi}$. Then, the quadratic form Y can be written as,

$$Y = \psi_1^2 \cdot \sum_{i=k-N}^{k} S_{1,i}^2 + \psi_2^2 \cdot \sum_{i=k-N}^{k} S_{2,i}^2 + \ldots + \psi_n^2 \cdot \sum_{i=k-N}^{k} S_{n,i}^2 +$$

$$+ 2 \cdot \psi_1 \cdot \psi_2 \cdot \sum_{i=k-N}^{k} S_{1,i} \cdot S_{2,i} + 2 \cdot \psi_1 \cdot \psi_3 \cdot \sum_{i=k-N}^{k} S_{1,i} \cdot S_{3,i}$$

$$+ \ldots + 2 \cdot \psi_{n-1} \cdot \psi_n \cdot \sum_{i=k-N}^{k} S_{n-1,i} \cdot S_{n,i} \qquad \text{(C.7)}$$

or

$$Y = \sum_{i=k-N}^{k} \left(\sum_{j=1}^{n} \psi_j \cdot S_{j,i} \right)^2 \qquad \text{(C.8)}$$

Since Y is equal to the sum of non-negative variables, it is equal to zero when all the variables are equal to zero. An equivalent statement will be that the set of N + 1 equations:

$$\sum_{j=1}^{n} \psi_j \cdot S_{j,i} = 0; \quad i = k-N, \ldots, k \qquad \text{(C.9)}$$

has a solution ψ_j^*, $j = 1,2,\ldots,n$, not equal to zero. In general, this will be the case for $N \geq n$.

The i^{th} transition function of the system of Equations (5.106) and (5.107) is,

$$\phi_{i,i}(t_j,t_k) = e^{\sum_{\ell=j}^{k-1} \beta_{i_\ell} \cdot \Delta t_\ell} \quad ; \qquad t_i < t_k,\ \beta_{i_\ell} > 0 \quad \forall \ell \tag{C.10}$$

where β_{i_k} is the i^{th} eigenvalue at time step k.

If, for k growing, the eigenvalues are finite; the elements of $\underline{\overline{H}}_r(t_k)$ are bounded; and $R(t_k)$ is not zero, then the second condition of Inequality (C.1) will not be violated. In this case the system is uniformly completely observable.

The system of Equations (5.106) and (5.107) is said to be uniformly completely controllable (Gelb, 1974) if there exist some positive integer N and two positive integers α_1, α_2 such that

$$\alpha_1 \cdot \underline{\overline{I}}_{nn} \leq \sum_{i=k-N}^{k-1} \underline{\overline{\Phi}}_r(t_k,t_{i+1}) \cdot \underline{\overline{\Gamma}}_i \cdot \underline{\overline{Q}}(t_i) \cdot \underline{\overline{\Gamma}}_i^T \cdot \underline{\overline{\Phi}}_r^T(t_k,t_{i+1})$$

$$\leq \alpha_2 \cdot \underline{\overline{I}}_{nn}; \qquad \alpha_2 < \infty \tag{C.11}$$

for all $k \geq N$.

Denote by $\underline{\overline{\Omega}}_{2_N}$ the matrix defined as

$$\underline{\overline{\Omega}}_{2_N} = \sum_{i=k-N}^{k-1} \underline{\overline{\Phi}}_r(t_k, t_{i+1}) \cdot \underline{\overline{\Gamma}}_i \cdot \underline{\overline{Q}}(t_i) \cdot \underline{\overline{\Gamma}}_i^T \cdot \underline{\overline{\Phi}}_r^T(t_k,t_{i+1}) \tag{C.12}$$

Let the $\underline{\overline{Q}}(t_i)$ matrix be diagonal with positive elements $q_{j,j}(t_i)$, $j = 1,2,\ldots,r$ and let the $(\ell,j)^{th}$ element of matrix $\underline{\overline{\Gamma}}_i$

be γ_{ℓ,j_i}, $\ell = 1,2,\ldots,n$, $j = 1,2,\ldots,r$. Then, given that the system is in canoninal form it holds

$$\omega_{2_{\ell,j}} = \sum_{i=k-N}^{k-1} \phi_{r_{\ell,\ell}} (t_k,t_{i+1}) \cdot \left(\sum_{\tau=1}^{r} \gamma_{\ell,\tau_i} \cdot q_{\tau,\tau}(t_i) \cdot \gamma_{\tau,j_i} \right)$$

$$\cdot \phi_{r_{j,j}}(t_k,t_{i+1}) \tag{C.13}$$

If $\underline{\psi}$ is any real vector with j^{th} element ψ_j, and the quadratic form:

$$Y = \underline{\psi}^T \cdot \underline{\overline{\Omega}}_{2_N} \cdot \underline{\psi} \qquad , \qquad \forall \underline{\psi} \neq 0 \tag{C.14}$$

is positive, then the matrix $\underline{\overline{\Omega}}_{2_N}$ is positive definite.

The above can be expressed as,

$$Y = \sum_{\ell=1}^{n} \sum_{j=1}^{n} \left[\sum_{i=k-N}^{k-1} \psi_\ell \cdot \phi_{r_{\ell,\ell}}(t_k,t_{i+1}) \cdot \right.$$

$$\left. \cdot \left(\sum_{\tau=1}^{r} \gamma_{\ell,\tau_i} \cdot q_{\tau,\tau} (t_i) \cdot \gamma_{\tau,j_i} \right) \cdot \phi_{r_{j,j}}(t_k,t_{i+1}) \cdot \psi_j \right] \tag{C.15}$$

Interchanging the summation operators yields,

$$Y = \sum_{i=k-N}^{k} \sum_{\tau=1}^{r} \left[\sum_{\ell=1}^{n} \sum_{j=1}^{n} \psi'_{\ell,\tau_i} \cdot \psi'_{\tau,j_i} \right] \tag{C.16}$$

where

$$\psi'_{\tau,j_i} = \sqrt{q_{\tau,\tau}(t_i)} \cdot \gamma_{\tau,j_i} \cdot \phi_{r_{j,j}}(t_k,t_{i+1}) \cdot \psi_j \ ;$$

$$j = 1,2,\ldots,n \qquad \text{(C.17)}$$

At this point the derivation is restricted by the requirement that $\gamma_{\ell,j_i} = \gamma_{j,\ell_i}$, $j,\ell = 1,2,\ldots,n$ for all i. This type of $\underline{\overline{\Gamma}}_i$ matrix is of use in this work. Under this condition,

$$\psi'_{\ell,\tau_i} = \psi'_{\tau,\ell_i} \quad , \quad \tau,\ell = 1,2,\ldots,n, \quad \forall i \qquad \text{(C.18)}$$

Equation (C.16) is equivalent now to,

$$Y = \sum_{i=k-N}^{k} \sum_{\tau=1}^{r} \left[\sum_{j=1}^{n} \psi'_{j,\tau_i} \right]^2 \qquad \text{(C.19)}$$

Based on Equation (C.19), Y is zero if and only if

$$\sum_{j=1}^{n} \psi'_{j,\tau_i} = 0 \ ; \quad \tau = 1,2,\ldots r, \quad i = k-N, \ldots, k \qquad \text{(C.20)}$$

For well behaved matrices $\underline{\overline{\Gamma}}_i$, $\underline{\overline{Q}}(t_i)$ and $\underline{\overline{\Phi}}_r(t_i,t_{i+1})$, it is clear that for $N \geq (n - r + 1)$ Equation (C.20) will give zero solution for the elements of the $\underline{\psi}$ matrix. For example if $\underline{\overline{\Gamma}}_i$

is diagonal with elements equal to unity, Equation (C.20) can be rewritten as

$$\sum_{j=1}^{n} \sqrt{q_{\tau,\tau}(t_i)} \cdot \phi_{r_{j,j}}(t_k, t_{i+1}) \cdot \psi_j = 0; \; \tau = 1,2,\ldots,r, \; i = k-N,\ldots,k \tag{C.21}$$

In this case one has to find the minimum N such that the resultant r + N simultaneous linear equations in the n variables ψ_j, j = 1,2,...,n, have the solution $\psi_j = 0$, j = 1,2,...,n, for time steps $k \geq N$.

For finite values of the elements of $\underline{\overline{\Gamma}}_i \cdot \underline{\overline{Q}}(t_i) \cdot \underline{\overline{\Gamma}}_i^T(t_i)$, it is clear from Equation (C.13) that the inequality in (C.11) is true, given that,

$$\phi_{j,j}(t_k, t_{i+1}) = e^{-\sum_{\ell=i+1}^{k-1} \beta_{i_\ell} \cdot \Delta t_\ell} \; ; \quad t_{i+1} < t_k,$$

$$\beta_{i_\ell} > 0 \qquad \forall \ell \tag{C.22}$$