AUTOMATIC PARAMETER ESTIMATION OF A LARGE CONCEPTUAL RAINFALL-RUNOFF MODEL: A MAXIMUM LIKELIHOOD APPROACH

by
PEDRO J. RESTREPO POSADA
and
RAFAEL L. BRAS

RALPH M. PARSONS LABORATORY
HYDROLOGY AND WATER RESOURCES SYSTEMS

Department of Civil Engineering
Massachusetts Institute of Technology

Report No. 267

This work was sponsored by
the Hydrologic Research Laboratory
of the National Weather Service,
U.S. Department of Commerce,
NA 80AA-H-00044.

March 1982

MIT

DEPARTMENT
OF
CIVIL
ENGINEERING

SCHOOL OF ENGINEERING
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Cambridge, Massachusetts 02139

R82-05 OSP 89310

AUTOMATIC PARAMETER ESTIMATION OF A LARGE CONCEPTUAL RAINFALL-RUNOFF MODEL: A MAXIMUM LIKELIHOOD APPROACH

by

PEDRO J. RESTREPO POSADA
and
RAFAEL L. BRAS

RALPH M. PARSONS LABORATORY
HYDROLOGY AND WATER RESOURCES SYSTEMS

Department of Civil Engineering
Massachusetts Institute of Technology

Report No. 267

This work was sponsored by the Hydrologic Research Laboratory of the National Weather Service, U. S. Department of Commerce, NA 80AA-H-00044.

March 1982

AUTOMATIC PARAMETER ESTIMATION OF A LARGE CONCEPTUAL RAINFALL-RUNOFF MODEL: A MAXIMUM LIKELIHOOD APPROACH

BY

PEDRO J. RESTREPO POSADA AND RAFAEL L. BRAS

ABSTRACT

A stochastic parameter estimation procedure to be applied to large conceptual rainfall-runoff models is proposed. The procedure is based on a maximum likelihood approach, which is enhanced to allow for the use of prior information about some of the parameters.

A stochastic model of base flow is proposed, and an algorithm for automatic identification of the time interval of base flow activity is developed. This algorithm, which permits the estimation of prior values for the base flow discharge coefficients, is successfully applied to two basins in the United States.

The application of the maximum likelihood estimation procedure to large conceptual rainfall-runoff models is divided into two stages. In the first stage the procedure is applied to a simplified version of the National Weather Service River Forecasting System model, the parameters of which are estimated by using synthetically generated data. This stage produces valuable information about the performance of stochastic parameter estimation procedures in large conceptual rainfall-runoff models. In the second stage the maximum likelihood procedure is applied to estimating the parameters of the National Weather Service River Forecasting System model for the Bird Creek and for the Cohocton River basins, respectively. In the first case a severe structural error arising from a defficient formulation of the channel causes some of the parameters to converge to unrealistic values. In the second case the stochastic model performs in a matematically correct but hydrologically unappealing fashion. The reasons for this problems are attributed to the non-identifiability of the threshold parameter of the upper zone tension water element and to interaction between several of the percolation function parameters.

C7438

ACKNOWLEDGEMENTS

This work was sponsored by the Hydrologic Research Laboratory of the National Weather Service, U. S. Department of Commerce, under Cooperative Agreement NA 80AA-H-00044. The financial support, as well as the valuable interaction with the HRL's staff is most greatfully achnowledged. Special mention is due to Dr. Michael Hudlow, Dr. Eric Anderson, Mr. George Smith, Dr. Edward Johnson, and Dr. Eugene Peck.

Administrative support was provided by M.I.T. through the Office of Sponsored Program, OSP 89310. All computer work was performed at the M.I.T. Information Processing Center.

Work was performed at the Ralph M. Parsons Laboratory, Water Resources and Environmental Engineering Division, M.I.T.

The comments and help of Professor Fred Schweppe, Professor Peter S. Eagleson, and Professor Daniele Veneziano is greatly appreciated. Discussions with Mr. Konstantine Georgakakos are also greatfully acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES

LIST OF SYMBOLS

Symbol	Description
a_1	Fraction of basin that becomes impervious when tension water requirements are met.
a_2	Fraction of basin permanently impervious.
A_1, A_2	Discrete-time parameters of the base flow model.
B	Input gain matrix.
d_k	Direction of search at iteration k.
d_1'	Lower zone primary instantaneous drainage coefficient.
d_1''	Lower zone supplementary instantaneous drainage coefficient.
d_u	Upper zone instantaneous drainage coefficient.
d_r	Defficiency ratio.
e_i	Evaporation from the "ith" element.
E[ ]	Expectation operator.
f_1	Overflow from the lower zone tension water.
f_u	Overflow from the upper zone tension water.
F	Fraction of percolating water to the free aquifers that goes to the primary aquifer.
g	Threshold function.
g_k	Gradient of $-\xi$ at iteration k.
G	Non-linear model error matrix.
G_c	Linearized complete-system input-gain matrix.
G_u	Unit hydrograph input-gain matrix.
H	Measurement matrix.

LIST OF SYMBOLS

Symbol	Description
i_f	Interflow.
i_r	Fraction of precipitation which becomes infiltrated.
L	Likelihood function.
L_w	Window's length.
M_1, M_2	Models of 1 and 2 state variables for identification of the interval of base flow .
n_x, n_z, n_r	Dimension of several vectors.
N_g	Number of non positive eigenvalues.
p'	Prior probability.
p"	Posterior probability.
P_b	Percolation base rate.
P_c	Percolation rate.
P_f	Fraction of percolated water assigned to the free water aquifers.
P_s	Slope changing parameter of the "S" curves.
P_x	Covariance matrix of the states.
P_z	Covariance matrix of the residuals.
P_r	Covariance of the parameter estimates. Percolating water.
P_r'	Prior covariance of the parameters.
P_0	Initial covariance matrix.
Q	Covariance matrix of the model error.
r	Fraction of the lower zone free water unavailable to supply lower zone tension requirements. Normalized residuals.

LIST OF SYMBOLS

Symbol	Description
r_b	Base flow
r_{bc}	Base flow, channel component.
r_{bn}	Base flow, non channel component.
r_d	Direct runoff
R	Coefficient of variation of the measurement error.
S_f	Shift parameter in the "S" curves.
S_k	Approximation to the inverse Hessian at iteration k.
S_R	Surface runoff.
S_s	Constant losses from the channel.
$S_1..S_5$	"S" curves.
t	Time step.
T_{ci}	Total channel inflow.
T_{co}	Total channel outflow.
T_r	Threshold for the normalized correlation coefficients.
u_1	Instantaneous precipitation input.
u_2	Instantaneous evapotranspiration demand.
v	Measurement error.
w	Noise term.
w	Weighting matrix. Model error vector.
$\underline{x}, \underline{x}$	State variable.
$\underline{x}_c$	Complete-system state vector.

LIST OF SYMBOLS

Symbol	Description
x_i^o	Parameter of the threshold function and "S" curves.
$\underline{x}_L$	Values at which the system is linearized.
$\underline{x}_0$	Initial conditions.
y	Defficiency ratio.
Z	Measurement vector. (Random vector).
$\underline{z}$	Realization of $\underline{Z}$.
$\underline{z}$	Predicted measurement.
α	Percolation function exponent.
α_k	Cummulant of steps at iteration k.
β	Component of the linearized total channel inflow.
β_u	Component of the linearized total channel inflow.
γ	Percolation function parameter.
δ	Kronecker's delta.
ε	Fraction of parameters for calculation of initial step size.
θ	Set of parameters.
θ^*	Maximum likelihood estimate of $\underline{\theta}$
κ	Normalizing constant.
ξ	Loglikelihood function.
ξ_b	Bias component of ξ.
ξ_o	Observations component of ξ.

LIST OF SYMBOLS

Symbol	Description
ξ_o'	Extended observations component of ξ_x.
ξ_x	Extended loglikelihood function.
ϕ	Non-linear system dynamics matrix.
μ	Fraction of base flow not appearing in river flow.
ν	Residuals.
σ_k^i	Step size in the linear search k.
τ_3	Approximation to the threshold function in x_3.
Φ_c	Linearized complete-system transition matrix.
Φ_u	Unit hydrograph transition matrix.

Chapter 1

INTRODUCTION AND REVIEW OF LITERATURE.

1.1 Introduction.

Flood forecasting requires the use of mathematical models to represent the very complex processes ongoing in a catchment during the rainfall-runoff process. According to the modeling philosophy, rainfall-runoff models can be divided into three categories: physical, conceptual, black box. Physical models are those that represent the process following the physical laws that govern it. Conceptual models do not follow the laws exactly, but have a structure that approximates the physical reality. The parameters of these models thus have some physical meaning. Black box models consist of a functional relationship between input and output, and have no physical meaning. Each of these three categories can be further subdivided into deterministic and stochastic models. Stochastic models are those in which the presence of random errors in the data and/or in the model is explicitly considered. Deterministic models do not take into account these errors.

The estimation of model parameters in real catchments can be done by experienced hydrologists by trial and error using manual procedures, with a reasonable degree of success. Since the number of skilled personnel capable of performing high quality estimation of parameters is very small, an automatic technique of parameter estimation is highly desirable.

This work will be directed toward the development and implementation of a methodology for automatic parameter estimation in stochastic rainfall-runoff conceptual models.

1.2 Review of the Literature.

1.2.1 Introduction.

A literature search of parameter estimation in rainfall-runoff models is necessary in order to clearly state the areas of the field in which research is needed. In order to take advantage of previous researchers' experience in parameter estimation, it is very important to search the respective literature in fields other than rainfall-runoff modeling. Therefore, a review of the research done in groundwater hydrology and water quality modeling will also be included here.

1.2.2 Parameter Estimation in Rainfall-Runoff Modeling.

Dawdy and O'Donnell (1965) published the first known paper on automatic parameter estimation of a conceptual rainfall-runoff model. The model chosen was developed by O'Donnell, and the nonlinear optimization technique was

Rosenbrock's (1960). This technique, also known as "rotating coordinates method", is a ridge-following algorithm in which the axes of coordinates are rotated so that one of them coincides with the ridge. These authors first generated a synthetic series of streamflows from a given rainfall-evaporation record. Then, they proceeded to evaluate the performance of the optimization algorithm in finding the correct set of parameters, when an initial set different from the one used for generating the flows was given. This particular technique of using synthetic series enabled these authors to study the performance under ideal circumstances, since no input, output or structural errors were present.

Chapman (1970) compared three non-linear optimization methods for a deterministic conceptual model developed for the Australian Water Resource Council. This author judged the Simplex method (Nelder and Mead, 1965) highly superior to the steepest descent and univariate minimization techniques when applied to conceptual models in real catchments.

Murray (1970) used Boughton's (1965) deterministic daily rainfall-runoff model to study a 20.2 km^2 catchment in North Wales. An automatic parameter estimation routine that made use of Rosenbrock's (1960) algorithm was included. The initial values for the parameters were obtained manually from the original record. Some parameters had to be constrained in order to prevent them from taking unrealistic values. Other parameters were left unchanged by the optimization algorithm.

Liou (1970) developed a computer program that automatically adjusted the parameters of the Stanford Watershed Model using a least squares criterion. The optimization method was based on a trial and error scheme which in essence was an automation of the manual procedure.

Grunewald and Dyck (1971) used two different optimization techniques, linearization of normal equations, LNE, and statistical optimization (a regression analysis) with three linear deterministic conceptual models.

Ibbitt and O'Donnell (1971) presented an excellent study of the problems involved in the automatic optimization of parameters in conceptual models. The model chosen for their study was Dawdy and O'Donnell's. They presented a heuristic discussion of some of the potential difficulties in the optimization of parameters in conceptual models. They discussed problems with saddle points, valleys and plateaus. Some concrete examples found in their specific case were illustrated, as the multiple optima situation. A useful discussion on the interaction between parameters was included.

Monro (1971) implemented a pattern search method to optimize the parameters of the National Weather Service Rainfall-Runoff model. The results showed that unrealistic values for the parameters were frequently obtained.

Ibitt (1972) studied the effect on parameter estimation of the random errors in the data. This author used Rosenbrock's optimization method, modified to handle the "peculiar fitting problems" associated with conceptual catchment models. He concluded that, despite the high influence of errors in the runoff on the value of the objective function, the final value of the parameters was not "much worse" than in the case of those obtained with error-free data.

Haan (1972) used a simple four parameter self-calibrating deterministic model to predict monthly runoff from daily rainfall in small catchments. The optimization procedure was a one-dimensional search of one parameter at a time, using prespecified steps.

Bultot and Dupriez (1976) developed a conceptual deterministic model with 21 seasonally varying parameters. The performance of the model was judged very satisfactory with manually fitted parameters.

Moore and Weiss (1976) reported the application of an Extended Kalman Filter for on-line estimation of the four parameters of a non-linear model. Their approach may be considered as a particular case of the state augmentation technique. It was validated in a real situation.

Johnston and Pilgrim (1976) compared the performance of two numerical techniques, Simplex and Davidon-Fletcher-Powell methods. This last method is one of the "Quasi Newton" methods in which an approximation of the inverse Hessian matrix of the objective function is constructed at every time step. These authors reported convergence of both methods to local minima. The optimization procedure was highly improved after the mathematical optimization techniques were modified to account for the special characterisitics of the conceptual rainfall-runoff models.

Singh (1977) used a modified version of Rosenbrock's algorithm to estimate the three parameters of Dooge's (1967) cascade of non-linear reservoirs model. Two of the parameters were assumed to be constant for basins of the same size. A regression equation between the third parameter and some physical measurements of the basin was developed.

Sugawara (1979) developed a trial and error procedure to find the optimal set of parameters for the tank model. This deterministic model is a cascade of linear and non-linear reservoirs. The fitting criterion was to minimize the deviations between the shapes of the predicted and the observed hydrograph and to minimize, simultaneously, the error in the predicted time to peak. The author claimed a high rate of convergence (less than 15 iterations).

Bras and Restrepo-Posada (1980) used the state augmentation technique for on-line estimation of two parameters of a four parameter non-linear conceptual catchment model. The results showed that filtering techniques were successful in minimizing the impact of noisy synthetic data on the estimation of parameters. The effect of the structural errors was

studied by generating synthetic data with a different model and then trying to estimate the parameters of the original model. It was shown that structural errors prevent the convergence of the parameters to a single value. In addition, these authors showed that simple heuristic techniques can also be successfully used in estimating the parameters of a model.

Fjeld and Aam (1980) published a paper on the use of estimation techniques with a hydrologic model which is applied to the prediction of daily runoff into a hydroelectric power station. The authors emphasized the advantages of a conceptual model over a linear ARMA model in modeling runoff under Nordic conditions. This was due to the fact that the inclusion of freezing and snow melting during the winter season implies either a change in the structure or large parametric changes on an ARMA model. The stochastic conceptual model used by these authors was based on the deterministic model developed by Bergstrom and Forsman (1973), and Bergstrom (1975), which is considerably simpler than that of the Stanford-type models. All parameters of the model were coarse calibrated with a least square criterion. Only the runoff discharge coefficients were "fine tuned" with a maximum likelihood criterion. The calibrated model has been operational at the Tonstad power station on the Sira Kvina river in southern Norway.

Soorooshian and Dracup (1980) used a power transformation technique such that the variance of the transformed flows is the same at every time step. The objective function was a least squares criterion of the transformed flows. These authors reported a good rate of convergence in the solution procedure.

Sooroshian (Unpublished report, 1981), working with a reduced set of parameters of the National Weather Service model, has shown the occurence of ridges and local optima in the loglikelihood function of that particular model. Further discussion of these issues is left to later chapters.

1.2.3 Stochastic Parameter Estimation in Groundwater Hydrology and Water Quality Modeling.

Currently, there is a large number of ongoing research dedicated to the problem of parameter identification in groundwater and water quality models. Unfortunately, most approaches to the problem have been from a deterministic point of view. In the area of stochastic parameter estimation are the works of Beck (1974, 1976) in which a maximum likelihood methodology was used to estimate the parameters of a water quality model. Moore and Jones (1977) developed a coupled Bayesian-Kalman Filter parameter estimation scheme for a dynamic water quality model. McLaughlin (1975), after comparing several deterministic optimization techniques, strongly recomended the use of linear filtering techniques in order to minimize the impact of noise in the measurements on the estimates of the parameters in groundwater models. Lin and

Yeh (1974) applied quasi-linearization and optimal control, to identify 11 parameters of an inhomogeneous aquifer.

Extensive literature on parameter estimation is available in fields other than water resources. The reader is referred to an excellent survey published by Astrom and Eykhoff (1970).

1.2.4 Summary of the Review.

The previous section revealed the following points:

1) There is a strong evidence of the unfavorable impact that data errors and structural errors have on the estimates of the parameters of conceptually-based rainfall-runoff models.

2) Stochastic conceptual rainfall-runoff models with a built-in filtering procedure have better performance than comparable deterministic models. Similarly, filtering techniques have been proven useful in minimizing the effect of data errors in the estimation of the parameters.

3) In spite of 1) and 2), few applications of stochastic parameters estimation to conceptual rainfall-runoff models have been recorded.

4) It has been shown that a good set of parameters can be obtained by means of heuristic procedures. Moreover, the performance of pure mathematical techniques on parameter estimation can be greatly improved by heuristically modifying such techniques in order to account for particular characteristics of conceptual rainfall-runoff models.

1.3 Scope and Organization of the Research.

This research aims to solve some of the problems discussed in the previous section. A methodology for stochastic parameter estimation in large conceptual rainfall-runoff models is developed in order to explicitly consider the problems posed by noise in the system and in the measurements on the parameter estimates. In addition, an algorithm is developed which identifies the interval of the records in which base flow is the main streamflow component, and estimates the corresponding parameters. This is done to circumvent observability problems that may lead to unrealistic parameter values. The final purpose of the research is to apply the methodology developed to real catchments. The rainfall-runoff model used is based on the National Weather Service River Forecasting Model, which will be described later on this work.

The research was divided into three stages. The first stage was the development of a stochastic parameters estimation procedure and the identification of possible problems associated with the use of the procedure. Due to the high cost of performing tests on the complete National Weather Service model a simplified version of the model was developed which was less expensive to run, but kept the basic structure and characteristics of the original.

The second stage was directed toward the development of an algorithm for automatic identification of the time interval of base flow activity. This task involved the implementation and verification of a stochastic model of Base Flow and the testing of the algorithm with real data.

The third and final stage applied the methodologies developed in the first two stages to estimate parameters of the National Weather Service river forecasting model in two river basins. The performance of the model under the parameters thus estimated was also studied.

1.4 Thesis Outline.

This thesis is divided into 8 chapters. Chapter 1 contains the introduction, literature review, scope and outline of the research, and outline of the thesis. In Chapter 2 several parameter estimation techniques are presented. Chapter 3 includes a description of the maximum likelihood parameter estimation technique as applied to the present research. The basics of the National Weather Service model, not covered elsewhere, and the simplified rainfall-runoff model are described on Chapter 4. The algorithm for automatic identification of base flow activity is presented in Chapter 5. Chapter 6 describes the identification of the parameters of the simplified rainfall-runoff model and the problems found. Chapter 7 is concerned with the application of the complete methodology to the Bird Creek and Cohocton River basins, using the National Weather Service model. The conclusions and recomendations are included in Chapter 8.

CHAPTER 2

PARAMETER ESTIMATION TECHNIQUES: A BRIEF SUMMARY.

2.1 Classification of Parameter Estimation Techniques.

Mathematical models of physical systems can be divided into two broad categories, according to whether the variables representing the state of the system are considered random variables or deterministic variables. These two categories are: 1) stochastic and 2) deterministic models. Similarly, the parameter estimation techniques are also divided into "probabilistic" or "deterministic". Probabilistic techniques are also known as "stochastic" techniques when the system under consideration is dynamic.

2.2 Brief Survey of Parameter Estimation Techniques.

Section 2.2 contains a brief description of some of the parameter estimation techniques, although by no means it is an exhaustive survey. Furthermore, not all of these techniques are applicable to all kind of problems.

2.2.1 Deterministic Methods.

2.2.1.1 Algebraic.

This method is applicable to those models for which it is possible to formulate a system of linear algebraic equations in which the parameters are the unknown and the coefficients of the equations are functions of the (known) state variables. If there are k_1 unknown parameters, k_1 measurements will be needed. A system of k_1 simultaneous equations with k_1 unknowns is solved (see McLaughlin, 1979, and Sagar, et al., 1975). The simplicity of this method is counterbalanced by its sensitivity to measurement errors, commonly yielding physically unrealistic parameter estimates (McLaughlin, 1979).

2.2.1.2 Mathematical Programming

This method is derived from the algebraic method. More measurement equations are used in order to solve the problems presented by the algebraic method. This will lead to a system of k_2 linear equations with k_1 unknowns, with $k_1 < k_2$. This implies that all equations can not be satisfied simultaneously and there will be an error in $k_2 - k_1$ equations. The algebraic equations are thus modified in order to include this error and a linear programming problem is set up, in the case of a linear objective function, such that a chosen function of the error terms is minimized. One of the main problems of this method is the frequent occurrence of degener-

acy, which has been reported by Kleinecke (1971) and McLaughlin (1975). A non linear programming technique will be required if the objective function is non-linear (Yoon and Yeh, 1976, Chang and Yeh, 1976).

2.2.1.3 Weighted Least Squares

The goal is to find a set of parameters, $\underline{\theta}$, such that the weighted sum of squares of the errors between the predicted and the measured values is minimized:

$$\underset{\underline{\theta}}{\text{Min}}\ \underline{\varepsilon}^{T}(\underline{\theta})\ \underline{W}(\underline{\theta})\ \underline{\varepsilon}(\underline{\theta}) \qquad (2.1)$$

where $\underline{\varepsilon}$ is the (column) vector containing the differences between the predicted and the measured values and $\underline{W}$ is the weighting matrix. It may be chosen by "engineering judgement" in a way such that more weight is placed on some measurements than on others. It is important to mention the fact that when the system under consideration has been modeled as a stochastic process, the weighting matrix that yields the smallest error is $\underline{W} = \underline{P}_z$, where $\underline{P}_z$ is the covariance matrix of the sampling error vector.

This is a very powerful and general approach and has been widely used in the past.

2.2.1.4 Heuristic Methods

This class of parameter estimation methods covers procedures based mainly on experiences with a given system. The estimation procedure is frequently carried out by trial and error, either manually or by means of an algorithm programmed into a computer. Alternatively, heuristic methods can be used to calculate the gain matrix for the feedback term in the stochastic approximations algorithm, which will be presented later in this chapter. Although the first reaction of many people to these methods is unfavorable, experience has shown that skillful and experienced technicians can get a good set of parameters by taking advantage of their knowledge of the physical system being modeled.

2.2.2 Stochastic Methods

2.2.2.1 State Augmentation

This simple technique can be applied in the case of stochastic dynamic models of the form (in the general, non-linear case),

$$\underline{x}(t+1) = \Phi(\underline{x}(t), \underline{\theta}(t), t) + G(\underline{x}(t), \underline{\theta}(t), \underline{w}, t) \qquad (2.2)$$

$$\underline{z}(t+1) = H(\underline{x}(t+1), \underline{\theta}(t+1), t+1) + \underline{v}(t+1) \qquad (2.3)$$

where

$\underline{x}(t)$ = state vector at time t
$\underline{\theta}(t)$ = unknown parameters vector, at time t
$\Phi(.,.,.)$ = non-linear state transition matrix
$G(.,.,.)$ = non-linear model error matrix
$\underline{w}(t)$ = noise term
$H(.,.,.)$ = non-linear measurement matrix
$\underline{v}(t)$ = measurement error
$\underline{z}(t)$ = vector containing the measurements at time t

For the most common case of constant parameters, the "state dynamics" equation of the parameters is,

$$\underline{\theta}(t+1) = \underline{\theta}(t) \tag{2.4}$$

Equations (2.2) through (2.4) are thus combined to yield:

$$\tilde{\underline{x}}(t+1) = \phi(\tilde{\underline{x}}(t),t) + G(\underline{x}(t),\underline{w}(t),t) \tag{2.5}$$

$$\underline{z}(t+1) = \tilde{H}(\tilde{\underline{x}}(t+1),\ t+1) + \underline{v}(t+1) \tag{2.6}$$

where

$$\tilde{\underline{x}}(t) = \begin{vmatrix} \underline{x}(t) \\ \underline{\theta}(t) \end{vmatrix}$$

$$\phi = \begin{vmatrix} \dfrac{\Phi(\underline{x}(t),\ \underline{\theta}(t),\ t)}{\underline{\theta}(t)} \end{vmatrix}$$

and $\tilde{H}$ is different from H only in the notation, since the parameter array $\underline{\theta}(t)$ is now included in $\tilde{\underline{x}}(t)$.

The state variables of the system represented by Equations (2.5) and (2.6) are commonly estimated by linear filtering techniques, with appropriate algorithms to deal with the non-linearity of these equations.

The simplicity of this method and the ease of implementation are highly counterbalanced by the following two problems. First, the non-linearity of Equations (2.5) and (2.6) decreases such linear systems' advantages as the propagation of Gaussian properties. Second, the number of operations necessary to carry out the state estimation by one of the most commonly used linear filtering techniques, known as the Extended Kalman Filter, is proportional to the third power of the number of state variables. Augmenting the state vector with the unknown parameters will thus cause costs which are proportional to the third power of the size of the resulting state vector.

2.2.2.2 Stochastic Approximations

This name applies to a family of iterative techniques that are used to find minima of functions without requiring too much knowledge about these functions. In the case of dynamic systems with constant parameters a parameter dynamics equation as Equation (2.4) is formulated. Estimates are a linear combination of old estimates and a correction term based on the difference between a predicted measurement from the system and the actual observation. The correction term is weighted with an appropriate gain matrix:

$$\underline{\theta}(t+1|t) = \underline{\theta}(t|t) \tag{2.7}$$

$$\underline{\theta}(t+1|t+1) = \underline{w}(t+1)\,[\underline{z}(t+1) - \hat{\underline{z}}(t+1)] + \underline{\theta}(t+1|t) \tag{2.8}$$

where

$\underline{\theta}(s|t)$ = estimate of the parameters' vector at time s, with measurements taken up to time t

$\underline{w}(t)$ = weighting matrix

$\underline{z}(t)$ = actual observation at time t

$\hat{\underline{z}}(t)$ = predicted measurement at time t

Equations (2.6) and (2.7) are used in combination with Equations (2.2) and (2.3), and with the estimation of states. The simplicity of this family of algorithms is the cause of their main disadvantage. Since, in general, little information about the system is built into the algorithm, the convergence is usually slow (Schweppe, 1973, p. 358).

2.2.2.3 Maximum Likelihood Estimation.

This is a very powerful, though conceptually simple, technique. The parameter estimation algorithms developed during the course of this research are based on the maximum likelihood criterion. The next chapter will be fully dedicated to the theory and implementation of this technique.

CHAPTER 3

MAXIMUM LIKELIHOOD PARAMETER ESTIMATION

3.1 Introduction

The method of maximum likelihood is widely used for parameter estimation. Its popularity stems from a set of valuable properties which will be outlined later on this chapter. The method is very well described in many statistics textbooks and the uninitiated reader is particularly refered to Benjamin and Cornell (1971) for a simple but complete description. We will, otherwise, assume that the reader is familiar with the essential concepts behind maximum likelihood procedures.

3.2. Maximum Likelihood Estimation

Let $\underline{Z}$ be a set of identically distributed random variables, and let

$$p_{\underline{Z}}(\underline{Z};\underline{\theta}) \tag{3.1}$$

be the joint probability density function of that set, in which $\underline{\theta}$ is the set of unknown parameters of that distribution. Given a set of measurements on $\underline{Z}$, say $\underline{z}$, the likelihood function of the sample $\underline{z}$ is defined by equation (3.2)

$$L(\underline{\theta}|\underline{z}) = p_{\underline{Z}}(\underline{z};\underline{\theta}) \tag{3.2}$$

Having defined the likelihood of the sample by Equation (3.2), the criterion of maximum likelihood can now be defined:

"The maximum likelihood estimator of $\underline{\theta}$ is the value $\underline{\theta}^*$ that maximizes Equation (3.2)"

For some probability density functions, it is more convenient to define the estimator in terms of the logarithm of the likelihood function, which is called the loglikelihood function, and which will be denoted in this work by ξ. Since the logarithm is a monotonically increasing function, the value $\underline{\theta}^*$ that maximizes the likelihood function also maximizes the loglikelihood function.

The maximum likelihood estimators, hereon $\underline{\theta}^*$, have a series of properties which are independent of the underlying distribution. As mentioned earlier, we will list these properties without a formal proof. The reader is again

refered to Benjamin and Cornell (1971) for examples of the applications and for additional references. For large n, where n is the number of data values, the estimator, $\underline{\theta}^*$, is approximately normal, with mean asymptotically equal to the true parameter value; that is, $\underline{\theta}^*$ is asymptotically unbiased. Its mean square error is asymptotically equal to

$$\mathrm{Var}[\theta_i^*] = -[nE[\partial^2\xi/\partial\theta_i^2]]^{-1}$$

where subscript i refers to a given element of vector $\underline{\theta}$.

Also asymptotically, the maximum likelihood estimator has the minimum expected squared error of all possible unbiased estimators: it is efficient. It also is consistent, since with increasingly high probability they will be arbitrarily close to the true values of the parameters. Finally, $\underline{\theta}$ makes maximum use of the information contained in the data; that is, they are sufficient.

3.3 Maximum Likelihood Estimation in Linear Dynamic Systems.

Let a discrete-time linear dynamic system be

$$\underline{x}(t+1) = \Phi\ \underline{x}(t) + B\ \underline{u}(t) + G\ \underline{w}(t) \qquad (3.3)$$

and the associated measurement equation

$$\underline{Z}(t) = H\ \underline{x}\ (t) + \underline{v}(t) \qquad (3.4)$$

where:

$\underline{x}(t)$:	State variable at time t. (An n_x vector)
Φ:	State transition matrix. ($n_x \cdot n_x$)
B:	Matrix relating the exogenous inputs to the state vector. ($n_x \cdot n_u$)
$\underline{u}(t)$:	Exogenous input. (An n_u vector)
G:	Matrix ($n_x \cdot n_x$)
$\underline{w}(t)$:	Discrete time white gaussian noise. (An n_x vector)
$\underline{Z}(t)$:	Measurements. (An n_z vector)
H:	Measurement matrix. ($n_z \cdot n_x$)
$\underline{v}(t)$:	White gaussian noise. (An n_z vector)

in which

$\underline{x}(0)$: Normally distributed initial conditions for the state vector

$E[\underline{w}(t)]=\underline{0}$

$E[\underline{v}(t)]=\underline{0}$

$E[\underline{w}(t)\underline{w}^T(s)]=Q\delta_{s,t}$

$E[\underline{v}(t)\ \underline{v}(s)^T]=R\delta_{s,t}$

$E[.]$ Expectation operator

$\delta_{i,j}$: Kronecker's delta

Superscript T indicates the transpose of a vector or a matrix

Assume now that some of the elements of the matrices Φ, B, H, Q, R are unknown and we want to apply the maximum likelihood criterion to estimate these elements of the system. To do this we have to form the likelihood function of the unknown parameters, given the measurements, $L[\underline{\theta}|\underline{z}(1),...\underline{z}(T)]$, or the loglikelihood function $\xi[\underline{\theta}|\underline{z}(1),....\underline{z}(T)]$. We therefore need the pdf of the output of the system, $p_{Z(1),...Z(T)}[\underline{z}(1),...\underline{z}(T)]$.

Linear systems have the property of propagating the gaussian quality of the inputs into the system. Since the measurements are also a linear function of the states, the set of $\underline{Z}(t)$, for $t > 1$, will also be normally distributed. Moreover, the probability density function of $\underline{Z}(t)$ given the measurements up to time T, can be calculated by linear estimation techniques. One of these techniques is the Kalman Filter, the description of which is standard material in modern estimation textbooks. The reader is referred in particular to Schweppe (1973) and to Gelb (1974) for formal presentation and development of the filter. We will limit our discussion to the presentation of the equations required for the calculation of ξ in linear dynamic systems, and in the next section for the non-linear dynamic systems.

The stochastic process $\underline{Z}_T$ is defined as:

$$\underline{Z}_T{}^T=[\underline{Z}(1). \ . \ . \ \underline{Z}(T)]$$

and it is assumed to be normally distributed.

Once the measurements $\underline{z}_T$ of the stochastic process $\underline{Z}_T$ have been made, the loglikelihood function of these measurements can be calculated from the joint Normal distribution of $\underline{Z}_T$,

$$2\xi(\underline{\theta}|\underline{z}_T] = \xi_b(\underline{\theta}|\underline{z}_T) + \xi_o(\underline{\theta}|\underline{z}_T) \qquad (3.5)$$

The terms ξ_b and ξ_o are known as the bias component and the observations component of the loglikelihood function. These terms are defined by

$$\xi_b(\underline{\theta}|\underline{z}_T) = -T\, n_z\, \ell n\, (2\pi) - \Sigma\, \ell n|P_z(t|t-1:\underline{\theta})| \qquad (3.6)$$

$$\xi_o(\underline{\theta}|\underline{z}_T) = -\,\Sigma \nu^T(t:\underline{\theta})\, P_z(t|t-1:\underline{\theta})^{-1}\, \nu(t:\underline{\theta}) \qquad (3.7)$$

where

$\nu(t:\underline{\theta}) = \underline{z}(t) - \hat{\underline{z}}(t|t-1:\underline{\theta})$

$\hat{\underline{z}}(t|t-1:\underline{\theta})$: Best estimate of $\underline{z}(t)$, given measurements up to t-1

$P_z(t|t-1:\underline{\theta}) = E[\underline{\nu}(t:\underline{\theta})\cdot\underline{\nu}(t:\underline{\theta})^T]$

From Equation (3.7), we see that ξ_o is composed by the sum of the squares of the residuals, normalized by their variances. This is an important term which will be used in one of the post-optimality tests to be defined later on. The values $\hat{\underline{z}}(t|t-1)$ and $P_z(t|t-1)$ are calculated by means of the well known Kalman Filter equations. For the sake of completeness, these equations are included here. In the following set of equations, the dependence of the state variables on the unknown set of parameters $\underline{\theta}$ is explicitly shown. However, for shortness in notation, the time and $\underline{\theta}$ dependence of the matrices Φ, H, B, Q, R, is not shown.

$$\hat{\underline{x}}(t+1|t+1:\underline{\theta}) = \Phi\, \underline{x}(t|t:\underline{\theta}) + B\, \underline{u}(t) + P_x(t+1|t:\underline{\theta})H^T P_z^{-1}(t+1|t:\underline{\theta})\cdot \underline{\nu}(t+1:\underline{\theta}) \qquad (3.8)$$

$$\hat{\underline{z}}(t+1|t:\underline{\theta}) = H[\Phi\hat{\underline{x}}(t|t:\underline{\theta}) + B\underline{u}(t)] \qquad (3.9)$$

$$P_x(t+1|t+1:\underline{\theta}) = P_x(t+1|t:\underline{\theta}) - P_x(t+1|t:\underline{\theta})H^T P_z^{-1}(t+1|t:\underline{\theta})\cdot HP_x(t+1|t:\underline{\theta}) \qquad (3.10)$$

$$P_x(t+1|t:\underline{\theta}) = \Phi P_x(t|t:\underline{\theta})\Phi^T + GQG^T \qquad (3.11)$$

$$P_z(t+1|t:\underline{\theta}) = R + HP_x(t+1|t:\underline{\theta})H^T \qquad (3.12)$$

with initial conditions

$$P_x(0|0:\underline{\theta}) = P_0$$
$$\underline{x}(0|0) = \underline{0}$$

System identification is per se an off-line procedure. This is, for some historic data one looks for the best set of parameters that fit a chosen model. In contrast to off-line procedures, some models require the use of "adaptive" parameters, so that short term inadequacies of the model in representing real life can be corrected "on-line", that is, in real time. This is particularly true for the case of linear models that approximate non-linear processes. This modification of the parameters to account for the most recent errors in the prediction necessarily hampers the long range forecasting capability of the models. It seems then natural that models that accurately represent a real process should be calibrated "off-line".

Some other considerations, such as the cost of the estimation scheme, play a role in deciding on whether to develop an on-line or an off-line parameter estimation procedure. As a rule of thumb, on-line parameter estimation schemes may be cheaper to use that off-line parameter estimation schemes. This is due to the fact that in an on-line scheme the time-steps to calculate the state variables are combined with successive approximations to the parameter's values, following a scheme similar to the stochastic approximations discussed in Section 2.2.2.2. In the case of maximum likelihood estimation of dynamic systems, however, there is no way of calculating exactly the loglikelihood function on-line, at a given time step T, without having to redo all the computations from t=1 to t=T. Since the cost of doing this will eliminate the cost-advantage of the on-line methods, approximations to the loglikelihood function must be used. This approach, however, invalidates the properties of the maximum likelihood estimators, since the true loglikelihood is not being used. In general, we can say that the better the approximation, the highest the cost. There is, therefore, a trade off between cost and accuracy. The approach to be followed in this work is the off-line procedure. Since the NWSRFS is a non-linear dynamic model, the maximum likelihood estimation procedure in non-linear dynamic systems is presented next.

3.4 Maximum Likelihood Estimation in Non-linear Dynamic Systems.

Non-linear dynamic systems do not propagate the gaussian property of the inputs through the system. Therefore the results of optimum linear estimation theory, in which the Kalman Filter is based, do not fully apply. However, by

linearization of the non-linear equations the results of the previous sections can be used. The degree of success in applying those results depends not only on the linearization scheme but on the number of measurements taken to calculate the likelihood function. A large number of measurements lessen the impact of the gaussian assumption, and hence, of the non-linearities in the system. (Schweppe, 1973, p. 442). There are different techniques by which the system can be linearized. Kitanidis and Bras (1980) used two of these techniques in linearizing the continuous-time equations of the NWSRFS. They utilized statistical linearization and Taylor's series expansion to the first order approximation. The system can then be written in a discrete-time linear form, as explained in the above reference, and the calculation of the loglikelihood function is performed as in the linear case.

3.5 Incorporation of Hydrologic Information into the Stochastic Parameter Estimation Procedure.

One of the major problems addressed in the automatic parameter estimation literature is the convergence of some prameters to 'unrealistic' values. This immediately suggests the idea of the existance of prior knowledge about the range of some of the parameters. If this is the case, this prior information should be incorporated within the parameter estimation procedure.

In a Bayesian framework, this amounts to maximizing the posterior probability which is proportional to the product of the likelihood function and the prior probability:

$$p''(\underline{\theta}) = \kappa \; L\,(\underline{\theta}|\underline{z}_T) \cdot p'(\underline{\theta}) \qquad (3.13)$$

in which

p":	posterior probability
L:	likelihood Function
κ:	normalizing constant
p':	prior probability

Taking logarithms in both sides of Equation (3.13),

$$\ell n\, p''(\underline{\theta}) = \ell n\, \kappa + \xi(\underline{\theta}|\underline{z}) + \ell n\, p'(\underline{\theta}) \qquad (3.14)$$

The problem can be formulated as

$$\underset{\underline{\theta}}{\text{Max}} \; [\xi(\underline{\theta}|\underline{z}) + \ell n(p'(\underline{\theta}))] \qquad (3.15)$$

Notice that since $\ell n\ \kappa$ is a constant it was taken out of Equation (3.15) without affecting the outcome of the maximization process. The importance of this approach is that not only the initial estimates, but also the quality of these estimates is taken into account. The acquisition of the prior values and variances of the parameters is the object of Chapter 5. Notice that Equation (3.15) was derived from a Bayesian point of view. We can, however, look at it from a Fisher's point of view, in which the prior information on the parameters is added to the information coming from the system's output, to form a joint loglikelihood function. This will be especially clear for the case of $p'(\underline{\theta})$ being a Normal distribution. In this case, the function p' is fully described by its first two moments, the prior values of the parameter's estimates as the first moment, $\underline{\theta}'$, and the prior covariance of these estimates as the second moment, P_θ'. Under the Normal assumption, the logarithm of the prior that enters in Equation (3.15) can be written,

$$2 \cdot \ell n\ p'(\underline{\theta}) = -n_\theta(2\Pi) - \ell n|P_\theta|^{-1} + 2\xi_{\theta'}(\underline{\theta}|\underline{\theta}') \qquad (3.16)$$

$$\xi_{\theta'}(\underline{\theta}|\underline{\theta}') = 1/2 \cdot (\underline{\theta}-\underline{\theta}')^T \cdot P_\theta^{-1} \cdot (\underline{\theta}-\underline{\theta}') \qquad (3.17)$$

where n_θ is the number of parameters with prior information. Of the three terms on the right hand side of the equal sign in Equation (3.16), the first two will remain constant under the posterior optimization of Equation (3.15) and can be dropped without affecting the outcome. The remaining term is the square of the deviations of the prior values from the final ones, weighted by the prior variance of estimation. This term acts as an extended observation component of the loglikelihood (ξ_o). The extended observations component will consist of the T measurements of discharge and n_θ observations on the parameters. We can combine Equations (3.15) and (3.17) in the form of Equation (3.5), to arrive at

$$2\xi_x(\underline{\theta}|\underline{z}_T,\underline{\theta}') = \xi_b(\underline{\theta}|\underline{z}_T) + \xi_o'(\underline{\theta}|\underline{z}_T,\underline{\theta}') \qquad (3.18)$$

$$\xi_o'(\underline{\theta}|\underline{\theta}') = \xi_o(\underline{\theta}|\underline{z}_T) + \xi_{\theta'}(\underline{\theta}|\underline{\theta}') \qquad (3.19)$$

In which

$\xi_x(\underline{\theta}|\underline{z}_T,\underline{\theta}')$: Extended loglikelihood of the parameters

given the discharge measurements and the prior estimates of the parameters

$\xi_o'(\underline{\theta}|\underline{\theta}')$: Extended observations component

Two final observations can be made about $\xi_\theta'(\underline{\theta}|\underline{\theta}')$: First, if the prior estimates of the parameters form an independent set, P_θ will be a diagonal matrix, and $\xi_\theta'(\underline{\theta}|\underline{\theta}')$ can be written as

$$\xi_\theta(\underline{\theta}|\underline{\theta}') = 1/2 \quad (\theta_i-\theta_i')^2/\sigma_{\theta i}^2 \qquad (3.20)$$

In which $\sigma_{\theta i}^2$ is the prior variance of estimation of the parameter θ_i. Second, there are no restrictions on the number of observations of each parameter. Each independent observation can be included weighted by its variance of estimation.

3.6 Details of The Maximum Likelihood Algorithm

3.6.1 Search Algorithm.

The optimization of the likelihood function requires the use of non linear, unconstrained optimization schemes. These non linear optimization algorithms are based on the systematic search for the maximum of the objective function along a sequence of straight line searches. In each of these searches, a one dimensional optimization problem is solved. A new direction of search is chosen once an optimum is found. The way in which the direction of these line searches is chosen defines the difference among the non linear optimization procedures.

The efficiency of a non linear optimization algorithm is measured in terms of the rate of convergence of the solution toward the optimum point. The most efficient algorithms are also the most demanding ones in terms of computational requirements. Therefore, there is a trade-off between cost of computation and efficiency of the algorithm. The ease of implementation of the algorithm is also a factor.

An algorithm that offers an excellent compromise between these three factors is the Davidon-Fletcher-Powell (hereafter refered to as DFP) algorithm (Luenberger, 1973) which was selected as the optimization algorithm for the problem at hand. The DFP method belongs to the class of quasi-Newton methods in which the inverse Hessian matrix is approximated iteratively starting from any positive definite matrix. The inverse Hessian calculated by DFP is exact for quadratic objective functions at the end of the optimization process. The way the algorith successively approximates the inverse Hessian will be presented below.

The DFP procedure has been developed for <u>minimization</u> of non linear functions. Since we are interested in the <u>maximization</u> of the loglikelihood function, we have

implemented the method to minimize the negative of the log-likelihood function. The method goes through the following steps.

1.) Start with any symmetric positive definite matrix S^o (In our case, the identity matrix), and any point $\underline{\theta}^o$. Set k=0

2.) Set $\underline{d}^k = -S^k \underline{g}^k$
where

$\underline{g}^k$ is the gradient of the negative loglikelihood function with respect to the parameters at iteration k.

S^k is the approximation to the inverse Hessian.

$\underline{d}^k$ is the direction of search.

3.) Minimize $-\xi(\underline{\theta}^k + \alpha^k \underline{d}^k)$ with respect to α^k, to obtain a new value of the parameters, $\underline{\theta}^{k+1}$. The corresponding gradient, $\underline{g}^{k+1}$ is calculated at this point. The new set of parameters is calculated at the optimum α^k, as

$$\underline{\theta}^{k+1} = \underline{\theta}^k + \underline{p}^k$$

where

$$\underline{p}^k = \alpha^k \underline{d}^k$$

4.) Compute the difference between the gradients at iterations k+1 and k, $\underline{q}^k$, which is used to approximate the inverse Hessian:

$$\underline{q}^k = \underline{g}^{k+1} - \underline{g}^k$$

Improve the approximation to the inverse Hessian by means of

$$S^{k+1} = S^k + \frac{p^k p^{kT}}{p^k q^k} - \frac{S^k q^k q^{kT} S^k}{q^{kT} S^k q^k}$$

Update k and go to step 2.

In our work, two additional issues were considered in implementing DFP. First, the approximation to the inverse Hessian S^k, computed by DFP is <u>exact</u> for quadratic objective functions. Since the loglikelihood is a non quadratic function, the approximate inverse Hessian at a point, after several iterations, may not resemble the real inverse Hessian at that point. Second, a linear search in a multi-parameter space may lead some of the parameters to unfeasible regions thus posing serious computational problems.

The solution to the first problem is to re-initialize the approximation to the inverse Hessian after a number of

linear searches, by setting the Hessian equal to the identity matrix. The solution to the second problem will be presented later in this section.

3.6.2 Gradient Calculation.

The analytical calculation of the gradient of the log-likelihood function with respect to the parameters of a linear problem is an easy task and represents a relatively small cost of computation. However, the analytical gradient of a highly non linear system, as ours, represents an extremely high computational burden. Due to this, we chose to calculate the gradient by the method of finite differences, which has been shown to perform satisfactorily in many maximum likelihood estimation cases. (Peterson, 1975).

3.6.3 Linear Search Optimization.

Once a direction of search $\underline{d}^k$ has been chosen at step 2.) of DFP, a one dimensional (along the given direction) optimization problem has to be solved. This problem is expressed as:

$$\underset{\alpha^k}{\text{Min}} \; -\xi(\underline{\theta}^k + \alpha^k \underline{d}^k) \tag{3.21}$$

In which ξ is the loglikelihood function, the superscript k refers to the kth iteration, and $\underline{\theta}^k$ is the vector of parameters at the kth iteration.

In general, the optimum point along this direction, α^{k*}, can not be found exactly, and a compromise must be made between accuracy of the solution and cost of computation. The procedure implemented here is such that the loglikelihood function is calculated successively for increasing values of α^k: $\alpha^k_1, \alpha^k_2, \ldots, \alpha^k_j$. These values are chosen such that $\alpha^k_1 = \sigma_1$, $\alpha^k_2 - \alpha^k_1 = \sigma_2, \ldots \alpha^k_{i+1} - \alpha^k_i = \sigma_{i+1}$, in which σ_i are a series of step sizes that follow the recursion:

$$\sigma_{i+1} = 2 \cdot \sigma_i \tag{3.22}$$

Figure 3.1 represents a hypothetical two parameter case of the linear search problem and its associated variables. The origin of the axis has been set at the value the parameters had at iteration k.

The initial value, σ_1 is chosen as a solution to

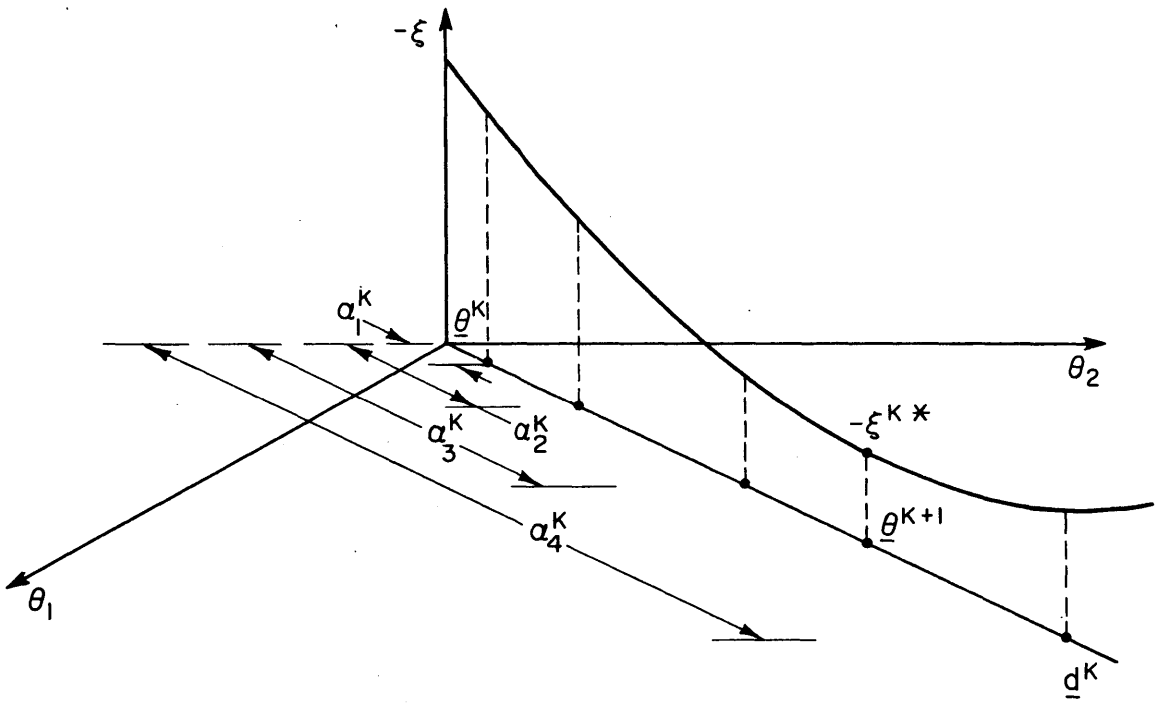


Figure 3.1 Definition Sketch for the Linear Search Optimization at Iteration k.

$$\text{Max } \sigma_1$$

Subject to

$$\sigma_1 d_m^k < \varepsilon \cdot \theta^k_m, \text{ for all } m$$

In which m is a subscript that indicates each of the parameters, and ε is a value set by the user (ε=0.1 in our case). This is done in order to keep the most sensitive parameters (those with a high d_m^k) from changing abruptly at the beginning of the linear search. The geometric growth of the series of σ_i, given by Equation (3.22) guarantees a rapidly increasing step size, thus decreasing the number of values of the objective function to be calculated until the optimum value along the direction of search has been located between two consecutive α^k_{j-1} and α^k_j. This minimum is located when $-\xi(\alpha^k_j) > -\xi(\alpha^k_{j-1})$. At this point, the last three values of $-\xi$: $(-\xi(\alpha^k_{j-2}), -\xi(\alpha^k_{j-1}), -\xi(\alpha^k_j))$, are used to calculate the minimum of the parabola passing through these points. (Figure 3.2). The point at which this minimum lies shall be called α^{k*}. The loglikelihood at this point $\xi(\alpha^{k*})$ is calculated and a new set of three points will be formed. The new set will include α^{k*}, and one of the ends of the previous set will be excluded. (Point "a" in Figure 3.2). This reduces the size of the interval in which the minimum point of the linear search lies. The criterion to indicate which point will be replaced is very clearly described by Luenberger (1973, pp 142-143) and guarantees that the optimum value of the objective function lies somewhere within the interval. A new parabolic interpolation is performed and the process of selecting new sets of three points and interpolating is repeated until the distance between the points surrounding the minimum is less than a given value, (10% of the step size, in our case), or until the number of quadratic interpolations exceeds a maximum (5, in our case).

A test to determine global optimality, (described later in this chapter) is performed at this stage. Should this point fail to pass the test, a new direction of search (step 2. of DFP) is calculated and the procedure is repeated.

Sometimes, in places near the global optimum, the value of the loglikelihood calculated at the first step in a linear search along the direction $\underline{d}^k$ is smaller than the value of the function at the beginning of that search. Should this be the case, a new search is performed in the direction of the gradient $\underline{g}^k$ and a new value of the loglikelihood function is calculated. If the condition persists, like in the case of

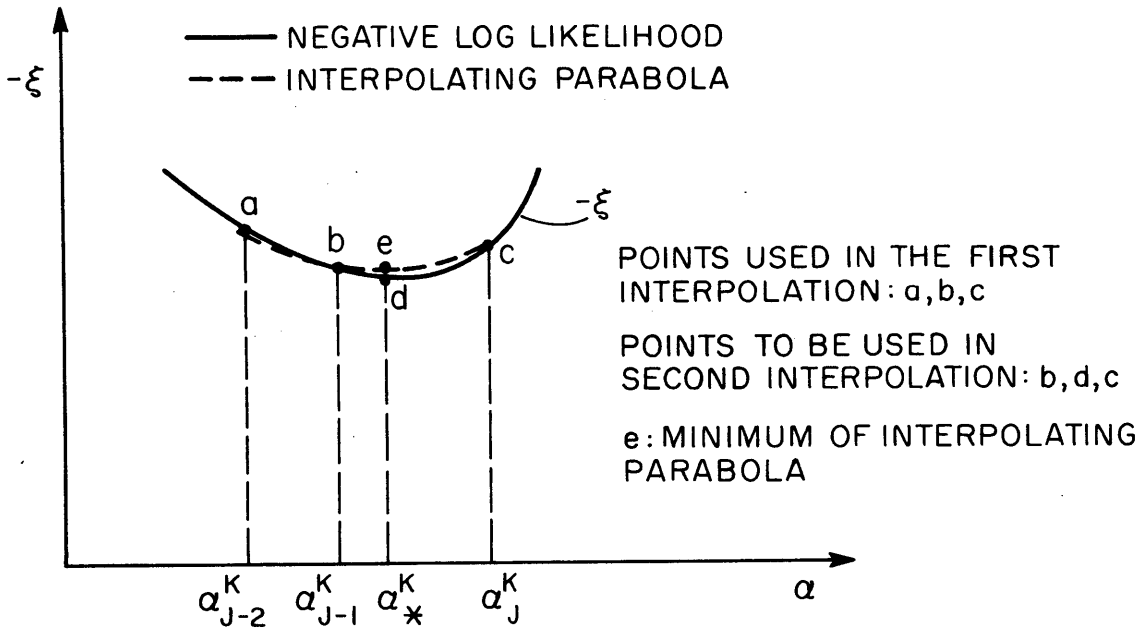


Figure 3.2 Definition Sketch for the Parabolic Interpolation Scheme.

point "b" in Figure 3.3 an additional value for the likelihood function is calculated at a third point located 0.1*step from the starting point. If this last point fails to improve the value of the loglikelihood, (Point "c" in Figure 3.3) it is assumed that global convergence has been achieved and the value of the parameters at the beginning of the iteration will be the maximum likelihood estimates for these parameters. If the third point improves the value of the loglikelihood, (Point "d" in Figure 3.3), the quadratic interpolation procedure mentioned above is performed, and the standard global convergency test, (described later on) will be applied.

3.6.4 Handling of Parameters' Infeasibility.

There are regions of the parameter's space, usually far from the global optimum, in which the direction of search may force some of the parameters to take physically infeasible values. For example, the coefficient of a linear reservoir may become negative which will make the dynamic system unstable. The numerical problems created by this infeasibility preclude the calculation of the loglikelihood function, which, no doubt, would be extremely small, thus making the boundary an obvious constrained optimum in the linear search. Figure 3.4 illustrates this situation.

A basic assumption of the maximum likelihood method is that there is a set of parameters for which the stochastic model corresponds exactly to the measurements, in a statistical sense. That is, there are no structural errors. Under these ideal conditions, the global optimum lies in a region in which all parameters are feasible. Assuming the loglikelihood is unimodal, there are starting points from which the convergency path to the optimum lies always in the feasible region. The immediate solution to the problem above would be to try different starting points, until global convergence is achieved by trial and error.

There are three problems connected with the trial and error solution. First, although it may be easy to find a good starting point in a case with few parameters to be estimated, this may not be the case in a high dimensional problem. Second, it requires the interaction between the user and the computer. Third, and a very important one in real life and especially in rainfall-runoff models, structural errors are unavoidable. Therefore, we must expect to find cases in which the optimum lies outside the feasible region. This would force the user to try a very large number of parameters while hopelessly and unknowingly trying to achieve convergency in a feasible region. Since one of the aims of our research is to develop a user independent parameter estimation procedure, the trial and error solution is not applicable here and a different one has been implemented. The adopted solution takes the following steps:

Step 1. When a parameter reaches the boundary, the value of the loglikelihood at a feasible point very close to

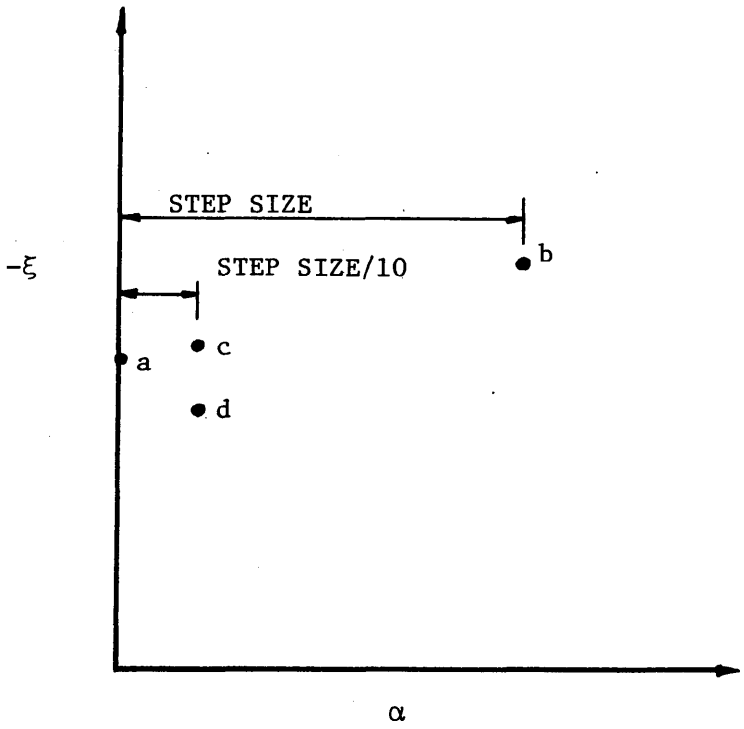
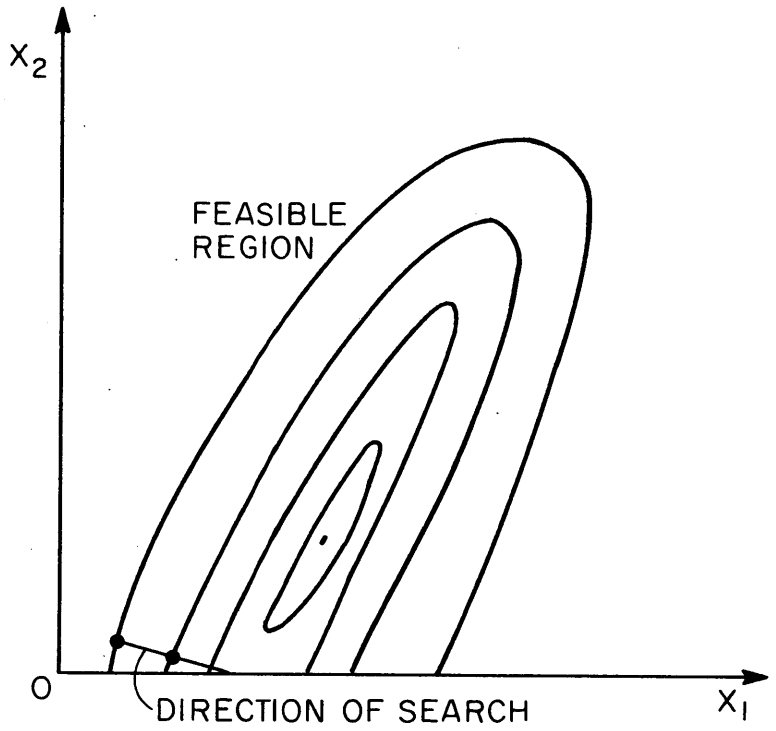
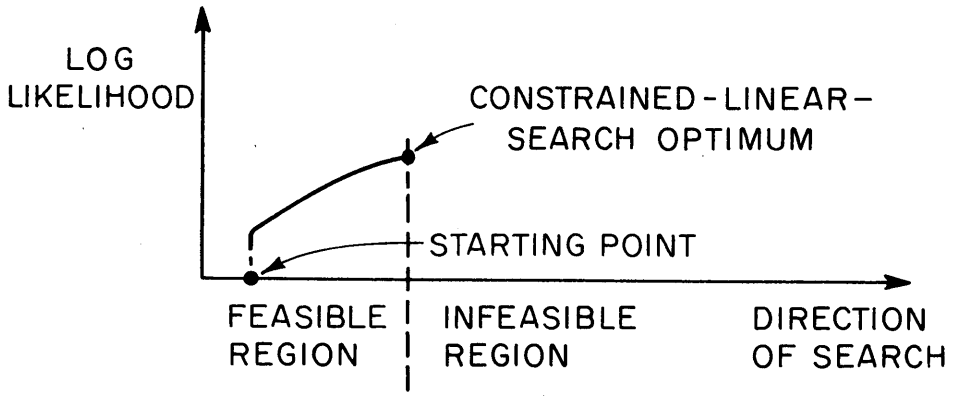


Figure 3.3 Values of the Loglikelihood in the Direction of the Gradient.



a) ISO CONTOURS OF THE LOG LIKELIHOOD FUNCTION



b) VARIATION OF THE LOG LIKELIHOOD FUNCTION ALONG THE DIRECTION OF SEARCH

Figure 3.4 Constrained Linear-Search Optimum Due to Bounds in the Parameter Values.

the boundary is calculated. This point is such that the binding parameter takes the (arbitrary) value of 10^{-10}, which is several orders of magnitude smaller than the normal range of the parameters of the model, but still within the feasible region. If the loglikelihood is smaller than the value of the loglikelihood computed in the previous step the maximum in the linear search lies somewhere in the feasible region and the standard quadratic interpolation procedure mentioned before is performed. If this is not the case, the procedure continues with step 2.

Step 2. The gradient at this point near the boundary is computed and it is determined whether or not a search in the direction of the gradient will move the binding parameter back into the feasible region. Should this be the case, DFP is reinitialized, starting in the direction of the gradient. If this is not the case, a DFP search is initiated in a subspace excluding the binding parameter. Within this subspace, the feasibility check is performed in every line search and the dimension of the subspace is then further decreased should new parameters tend to move into the non feasible regions. Given the high cost of calculating the gradient, its computation at the binding point may be skipped and the DFP will proceed with the search in the subspace excluding the binding parameter.

Step 3. At the end of every linear search, the gradient of the loglikelihood function with respect to all parameters is calculated and it is determined which bounded parameters will move into the feasible region, thus increasing the size of the subspace.

3.6.5 Global Optimality Criterion.

In the neighborhood of the optimum point, the loglikelihood function can be approximated by a quadratic surface. It has been shown by (Edwards, 1972) that the loglikelihood function evaluated at two standard deviations of the parameters away from the optimum is exactly two units smaller than the maximum. Therefore, if a linear search fails to improve the loglikelihood function by more than two units it is likely that the parameters are closer than two standard deviations from their maximum likelihood values. Since any improvement in the loglikelihood function thereafter will be very small, the last point in the linear search can be taken as the maximum likelihood estimate of the parameters. This criterion has been succesfully used in the General Purpose System Identifier and Evaluator (GPSIE), developed by (Peterson, 1976). Our own experience with estimating the parameters of the conceptual rainfall-runoff models shows that sometimes the difference between the loglikelihood functions at two consecutive iterations may be less than two units, yet the maximum of the likelihood surface is not near. We adopted a stricter criterion which calls for the global search to be suspended after the

optimization algorithm fails to improve the loglikelihood function by more than 0.01 units. The goodness of this estimate can be measured by several criteria, taken also from the above reference, and briefly described below:

1.) Positive definiteness of the information matrix, F, which is aproximated by the negative Hessian. This is a necessary condition for optimality.

2.) The elements in the diagonal of the parameters covariance matrix, approximated by the inverse of F, give us the lower bounds on the variances of the parameter's estimates, a measure the quality of these estimates.

3.) The value of the loglikelihood function at two standard deviations of the parameters away from the optimum point should be two units less than the value at the optimum.

4.) The expected value of the sum of squares of the normalized residuals, equation (3.7), equals the number of degrees of freedom, i. e., the number of measurements minus the number of parameters to be estimated from these data.

5.) The lag-0 correlation coefficient of the normalized residuals should be close to its expected value, 1.0, and the lag-j ($j \neq 0$) correlation coefficient should be close to zero. This a statement of residuals whiteness. The standard deviations of these correlation coefficients are also calculated, and the normalized correlation coefficients, defined as the estimated correlation coefficients divided by their respective standard deviations, are printed. These give an idea of the deviation of the correlation coefficients away from their expected value, in units of standard deviations. The correlation coefficients, $\rho(i)$, are estimated in our scalar measurements case by $\hat{\rho}(i)$, which is computed by

$$\hat{\rho}(j) = \sum_{t=1}^{T-j} [\nu(t) \cdot \nu(t-j)]/(T-j), j=0,4$$

where T is the total number of measurements and j is the lag

It can be shown that the standard deviation of the estimates of the correlation coefficients, σ_ρ is (Peterson, 1975):

$$\sigma_\rho(i) = \frac{\rho(i) - E[\rho(i)]}{\frac{1 + E[\rho(i)]}{T} - \frac{i-1}{T^2}}$$

$$\text{In which } E[\rho(i)] = \begin{cases} 1 \text{ for } i = 0 \\ 0 \text{ for } i > 0 \end{cases}$$

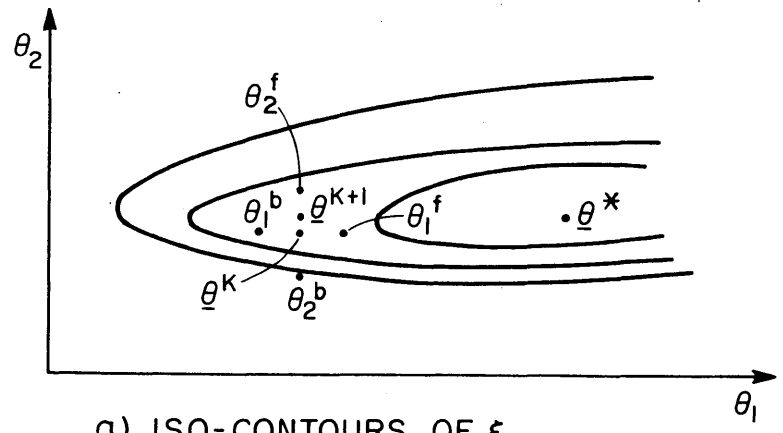
The Durbin and Watson Statistics, (Langseth and Bras, 1979), is also calculated to test the whiteness of residuals.

A problem associated with this convergency criterion was detected in the last stage of this work and appears when the loglikelihood is very sensitive to some of the parameters, and relatively insensitive to the others. This problem is illustrated, for a two parameter case, in Figure 3.5. In that figure, the superscripts f and b denote, respectively, the forward and backward perturbation of the parameters. These perturbations are used to calculate the gradient of the loglikelihood function with respect to the parameters by a central differences scheme. In Figure (3.5 -a), $\underline{\theta}^k$ is the value of the parameter vector at iteration k.

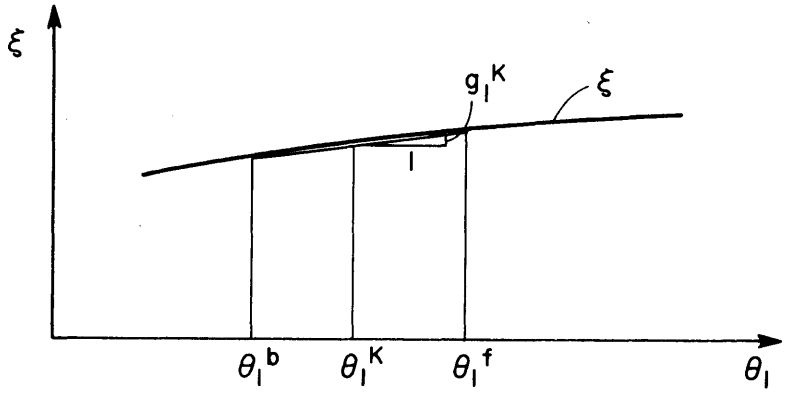
For the case we are presenting, the gradients computed by finite differences will be such that

$$|g^k{}_2| >> |g^k{}_1|$$

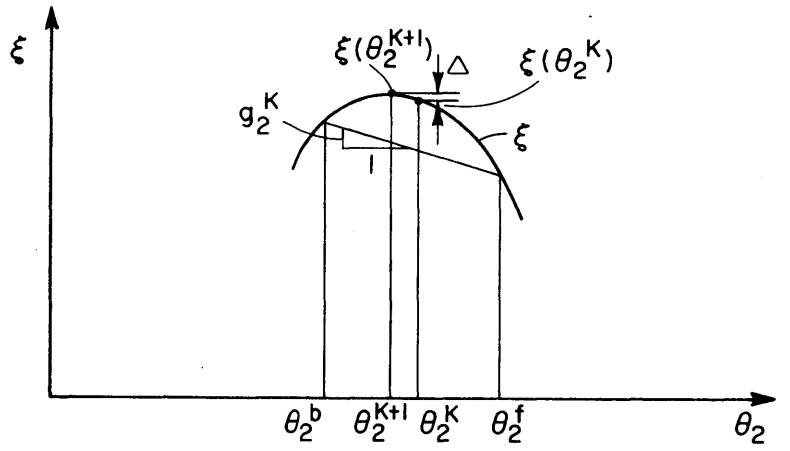
This causes the direction of linear search, $\underline{d}^k$, to be mostly in the direction of θ_2. We found that in several cases the difference between the values of the loglikelihood at the end of the linear search, $\xi(\underline{\theta}^{k+1})$, and that at the beginning, $\xi(\underline{\theta}^k)$, ($\Delta$ in Figure 3.5-c), was smaller than the global convergency criterion presented above. This caused the non linear optimization process to stop at points far from the global optimum. The process was re-started at the point of convergency after stopping the optimization of those parameters for which the values of the forward and backward differences did not improve the value of the loglikelihood. (Figure 3.5-c). An automatic solution to this problem was implemented by means of which only those parameters which at the end of a linear search improve the value of the loglikelihood in either the forward or the backward direction are included in the following linear search optimization. Those parameters that do not improve the value of the loglikelihood in either direction are left unchanged regardless of the magnitude of their gradient. Since this solution converts the DFP method into a blend of DFP and pattern search, we do not believe it to be an optimal solution. We will be addressing this topic again in the last chapter, Conclusions and Recomendations.



a) ISO-CONTOURS OF ξ



b) ξ IN THE DIRECTION OF θ_1



c) ξ IN THE DIRECTION OF θ_2

Figure 3.5 Calculation of the Gradient by a Central Finite Differences Scheme.

CHAPTER 4

THE CONCEPTUAL RAINFALL-RUNOFF MODELS

4.1 Introduction.

This chapter describes the stochastic conceptual rainfall-runoff models employed in this work. The first model, which will be presented in section 4.2, is the model for which the parameter estimation procedure was developed. The second model, described in section 4.3, was developed as a simpler version of the previous one with the purpose of helping in the implementation of the parameter estimation procedures at a reasonable cost.

4.2 Continuous Form of the NWSRFS Model

The stochastic conceptual rainfall-runoff model selected for application of the parameter estimation procedure developed in this work is esentially the model presented by Kitanidis and Bras (1980). These authors applied a linear filtering technique (Kalman Filter) to a stochastic version of the deterministic conceptual Sacramento model in use by the National Weather Service. This agency has developed a large system of programs for river flows forecasting. This system is known by the acronym NWSRFS, which stands for National Weather Service River Forecasting System, a part of which is the Sacramento model. In this work the term "NWSRFS model" will be used instead of the term "Sacramento model". The NWSRFS model was originally published as a modification of the Stanford Watershed Model IV, based on the work by Crawford and Linsley (1966). The elements of the NWSRFS model are schematized in Figure 4.1. The model by Kitanidis and Bras (1980) has been sligthly modified from its original form, as reported by Georgakakos et al. (1980). The single linear reservoir channel router used by Kitanidis and Bras (1980) has been replaced by the discrete-time linear routing scheme developed by TASC (1980).

Since the NWSRFS model is well described elsewhere (NWSRFS Hydro-31, 1976), it will not be extensively discussed here. We will present only the continuous-time equations which are the basis for the state-space formulation required by the estimation techniques presented in Chapter 3. These equations correspond to Equations (21) through (26), p. 1029 in the Kitanidis and Bras paper. Equations (27) and (28) in the same reference are the continuous time state space equations for the linear reservoir model of the channel, which, as mentioned above, will be replaced by a different routing scheme.

4.2.1 Continuous-Time Equations for the NWSRFS Model (Soil Moisture Phase).

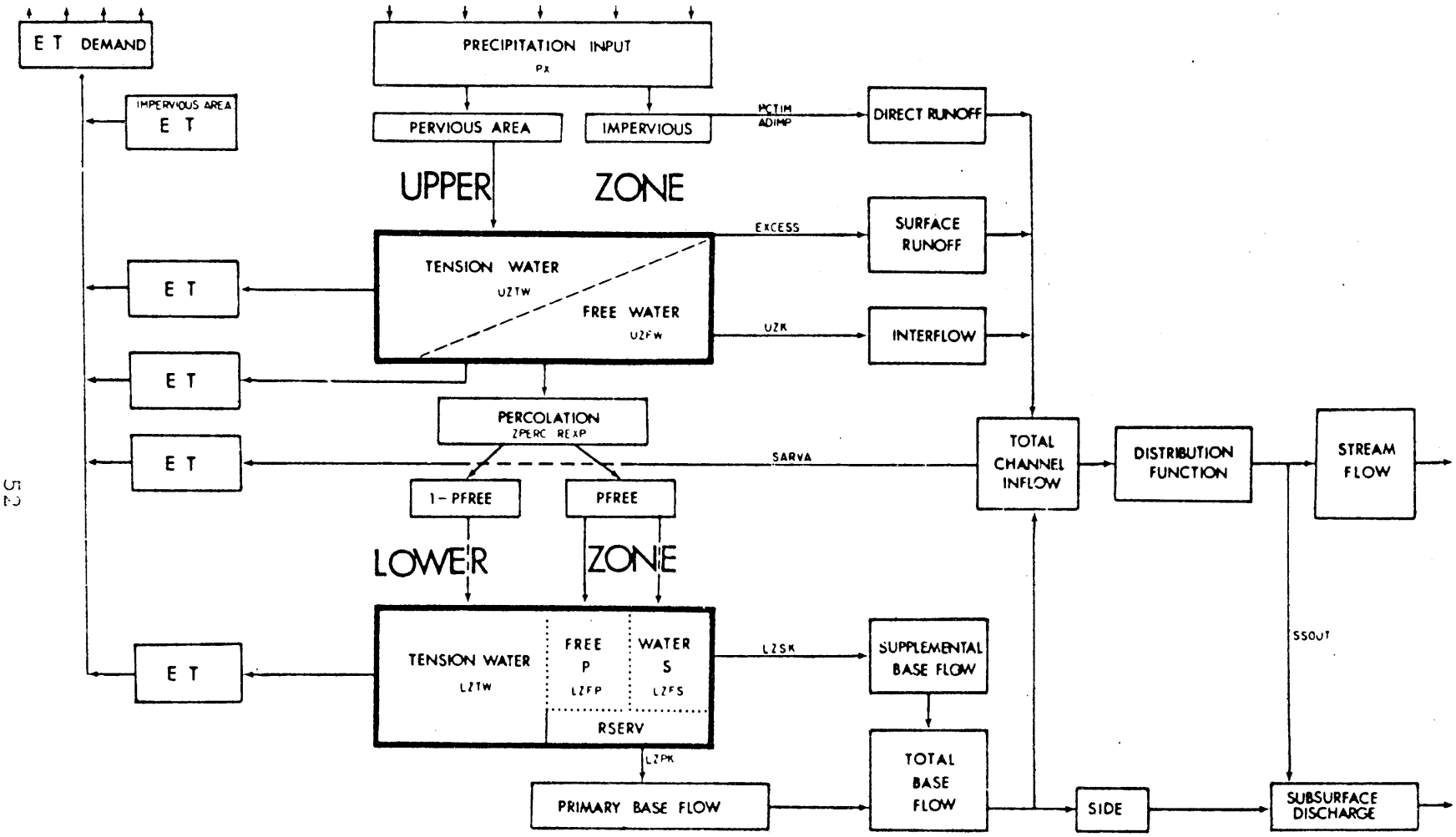


Figure 4.1 Schematic Diagram of the National Weather Service River Forecasting Model (Sacramento Model).

Before presenting the equations of the soil part of the model, it is convenient to introduce the reader to the notation to be used. This notation is presented in Table 4.1, and is adapted from the notation proposed by TASC (1980).

In order to make the notation compact and readable, we shall define a series of intermediate variables and functions that will be used in the state space formulation to follow. The threshold functions are defined by

$$g(x,x^{o},u) = \begin{cases} -u, & x=x^{o},\ u>0 \\ 0, & \text{otherwise} \end{cases}$$

Two constants that will be used are:

$$C_1 = d_1' x_4^{o} + d_1'' x_5^{o}$$

$$C_2 = x_4^{o}/(x_4^{o}+x_5^{o})$$

Other common functions are:

$$y = 1-[(x_3+x_4+x_5)/(x_3^{o}+x_4^{o}+x_5^{o})]$$

$$F = C_2[2(1-x_4/x_4^{o})]/[2-x_4/x_4^{o}-x_5/x_5^{o}]$$

$$\tau_3 = (1-P_f)[1-(x_3/x_3^{o})^{m}]$$

$$P_c = C_1(1+\gamma y^{\alpha})x_2/x_2^{o}$$

The governing simultaneous differential equations for the various storage elements of the NWSRFS model follow. The upper zone is defined by:

$$dx_1/dt = g(x_1,x_1^{o},u_1-u_2\cdot(x_1/x_1^{o})) + u_1 - u_2\cdot x_1/x_1^{o} \qquad (4.1)$$

$$dx_2/dt = -g(x_1,x_1^{o},u_1-u_2\cdot(x_1/x_1^{o})) - d_u x_2 - P_c$$

$$+g(x_2,x_2^{o},h_f(u_\ell)u_\ell) \qquad (4.2)$$

SYMBOL	NWS ACRONYM	DESCRIPTION
x_1	UZTWC	Upper zone tension water content (mm)
x_2	UZFWC	Upper zone free water content (mm)
x_3	LZTWC	Lower zone tension water content (mm)
x_4	LZFPC	Lower zone free primary water content (mm)
x_5	LZFSC	Lower zone free supplementary water content (mm)
x_6	ADIMC	Additional impervious storage (mm)
$x_7, x_8 \cdots$		Unit Hydrograph states (mm)
	Inputs	
u_1	PX6/DT	Instantaneous moisture input (mm/h)
u_2	EDMND/DT	Instantaneous evapotranspiration demand (mm/h)
	Parameters	
x_1^o	UZTWM	Upper zone tension water capacity (mm)
x_2^o	UZFWM	Upper zone free water capacity (mm)
x_3^o	LZTWM	Lower zone tension water capacity (mm)
x_4^o	LZFPM	Lower zone free primary capacity (mm)
x_5^o	LZFSM	Lower zone free supplementary capacity (mm)
d_u	-ℓn(1-UZK)/24	Upper zone instantaneous drainage coeff. (h^{-1})
d_1'	-ℓn(1-LZPK)/24	Lower zone primary instantaneous drainage coeff. (h^{-1})
d_1''	-ℓn(1-LZSK)/24	Lower zone supplementary instantaneous drainage coefficient (h^{-1})
γ	ZPERC	Parameter in percolation function
α	REXP	exponent in percolation function
P_f	PFREE	Fraction of percolated water assigned to the free water aquifers
μ	SIDE	Fraction of base flow not appearing in river flow
a_1	ADIMP	Fraction of basin that becomes impervious when tension water requirements are met
a_2	PCTIM	Fraction of basin permanently impervious
r	RSERV	Fraction of the lower zone free water unavailable to supply lower zone tension requirements
	Other variables and functions	
g		Threshold functions
τ_3		Approximation of the threshold function in x_3
P_c		Percolation rate
F		Fraction of percolating water to the free aquifers that goes to the primary aquifer
y		Defficiency ratio
T_{ci}		Total channel inflow

TABLE 4.1

NOTATION

In Equation (4.2) u_ℓ is the net inflow rate equal to the sum of all terms in the right of Equation (4.2) excluding the function $g(\ ,\ ,\ h_f(u_\ell)u_\ell)$, where

$$h_f(u_\ell) = 1 \quad u_\ell > 0$$
$$h_f(u_\ell) = 0 \quad u_\ell < 0$$

The lower zone obeys,

$$dx_3/dt = P_c\tau_3 - u_2(1-x_1/x_1^o)\cdot[x_3/(x_1^o+x_3^o)] \tag{4.3}$$

$$dx_4/dt = -d_1'x_4 + P_f(1-\tau_3)\cdot F \tag{4.4}$$

$$dx_5/dt = -d_1''x_5 + P_f(1-\tau_3)\cdot(1-F) \tag{4.5}$$

$$dx_6/dt = u_1 + [(x_6-x_1^o)/x_3^o]^2 g[x_1, x_1^o, u_1 - u_2(x_1/x_1^o)] \tag{4.6}$$

The total inflow to be routed in the channel is:

$$T_{ci} = u_1\cdot a_2 - [(x_6-x_1^o)/x_3^o]^2\cdot g(x_1, x_1^o, u_1 - u_2x_2/x_2^o]$$

$$a_1 - g(x_2, x_2^o, h_f(u_\ell)u_\ell)\cdot(1-a_1-a_2) -$$

$$C\cdot g(x_2, x_2^o, h_f(u_\ell)u_\ell)\cdot[1-(x_6-x_1^o)/x_3^o]^2\cdot a_1 +$$

$$(1-a_1-a_2)[d_u x_2 + (d_1'x_4 + d_1''x_5)/(1+\mu) \tag{4.7}$$

$$C = \begin{cases} 1 & \text{if } x_6 < x_1^o + x_3^o \\ 0 & \text{if } x_6 > x_1^o + x_3^o \end{cases}$$

4.2.2 Discrete-Time Reduced Order Linear Channel Routing Scheme.

The Analytic Sciences Corporation (TASC, 1980) suggested a discrete-time linear routing scheme with the form:

$$\underline{x}_u(k) = \Phi_u\underline{x}_u(k-1) + G_u \cdot T_{ci}(k-1) \tag{4.8}$$

and a measuring equation of the form

$$z(k) = \underline{H}(k)\ \underline{x}_u(k) + v(k) \qquad (4.9)$$

in which $\underline{x}_u(k)$ is the vector containing the unitgraph states at time step k, Φ_u and G_u are constant coefficient matrices determined from the unit hydrograph of the basin under consideration. The procedure to identify these matrices from an available unit hydrograph is described by TASC (1980).

The continuous time soil phase equations (4.1) to (4.6) are linked to the discrete time unit hydrograph channel (Equation 4.8) by means of the total channel inflow, T_{ci}, given by Equation (4.7). For the prediction step the soil phase equations are integerated through a complete time step of duration DT=0.25 days. The acumulated T_{ci} during that time interval is then used in Equation (4.8) to compute the predicted discharge.

The filtering step requires the estimation of the covariance matrix of the complete system, soil phase and channel, which, as we mentioned above, are continuous and discrete time models, respectively. Since the filter implemented by Kitanidis and Bras is a discrete time linear filter, (the fact that the measurements are taken at discrete time intervals requires the use of a discrete time filter), we have to find a discrete time linear approximation to the complete soil phase and channel systems. That linear approximation was developed by Georgakakos et al. (1980) and is presented here. To put this presentation into perspective, we first introduce the linear approximation of the complete system, and then we will explain how the different components of that approximation are computed.

The discrete time linear approximation of the system will take the form

$$\underline{x}_c(k) = \Phi_c\ \underline{x}_c(k-1) + G_c \cdot u(k-1) + F_o \qquad (4.11)$$

where:

$\underline{x}_c(k) = \begin{bmatrix} \underline{x} \\ \underline{x}_u \end{bmatrix}$ Complete-system state vector at k=T•DT.

Φ_c: Linearized complete system transition matrix.

G_c, F_o: Matrices obtained from the linearization and integration schemes, to be introduced later.

We are going to proceed to the derivation of Equation (4.11) in two steps. In the first step we are going to explain how the linearization and integration of the linearized equations of the soil phase model is performed. In the second step we will concentrate on the linearization of the channel model.

For the linearization and integration of the linearized equations of the soil phase model, we follow the schemes of Kitanidis and Bras (1980). Let the system of continuous time non linear differential equations of the soil phase be represented in a generic form as:

$$dx_i/dt = f_i(\underline{x},u,t) + w_i$$

We want the system to be linearized in the form:

$$d\underline{x}/dt = A\ \underline{x}(t) + B\ u(t) + \underline{w}(t) \qquad (4.12)$$

where

$$A_{ij} = \partial f_i/\partial x_j|_{x=x_L},\ j=1,6,\ i=1,6$$

$$B_i = \partial f_i/\partial u_1|_{x=x_L},\ i=1,6$$

Notice that only the derivative with respect to u_1 is required, since the evapotranspiration demand u_2 was assumed to be deterministic. Notice also that the matrices A and B are evaluated at a nominal value $\underline{x}_L$, which in Kitanidis and Bras' application corresponds to

$$\underline{x}_L = (\underline{x}(k+1) + \underline{x}(k))/2.$$

The integration of the system of equations (4.12) is done with the integration scheme proposed by Kitanidis and Bras (Equation 54, p. 1032, 1980), to arrive at

$$\underline{x}(k+1) = \Phi_s\ \underline{x}(k) + G_s\ u(k) + \underline{w}(k) \qquad (4.13)$$

where

$$\Phi_s(k)=(I-A\cdot DT/2+A^2\cdot DT^2/12)^{-1}(I+A\cdot DT/2+A^2\cdot DT^2/12)$$

$$G_s(k) = (I-A\cdot DT/2+A^2\cdot DT^2/12)^{-1}\cdot DT\cdot B$$

and $\underline{w}(k)$ is now a discrete time White Gaussian Noise.

The second step is the linearization of the channel model equation with respect to the soil phase states and the input, since Equation (4.8) is already linear on the unit hydrograph states. We can express the total channel inflow (Equation 4.7) in compact form as

$$T_{ci} = f_c(\underline{x}, u_1, t) \qquad (4.14)$$

Linearizing Equation (4.14) with respect to $\underline{x}$ and u_1 we obtain

$$T_{ci} \simeq f_o + \underline{\beta}^T \cdot \underline{x}(k) + \beta_u \cdot u_1(k) \qquad (4.15)$$

where

$$f_o = f_c(\underline{x}_L, u_1(k), k) - \underline{\beta}^T \cdot \underline{x}_L - \beta_u \cdot u(k)$$

$$\underline{\beta}^T = \partial f_c / \partial \underline{x} \big|_{x=x_L}$$

$$\beta_u = \partial f_c / \partial u_1 \big|_{u_1=u_1(k)}$$

By replacing Equation (4.15) into Equation (4.8) we arrive at

$$\underline{x}_u(k) = \Phi_u \cdot \underline{x}_u(k-1) + G_u(\underline{\beta}^T \cdot x(k-1) + \beta_u \cdot u_1(k-1) + f_o) \qquad (4.16)$$

Equations (4.13) and (4.16) can now be expressed in the form of Equation (4.11)

$$\begin{bmatrix} \underline{x} \\ \underline{x}_u \end{bmatrix}_k = \begin{bmatrix} \Phi_s & \underline{0} \\ G_u \, \underline{\beta}^T & \Phi_u \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{x}_u \end{bmatrix}_{k-1} + \begin{bmatrix} G_s \\ G_u \, \beta_u \end{bmatrix} u(\kappa-1) + \begin{bmatrix} W(k) \\ G_u \, f_o \end{bmatrix} .$$

4.3 The Simplified Rainfall-Runoff Model.

A simplified version of the NWSRFS model was developed during the first stage of the research. The motivation was mostly cost reduction. As in the original model, the state variables of the simplified model will be the water content in the different storage elements. We will use frequently the term "storage element" or simply "element" to refer to the state variables. In this section, we will explain the simple model, following Figure 4.2 as a guide. The continuous time equations, the description of the functional relationships among the state variables, and the parameters of the model will be presented in section 4.3.3. Before going into the description of the model itself, it is important to mention the differences between the simplified model and the Kitanidis and Bras model.

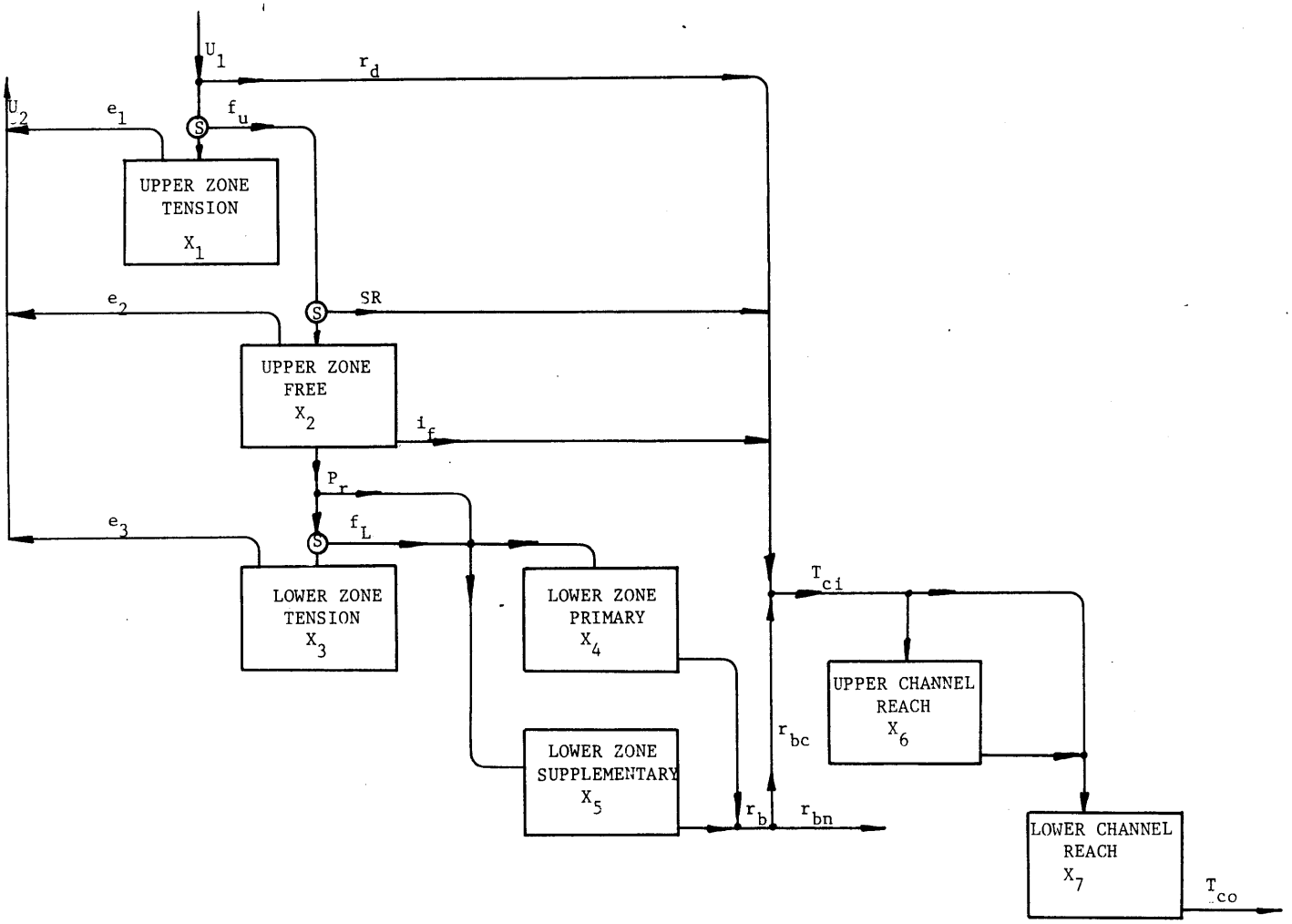


Figure 4.2 Schematic Diagram of the Simplified Rainfall-Runoff Model.

4.3.1 Differences Between the Simplified Model and Kitanidis and Bras' Model.

The first change was to modify the "S" curves which were proposed by Kitanidis and Bras (1980, p. 1031), as an option to the threshold functions. Two changes were made on these curves: 1) the function itself was changed from an arc-tangent to a hyperbolic tangent, and 2) the curves will now be included in the model so that they act before the water goes into the corresponding element. In this form, the contents of the storage elements are not constrained to be below a maximum value. In addition, the values of the "S" curves are used to measure the level of " saturation" in each element. This last point is a significant difference in the notations of both models. For the NWSRFS model saturation means the relative water content in a storage element, x_i/x_i^o, in which the subscript "i" refers to any of the storage elements.

The second change was to reduce the number of state variables in the soil moisture accounting part of the model, by eliminating the additional impervious content, ADIMC, and its associated parameter, a_1.

The third change was to model the channel by two non-linear reservoirs, following the model presented by Georgakakos and Bras (1980).

The fourth and last change was to replace the function that divides the water flowing to the lower zone free primary and supplementary reservoirs.

4.3.2 Model Description

There are two inputs to the model, both of them assumed to be deterministic inputs: the precipitation rate, u_1, and the evapotranspiration demand, u_2. u_1 is divided into the direct runoff, r_d, and the fraction of precipitation which becomes infiltrated, i_r. r_d goes directly to the channel, while i_r is divided by the "S_1" curve in a portion that drains into the upper zone tension water element, x_1, and a portion that overflows from the upper zone tension element, f_u. If the content of this element is very low, the value of the "S_1" curve will be close to zero, and most of the infiltrated water will go into x_1. On the other hand, if x_1 is high, S_1 will approach the saturation value of 1.0, and most of the water will overflow and go into the upper zone free water element. The only direct output from the tension water is the evaporation term, e_1.

The overflow from the upper zone tension is routed

through the second "S" curve, S_2, and divided between surface runoff, SR, and the part that will drain into the upper zone free water, x_2. In a similar manner as in the previous case, if x_2 is low, $S_2 \simeq 0.0$, and SR≃0.0. Three outputs are possible from the upper zone free water element: 1) the evaporation output, e_2, which tries to satisfy demand not fulfilled by x_1; 2) the interflow, i_f, which is modeled as a linear reservoir, following the NWSRFS model, and drains to the channel; and 3) the percolation function, which, being considered the "heart" of the NWSRFS model was modeled in exactly the same way as in the original model.

The percolating water, P_r, is divided into the fraction that drains directly into the lower zone free water elements and the part acted on by the "S" curve of the lower zone tension water element, S_3. When S_3 approaches saturation, the overflow from the lower zone tension water element, f_1, increases and joins the free-percolating part. The remaining drains into the tension water element, x_3, which has only one output, the evaporation term, e_3, that tries to satisfy demand not fulfilled by the two previous elements.

The water percolating to the lower zone free water elements, accounting for the base flow of the basin, is now divided between the primary and supplementary reservoirs, x_4 and x_5, respectively. The criterion for dividing the water between both elements was proposed by Georgakakos et al. (1980) and is reproduced later in this chapter. In the original NWSRFS model these two reservoirs do not have threshold functions. The maximum amount of water they hold is indirectly controlled by the saturation ratio of the reservoirs' content through the percolation function. Since in the simplified model the saturation of the elements is measured by the value of the "S" curve, we have used "S" curves in each of these reservoirs. Both reservoirs are linear, and the output from them forms the total base flow, r_b. This is divided into the base flow which is a component of the total channel inflow, r_{bc}, and the part that does not go to the channel, r_{bn}. The total channel inflow, T_{ci} is divided into two components, each draining into one of the two reaches of the non-linear channel, x_6 and x_7. The output from the upper non-linear channel goes to the second reach of the channel, and the output from this becomes the total channel outflow, T_{co}. No evapotranspiration from riparian vegetation is considered and no losses from the channel are taken into account.

4.3.3 Governing Simultaneous Differential Equations of the Simplified Model.

We made reference in the previous section to two functions which the simplified model employs and which have not yet been introduced. These are the "S" curves and the function that distributes water between the lower zone free primary and supplementary reservoirs (The F function). Before proceeding with the presentation of the simultaneous differential equations that mathematically describe the simplified model, we are going to introduce the "S" curves and the "F" function.

The "S" curves used in the simplified model are defined by Equation (4.17), and are presented in Figure 4.3

$$S_i = 0.5 \cdot [\tanh(w) + 1] \qquad (4.17)$$

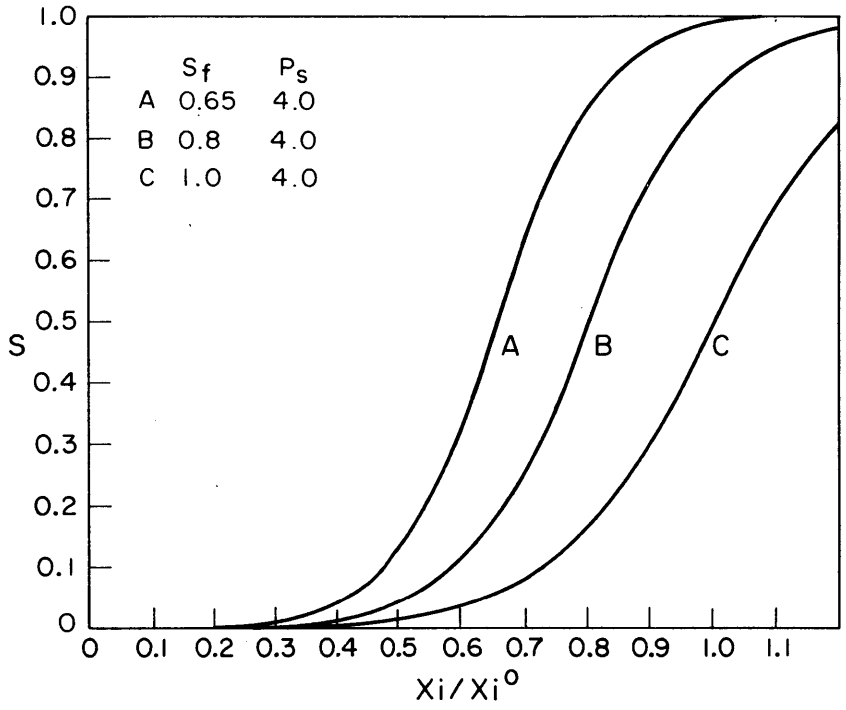
where

$$w = [(x_i/x_i^o)/S_f - 1] \cdot P_s \qquad (4.18)$$

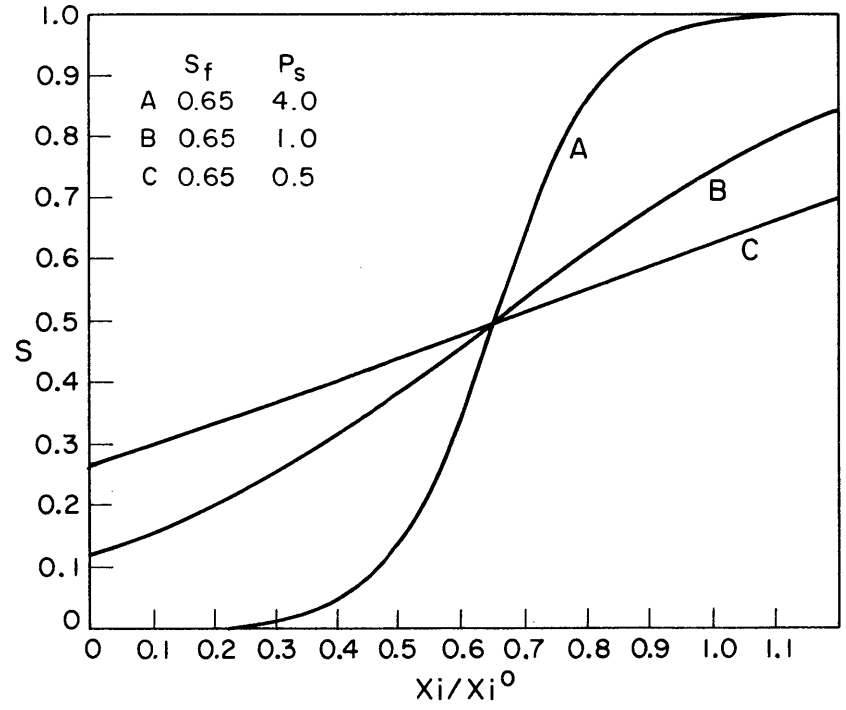
where

- w: Dimensionless water content.
- x_i: State variable "i"
- x_i^o: Parameter of the curve coresponding to state "i"
- S_f: Parameter to shift the midpoint of the curve (set to 0.65)
- P_s: Parameter to alter the slope of the curve (Set to 4.0)

The parameters x_i^o, i=1,5, are to be estimated. Two parameters were included in order to make the S curve to approximate closely the threshold function. The first one, S_f, shifts the midpoint of the curve. For S_f=1, S=0.5 for x_i/x_i^o = 1. With S_f<1, it is possible to get S≈1, for x_i/x_i^o≈1, thus simulating the effect of the original threshold in the NWSRFS model. The parameter P_s was arbitrarily set as 4.0 for all "S" curves. The effect of this parameter is to change the slope of the curve. For P_s=1 the slope of the curve is unity, at x_i/x_i^o=S_f. A value of P_s greater than one, will make the slope steeper. The opposite will hold for a value of P_s smaller than one.



a) FIXED P_s, VARYING S_f



b) FIXED S_f, VARYING P_s

Figure 4.3 "S" Curves for Different values of the Parameters S_f and P_s.

The reason why the original F function in the NWSRFS model was not used stems from the fact that the original function may attain an undetermined zero over zero value in the stochastic model. This condition can not appear in the deterministic model. The original function was given by

$$F(d_4,d_5) = C_2 \cdot d_4/(d_4+d_5) \qquad (4.19)$$

where

$$d_4 = (1 - x_4/x_4^{\circ}) \qquad (4.20)$$

$$d_5 = (1 - x_5/x_5^{\circ}) \qquad (4.21)$$

and C_2 was defined in Section 4.2.1.

In that form, Equations (4.20) and (4.21) represent the deficiency ratios, or relative dryness of the respective elements. However, updating the state variables after the measurements of the discharge have been obtained may cause both d_4 and d_5 to be simultaneously zero, which will cause an indetermination in the function F. To prevent this from happening, a different function was proposed by Georgakakos et al. (1980)

Let $F(S_4,S_5)$ be the function that divides the total percolating water to the lower zone free water elements, such that, at the two extremes, F=1 will direct the water to the primary reservoir, and F=0 will direct the water to the supplementary reservoir. Notice also that the function is now explicitly defined in terms of the saturation values given by the "S" curves of the lower zone reservoirs, instead of the reservoirs' contents themselves.

Define the following policy: for $x_4 = 0.0$, all water coming into the free water reservoirs (free precolation + overflow from tension water) will go to the primary reservoir:

$$F(0,S_5) = 1, \text{ all values of } S_5 \qquad (4.22)$$

When both reservoirs are full, the amount of water sent to the primary reservoir is equal to the water that leaves the reservoir. Under saturated conditions, the maximum percolating water to both reservoirs is set as

$$P_b = d_1' \cdot x_4^{\circ} + d_1'' \cdot x_5^{\circ} \qquad (4.23)$$

in which d_1' and d_1'' are the coefficients of the linear reservoirs. According to this policy, for $S_1=S_2=1$ (both reservoirs are fully saturated),

$$F(1,1)\cdot P_b = d_1'\cdot x_4^{o} \tag{4.24}$$

or

$$F_{11} = F(1,1) = d_1'\cdot x_4^{o}/P_b \tag{4.25}$$

Finally,

$$F(1,0) = 0.0 \tag{4.26}$$

A function that fulfills these conditions is given by Equation (4.27), and is represented in Figure 4.4.

$$F(S_4,S_5) = (F_{11}\cdot S_5-1)\cdot S_4 + 1 \tag{4.27}$$

Given the above definitions of S_i and $F(S_4,S_5)$ it is now possible to give the governing equations of the simplified model.

Upper zone tension water, x_1:

$$dx_1/dt = i_r\cdot(1-S_1) - e_1 \tag{4.28}$$

$$i_r = u_1 - r_d \tag{4.29}$$

$$r_d = u_1\cdot a_2 \tag{4.30}$$

$$e_1 = u_2\cdot S_1 \tag{4.31}$$

where:

u_1: Precipitation rate (Input)

u_2: Evapotranspiration demand (Input)

S_1: "S" curve corresponding to x_1 ($S_1(x_1,x_1^{o})$

i_r: Rate of infiltration

r_d: Direct runoff

a_2: Impervious area as a percent of total area (Parameter)

x_1^{o}: Parameter of S_1

Upper zone free water, x_2:

$$dx_2/dt = (1-S_2)\cdot f_u - i_f - e_2 - P_r \tag{4.32}$$

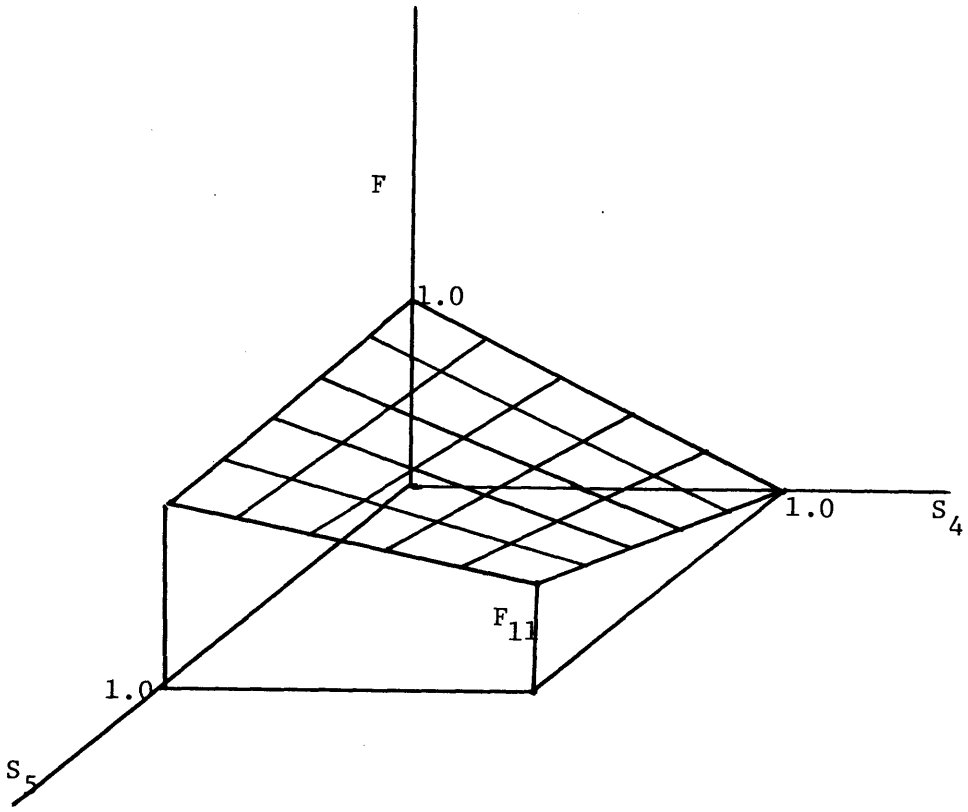


Figure 4.4 Percolating Water Distribution Function Between the Lower Zone Primary and Supplementary Reservoirs.

$$f_u = i_r \cdot S_1 \tag{4.33}$$

$$i_f = d_u \cdot x_2 \tag{4.34}$$

$$e_2 = u_2 \cdot (1-S_1) \cdot S_2 \tag{4.35}$$

$$P_r = P_b \cdot S_2(1 + \gamma \cdot d_r^{\alpha}) \tag{4.36}$$

$$P_b = d_1' \cdot x_4^{o} + d_1' \cdot x_5^{o} \tag{4.37}$$

$$d_r = 1 - (S_3 + S_4 + S_5)/3 \tag{4.38}$$

where:

S_2: "S" Curve: $S_2(x_2, x_1^{o})$

x_1^{o}: Parameter of the "S" curve

f_u: Overflow from x_1

i_f: Interflow

e_2: Evaporation from x_2

P_r: Percolation to the lower zone

d_u: Interflow's parameter

P_b: Maximum rate of percolation under complete saturation of the lower zone

d_1': Parameter for the lower zone primary reservoir

x_4^{o}: Parameter of S_4

d_1': Parameter for the lower zone supplementary reservoir

x_5^{o}: Parameter of S_5

d_r: Defficiency ratio

γ: Parameter of the percolation function

S_3: "S" curve of x_3: $S_3(x_3, x_3^{o})$

x_3^{o}: Parameter of S_3

S_4: "S" curve of x_4: $S_4(x_4, x_4^{o})$

x_4^{o}: Parameter of S_4

S_5: "S" curve of x_5: $S_5(x_5, x_5^{o})$

x_5^{o}: Parameter of S_5

Lower zone tension water element, x_3:

$$dx_3/dt = P_r \cdot (1 - P_f) \cdot (1 - S_3) - e_3 \quad (4.39)$$

$$e_3 = u_2 \cdot (1-S_1) \cdot (1-S_2) \cdot S_3 \quad (4.40)$$

where

e_3: Evaporation from x_3

Lower zone free primary reservoir, x_4

$$dx_4/dt = [P_r \cdot P_f + P_r \cdot (1-P_f) \cdot S_3] \cdot F - r_b' \quad (4.41)$$

$$r_b' = d_1' \cdot x_4 \quad (4.42)$$

where

P_f: Parameter that indicates the amount of water flowing directly to the free water zones

F: Function that controls the filling of the lower zone reservoirs.

Lower zone free supplementary reservoir, x_5:

$$dx_5/dt = [P_r \cdot P_f + P_r(1-P_f).S_3] \cdot (1-F) - r_b'' \quad (4.43)$$

$$r_b'' = d_1'' \cdot x_5 \quad (4.44)$$

Upper channel reach, x_6:

$$dx_6/dt = T_{ci} \cdot (1 - Dc) - f_c \quad (4.45)$$

$$f_c = Ck \cdot x_6^{m'} \quad (4.46)$$

$$T_{ci} = r_d + SR + i_f + r_{bc} \quad (4.47)$$

$$r_{bc} = (r_b' + r_b'')/(1 + \mu) \quad (4.48)$$

where

Dc: Fraction of T_{ci} flowing to the

second reach (Parameter)

f_c: Outflow from the upper channel reach

μ: Fraction of groundwater not flowing to the channel (Parameter)

Ck: Parameter of the channel

m': Exponent of the non-linear channel

Lower channel reach, x_7

$$dx_7/dt = T_{ci} \cdot Dc + f_c - T_{co} \quad (4.49)$$

$$T_{co} = 0.5 \cdot Ck \cdot x_7^{m'} \quad (4.50)$$

Notice that the channel coeffient in the lower reach of the channel is exactly one half of the coefficient in the upper reach. If the model were going to be applied to a real catchment, the channel coefficients in both reaches should be allowed to vary arbitrarily in order to better fit a real channel. Since the purpose of this model was to be used only with simulated data this simplification was used in order to decrease the number of parameters to be estimated.

CHAPTER 5

EXTRACTION OF PRIOR INFORMATION FROM HYDROLOGICAL RECORDS

5.1 Introduction

A manual calibration procedure for the parameters of the National Weather Service model is very well described by Peck, (1976). In that report, the parameters are classified into five categories. The first category groups the parameters which are easily computed from the observed hydrograph and precipitation. The second category groups parameters more difficult to estimate from the same records. The third category has only one parameter, the area of the basin, which is measured from maps of the basin. The fourth category covers the percolation function parameters, which are estimated from neighboring basins which have been previously simulated. The fifth category covers three parameters for which only nominal starting values are given.

From Peck's classification of the parameters' observability we can see that prior information on the parameters in the first two categories is available in portions of the discharge records. The prior information on the parameters of the fourth category comes from other sources, in this case neighboring basins for which the parameters have already been estimated. We are interested in extracting the prior information which is contained in portions of the records themselves.

The estimation of the parameters under the first two categories take advantage of the fact that some elements of the model would be inactive during some periods of time, thus making the parameters related to the active elements highly observable. Following that idea it is then possible to develop stochastic models of small dimensionality that represent one or two of the elements that may be active at a particular period. The problem of parameter estimation may then be posed as series of hypothesis tests in which not only the soil parameters but also the error terms and the time periods for which the models (sub sets of the larger model) apply have to be identified. The ability to identify the time periods for which the models apply is indeed the key issue in justifying the development of a stochastic model to be used in estimating parameters that are, otherwise, easily identified "by eye" from the records. In addition, the maximum likelihood approach in which the algorithm is based provides the variance of estimation of the parameters, which, together with the value of the parameters fully describes the prior normal distribution of the parameters (See Chapter 3). We can also obtain the variances of the model error and the measurements error terms with the stochastic procedure. While the identification of the last two variances require the use of a stochastic parameter estimation technique, the variance of the parameter estimates can be approximated by repeatedly estimating the pa-

rameters in different time intervals, using a deterministic procedure. The larger the number of estimates for each parameter, the better the approximation of the variance. This will, nevertheless, increase the costs of employing techniques alternate to maximum likelihood parameters estimation, thus reducing the cost advantage these alternate techniques may have had over maximum likelihood.

This chapter deals with the presentation of an algorithm for automatically identifying the region of base flow and related parameters using the maximum likelihood procedure. The resulting parameters and estimation variances will then be prior information in estimation of parameters of the full NWSRFS model. Chapter 3 discussed how to incorporate this prior information in the maximum likelihood procedure. The resulting increased observability will hopefully reduce the common problem of parameters achieving non realistic values during automatic calibration.

5.2 Stochastic Model of Base Flow

A schematic diagram on which the stochastic base flow model is based, is presented on Figure 5.1. In developing the model, the following assumptions are made:

1. The only input to the system is white Gaussian noise. The upper zone free water element is fully depleted implying that there is no percolation or interflow.
2. There is no transfer of water from the lower zone free water elements to the lower zone tension water element.
3. There is no evapotranspiration from the channel and riparian vegetation.
4. The ratio of unobserved discharge to observed discharge (μ) is taken as a nominal value.
5. The fixed rate of discharge lost from the total channel flow (NWSRFS acronym SSOUT) is negligible.
6. Due to the slow varying conditions of the river discharge during low flow periods, channel routing may be disregarded.

The first assumption is justified by the very definition of base flow. This is the only component of the discharge when there are no inputs to the system and when there is no interflow. Assumptions 2 and 3 are also made in the Kitanidis and Bras (1980) model. These authors showed that the modeling of these conditions was of negligible importance. Assumption 4 is an important one, since the influence of parameter μ is not negligible by any means. This assumption, however, will be fully justified later in this section after the model of base flow and its associated parameters are introduced. Assumption 5 is justified in all cases in which there is not strong contradictory geological evidence. If this evidence is present, the model of base flow proposed here will not be directly applicable. A modification to it, suggested in

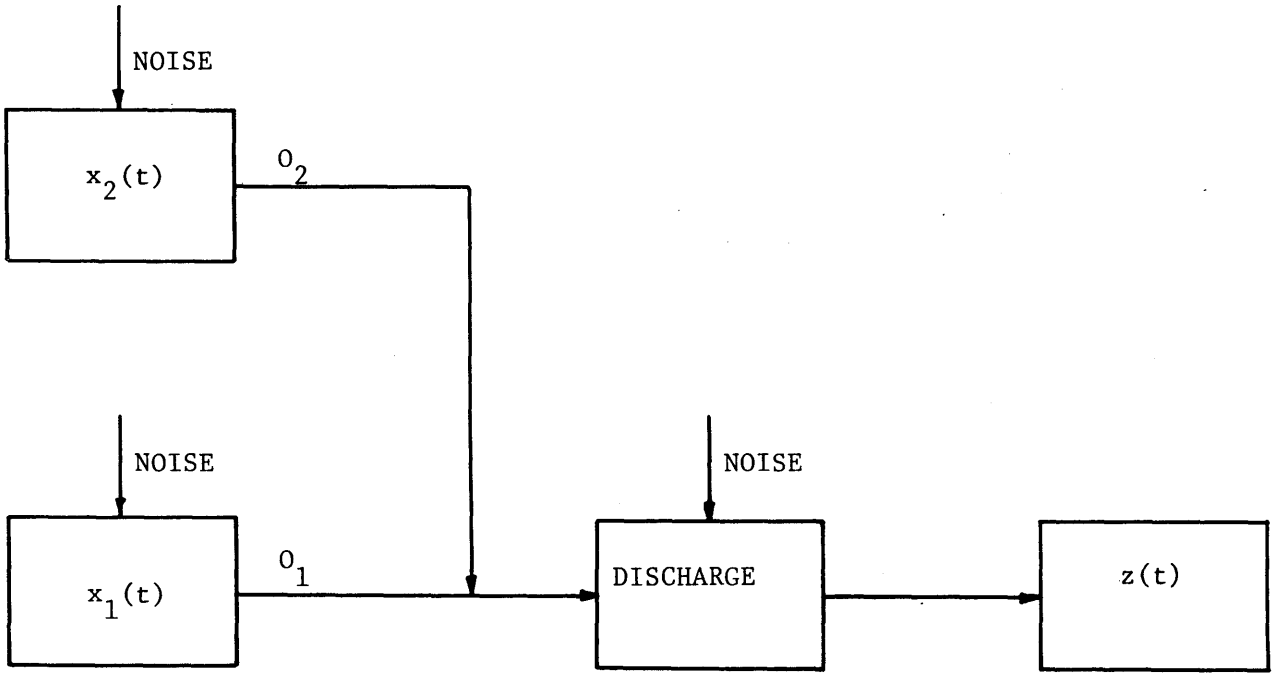


Figure 5.1 Schematic Representation of the Stochastic Model of Base Flow.

the final chapter, will be able to handle this case. The tests built-in in the model will assure that the procedure to be presented here does identify the base flow interval and its associated parameters which follow the previous set of assumptions.

The procedure we are going to follow in deriving the stochastic differential equations of the model of base flow involves the differentiation and integration of a system of differential equations. Since the first of our assumptions state that the only input to the system is white gaussian noise, we should start from a system of stochastic differential equations. However, since white gaussian noise is a non differentiable process, we will base our derivation on a differentiable deterministic system. Moreover, since the discharge measurements that will serve to update the states are available at discrete time steps, the final governing equations will be in the form of discrete-time difference stochastic equations. These equations will be derived by integration of the continuous-time deterministic differential equations and by addition of discrete-time white gaussian noise processes.

The continuous time deterministic differential equations describing the base flow model are:

$$d\tilde{x}_1/dt = I_1(t) - O_1(t) = - d_1' \tilde{x}_1(t) \qquad (5.1)$$

$$d\tilde{x}_2/dt = I_2(t) - O_2(t) = - d_1'' \tilde{x}_2(t) \qquad (5.2)$$

with initial conditions:

$$\tilde{x}_1(0),\ \tilde{x}_2(0)$$

In the above equations the subscript 1 refers to the water content in the lower zone free primary water element, and the subscript 2 represents the water content in the lower zone suplementary water element. The two (unknown) parameters d_1' and d_1'' correspond to the linear reservoir constants d_1' and d_1'' respectively.

Since the measurements are taken at discrete time intervals, Δt, rather than continuously, the system can be modeled in a discrete-time equivalent form by direct integration of Equations (5.1), and (5.2)

After integration from time t to time t + Δt, the discrete time stochastic system equations become:

$$\tilde{x}_1(t+\Delta t) = \exp(-d_1'\Delta t)\cdot\tilde{x}_1(t) + \tilde{W}_1(t) \qquad (5.3)$$

$$\tilde{x}_2(t+\Delta t) = \exp(-d_1''\Delta t)\cdot\tilde{x}_2(t) + \tilde{W}_2(t) \qquad (5.4)$$

$$z(t) = O_1(t) + O_2(t) + V(t)$$

$$= d_1' \tilde{x}_1(t) + d_1'' \tilde{x}_2(t) + V(t) \qquad (5.5)$$

In which $\tilde{W}_i(t)$ is a discrete time white gaussian noise process.

Notice that the original parameters appear now in a highly non linear form, in equations (5.3) and (5.4). This poses no problem for the parameter estimation procedure, which is, anyway, the optimization of a nonlinear function.

Since the discharge of the river is the rate of change in the content of the linear reservoirs, which is an indirect measurement of the states, we chose to change variables so the new state variables are observed directly.

Let us define

$$x_1(t) = d\tilde{x}_1(t)/dt \qquad (5.6)$$

$$x_2(t) = d\tilde{x}_2(t)/dt \qquad (5.7)$$

as the rates of discharge of both reservoirs, respectively.

By differentiating Equations (5.1) and (5.2), replacing equations (5.6) and (5.7) into equation (5.5), the continuous time system equation in the new state space becomes,

$$dx_1(t)/dt = -d_1' x_1(t) \qquad (5.8)$$

$$dx_2(t)/dt = -d_1'' x_2(t) \qquad (5.9)$$

Integrating Equations (5.8) and (5.9) from time t to time t+Δt and adding the discrete time white gaussian noise processes, the discrete time system equations become,

$$x_1(t + \Delta t) = A_1 x_1(t) + W_1(t) \qquad (5.10)$$

$$x_2(t + \Delta t) = A_2 x_2(t) + W_2(t) \qquad (5.11)$$

$$z(t) = x_1(t) + x_2(t) + v(t) \qquad (5.12)$$

In which

$$A_1 = \exp(-d_1'\Delta t) \tag{5.13}$$

$$A_2 = \exp(-d_1''\Delta t) \tag{5.14}$$

$W_1(t)$, $W_2(t)$ and $v(t)$ are discrete time independent white gaussian noises with zero mean and unknown variances, Q_{11}, Q_{22}, and $R_1(t)$, respectively, which must be estimated.

Q_{11} and Q_{22} will be assumed to be constant, while the modeling of the variance of $v(t)$ follows the research by Georgakakos and Bras (1980), in which it is assumed that the measurement noise ($E[v(t)^2]$) is proportional to the discharge. It will be assumed that at any time t, $R_1(t)=R \cdot z^2(t)$, and R is the parameter to be estimated.

5.2.1 Observability of the Model.

The system dynamics represented by Equations (5.10) and (5.11), and the measurement equation (5.12), can be represented in matrix form as

$$\underline{x}(t + \Delta t) = \underline{A}\ \underline{x}(t) + \underline{W}(t) \tag{5.15}$$

$$z(t) = \underline{H}\ \underline{x}(t) + v(t) \tag{5.16}$$

where

$$\underline{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{5.17}$$

$$\underline{A} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \tag{5.18}$$

$$\underline{w}(t) = \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} \tag{5.19}$$

$$H = [1\ 1] \tag{5.20}$$

The observability of a model defines the ability to determine the sequence of state variables $\underline{x}(1)$, $\underline{x}(2)$,..,$\underline{x}(k)$, from a given set of measurements $z(1)$, $z(2)$,...,$z(k)$. There are two tests to check observability in a system. A deterministic test, and a stochastic test. As we will show next, both tests yield the same results when applied to the model of base flow. The deterministic observability test consists in determining the rank of the matrix Ξ, defined as

$$\Xi = [\underline{H}^T : \underline{H}^T\underline{A}^T : \underline{H}^T\underline{A}^{2T} \ldots] \tag{5.21}$$

(For details, see Gelb, 1974, pag. 67)
In our case,

$$|\Xi| = \begin{vmatrix} 1 & A_1 \\ 1 & A_2 \end{vmatrix} = A_2 - A_1 \tag{5.22}$$

Which implies that the system is non observable, in the deterministic sense, only if the two coefficients of the linear reservoirs are identical.

The stochastic observability test is based on the existance of the states' covariance matrix, in the absence of process noise and a priori information about the state variables. (See Gelb, 1974, pp. 131-132, for details). Under these conditions, the inverse covariance matrix can be shown to be expressed by

$$P_x^{-1}(t) = \int_o^t \underline{A}^T(t,\tau)H^T R^{-1}H \underline{A}(t,\tau)\, d\tau \tag{5.23}$$

In which the matrix $\underline{A}$ is the state transition matrix, which is defined by Equations (5.13), (5.14) and (5.15). Notice that here we have included explicitly the dependency of $\underline{A}$ on the interval $\Delta t = t - \tau$. The stochastic observability criterion establishes that the system is stochastically observable, if for some $t > 0$ the integral in Equation (5.23) is positive definite and bounded. For linear stationary systems like ours the criterion corresponds exactly to the deterministic observability criterion (Smith, G. L., 1965, pp. 350-359).

5.2.2. Weighted Least Squares Parameter Estimation. (WLS)

Since the method proposed in this chapter is presented as an alternative to manual estimation methods, it is convenient to compare the estimates of the parameters obtained with maximum likelihood, manual estimation methods and a popular parameter estimation method, weighted least squares. Minimizing the weighted sum of squares of the observed discharges minus the predicted discharges is a very common criterion of parameter estimation. The criterion can be simply stated as

$$\underset{\underline{\theta}}{\text{Min}}\ J = \sum_{t=1}^{T} (z(t) - z(\underline{\theta}, t))^2\, w(t) \tag{5.24}$$

In which $\underline{\theta}$ is the unknown set of parameters, $z(t)$ are the observed discharges, $(\underline{\theta},t)$ are the predicted discharges, and $w(t)$ are the weights.

A set of weights consistent with the modeling of the measurement error is

$$w(t) = R/z^2(t) \qquad (5.25)$$

since it resembles the modeling of the standard deviation of the measurement errors being proportional to the discharge, as used in the stochastic model.

The WLS criterion which will be applied here can then be rewritten as

$$\text{Min } J = R \sum_{t=1}^{T} [(z(t) - \hat{z}(\underline{\theta},t))/z(t)]^2 \qquad (5.26)$$

5.2.3. Maximum Likelihood Parameter Estimation. (MLPE)

This section discusses some characteristics of the loglikelihood function for the model under consideration. The study of these characteristics is an invaluable tool in defining the strategy to follow when a problem-oriented, rather than a general optimization scheme is being developed.

A convenient way to study the features of the loglikelihood function is through its graphical representation. A three dimensional isometric view of the function of two parameters, combined with a contour lines plot of the same, provides global appreciation of the characteristics of the function over a wide range of the two parameters simultaneously. The following sections include these kind of plots, which were obtained by generating <u>synthetic</u> discharges with the set of parameters

$$A = \begin{bmatrix} 0.7 & 0 \\ 0 & 0.3 \end{bmatrix} \qquad \underline{x}(0) = \begin{vmatrix} 4 \\ 4 \end{vmatrix}$$

$$Q_{11} = Q_{22} = 10^{-6} \quad E[\nu^2] = 10^{-4} \; z^2$$

Next, two of the parameters at a time were varied within a range, (described in the corresponding section), forming a 40 X 40 grid. The remaining parameters were kept constant at their <u>real</u> values.

5.2.3.1 Loglikelihood of the Model Dynamics' Parameters.

The loglikelihood function of the two model dynamics parameters is shown in Figures 5.2 and 5.3. The reader should notice that the interval between contour lines varies so that important features such as the optima can be easily located. The two parameters were varied linearly from $A_1=A_2=0.1$, to $A_1=A_2=1.0$. Two characteristics of the loglikelihood function are immediately evident. First, the function is symmetric with respect to the plane $A_1=A_2$. Second, there is a cusp in the function along the same plane.

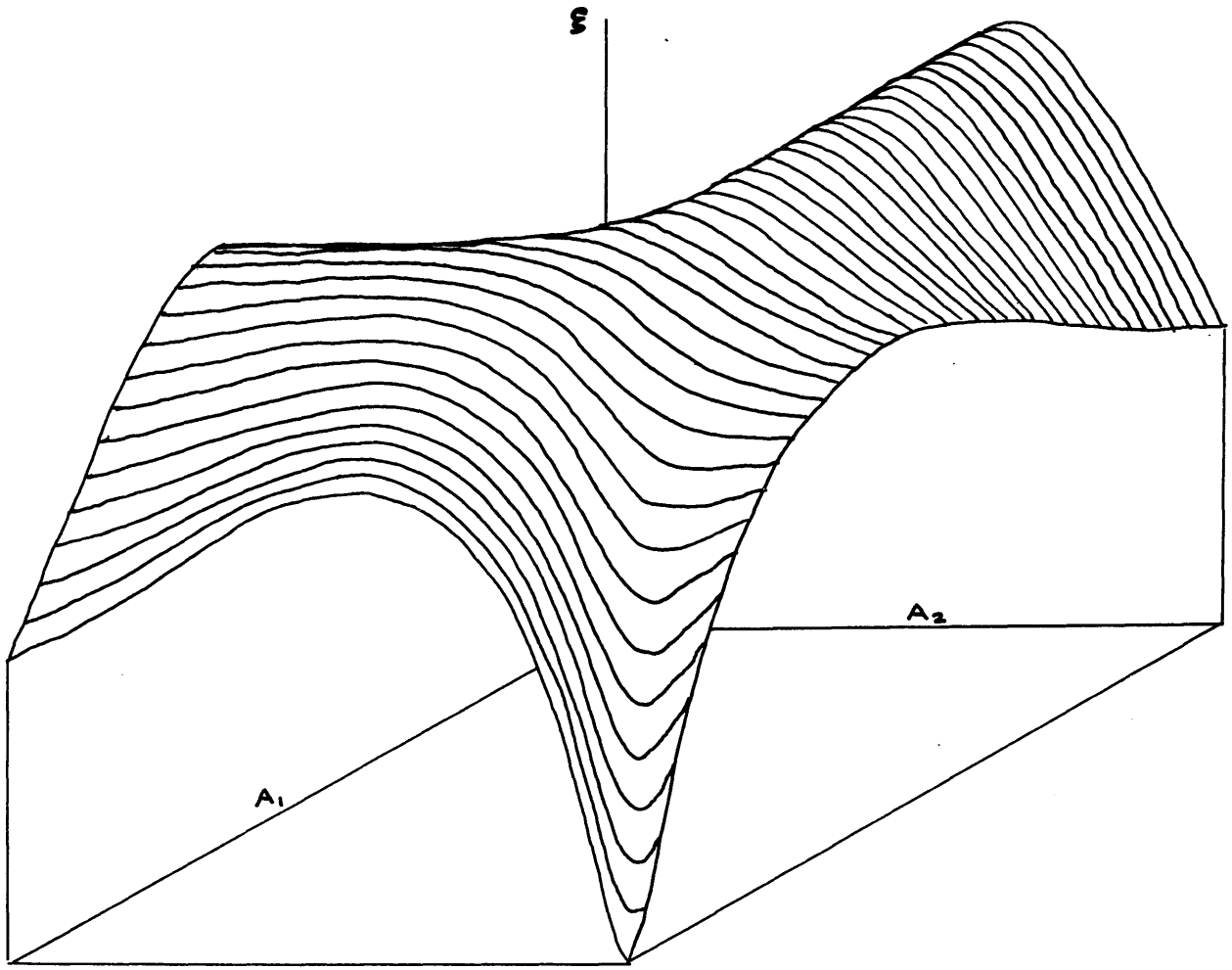


Figure 5.2 Isometric View of the Loglikelihood Function of the Model Dynamics Parameters.

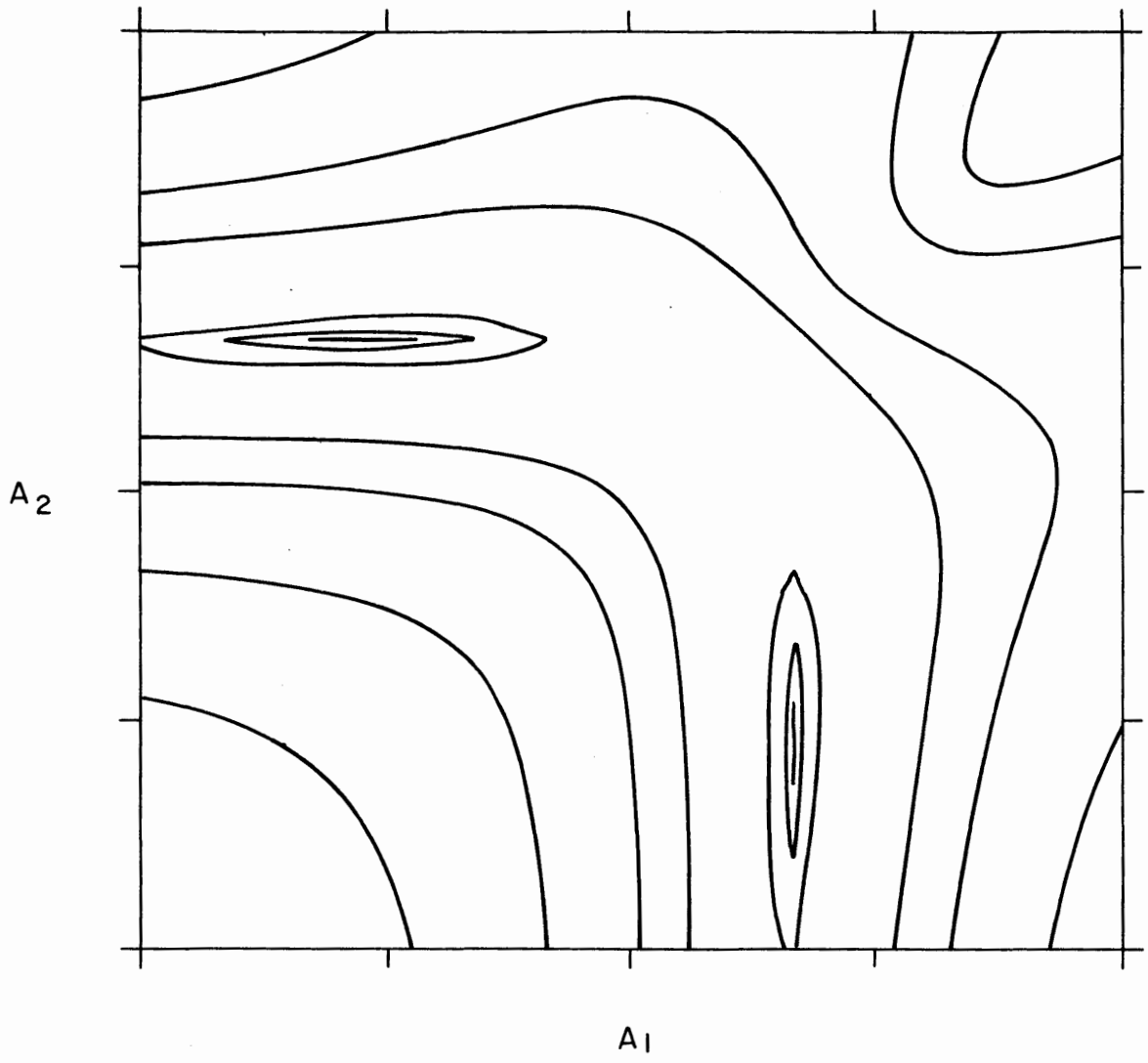


Figure 5.3 Contour Lines of the Loglikelihood Function of the Model Dynamics Parameters.

The symmetry appears when both initial conditions and both error terms are, respectively, identical. This symmetry is an important consideration. Any optimization procedure requires the choice of a starting point, an initial guess for the parameter's values, A_{1o}, A_{2o}, etc. A logical choice is to select $A_{1o} > A_{2o}$, which is equivalent to $d_1'_o < d_1"_o$, since the suplementary reservoir is faster draining than the primary one. Therefore, we expect the MLPE procedure to converge to a point such that the last inequality is not violated. During the non linear optimization procedure, it is possible that the search path crosses the symmetry plane and convergency is achieved on the maximum located at the opposite side of the symmetry plane from the starting point, with the consequence that $A_1^* < A_2^*$, or, equivalently, $d_1'^* > d_1"^*$, where the * denotes ML estimates. Should this be the case, the system dynamics parameters and other parameters being estimated, must be respectively interchanged.

The cusp in the loglikelihood function does not represent any problem in the MLPE procedure. It marks exactly the plane in which the state variables are non-observable.

5.2.3.2 Loglikelihood of the Model Error's Parameters.

The loglikelihood function of the two model error parameters is shown in Figures 5.4 and 5.5. The data for these figures was obtained by geometrically varying the model error terms from 10^{-10} to 10^{-3}.

Two important features are immediately distinguished. First, there is a plateau covering a wide range of the parameters' values, on which the loglikelihood function is almost insensitive to the model error parameters. Second, the loglikelihood function has a ridge that crosses the maximum point. The top of the ridge is esentially a horizontal line, which is parallel to one coordinate axis, for a very large range of the parameter's values. For some others cases, in which the parameters used in generating the synthetic data were different from the current ones, it has been observed that the plateau is bounded by a ridge on two of its sides. (Fig. 5.6).

The implications of the above features are the following. First, both model error parameters are non identifiable if the starting point lies on the plateau. Second, if the final convergency is on the ridge, the Hessian will be singular. This implies that one of the most valuable results of the MLPE technique, the lower bound of the parameters covariance matrix, which is defined as the inverse of the negative Hessian, can not be obtained.

These problems indicate that an alternative to estimating these parameters may be the selection of nominal values for them. How good that alternative is must be judged

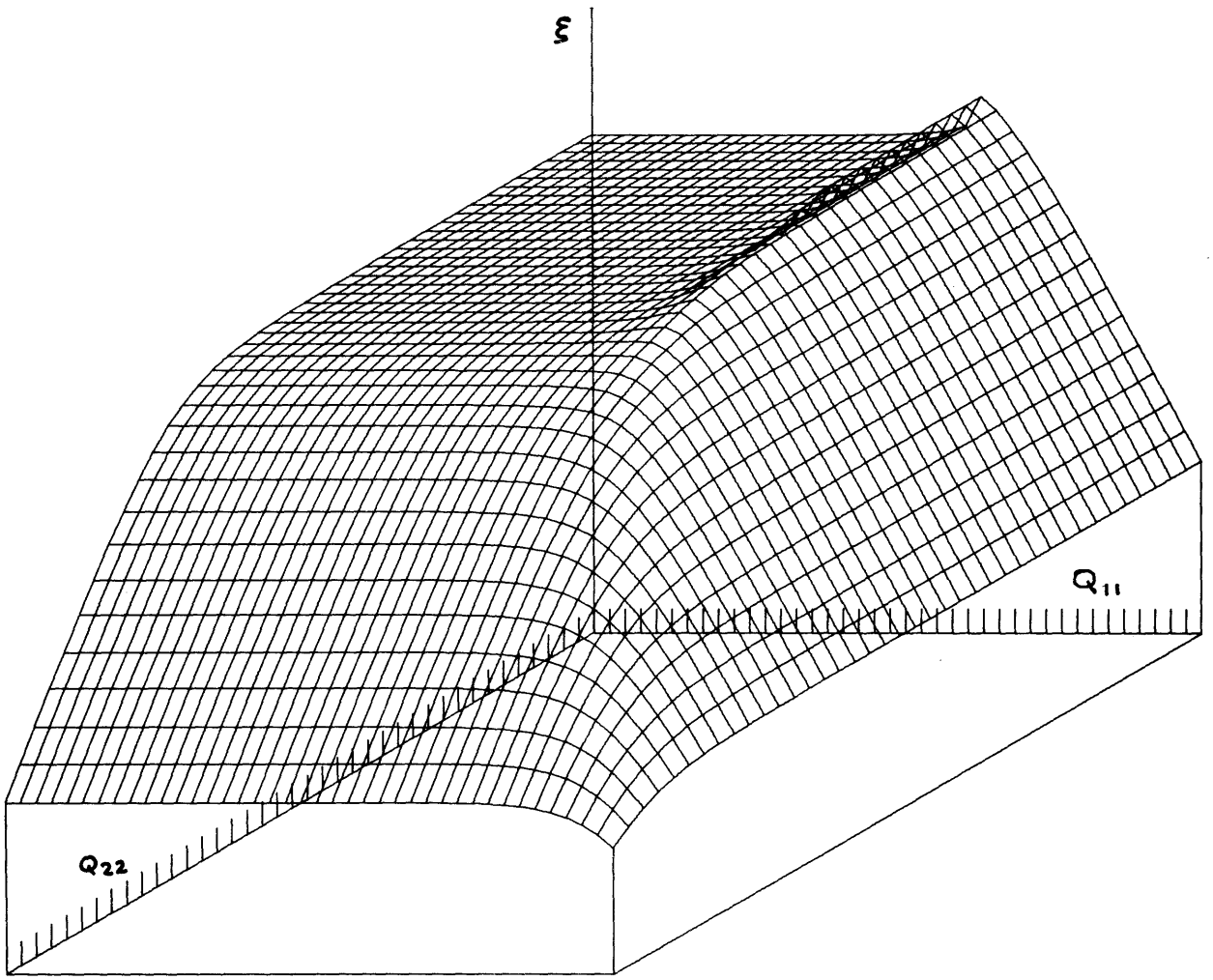


Figure 5.4 Isometric View of the Loglikelihood Function of the Model Error Parameters.

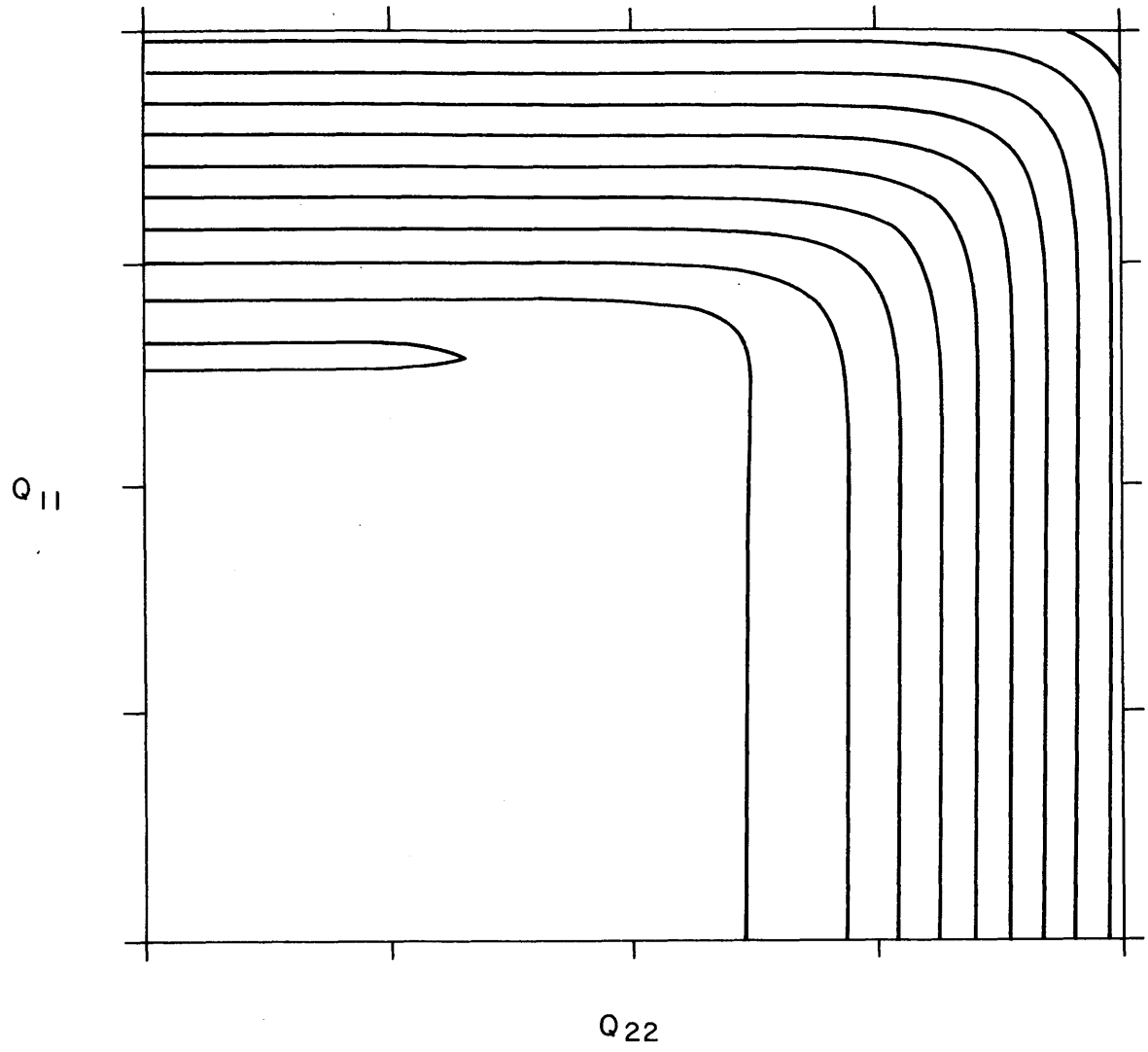


Figure 5.5 Contour Lines of the Loglikelihood Function of the Model Error Parameters.

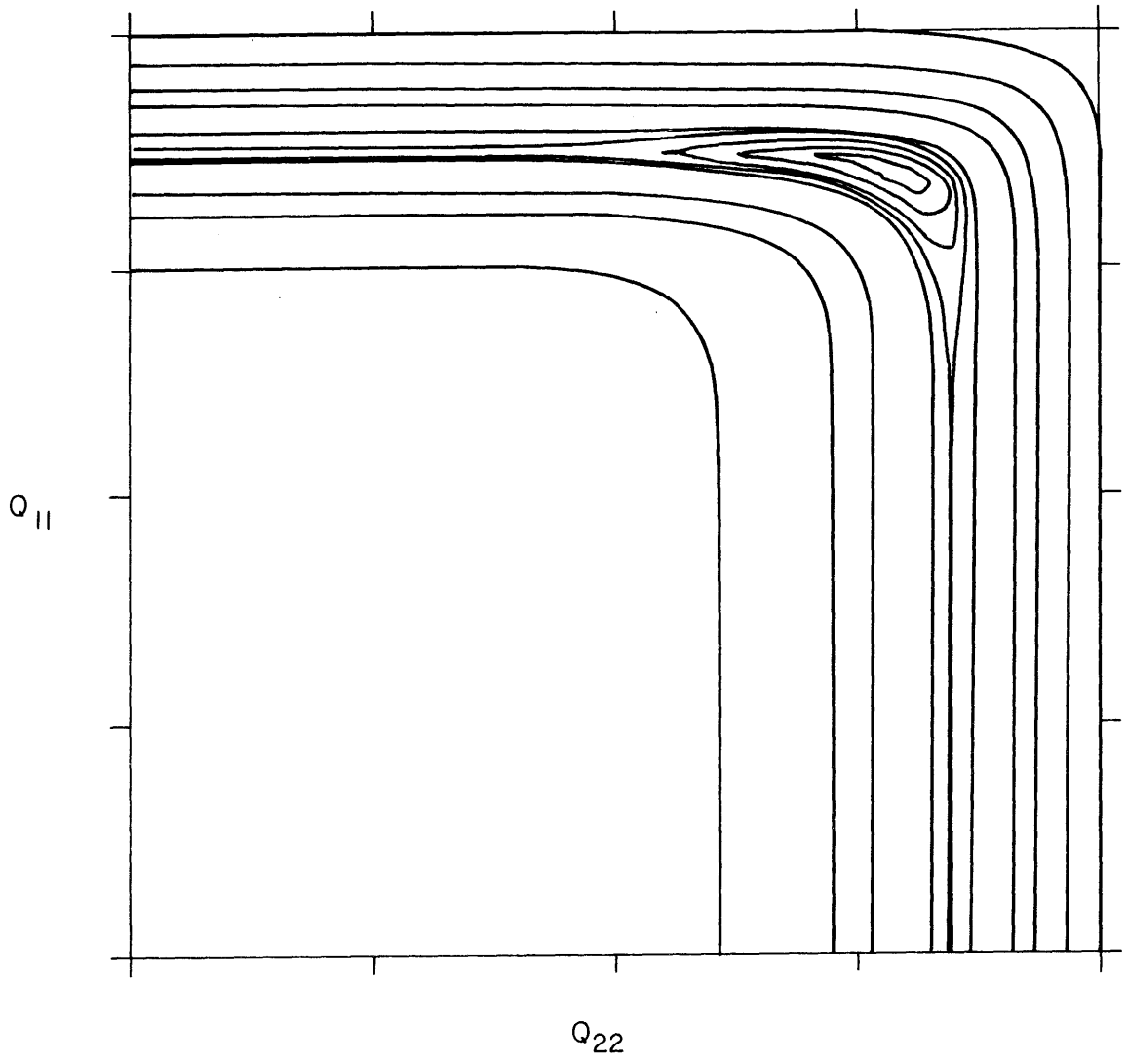


Figure 5.6 Contour Lines of the Loglikelihood Function of the Model Error Parameters with and Edge Around the Plateau.

by the effect of the value chosen for the model error terms on the other parameters estimates.

5.2.3.3 Estimation of the Initial Conditions of the State Variables.

The computational cost of non linear optimization procedures grows considerably according to the number of variables, thus making it desirable to keep the number of parameters to be estimated at a minimum. With this in mind, one might be tempted to assume nominal values for the initial conditions of the state variables instead of obtaining maximum likelihood estimates. The reasoning for this action may be that in optimal estimation theory, the Kalman filter will be able to "forget" about the initial conditions, if a large initial covariance matrix is given, and quickly converge on good estimates for the state variables. The fallacy of this approach lies on the fact that when the parameters are unknown, the estimation is NON optimal and the above property does not apply. Furthermore, a large initial covariance matrix will "blame" a large part of the prediction error on the state variables, rather than on the parameters. This will cause convergency of the MLPE to parameters' values which are close to the starting point, and probably, far from the real ones.

A very small initial covariance matrix, on the other hand, indicates very large confidence on the initial conditions of the state variables. This will make the current parameter estimates mostly responsible for the prediction errors, thus increasing the sensitivity of the loglikelihood function to the parameters. But, if the initial conditions are not close enough to the real ones, the parameters estimates will have to compensate for the bad initial conditions in the state variables, thus converging to values which are far from the real ones.

The consequence of the above arguments is that the initial conditions for the state variables must form part of the set of unknown parameters and must be jointly estimated with the remaining parameters of the model. Now that the need for estimating the initial conditions of the state variables as unknown parameters has been justified, we can justify assumption No. 4 at the beginning of the chapter.

Remembering that μ is defined as the ratio of unobserved discharge to observed discharge, including μ in the measuring equation of the base flow model leads to

$$z(t) = (x_1(t) + x_2(t))/(1+\mu) + v(t) \qquad (5.27)$$

or, in matrix notation,

$$z(t) = \underline{H}\ \underline{x}(t) + v(t) \qquad (5.28)$$

where $\underline{H} = [m \quad m]$
and $\bar{m} = 1/(1+\mu)$

For reasons of clarity, we will base our discussion below on the lower zone primary reservoir only. Although only one state variable will be included, the subscripts are kept. The conclusions will make it evident that the discussion is valid for the complete model of base flow.

At time t=1 x(1) and z(1) are given respectively by

$$x_1(1) = A_1 x_1(0) + W_1(0) \tag{5.29}$$

$$z(1) = m[A_1 \, x_1(0) + W_1(0)] + v(1) \tag{5.30}$$

The discharge at t=2 becomes

$$z(2) = m \, x_1(0) \, A_1^2 + m \, A_1 \, W_1(0) + m \, W_1(1) + v(2) \tag{5.31}$$

Taking expected values on both sides of the last two equations we arrive at

$$E[z(1)] = m \, x_1(0) \, A_1 \tag{5.32}$$

$$E[z(2)] = m \, x_1(0) \, A_1^2 \tag{5.33}$$

The variance of the discharges is obtained by subtracting the expected value from the measurement, squaring and taking expectations:

$$\sigma^2(1) = m^2 \, Q^2 + R^2 \tag{5.34}$$

$$\sigma^2(2) = m^2 \, Q^2 \, (A_1^2 + 1) + R^2 \tag{5.35}$$

Generalizing for t > 2, we can see from Equations (5.32) and (5.33) that the discharge contains only information about the <u>product</u> of the parameters m and $x_1(0)$, but not about the individual parameters. Hence, these parameters can not be jointly estimated. Similarly, from Equations (5.34) and (5.35) we see that the variance of the discharges contain information about the <u>product</u> of m and Q, but not about the individual parameters. It follows that m and Q are not jointly identifiable.

The solution to this problem is to fix m, (or equivalently fix μ) to some value, and estimate the other parameters with the awareness that the estimated parameters will be dependent on the value chosen for m. The largest value this parameter can take is unity, ($\mu = 0$), for the case in which all the groundwater discharge is observed through the channel. This implies that our estimates for the initial conditions will actually be lower bounds for these parameters. This was the criterion followed in this work.

5.2.4 Parameters Observability and Positive Definiteness of the Observation Matrix.

Earlier in this chapter we examined the observability of the states of the system when the parameters are known. In this section we are interested in examining the observability conditions for the unknown parameters. The discussions of this section are essential for the development of the algorithm to automatically identify the time interval of base flow activity which will be presented later on this chapter.

We say that a parameter θ_i is locally observable in the neighborhood of a point $\underline{\theta}_o$ when the first two partial derivatives of the loglikelihood function $\xi(\underline{\theta})$ in the direction of θ_i, evaluated at $\underline{\theta}_o$ do not vanish simultaneously.

Let us assume that we are trying to estimate the parameters of the base flow model (herein M_2) in an interval that corresponds to primary base flow exclusively. Since the discharge contains no information regarding the parameters of the lower zone supplementary reservoir, we would expect that all the partial derivatives of ξ with respect to the lower zone supplementary parameters would be zero under the non observability conditions presented above. In the way M_2 is formulated, however, the lower zone reservoirs are constantly driven by white gaussian noise, which would make them observable all the time, and consequently, the maximum likelihood parameter estimation procedure always computes non zero gradients with respect to the parameters of that model. Here we have, then, an apparent contradiction. We tell the model that the parameters are observable but we know that the discharge contains no information about the parameters of the supplementary reservoir for part of the data. This apparent contradiction is solved by the maximum likelihood parameter estimation by driving the coefficient of the supplemental reservoir, d_1", to zero. At that point, the first and second partial derivatives of the loglikelihood with respect to $x_1(0)$, (the initial condition), will be zero, since that parameter will have no effect on the likelihood function. This will cause the information matrix of M_2 to be non positive definite, making non positive definiteness a necessary but not

sufficient condition for non-informativeness. It is not a sufficient condition since convergence to a non optimum point, such as a saddle point, will cause the same non positive definiteness.

5.2.5. Model Validation.

5.2.5.1 Selected Data.

A dry period covering the months of July, 1957 and August, 1957 for the Bird Creek basin was selected to test the model under real conditions. The discharge during that period is shown on Figure 5.7.

Since there is an interval in which the semilogarithmic plot of the discharge can be approximated by two straight lines, "a" and "b" in Figure 5.7, it may be assumed that the discharge in that interval is indeed due to base flow. The slopes of these lines are estimates of the system dynamics parameters, A_1 and A_2, from which the coefficients of discharge of the linear reservoirs, d_1' and d_1'', respectively, can be calculated with Equations (5.13) and (5.14). The values for d_1' and d_1'' thus calculated are presented on Table 5.1, together with the values for the same parameters calibrated by the National Weather Service staff, and with the results obtained with the WLS and MLPE techniques.

5.2.5.2 Weighted Least Squares.

The WLS procedure defined by Equation (5.26) was performed with real data. The predicted values were calculated by means of a deterministic model which is described by Equations (5.15) to (5.20), when the model error parameters are set equal to zero. The discharges predicted with the deterministic model with the parameters estimated by WLS, are compared to the measured ones in Figure 5.8

5.2.5.3 Maximum Likelihood Estimation.

The parameters estimates and the corresponding coefficients of variation obtained with the MLPE procedure are presented in Table 5.1. The correlation matrix, and the serial correlation of the residuals are included in Table 5.2. Figure 5.9 shows the predicted and the measured discharges, in which the parameters estimated by MLPE have been used.

The lower bounds on the error for the parameter estimates (Table 5.1) are under 15% of the parameter values. The correlations among the parameter estimates are also small. Both results indicate that good estimates of the parameters were obtained. Comparing the parameter estimates by the different procedures we see that the estimates of d_1' given by the NWS, the visual fitting performed by the author and that

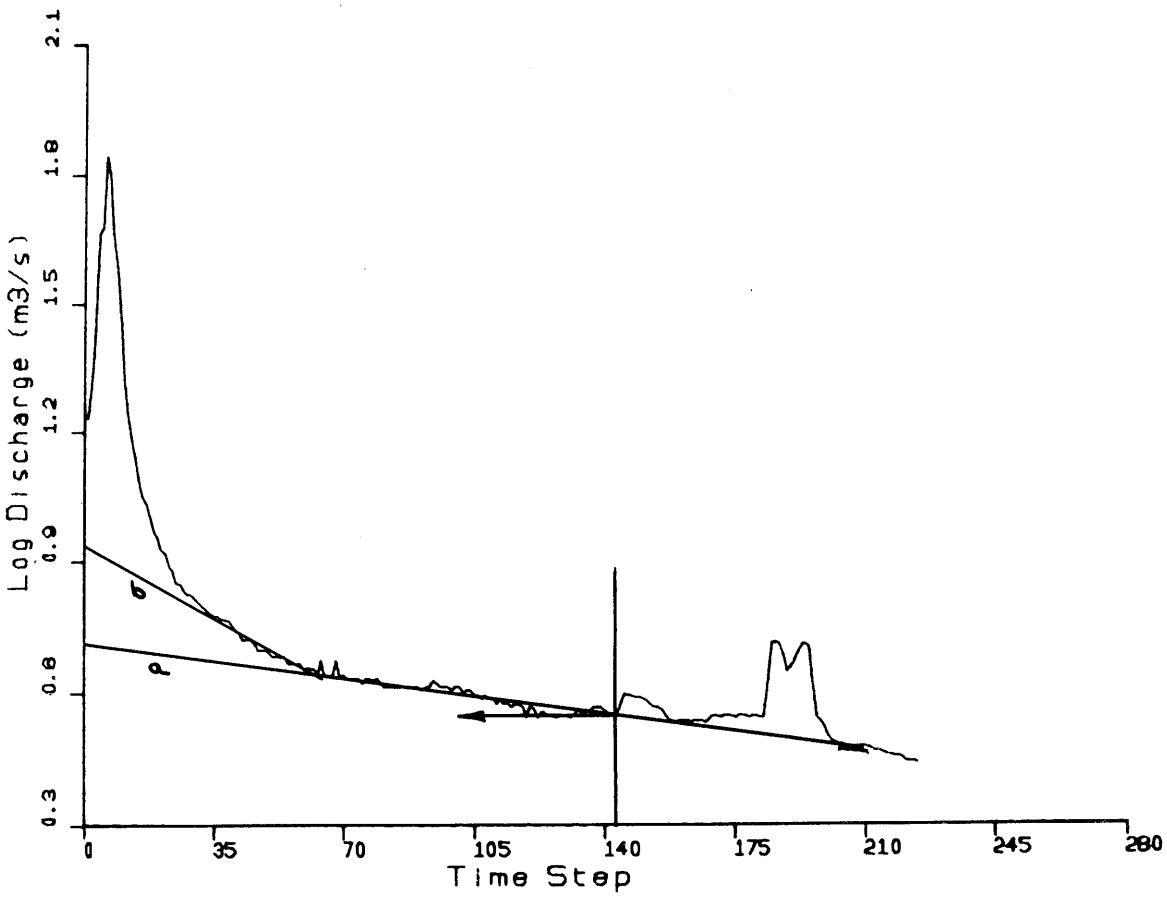


Figure 5.7 Semilogarithmic Plot of Bird Creek Discharges for July and August 1957.

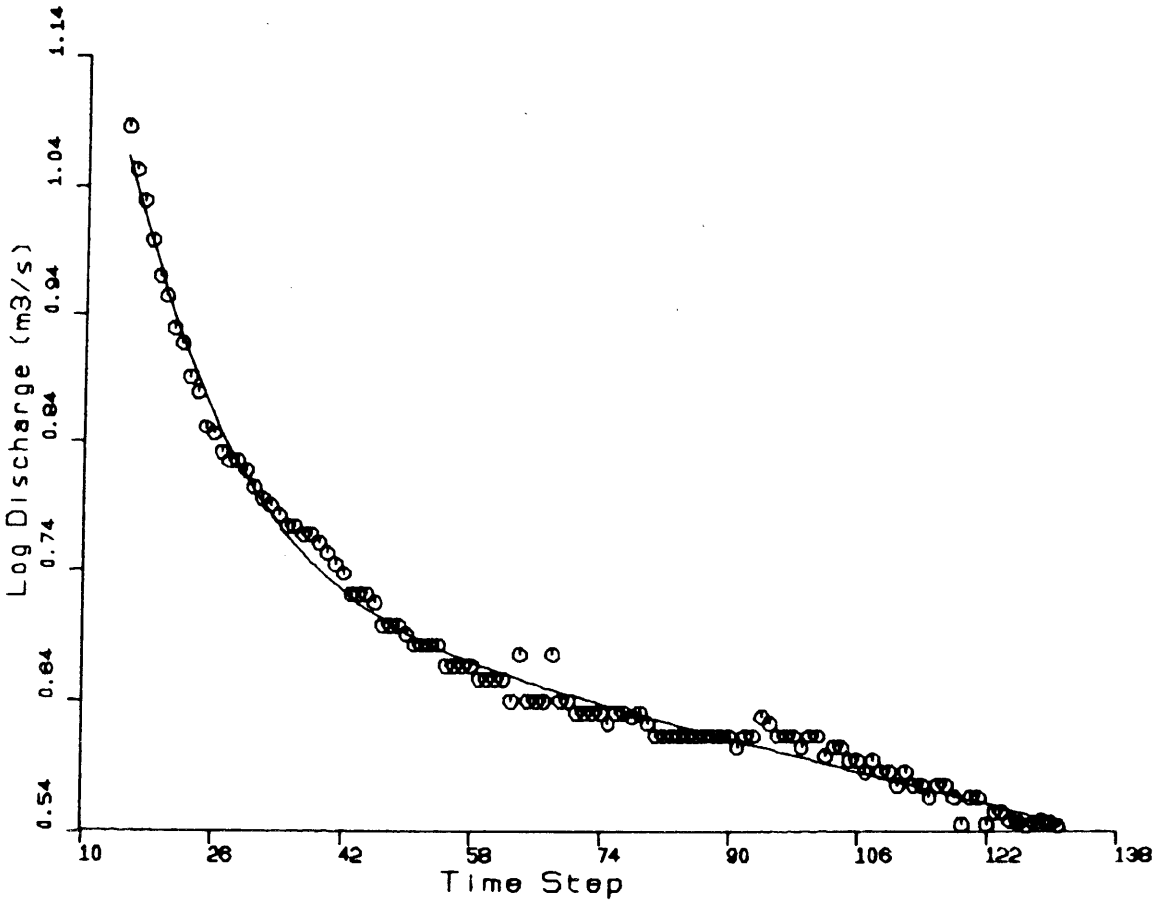


Figure 5.8 Six-Hour Lead Forecast of Bird Creek Discharges Using a Deterministic Base Flow Model With Parameters Estimated by WLS.

Parameter	NWS	Visual	WLS	ML	CV
d_1'	0.013	0.013	0.014	0.020	0.12
d_1''	0.126	0.338	0.390	0.550	0.06
$X_1(0)$	-	-	5.35	6.04	0.06
$X_2(0)$	-	-	6.89	7.00	0.05
Q_{11}	-	-	-	1.96×10^{-3}	0.15
R	-	-	-	1.79×10^{-4}	0.13

Note:

NWS: National Weather Service

WLS: Weighted Least Squares

ML: Maximum Likelihood

CV: Coefficient of Variation

Table 5.1 Parameter Estimates by Different Procedures

Correlation Matrix

	d_1'	d_1''	$X_1(0)$	$X_2(0)$	Q_{11}	R
d_1'	1					
d_1''	0.41	1				
$x_1(0)$	0.37	0.54	1	S Y M M E T R I C		
$x_2(0)$	-0.27	-0.10	-0.11	1		
Q_{11}	-0.07	0.54	0.02	0.03	1	
R	0.06	-0.03	-0.01	-0.03	-0.13	1

Serial Correlation Coefficients of the Normalized Residuals

Lag	0	1	2	3	4
ρ	1.01	-0.05	0.03	0.00	0.34
r	0.04	-0.54	0.30	0.03	3.72

Sum of Squares Test

Number of Data Points:	117
Number of Parameters:	6
Expected Sum of Squares:	111
Standard Deviation:	14.9
Computed Sum of Squares:	116.54
Deviation (Units of S. Dev):	0.37
Durbin and Watson Statistic:	1.022

Table 5.2 Maximum Likelihood Post Optimality Analysis.

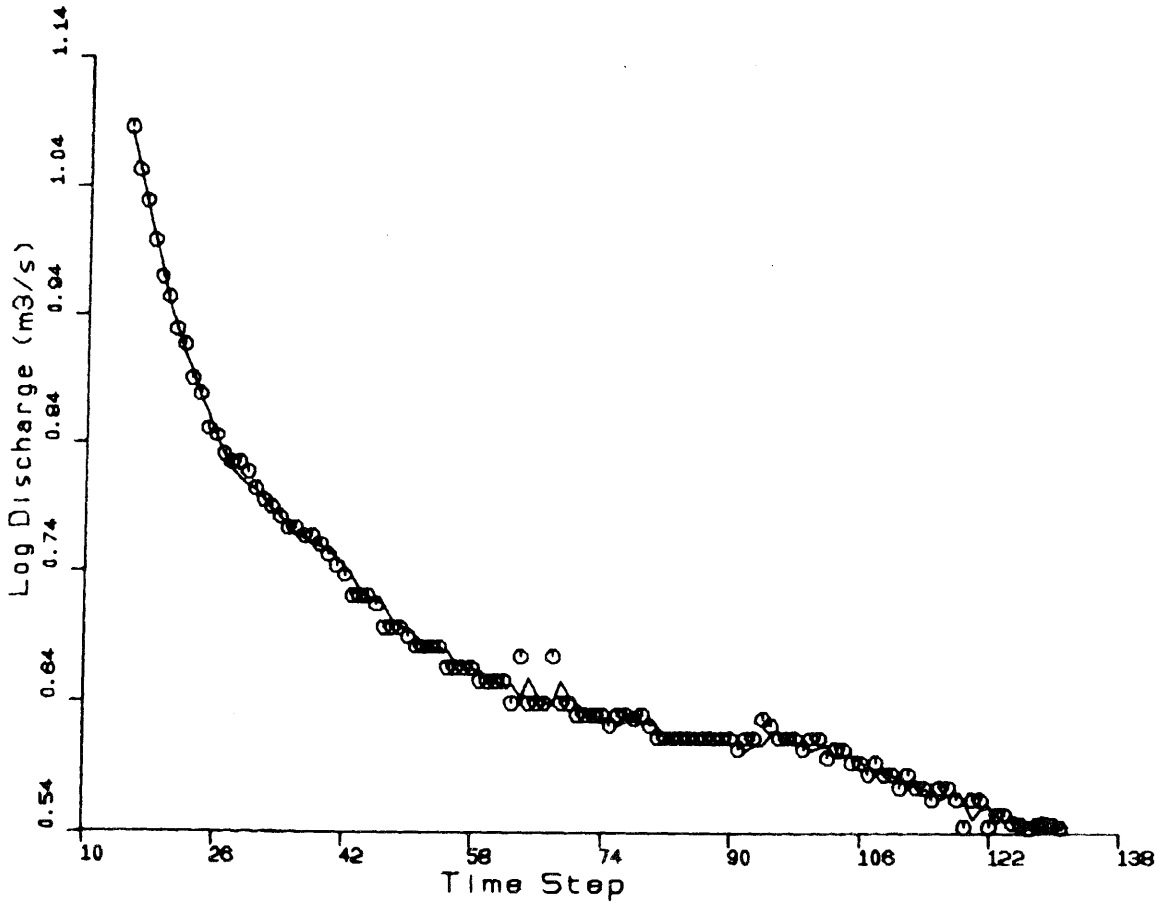


Figure 5.9 Six-Hour Lead Forecast of Bird Creek Discharges using a Stochastic Base-Flow Model with Parameters Estimated by ML.

obtained by WLS are essentially identical, while differing considerably from the MLPE estimate. This can be explained because a visual fitting is an approximated least squares. (The person performing the fitting tries to minimize the deviations between the predicted and the observed discharges). The method of maximum likelihood includes, in addition to the sum of the squares, the bias term of the likelihood function. We should not expect parameter estimates obtained from different estimation criteria to be identical. A wider range of values was obtained for the estimates of d_1", with the smallest value given by the NWS, and the highest given by MLPE. This wide range of values for that parameter would make us think that that parameter would not be highly observable. But paradoxically, the coefficient of variation in the maximum likelihood estimate of this parameter is only 6%.

The quality of the point of convergency for the maximum likelihood estimates can be judged by the comparison between the calculated sum of squares of the normalized residuals and its expected value and by the degree of independence of the residuals. The results indicate that the calculated sum of squares of the normalized residual is only 0.37 units of standard deviation from its expected value. This indicates that the point of convergency is very close to the ML point. The estimates of the serial correlation coefficients of the normalized residuals are excellent, with the exception of $\rho(4)$, which is 3.7 units of standard deviation off its expected value. The reason for this is that Bird Creek's discharge records are good, except during low flows, (i. e. during the period of base flow activity) in which the instantaneous discharges are manually estimated from the daily average. This is done by letting the instantaneous discharge be equal to the daily average instantaneous discharge. Since we have four "measurements" per day, the first two measurements are, on the average, under estimated, and the last two measurements are, also on the average, over estimated. (Figure 5.10) This introduces an error with a periodicity of four time steps (24 h) which is the cause of the very high $\rho(4)$. The Durbin and Watson statistic gives another measure of the independence of the residuals. A value of that statistic in the neighborhood of 2 would indicate a good degree of independence. The value of 1.022 indicates that the residuals are not independent. The problem, again, is blamed on the reconstruction of the instantaneous discharge from the daily average.

Up to this point we have shown how the maximum likelihood parameter estimation procedure performs when estimating the parameters of the base flow model in an interval of base flow. A question we would like to address now is how the procedure performs when estimating the parameters of the base flow model in an interval which is not a base flow interval. To answer this question we applied the maximum likelihood estimation procedure to estimate the parameters of the base

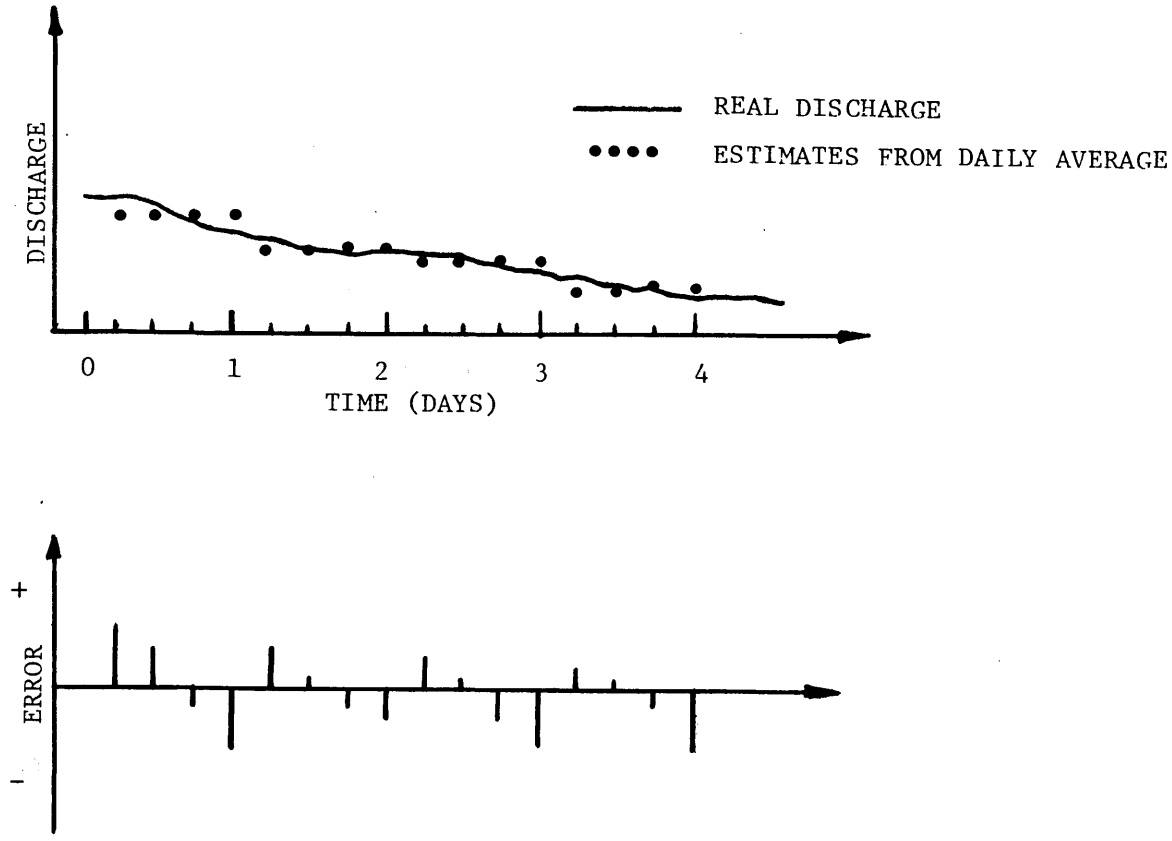


Figure 5.10 Error Introduced in the Discharge Measurements by Estimating the Instantaneous Discharges from Daily Averages

flow model in an interval that "looks like" an interval of base flow (Figure 5.11). Notice the two straight lines which fit the data reasonably well. Table 5.3 presents the parameter estimates by visual fitting, weighted least squares and maximum likelihood. All of these estimates are reasonably close, and the parameter estimates obtained from WLS and MLPE reproduce the discharges in a reasonable way. (Figures 5.12 and 5.13). The comparison between the computed sum of squares and its expected value, Table 5.4, indicates that point of convergency was close to the maximum likelihood point. The analysis of the normalized residuals indicates, however, that the residuals are highly correlated at lags 1 through 4 an indication that the filter is performing poorly. This result makes it clear that even though the discharge looks like being due to base flow, and that the models are able to reproduce the discharge reasonably well, that period is probably not an interval of base flow.

5.3. A Stochastic One-linear-Reservoir Model, M_1.

We will call M_1 a sub-model of the stochastic model of base flow, in which only the lower zone primary reservoir is active. The discussion included in the previous sections regarding the observability of the model noise parameters still applies.

The Stochastic One Linear Reservoir Model is defined by the state-space equation:

$$x_1(t+1) = A_1 x_1(t) + W_1(t) \tag{5.36}$$

with a measuring equation defined by:

$$z(t) = x_1(t) + v(t) \tag{5.37}$$

where:

$x_1(t)$:	Discharge at time t
A_1:	Unknown reservoir's parameter
$x_1(0)$:	Unknown initial condition
$W_1(t)$:	Discrete-time white gaussian noise
$z(t)$:	Observations at time t
$v(t)$:	Discrete-time white gaussian noise

The statistics of the white noise are:

$$E[W_1(t)] = 0$$

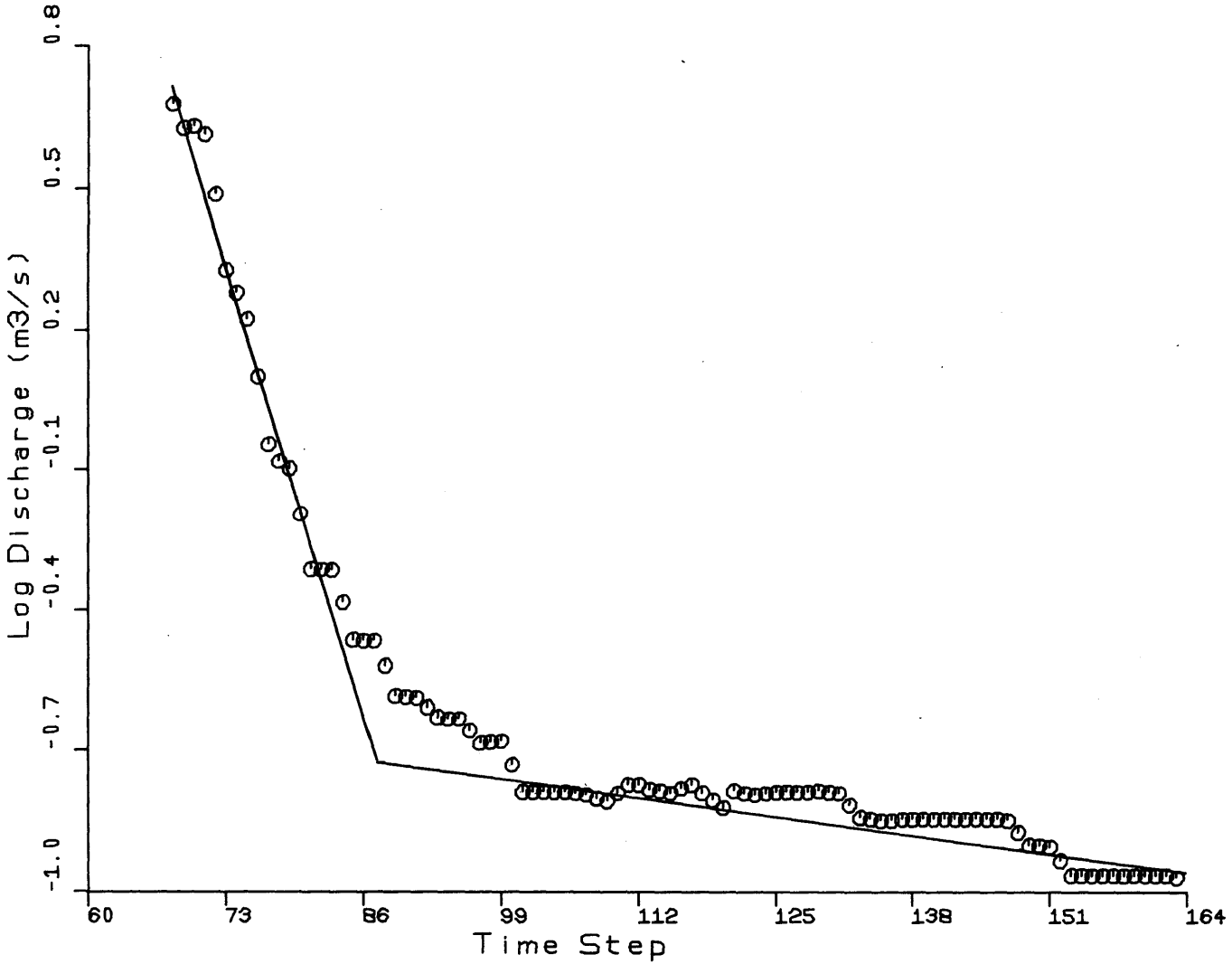


Figure 5.11 Semilogarithmic Plot of Bird Creek Discharges During a Non-Base-Flow Interval.

Parameter	Visual	WLS	ML	CV
A_1	0.99	0.992	0.973	0.011
A_2	0.88	0.811	0.822	0.005
$X_1(0)$	-	0.239	0.165	0.170
$X_2(0)$	-	0.709	6.530	0.029
R	-	-	0.002	0.109

Note:

WLS: Weighted Least Squares

ML: Maximum Likelihood

CV: Coefficient of Variation

Table 5.3 Parameter Estimates by Different Procedures.

Non Base Flow Interval.

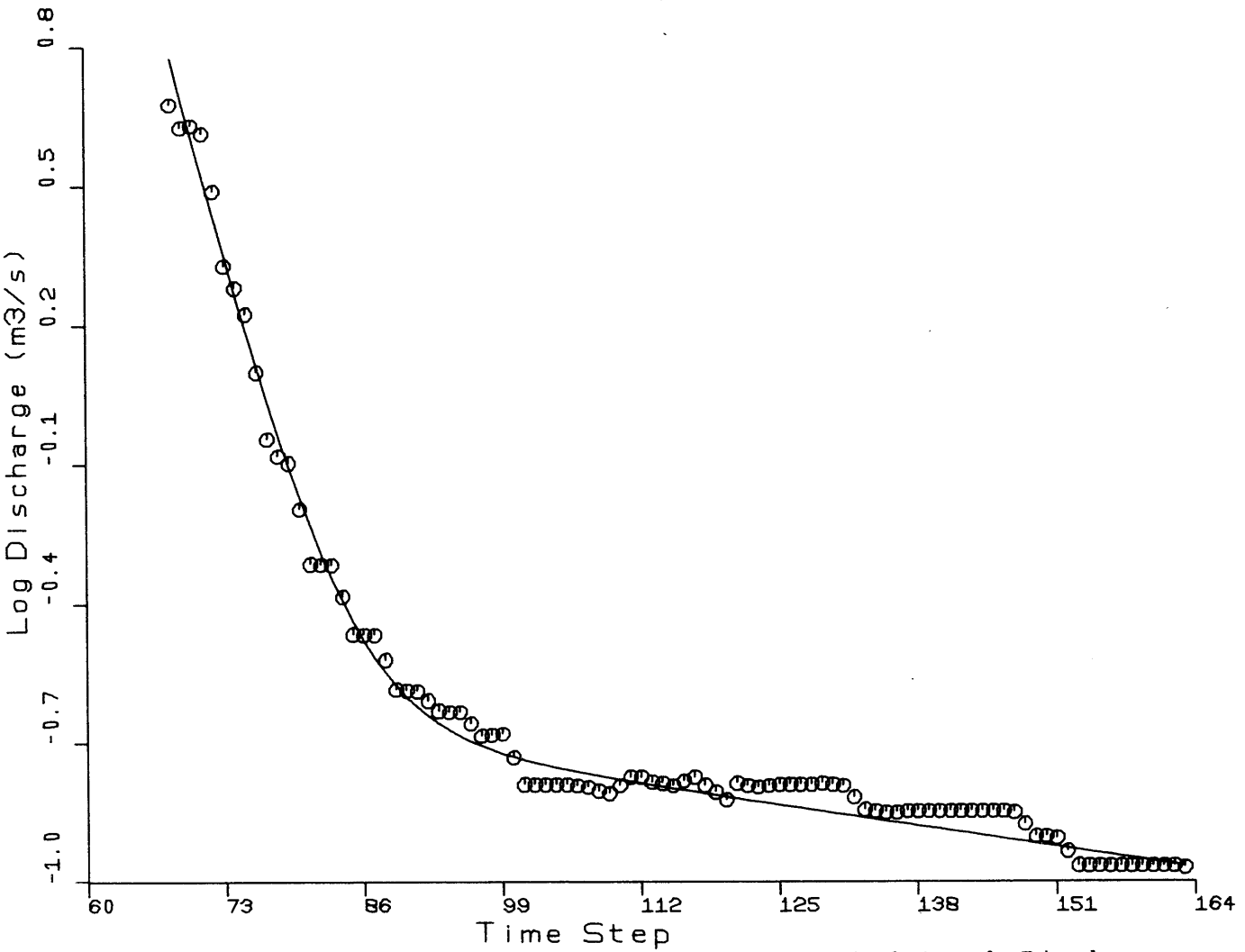


Figure 5.12 Six-Hour Lead Forecast of Bird Creek Discharges Using a Deterministic Base Flow Model With Parameters Estimated by WLS. Non-Base-Flow Interval.

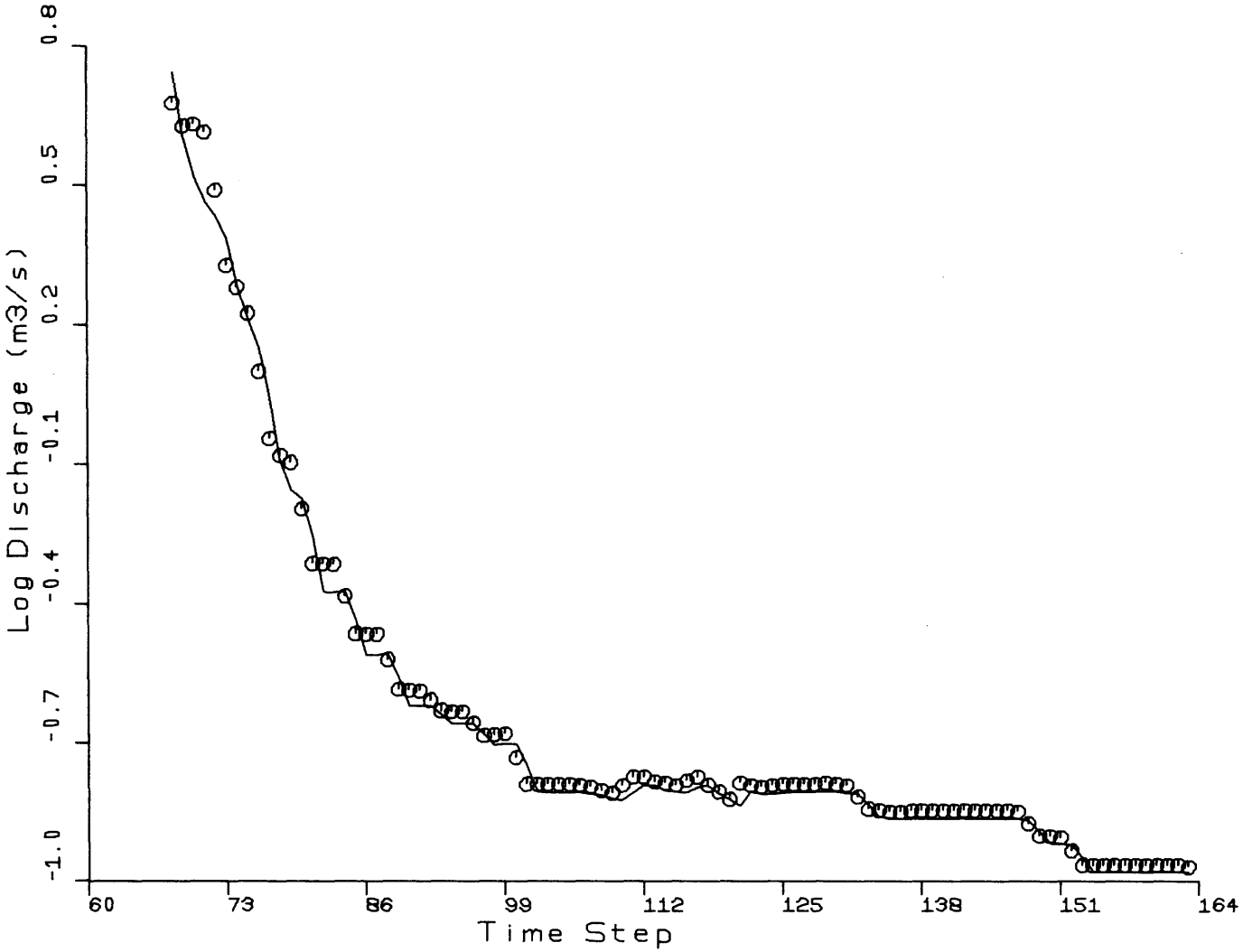


Figure 5.13 Six-Hour Lead Forecast of Bird Creek Discharges Using a Stochastic Base Flow Model With Parameters Estimated by ML. Non-Base-Flow Interval.

Correlation Matrix

	A_1	A_2	$x_1(0)$	$x_2(0)$	R
A_1	1.0				
A_2	-0.41	1.0			
$x_1(0)$	-0.31	-0.30	1.0		
$x_2(0)$	0.40	-0.67	-0.07	1.0	
R	-0.15	-0.01	-0.23	-0.08	1.0

Serial Correlation Coefficients of the Normalized Residuals

Lag	0	1	2	3	4
ρ	1.04	0.42	-0.31	-0.14	0.22
r	0.30	4.14	-3.05	-1.42	2.15

Sum of Squares Test

Number of Data Points:	96
Number of Parameters:	5
Expected Sum of Squares:	91
Standard Deviation:	13.5
Computed Sum of Squares:	99.1
Deviation (Units of S. Dev):	0.6

Table 5.4 Maximum Likelihood Post Optimality Analysis.

Non Base Flow Interval.

$$E[W_1^2(t)] = Q$$
$$E[v(t)] = 0$$
$$E[v^2(t)] = z^2(t)\ R^2$$
$$E[v(t)W_1(t)] = 0$$

where:

R: Unknown parameter
Q: Unknown parameter

This model will play a role in the identification of the interval of base flow activity to be discussed next.

5.4. Description of the Algorithm for Identification of the Time Interval of Base Flow Activity.

The identification of the time interval of base flow activity is done by means of an iterative procedure which examines the records through a "window" of varying length, L_w. The window has an initial length, I_{Lw}, defined by the user, which is increased by a number of D_{Lw} data points at every iteration. The window is set such that the last value in the window corresponds to the smallest discharge in a particular year. (See Figure 5.14). The rationale behind this is that if in a given year there is a time interval of base flow activity, the lowest discharge should be at, or close to, the end of that interval.

Two shortcomings of that approach are clear. First, it is possible that due to large errors in the discharge measurements, the smallest discharge in a year and hence the position of the end of the window, may fall in an area in which the base flow is not the dominant feature. Second, if more than one possible base flow intervals are present in a year the algorithm will identify only that one wich has the smallest discharge. This problem may be solved by dividing the data into several parts, each containing only one "suspected" interval of base flow.

At every iteration, the algorithm employs two stochastic models, M_1 and M_2, as defined in Sections 5.3 and 5.2 respectively. M_1 is the stochastic one linear reservoir model and M_2 is the stochastic model of base flow. The testing of two alternate models is required because towards the end of the base flow activity only the lower zone primary reservoir is active. The discharge then contains no information about the lower zone supplementary reservoir parameters, thus making imposible the successful identification of the supplementary reservoir's parameters. (The detailed description of the behavior of the model under the non-observability of the supplementary reservoir's parameters was explained in Section 5.2).

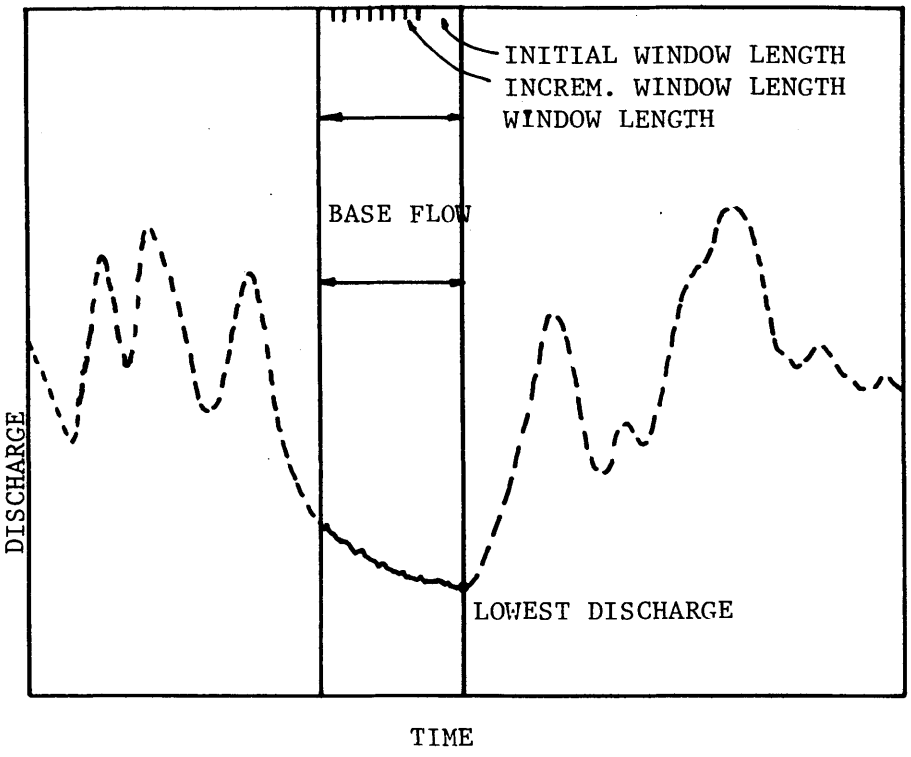


Figure 5.14 Schematic Diagram of the Varying-Size Window.

Therefore, in a particular iteration, the successful identification of the parameters of a one linear reservoir model combined with the unsuccessful identification of the parameters of the two reservoir model, indicates that the window is, at that particular iteration placed over the final part of the interval of base flow activity.

At every iteration, the models are tested according to three rules . There are three possible outcomes from these tests: 1) M_1 and M_2 are rejected; 2) M_1 is accepted and M_2 is rejected; and 3) M_1 is rejected, and M_2 is accepted. Upon examination of the results, once the maximum number of iterations has been completed, the time interval of base flow activity will be defined as the one corresponding to the largest window for which M_2 was accepted. A flow diagram of the algorithm is shown in Figure 5.15.

5.4.1. Tests for Model Acceptance or Rejection.

Test No. 1. is based on the positive definiteness of the information matrix. If a model has a non-positive definite information matrix the model is rejected. The conditions under which a non-positive definite Information matrix may arise were discussed in Section 5.2.

Test No. 2. is based in the evaluation of the likelihood of both models. It is used to decide between M_1 or M_2 in a given window, in case that both models have passsed test No. 1. Let ξ^*_i be the loglikelihood of model "i" at its optimum value, for i=1,2. Then the model rejected is the one with the smallest ξ^*_i.

The last test, No. 3, looks at the normalized serial correlation coefficients of the residuals, which were defined in Chapter 3. The normalized serial correlation coefficients of the residuals measure the number of standard deviations at which the estimated correlations are from their theoretical values. In this form, these statistics are an indicator of the performance of the stochastic model. To apply this test, the user specifies a threshold value for the normalized correlation coefficients, T_r, and a maximum number, N_r, of these coefficients that he would allow to go beyond that threshold without rejecting the model. As an illustration, assume that T_r = 4.0 units of standard deviation and N_r = 2 have been chosen, and that the following normalized serial correlation coefficients have been obtained:

	r(0)	r(1)	r(2)	r(3)	r(4)
No. 1	4.2	5.1	4.1	1.3	0.8
No. 2	2.1	1.2	1.5	0.9	0.7

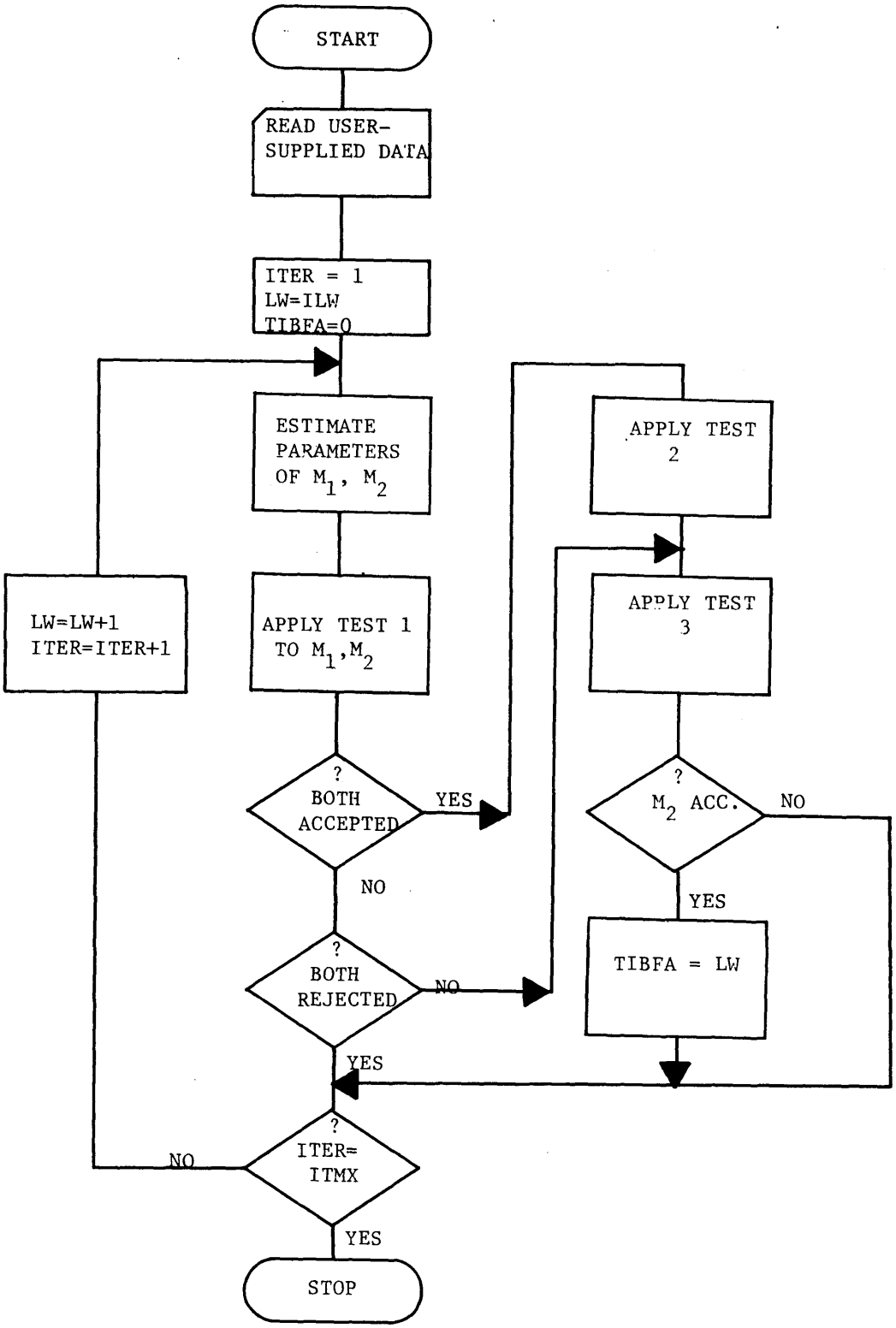


Figure 5.15 Flow Chart of the Algorithm for Identification of the Time Interval of Base Flow Activity.

In series No. 1, the model will be rejected, since three coefficients exceeded the threshold of 4.0, while the model corresponding to series No. 2 will be accepted.

5.5. Examples of Applications of the Algorithm.

5.5.1 Application to Bird Creek, near Sperry, Oklahoma.

As was already explained in Section 5.2.5.3 the data available for Bird Creek consisted of "Best Estimate Instantaneous Discharge", QINE, in the NWSRFS convention for series' denominations. That series is formed by the observed instantaneous discharges during the high flows, while the instantaneous discharges during the low flow periods are estimated from their average daily values. This causes very high lag-4 correlation coefficients. With this in mind, we decided not to include the lag-4 correlation coefficient in test No. 3. The threshold for rejection of the models in this test, T_r, was set at 1.5 standard deviations. A summary of the results of applying the proposed procedure to Bird Creek are presented in Table 5.5. The different variables that appear in that table are: Iteration No. (I_t), the model (M), the window's length at each iteration (L_w), the number of non positive eigenvalues in the Hessian (N_g), the loglikelihood function (ξ), the normalized correlation coefficients (r(0) through r(4)), the model accepted at the end of each iteration, and the final decision. To reduce the computational costs, the estimation of parameters for each interval is divided into two stages. In the first stage, a weighted least squares estimate of the model dynamics parameters and initial conditions is obtained. If all these parameters are in the feasible region, a maximum likelihood estimate of these parameters and the error terms are obtained in a second stage. Since the weighted least squares procedure uses a deterministic model, its use is considerably less expensive than trying to use maximum likelihood in a single stage. When some of the parameters converge to zero in the WLS stage, the maximum likelihood estimation is not performed and the values of $-\xi$ and r(0)...r(4) do not appear in Table 5.5.

The data chosen for the testing of the algorithm was the same data which was used for the validation of the stochastic model of base flow. The selected interval covered the months of July and August, 1957. A plot of the discharges was reproduced in Figure 5.7.

The analysis of the results in Table 5.5 show how the algorithm performs. From iteration 1 through iteration 8, M_2 is rejected by Test No. 1, while M_1 passes Tests Nos. 1 and 3. (Test No. 2 is applied only when both M_1 and M_2 pass Test No. 1). After iteration 9, both models pass Test No. 1.

Ic	M	Lw	ξ	r(0)	r(1)	r(2)	r(3)	r(4)	n_g	M^*
1	1	50	66.0	0.01	-0.09	-1.03	1.00	1.86	0	1
	2	50							2	
2	1	55	74.0	-0.03	-0.05	-1.08	0.99	1.80	0	1
	2	55							2	
3	1	60	80.4	0.05	0.11	-1.16	0.85	1.25	0	1
	2	60							2	
4	1	65	89.1	-0.02	0.08	-1.17	0.91	1.43	0	1
	2	65							2	
5	1	70	82.3	-0.04	0.39	-0.35	-0.11	0.71	0	1
	2	70							2	
6	1	75	82.5	-0.02	-0.45	-0.15	-0.12	2.63	0	1
	2	75							2	
7	1	80	89.5	-0.02	-0.48	-0.20	-0.19	2.79	0	1
	2	80							1	
8	1	85	94.1	-0.03	-0.46	-0.16	-0.06	3.15	0	1
	2	85							2	
9	1	90	95.3	-0.08	-0.28	-0.10	0.15	3.42	0	2
	2	90	101.9	-0.02	-0.48	-0.39	-0.18	3.14	0	
10	1	95	96.6	-0.04	-0.27	0.16	0.33	3.96	0	2
	2	95	108.1	-0.14	-0.66	-0.64	-0.28	3.27	0	
11	1	100	101.0	0.01	-0.31	0.36	0.57	4.17	0	2
	2	100	110.7	-0.11	-0.21	-0.29	-0.20	3.07	0	
12	1	105	102.3	-0.03	-0.27	0.59	0.69	4.29	0	2
	2	105	117.6	-0.07	-0.21	-0.31	-0.15	3.19	0	
13	1	110	105.3	0.03	-0.28	0.70	0.86	4.76	0	1
	2	110	123.0	-0.04	-0.34	-0.46	-0.16	3.37	1	
14	1	115	82.6	0.78	0.09	3.54	1.28	5.30	0	2
	2	115	118.0	-0.09	-0.17	0.35	0.00	3.41	0	
15	1	120	32.8	-0.81	0.25	2.36	0.65	1.93	1	2
	2	120	121.8	-0.06	-0.58	0.26	-0.04	3.70	0	
16	1	125	-77.5	-0.06	7.42	5.24	3.41	2.87	0	
	2	125	63.6	-0.22	3.13	2.97	1.37	3.14	0	
17	1	130	-170.2	0.05	9.99	7.73	5.37	3.20	0	
	2	130	-52.0	-0.02	4.86	2.08	0.66	-1.06	1	

Table 5.5 Summary of the Algorithm's Results for Bird Creek.

Test No. 2 indicates that M_1 should be rejected, since $-\xi_1 >- \xi_2$, and M_2 passed test No. 3. After iteration 16 both models are rejected. Notice however that at iteration 13 M_2 has a non positive definite Hessian and it is rejected. The analysis of the Hessian revealed that the negative eigenvalue is almost zero. This may indicate that the non possitive definiteness of the Hessian may be due to numerical inaccuracies rather than to convergency to a non optimal point. This conclusion is supported by the fact that the parameter estimates in iteration 13 are quite close to the estimates in iterations 12 and 14. Therefore, according to our definition, the time interval for base flow activity, is the one covered by iteration No. 15.

We have included in Figure 5.16 the values of the parameter estimates for different iterations. Notice that the model error term (Q_{11}) remains essentially constant for iterations 9 through 13, after which it increases by a factor of 2. Notice also that the standard deviations of the model dynamics' parameters decrease to a minimum at iteration 12. Moreover, the parameter estimates at iterations 11 through 13 were considerably closer to the NWS estimates of the same parameters than the maximum likelihood estimates at iterations 14 and 15. If we combine all these factors, we could establish an alternate definition of the interval of base flow as the interval before the model error starts to increase significantly. This behavior must be corroborated by analyzing different basins. However, in the application to the Cohocton River, presented in the next section, M_2 was accepted only at one iteration. Further research is required before an alternate definition is adopted.

5.5.2 Application to Cohocton River, near Campbell, NY.

This river is the only river avilable to us for which measured instantaneous discharge (QIN) is available for both high and low flows. Part of the river discharge is sometimes diverted for hydropower production. (The volumes of daily diversions are tabulated as USGS station 01528700). Since this diversion may seriously affect the estimation of the parameters of the base flow, we were restricted to work with a year for which there was no diversion during the low flows. Accordingly, a period covering the months of August, 1970 through October, 1970, was selected. The corresponding discharges are shown in Figure 5.17. A summary of the results at every iteration is shown in Table 5.6.

An analysis of the results in Table 5.6 shows that both models were rejected for the first 4 iterations. M_1 was accepted in iteration No. 5, and M_2 was accepted in iteration No. 6.

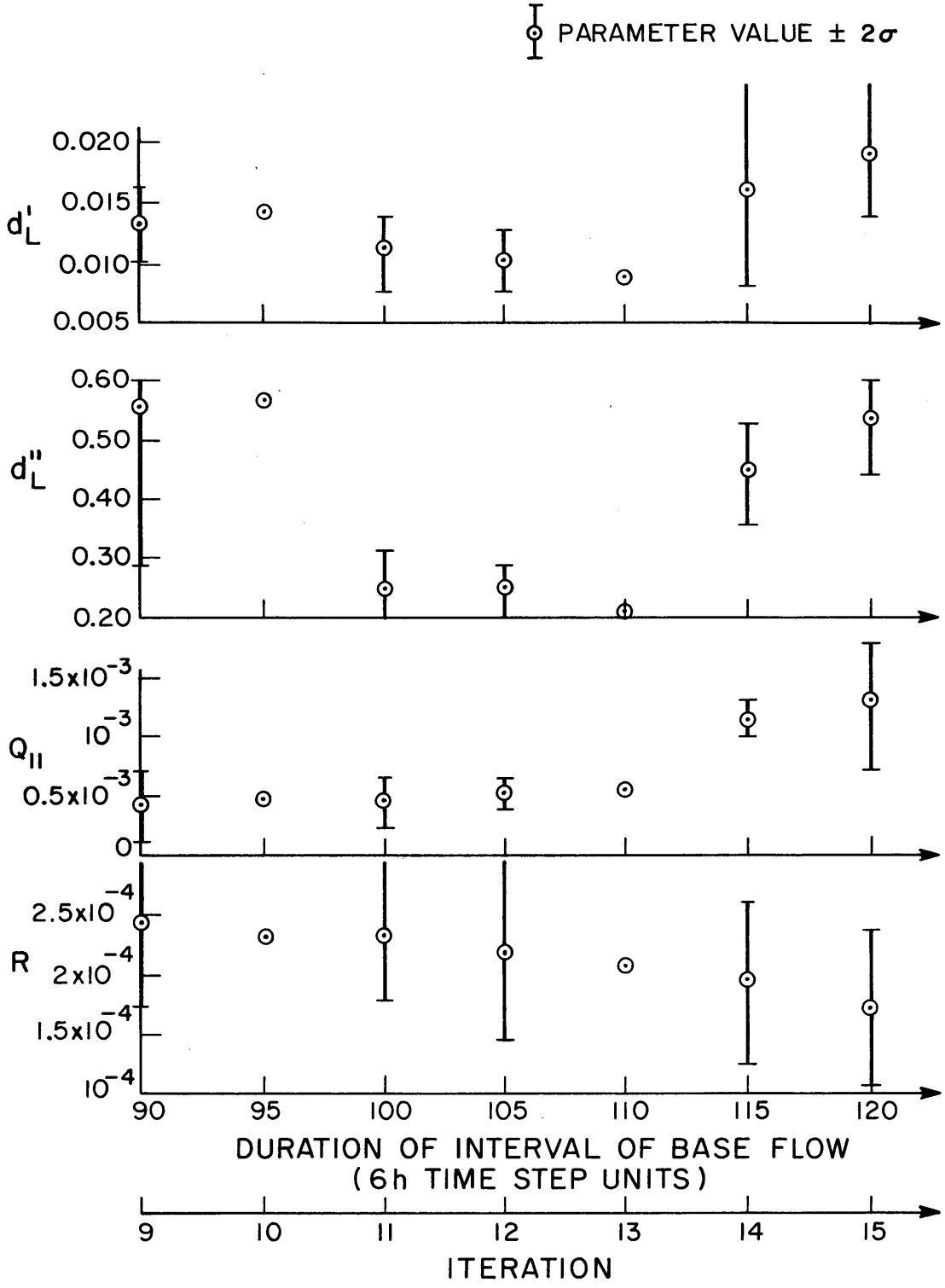


Figure 5.16 Base Flow Parameter Values for Several Iterations.

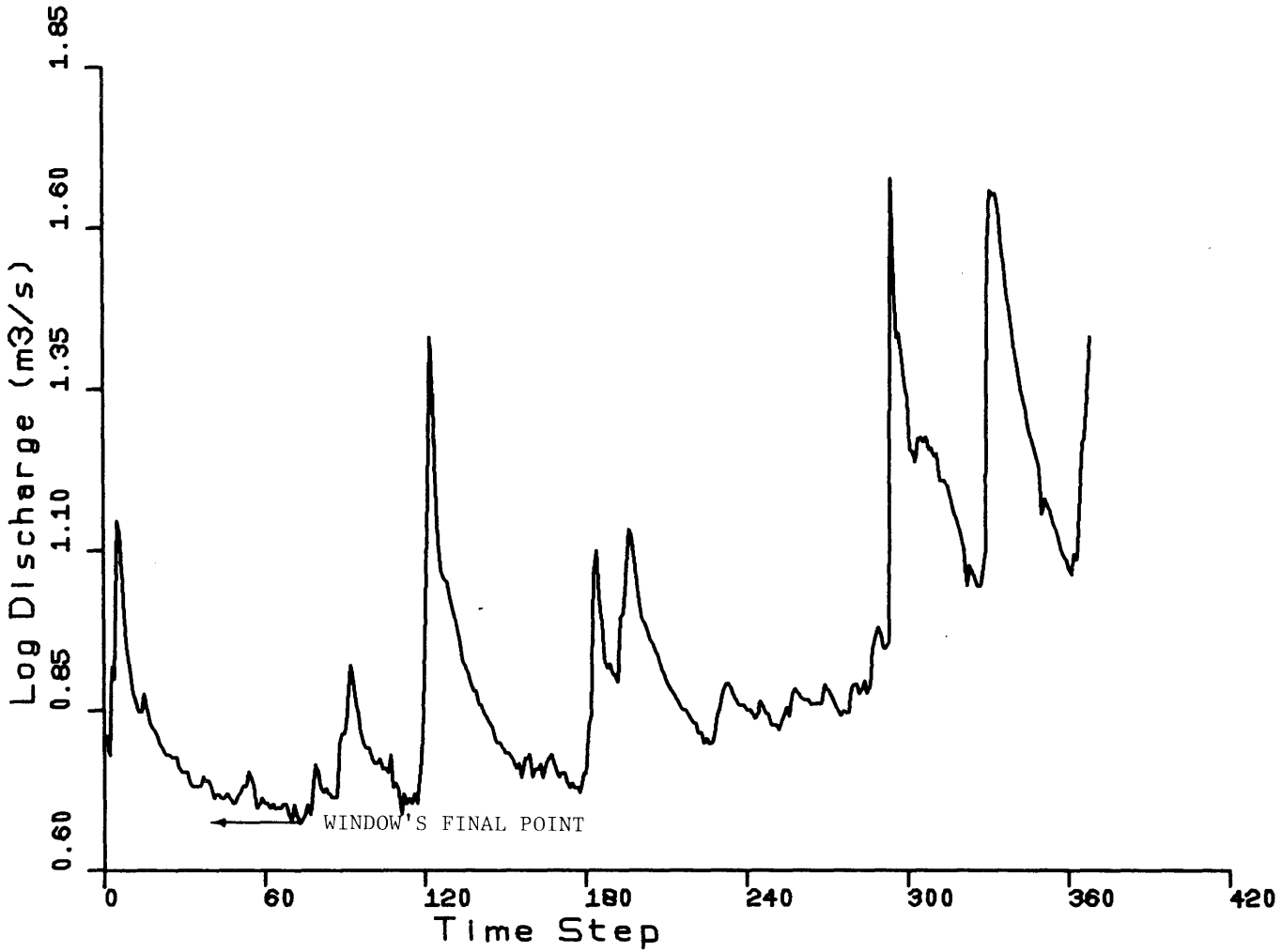


Figure 5.17 Semilogarithmic Plot of the Cohocton River Discharges for August through October, 1970.

Ic	M	Lw	ξ	r(0)	r(1)	r(2)	r(3)	r(4)	n_g	M^*
1	1	14	12.7	0.04	-0.23	-2.09	0.22	1.78	1	
	2	14	12.4	0.10	-0.21	-2.05	0.29	1.83	1	
2	1	24	12.5	-0.08	1.69	-0.50	-1.37	-0.45	1	
	2	24	12.0	-0.09	1.75	-0.40	-1.30	-0.45	1	
3	1	34	18.5	0.02	0.21	-1.65	-0.90	1.11	1	
	2	34							2	
4	1	44	19.2	-0.12	3.94	1.80	0.89	0.39	1	
	2	44							1	
5	1	54	34.0	-0.04	-0.15	-1.61	-0.48	0.49	0	1
	2	54	35.3	-0.01	-0.14	-1.40	-0.47	0.43	2	
6	1	64	29.4	0.06	1.35	-2.02	-0.77	1.17	0	2
	2	64	31.8	-0.05	1.50	-2.23	-1.42	0.22	0	

Table 5.6 Summary of the Algorithm's Results for the Cohocton River

The base flow activity corresponds to the period covered by iteration No. 6, according to the definition proposed earlier in this chapter.

5.6. Conclusions.

This chapter has presented an algorithm for automatically identifying the time interval of base flow activity. This algorithm was based on the maximum likelihood estimation of the parameters of the stochastic model of base flow, also presented in this chapter. Two applications of the algorithm were successfully carried out, which demonstrated the feasibility of this approach. The parameters determined from the application of this algorithm will be used to form the prior likelihood to be used in the global optimization program, according to the scheme presented in Chapter 3.

A discussion on the characteristics of the loglikelihood function was included and served to justify the choice of nominal values for the model error parameters.

The results showed that the visual, WLS and MLPE approaches yielded comparable parameters estimates. The advantage of MLPE is in providing independent measures of the quality of the parameter estimates.

CHAPTER 6

APPLICATION OF MAXIMUM LIKELIHOOD PARAMETER ESTIMATION TO THE SIMPLIFIED RAINFALL-RUNOFF MODEL

6.1. Introduction.

The implementation of the stochastic parameter estimation procedure, which was outlined in Chapter 1, included the development of a simplified rainfall-runoff model as an inexpensive way to research the features of estimating the parameters of conceptual rainfall-runoff models. This chapter is dedicated to the discussion of those results.

The results presented in this chapter are divided into two categories:

1. special features and problems of the estimation procedure; and
2. evaluation of the performance of the parameter estimation technique.

The identification of special features and problems of the estimation procedure must be done in a systematic way. This was done by dividing the model parameters into three classes and studying the features corresponding to each class of parameters and to combinations of several classes. These classes are:

1. parameters of the soil moisture accounting model.
2. initial conditions for the state variables.
3. model error and measurement error parameters.

Since this stage of the research was dedicated to explore the characteristics of maximum likelihood estimation in conceptual rainfall runoff models, it was important to perform the experiments, or test runs, under controlled conditions. This was achieved by generating a synthetic series of discharges from the precipitation and evapotranspiration input to a real catchment. In this form it was possible to study the performance of the estimation procedure under ideal conditions, since the real parameters of the model were known. Although a very large number of computer runs were made, it would be pointless to describe each one of these runs in detail since many of the results showed similar features. Instead, we are going to cover the major problems and features found in the application of the maximum likelihood procedure.

6.2 Problems Found.

6.2.1 Discontinuity in the Loglikelihood Function.

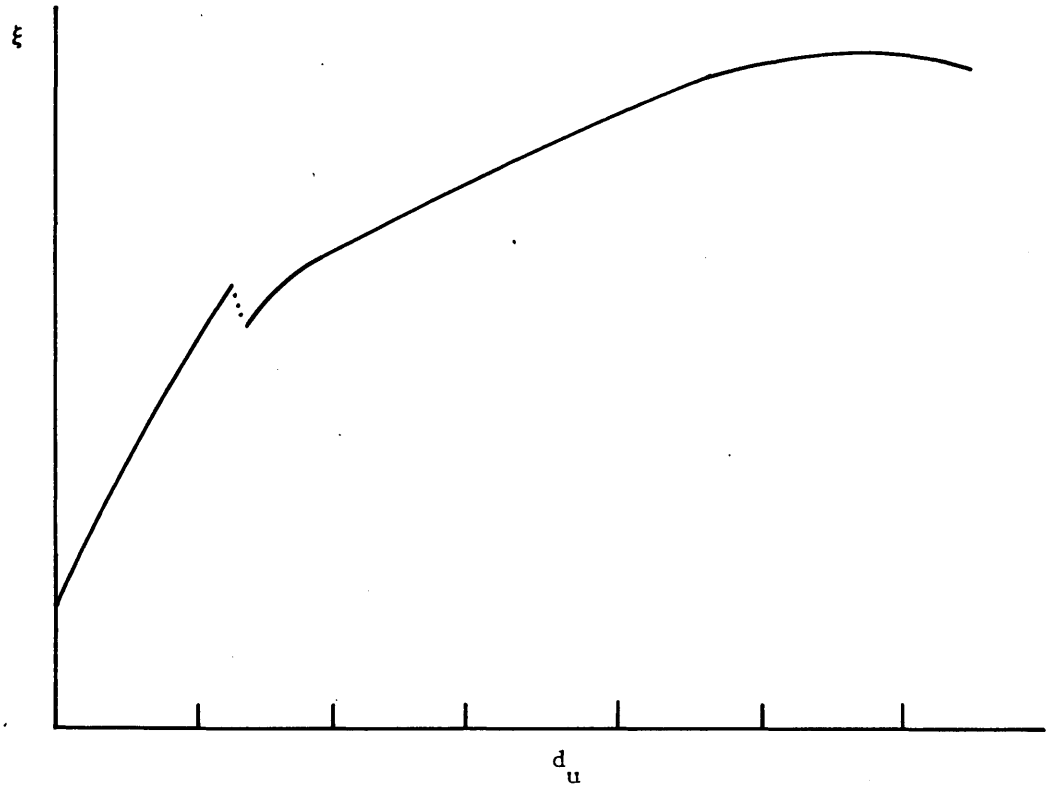
The builders of the model devised a variable time step integration scheme that insures high accuracy results. The integeration step is dependent on the amount of percolation

between the upper and the lower zones. For periods of time in which there is no percolation (or very little), and there is no surface runoff (controlled by the upper zone free water content, x_2), the time step is kept as 0.25 days, which is also the time between observations. When there is percolation the time interval is subdivided into NINC subintervals, each with a duration of DT=0.25/NINC days. NINC was arbitrarily chosen as the highest integer resulting from dividing x_2 plus the precipitation that becomes infiltrated by 5.00. This essentially limits major transfers of water to 5.0 mm or less during each time step.

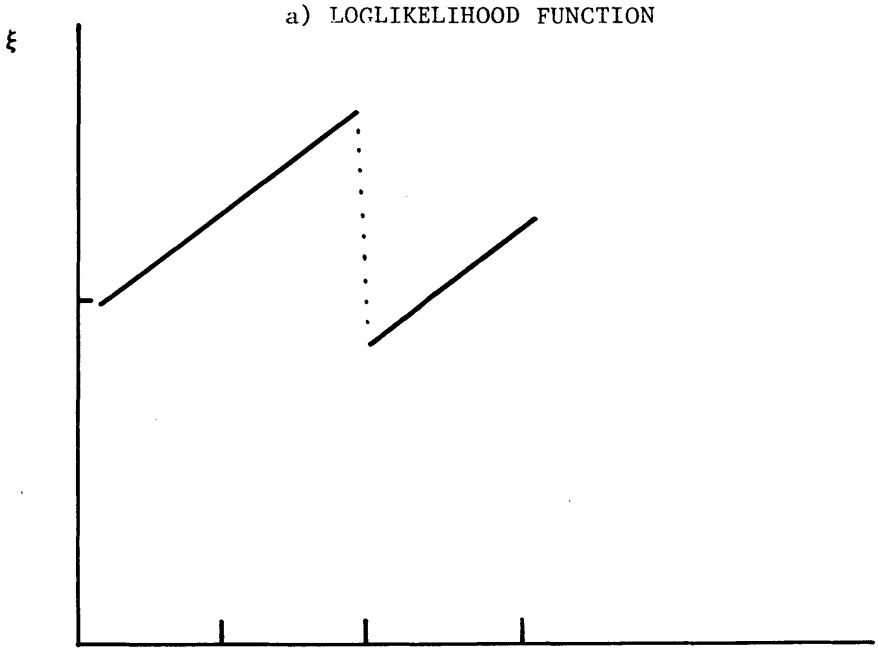
This scheme works very well in normal runs, without estimation of parameters. But when parameters are being estimated, the following situation may arise. Let us take for example the case of estimating d_u, the constant for the linear reservoir in the upper zone free water element, which determines the interflow. For some value of that parameter, θ_1, assume that at some time step x_2 is sligthly larger than 5.0 mm, and, for simplicity, there is no precipitation. This situation will result in NINC=2, which yields DT = 0.125 day. Assume the loglikelihood then takes the value ξ_1. Continuing with the linear search, a new parameter value θ_2 is tried. It may then happen that x_2 is now slightly less than 5.0 mm, leading to NINC=1, and DT = 0.25 days. With a larger time step, the integration is less accurate, the residual will be larger, and $\xi(\theta_2)<\xi(\theta_1)$. An actual occurence of this phenomenon is presented in Figure 6.1. A non linear optimization procedure, with a step size in the linear search not large enough to "jump" over the discontinuity, would converge to the peak of the discontinuity.

Although the example was done with one parameter, in fact several parameters affect the content of the upper zone free water element, thus causing ridge-like discontinuities in the loglikelihood function which may prevent the achievement of global optimality and will cause high correlations among the estimates of the parameters involved. An "artistic conception" of this situation is visualized in Figure 6.2. This figure shows a hypothetical loglikelihood function of any two parameters which affect the content of x_2. The global optimum, ξ^* may not be achieved if the search follows the dashed line in that figure. In this case, the discontinuity "traps" the search path and forces convergency to the point ξ', far from the global optimum ξ^*.

It is interesting to notice that the occurence of ridges in the objective function of parameter estimation procedures has been reported in the literature. In particular,



a) LOGLIKELIHOOD FUNCTION



b) DETAIL OF THE LOGLIKELIHOOD FUNCTION IN THE NEIGHBORHOOD OF THE DISCONTINUITY

Figure 6.1 Discontinuity in the Loglikelihood Function of d_u due to the Variable Time-Step Integration Scheme.

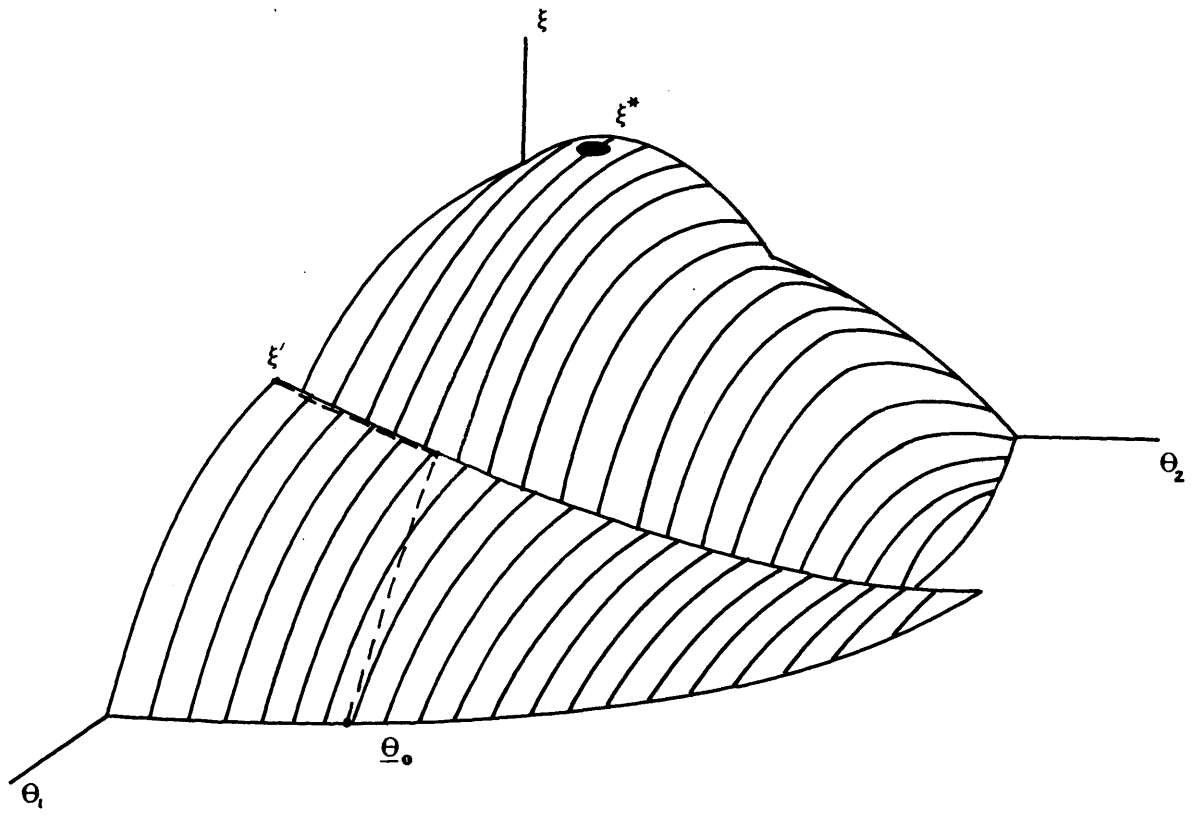


Figure 6.2 Conceptualization of a Ridge-like Discontinuity in the Loglikelihood Function of Two Parameters.

Soorooshian (1981), working with a small set of parameters of the NWSRFS model has reported the presence of these ridges when using maximum likelihood, as well as weighted least squares criteria. However, we can not claim that the variable time integration scheme is the only cause of the ridges. These ridges will also be caused by parameters' interdependency, as will be seen in the next section.

The solution to this problem is a rather simple one. We recognize the fact that, due to the non-linearities involved in the simulation of the soil moisture by the NWSRFS model when the percolation is active, a small time step is required to keep an accurate track of the process. On the other hand, the choice of a small, constant, time integration scheme would be wasteful in those periods in which the non-linearities are not so acute. Therefore, it is necessary to choose a scheme which allows for variation of the time increment according to the non-linearity of the process, yet independent of the parameters of the model. A very simple solution is to make NINC to be dependent on the precipitation in the last n_t hours, in which n_t should be specified by the user. In this form, we still have an indirect influence of the amount of percolating water on the selection of the time interval for integration.

6.2.2 Parameter Interaction.

In studying the characteristics of the joint identification of some of the soil system parameters and some of the initial conditions for the state variables we detected a case of parameters interaction, also known as parameters interdependency. The set of parameters that was used is shown in Table 6.1. That table gives the real value for the first three parameters. The real values for $x_1(0)$ and $x_2(0)$, which were synthetically generated, were not recorded. The same table shows the initial value of the parameters, and the value at which the parameter estimation procedure converged. Also shown is the value of the initial covariance matrix of the state variables, $P_{0_{ij}}$. The term δ_{ij} is the Kronecker's delta, and indicates that the initial covariance matrix was chosen as a diagonal matrix with the elements in the main diagonal equal to 10.0. This value corresponds to a coefficient of variation of the initial state variables of the order of 10%, which indicates a high certainty on the value of the initial state variables. This is important in the context of estimating the initial conditions of the state variables, as discussed in the previous chapter. The results can be qualified as very satisfactory, according to the closeness of the parameter estimates from their real values, and also according to the post optimality tests discussed in previous chapters which are shown in Table 6.2. The coefficients of variation calculated by the program, are, with the exception of the coefficient of

Param.	Real Value	Initial Value	Final Value	C.V. (%)
x_1^o	120.0	100.0	143.6	8.0
x_2^o	15.0	10.0	15.2	2.2
d_u	0.3	0.5	0.27	3.8
$x_1(0)$	-	50.0	90.1	8.6
$x_2(0)$	-	2.0	1.1	107.0

$$P_{Oij} = 10\ \delta_{ij}$$

Table 6.1 Maximum Likelihood Estimates for Five Parameters With Synthetic Data. Simplified Model.

Correlation Matrix

	x_1^o	x_2^o	d_u	$x_1(0)$	$x_2(0)$
x_1^o	1				
x_2^o	0.41	1	SYMMETRIC		
d_u	-0.13	-0.60	1		
$x_1(0)$	0.97	0.47	-0.20	1	
$x_2(0)$	-0.01	0	-0.01	-0.01	1

Serial Correlation Coefficients of the Normalized Residuals

Lag	0	1	2	3	4
ρ	1.01	-0.07	0.02	0.01	-0.02
r	0.08	-1.06	0.30	0.20	-0.35

Sum of Squares Test

Number of Data Points:	200
Number of Parameters:	5
Expected Sum of Squares:	195
Standard Deviation:	19.8
Computed Sum of Squares:	201.5
Deviation (Units of S. Dev):	0.33

Table 6.2 Post Optimality Results. Five Parameters, Synthetic Data.

variation of the estimate of $x_2(0)$, very low. The final estimate for x_1^o, however, is about 20% higher than the real value of the parameter. This contrasts sharply with the lower estimate of the error of estimation of 8% given by the coefficient of variation. This problem is related to parameters' interaction as will be discussed below. The high coefficient of variation of the estimate of $x_2(0)$ means that little information about this parameter was available in the data. This is due to the very fast varying nature of x_2, which makes the system quite insensitive to its initial condition.

The calculated sum of squares is only 0.33 standard deviations from its theoretical value, which indicates that the point of convergency is very close to the true optimum. The same conclusion is supported by the computed correlogram and the normalized correlogram.

The outstanding feature of the results of that table is the very high correlation (.97) between the estimates of x_1^o and $x_1(0)$, which are the "S" curve parameter for the upper zone tension water content and the initial condition for the same storage element. This high correlation implies that the information about both parameters is such that

$$x_1(0) \simeq k.x_1^o \qquad (6.1)$$

in which k is a proportionality constant. This interdependency between parameters causes a ridge in the loglikelihood function. This ridge "traps" the direction of search in the non-linear optimization process, which decreases the efficiency of gradient-based optimization procedures. We can visualize the location of this ridge in Figure (6.3). In that figure, the numbers on the line correspond to the iteration's number, and the position corresponds to the values the parameters had at the end of each linear search. After the sixth iteration all subsequent values of the parameters lie close to a straight line which is the location of the ridge.

The interaction between parameters not only adds cost to the optimization process but it is also an indication of overparameterization (one parameter is functionally dependent on the other), which prevents the simultaneous identification of both parameters (i. e., the product of both parameters, or the ratio between the parameters is identifiable, but not the individual parameters). It is important, then, to determine the cause of this interdependency. Several other runs were made to investigate the correlation between the above mentioned parameters. Although the detailed results will not be discussed here, it was shown that the correlation persisted even with 1): the inclusion of additional measurements, (T=400), and 2) starting from points far from the ridge.

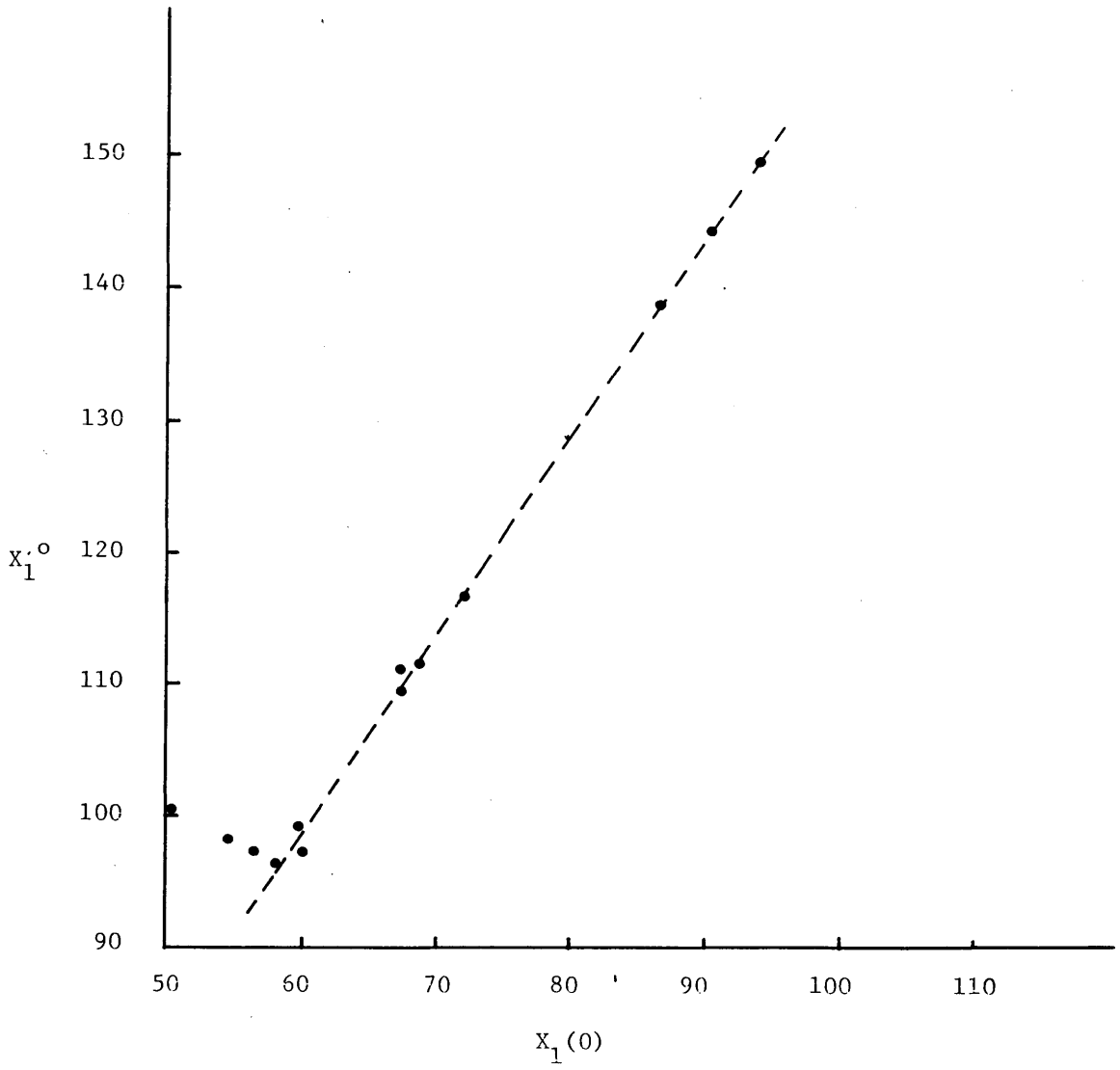


Figure 6.3 Alignement of the Values of the Parameters Along a Straight Ridge, for Several Iterations.

As mentioned earlier in this chapter, the testing of the parameter estimation procedure on the simplified model was done with synthetically generated streamflows. This made it possible to examine the state variables from which the streamflows were generated. This analysis revealed that the upper zone tension water content is a very slowly varying element. Figure 6.4 presents the S_1 curve, in which w_1 is the normalized water content. We can distinguish three zones in which the curve has been divided. These zones correspond to the dry, intermediate and wet regions. To understand the mechanics of the S_1 curve, we are going to examine, separately, the processes of filling and depleting. In filling: when $0<w_1<a$, the element is very dry, $S_1 \approx 0$, and most of the incoming water will be directed to the upper zone tension water element, which will make w_1 to approach rapidly the intermediate zone, B. For $w_1>b$, the opposite is true: $S_1 \approx 1$, and most of the incoming water will be diverted to the free water element, thus making w_1 to remain close to b. In drying, the evaporation part of the model is also controlled by the S curve. For $w_1>b$, $S_1 \approx 1$, and the evaporation will be high, thus making w_1 to approach the intermediate zone. For $a<w_1<b$ the value of S_1 decreases rapidly and so does the evaporation from that element. The net effect of both filling and drying is that w_1 tends to remain in the intermediate region. Figure 6.5 presents the variation of the synthetically generated x_1 vs. time. The two horizontal lines represent the limits of the corresponding intermediate region of the "S_1" curve, and the letters A, B, ands C identify the corresponding regions in Figure 6.4. The two points t_o and t_1 delimit the ends of the interval that was used for estimating the parameters of the simplified model. Studying that interval, ($t_o<t<t_1$), it is clear that the state variable at any time step, $x_1(t)$, is highly correlated with the initial condition for the state, $x_1(t_o)$ in that interval. In other words, letting

$$x_1(t) = x_1(t_o) \qquad t>0 \tag{6.2}$$

would not make much of a difference in the simulated discharges. Now, in the model, x_1^o and $x_1(t)$ appear only as the ratio $x_1(t)/x_1^o$ in the S_1 curve. By using Equation (6.2)

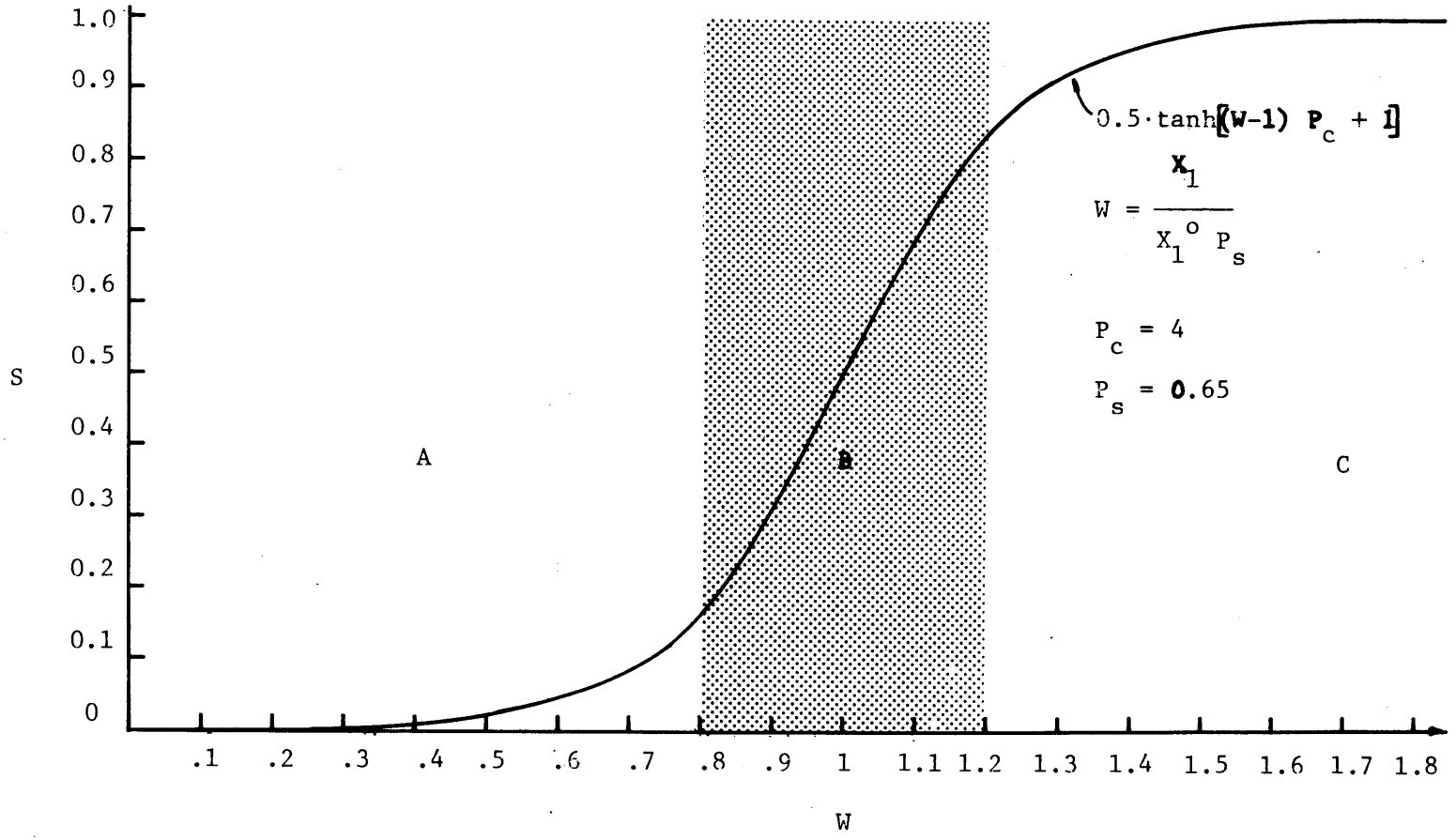


Figure 6.4 Values of the "S" Curve as a Function of the Normalized Water Content.

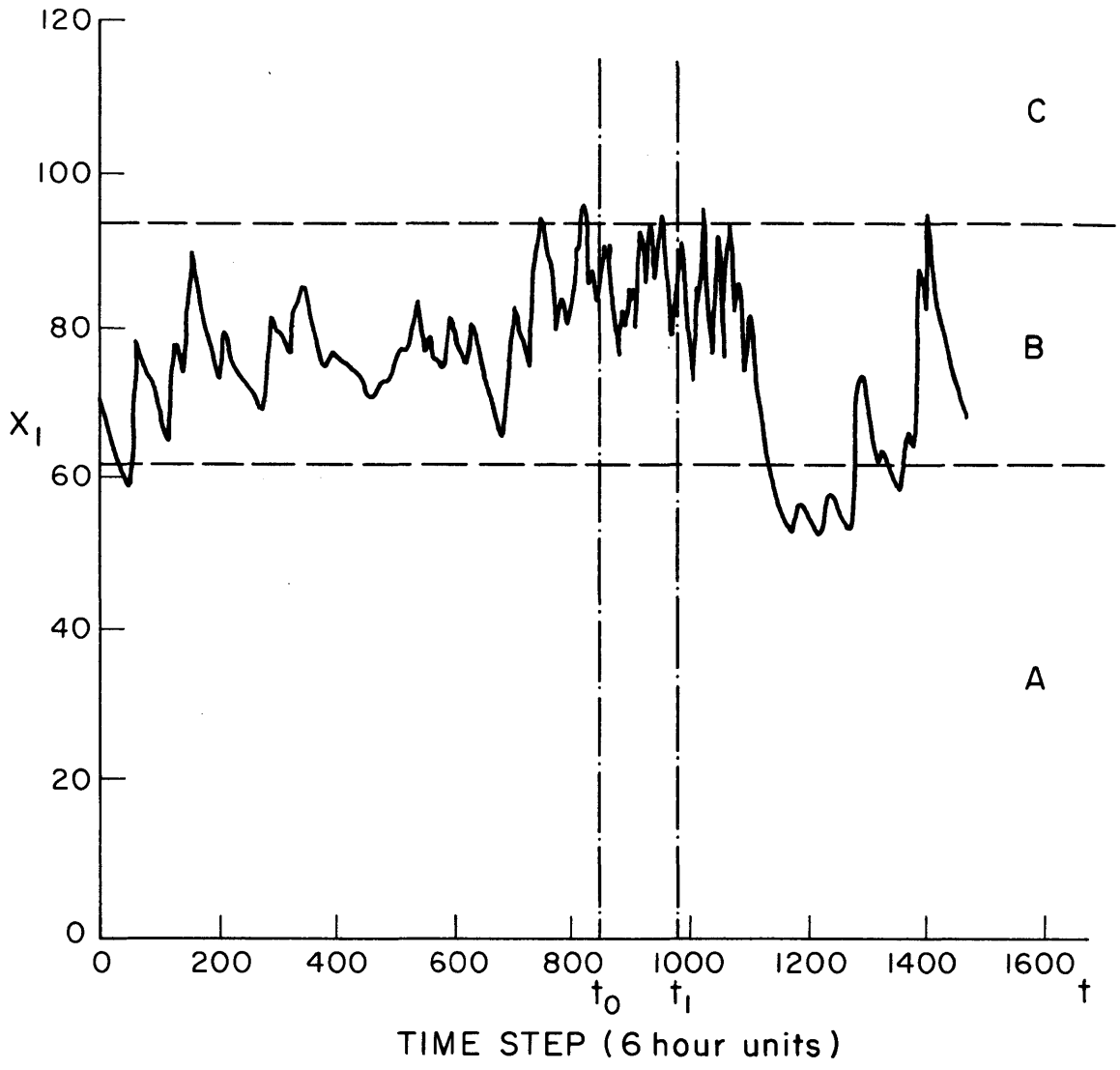


Figure 6.5 Time Variation of the Synthetically Generated x_1.

in the previous ratio it follows that the ratio $x_1(t_o)/x_1^o$, and not the individual parameters are identifiable. The value of that ratio corresponds to the constant k in equation (6.1).

Based on the above scenario, we assume that the use of the "S" curve, under the current formulation for both filling and drying of the upper zone tension water element is the probable cause of the correlation problem. The problem is particularly acute in the case of simulated discharges. In real life, however, the problem may be different since the upper zone of the soil (modeled by the upper zone tension and free water elements) does become dry. It follows that there is a defficiency of the model in capturing the behavior of the real counterpart. This structural error may lead to poor forecasting ability, and to biases in the estimates of the parameters.

The solution to the interaction problem could be a reparameterization of the model by eliminating a redundant parameter. Since in our case the two parameters involved belong to different classes of parameters (soil moisture and initial conditions), this reparameterization would mandate considerable changes in the structure of the model, such as changing the state variable No. 1 from accounting for absolute amount of water (mm) to relative water content (dimensionless). Since the problem was found to be caused by the S_1 curve, another solution may be the replacement of the S curves by different functions. The next chapter will present the results of estimating the parameters of the NWSRFS model. It will be shown that a different source of correlation between the same parameters was caused by the non observability of the upper zone tension water content. It is thus not immediate that a change of the "S" curves for different functions would solve the correlation problem.

6.2.3 Estimation of the Variances of the Model Error Terms

Identifying the variances of the model error terms is one of the most difficult remaining problems to solve. We showed already in Chapter 5 the problems of estimating the variances of the noise corresponding to the base flow elements. This section will address a multiple local optima situation

We will concentrate our attention on the characteristics of the loglikelihood functions of the variance of the error of the last reach of the non linear channel (Q_{77}). Figure 6.6 shows the loglikelihood function of that parameter, for values ranging from 0.1 to 100. This curve was generated using synthetic data, while setting all other parameters of the model at their real values. The presence of three local optima is a disturbing feature, for which there is not a definite explanation. For clarity, these local optima are numbered

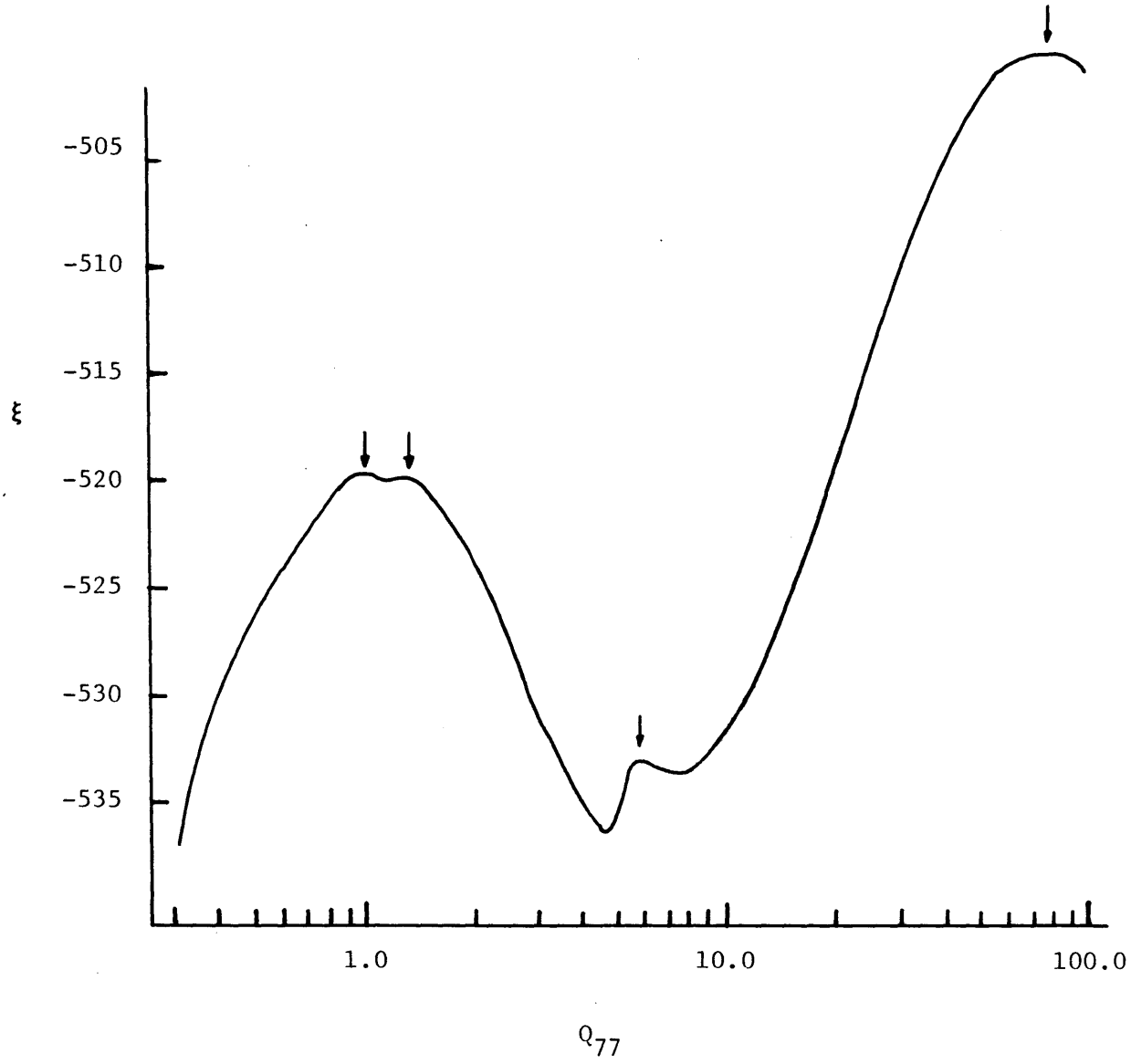


Figure 6.6 Multiple Optima in the Loglikelihood of Q_{77}.

in the figure. The global optimum, $Q_{77} = 75$ is close to the real value for this parameter ($Q_{77} = 100$). The possibility of convergence to the other optima is complicated by the fact that the curvature of the loglikelihood function at these points is very large (notice the logarithmic scale), which would yield a high information matrix and a very small bound for the estimation variance of that parameter. This small estimation variance misleadingly indicates a very high confidence on that estimate.

In an attempt to find a reason for this behavior we examined the components of the loglikelihood function. For simplicity we will eliminate the measurements $\underline{z}_T$ and the remaining parameters from the notation, and we will represent Q_{77} by θ.

In Chapter 3 we saw that the the loglikelihood function without prior information is formed by

$$2\xi(\theta) = -\xi_o(\theta) - \xi_b(\theta) \tag{6.3}$$

The first order necessary conditions for optimality state that

$$\left.\frac{d\xi}{d\theta}\right|_{\theta=\theta^*} = 0 \tag{6.4}$$

Using Equation (6.4) into Equation (6.3) we get

$$\left.\frac{d\xi_o}{d\theta}\right|_{\theta=\theta^*} = -\left.\frac{d\xi_b}{d\theta}\right|_{\theta=\theta^*} \tag{6.5}$$

In the case of convex functions Equation (6.5) will hold at only one point and we will have a single optimum. Figure 6.7 shows two local minima in the negative loglikelihood function which are due to the non convexity of the normalized sum of squares. The conditions expressed by equation (6.5) are fulfilled at $\theta=\theta_1$ and $\theta=\theta_2$. To check the convexity of the components of ξ we made high resolution plots of ξ_o and ξ_b. We found out that while ξ_b was a convex increasing function , ξ_o was non convex and has a discontinuity in its first derivative. A plot of the normalized sum of squares, (ξ_o), for several values of Q_{77}, is presented in Figure 6.8. The observed discontinuity in the derivative of the normalized sum of squares with respect to the parameter is the cause of the third local optimum in Figure 6.8. Since there is no explicit calculation in the model of that derivative, the only causes

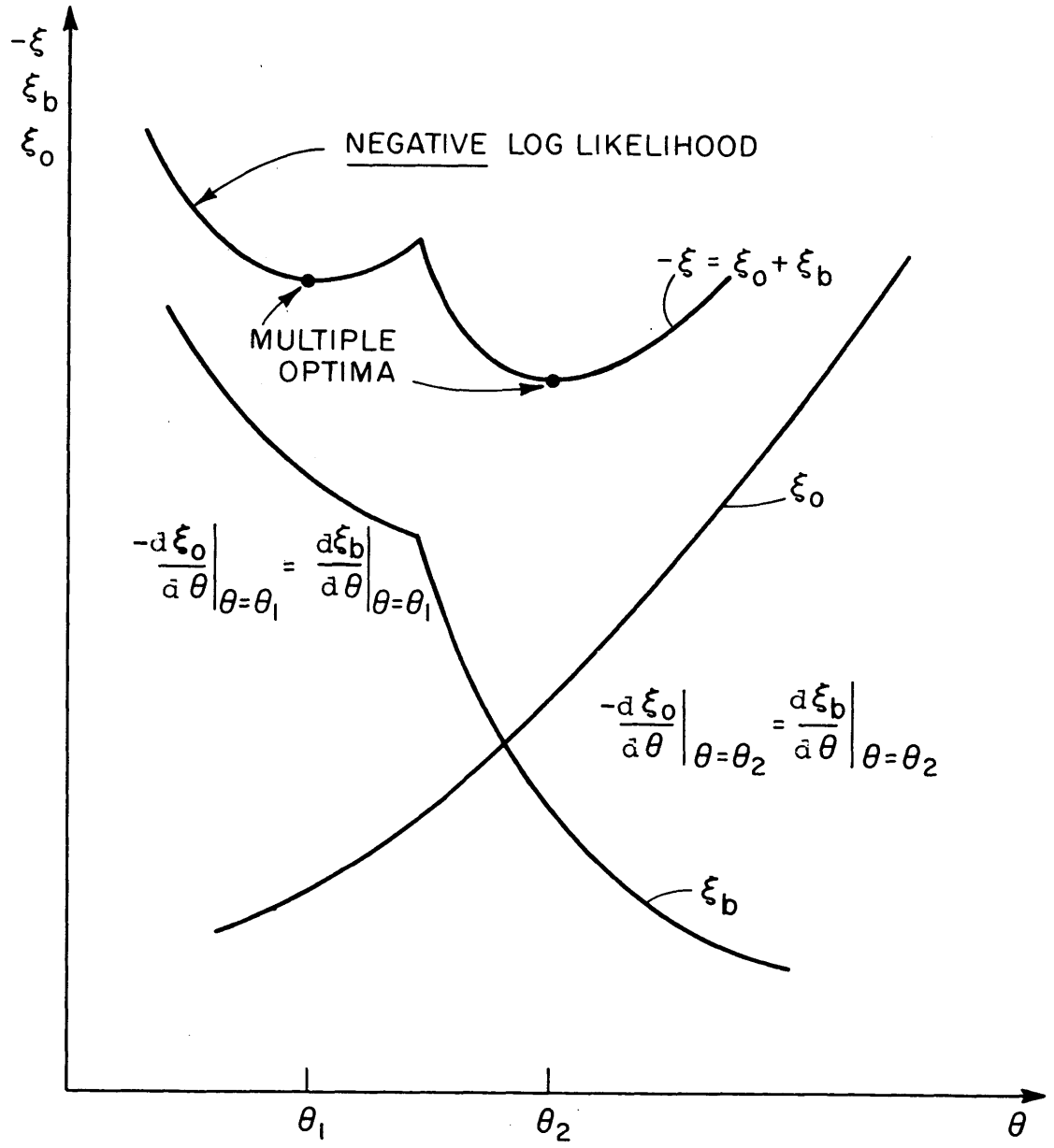


Figure 6.7 Multiple Optima in the Loglikelihood Function Due to Non-convexity in the Normalized Sum of Squares.

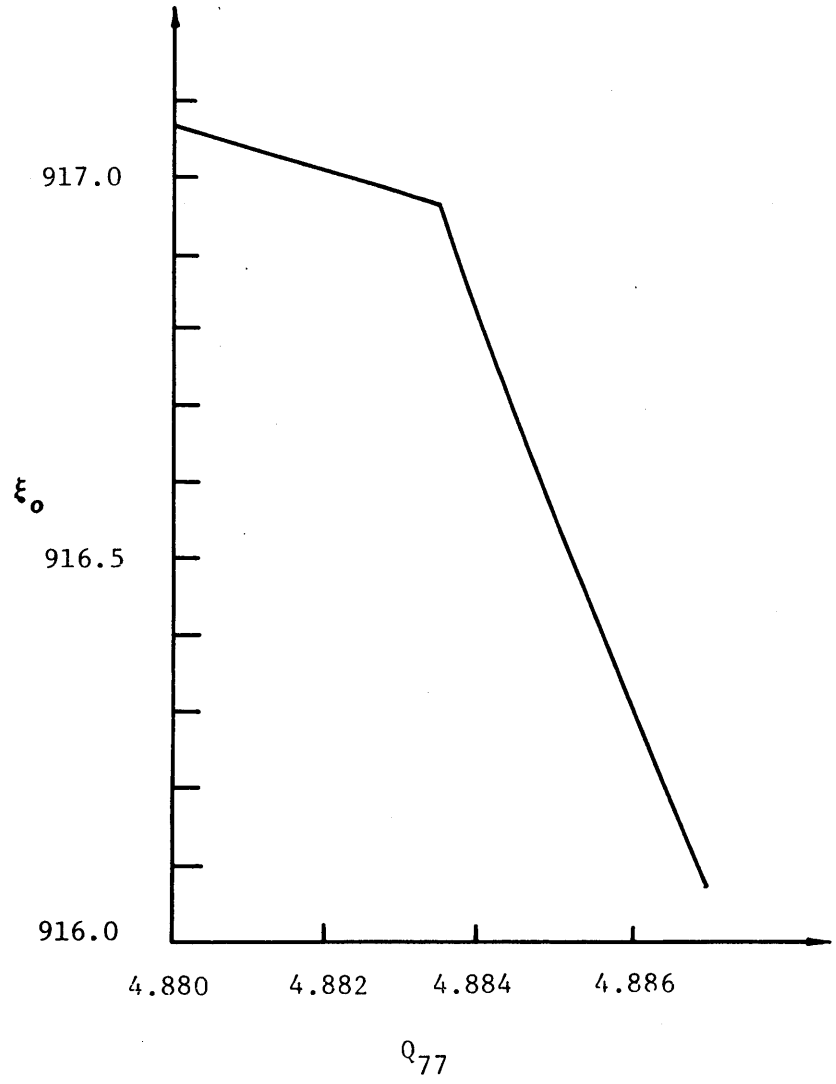


Figure 6.8 Discontinuity in $d\xi_o/d(Q_{77})$.

for its existance are either numerical inaccuracies or errors introduced by the Taylor's series expansion of the non linear channel. This last explanation is supported by the fact that the loglikelihood function of the linear channel of the NWSRFS model, to be presented in the next chapter exhibits a completely different behavior.

6.3. Results of the Joint Identification of the State Variables and Parameters of the Soil Moisture Accounting Model.

In the previous sections we presented three problems that were isolated during the first stage of the research, two of which were left unsolved, waiting for additional research to be done during the last stage of the work, which is covered in Chapter 7. The first of these two problems was the high correlation between the estimates of x_1^o and $x_1(0)$, and the second was the the occurence of multiple optima for the likelihood of the model error term. The final experiment performed with the simplified model consisted of the joint estimation of the soil system parameters and three initial conditions. Since the model has seven state variables, there were four initial conditions that were not estimated: $x_1(0)$ (due to the correlation problem), $x_5(0)$ and the initial content of the two channel reaches (due to the fact that the very fast varying nature of the corresponding variables makes the loglikelihood insensitive to these parameters). In performing the experiment, the parameters and initial conditions which were not estimated were set at their real values, those used in the generation of the synthetic data. The initial values, real values, estimated values and coefficients of variation of the parameters which were estimated are shown on Table 6.3

The estimates of the parameters, judged objectively by the coefficient of variation of the lower bound of the estimates, and subjectively by the closeness to their real values, can be judged as good. The calculated sum of squares of the normalized residuals is, however, more than 2 standard deviations from its expected value, which indicates that the parameters estimates can still be improved. This may be due to the fact that not all the parameters of the model were estimated, therefore they may not be at their maximum likelihood point. An important observation concerning the accuracy of the lower bound in the variance of estimation of the parameters must be made here. The real value for d_1' was 0.0113, whereas the value estimated by the maximum likelihood procedure was 0.029, which is 156% in error, compared with the real value. The coefficient of variation calculated from the lower bound in the variance of estimation yielded a very optimistic 2.0%. This clearly indicates that the coefficient of variation of the estimates should be looked at carefully before drawing conclusions regarding the quality of the estimates.

Param.	Real Value	Initial Value	Final Value	C.V. (%)
a_2	0.001	0.002	0.0015	7.1
x_1^o	120.0	100.0	127.2	3.6
x_2^o	15.0	10.0	13.7	3.4
x_3^o	160.0	120.0	126.2	3.4
x_4^o	140.0	120.0	119.2	1.2
x_5^o	14.0	10.0	13.0	3.3
d_u	0.3	0.2	0.32	5.0
d_1'	0.011	0.02	0.03	2.0
d_1''	0.126	0.2	0.149	3.6
γ	48.0	20.0	29.8	3.0
μ	3.55	3.0	4.32	2.9
$x_2(0)$		5.0	6.39	4.6
$x_3(0)$		90.0	81.2	7.8
$x_4(0)$		10.0	11.1	10.0

Serial Correlation Coefficients of the Normalized Residuals

Lag	0	1	2	3	4
ρ	0.70	-0.06	0.04	-0.03	0.01
r	-2.95	-0.89	0.51	-0.39	0.16

Sum of Squares Test

Number of Data Points:	200
Number of Parameters:	14
Expected Sum of Squares:	186
Standard Deviation:	19.3
Computed Sum of Squares:	140.9
Deviation (Units of S. Dev):	2.3

Table 6.3 Parameter Estimates and Post Optimality Results for 14 Parameters, Synthetic Data. Simplified Model.

6.4. Conclusions.

This chapter presented the characteristics of the estimation of parameters for the simplified rainfall runoff model. It was shown that one of the causes of ridges in the objective functions of parameter estimation procedures in rainfall runoff models is due to a variable time-step integration scheme. When estimating some of the soil parameters jointly with the initial conditions for the state variables, a large correlation between two of the parameters was detected. The reason for the problem was traced to the use of the continuous approximation to the threshold elements which is known in this report as the S curve. We also showed the presence of multiple optima with respect to one of the parameters of the model error terms. Finally, we presented the results obtained with the simultaneous estimation of 14 parameters. From these results, it is clear that theoretical errors resulting from the maximum likelihood technique, being only lower bounds, may tend to be unrealistically low.

CHAPTER 7

APPLICATION OF MAXIMUM LIKELIHOOD PARAMETER ESTIMATION TO THE NWSRFS MODEL

7.1. Introduction.

In the previous chapter we examined the performance of the maximum likelihood parameter estimation procedure under ideal conditions, since the discharge data used had been synthetically generated from real precipitation and evapotranspiration demand records. In this form, we were certain to comply with two hard-to-meet assumptions of maximum likelihood estimation. First, there is a set of parameters for which the model fits the data. Second, the system and measuring errors were pure white gaussian noise processes. When applying the maximum likelihood estimation criterion to estimate the parameters of a model representing a real basin the two conditions above are not strictly met. The one-dimensional, lumped parameter conceptualization of the physical process of water movement in the soil certainly violates the first assumption, and, consequently, there is no "real" set of parameters with which we can compare our estimates. The second assumption is also violated very commonly, since pure white gaussian processes do not exist in nature. We can at best hope for a reasonable fascimile.

We can see that while the use of synthetic data is an important tool in finding some of the features of a parameter estimation procedure, the real test of such procedure is given by its use with real data. The object of this chapter is to describe the characteristics of the maximum likelihood parameter estimation procedure when applied to two basins in the United States. The basins are the only ones for which instantaneous discharges measured every six hours were available. These are the same basins in which the model of base flow was tested (Chapter 5).

This chapter is divided into four major sections. Section 7.1 contains the Introduction. Section 7.2 describes the application of the procedure to estimate the parameters of Bird Creek. Section 7.3 covers the application of the maximum likelihood parameter estimation procedure to estimate the parameters of the Cohocton River. The conclusions are presented in Section 7.4.

An important factor in parameter estimation is the cost of the algorithm. The rest of the present section will give some figures reflecting the cpu usage of the model as implemented. The procedure was programmed in Fortran IV, on a Honeywell Level 68/DPS computer, running under the Multics operating system. In this environment, the computer takes about 22 seconds of cpu to simulate <u>one month</u> of six-hourly discharges. This is equivalent to approximately 7 seconds of cpu on an IBM 370/168, running under os/vs1 as the operating

system. As was seen in Chapter 3, the non linear optimization algorithm which was implemented requires the calculation of the gradient of the loglikelihood function, to determine the direction of search. A calculation of the gradient by central differences takes $2n_\theta$ computations of the loglikelihood function, where n_θ is the number of parameters being estimated. Since each computation of one value of the loglikelihood function requires the performing of a simulation run, it follows that the total cost of calculating the gradient is about $44n_\theta$ cpu seconds per parameter on the Honeywell 68 computer, or about $14n_\theta$ seconds on the IBM 370/168 machine. For a simultaneous estimation of 20 parameters, this would be translated into 880 secs. of cpu per month of data on the Honeywell, or 280 secs. of cpu on the IBM. This cost certainly places financial constraints on the extent of the data we will examine. We decided to initially test the procedure with one month of 6-hourly precipitation and discharge values, and to use the information on the estimation error to decide if additional data is required.

7.2 Maximum Likelihood Parameter Estimation of the NWSRFS Model in the Bird Creek Basin.

7.2.1 Introduction and Data Selection.

The characteristics of Bird Creek's records were described in Chapter 5, where we estimated the parameters of the base flow. In that chapter, we mentioned that these records are considered of good quality for high flows, while the instantaneous discharges for low flows are actually reconstructed from the average daily flow records. This data manipulation must be kept in mind when evaluating the results of this section. In addition to the quality of the records, the cost of performing the parameter estimation runs is another factor in determining the length of the interval of the records to be used in the estimation. Since the cost criterion expressed in Section 7.1 restricted us in principle to use only one month of data, we had to choose a month for which we believed that all the elements of the model, and hence its parameters, were going to be active. This lead us to select a month presenting very high discharges. Simulation runs performed with the deterministic model and manually calibrated parameters indicated that the year of 1957 was exceptionally well simulated by the deterministic model. In particular, the month of April, 1957 also fulfills the requirement of having high flows (Figure 7.1). Therefore, this month was initially selected for the experiments.

We shall compare our results with the parameters calibrated by the National Weather Service Hydrologic Research Laboratory (NWSHRL) staff, but three warnings should be given in advance. First, the model for which these parameters were

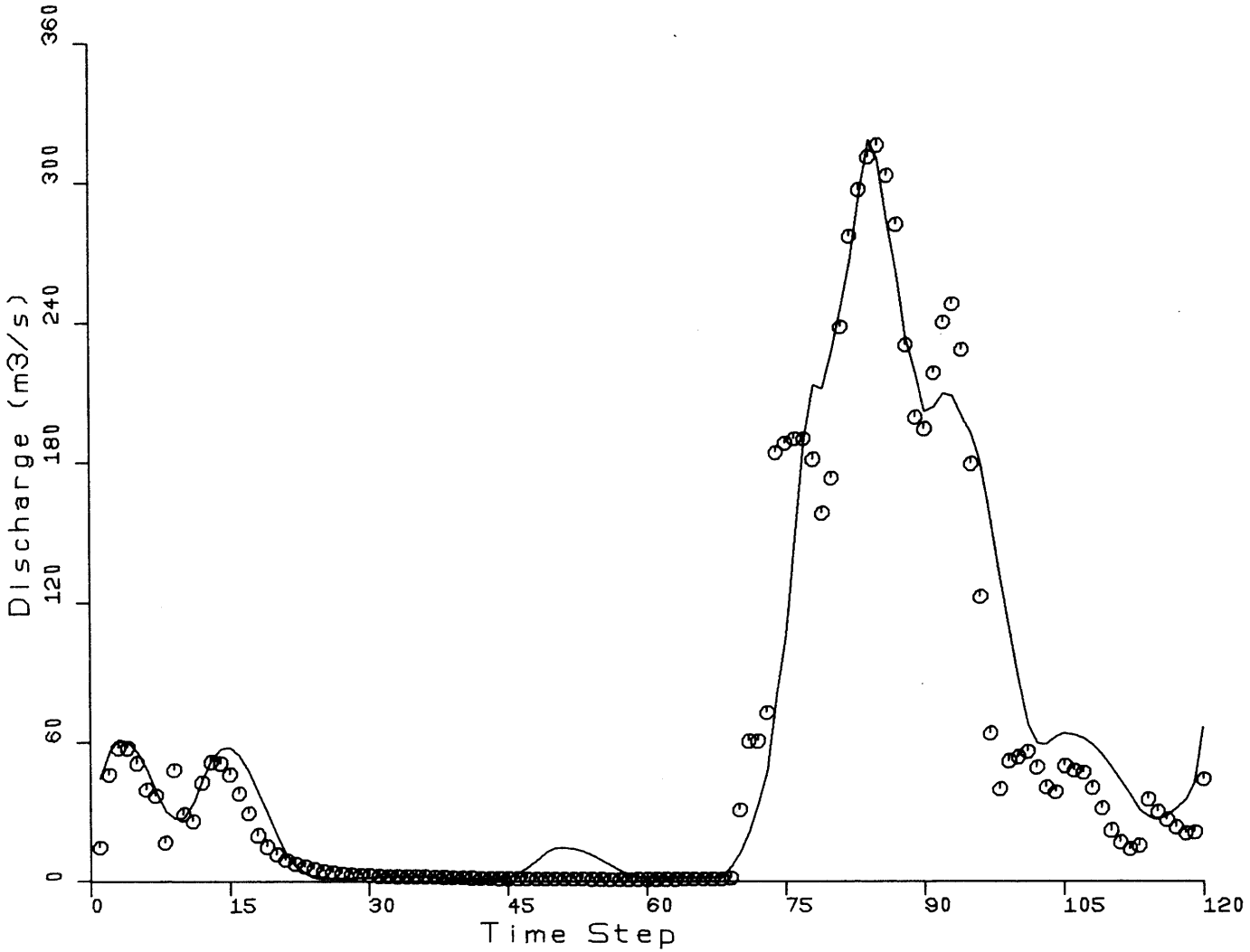


Figure 7.1 Six-Hour Lead Forecasts of Bird Creek Discharges by the Deterministic NWSRFS model with NWS-estimated Parameters, for April 1957.

calibrated was a deterministic model. This implies that the parameters had the sole responsibility of accounting for the errors between the forecasted and the measured flows. Second, the deterministic part of the stochastic model has some differences with the original NWSRFS model. These differences were already described in Chapter 4 of this work. Third, there is no set of "real" parameters. A good manual calibration is just a reasonable alternative. This combination of factors alone will almost guarantee that some of the maximum likelihood estimated parameters will converge to different values, even if the same data are used for the calibration.

7.2.2 Parameter Estimation Without Prior Information.

The first results we will present are those corresponding to the estimation of parameters without the use of prior information about. In other words, we will carry out the estimation of parameters by a standard maximum likelihood procedure. Table 7.1 presents a list of the parameters that were estimated. The second column presents the value of the parameters at the beginning of the estimation procedure. This set of values corresponds to the values which the staff of the National Weather Service Hydrologic Research Laboratory had previously calibrated for the same basin. It was felt that, given the quality of the predictions in the off line simulation (Figure 7.1), the initial value of the parameters should be reasonably close to the maximum likelihood estimate of the same parameters. This should result in faster convergency to the optimum than when starting from arbitrary values. Also shown in Table 7.1 are the maximum likelihood estimates of the parameters and the lower bound of the coefficient of variation of these estimates, obtained with one month of data and with four months of data. The reader may notice that the initial conditions of the state variables are not among the estimated parameters. Although this omission aparently contradicts our previous assessment of the need of estimating the initial conditions for the state variables jointly with the remaining parameters, we chose a different approach because the quality of the predictions of the off line run for the calibration month indicated to us not only that the parameters should be relatively close to their maximum likelihood estimates, but also that the initial conditions for the state variables should also be close to their real values. We decided to decrease the number of parameters to be estimated (and the cost), by setting those initial conditions at the values given by the off line simulation.

Analizing the results of column (3) in Table 7.1 we see that for five of the parameters the final values are given as N. F. (Which stands for Non Feasible). This means that these parameters converged to zero or tried to become negative. These parameters are: a_2, $d_1"$, d_1', P_f, μ. Since the month chosen for the estimation contains high discharges, it

Param. (1)	NWS Estim. (2)	April Estim. (3)	1957 C.V. (4)	April-July Estim. (5)	1957 C.V. (6)
x_1°	120.0	109.8	0.02	122.5	0.09
x_2°	15.0	14.5	13.3	14.7	0.13
d_u	0.3	6.36	0.21	0.75	0.20
a_2	0.001	NF	-	0.013	0.68
a_1	0.17	0.79	0.05	0.14	0.31
γ	48.0	48.0	7.33	38.2	1.09
α	2.10	2.37	25.47	1.09	0.66
x_3°	160.0	159.2	0.17	160.0	0.51
x_5°	14.0	14.0	23.18	12.5	2.40
x_4°	140.0	140.0	27.63	140.0	0.61
d_1''	0.126	NF	-	0.08	1.71
d_1'	0.013	NF	-	0.011	1.16
P_f	0.02	NF	-	0.02	25.21
μ	3.5	3.57	-	5.17	0.83
R		0.17	0.17	0.25	0.12
Q_{11}		1.25	0.15	1.33	0.14
Q_{22}		0.0018	1.45		

NF: Non Feasible.

Table 7.1 Maximum Likelihood Parameter Estimates Without Prior Information for the Bird Creek Basin.

is not surprising tht the two coefficients of the base flow, d_1' and d_1'' attained non feasible values. Base flow, which is the most direct source of information about those parameters, is a very small component of the discharge. While the total measured discharge reached a peak value of 311 m^3/s, the base flow is only of the order of $1m^3/s$. More surprising perhaps is the fact that a_2 also attained a non feasible value. This parameter acts only when there is precipitation, and we would expect that, in the absence of structural errors, the high discharges on a humid month would contain a large amount of information about this parameter. The parameters P_f and μ actually converged to zero. Strictly speaking this is a feasible value for these parameters. They were signaled as N. F. because the convergence to zero may be related to the values at which other parameters converged, as we explain in the next paragraph.

From the remaining parameters, only d_u and a_1 converged to a value significantly different from the initial one. The rest of parameters converged to values identical or close to their initial value. Two explanations are possible. First, the model is insensitive to these parameters, or, second, the initial value did correspond to the maximum likelihood estimate of these parameters. An answer to that question could be given by the coefficient of variation of the parameter estimates, column (4). A very high coefficient of variation means that the confidence on the parameter estimate is very small which implies that the data contained little or no information about these parameters, and, consequently, the model is insensitive to these parameters. If the coefficient of variation is very small, the values of these parameters at the point of convergency do correspond to their maximum likelihood estimates. We will be discussing this further. The final values for d_u and a_1 may be due to the model trying to compensate for the non feasibility of a_2. Both d_u and a_1 play an active role during high moisture conditions. The values to which these parameters converged may indicate that the model is trying to make up for the loss of the additional direct runoff, which is controlled by a_2, which converged to zero. Both parameters converged to very high values when compared with the manually calibrated ones. These high values, drastically reduce the role played by the lower zone elements. The high value for a_1 decreases dramatically the fraction of water that becomes infiltrated. Of the fraction that becomes infiltrated, a large part will appear as interflow, due to the high value of d_u, thus largely decreasing the amount of water that percolates to the lower zone. Under these circumstances

the only active parts of the model are the upper zone elements and the channel. This caused the lower zone coefficients of discharge of the linear reservoirs, d_1' and d_1'' to be driven to zero since these parameters are not playing any role. In the absence of discharge from the lower zone, the other lower zone parameters become unidentifiable, which means that these parameters can converge to any value without changing the value of the loglikelihood. This caused μ to converge to zero, and x_4^o and x_5^o to remain unchanged.

Notice also that the value of 6.36 at which d_u converged may be non feasible in some cases. d_u is defined as the daily fraction of the upper zone free water element that goes to the channel as interflow. If the time interval of integration of the soil phase equations is 1 day, the maximum feasible value which d_u can take would be 1.0/day. If the time interval is smaller than one day, 6 hours, for example, the maximum feasible value for d_u will be 4.0, or 1/(0.25 day). Therefore, a value of 6.36/day would be feasible for those cases in which the time interval of integration is smaller than 1./6.36 days, or 3.8 hours. Notice that since we estimated that parameter using a month with very high flows, the time interval of integration, which depends on the amount of precipitation, (Section 6.2.1), is certainly smaller than 0.25 days. Nonetheless, this very high value and the values at which the other parameters converged indicate that one month of data do not supply sufficient information for obtaining meaningful estimates of the parameters in Bird Creek. A close look at column (4) of Table 7.1 confirms this conclusion. Some of the coefficients of variation in this column are "astronomically" high. Some others are reasonably small, but those correspond to the parameters that converged to values very different from the values calibrated by the NWS, which is, at least, suspicious. Based on these findings an additional estimation run using four months of data which included a large portion of low flows was made. The months selected for this run were the months of April through July, 1957, in which a more representative sample of the range of discharges in Bird Creek is given. The measured discharges in these months are presented in Figure 7.2. From the results in columns (5) and (6) of Table 7.1 we can see that not only the parameter estimates are considerably closer to those of the NWS but the errors of estimation are within acceptable ranges, although the parameters related to the lower zone still have very large errors of estimation. It is important to notice that all parameters converged to feasible values. Only a_2 is one order of magnitude larger than the value estimated by the NWS. We hypothesize that the reduced order unit hydrograph model for the channel may be, in part, responsible for this

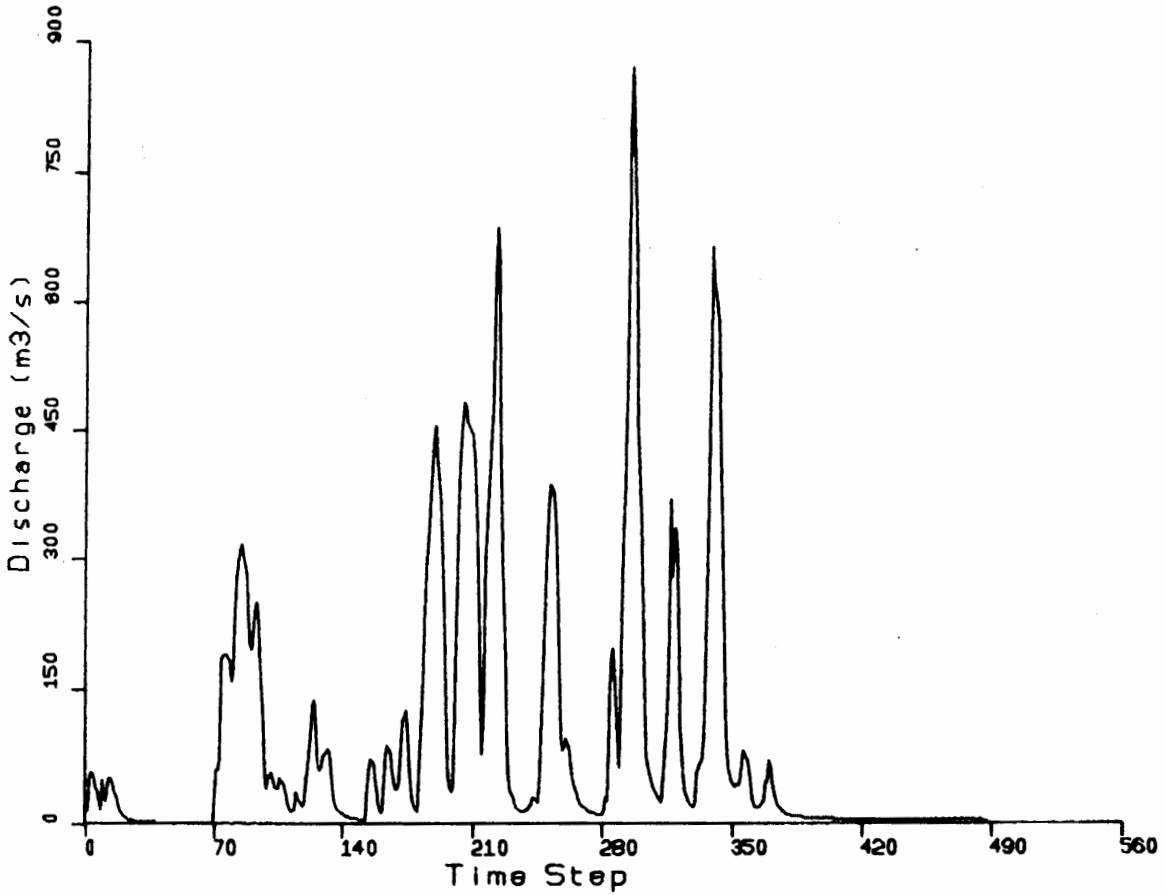


Figure 7.2 Bird Creek Discharges from April through July, 1957.

situation. In Figure 7.3 we have included the three state unit hydrograph used in our version of the NWSRFS model and the original Bird Creek hydrograph which was used for calibration of the model by the NWS. Notice that the three state hydrograph is unable to reproduce the magnitude of the peak of the original hydrograph. This inherent structural error causes some of the parameters to converge to values different from the true ones. (Bras and Restrepo-Posada, 1980). It was also shown by these authors that the value of convergence of the parameters depends on the time interval used for the parameter estimation, in such a form that the parameter estimates optimize the performance during the interval of parameter identification, but not in a global sense. In the NWSRFS model, a_2 produces the fastest possible responding output to the precipitation input, since it directs a fraction of that input directly to the channel as impervious runoff. It is reasonable to speculate that the model, trying to compensate for the smaller peak in the prediction of the hydrograph because of the channel defficiency, tries to direct a larger part of the input to the channel thus forcing a_2 to converge to a value considerably larger than the previously calibrated one. Notice also that we have obtained value of 25 to the coefficient of variation of P_f. This is a classical example of non observability about a parameter. This essentially means that the model will do as well with any value of P_f during the calibration period. The reason for this is based in the period of records chosen for estimating the parameters. A very high precipitation insures that the lower zone tension water element will be at its maximum capacity during a large part of the time. This implies that all percolating water will go directly to the lower zone free water elements thus making P_f unimportant.

The criteria for optimality are included in Table 7.2. The sum of squares test indicates that the computed sum of squares is very close to the expected value, which is a good indicator of the parameters being at the optimum point. The performance of the filter is analized through the statistics of the residuals. An optimally performing filter should produce residuals which are independent. The degree of independence of the residuals is measured by two different tests, already mentioned in Chapter 3. A Durbin and Watson statistic close to 2.0 indicates that the residuals are uncorrelated. In our case, a value of 0.8 indicates that the residuals are highly correlated. This is confirmed by the lag-1 correlation coefficient, which at a value of 0.52 is 11.5 standard deviations from its expected value of 0.0. This type of suboptimal performance is caused by the presence of structural errors in the system. These structural errors can be divided into two categories. First, errors due to parameters that are

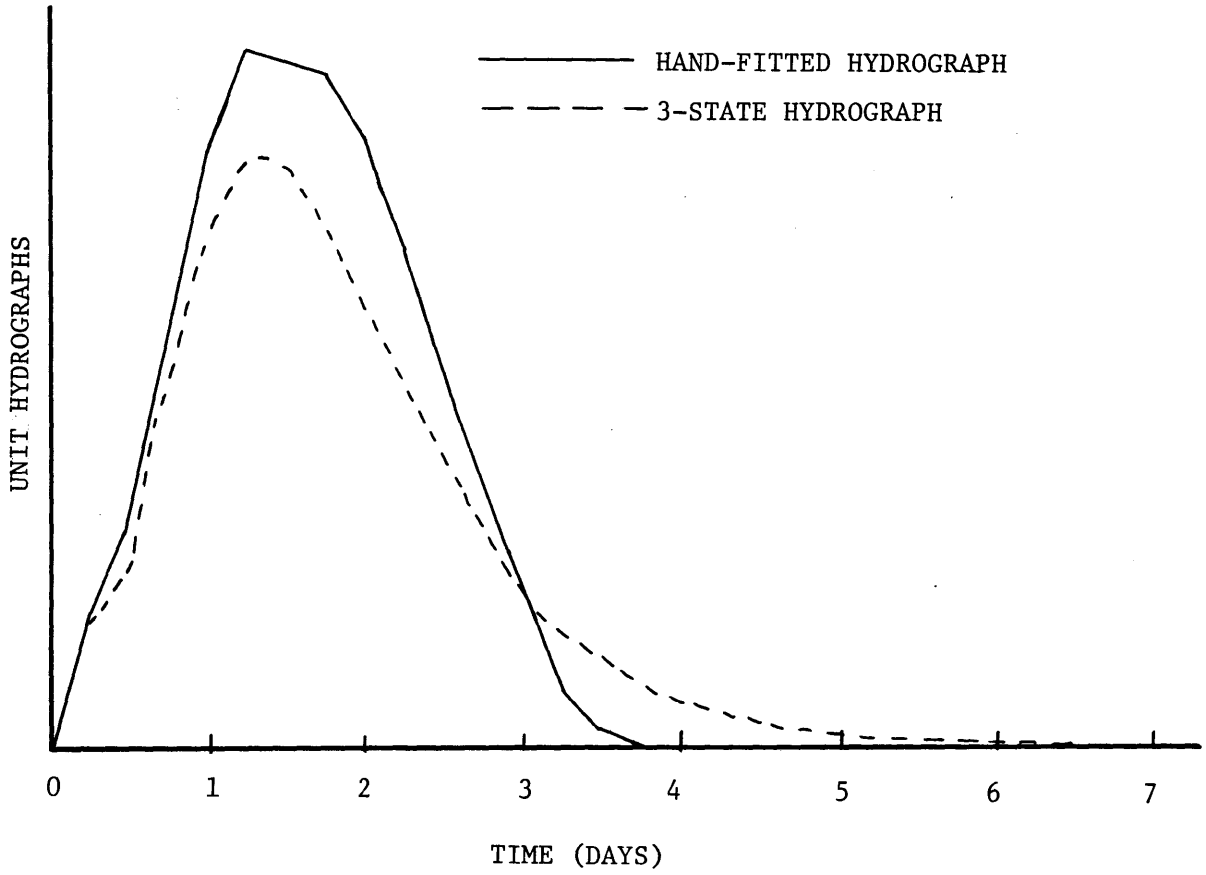


Figure 7.3 Three State Reduced-Order Unit Hydrograph for Bird Creek.

Serial Correlation Coefficients of the Normalized Residuals

Lag	0	1	2	3	4
ρ	0.90	0.52	0.14	-0.14	-0.26
r	-1.51	11.55	3.13	-3.07	-5.66

Sum of Squares Test

Number of Data Points:	488
Number of Parameters:	16
Expected Sum of Squares:	472
Standard Deviation:	31
Computed Sum of Squares:	440
Deviation (Units of S. Dev):	1.0
Durbin and Watson Statistic:	0.8

Table 7.2 Post Optimality Analysis. Parameter Estimates Without Prior Information, Bird Creek Basin, April-July 1957.

not at their true value. The errors in this category are solved by means of the parameter identification procedure. The second category covers the cases in which there is no set of parameters that fits the data in a statistical sense. Notice that even if we disregard the lumped characteristics of the model, the structural errors we have pointed out in the channel are sufficient to cause this type of suboptimal performance. Two more factors may also be a source of structural errors. These factors are the modeling of the model error terms as constant over time, and the assumption of independence between the model error terms. The effect of these factors on maximum likelihood parameter estimation has not yet been addressed.

7.2.3 Parameter Estimation With Prior Information.

The use of prior information in the maximum likelihood estimation of parameters requires the computation of the extended sum of squares, introduced in Chapter 3 as ξ_x. The calculation of this term requires the prior value of the parameters and the corresponding error of estimation. We obtained prior estimates for d_1' and d_1'' for Bird Creek from the model validation results presented in Chapter 5. These values and the corresponding errors of estimation were shown in Table (5.1). The results of estimating the parameters of the model by using prior information for the values of d_1' and d_1'' are presented in Table 7.3. For comparison with the previous results and with the values estimated by the NWSHRL, the results of estimating the parameters without prior information are also included. In general, the parameter estimates with prior information are close to the estimates without prior information, with few exceptions. In the estimates with prior information, a_2 is even higher than the corresponding estimate without prior information, but a_1 compensates by being smaller. The major change is in the estimate for d_1'', which moves closer to the prior value of that parameter, and away from the NWS estimate. Another effect of the use of prior information is the overall reduction in the lower bound of the estimation error, measured by the coefficient of variation. The only exception is the estimate of P_f which now has an infinite variance. The optimality measurements are presented in Table 7.4. There are not significant changes between these results and the corresponding results for the estimation without prior information.

7.2.4 Verification of the Parameter Estimates for Bird Creek

The results discussed in the previous section showed that the stochastic model is not performing optimally during the calibration period. It is important, nevertheless, to ap-

Param.	NWS	April-July 1957 (No Prior)		April-July 1957 (Prior)	
	Estim.	Estim.	C.V.	Estim.	C.V.
(1)	(2)	(3)	(4)	(5)	(6)
x_1^o	120.0	122.5	0.09	122.5	0.09
x_2^o	15.0	14.7	0.13	14.7	0.13
d_u	0.3	0.75	0.20	0.86	0.001
a_1	0.001	0.013	0.68	0.02	0.61
a_1	0.17	0.14	0.31	0.13	0.29
γ	48.0	38.2	1.09	38.0	0.84
α	2.10	1.09	0.66	1.14	0.46
x_3^o	160.0	160.0	0.51	160.0	0.50
x_5^o	14.0	12.5	2.40	12.5	0.55
x_4^o	140.0	140.0	0.61	140.0	0.37
d_1''	0.126	0.08	1.71	0.55	0.45
d_1'	0.013	0.011	1.16	0.015	0.70
P_f	0.02	0.02	25.21	0.0	∞
μ	3.5	5.17	0.83	5.18	0.25
R		0.25	0.12	0.24	0.12
Q_{11}		1.33	0.14	1.35	0.10

Table 7.3 Maximum Likelihood Parameter Estimates With Prior Information for the Bird Creek Basin, April-July 1957.

Serial Correlation Coefficients of the Normalized Residuals

Lag	0	1	2	3	4
ρ	0.91	0.55	0.17	-0.12	-0.25
r	-1.38	12.11	3.74	-2.59	-5.56

Sum of Squares Test

Number of Data Points:	488
Number of Parameters:	16
Expected Sum of Squares:	472
Standard Deviation:	31
Computed Sum of Squares:	444
Deviation (Units of S. Dev):	0.9
Durbin and Watson Statistic:	0.8

Table 7.4 Post Optimality Analysis. Parameter Estimates With Prior Information, Bird Creek Basin, April-July 1957.

ply the model to periods of records which were not used in the parameter estimation and to compare the performance of the model with the two sets of parameters, the one manually calibrated by the National Weather Service, and the other obtained from the maximum likelihood estimation with prior information. We selected the water year 1958-1959 which previous experience showed to be difficult to simulate. Four different simulations were made. The first two series of results that we are going to show correspond to deterministic simulations, each one performed with each set of parameters. We will be referring to these deterministic simulations as "off line runs". The second series of results correspond to stochastic simulations in which the linearized Kalman filter has been used to update the states. These simulations will be referred to as "on line runs". For shortness, we will call OFNWS and OFML the off line runs made with the National Weather Service and the maximum likelihood estimated parameters, respectively. The on line runs will be called ONNWS and ONML. The input data used for these runs are included in Appendix A.

From the water year 1958-1959 we selected three months to illustrate the characteristics of the simulations carried on with the two set of parameters. The general characteristic of the parameters estimated with maximum likelihood is that, in the Bird Creek case, the model over predicted the small hydrographs, predicted well hydrographs in an intermediate range, and under predicted the large hydrographs. The under prediction of small hydrographs can be seen in the off line simulation for the month of February, 1959. The OFNWS run, Figure 7.4 predicts reasonably well a small hydrograph starting around the 10th of the month. The discharge predicted by OFML is much larger, which can also be seen in Figure 7.4. The on line runs correct the predictions with the final result that ONML predicts the peak better than ONNWS. (Figure 7.5). The simulation of the discharges for the month of March is an example of an intermediate hydrograph, beginning on the 25th of the month, which is better predicted by the OFML than by the OFNWS, (Figure 7.6). Both filtered predictions, Figure 7.7, correct the over prediction of the small hydrograph at the beginning of the month, with the ONML showing a better prediction of the rising limb of the final hydrograph, but a worse prediction of the peak flow than ONNWS. Finally, the month of July presents a case in which all large hydrographs are under predicted both by OFNWS and OFML, with the larger errors being produced by the latter. (Figures 7.8). The ONNWS run predictions are good regarding the magnitude of the peak, but are lagging one time step. (Figure 7.9). This lag in the prediction of the hydrograph is typical of the on line runs for the complete year. Notice that the lag essentially indicates that the model under predicts the rising limb of the hydrograph, while over predicting the falling limb. If we recall the unit hydrograph model, Figure 7.3, we see that the three state unit hydrograph model does exactly that: the

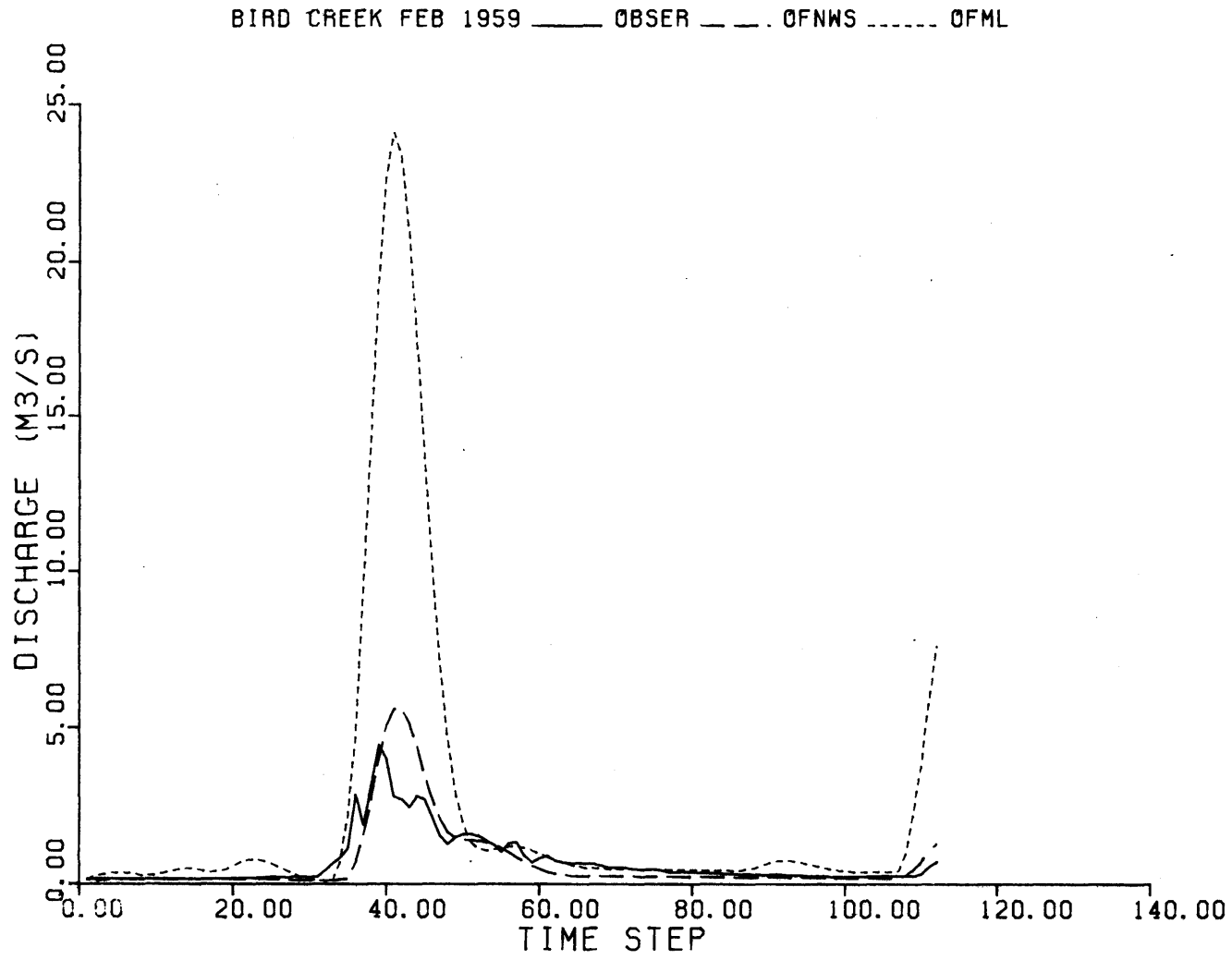


Figure 7.4 Six-Hour Lead Off-Line Forecasts for Bird Creek, February 1959.

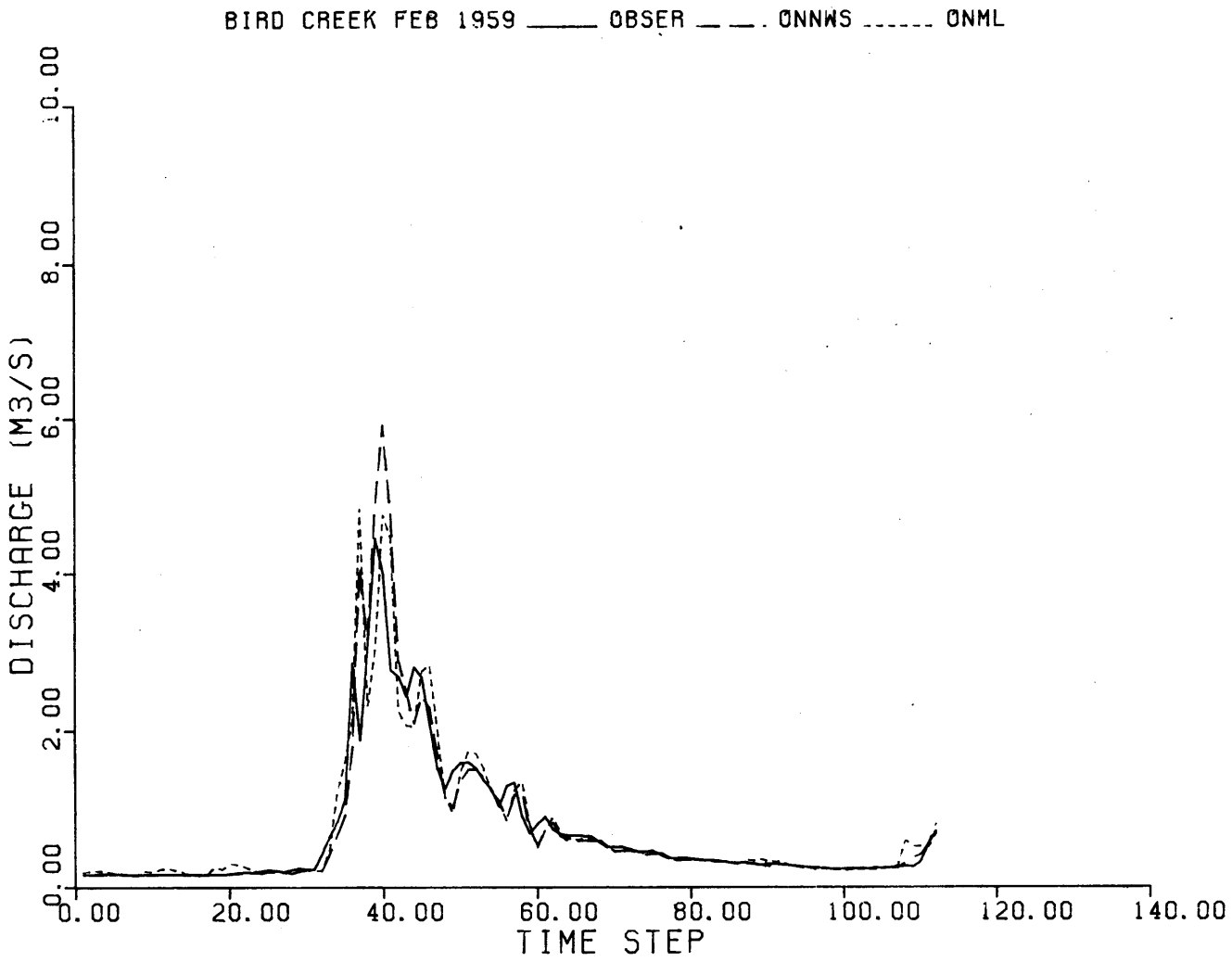


Figure 7.5 Six-Hour Lead On-Line Forecasts for Bird Creek, February 1959.

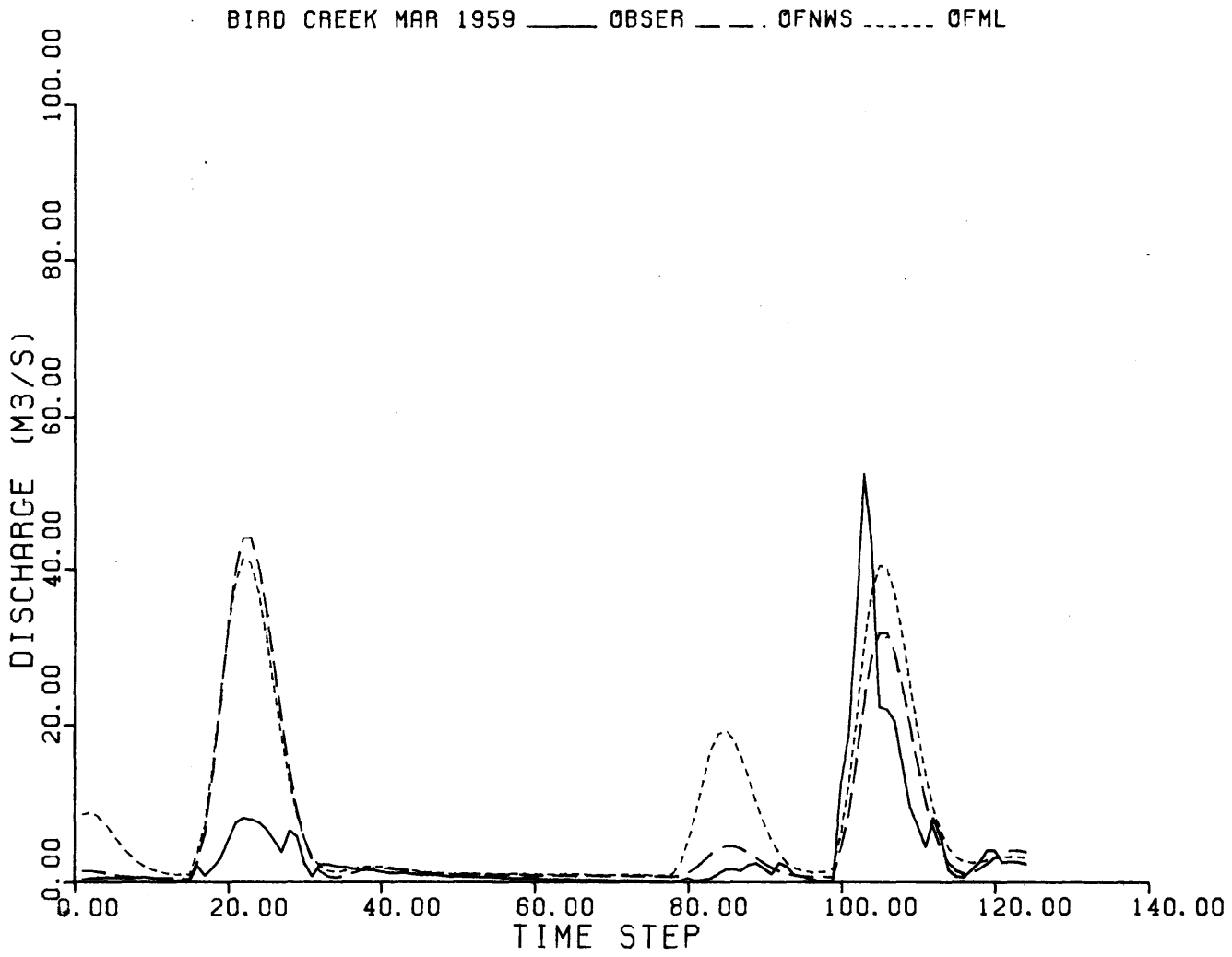


Figure 7.6 Six-Hour Lead Off-Line Forecasts for Bird Creek, March 1959.

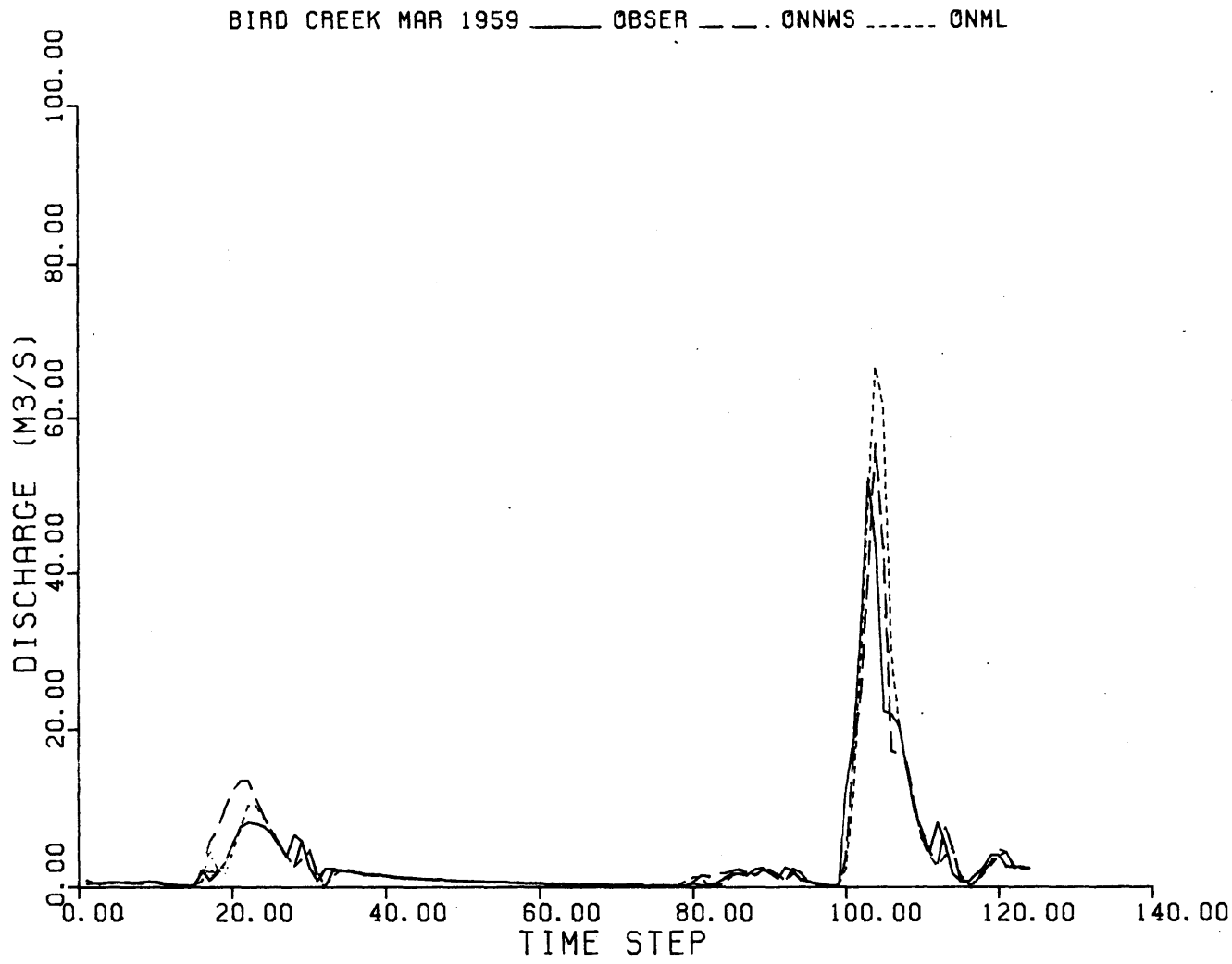


Figure 7.7 Six-Hour Lead On-Line Forecasts for Bird Creek, March 1959.

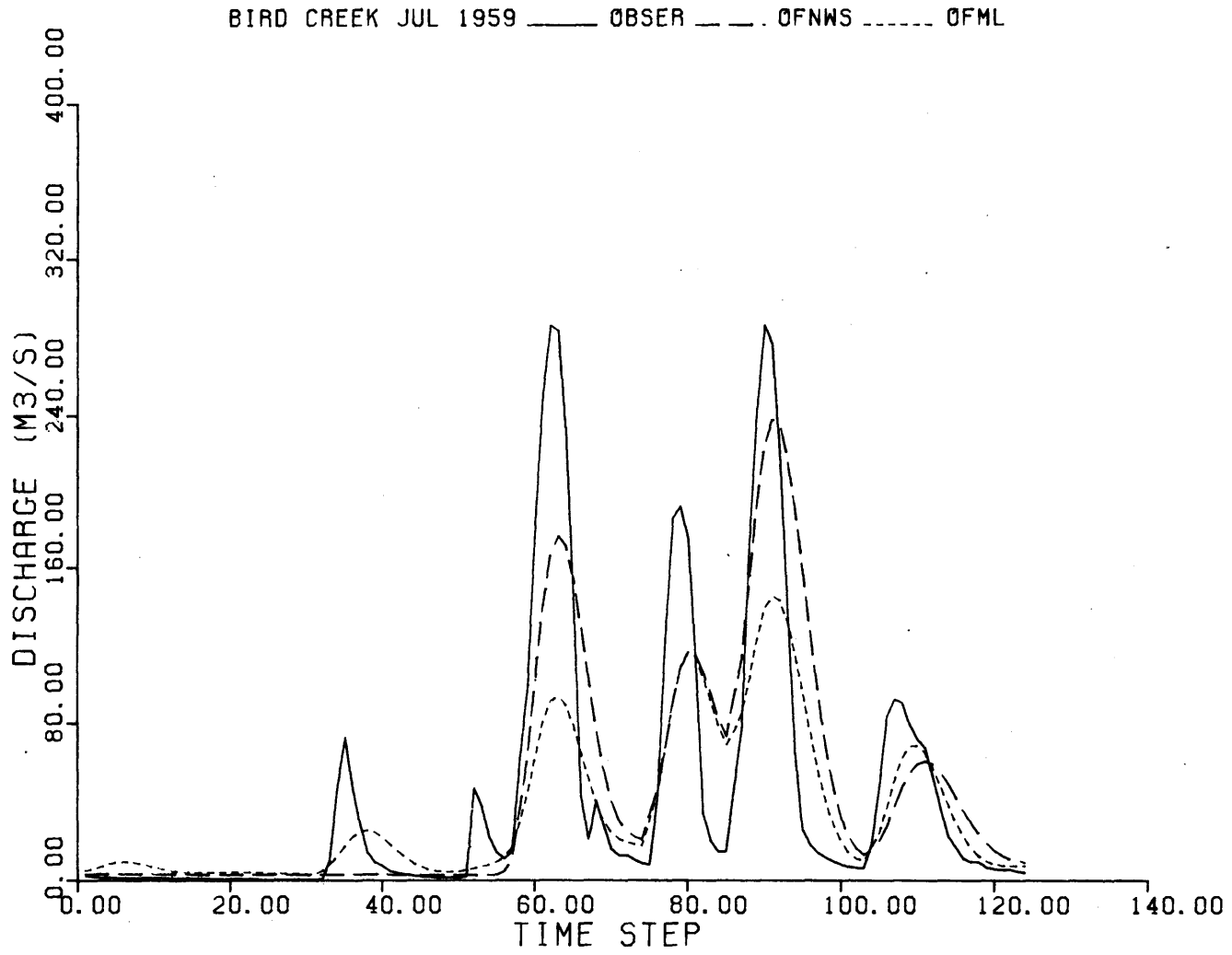


Figure 7.8 Six-Hour Lead Off-Line Forecasts for Bird Creek, July 1959.

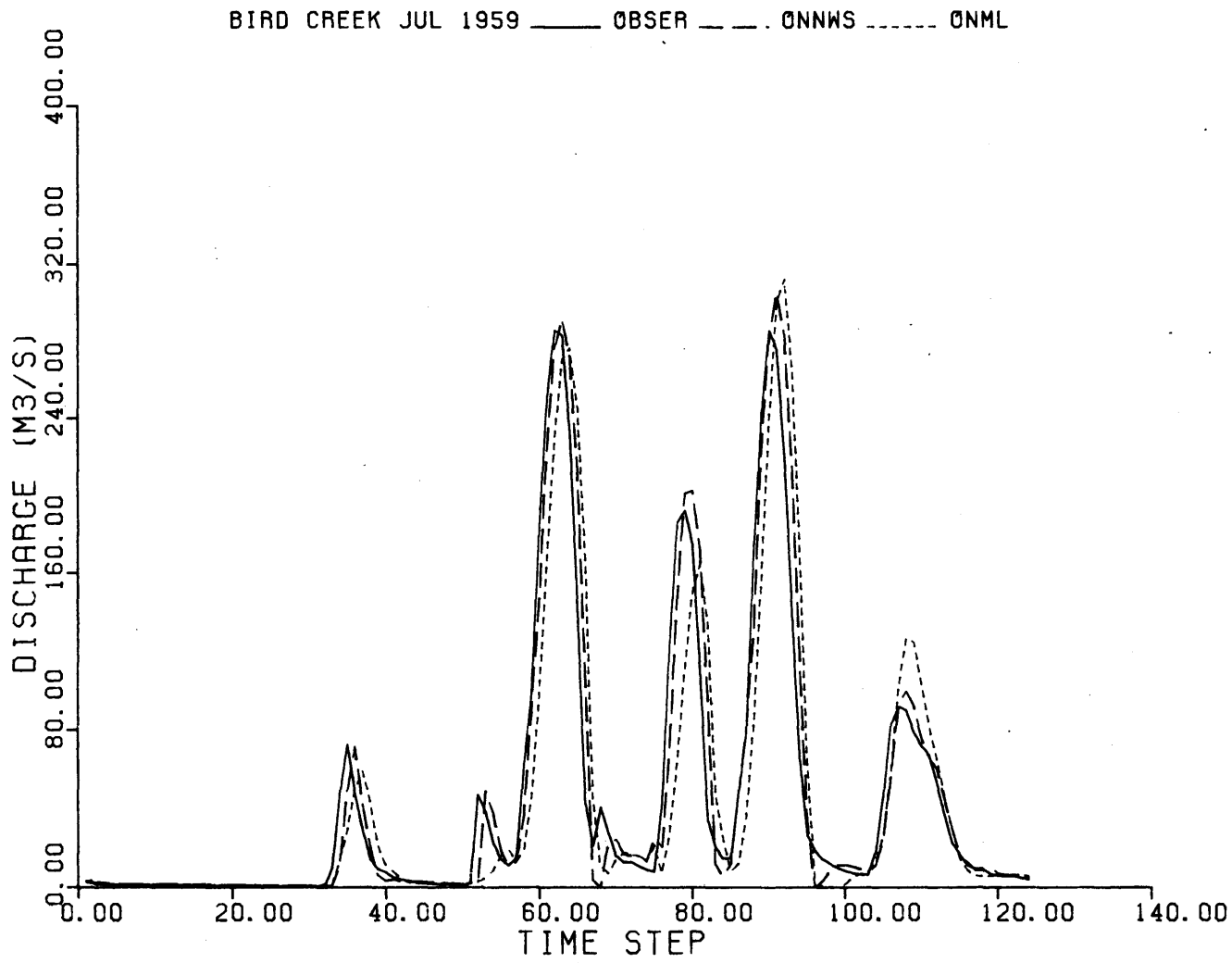


Figure 7.9 Six-Hour Lead On-Line Forecasts for Bird Creek, July 1959.

rising limb of the unit hydrograph is under predicted and the falling limb is over predicted. This indicates that the channel model is certainly one of the causes of the structural errors. The ONML run, Figure 7.9, also shows the lag in the prediction, with a worse prediction of the hydrographs' peak. The figures corresponding to the remaining months of the four simulation runs are included in Appendix B. In addition to the graphical summary of the runs we have made, we are including a statistical summary, produced by the standard NWSRFS.

The statistical analysis of the simulations, included in the NWSRFS model is based on the errors in the predictions of daily volumes of runoff, while the maximum likelihood parameter estimates were obtained from instantaneous discharges taken every six hours. A proper statistical analysis should be based on the analysis of the results at the same level of aggregation as the one used in the parameter identification procedure. However, since the performance of all models will be compared using the same type of statistics, we believe that, for that purpose, the use of the daily volumes of runoff is adequate.

The statistical summary of the off line runs is presented in Table 7.5. The difference between the simulated mean and the observed mean for the off line run with parameters estimated by maximum likelihood indicates that this model under predicts the discharges. Moreover, the overall performance of the model with maximum likelihood estimated parameters is worse than the performance of the model with the original parameters, according to the percent absolute error and the percent root mean square error, (RMS), although in both cases the error is very high (181.46% and 193.53%). These two statistics, the percent absolute error and the percent root mean square error are calculated in the NWSRFS as percent of the annual average flow. This places equal weighs on the individual predictions, regardless of the magnitude of the observed flows. In other words, a 10 m^3/s day error in the prediction of a 100 m^3/s day flow, (a 10% error), will have the same weight than a 10 m^3/s day error in the prediction of a 1 m^3/s day flow (a 1000% error). These statistics are a good measure of the performance of a model if the fitting criterion has been the minimization of either the average absolute error or the RMS error. If we assume that a person performing a manual fitting of a model will pay more attention to the prediction of the high flows than to the prediction of the low flows, the two statistics we are discussing will adequately measure the performance of the model. These two statistics may not reflect the performance of a model whose parameters have been estimated by the method of maximum likelihood, because this method weighs all errors according to their expected standard deviations, which, in a non stationary system like

Annual Averages	NWS	ML
Simulated Mean (m^3/s day)	9.605	9.444
Observed Mean (m^3/s day)	9.180	9.180
Percent Bias	4.62	2.87
Bias (mm)	5.710	3.547
Maximum Error (Sim-Obs)(m^3/s day)	-123.978	-175.075
Percent Absolute Error	65.47	65.85
Percent RMS Error	181.46	193.53

a) Daily Volume Error Statistics

Flow Interval (m^3/s day)	Number of cases	Percent RMS Error NWS	ML
0.00 - 0.15	87	1423.60	1510.48
0.15 - 0.70	109	547.47	807.34
0.70 - 7.00	113	161.33	241.95
7.00 - 14.00	16	223.68	167.78
14.00 - 42.00	18	90.16	61.57
42.00 - 140.00	16	62.81	47.81
140.00 and above	9	42.66	56.53

b) Distribution of Errors by Flow Interval.

Table 7.5 Statistical Analysis of Prediction of Daily Volumes of Runoff with Off Line Simulations. Bird Creek, Water Year 1958-1959.

the NWSRFS model, vary according to the magnitude of the flow. Table 7.5-b presents the same statistics grouped according to the magnitude of several intervals of flow. The results of Table 7.5-b are plotted in Figure 7.10 to provide a visual comparison. We can see that both models performed very poorly in predicting the flows under 0.70 m^3/s day, and reasonably well for high flows. In the same table, we see that the model with the parameters estimated by maximum likelihood performed better than the model with the original parameters in predicting the flows in the intermediate ranges of 7.0 to 140.0 m^3/s day. Another dissaggregation of the percent absolute error and the percent RMS error is shown in Table 7.6. In that table we show the two statistics computed for every months of the year. The percents are calculated relative to the average monthly runoff. The annual averages shown in that table are the arithmetic average of the values in the columns above them. Therefore, they differ from the annual percent absolute error and annual percent RMS error reported in Table 7.5.

The statistical summary for the on-line runs is presented in Table 7.7. In that table we see that the values of the percent absolute error and the percent RMS error for the on line runs are smaller than the errors for the off line runs which were shown in Table 7.5. The percent RMS error for the on line run with the maximum likelihood estimated parameters is bigger than the same statistic for both on line and off line runs done with the original parameters. We have explained already why this statistic may not reflect very well the performance of a model whose parameters have been fitted with a maximum likelihood criterion. The disaggregation of the percent RMS error by interval of flow, presented in Table 7.7-b shows that the percent RMS for the on line run with the maximum likelihood estimated parameters tend to stay in a range between 25% and 65% of the flow, with exception of the very low flows, which although showing a 144% RMS error in the prediction, perform better than the model with the original parameters. Figure 7.11 shows the percent RMS error by flow interval for both on line runs, for ease of comparison. The percent RMS for the on line run with the original parameters clearly shows a better performance of the model in predicting the high flows than in predicting the low flows, thus confirming our assesment about the performance of models with manually fitted parameters. The disaggregation of percent absolute error and percent RMS error by months is shown in Table 7.8.

7.3 Maximum Likelihood Parameter Estimation of the NWSRFS model in the Cohocton River Basin.

The experience gained in estimating the parameters of Bird Creek indicated that one month of data was not enough for

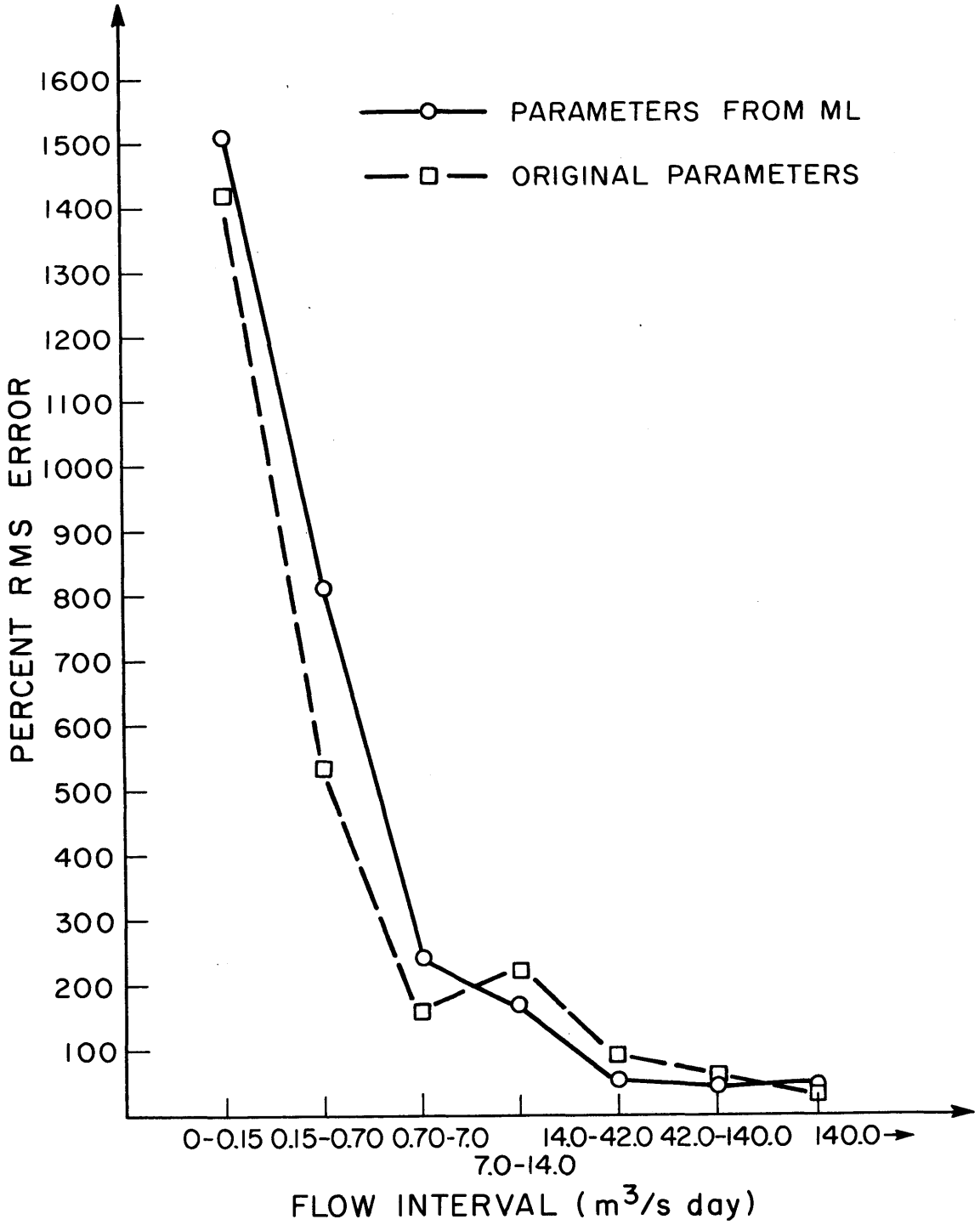


Figure 7.10 Percent RMSError by Interval of Flow. Bird Creek, Off-line Runs.

	National Weather Service		Maximum Likelihood	
MONTH	Percent Abs Err	Percent RMS	Percent Abs Err	Percent RMS
October	616.45	629.02	683.02	822.57
November	107.18	213.29	83.34	124.94
December	73.84	79.11	164.49	362.97
January	23.34	27.89	158.69	329.98
February	41.45	85.15	242.71	658.76
March	114.41	244.89	153.64	263.62
April	67.76	117.80	117.03	178.90
May	58.52	95.69	51.79	89.55
June	81.58	129.47	85.91	110.07
July	55.61	83.44	51.66	97.58
August	312.57	321.48	328.36	335.72
September	82.92	194.10	79.64	231.99
Average	136.30	185.11	183.36	300.55

Table 7.6. Percent Absolute Error and Percent RMS Error for the Off Line Runs by Month. Bird Creek Water Year 1958-1959.

Annual Averages	NWS	ML
Simulated Mean (m^3/s day)	9.297	7.441
Observed Mean (m^3/s day)	9.180	9.180
Percent Bias	1.27	-18.94
Bias (mm)	1.573	-23.398
Maximum Error (Sim-Obs)(m^3/s day)	41.168	-168.143
Percent Absolute Error	18.37	44.39
Percent RMS Error	62.13	186.86

a) Daily Volume Error Statistics

Flow Interval (m^3/s day)	Number of cases	Percent RMS Error NWS	ML
0.00 - 0.15	87	213.64	144.23
0.15 - 0.70	109	88.71	55.47
0.70 - 7.00	113	36.07	26.62
7.00 - 14.00	16	34.19	49.06
14.00 - 42.00	18	24.11	47.87
42.00 - 140.00	16	29.56	64.27
140.00 and above	9	11.00	51.37

b) Distribution of Errors by Flow Interval.

Table 7.7 Statistical Analysis of Prediction of Daily Volumes of Runoff with On Line Simulations. Bird Creek, Water Year 1958-1959.

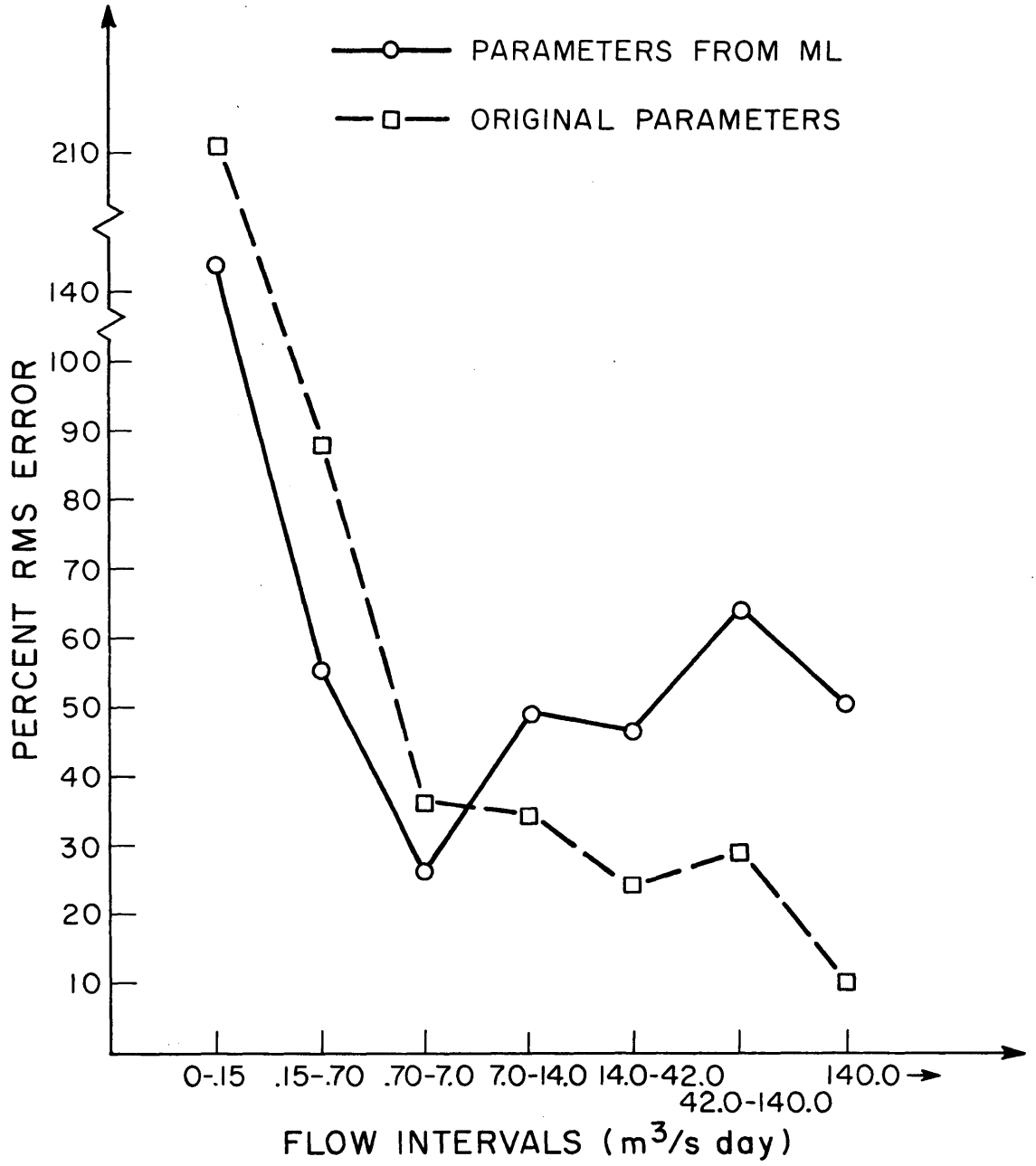


Figure 7.11 Percent RMS Error by Interval of Flow. Bird Creek, On-line Runs.

	National Weather Service		Maximum Likelihood	
MONTH	Percent Abs Err	Percent RMS	Percent Abs Err	Percent RMS
October	94.26	95.09	82.07	99.60
November	16.95	48.30	9.58	15.47
December	10.88	11.91	23.22	57.30
January	3.99	4.78	12.16	24.44
February	16.74	38.99	10.25	17.56
March	29.28	64.58	28.04	90.20
April	15.65	33.73	23.03	55.90
May	19.27	33.81	63.52	124.41
June	20.98	37.50	20.21	40.06
July	15.43	28.42	38.91	66.80
August	47.37	49.76	34.55	38.41
September	20.64	63.34	17.57	54.49
Average	25.96	42.52	30.25	57.05

Table 7.8. Percent Absolute Error and Percent RMS Error for the On Line Runs by Month.
Bird Creek Water Year 1958-1959.

obtaining meaningful parameter estimates. We decided to use at least two months of data for estimating the parameters of the Cohocton River. The months selected were May and June of 1969, which not only have a good range of high and low flows (Figure 7.12), but are also free from snowfall and snowmelt. This last factor enables us to obtain parameter estimates which are independent of the influence of the calibration of the snowmelt model and errors in the snowfall measurements. Table 7.9 presents the original parameters used by the National Weather Service, which were also used by Kitanidis and Bras (1980) for the validation of the stochastic version of the NWSRFS model. In that table we also include the initial value of the parameters, the value at which the parameters converged, and the coefficient of variation of the estimated parameters. Notice that, in contrast with the parameter estimation for Bird Creek, Section 7.2, the initial conditions of some of the state variables were also estimated. Since the prior information about the coefficients of the lower zone linear reservoirs indicated that these coefficients were essentially identical to the values estimated by the National Weather Service, we decided to leave these parameters fixed at their prior values. Some of the parameters converged to values significantly different to those calibrated by the National Weather Service. In particular, the value of x_1^o estimated by the NWS as 100, converged to a value of 39.4. We may notice, however, that x_1 is a state variable that is depleted only by evaporation and which is observable only when there is outflow from that element into the upper zone free water element. That occurs in the original formulation of the model when $x_1(t)=x_1^o$. In our case, the threshold corresponding to the upper zone tension water element has been statistically linearized. In this form, the outflow from the element depends actually on the probability that the water contents during a time interval exceeds the value of the threshold. Since the probability distribution of the state variable x_1 has been approximated by a normal distribution, there is always, in theory, a finite probability of the threshold being exceeded. In Kitanidis and Bras' model, the prior variance of x_1 depends on the precipitation. If there is no precipitation, the variance is very small, which essentially converts the normal distribution into a spike thus decreasing the probability of exeeding the threshold to virtually zero. This makes the water content in the upper zone tension water element only observable when there is precipitation and shortly after. This causes the model to be sensitive in most cases to the difference between x_1^o and $x_1(t)$, rather than to the full value of

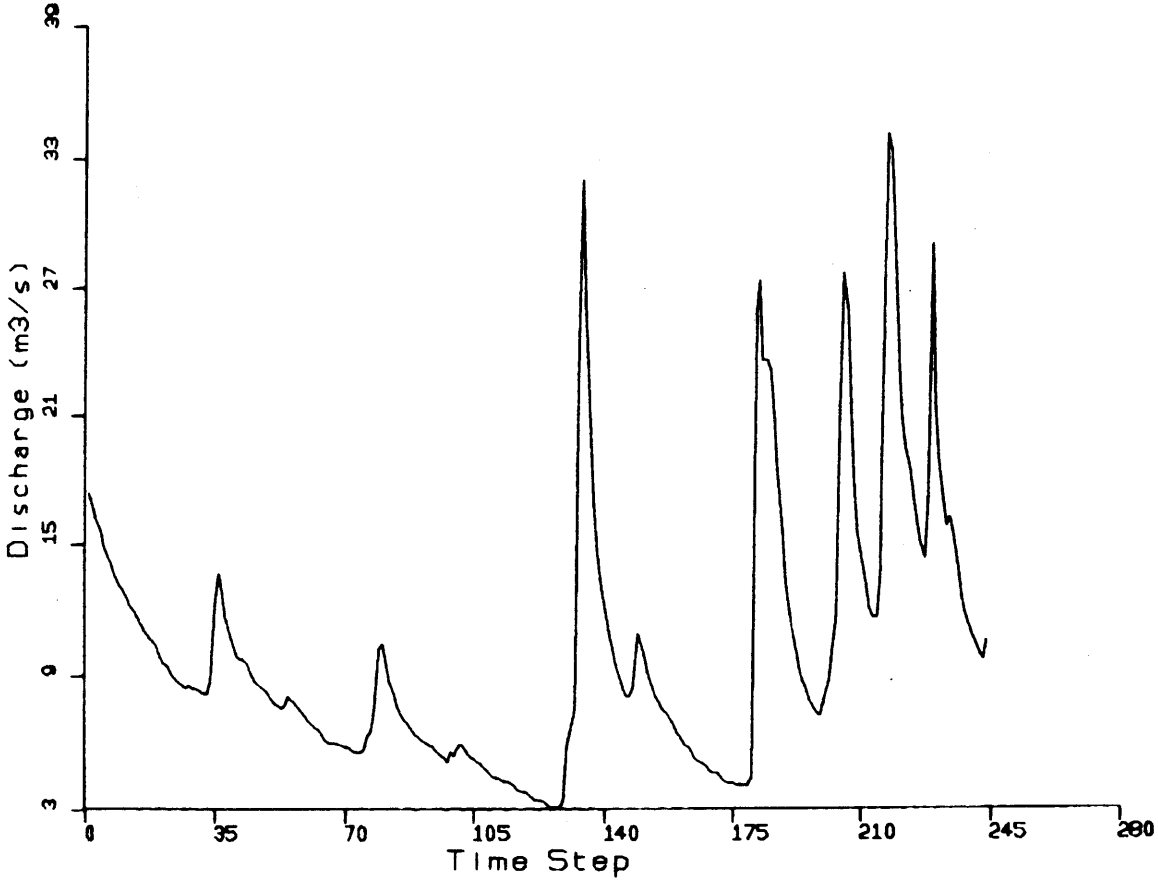


Figure 7.12 Cohocton River Discharges for May and June, 1969.

Param.	NWS Estimate	Initial Value	May-June Estimate	1969 C.V.
x_1^o	100.0	50.0	39.4	0.09
x_2^o	25.0	10.0	16.5	0.17
d_u	0.3	0.3	0.20	0.13
a_2	0.004	0.2	0.0	-
a_1	0.010	0.02	0.0	-
γ	25.0	25.0	22.6	7.5
α	1.8	1.8	0.8	8.3
x_3^o	200.0	80.0	80.9	8.2
x_5^o	100.0	20.0	17.6	9.2
x_4^o	75.0	50.0	50.0	6.9
$x_1(0)$	-	10.0	6.16	1.3
$x_2(0)$	-	1.0	1.25	3.0
$x_3(0)$	-	60.0	81.1	18.7
$x_4(0)$	-	1.0	1.25	3.0
$x_5(0)$	-	40.0	28.2	3.5
μ	0.03	0.02	0.0	-
R	-	0.1	0.025	0.05
Q_{11}	-	0.3	1.29	0.15
Q_{22}	-	0.1	0.05	0.31

Table 7.9 Maximum Likelihood Parameter Estimates for the Cohocton River Basin.

x_1^o. For example, a precipitation of 20 mm which does not produce interflow, only indicates that $x_1(t)$ was at least 20 mm smaller than x_1^o. To see this more clearly we are going to examine the relationship between x_1^o and $x_1(0)$, which is the initial condition of $x_1(t)$.

Since $x_1(t)$ is observable only when there is precipitation, the filtering stage of the state estimation process will not correct the state variable between t=0 and the arrival of the first storm, $t=t_1$. This implies that there is a deterministic relationship between $x_1(0)$ and $x_1(t_1)$, unless $x_1(t_1)=0.0$. (Notice that the evapotranspiration demand is modeled as a deterministic input). Now, if the model observes only the difference $x_1^o - x_1(t_1)$, given the deterministic relationship between $x_1(0)$ and $x_1(t_1)$, we would expect the deterministic model to be sensitive to the difference $x_1^o - x_1(0)$, which implies that the estimates of x_1^o and $x_1(0)$ should be highly correlated. To verify this hypothesis we followed the procedure which was suggested in Chapter 6 to examine the interdependency between the same two parameters when the threshold was substituted by the smooth "S" curve. That procedure consisted in plotting the successive values the parameters take during the non linear optimization process. The basic idea is that if the parameters are highly interdependent, the successive values the parameters take during the non linear optimization process should lie along a straight line that marks the position of the ridge. Therefore, if the model is sensitive to the difference between two parameters, the parameter values at different iterations should lie along a straight line with the form

$$\theta_1 - \theta_2 = \Delta$$

where θ_1 and θ_2 are the two parameters and Δ is the difference actually observable.

Now, if we account for the fact that the model is sensitive to the difference between one of the parameters (x_1^o) and a <u>function</u> of the other parameter ($x_1(t_1) \simeq \lambda x_1(0) + \Delta'$), we should find a relationship of the form

$$x_1^{\circ} - \lambda x_1(0) = \Delta$$

or,

$$x_1^{\circ} = \lambda x_1(0) + \Delta$$

To test the effect of the interaction between x_1° and $x_1(t)$ in a deterministic model, we performed a least squares estimation of x_1°, $x_1(0)$ and another parameter, d_u, to add one more dimension to the parameter space. The initial values of the parameters was the maximum likelihood estimate. The different pairs of values x_1° and $x_1(0)$ took at the end of 21 linear searches are plotted in Figure 7.13. These different values lie along the line with equation $x_1^{\circ}=0.676x_1(0)+45.9$. Additional insight into the presense of the ridge is given by the contour lines of the objective function. The contours of the weighted least squares objective function near the maximum likelihood optimum are shown in Figure 7.14. These contours were obtained with all parameters fixed at their maximum likelihood estimates, while the weighted least squares function was calculated for different values of x_1° and $x_1(0)$. The elongated shape of the contours delimits the ridge, which coincides with the straight line in Figure 7.13 This result confirms the hypotheses of the deterministic model being sensitive to the difference between x_1° and $x_1(t_1)$.

Since interdependency between parameters is inherent to the model it affects all methods of parameter estimation, regardless of the fitting criterion. However, the nature of the interaction is different for the maximum likelihood procedure. This is due to the fact that the maximum likelihood criterion includes a filtering stage which affects $x_1(t_1)$. The action of this filtering stage can not be predicted in advance. We can only say that it affects the approximate linear relationship between $x_1(t_1)$ and $x_1(0)$ that we assumed above. Figure 7.15 shows the contour lines of the loglikelihood function in the neighborhood of the optimum. These were generated by changing the values of the parameters x_1° and $x_1(0)$, while the remaining parameters were left at their maximum likelihood estimates. We first notice that the final point of convergence (ξ') was close to the global optimum, (ξ^*), and that there is

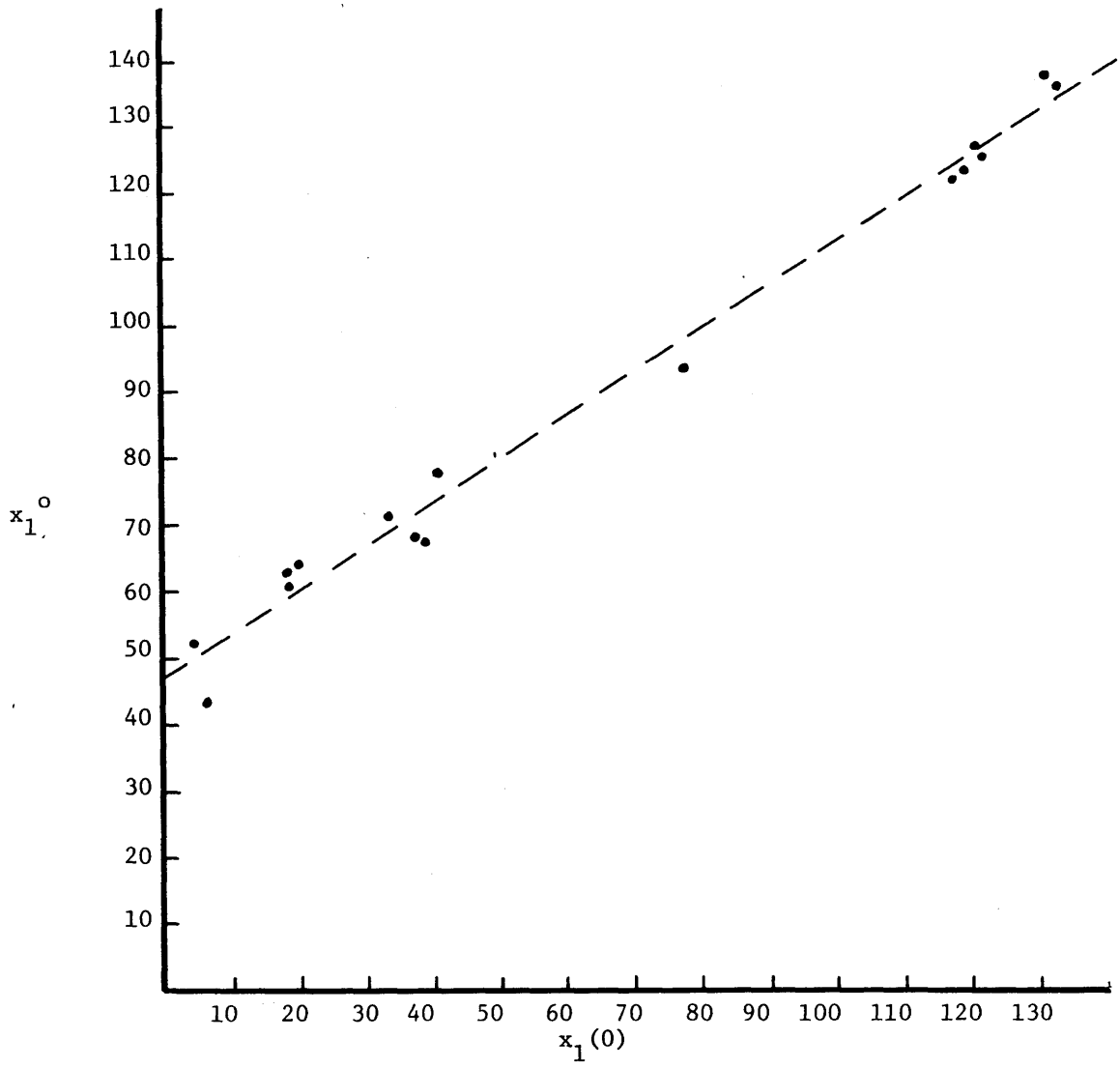


Figure 7.13 Alignment of Estimates of x_1^o and $x_1(0)$ at Several Iterations.

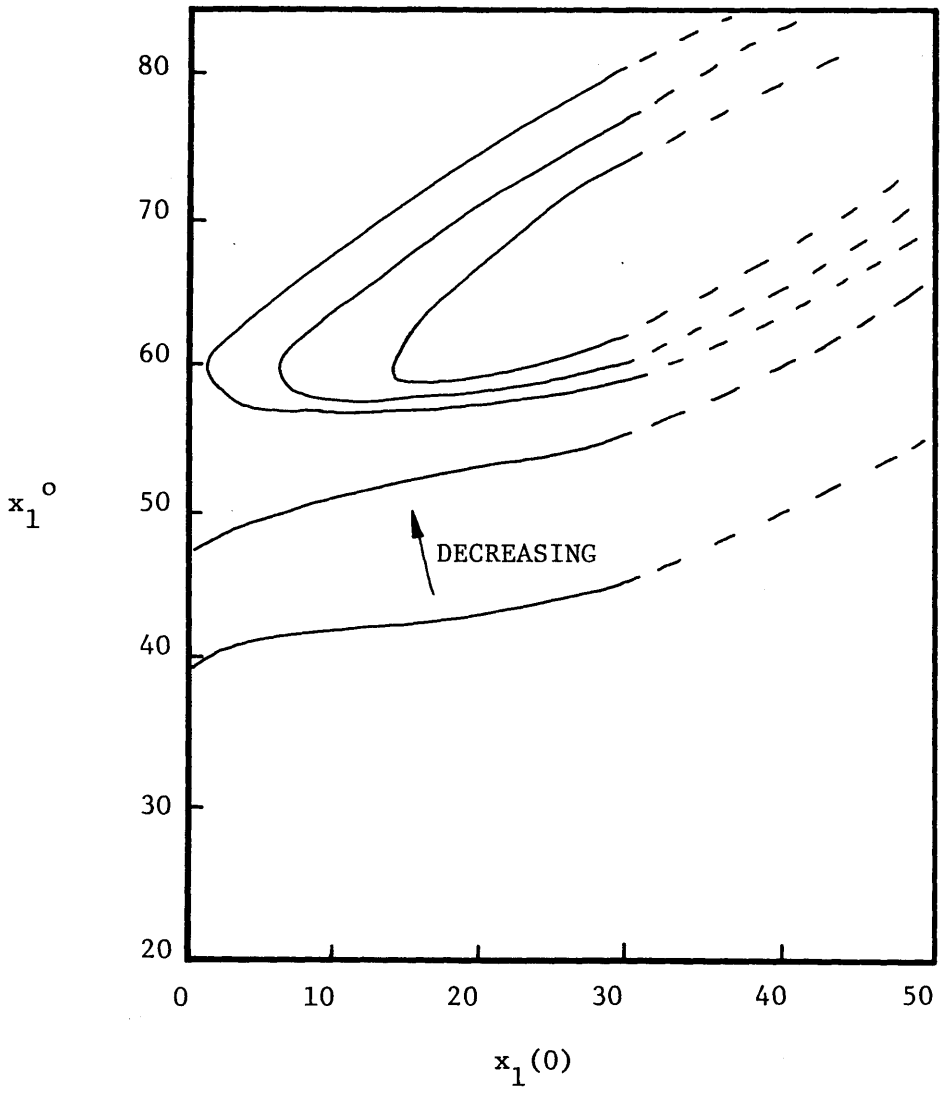


Figure 7.14 Contour Lines of the WLS Objective Function Near the ML Optimum.

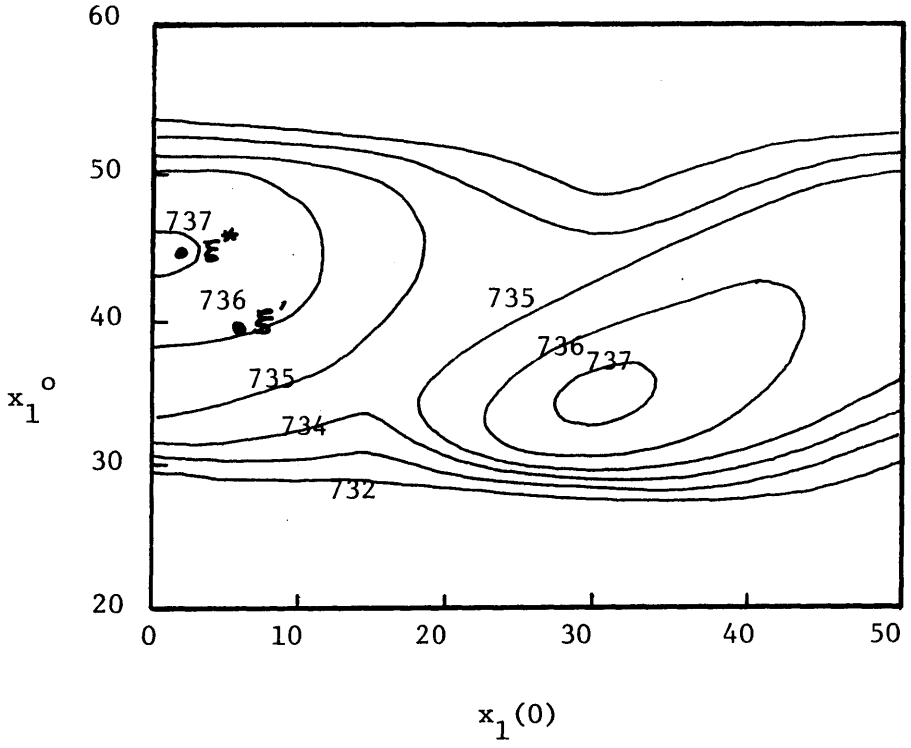


Figure 7.15 Contour Lines of the Loglikelihood Function in the Neighborhood of the Optimum.

a local optimum at $x_1(0)=30$, $x_1^o=35$. Interestingly enough, the reader may notice that the contours of the loglikelihood do not change abruptly at the value of the threshold ($x_1(0) = x_1^o$). This is due to the fact that the threshold function of the upper zone element is statistically linearized, which sends the excess $x_1(0)-x_1^o>0$ to the upper zone free water element. The low and high variances of estimation of x_1^o and $x_1(0)$, respectively, are explained by the curvature of the contour lines around ξ^*. The loglikelihood varies considerably along the direction of x_1^o, while the contour lines are essentially parallel to the direction of $x_1(0)$. The reader may also notice that the estimated correlation coefficient between x_1^o and $x_1(0)$ of 0.61, Table 7.10, implies that those two parameter estimates are not highly correlated. This is also consistent with the absence of a dominant ridge in the loglikelihood function, as was shown in Figure 7.15.

This problem can only be solved if the total magnitude of x_1^o were observable during the data interval selected for model calibration. The total magnitude of x_1^o would be observable only during a period of high precipitation following a long dry spell, during which x_1 and x_2 can be safely assumed to be empty. The incoming precipitation would first satisfy the tension water requirements with the excess water flowing into the upper zone free water element. When the latter becomes saturated surface runoff will be generated. The predicted discharge will thus contain information about the total magnitude of x_1^o and x_2^o. A value of x_1^o smaller than the true one will cause a large fraction of the precipitation to flow into the channel as interflow, and eventually as surface runoff, causing an over prediction of the discharge. Even though we tried to include the case of a high discharge following a recession curve in the period selected for calibration, the verification runs will clearly show that the true value of x_1^o is larger than the one estimated. This finding may seem paradoxical given the small coefficient of variation which was obtained for the estimate of x_1^o. We must remember that the variance of estimation computed by the maximum likelihood procedure is just a lower bound of that variance. The actual error may be larger.

	x_1^o	x_2^o	d_u	γ	α	x_3^o	x_5^o	x_4^o	R	$X_1(0)$	$X_2(0)$	$X_3(0)$	$X_4(0)$	$X_5(0)$	Q_{11}	Q_{22}
x_1^o	1.00															
x_2^o	0.00	1.00														
d_u	0.06	-0.72	1.00													
γ	0.08	0.18	0.07	1.00												
α	-0.07	-0.22	0.00	-0.83	1.00											
x_3^o	-0.13	-0.07	0.15	0.65	-0.30	1.00			SYMMETRIC							
x_5^o	-0.07	-0.23	-0.01	-0.92	0.97	-0.36	1.00									
x_4^o	-0.15	0.02	-0.16	-0.79	0.55	-0.94	0.59	1.00								
R	-0.02	-0.14	0.18	-0.04	-0.05	-0.11	-0.02	0.11	1.00							
$X_1(0)$	0.61	-0.50	0.56	-0.04	-0.02	0.06	0.02	-0.09	-0.16	1.00						
$x_2(0)$	-0.01	-0.14	0.01	-0.42	0.53	-0.01	0.53	0.19	-0.05	-0.02	1.00					
$x_3(0)$	-0.21	0.14	-0.20	-0.57	0.26	-0.96	0.30	0.89	0.09	-0.17	-0.04	1.00				
$x_4(0)$	0.01	0.25	-0.18	0.29	-0.55	-0.24	-0.51	-0.02	0.06	-0.08	-0.70	0.35	1.00			
$x_5(0)$	0.11	-0.09	0.21	0.52	-0.34	0.56	-0.40	-0.47	0.01	0.14	-0.02	-0.69	-0.42	1.00		
Q_{11}	0.09	0.77	-0.86	0.14	-0.13	-0.07	-0.16	0.04	-0.10	-0.61	-0.11	0.16	0.27	-0.18	1.00	
Q_{12}	0.10	0.49	-0.77	-0.13	-0.04	-0.18	0.02	0.16	0.33	-0.44	-0.00	0.17	0.21	-0.22	0.67	1.00

Table 7.10 Correlation Matrix of the Parameter Estimates
Cohocton River

Some of the parameter estimates have very large coefficients of variation. In some instances these large coefficients of variation are caused by parameter interdependency. Table 7.10 presents the computed correlation coefficients between the parameter estimates. Notice that the highly correlated parameters, with the exception of α, remained at their starting values, which confirms the fact of the loglikelihood function not being sensitive to these parameters. The parameters with the highest correlations are precisely those which enter in the percolation function. We point out the 0.97 correlation coefficient between α and x_5^o, the -0.96 between x_3^o and $x_3(0)$, -0.94 between x_3^o and x_4^o, -0.92 between γ and x_5^o. As we explained in Chapter 4, these high correlations imply the occurence of ridges in the loglikelihood function, along which the value of the loglikelihood function does not change significantly. This explains the high variances of estimation of γ, α, x_3^o, x_4^o, x_5^o. The reason why α changed significantly from its starting value and still showed such high coefficient of variation is explained by the fact that the starting value of α was probably far from the ridge in which α lies. Once the optimization process started, α was trapped by the ridge which produced the high correlation shown.

Due to the high variances of estimation of the percolation function parameters, it is not surprising to see that the point of convergency for these parameters is very different from the value calibrated by the NWS. In particular, the 0.8 estimate of α not only is different from the NWS estimate, but also changes the curvature of the percolation function, as shown in Figure 7.16. For $\alpha>1$ the percolation rate tends to approach the percolation base rate faster than for $\alpha=1$. The opposite is true for $\alpha<1$. The net effect is that a higher percolation rate keeps the upper zone free water element at lower levels, thus decreasing the volumes of interflow and surface runoff. Notice that since this result compensates the high volumes of surface runoff and interflow which are due to having a small x_1^o, we would expect to find some degree of interdependency between x_1^o and α. However, the correlation coefficient measures only the degree of <u>linear</u> dependency, without shedding any light on the occurrence of non linear dependencies. We may say that, in general, the value for α should be larger than one, in order to resemble the actual percolation process ocurring in nature.

In the application to the Cohocton River, the model displayed little sensitivity to the initial conditions of the state variables as can be inferred from their coefficients of

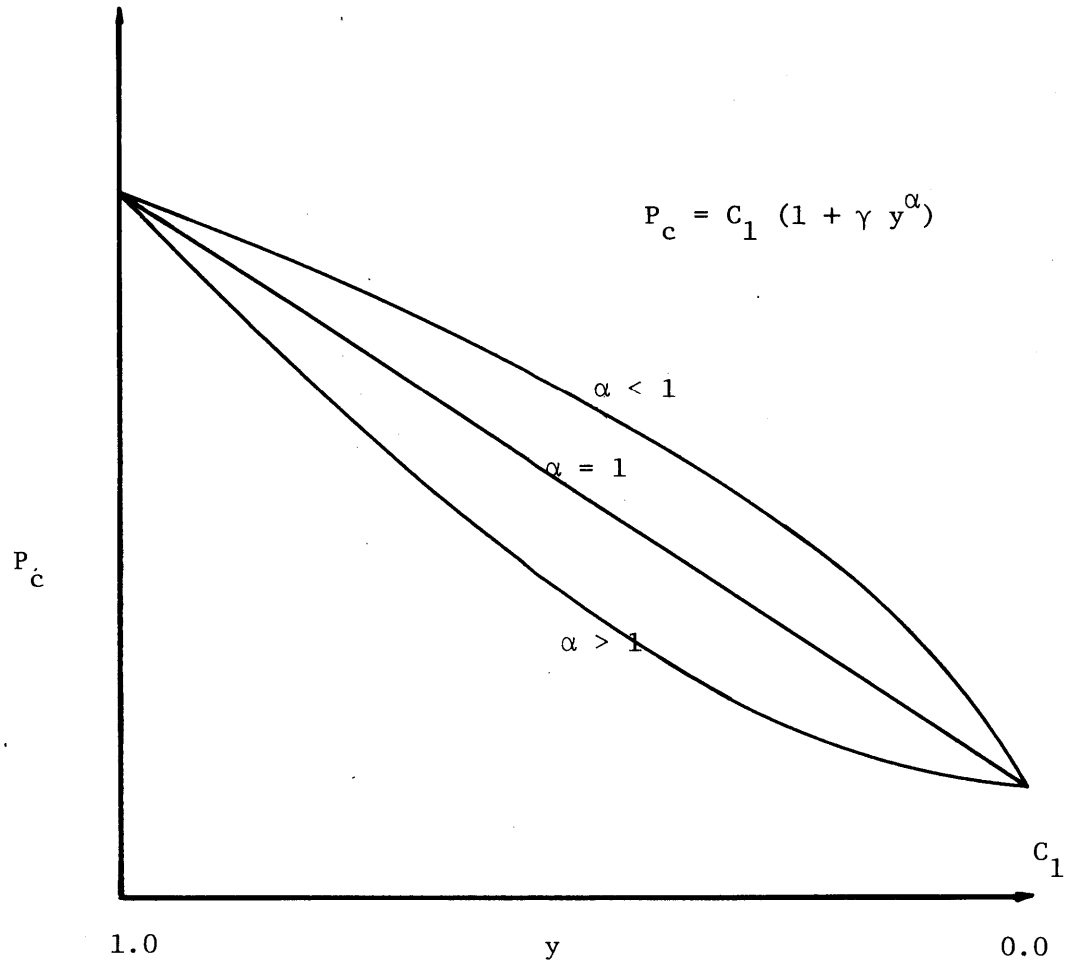


Figure 7.16 Percolation Function.

variation. Among these, $x_1(0)$ and $x_3(0)$ are highly correlated with the corresponding threshold parameter, x_1^o and x_3^o, which explains the high coefficients of variation for those estimates. Since no significantly high correlations were found between the remaining initial conditions estimates and other parameter estimates, the high coefficient of variation indicates that the loglikelihood is not sensitive to these initial conditions. In fact, if we examine the contribution to the runoff from x_2, x_4 and x_5, (interflow, primary base flow and supplementary base flow), at the beginning of the calibration interval, we find that the primary contributor is the supplementary base flow (73%), followed by the interflow (22%), while the primary base flow is just a 5% of the total runoff. This implies that, given the small contribution of the primary base flow to the total runoff, the model can not be sensitive to that initial condition, during the period of calibration. The value of x_2 is very small to begin with, which implies that the element will become dry quickly due to percolation and interflow. Therefore, the model is not very sensitive to this initial condition either.

Three parameters converged to zero, and since that value is quite close to their previous estimate, they are not reported as "Non Feasible" as was the case with some of the Bird Creek parameters.

7.3.1 Verification of the Parameter Estimates for Cohocton River.

Verification was performed with water year 1969-1970, which was the same year used by Kitanidis and Bras (1980) for the verification of the on line river forecasting procedures. We will be presenting the results corresponding to off line runs and on line runs in which the parameters of the model were estimated by the National Weather Service and by the maximum likelihood procedure. As before, we will be referring to these runs as OFNWS, OFML, ONNWS, ONML. The parameters of the stochastic part of the model used for the ONNWS simulation were those manually estimated by Kitanidis and Bras. The input data used for these runs is included in Appendix A. We selected four months of the year which we considered were representative of the features exhibited by the model running with the different sets of parameters. These months are December, 1969, January, July and August 1970. For the month of December, Figure 7.17, OFNWS predicts the rising limb and the timing of the peak of the hydrograph starting on the 10th about six hours earlier than the actual occurrence. The actual magnitude of the peak is over predicted by about 20%. The corresponding month for the OFML run, Figure 7.17 not only show the same earlier prediction of the rising limb of the hydrograph, but the peak is now over predicted by 508% of the

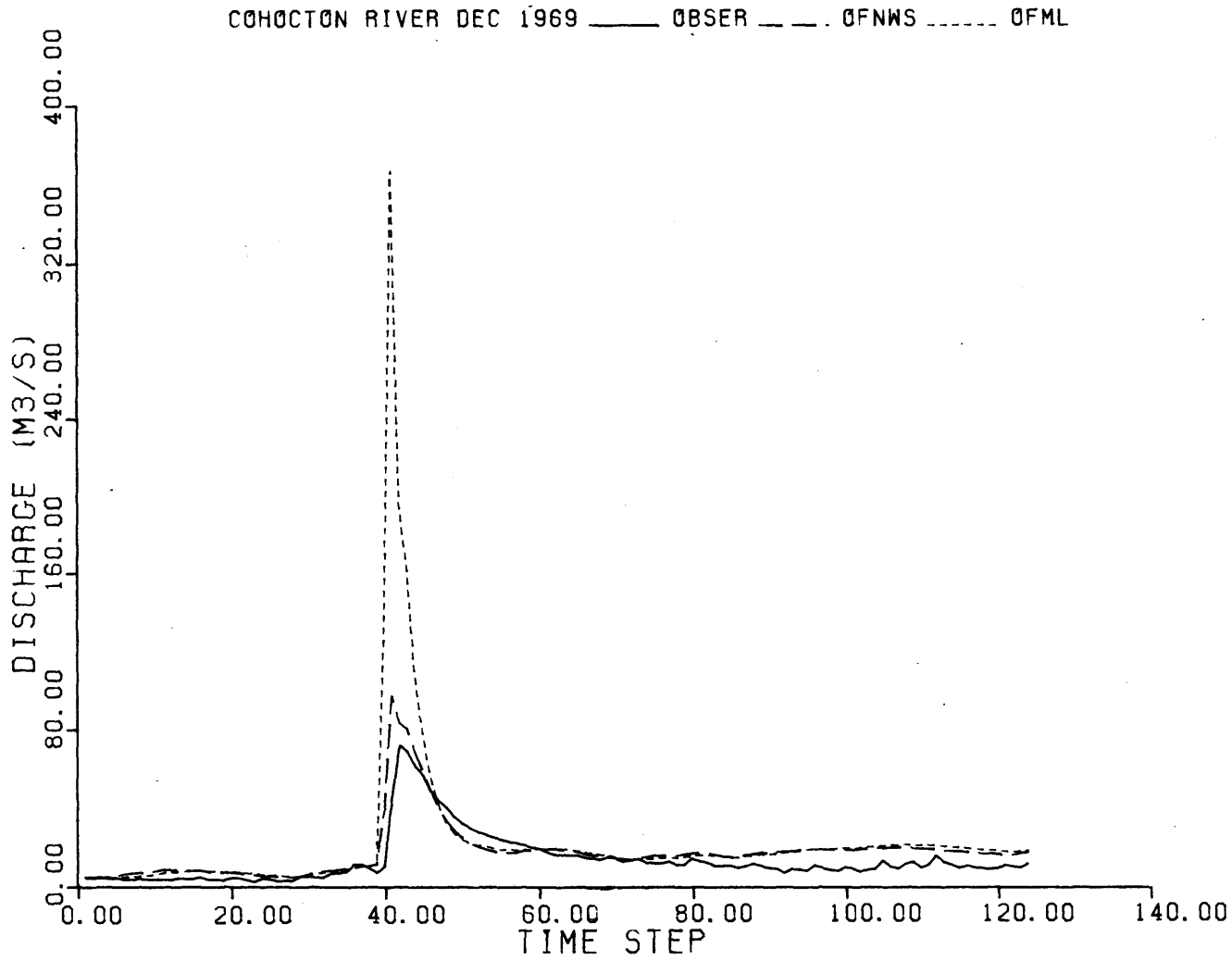


Figure 7.17 Six-Hour Lead Off-Line Forecasts for the Cohocton River, December 1969.

measured discharge. This is a case in which the role of the small estimated value of x_1^o is very important. The correction action taken by the filter is seen in the ONNWS and ONML runs, Figure 7.18. In the ONNWS case, the corrective action is taken two time steps after the rising started, producing a zig zagged hydrograph that, although odd in shape, reproduces well the falling limb. The groundwater recession curve is over estimated. A more pronounced zig zag effect is displayed by ONML, but the ground water recession curve is well estimated. The one step ahead predictions for the month of January with the off line runs are shown in Figure 7.19. Both results are quite similar, and over predict the discharge during most of the month. Towards the end of the month both OFNWS and OFML fail to predict the increase in the discharge. We speculate that in this case the errors may be due to ice jamming, to over prediction of the snow melt by the snowmelt model to freezing of the ground, or to a combination of these factors. The corrective action taken by the filter can be seen in Figure 7.20, for the ONNWS and ONML runs. In the former, the prediction of the receding discharge is corrected, although still shows a large correlation in the residuals. Towards the end, the model increases the predicted discharge but still fails to raise the discharge to the measured levels. A much better performance is displayed throughout the complete month by ONML. The month of July, 1970 presents a completely different behavior between the forecastings produced by the OFNWS and OFML runs. (Figure 7.21). The former essentially missed all the small hydrographs, while the latter over predicted them. Again, we believe the reason of this behavior is the wrong estimate of x_1^o. The fact that OFNWS misses the hydrographs may indicate, however, that the NWS estimate for x_1^o is larger than the real value, or that the real evapotranspiration was considerably smaller than the one calculated from the annual averages. In both cases, the effect is that x_1 is able to hold all the incoming precipitation without becoming saturated. The on line runs for the same month are shown in Figure 7.22. The ONNWS shows the hydrograph predictions lagging with respect to the observed ones, while the ONML displays the already familiar zig zag. Finally, the month of August, 1970 is the most important example of the effect of the small value of x_1^o. The off line runs for that month are presented in Figure 7.23. Towards the end of the month, at the end of the recession curve there is a measured discharge that peaks at 17.2 m^3/s. The OFNWS predicts a peak of 20.2 m^3/s, while OFML predicts a peak of 445 m^3/s. The ONNWS and ONML simulations are shown in Figure 7.24. The former misses the timing for the

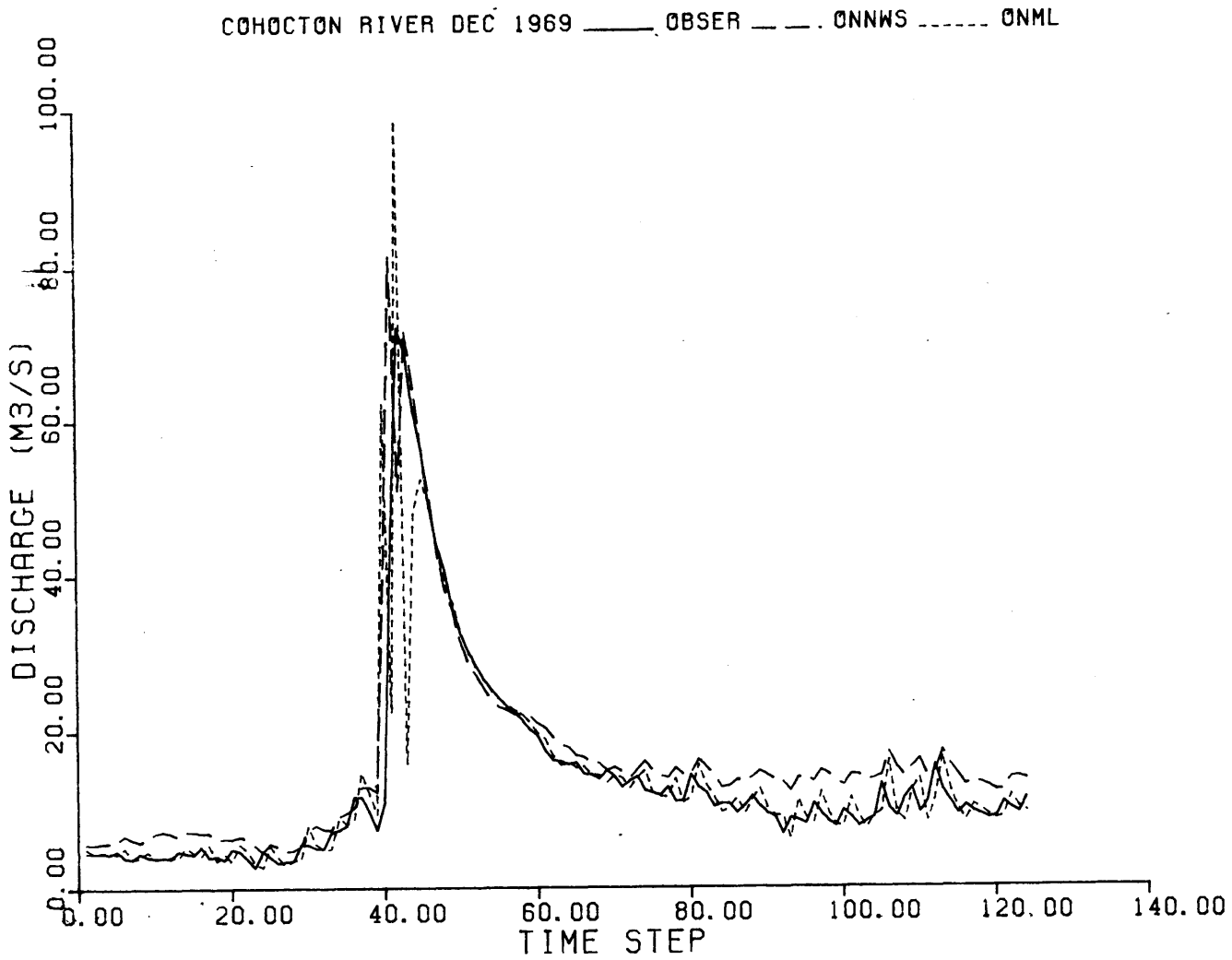


Figure 7.18 Six-Hour Lead On-Line Forecasts for the Cohocton River, December 1969.

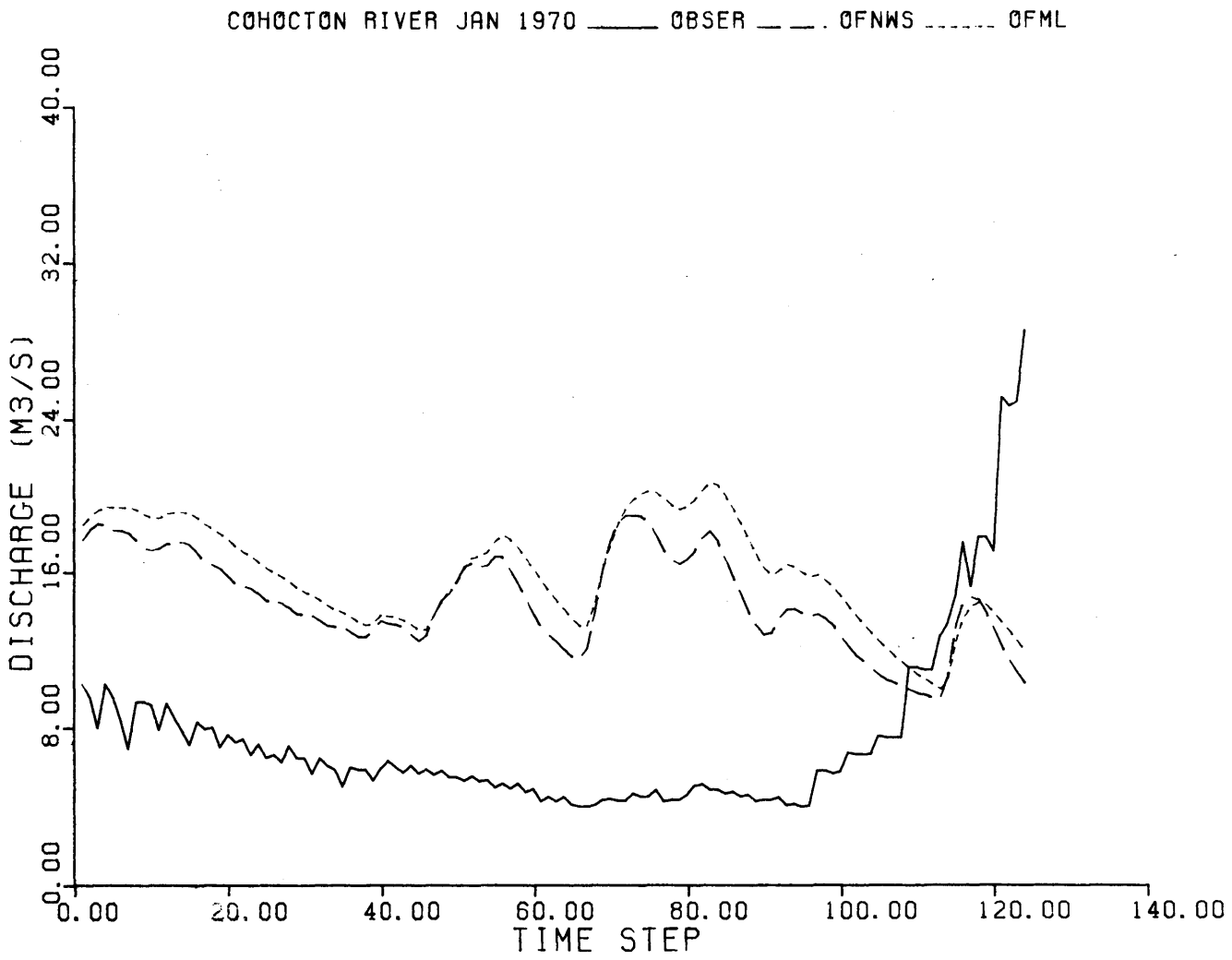


Figure 7.19 Six-Hour Lead Off-Line Forecasts for the Cohocton River, January 1970.

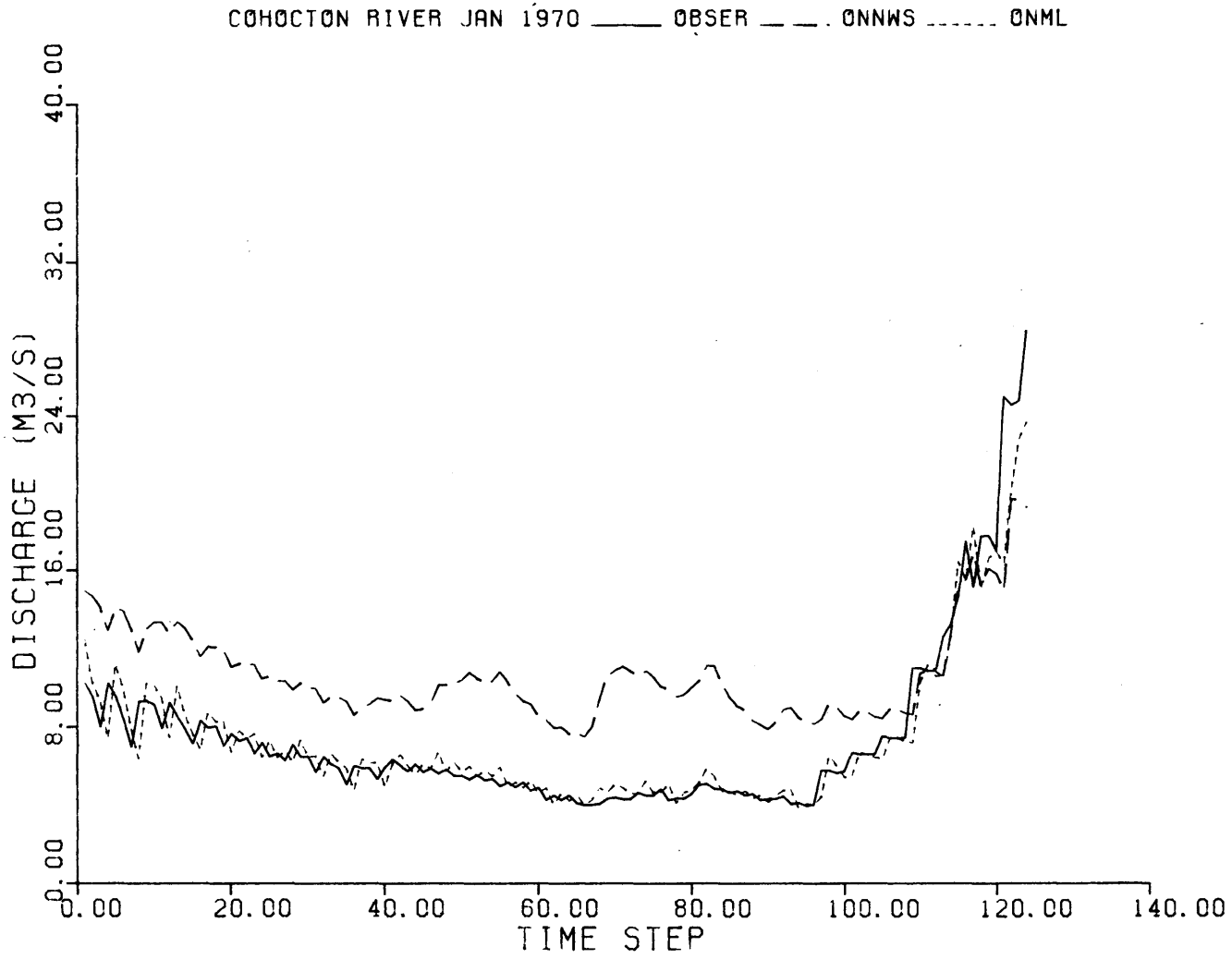


Figure 7.20 Six-Hour Lead On-Line Forecasts for the Cohocton River, January 1970.

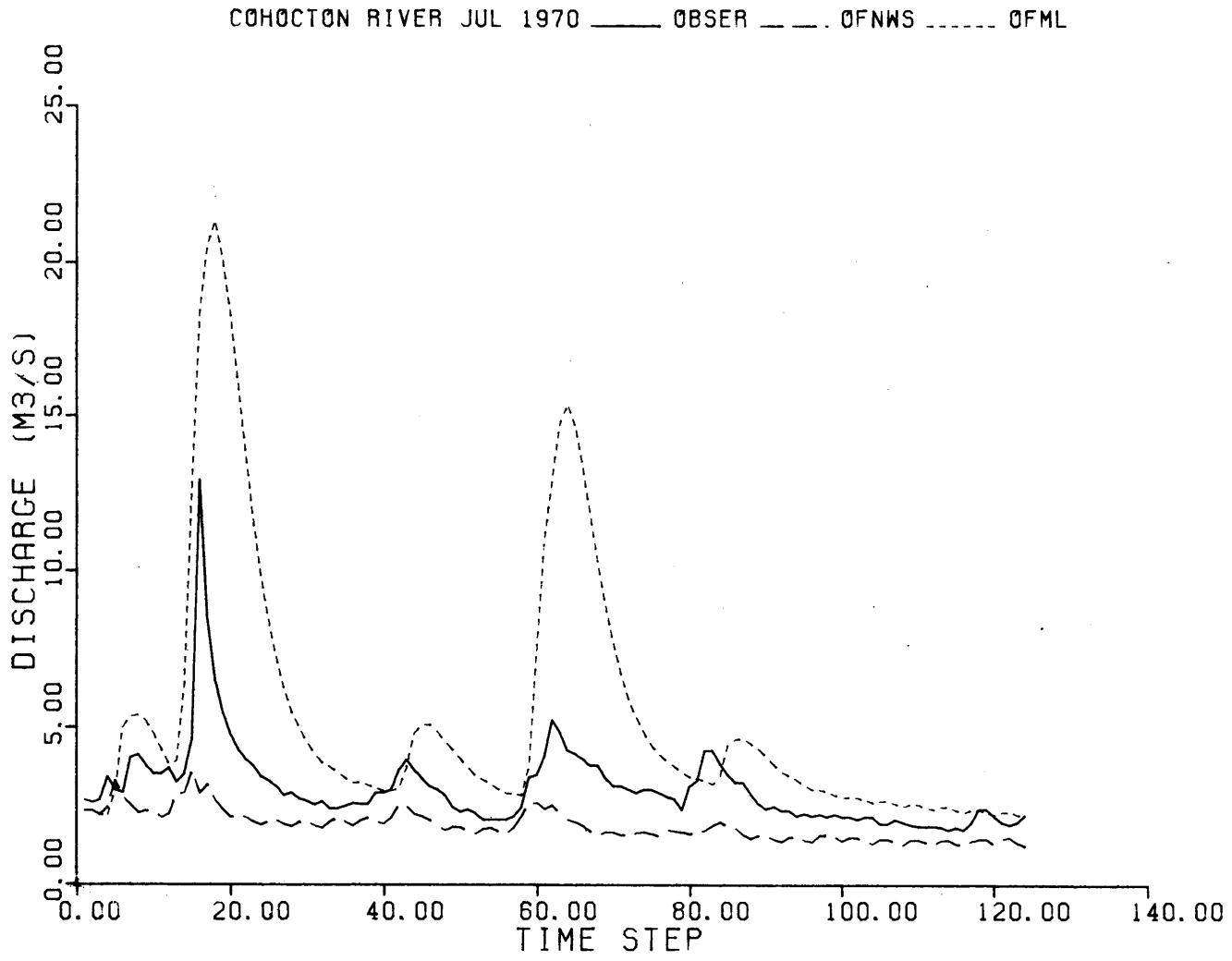


Figure 7.21 Six-Hour Lead Off-Line Forecasts for the Cohocton River, July 1970.

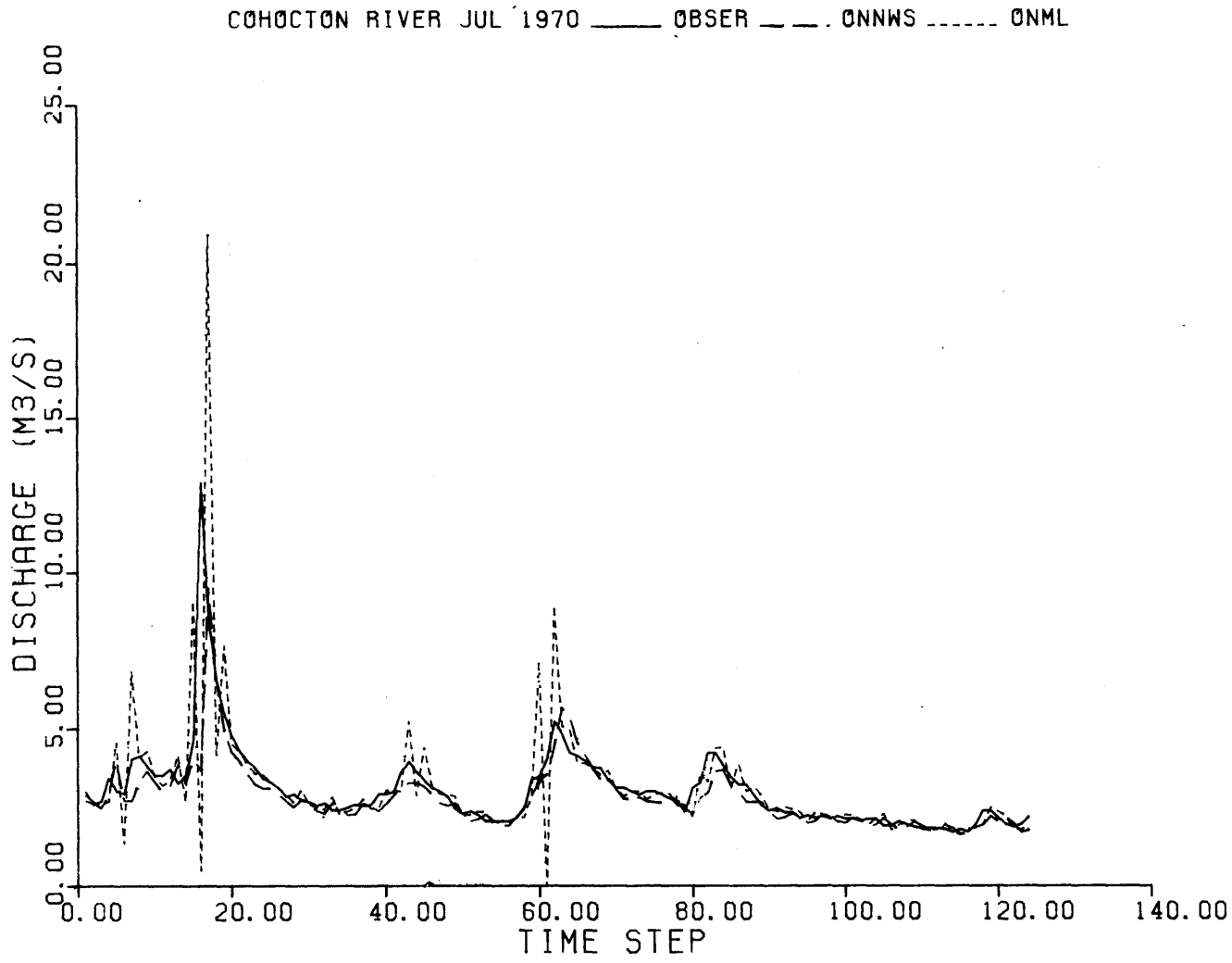


Figure 7.22 Six-Hour Lead On-Line Forecasts for the Cohocton River, July 1970

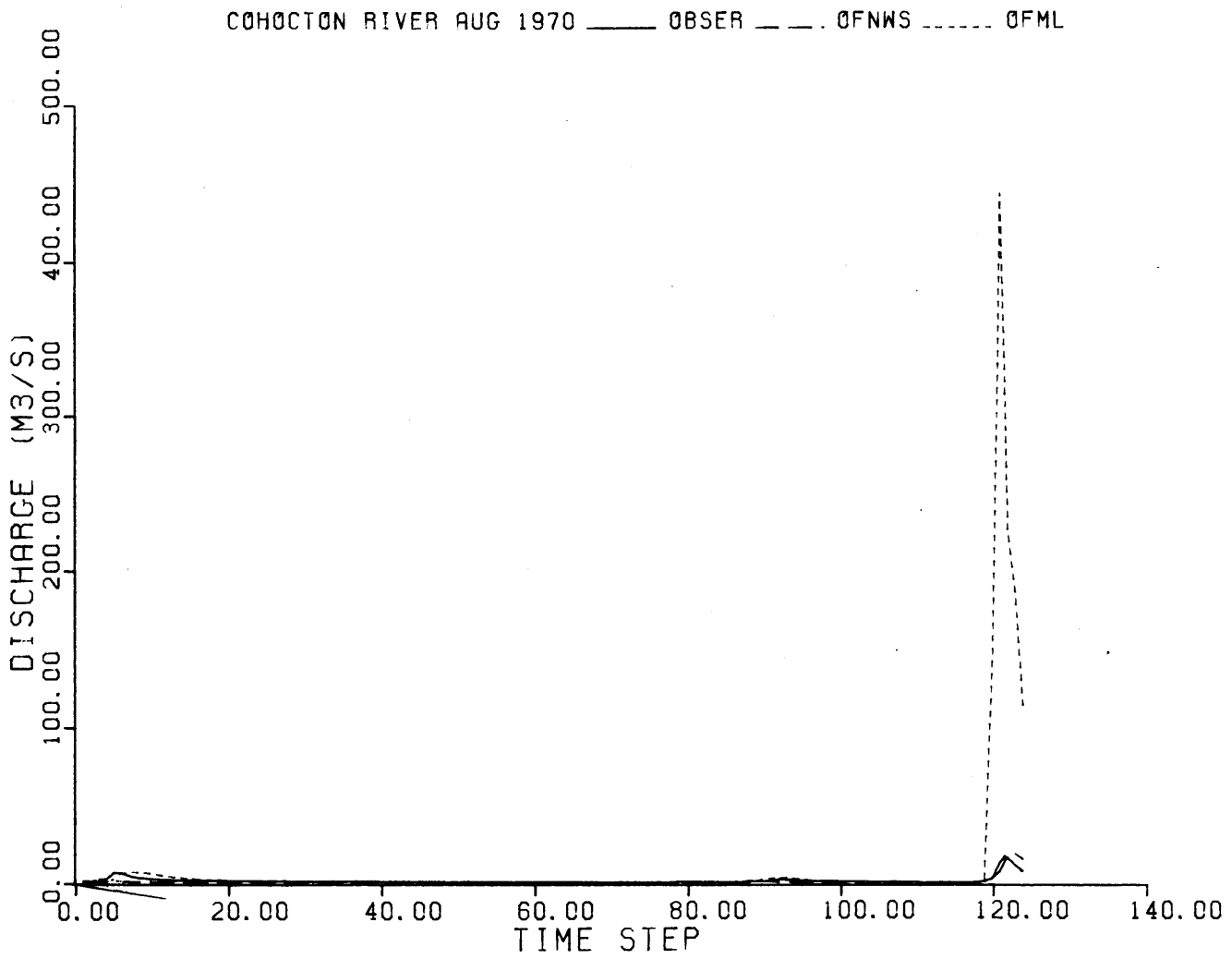


Figure 7.23 Six-Hour Lead Off-Line Forecasts for the Cohocton River, August 1970.

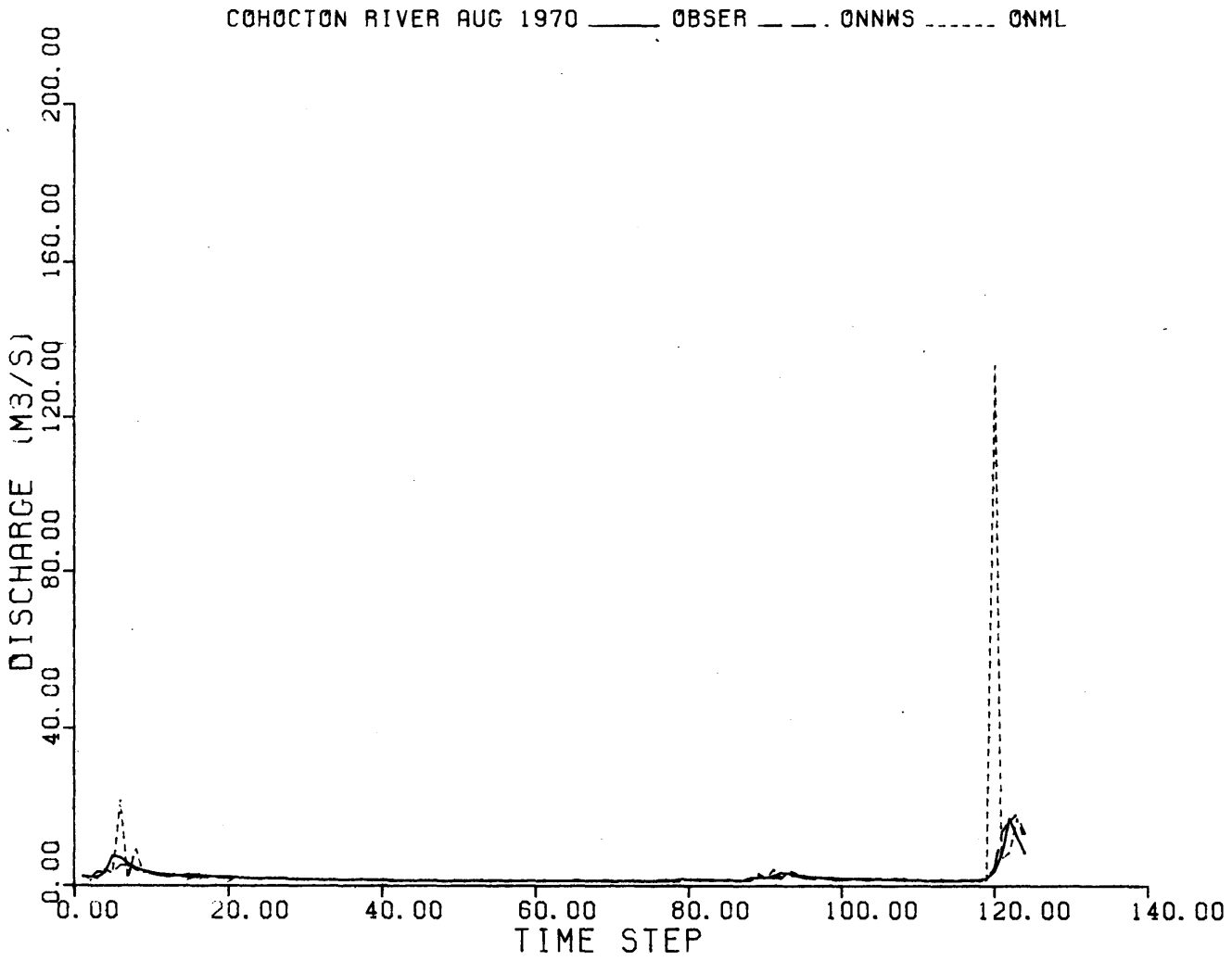


Figure 7.24 Six-Hour Lead On-Line Forecasts for the Cohocton River, August 1970.

peak, but the predicted hydrograph is considered "satisfactory". The latter over predicts the rising limb of the hydrograph, correct itself and predicts the rest of the hydrograph one step behind. We have included in Table 7.11 the computed values of the components of the runoff for the different models. We can see that the major difference between the models with the NWS estimated parameters and the models with the maximum likelihood estimated parameters is the amount of surface runoff. The simulated discharges with the different procedures for the remaining months of the water year 1969-1970 are included in Appendix B.

The performance summary of the off line runs, measured by the percent absolute error and the percent RMS error are included in Table 7.12. Although the bias and the percent bias are smaller in the ML case, the maximum error of 236 m^3/s is considerably larger than the maximum error with the NWS estimated parameters of 90.7 m^3/s. The percent absolute error and the percent RMS error with the NWS estimated parameters are also smaller than the corresponding values with the ML estimated parameters. The distribution of the percent RMS errors by interval of flow, are shown in Table 7.12-b, and are plotted in Figure 7.25. In general, the model with the original parameters without a filter out-performs the model with the maximum likelihood estimated parameters. The distribution of the errors by month is presented Table 13. The outstanding feature in these results is the 2111.7% RMS error for the month of August, with the maximum likelihood estimated parameters, which was already analyzed.

The statistical summary for the on line runs is presented in Table 7.14. In this case the run with the maximum likelihood estimated parameters out-performs the on line run with the NWS estimated parameters. The distribution of the errors per interval of flow is presented in Table 14 -b. With the exception of the very low flows, (0-2m^3/s day), the model with the maximum likelihood estimated parameters performed better than the model with the original parameters. These results are also presented in Figure 7.26. The distribution of errors by month, Table 7.15 shows that the first month has a very large error in both cases. This is due to the fact that the on line runs were started at that particular month with arbitrary initial conditions. If the first month is excluded from the computation of the annual average at the bottom of the table, the computed values for the RMS error would be 33.05% and 37.28% for the maximum likelihood estimated and the NWS estimated parameters, respectively.

7.4 Conclusions.

This chapter has presented the results of applying the maximum likelihood estimation estimation procedure to two basins in the United States. In the case of the application to

MONTH	TOTAL-RO	IMPV-RO	DIRECT-RO	SURF-RO	INTERFLOW	BASEFLOW (CHANNEL COMP)
Oct.	2.8	0.3	0.0	0.0	0.0	2.8
Nov.	20.5	0.4	0.3	0.2	8.7	11.0
Dec.	39.3	0.4	0.6	2.3	15.7	20.4
Jan	31.7	0.2	0.3	0.0	8.6	22.7
Feb.	36.5	0.2	0.4	0.0	12.6	23.3
Mar.	20.1	0.1	0.1	0.0	3.4	16.5
Apr.	13.9	0.2	0.1	0.0	2.3	11.5
May	11.8	0.3	0.2	0.1	4.3	7.3
June	8.3	0.2	0.0	0.0	0.0	8.6
July	4.0	0.4	0.0	0.0	0.0	4.1
Aug.	4.3	0.4	0.0	0.0	1.4	3.0
Sept.	7.9	0.3	0.1	0.0	2.4	5.5
TOTAL	201.2	3.5	2.0	2.6	59.4	136.7

a) OFF LINE NWS PARAMETERS

MONTH	TOTAL-RO	IMPV-RO	DIRECT-RO	SURF-RO	INTERFLOW	BASEFLOW (CHANNEL COMP)
Oct.	4.9	0.0	0.0	0.0	2.2	2.8
Nov.	30.5	0.0	0.0	8.0	14.4	8.1
Dec.	50.0	0.0	0.0	15.4	20.3	14.3
Jan.	34.3	0.0	0.0	0.0	15.6	18.6
Feb.	43.4	0.0	0.0	0.4	23.8	19.2
Mar.	20.8	0.0	0.0	0.0	6.2	14.5
Apr.	16.9	0.0	0.0	0.1	5.7	11.2
May	18.7	0.0	0.0	2.4	8.0	8.3
June	8.2	0.0	0.0	0.0	0.3	8.2
July	11.6	0.0	0.0	0.1	5.7	6.0
Aug.	27.0	0.0	0.0	19.4	3.9	4.1
Sept.	11.5	0.0	0.0	0.1	6.7	4.9
TOTAL	277.8	0.0	0.0	46.0	112.9	120.1

b) OFF LINE ML PARAMETERS

MONTH	TOTAL-RO	IMPV-RO	DIRECT-RO	SURF-RO	INTERFLOW	BASEFLOW (CHANNEL COMP)
Oct.	16.3	0.3	0.0	0.0	2.2	14.1
Nov.	32.3	0.4	0.4	0.5	13.1	18.0
Dec.	45.9	0.4	0.6	2.8	17.9	24.2
Jan.	34.8	0.2	0.3	0.0	9.6	24.8
Feb.	39.3	0.2	0.4	0.0	13.7	24.9
Mar.	21.3	0.1	0.1	0.0	3.6	17.5
Apr.	20.1	0.2	0.1	0.0	3.4	16.6
May	14.9	0.3	0.2	0.1	4.6	10.0
June	9.4	0.2	0.0	0.0	0.0	9.8
July	5.4	0.4	0.0	0.0	0.4	5.1
Aug.	5.1	0.4	0.0	0.0	1.5	3.6
Sept.	8.1	0.3	0.1	0.0	2.3	5.7
TOTAL	252.8	3.5	2.2	3.5	72.2	174.4

C) ON LINE NWS PARAMETERS

MONTH	TOTAL-RO	IMPV-RO	DIRECT-RO	SURF-RO	INTERFLOW	BASEFLOW (CHANNEL COMP)
Oct.	20.0	0.0	0.0	13.0	3.6	3.5
Nov.	17.5	0.0	0.0	1.2	8.3	8.0
Dec.	37.1	0.0	0.0	8.2	14.9	14.1
Jan	17.2	0.0	0.0	0.0	7.4	9.8
Feb.	41.9	0.0	0.0	3.0	19.1	19.8
Mar.	35.8	0.0	0.0	0.0	19.4	16.3
Apr.	73.0	0.0	0.0	1.4	44.6	27.0
May	26.2	0.0	0.0	1.2	7.8	17.4
June	10.7	0.0	0.0	0.0	0.9	10.1
July	7.9	0.0	0.0	0.0	2.8	5.4
Aug.	23.0	0.0	0.0	17.4	2.6	3.6
Sept.	5.9	0.0	0.0	0.0	2.0	4.3
TOTAL	316.1	0.0	0.0	45.4	133.4	139.2

d) ON LINE ML PARAMETERS

Table 7.11 Distribution of the Total Runoff Components by Month for the Different Simulations

Annual Averages	NWS	ML
Simulated Mean (m^3/s day)	7.715	10.656
Observed Mean (m^3/s day)	11.087	11.087
Percent Bias	-30.41	-3.89
Bias (mm)	-87.376	-11.167
Maximum Error (Sim-Obs)(m^3/s day)	-90.716	236.087
Percent Absolute Error	63.89	80.68
Percent RMS Error	135.07	195.85

a) Daily Volume Error Statistics

Flow Interval (m^3/s day)	Number of cases	Percent RMS Error NWS	ML
0.00 - 2.00	68	21.97	155.48
2.00 - 5.00	98	71.25	191.64
5.00 - 10.00	84	56.74	106.00
10.00 - 20.00	64	47.02	240.00
20.00 - 30.00	19	60.94	85.73
30.40 - 40.00	12	67.46	70.61
40.00 and above	20	80.17	101.35

b) Distribution of Errors by Flow Interval.

Table 7.12 Statistical Analysis of Prediction of Daily Volumes of Runoff with Off Line Simulations. Cohocton River, 1969-1970.

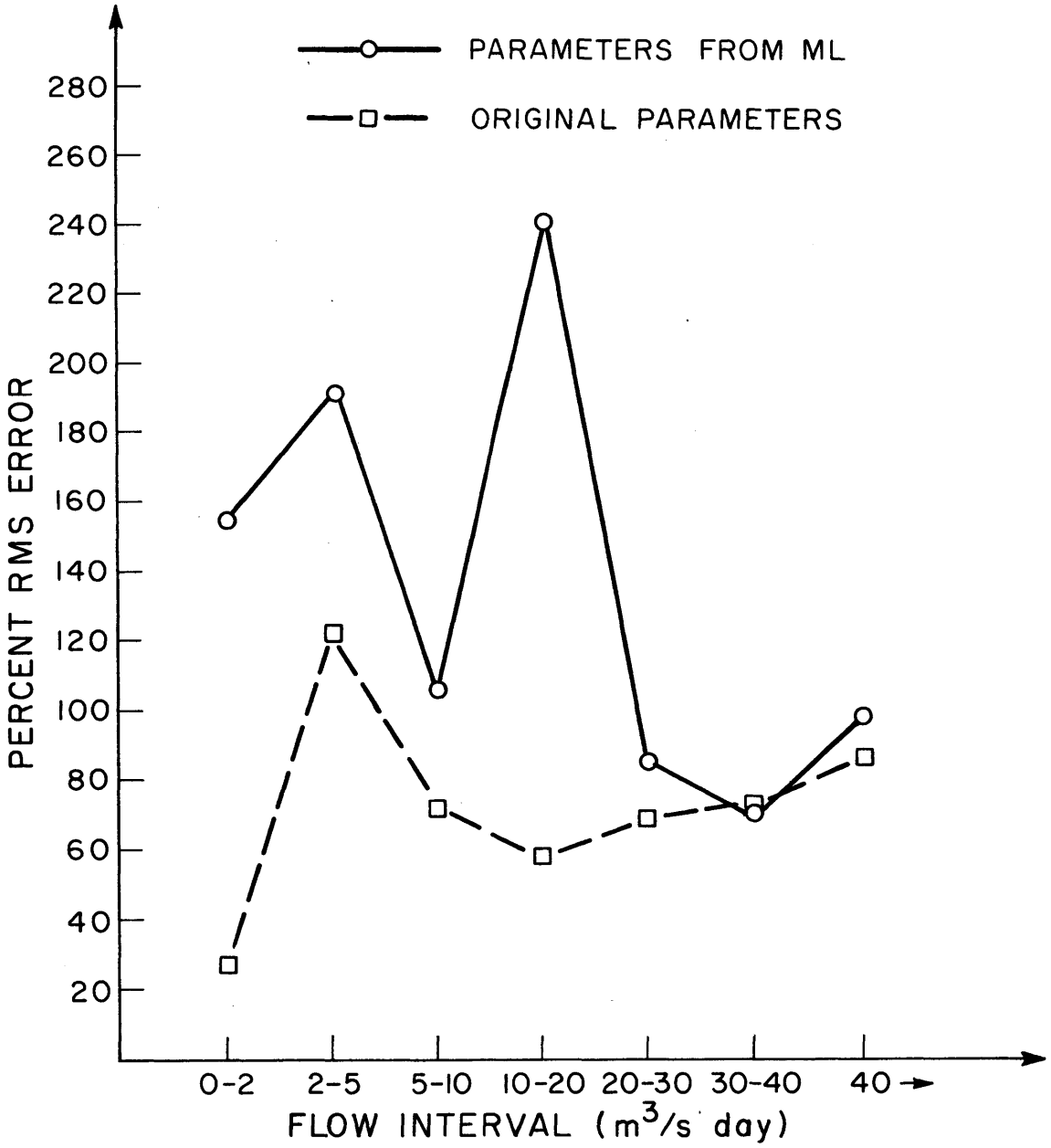


Figure 7.25 Percent RMS Error by Interval of Flow. Cohocton River, Off-line Runs.

	National Weather Service		Maximum Likelihood	
MONTH	Percent Abs Err	Percent RMS	Percent Abs Err	Percent RMS
October	19.18	25.03	54.22	122.64
November	51.36	105.07	129.92	288.84
December	42.42	53.96	79.21	219.31
January	116.76	124.75	131.61	140.35
February	45.74	63.41	50.52	67.79
March	59.48	126.34	57.83	122.69
April	84.57	96.60	81.34	93.87
May	59.74	63.20	54.29	72.09
June	22.08	30.89	21.91	30.07
July	37.66	46.44	80.35	134.63
August	34.05	69.26	434.77	2111.70
September	62.20	102.34	169.76	390.45
Average	52.99	75.56	112.14	316.20

Table 7.13. Percent Absolute Error and Percent RMS Error for the Off Line Runs by Month.
Cohocton River, Water Year 1969-1970.

Annual Averages	NWS	ML
Simulated Mean (m^3/s day)	10.405	10.787
Observed Mean (m^3/s day)	11.087	11.087
Percent Bias	-6.15	-2.70
Bias (mm)	-17.676	-7.762
Maximum Error (Sim-Obs)(m^3/s day)	-42.831	-40.932
Percent Absolute Error	30.32	11.27
Percent RMS Error	58.90	35.99

a) Daily Volume Error Statistics

Flow Interval (m^3/s day)	Number of cases	Percent RMS Error NWS	ML
0.00 - 2.00	68	127.59	275.87
2.00 - 5.00	98	75.24	15.62
5.00 - 10.00	84	35.23	14.32
10.00 - 20.00	64	29.88	17.11
20.00 - 30.00	19	28.01	8.80
30.40 - 40.00	12	31.93	17.74
40.00 and above	20	36.79	21.64

b) Distribution of Errors by Flow Interval.

Table 7.14 Statistical Analysis of Prediction of Daily Volumes of Runoff with On Line Simulations. Cohocton River, 1969-1970.

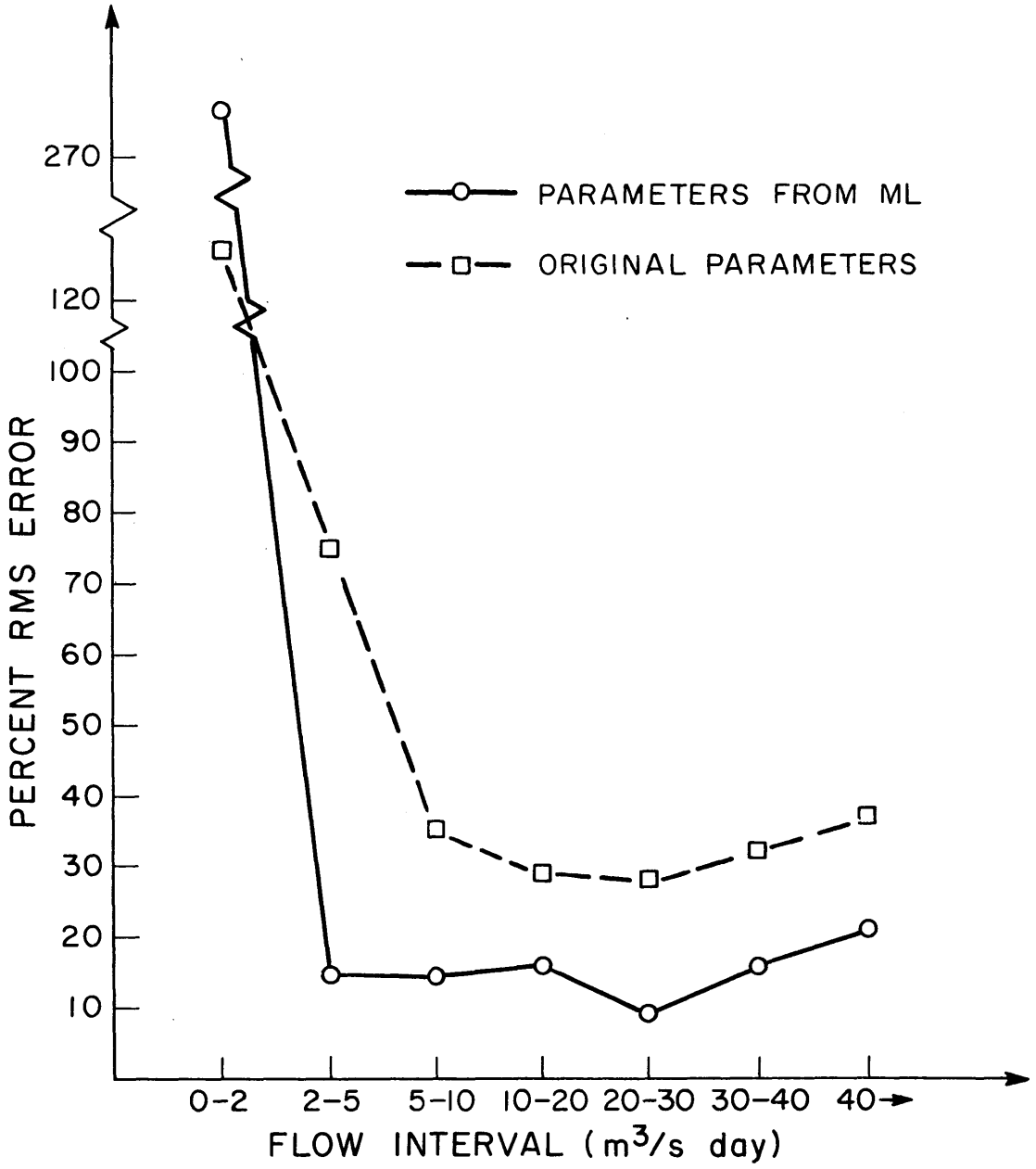


Figure 7.26 Percent RMS Error by Interval of Flow. Cohocton River, On-line Runs.

	National Weather Service		Maximum Likelihood	
MONTH	Percent Abs Err	Percent RMS	Percent Abs Err	Percent RMS
October	148.52	186.71	69.46	332.11
November	54.21	79.69	8.30	14.19
December	25.64	29.24	8.82	16.37
January	55.45	59.66	7.21	14.32
February	18.46	25.34	9.13	13.89
March	26.88	55.70	14.55	40.34
April	34.42	40.19	11.27	19.39
May	19.72	21.22	3.30	4.46
June	8.20	10.74	5.43	6.25
July	5.99	10.31	6.66	13.54
August	14.82	40.77	60.31	204.30
September	25.56	37.13	9.77	16.48
Average	36.48	49.73	17.85	57.97

Table 7.15. Percent Absolute Error and Percent RMS Error for the On Line Runs by Month.
Cohocton River, Water Year 1969-1970.

Bird Creek, it was shown how structural errors in the model of the channel cause some of the parameters to converge to unrealistic values. The use of prior information about some of the parameters decreased the variance of estimation of the remaining parameters. The bias in the parameters, introduced by the channel model caused the model to perform worse than the previously estimated parameters, for off line and on line runs. The application to the Cohocton river detected a series of strong correlations between several of the parameters, all of them associated with the percolation function. These high correlations were the cause of very high variances in the parameter estimates. This application also showed that one of the parameters, x_1^o, converged to the maximum likelihood estimate of that parameter corresponding to the period of data used in the estimation. This estimate, however, proved to be smaller than the best estimate of the parameter in the verification runs, since the model consistently over predicted the amount of interflow and surface runoff. These over predictions forced the on line model to correct the hydrographs in a zig zagued fashion, which although mathematically correct is hydrologically unappealing.

We found the the maximum likelihood procedure to behave very differently in each of the two basins tested. In the application to the Cohocton River, we found that x_1^o converged to an optimum point which applied only to the calibration period, while the estimate of the same parameter for Bird Creek checked very well with the NWS estimate of the same parameter. The very high correlation between several parameters in the Cohocton River case was not found in the Bird Creek application. In the latter, the absense of high correlations between the parameter estimates must be cautiously evaluated because of the structural errors in the channel.

The length of the data interval used in the calibration of Bird Creek had a strong influence in the parameter estimates and in the variances of these estimates. Due to the very high cost of performing the experiments several questions remain unanswered, such as the influence of of the initial guess of the parameters on the final parameter estimates. We do not know whether the correlation between the percolation function parameters will re-appear when applying the maximum methodology to different basins. The effect of the non stationarity of the model error terms, and of the independence between the same error terms on the parameter estimates remains to be studied. These points will be addressed again in the next chapter, Conclusions and Recomendations.

CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS

8.1 Summary of Results

This work has studied the application of maximum likelihood, in conjunction with the use of prior information on some of the parameters, to the estimation of the parameters of the National Weather Service River Forecast System soil phase model.

The prior information is built into the stochastic parameter estimation procedure in a form that permits subjective information to be used in a mathematically sound form in combination with information coming from the discharge records.

A simple stochastic model is proposed for modeling the base flow of a basin. This model is used in an iterative fashion to determine the parameters of the base flow discharge coefficients, and at the same time identify the interval of discharges that corresponds to base flow. In this form prior information about these parameters is obtained which can be used in the global estimation procedure. The model and the procedure are tested in two basins in the United States. The robustness of the algorithm in identifying the time interval of base flow activity is shown with an example of how the model rejects an interval that "looks like" an interval of base flow.

A non linear constrained optimization procedure, based on the Davidon-Fletcher-Powell method and on a gradient projection algorithm is developed. The algorithm had to be adapted to the special characteristics of the rainfall runoff model, in order to prevent convergence of the parameters to points far from the global optimum, by selectively including or excluding parameters from the linear searches that are part of the Davidon-Fletcher-Powell algorithm.

In order to minimize the implementation costs of the parameter estimation procedure a simplified version of the NWSRFS model has been proposed. In this simple version the threshold curves which cause the highest non linearities in the original system have been replaced by smooth "S" curves. The simplified model, by keeping the general structure of the original model, helps in finding a major source of discontinuities in the loglikelihood function. These discontinuities cause the parameters to converge to local optima. A change in the criterion for calculation of the number of time subintervals required for the integration of the non linear soil moisture equations is proposed as a solution to the problem. A very high correlation between the estimates of two of the parameters of the model is traced to the "S" curve corresponding to the upper zone tension water element. This indicates that some modifications to the incorporation of the "S" curve into the model must be carried on, if the

simplified model is to be used in the prediction of real flows. These modifications are outlined in the next section.

The maximum likelihood parameter estimation procedure was applied to two basins with different physiographic characteristics. In the first case, the application to Bird Creek was seriously affected by biases in the parameters introduced by the model of the channel. The use of prior information about two of the parameters produced an overall reduction on the estimation variance of the remaining parameters. In the application to the Cohocton River some of the parameters were unidentifiable during the period of calibration. Very high interdependency was found between several parameters. Especially afected were some of the parameters entering in the percolation function. Furthermore, it was shown that, due to the unobservability of the upper zone tension water content, the model was sensitive to the difference between the maximum upper zone tension water content and the initial condition of the associated state variable.

8.2 Recommendations

In order to eliminate the biases in the parameter estimates due to inaccuracies in the precipitation and snow melt model, and to the channel model, the state variables of the three models, (precipitation-snow melt, soil phase and channel) must be jointly estimated, and the parameters of the three models should be jointly identified.

There is recent evidence in the literature (Kuczera, 1982) on the improvement on the parameter estimates by using additional time series data in addition of the input-output series currently being used. Kuczera studied the effect of groundwater depth data averaged over two months in the parameter estimates of a simple two-state conceptual model used to predict two-month averaged discharges. This author concluded that the use of groundwater depth information helped in obtaining consistent estimates for one of the groundwater related parameters. Although groundwater depth supplies little information for six hour runoff forecasting, the same principle can be applied here. In particular, an effort should be made of including snow depth information and satellite measurements of areal averaged short wave radiation in the 1.55 cm wave length band. The former will improve the estimation of snow melt. The second has been shown to be correlated with the soil moisture of the upper layers of the soil, (Blanchard et al., 1981). For application to the NWSRFS model, the short wave emissivity data must be first related to the upper zone tension water content and to the upper zone free water content. In this form the problems of non observability of the upper zone tension water content and the identifiability of the upper zone tension water maximum may be solved. In addition, the additional information will hopefully improve the estimates of the remaining parameters of the soil phase model, since the direct measurements of the state variables may elim-

inate the inconsistent estimates obtained due to biases in the channel model.

Since the availability of satellite data as a routine may take many years to be implemented, short term solutions to the problem of observability of the upper zone tension water element must be found. In its present form, the NWSRFS model is an accurate conceptualization of the soil moisture movement. But, is it really worthy to have a good description of a process whose parameters can not be estimated.? Modifications of the model's current structure must be studied, such that the upper zone tension water element is permanently observable. Here we have a trade-off between the high accuracy of a model with some parameters which are non identifiable, and a less accurate model which has identifiable parameters and observable state variables. One such modification of the upper zone tension water element which will make this element observable has been made in the HBV model (Fjeld and Aam, 1980, Bergstrom, 1975, Bergstrom and Forsman, 1973). In this model the upper zone tension water permanently drains into the upper zone free water element, according to a power function of the relative fullness of the upper zone tension water element. In this form, that state variable is permanently observable, which will hopefully solve the problem of the identifiability of x_1^o.

Further research into the interaction of the percolation function parameters is necessary, by applying the maximum likelihood parameter estimation procedure to other basins. If the interaction between these parameters persists, the percolation function must be modified such that the redundancies are eliminated.

The stochastic model of base flow can be enhanced to account for constant water losses from the channel. One approach is outlined here. The number of state variables remains in two, and the number of unknown parameters to be estimated by maximum likelihood is increased by one. The governing equations for the proposed approached are included in Appendix C.

The ideas of Section 5.6.1 regarding the use of the variation in the estimate of the model error term to define more accurately the interval of base flow must be tested in different basins.

The modifications which were required in the non linear optimization algorithm converted the method into a blend of pattern search and gradient based techniques. No comparison exists between the performance of this mixed procedure and the performance of a pattern search technique. This comparison should be carried out and the procedure changed if a pure pattern search technique proves to be more efficient in achieving global convergency then the algorithm here implemented.

The sensitivity of the point of convergency to the initial estimates of the parameters must also be studied, since the possibility of convergence to local optima exists.

The importance of the stochasticity of the potential evapotranspiration and the non stationarity of the model noise in the state estimation and in the parameter identification must be investigated. The former can be modeled similarly to the rainfall plus snowmelt input to the soil phase model. The latter can be related to the state variables such that the variance of the model error terms is functionally related to the state variables themselves.

Finally the high cost of using maximum likelihood to identify the parameters of the NWSRFS model was a major reason for leaving many unanswered questions. Least expensive, though approximate ways, of calculating the loglikelihood function, such as the use of constant gains for the filtering of the states and propagation of the covariance must be seriously considered and evaluated. It must be determined if the model is consistently insensitive to some of the parameters. If this is the case, the insensitive parameters must be identified and given nominal values. If it is determined that the model is relatively insensitive to some of the parameters, those parameters can be roughly estimated by a less expensive method such as weighted least squares, after which the most sensitive parameters must be "fine tuned" by maximum likelihood.

REFERENCES

Astrom, K. and Eykhof, D., "System Identification, a Survey", IFAC Symposium on Identification and Process-Parameter Estimation, Prague, June 1970.

Beck, M. B., "Maximum Likelihood Identification Applied to DO-BOD-Algae Models for a Freshwater Stream", Report 7431(c) Lund Institute of Technology, Division of Automatic Control, Sweden, 1974.

Beck, M. B., "Identification and Parameter Estimation of Biological Process Models", paper presented at the IIASA/WMO Workshop on Recent Developments in Real-Time forecasting/Control of Water Resource Systems, October 18-21, Laxenberg, Austria, 1976.

Benjamin. J. R. and Cornell, C. A. Probability Statistics and Decision for Civil Engineers, McGraw Hill Book Company 1970

Bergstrom, S., "The Development of a Snow Routine for the HBV-2 Model", Nordic Hydrology, Vol. 6 No. 2, 1975.

Bergstrom, S. and Forsman, A. "Development of a Conceptual Deterministic Rainfall Runoff Model", Nordic Hydrology, Vol. 4, No. 3, 1973.

Blanchard, B. J., McFarland, M. J., Schmugge, T. J. and Rhades, E., "Estimation of Soil Moisture with API Algorithms and Microwave Emission", Water Resources Bulletin, pp. 767-774, Vol 17 No. 5, October 1981.

Boughton, W. C., "A New Simulation Technique for Estimating Catchment Yield", Manly Vale, N.S.W., New South Wales University, Water Researcg Laboratory, Rep. #78, 1965.

Boultot, F. and Dupriez, G. L., "Conceptual Model for an Average-Sized Catchment Area, II Estimate of Parameters, Validity of Model, Applications', Journal of Hydrology, V29, 3/4, pp. 273-292, April 1976.

Bras R. L. and Restrepo-Posada, P. J. "Real Time, Automatic Parameter Calibration in Conceptual Runoff Forecasting Models", presented at the Third International Symposium on Stochastic Hydraulics, Tokyo, Japan, Aug. 5-7, 1980.

Chapman, T. G., "Optimization of a Rainfall-Runoff Model for an Arid Zone Catchment", presented at the symposium of the Results of Research on Representative and Experimental Basins, Wellington, N. Z., Dec. 1-8, 1970, Publ No. 97, AIHS, 1972.

Dawdy, D. R. and O'Donnell, T., "Mathematical Models of Catchment Behavior", Proc. ASCE, Vol 91 No. HY4, pp.123-137, 1965.

Dooge, J. C. I., "A New Approach to Non-Linear Problems in Surface Water Hydrology: Hydrologic Systems with Uniform Non-linearity", IASH Pub. 76 409-413, 1967.

Edwards, A. W. F. Likelihood, Cambridge University Press, 1972.

Fjeld M. and Aam, S. "An Implementation of Estimation Techniques to a Hydrological Model for Prediction of Runoff to a Hydroelectric Power Station", IEEE Trans. on Autom. Control, Vol. AC-25, No. 2, pp.151-163, April 1980.

Georgakakos, K. P. and Bras, R. L. "A Statistical Linearization Approach to Real Time Non-Linear Flood Routing", M. I. T. Ralph M. Parsons Lab. T. R. No. 256, June 1980.

Georgakakos, K. P., Restrepo-Posada P. J. and Bras, R. L., "On-Line River Discharge Forecasting Using Filtering and Estimation Theory" Unpublished Progress Report, January 1980.

Gelb, A. ed. Applied Optimal Estimation. MIT Press, Cambridge Massachusetts, 1974.

Grunewald, V. and Dick, S., "On the Application of Optimization Techniques to Conceptual Catchment Models", Mathematical Models in Hydrology, Proc. Warsaw Symposium IAHS/UNESCO, Poland, July 1971.

Haan, C. T., "A Water Yield Model for Small Watersheds", Water Resources Research, Vol 8, No. 1, pp.58-69, Feb 1972.

Ibbit, R. P. and O'Donnell, T., "Fitting Methods for Conceptual Catchment Models", ASCE Proc. Journal of Hydraulics Division, Vol 97, No. HY9, pp. 1331-1342, Sept. 1971.

Johnston P. R. and Pilgrim, D. H., "Parameter Optimization for Watershed Models", Water Resources Research, Vol. 12, No. 3, pp 477-486, June 1976.

Kitanidis, P. K. and Bras, R. L., "Real-Time Forecasting with a Conceptual Hydrologic Model. I. Analysis of Uncertainty", Water Resources Research, Vol. 16, No. 6, pp. 1025-1033, December 1980.

Kleinecke, D., "Use of Linear Programming for Estimating Geohydrologic Parameters of Groundwater Basins", Water Resoures Research, Vol. 7, 1971.

Kuczera, G. "On the Relationship Between the Reliability of Parameter Estimates and Hydrologic Time Series Data Used in Calibration", to appear in Water Resources Research, 1982.

Langseth, D. E. and Bras, R. L. "Empirical Temperature Forecasting: Extensions of the Model Output Statistics Method". MIT Ralph M. Parsons Lab. Tech. Rep. No. 257, June 1980.

Lin, A. C. and Yeh, W. W.-G., "Identification of Parameters in an Inhomogeneous Aquifer by use of the Maximum Principle of Optimal Control and Quasi-Linearization", Water Resources Research, Vol 10, No. 4, pp.829-838, Aug 1974.

Liou, E. Y., "OPSET: Program for Computerized Selections of Watershed Parameter Values for the Stanford Watershed Model", Water Resources Institute Research Report No. 34, 1970.

Luenberger, D. G., Introduction to Linear and Nonlinear Programming. Addison-Wesley Publishing Company, 1973.

McLaughlin, D., Investigation of Alternative Procedures for Estimating Groundwater Parameters", Water Resources Engineers, NTIS Pub. No. PB 239 952, January 1975.

Monro, J. C., "Direct Search Optimization in Mathematical Modeling and a Watershed Model Application", NOAA Technical Memorandum NWS HYDRO-12, 1971.

Moore, R. J. and Jones, D. A., "Coupled Bayesian-Kalman Filter Estimation of Parameters and States of a Dynamic Water Quality Model", presented at International Federation for Information Processing, Working Conference on Modelling and Simulation of Land, Air and Water Resources Systems, Ghent, Belgium, Aug.-Sept., 1977.

Moore, R. J. and Weiss, G., "Real Time Parameter Estimation of a Non-linear Catchment Model, Using Extended Kalman Filters", presented at IIASA/WMO Workshop on Recent Developments in in Real Time Forecasting/Control of Water Resource Systems", Laxenburg, Austria, Oct. 18-21, 1976.

Murray, D. L., "Boughton's Daily RainfallpRunoff Model Modified for the Brening Catchment", International Association of Scientific Hydrology, Pub. No. 96, pp. 144-161, 1970.

Nelder J. A. and Mead, R., "A Simplex Method for Function Minimization", Computer J., V 1, pp. 308-313, 1965.

Peck. E. L. "Catchment Modeling and Initial Parameter Estimation for the National Weather SErvice River Forecasting System", NOAA Technical Memorandum NWS Hydro-31, Washington, D. C., June 1976.

Peterson, D. W. "Hypothesis, Estimation and VAlidation of Dynamic Social Models...", Ph. D. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology, 1975.

Peterson, D. W. "GPSIE- General Purpose System Identifier and Evaluator", User's Manual, Draft, 1976.

Sagar, B., Yakowitz, S., Duckstein, L., "A Direct Method for the Identification of Parameters of Dynamic Non-homogeneous Aquifers", Water Resources Research, Vol. 11 No. 4, pp. 563-570, 1975.

Rosenbrock, H. H., "An Automatic Method for Finding the Greatest or the Least Value of a Function", Computer J. V 3, pp. 175-184, 1960.

Schweppe, F. C. Uncertain Dynamic Systems, Adison Wesley Publ., 1973.

Sooroshian, S. Unpublished Progress Report, 1981

Sooroshian, S. and Dracup, J. A., "Stochastic Parameter Estimation Procedures for Hydrologic Rainfall-Runoff Models: Correlated and Heteroscedastic Error Cases", Water Resources Research, Vol. 16 No. 7, PP. 430-442, April 1980.

Sugawara, M., "Automatic Calibration of the Tank Model," Hydrological Sciences Bulletin", Vol 24, No. 3, Sept. 1979.

Singh, V. P., "Estimation of Parameters of a Uniformly Non-linear Surface Runoff Model, "Nordic Hydrology, Vol 8 No. 1, pp. 33-46, 1977.

The Analytic Sciences Corporation, TASC, "Reduced-Order Unit Hydrograph Program Documentation", TR-1480-2, April 1980.

APPENDIX A

LISTING OF CARD DECKS FOR SIMULATION RUNS

```
VERIFICATION RUN. OFF LINE, ORIGINAL PARAMETERS
COHOCTON-RIVER AT CAMPBELL-NY
   10 1968    9 1970
    1    1    3    1    1    0    1    0    0    0    0     LAB FILE 1C
    1    1    0    1    1    0    1
LAB FILE 1C MAP    6  4319-CAMPBELL, NY- 4/28/77-1434.7016  COHOCTON R, CAMPBELL
LAB FILE 1C QME   24  3742-    01529500- 1/ 8/77-1730.3780  COHOCTON R. CAMPBELL
LAB FILE 1C QIN    6  6126-CAMPBELL, NY- 9/16/77-1547.7593  COHOCTON R, CAMPBELL
CAMPBELL, NY   CAMPBELL BASIN       1.00 1.00 100.  25.  .30 .004  .01  .01
CAMPBELL, NY         25.0  1.8 200. 100.  75.  .05 .002   02  0.3 0.03
CAMPBELL, NY                  36  51  89 160 270 335 375 340 250 150  94  53
CAMPBELL, NY          60.  10. 150.  10.  70. 210.
    0    0    1    0    1    0
LAB FILE 1C MAT    6  4801-         001- 4/28/77-1434.8748  CAMPBELL BASIN, N.Y.
CAMPBELL, N.Y. MAT       450. -0.4 -0.6 -0.7 -0.6
CAMPBELL, NY               001
                         450.   1. 1.20   1.  .32  .05  .06  40.  3. 42.3
                        0.   0.   0.   0.   0.   0.   0.
                        0.   1.  .05  .20
                       .03  .07  .13  .20  .27  .35  .45  .57  .73
CAMPBELL, N.Y.          01529500        1217.    0    0    6    0    0    0
    01529500            01529500        CAMPBELL, NY            0    1    7
         2         5        10        20        30        40
      3.01
                                  1    0    1    1       100
                                .04  .56  .28  .08  .04
                                  1    1    1    1    1
   0   0   0   0   0   3   0
0.279E00       0.468
0.293E00       0.862E00
0.660E00       1.445E00
0.842E00       3.675E00
0.776E00       3.803E00
1.020E00      12.986E00
0.625E00       4.176E00
0.571E00       6.305E00
0.660E00       4.686E00
0.591E00       4.391E00
0.689E00       3.666E00
0.418E00       2.041E00
     3.   4.
   9 2
    1.E-12    1.E-06   0.25E00 0.0351E00 0.0625E00 0.5625E00    5.E-01
    0.3E00    1.0E03
         0         0         0         0         0         0         0
7.4687359E-012.0566864E-01-1.288279E-02-3.922349E-03-3.017193E-03
-5.587757E-01-3.146569E-01-9.930304E-02-2.258615E-02-1.887379E-02
2.5913744E-01-7.352188E-01 5.244121E-01-2.732608E-01-1.450422E-01
-1.673535E-01 3.547002E-01 5.796233E-01 2.985887E-01-6.398414E-01
1.4403226E-01-3.316243E-01-3.442156E-01 7.158802E-01-2.363407E-01
            0         0         0         0         0
-6.322069E-01-7.601960E-01 5.909361E-02-9.304722E-02 6.733289E-02
-2.474841E+01 1.095328E+01-1.150028E-01-8.536949E-02-5.521520E-02
.SK
```

```
VERIFICATION RUN. OFF LINE, PARAMETERS FROM MITSCP, MAY-JUNE69
COHOCTON-RIVER AT CAMPBELL-NY
   10 1968    9 1970
    1    1    3    1    1    0    1    0    0    0    0     LAB FILE 1C
    1    1    0    1    1    0    1
LAB FILE 1C MAP    6  4319-CAMPBELL, NY- 4/28/77-1434.7016  COHOCTON R, CAMPBELL
LAB FILE 1C QME   24  3742-    01529500- 1/ 8/77-1730.3780  COHOCTON R. CAMPBELL
LAB FILE 1C QIN    6  6126-CAMPBELL, NY- 9/16/77-1547.7593  COHOCTON R, CAMPBELL
CAMPBELL, NY   CAMPBELL BASIN       1.00 1.00 39.4 16.5  .20 .000  .00  .01
CAMPBELL, NY         22.6  0.8 80.9  18.  50.  .05 .002  .02  0.3 0.00  .00
CAMPBELL, NY                 36  51  89 160 270 335 375 340 250 150  94  53
CAMPBELL, NY           60.  10. 150.  10.  70. 210.
    0    0    1    0    1    0
LAB FILE 1C MAT    6  4801-         001- 4/28/77-1434.8748  CAMPBELL BASIN, N.Y.
CAMPBELL, N.Y. MAT       450. -0.4 -0.6 -0.7 -0.6
CAMPBELL, NY                  001
                         450.   1. 1.20   1.  .32  .05  .06  40.  3. 42.3
                      0.   0.   0.   0.   0.   0.   0.   0.
                      0.   1.  .05  .20
                     .03  .07  .13  .20  .27  .35  .45  .57  .73
CAMPBELL, N.Y.         01529500        1217.    0    0    6    0    0    0
    01529500           01529500        CAMPBELL, NY            0    1    7
        2         5        10        20        30        40
     3.01
                                  1    0    1    1       100
                                .04  .56  .28  .08  .04
                                  1    1    1    1    1
   0   0   0   0   0   3   0
0.279E00        0.468
0.293E00        0.862E00
0.660E00        1.445E00
0.842E00        3.675E00
0.776E00        3.803E00
1.020E00       12.986E00
0.625E00        4.176E00
0.571E00        6.305E00
0.660E00        4.686E00
0.591E00        4.391E00
0.689E00        3.666E00
0.418E00        2.041E00
     3.   4.
   9 2
    1.E-12   5.1E-02   0.25E00 0.0351E00 0.0625E00 0.5625E00    1.E-07
    1.3E00   40.7E00
         0         0         0         0         0         0         0
7.4687359E-012.0566864E-01-1.288279E-02-3.922349E-03-3.017193E-03
-5.587757E-01-3.146569E-01-9.930304E-02-2.258615E-02-1.887379E-02
2.5913744E-01-7.352188E-01 5.244121E-01-2.732608E-01-1.450422E-01
-1.673535E-01 3.547002E-01 5.796233E-01 2.985887E-01-6.398414E-01
1.4403226E-01-3.316243E-01-3.442156E-01 7.158802E-01-2.363407E-01
          0         0         0         0         0
-6.322069E-01-7.601960E-01 5.909361E-02-9.304722E-02 6.733289E-02
-2.474841E+01 1.095328E+01-1.150028E-01-8.536949E-02-5.521520E-02
.SK
```

```
VERIFICATION RUN. ON LINE, ORIGINAL PARAMETERS
COHOCTON-RIVER AT CAMPBELL-NY
   10 1969    09 1970
    1    1    3    1    1    0    1    0    0    0    0     LAB FILE 1C
    1    1    0    1    1    0    1
LAB FILE 1C MAP    6  4319-CAMPBELL, NY- 4/28/77-1434.7016  COHOCTON R, CAMPBELL
LAB FILE 1C QME   24  3742-    01529500- 1/ 8/77-1730.3780  COHOCTON R. CAMPBELL
LAB FILE 1C QIN    6  6126-CAMPBELL, NY- 9/16/77-1547.7593  COHOCTON R, CAMPBELL
CAMPBELL, NY   CAMPBELL BASIN       1.00 1.00 100.  25.  .30 .004  .01  .01
CAMPBELL, NY          25.0  1.8 200. 100.  75.  .05 .002   02  0.3 0.03
CAMPBELL, NY                   36  51  89 160 270 335 375 340 250 150  94  53
CAMPBELL, NY           60.  10. 150.  10.  70. 210.
    0    0    1    0    1    0
LAB FILE 1C MAT    6  4801-         001- 4/28/77-1434.8748  CAMPBELL BASIN, N.Y.
CAMPBELL, N.Y. MAT        450. -0.4 -0.6 -0.7 -0.6
CAMPBELL, NY                 001
                          450.   1. 1.20   1.  .32  .05  .06  40.  3. 42.3
                        0.   0.   0.   0.   0.   0.   0.   0.
                        0.   1.  .05  .20
                       .03  .07  .13  .20  .27  .35  .45  .57  .73
CAMPBELL, N.Y.           01529500       1217.    0    0    6    0    0    0
    01529500             01529500        CAMPBELL, NY            0    1    7
         2         5         10        20        30        40
      3.01
                                   1    0    1    1       100
                                 .04  .56  .28  .08  .04
                                   1    1    1    1    1
   0   0   0   0   0   0   0
0.279E00        0.468
0.293E00        0.862E00
0.660E00        1.445E00
0.842E00        3.675E00
0.776E00        3.803E00
1.020E00       12.986E00
0.625E00        4.176E00
0.571E00        6.305E00
0.660E00        4.686E00
0.591E00        4.391E00
0.689E00        3.666E00
0.418E00        2.041E00
     3.   4.
   9 2
    1.E-12    1.E-06   0.25E00 0.0351E00 0.0625E00 0.5625E00    5.E-01
    0.3E00    1.0E03
         0         0         0         0         0         0         0
7.4687359E-012.0566864E-01-1.288279E-02-3.922349E-03-3.017193E-03
-5.587757E-01-3.146569E-01-9.930304E-02-2.258615E-02-1.887379E-02
2.5913744E-01-7.352188E-01 5.244121E-01-2.732608E-01-1.450422E-01
-1.673535E-01 3.547002E-01 5.796233E-01 2.985887E-01-6.398414E-01
1.4403226E-01-3.316243E-01-3.442156E-01 7.158802E-01-2.363407E-01
         0         0         0         0         0
-6.322069E-01-7.601960E-01 5.909361E-02-9.304722E-02 6.733289E-02
-2.474841E+01 1.095328E+01-1.150028E-01-8.536949E-02-5.521520E-02
```

```
VERIFICATION RUN. ON LINE, PARAMETERS FROM MITSCP, MAY-JUNE69
COHOCTON-RIVER AT CAMPBELL-NY
   10 1969    9 1970
    1    1    3    1    1    0    1    0    0    0    0     LAB FILE 1C
    1    1    0    1    1    0    1
LAB FILE 1C MAP    6  4319-CAMPBELL, NY- 4/28/77-1434.7016  COHOCTON R, CAMPBELL
LAB FILE 1C QME   24  3742-    01529500- 1/ 8/77-1730.3780  COHOCTON R. CAMPBELL
LAB FILE 1C QIN    6  6126-CAMPBELL, NY- 9/16/77-1547.7593  COHOCTON R, CAMPBELL
CAMPBELL, NY   CAMPBELL BASIN       1.00 1.00 39.4 16.5  .20 .000  .00  .01
CAMPBELL, NY         22.6  0.8 80.9  18.  50.  .05 .002  .02  0.3 0.00  .00
CAMPBELL, NY                   36  51  89 160 270 335 375 340 250 150  94  53
CAMPBELL, NY          00.  10. 150.  10.  70. 210.
    0    0    1    0    1    0
LAB FILE 1C MAT    6  4801-         001- 4/28/77-1434.8748  CAMPBELL BASIN, N.Y.
CAMPBELL, N.Y. MAT        450. -0.4 -0.6 -0.7 -0.6
CAMPBELL, NY                 001
                          450.   1. 1.20   1.  .32  .05  .06  40.  3. 42.3
                        0.   0.   0.   0.   0.   0.   0.   0.
                        0.   1.  .05  .20
                       .03  .07  .13  .20  .27  .35  .45  .57  .73
CAMPBELL, N.Y.           01529500       1217.    0    0    6    0    0    0
    01529500             01529500        CAMPBELL, NY            0    1    7
        2         5         10        20        30        40
     3.01
                                  1    0    1    1       100
                                .04  .56  .28  .08  .04
                                  1    1    1    1    1
   0   0   0   0   0   0   0
0.279E00        0.468
0.293E00        0.862E00
0.660E00        1.445E00
0.842E00        3.675E00
0.776E00        3.803E00
1.020E00       12.986E00
0.625E00        4.176E00
0.571E00        6.305E00
0.660E00        4.686E00
0.591E00        4.391E00
0.689E00        3.666E00
0.418E00        2.041E00
     3.   4.
   9 2
    1.E-12   5.1E-02   0.25E00 0.0351E00 0.0625E00 0.5625E00    1.E-07
    1.3E00   40.7E00
         0         0         0         0         0         0         0
7.4687359E-012.0566864E-01-1.288279E-02-3.922349E-03-3.017193E-03
-5.587757E-01-3.146569E-01-9.930304E-02-2.258615E-02-1.887379E-02
2.5913744E-01-7.352188E-01 5.244121E-01-2.732608E-01-1.450422E-01
-1.673535E-01 3.547002E-01 5.796233E-01 2.985887E-01-6.398414E-01
1.4403226E-01-3.316243E-01-3.442156E-01 7.158802E-01-2.363407E-01
         0         0         0         0         0         0
-6.322069E-01-7.601960E-01 5.909361E-02-9.304722E-02 6.733289E-02
-2.474841E+01 1.095328E+01-1.150028E-01-8.536949E-02-5.521520E-02
.SK
```

```
ON LINE RUN, UNITGRAPH MODEL, PARAMETERS FROM MCP.
BIRD CREEK NEAR SPERRY,OKLAHOMA
   10 1958   09 1959
    0    1    3    1    1    0    1    0    0    0    0     LAB FILE 3C
    1    1    1    1    1    0    1
LAB FILE 3C MAP    6  5766-BIRD CREEK  - 7/ 7/78-1854.9825  WMO MAP(ADJUSTED)
LAB FILE 3C PTPE  24  6104-BIRD CREEK  - 7/ 7/78-1856.6403  POINT POTENTIAL ET
LAB FILE 3C QME   24  6190-    07177500- 7/ 7/78-1856.7513  BIRD CR,NEAR SPERRY
LAB FILE 3C QINE   6  6288-USGS07177500- 7/ 7/78-1856.7847  BIRD CREEK--SPERRY
BIRD CREEK     BIRD CREEK MAP          1.0  1.0 120.  15.  0.3 .001  .17 .001
BIRD CREEK             48.  2.1 160.  14. 140. .126 .013  .02  .30 3.55
BIRD CREEK     BIRD CREEK      .33 .33 .31 .32 .52 .67 .69 .69 .66 .62 .42 .33
BIRD CREEK            102.   0. 140.   0.  10. 221.
BIRD CREEK--SPERRY  USGS07177500        2344.    0    0    6    0    0    0
USGS07177500            07177500        USGS07177500                0    1   14
       .15       0.7       7.0       14.       42.       140.
       3.0
BIRD CREEK--SPERRY                0  0.0    1    1      750.
BIRD CREEK--SPERRY             .023  .04 .072  .11  .13 .128 .125 .114 .095  .07
BIRD CREEK--SPERRY              .05  .03  .01 .003
BIRD CREEK--SPERRY                1    1    1    1    1    1    1    1    1    1
BIRD CREEK--SPERRY                1    1    1    1
   0   0   0   0   0   0   0
    0.358984E 00    0.257818E 01
    0.150761E 01    0.192790E 02
    0.690823E 00    0.109351E 02
    0.187018E 01    0.248343E 02
    0.272435E 00    0.123490E 01
    0.127052E 01    0.235601E 02
    0.573548E-01    0.764108E-01
    0.207403E 00    0.174053E 01
    0.159774E 00    0.314154E 00
    0.839839E-01    0.169653E 00
    0.255548E 00    0.245921E 01
    0.573548E 00    0.470271E 01
     3.   4.
   9 2
    1.E-12     1.E-2    2.5E-1    4.6E-2   1.95E-1    4.9E-1     5.E-1
   0.300E0     10.E0
   1.96E-2    1.1E-3     1.E-3   8.79E-2   1.56E-2    1.4E-1     1.E-2
 0.924498E 00 0.127624E 00-0.120877E-01 0.147609E-02-0.221878E-02 0.140791E-02
-0.320435E 00 0.749457E 00 0.133947E 00-0.723106E-02 0.118776E-01-0.550329E-02
-0.129278E 00-0.570561E 00 0.571302E 00 0.777237E-01-0.106783E 00 0.831124E-01
-0.595580E-01-0.116204E 00-0.293226E 00 0.638737E 00 0.656881E 00 0.128785E 00
-0.937171E-01-0.199812E 00-0.421724E 00-0.687642E 00 0.267024E 00 0.346402E 00
-0.909697E-01-0.141622E 00-0.502120E 00 0.206233E 00-0.529903E 00-0.305710E 00
 0.358971E 00 0.563439E 00 0.553704E 00 0.185776E 00 0.301332E 00 0.265058E 00
 0.286981E 02-0.179404E 02 0.413898E 01-0.368092E 00 0.570344E 00-0.327956E 00
-0.327000E-03-0.100000E-03-0.100000E-03-0.100000E-03-0.100000E-03-0.100000E-03
```

```
ON LINE RUN, UNITGRAPH MODEL, PARAMETERS FROM MITSCP, APRIL-JULY57 (PRIOR
BIRD CREEK NEAR SPERRY,OKLAHOMA
   10 1958   09 1960
    0    1    3    1    1    0    1    0    0    0    0     LAB FILE 3C
    1    1    1    1    1    0    1
LAB FILE 3C MAP    6  5766-BIRD CREEK  - 7/ 7/78-1854.9825  WMO MAP(ADJUSTED)
LAB FILE 3C PTPE  24  6104-BIRD CREEK  - 7/ 7/78-1856.6403  POINT POTENTIAL ET
LAB FILE 3C QME   24  6190-    07177500- 7/ 7/78-1856.7513  BIRD CR,NEAR SPERRY
LAB FILE 3C QINE   6  6288-USGS07177500- 7/ 7/78-1856.7847  BIRD CREEK--SPERRY
BIRD CREEK     BIRD CREEK MAP         1.0  1.0122.5 14.7 0.86 .019  .13 .001
BIRD CREEK             38.  1.1 160. 12.5 140. .549.0155.0000  .30 5.18
BIRD CREEK     BIRD CREEK      .33 .33 .31 .32 .52 .67 .69 .69 .66 .62 .42 .33
BIRD CREEK           102.   0. 140.   0.  10. 221.
BIRD CREEK--SPERRY  USGS07177500        2344.    0    0    6    0    0    0
USGS07177500            07177500        USGS07177500             0    1   14
       .15       0.7       7.0       14.       42.       140.
       3.0
BIRD CREEK--SPERRY                0  0.0    1    1       750.
BIRD CREEK--SPERRY             .023  .04 .072  .11  .13 .128 .125 .114 .095  .07
BIRD CREEK--SPERRY              .05  .03  .01 .003
BIRD CREEK--SPERRY                1    1    1    1    1    1    1    1    1    1
BIRD CREEK--SPERRY                1    1    1    1
   0   0   0   0   0   0   0
    0.358984E 00    0.257818E 01
    0.150761E 01    0.192790E 02
    0.690823E 00    0.109351E 02
    0.187018E 01    0.248343E 02
    0.272435E 00    0.123490E 01
    0.127052E 01    0.235601E 02
    0.573548E-01    0.764108E-01
    0.207403E 00    0.174053E 01
    0.159774E 00    0.314154E 00
    0.839839E-01    0.169653E 00
    0.255548E 00    0.245921E 01
    0.573548E 00    0.470271E 01
     3.   4.
   9 2
    1.E-12     1E-14    2.5E-1    4.6E-2   1.95E-1    4.9E-1     3.E-4
   1.3523       4.12
   1.96E-2    1.1E-3     1.E-3   8.79E-2   1.56E-2    1.4E-1     1.E-2
 0.924498E 00 0.127624E 00-0.120877E-01 0.147609E-02-0.221878E-02 0.140791E-02
-0.320435E 00 0.749457E 00 0.133947E 00-0.723106E-02 0.118776E-01-0.550329E-02
-0.129278E 00-0.570561E 00 0.571302E 00 0.777237E-01-0.106783E 00 0.831124E-01
-0.595580E-01-0.116204E 00-0.293226E 00 0.638737E 00 0.656881E 00 0.128785E 00
-0.937171E-01-0.199812E 00-0.421724E 00-0.687642E 00 0.267024E 00 0.346402E 00
-0.909697E-01-0.141622E 00-0.502120E 00 0.206233E 00-0.529903E 00-0.305710E 00
 0.358971E 00 0.563439E 00 0.553704E 00 0.185776E 00 0.301332E 00 0.265058E 00
 0.286981E 02-0.179404E 02 0.413898E 01-0.368092E 00 0.570344E 00-0.327956E 00
-0.327000E-03-0.100000E-03-0.100000E-03-0.100000E-03-0.100000E-03-0.100000E-03
```

```
OFF LINE RUN, UNITGRAPH MODEL, PARAMETERS FROM MCP.
BIRD CREEK NEAR SPERRY,OKLAHOMA
   10 1958   09 1959
    0    1    3    1    1    0    1    0    0    0    0     LAB FILE 3C
    1    1    1    1    1    0    1
LAB FILE 3C MAP    6  5766-BIRD CREEK  - 7/ 7/78-1854.9825  WMO MAP(ADJUSTED)
LAB FILE 3C PTPE  24  6104-BIRD CREEK  - 7/ 7/78-1856.6403  POINT POTENTIAL ET
LAB FILE 3C QME   24  6190-    07177500- 7/ 7/78-1856.7513  BIRD CR,NEAR SPERRY
LAB FILE 3C QINE   6  6288-USGS07177500- 7/ 7/78-1856.7847  BIRD CREEK--SPERRY
BIRD CREEK     BIRD CREEK MAP         1.0  1.0 120.  15.  0.3 .001  .17 .001
BIRD CREEK            48.  2.1 160.  14. 140. .126 .013  .02  .30 3.55
BIRD CREEK     BIRD CREEK      .33 .33 .31 .32 .52 .67 .69 .69 .66 .62 .42 .33
BIRD CREEK           102.   0. 140.   0.  10. 221.
BIRD CREEK--SPERRY  USGS07177500        2344.    0    0    6    0    0    0
USGS07177500            07177500        USGS07177500              0    1   14
       .15       0.7       7.0       14.       42.       140.
       3.0
BIRD CREEK--SPERRY                  0  0.0    1    1      750.
BIRD CREEK--SPERRY                .023  .04 .072  .11  .13 .128 .125 .114 .095  .07
BIRD CREEK--SPERRY                 .05  .03  .01 .003
BIRD CREEK--SPERRY                   1    1    1    1    1    1    1    1    1    1
BIRD CREEK--SPERRY                   1    1    1    1
   0   0   0   0   0   3   0
    0.358984E 00    0.257818E 01
    0.150761E 01    0.192790E 02
    0.690823E 00    0.109351E 02
    0.187018E 01    0.248343E 02
    0.272435E 00    0.123490E 01
    0.127052E 01    0.235601E 02
    0.573548E-01    0.764108E-01
    0.207403E 00    0.174053E 01
    0.159774E 00    0.314154E 00
    0.839839E-01    0.169653E 00
    0.255548E 00    0.245921E 01
    0.573548E 00    0.470271E 01
     3.   4.
   9 2
    1.E-12     1.E-2    2.5E-1    4.6E-2   1.95E-1    4.9E-1     5.E-1
   0.300E0     10.E0
   1.96E-2    1.1E-3     1.E-3   8.79E-2   1.56E-2    1.4E-1     1.E-2
 0.924498E 00 0.127624E 00-0.120877E-01 0.147609E-02-0.221878E-02 0.140791E-02
-0.320435E 00 0.749457E 00 0.133947E 00-0.723106E-02 0.118776E-01-0.550329E-02
-0.129278E 00-0.570561E 00 0.571302E 00 0.777237E-01-0.106783E 00 0.831124E-01
-0.595580E-01-0.116204E 00-0.293226E 00 0.638737E 00 0.656881E 00 0.128785E 00
-0.937171E-01-0.199812E 00-0.421724E 00-0.687642E 00 0.267024E 00 0.346402E 00
-0.909697E-01-0.141622E 00-0.502120E 00 0.206233E 00-0.529903E 00-0.305710E 00
 0.358971E 00 0.563439E 00 0.553704E 00 0.185776E 00 0.301332E 00 0.265058E 00
 0.286981E 02-0.179404E 02 0.413898E 01-0.368092E 00 0.570344E 00-0.327956E 00
-0.327000E-03-0.100000E-03-0.100000E-03-0.100000E-03-0.100000E-03-0.100000E-03
```

```
OFF LINE RUN, UNITGRAPH MODEL, PARAMETERS FROM MITSCP, APRIL-JULY57(PRIOR
BIRD CREEK NEAR SPERRY,OKLAHOMA
   10 1958   09 1960
    0    1    3    1    1    0    1    0    0    0    0     LAB FILE 3C
    1    1    1    1    1    0    1
LAB FILE 3C MAP    6  5766-BIRD CREEK  - 7/ 7/78-1854.9825  WMO MAP(ADJUSTED)
LAB FILE 3C PTPE  24  6104-BIRD CREEK  - 7/ 7/78-1856.6403  POINT POTENTIAL ET
LAB FILE 3C QME   24  6190-    07177500- 7/ 7/78-1856.7513  BIRD CR,NEAR SPERRY
LAB FILE 3C QINE   6  6288-USGS07177500- 7/ 7/78-1856.7847  BIRD CREEK--SPERRY
BIRD CREEK     BIRD CREEK MAP        1.0  1.0122.5 14.7 0.86 .019  .13 .001
BIRD CREEK            38.  1.1 160. 12.5 140. .549.0155.0000  .30 5.18
BIRD CREEK     BIRD CREEK      .33 .33 .31 .32 .52 .67 .69 .69 .66 .62 .42 .33
BIRD CREEK           102.   0. 140.   0.  10. 221.
BIRD CREEK--SPERRY  USGS07177500        2344.    0    0    6    0    0    0
USGS07177500            07177500        USGS07177500             0    1   14
       .15       0.7       7.0       14.       42.      140.
       3.0
BIRD CREEK--SPERRY                0  0.0    1    1      750.
BIRD CREEK--SPERRY             .023  .04 .072  .11  .13 .128 .125 .114 .095  .07
BIRD CREEK--SPERRY              .05  .03  .01 .003
BIRD CREEK--SPERRY                1    1    1    1    1    1    1    1    1    1
BIRD CREEK--SPERRY                1    1    1    1
   0   0   0   0   0   3   0
    0.358984E 00    0.257818E 01
    0.150761E 01    0.192790E 02
    0.690823E 00    0.109351E 02
    0.187018E 01    0.248343E 02
    0.272435E 00    0.123490E 01
    0.127052E 01    0.235601E 02
    0.573548E-01    0.764108E-01
    0.207403E 00    0.174053E 01
    0.159774E 00    0.314154E 00
    0.839839E-01    0.169653E 00
    0.255548E 00    0.245921E 01
    0.573548E 00    0.470271E 01
     3.   4.
   9 2
    1.E-12     1E-14    2.5E-1    4.6E-2   1.95E-1    4.9E-1     3.E-4
   1.3523       4.12
   1.96E-2    1.1E-3     1.E-3   8.79E-2   1.56E-2    1.4E-1     1.E-2
 0.924498E 00 0.127624E 00-0.120877E-01 0.147609E-02-0.221878E-02 0.140791E-02
-0.320435E 00 0.749457E 00 0.133947E 00-0.723106E-02 0.118776E-01-0.550329E-02
-0.129278E 00-0.570561E 00 0.571302E 00 0.777237E-01-0.106783E 00 0.831124E-01
-0.595580E-01-0.116204E 00-0.293226E 00 0.638737E 00 0.656881E 00 0.128785E 00
-0.937171E-01-0.199812E 00-0.421724E 00-0.687642E 00 0.267024E 00 0.346402E 00
-0.909697E-01-0.141622E 00-0.502120E 00 0.206233E 00-0.529903E 00-0.305710E 00
 0.358971E 00 0.563439E 00 0.553704E 00 0.185776E 00 0.301332E 00 0.265058E 00
 0.286981E 02-0.179404E 02 0.413898E 01-0.368092E 00 0.570344E 00-0.327956E 00
-0.327000E-03-0.100000E-03-0.100000E-03-0.100000E-03-0.100000E-03-0.100000E-03
```

APPENDIX B

PLOTS OF SIMULATED DISCHARGES

DURING THE VERIFICATION RUNS

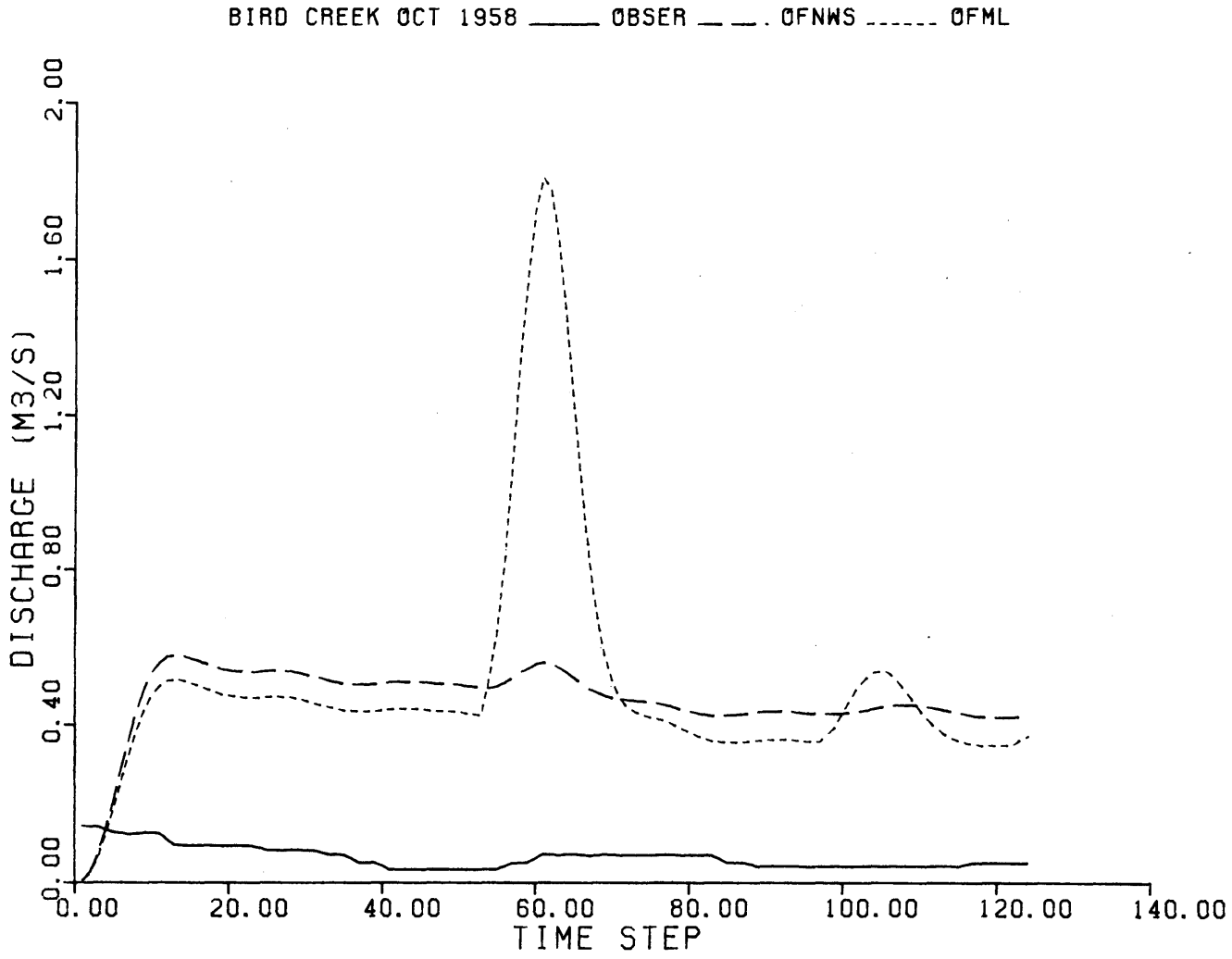


Figure B.1 Six-Hour Lead Off-Line Forecasts for Bird Creek, October 1958.

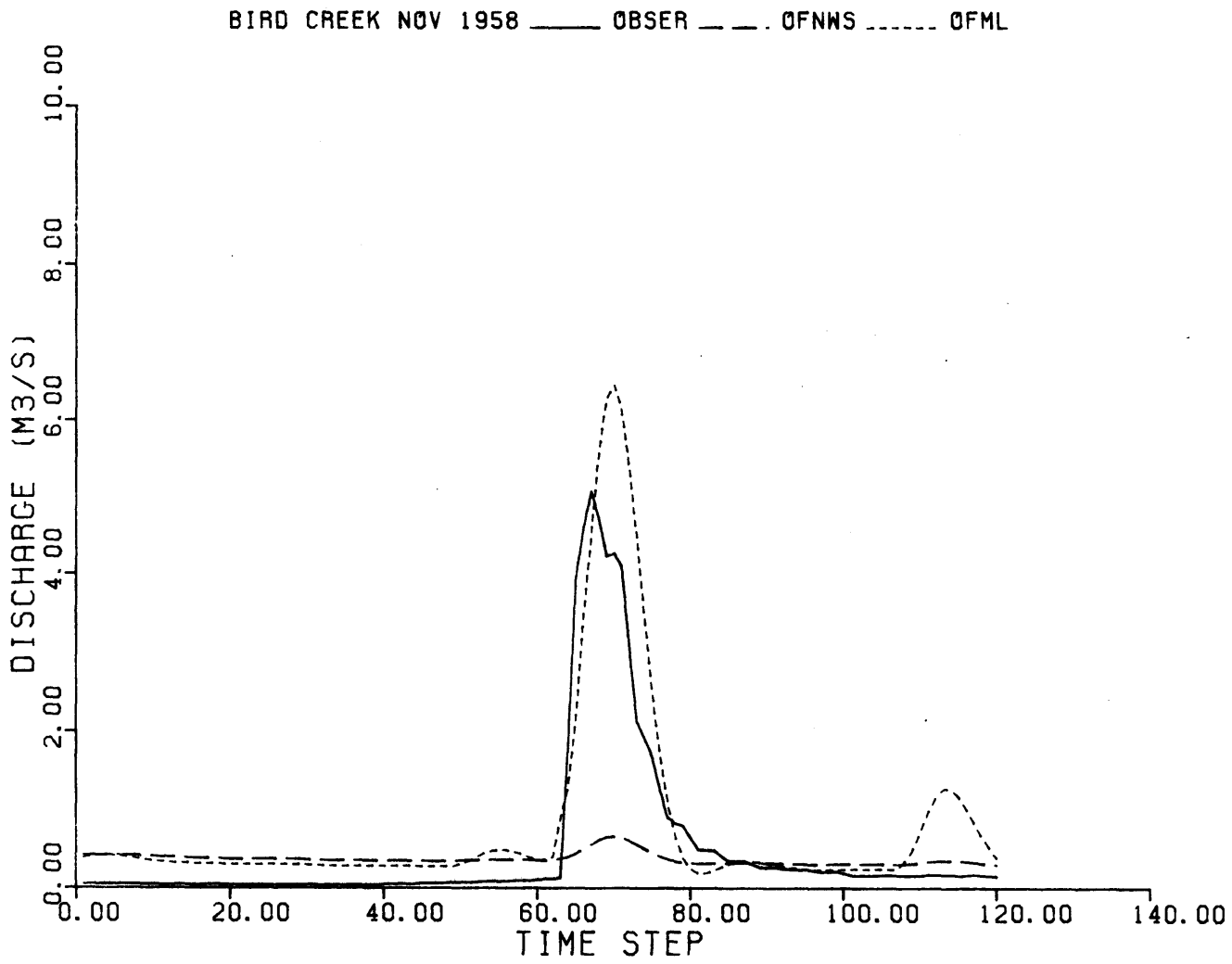


Figure B.2 Six-Hour Lead Off-Line Forecasts for Bird Creek, November 1958.

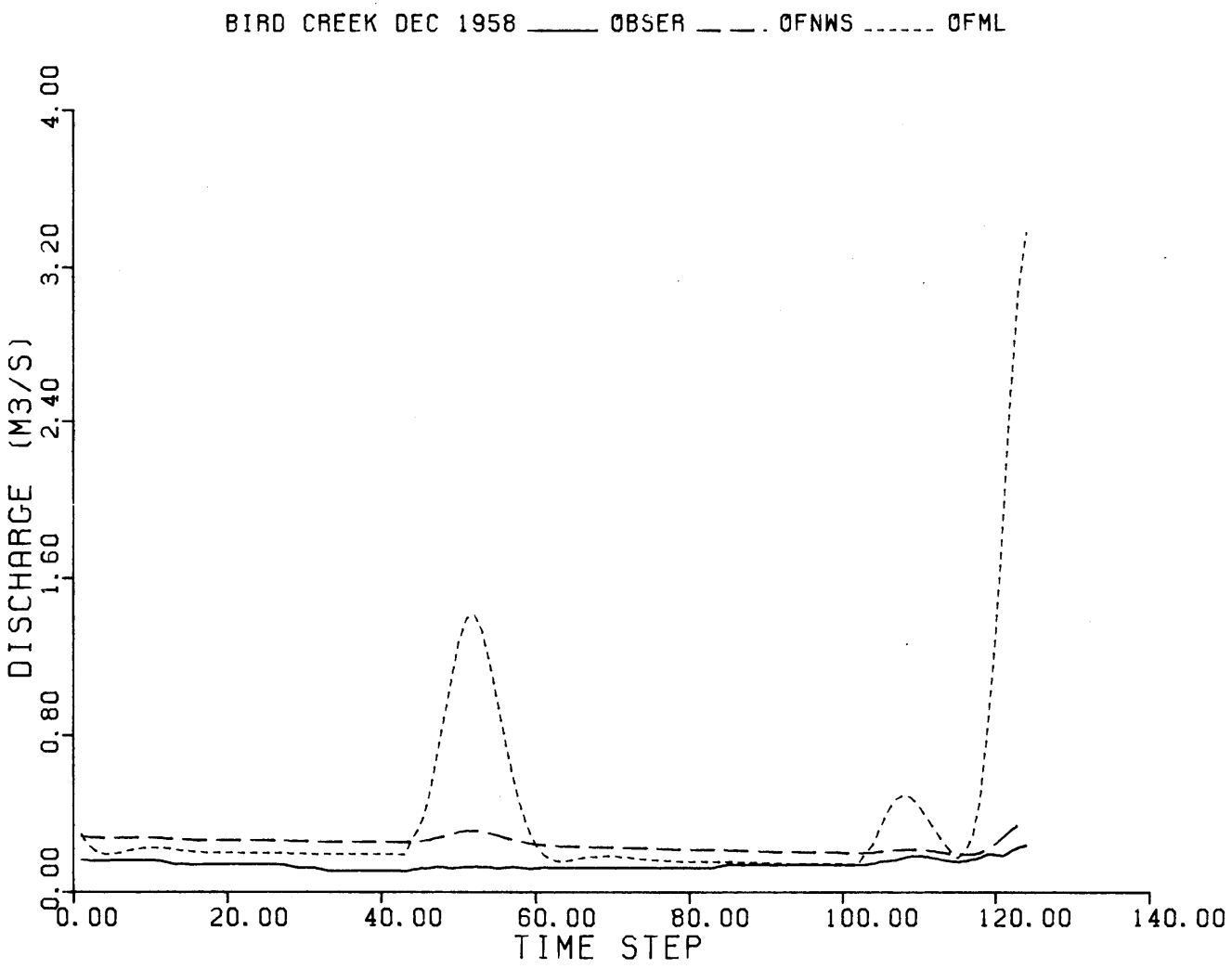


Figure B.3 Six-Hour Lead Off-Line Forecasts for Bird Creek, December 1958.

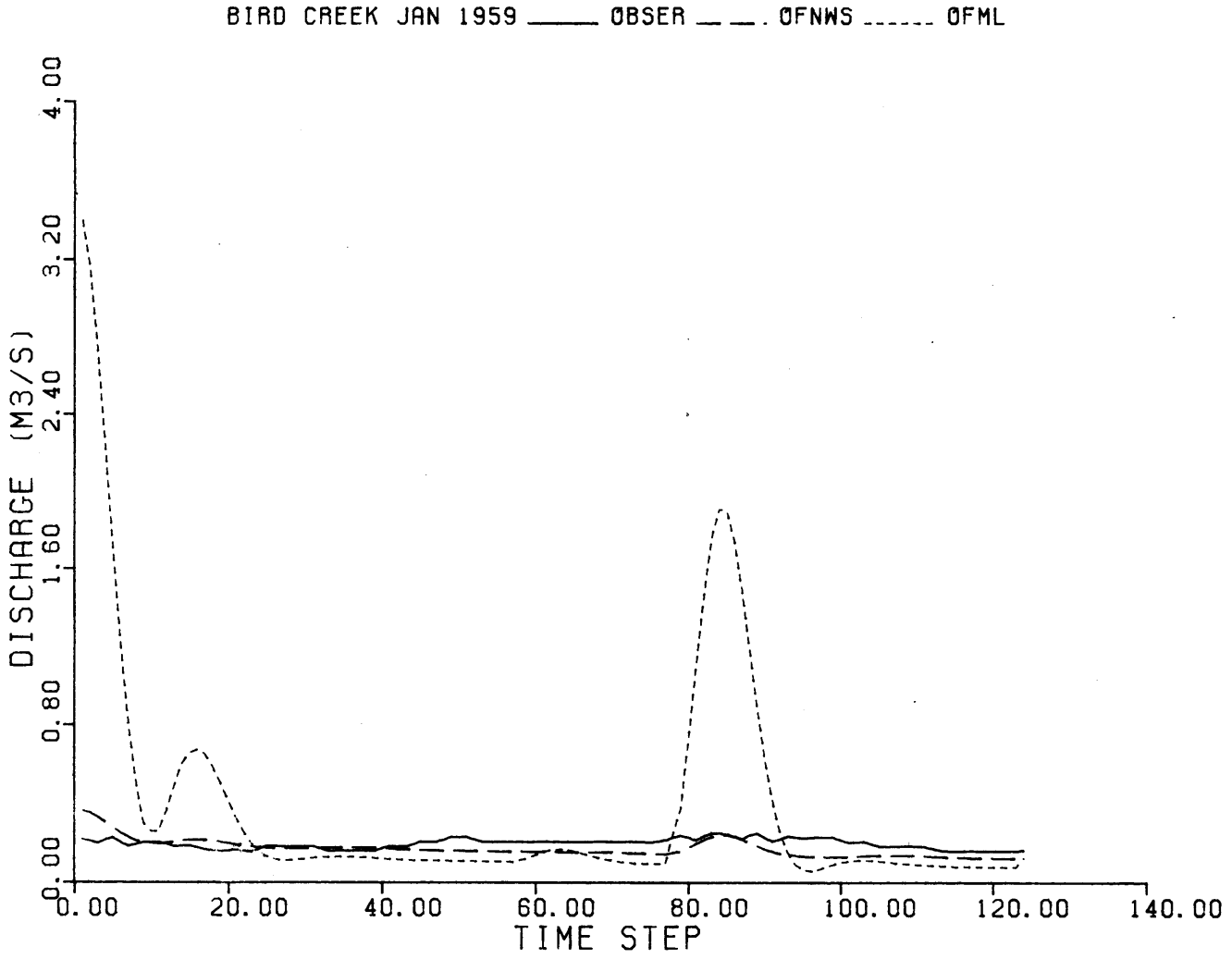


Figure B.4 Six-Hour Lead Off-Line Forecasts for Bird Creek, January 1959.

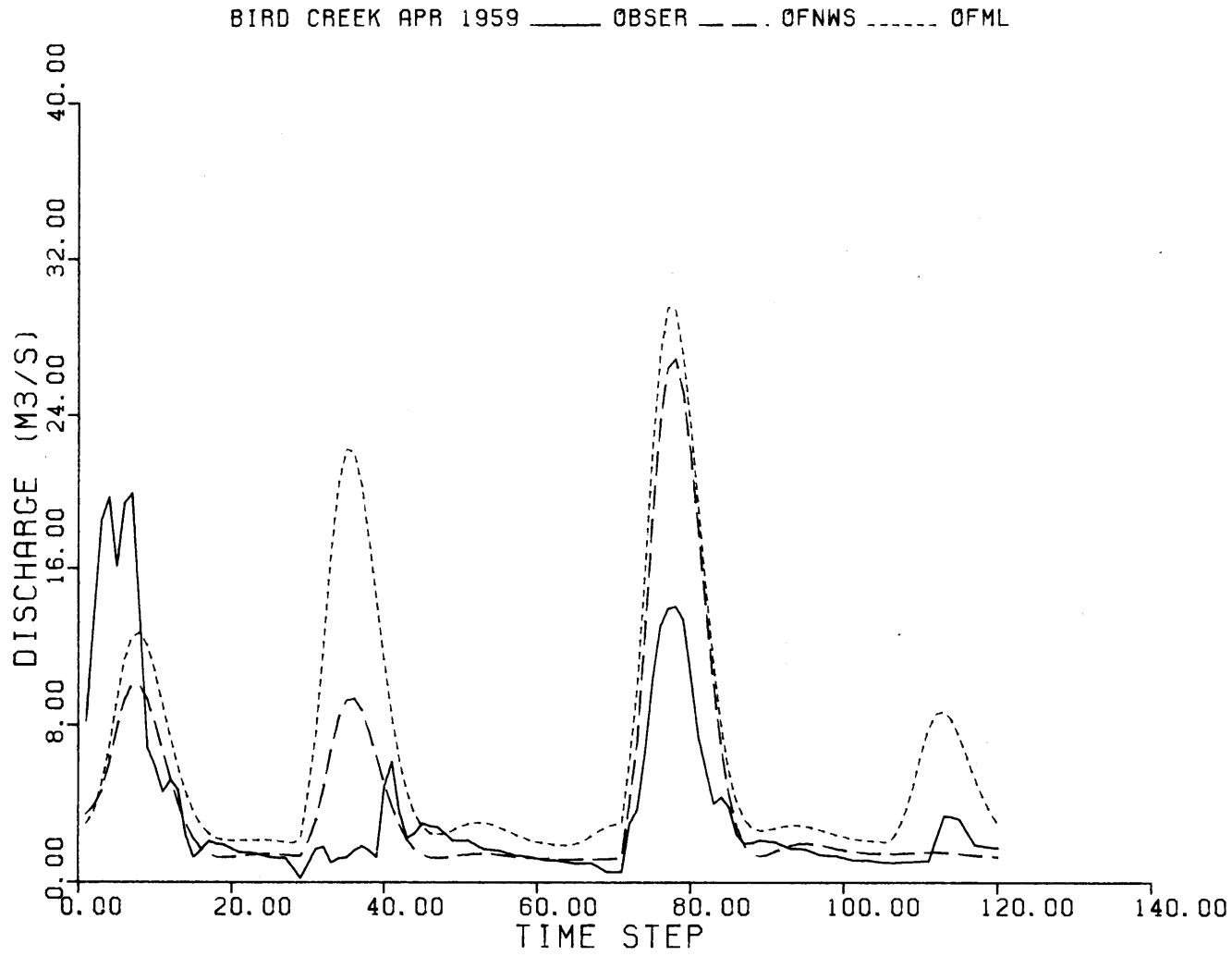


Figure B.5 Six-Hour Lead Off-Line Forecasts for Bird Creek, April 1959.

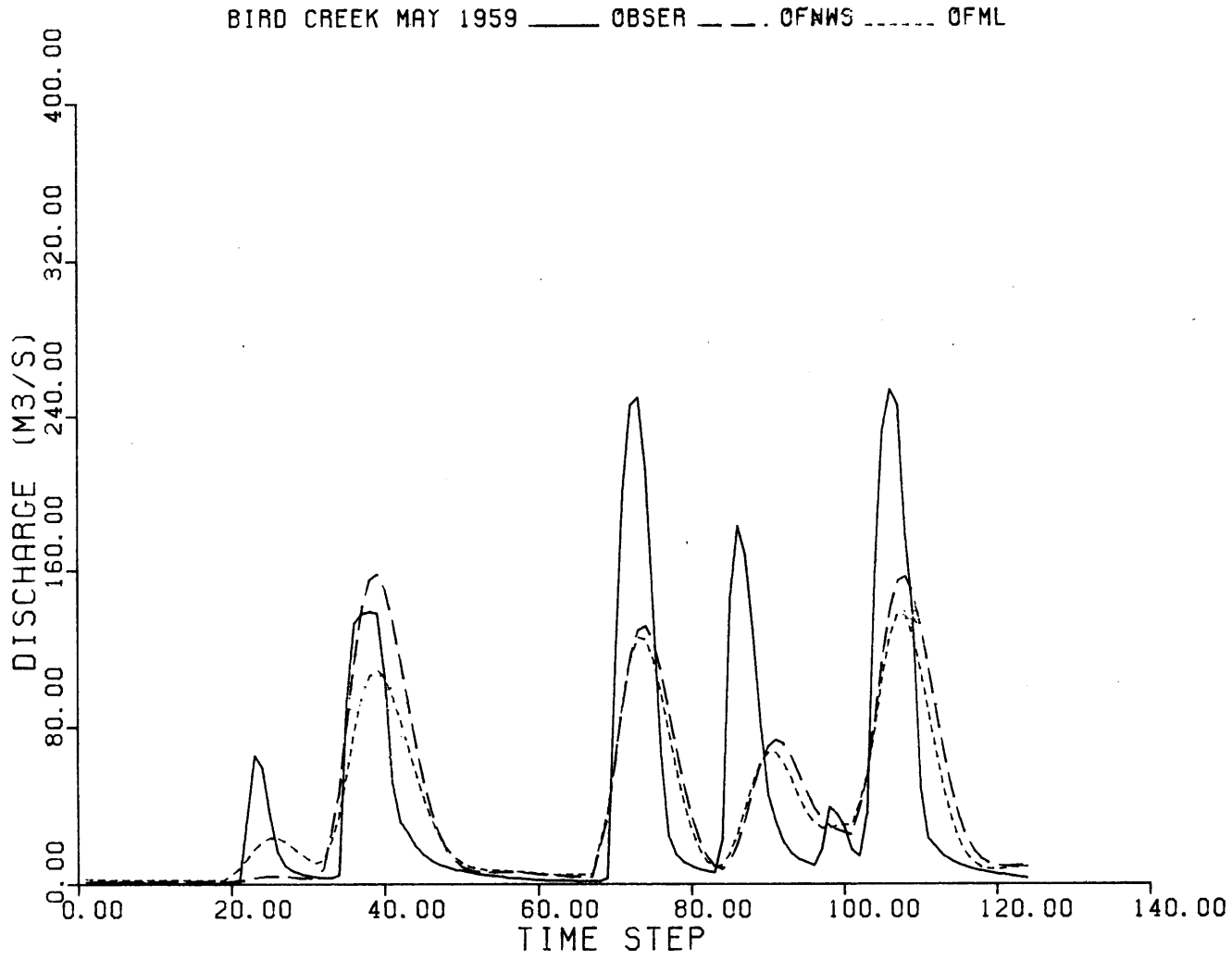


Figure B.6 Six-Hour Lead Off-Line Forecasts for Bird Creek, May 1959.

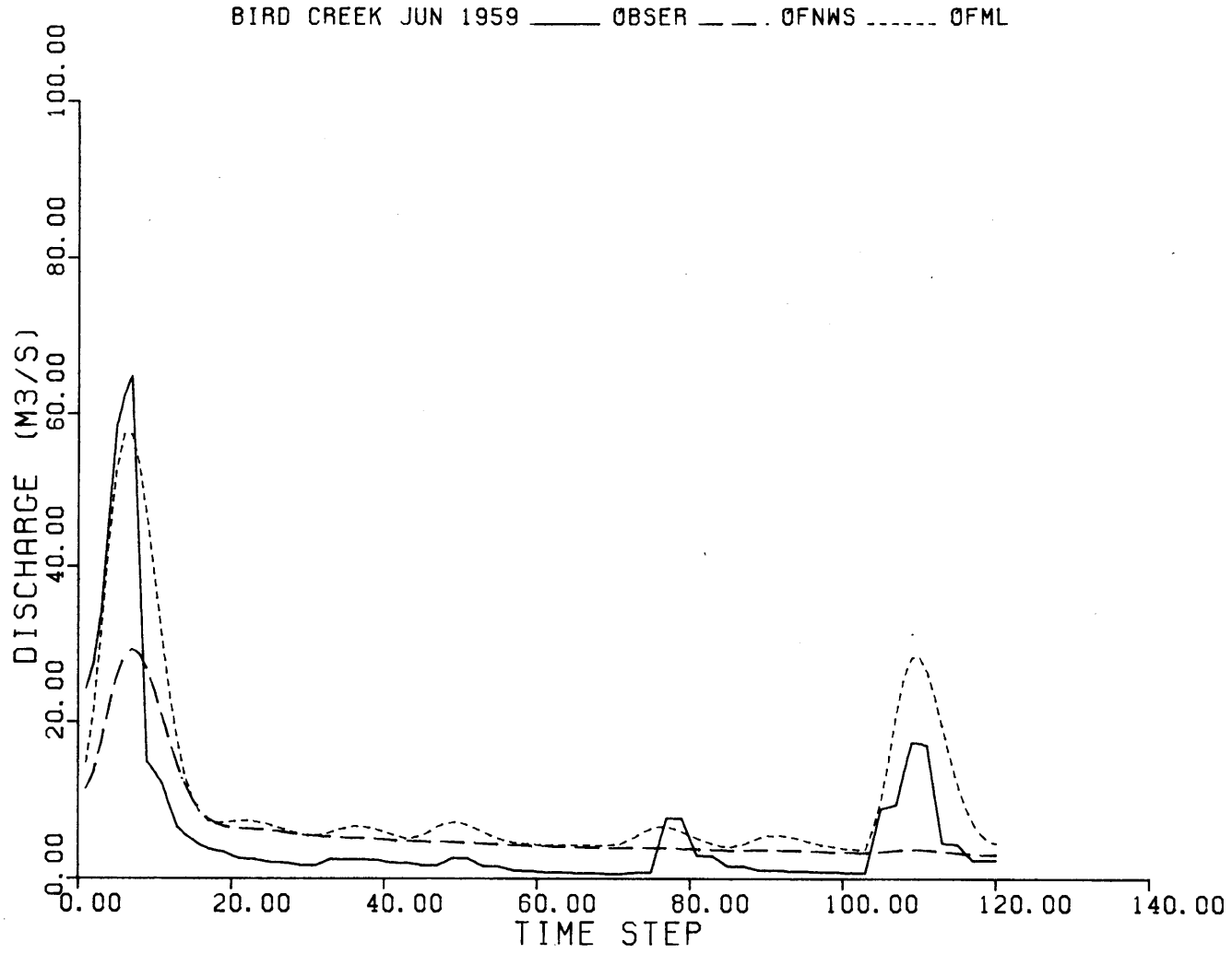


Figure B.7 Six-Hour Lead Off-Line Forecasts for Bird Creek, June 1959.

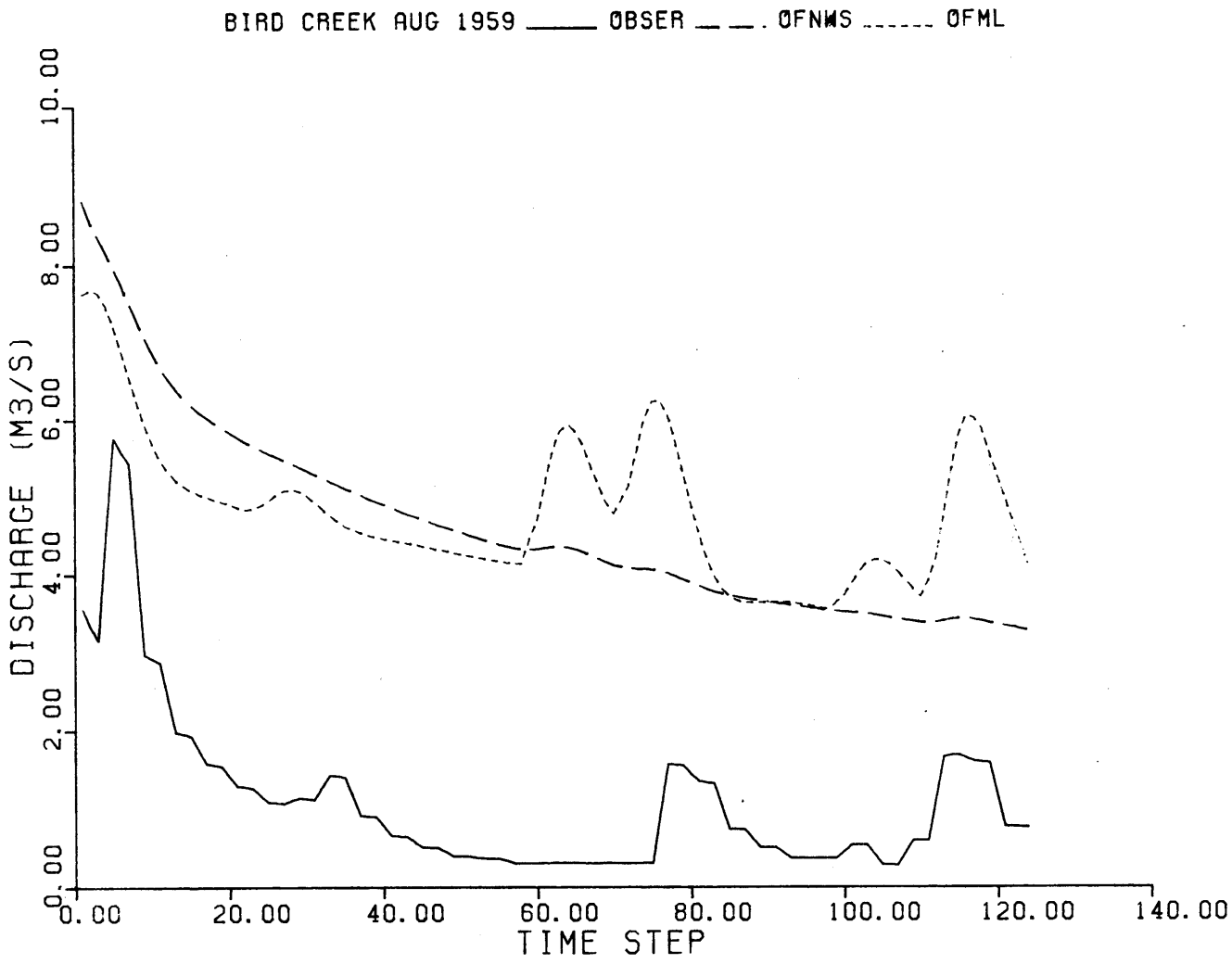


Figure B.8 Six-Hour Lead Off-Line Forecasts for Bird Creek, August 1959.

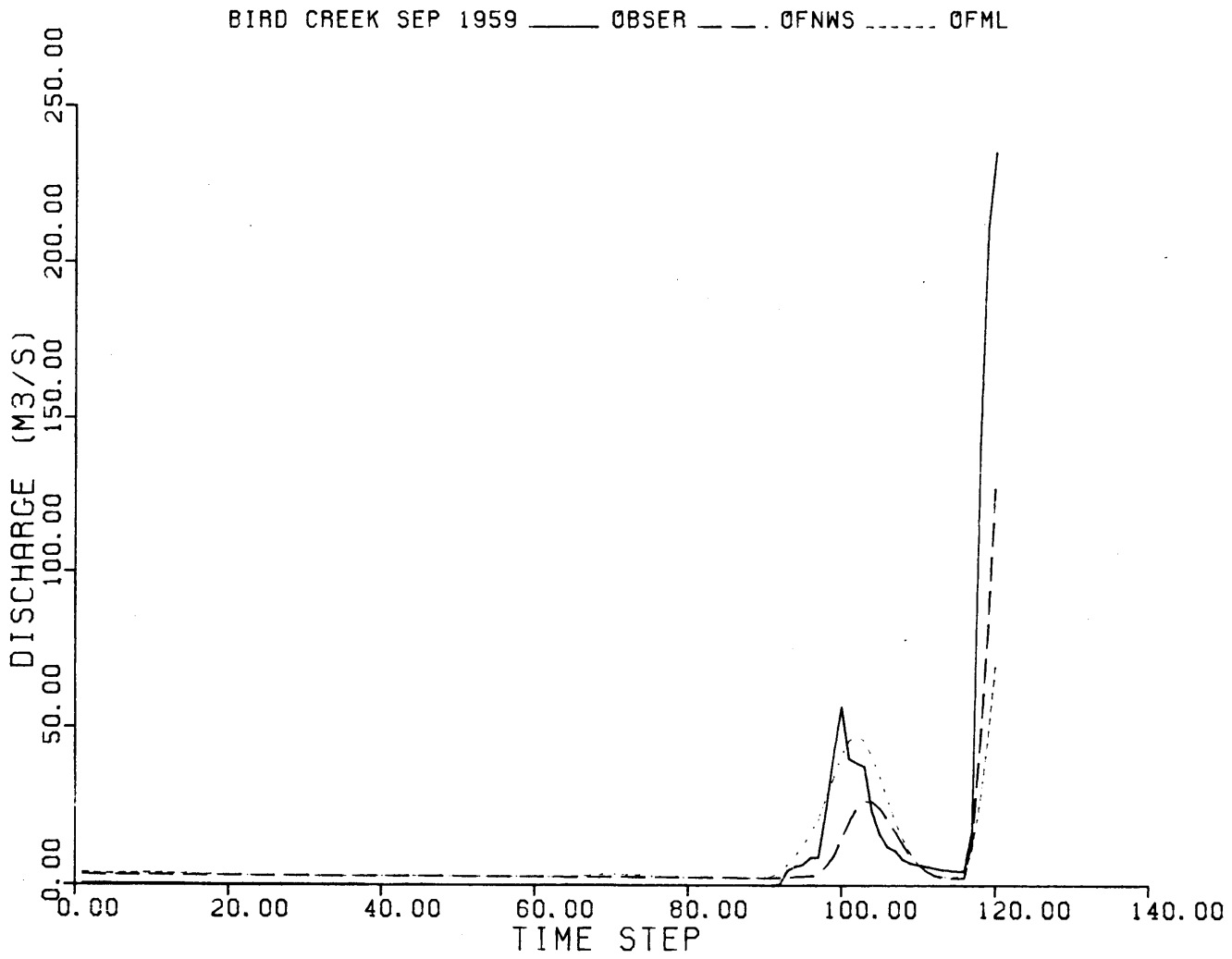


Figure B.9 Six-Hour Lead Off-Line Forecasts for Bird Creek, September 1959.

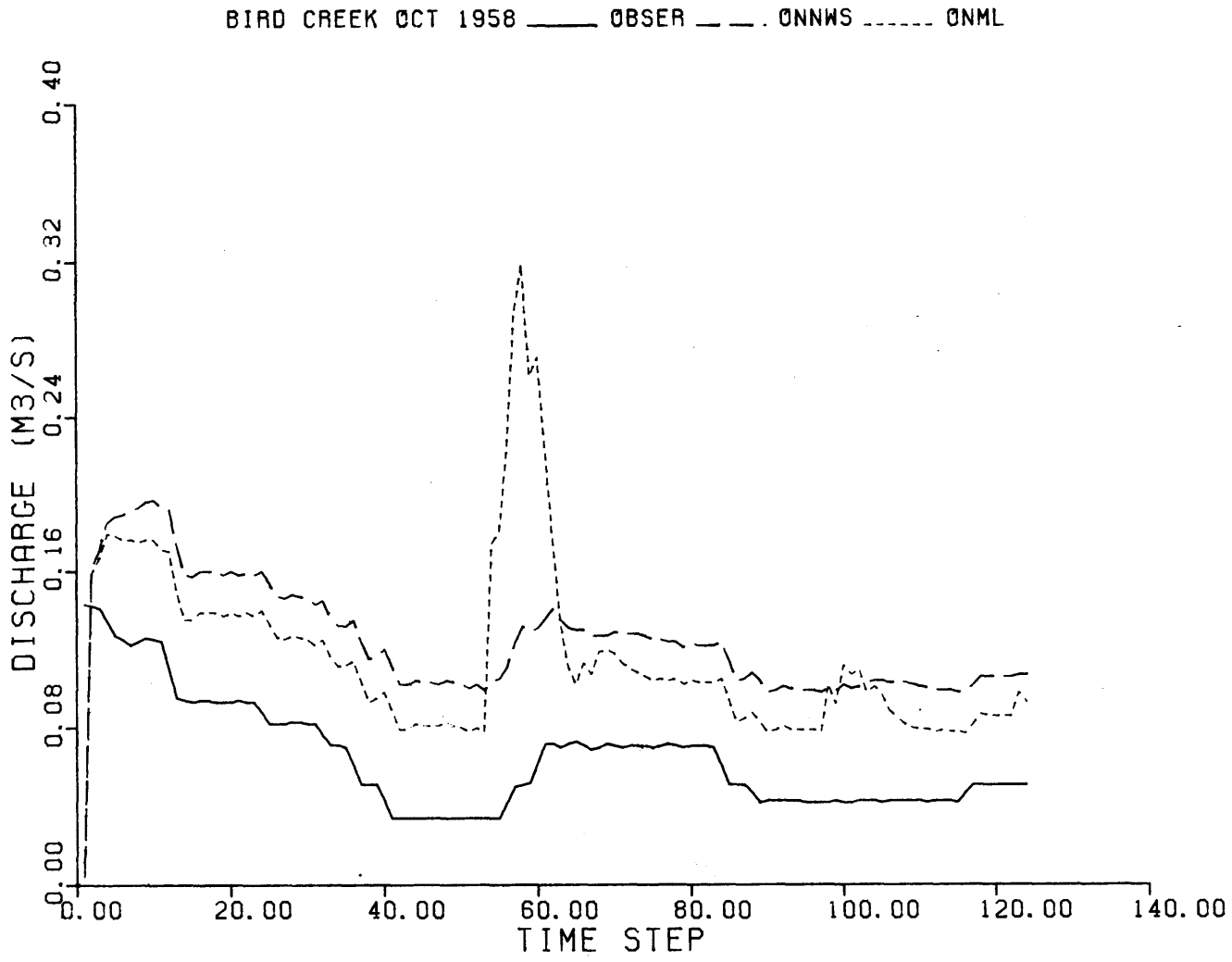


Figure B.10 Six-Hour Lead On-Line Forecasts for Bird Creek, October 1958.

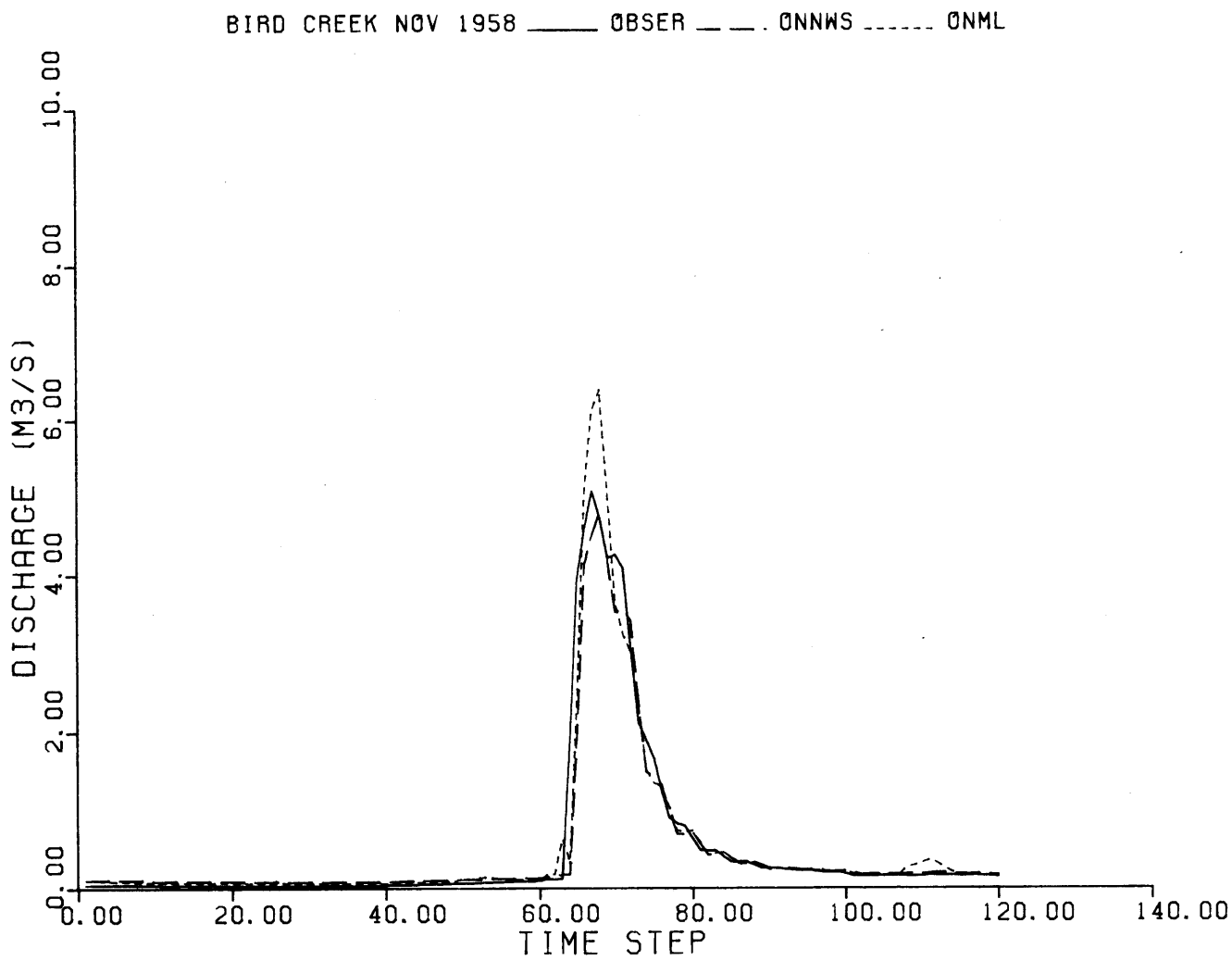


Figure B.11 Six-Hour Lead On-Line Forecasts for Bird Creek, November 1958.

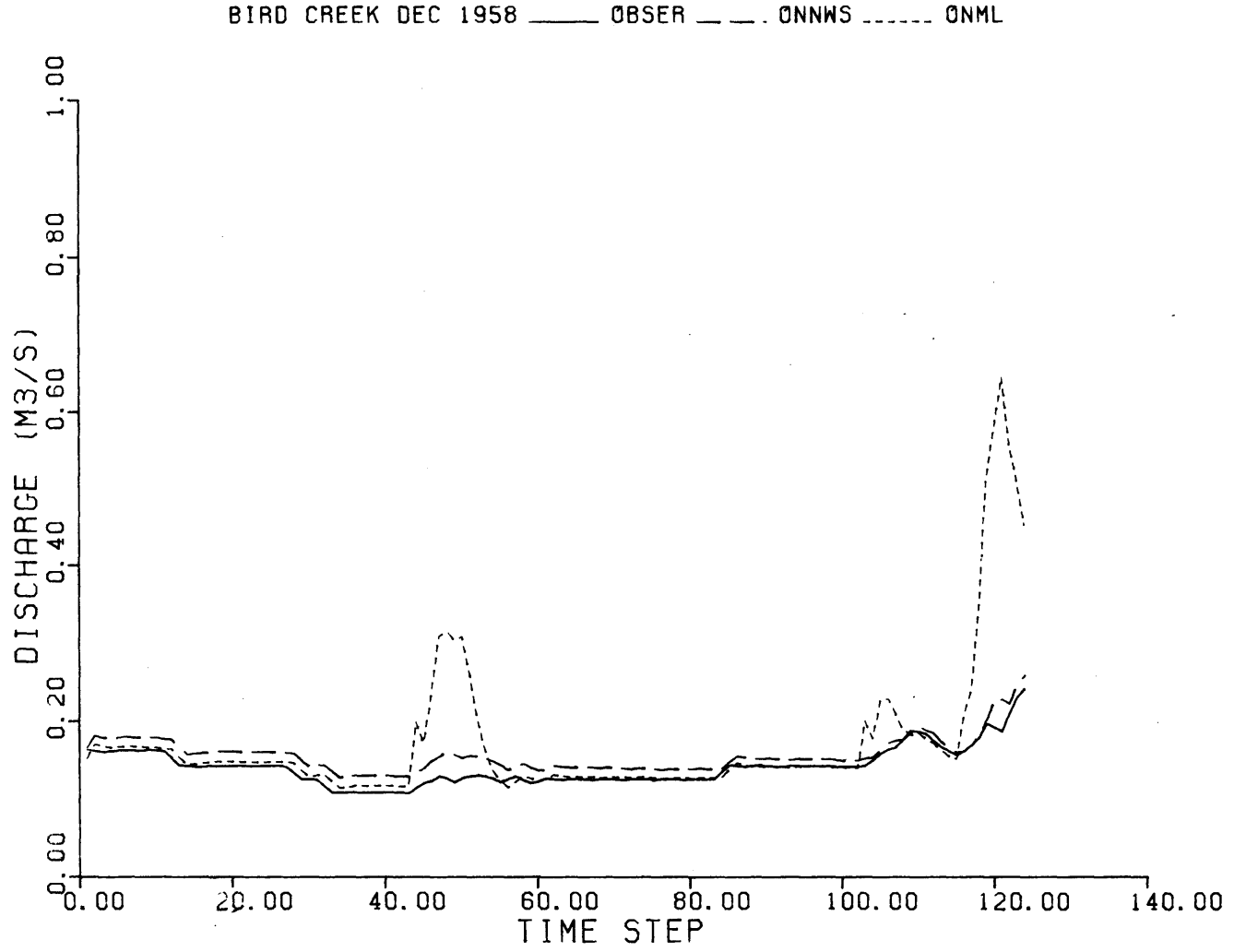


Figure B.12 Six-Hour Lead On-Line Forecasts for Bird Creek, December 1958.

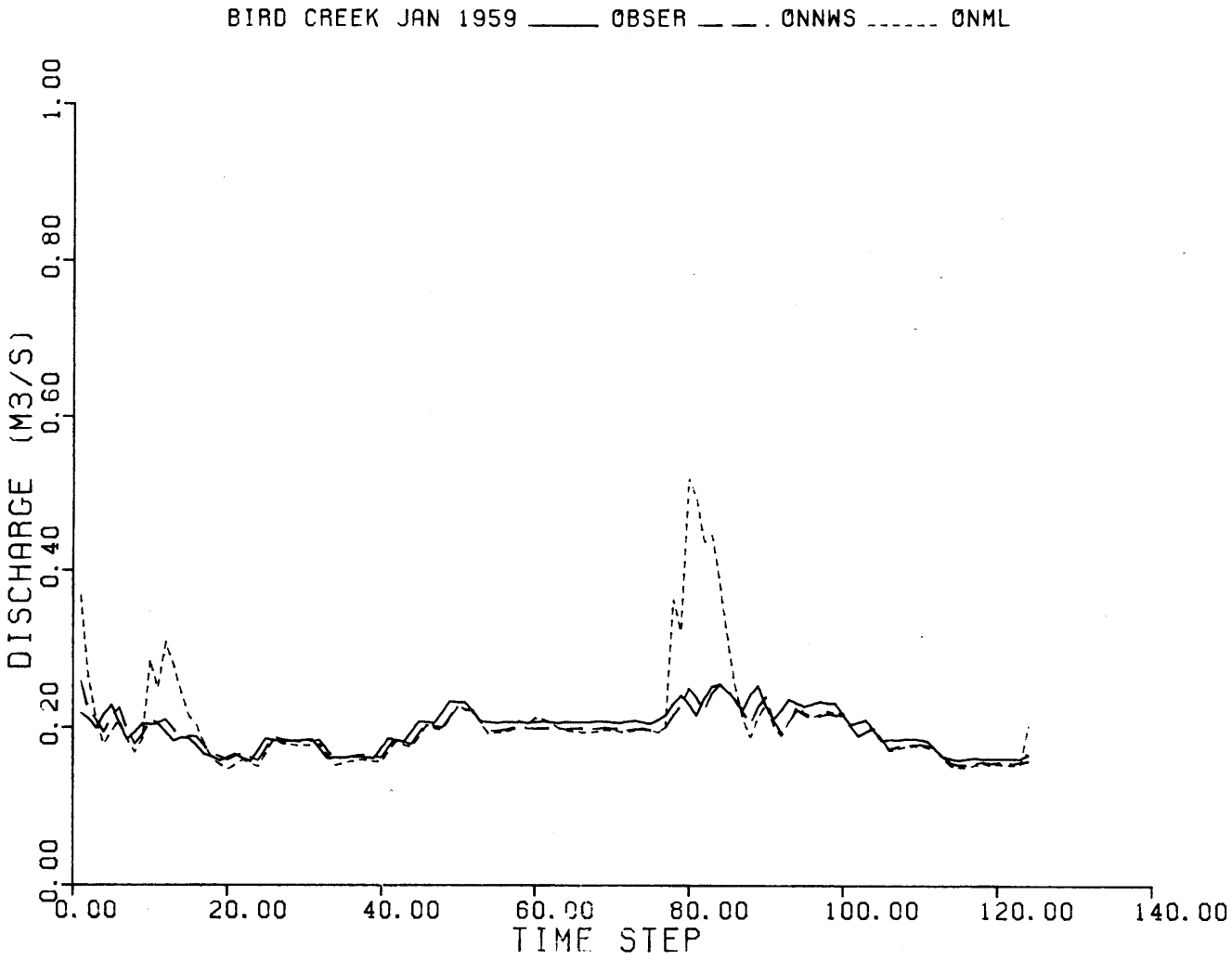


Figure B.13 Six-Hour Lead On-Line Forecasts for Bird Creek, January 1959.

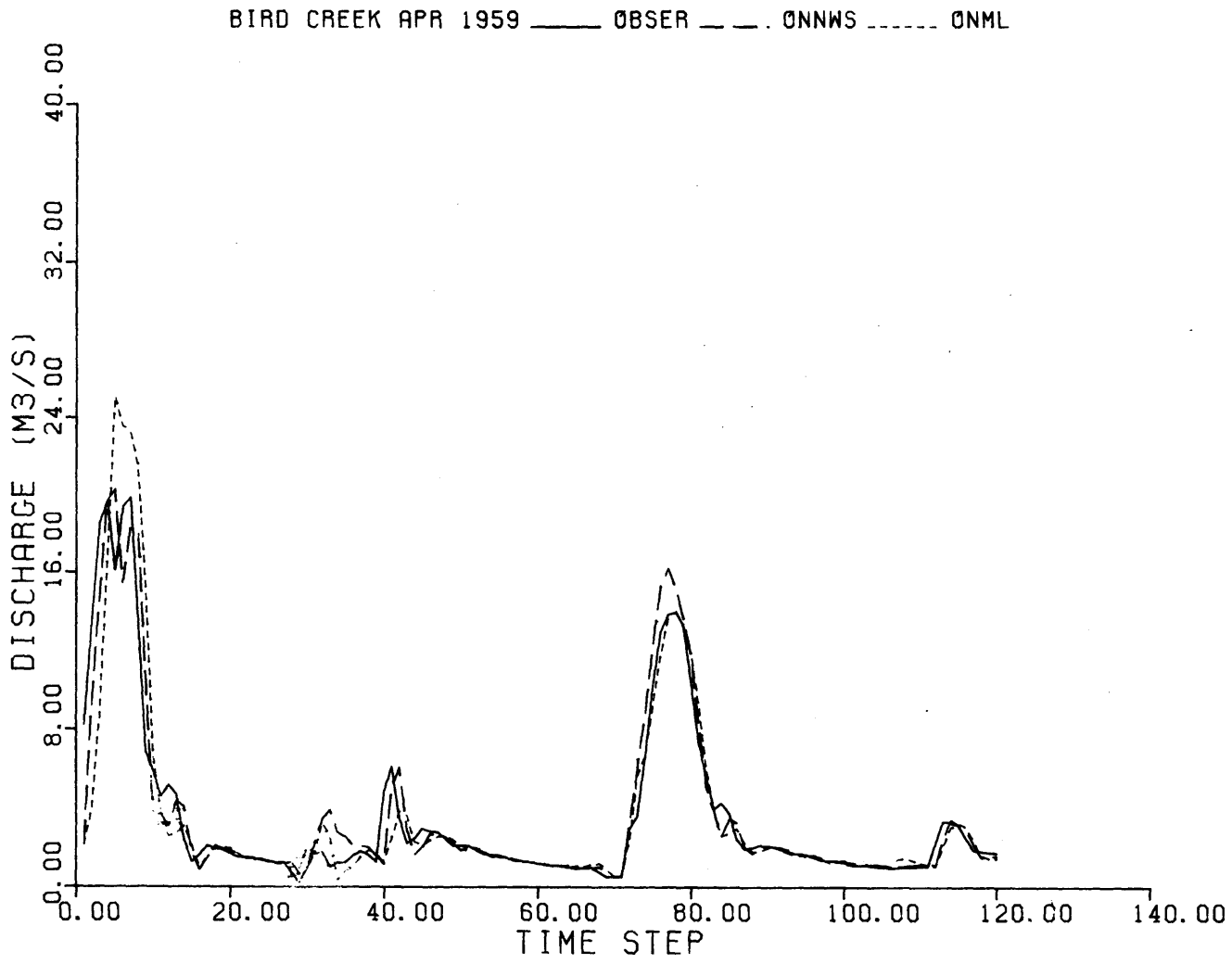


Figure B.14 Six-Hour Lead On-Line Forecasts for Bird Creek, April 1959.

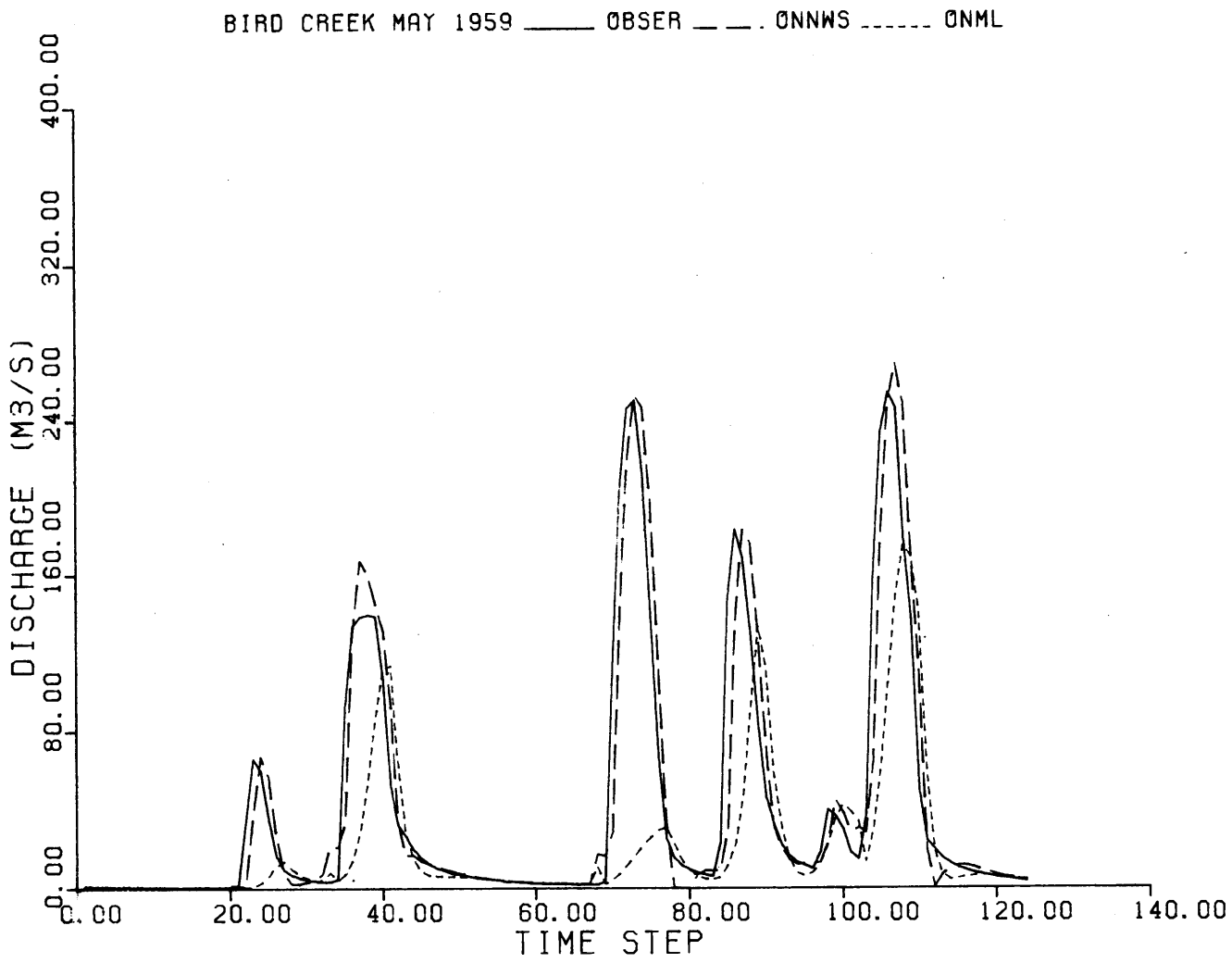


Figure B.15 Six-Hour Lead On-Line Forecasts for Bird Creek, May 1959.

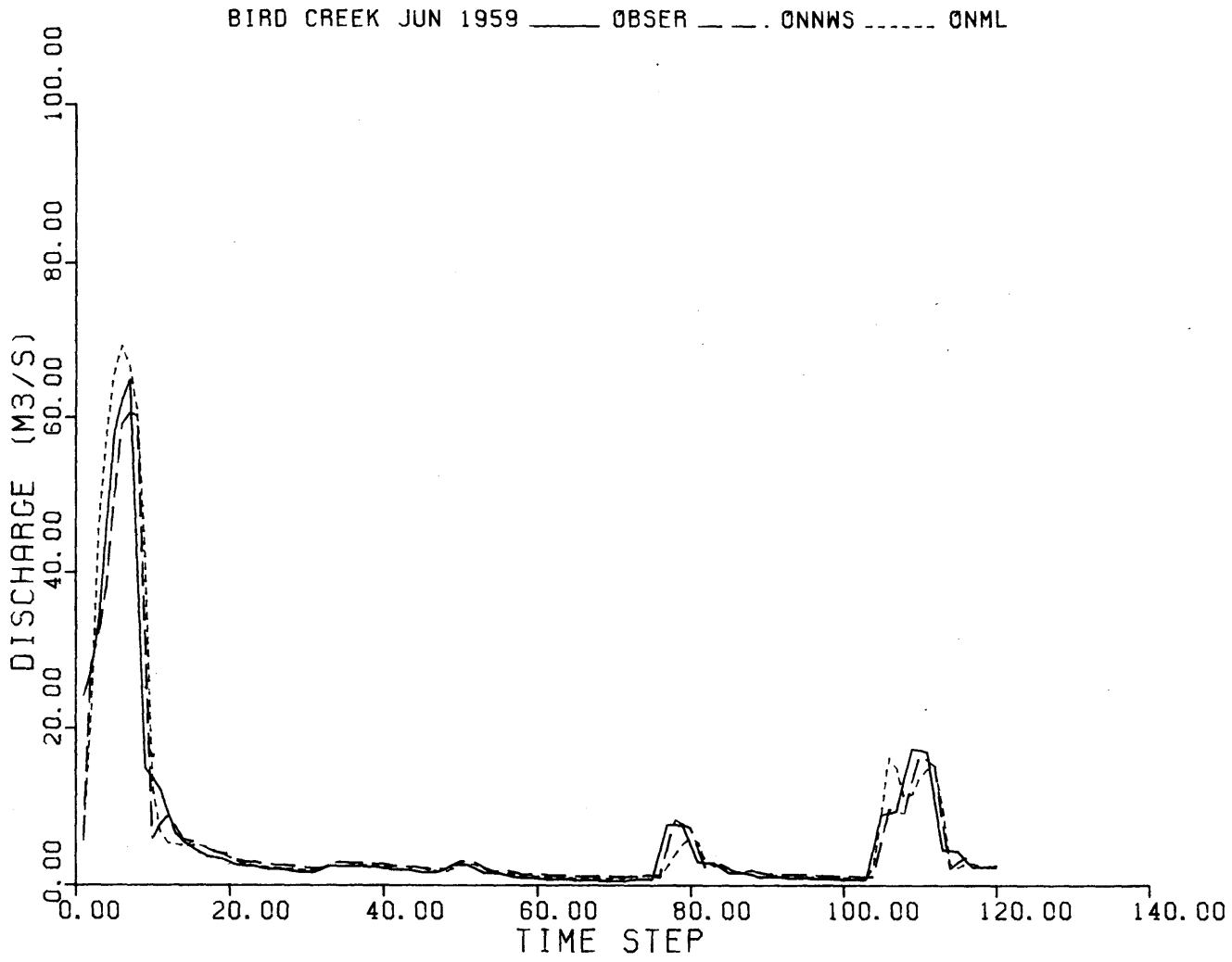


Figure B.16 Six-Hour Lead On-Line Forecasts for Bird Creek, June 1959.

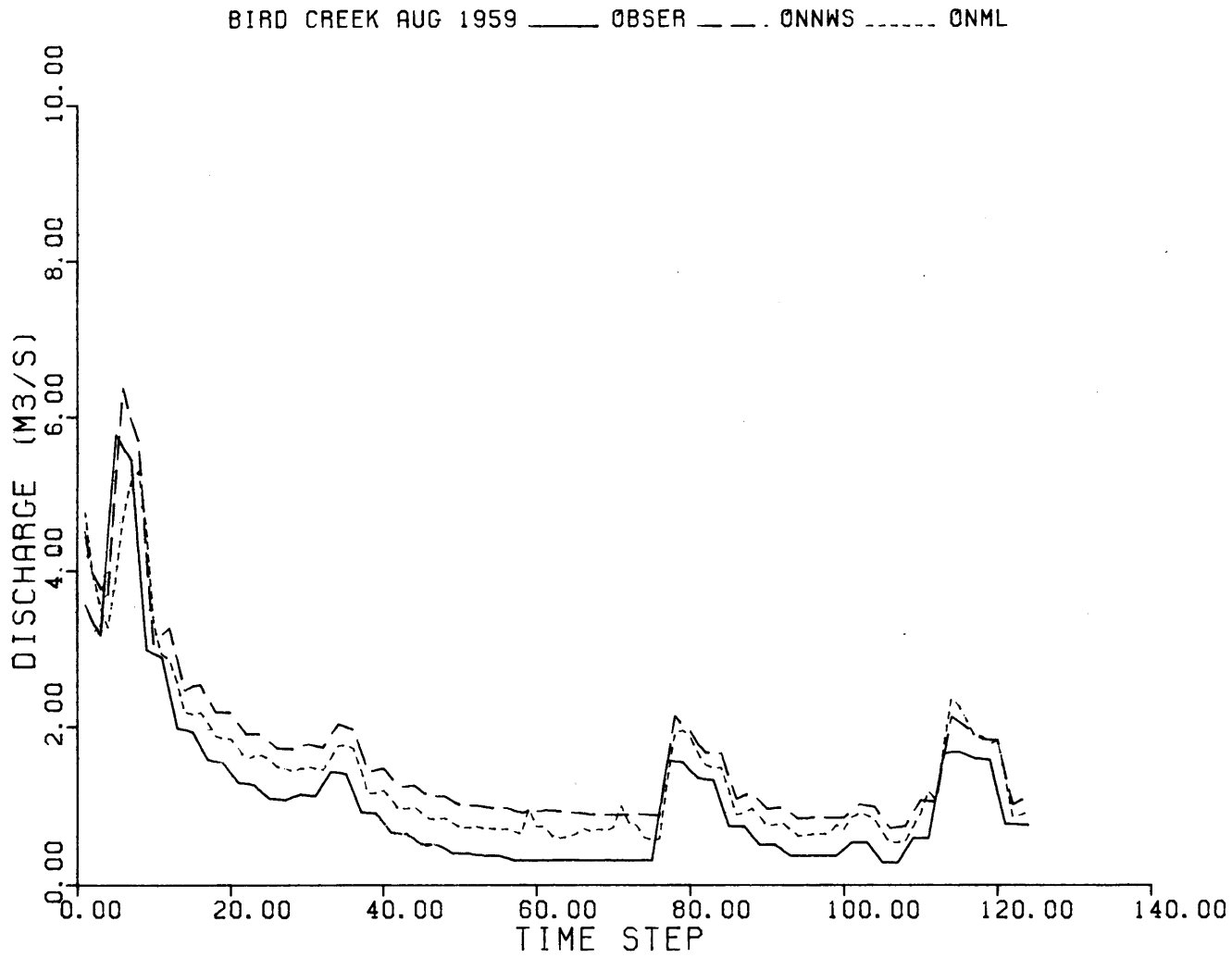


Figure B.17 Six-Hour Lead On-Line Forecasts for Bird Creek, August 1959.

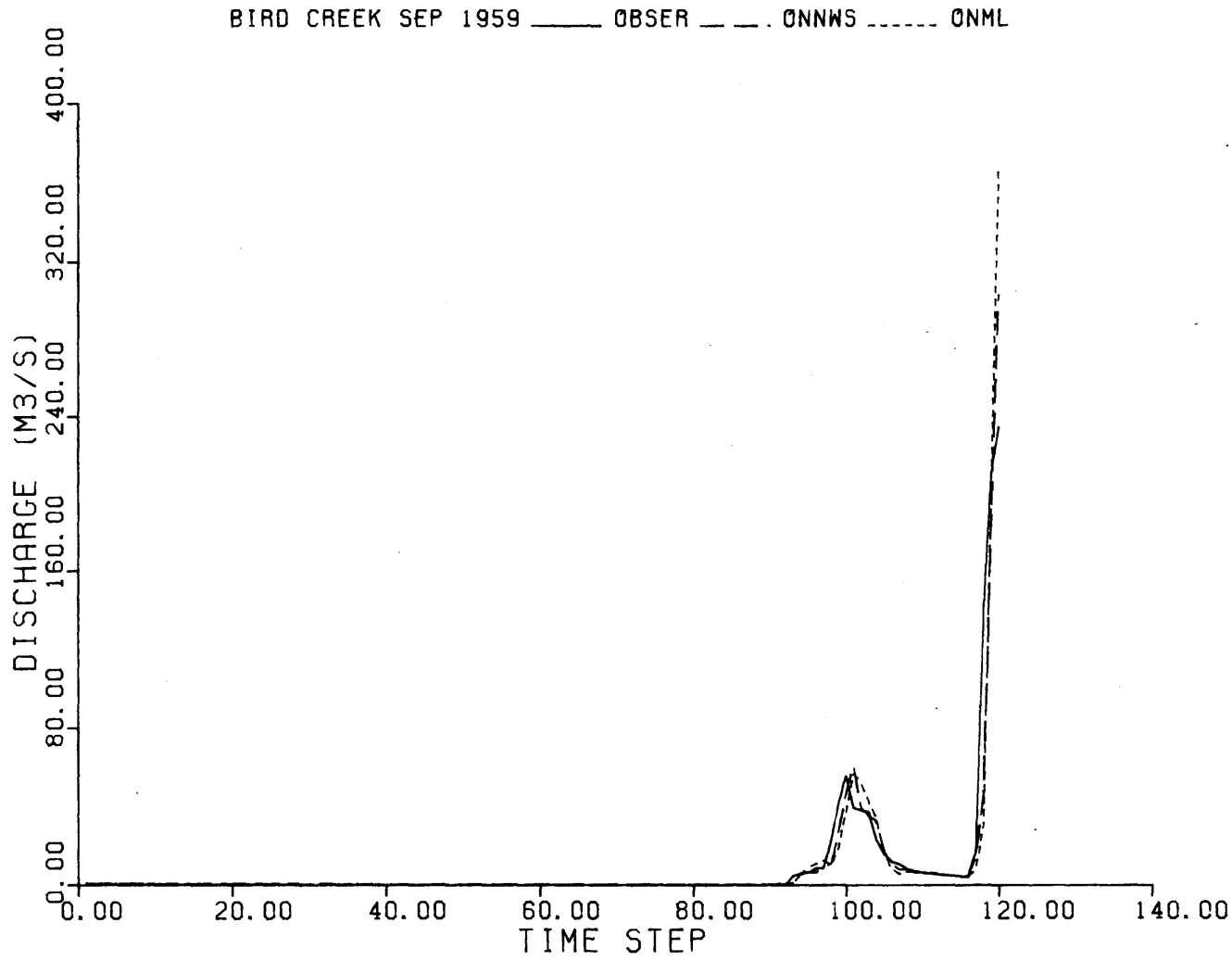


Figure B.18 Six-Hour Lead On-Line Forecasts for Bird Creek, September 1959.

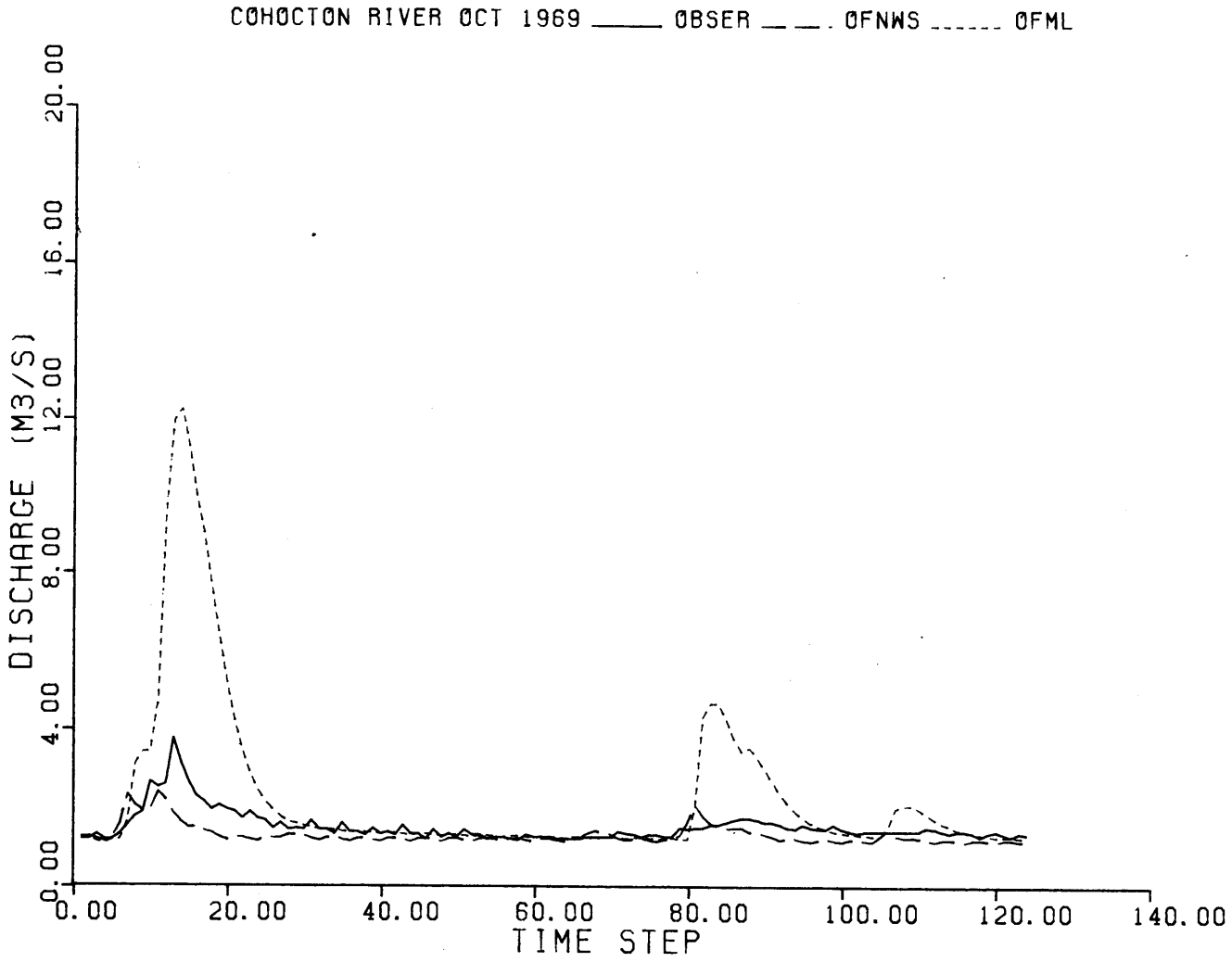


Figure B.19 Six-Hour Lead Off-Line Forecasts for the Cohocton River, October 1969.

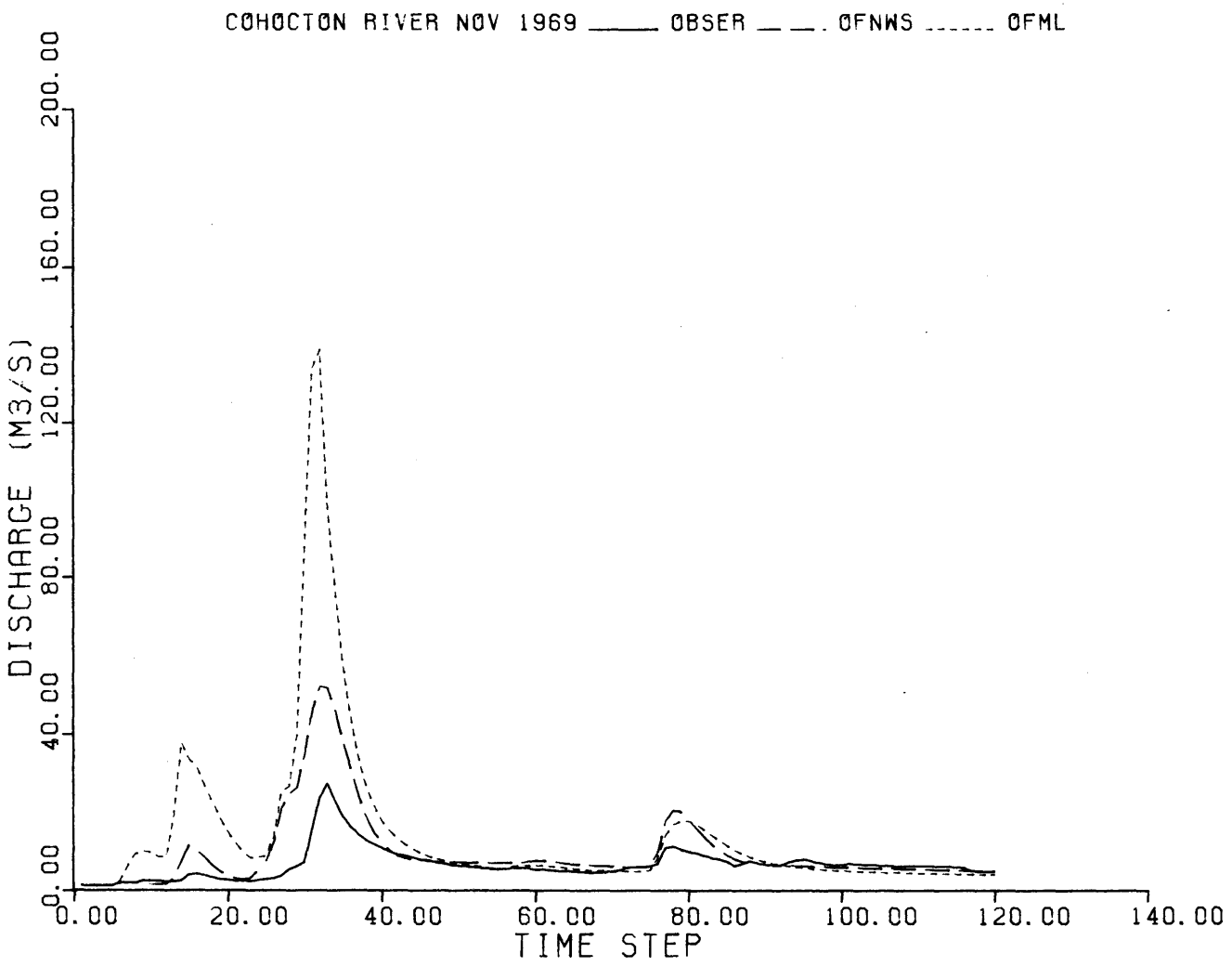


Figure B.20 Six-Hour Lead Off-Line Forecasts for the Cohocton River, November 1969.

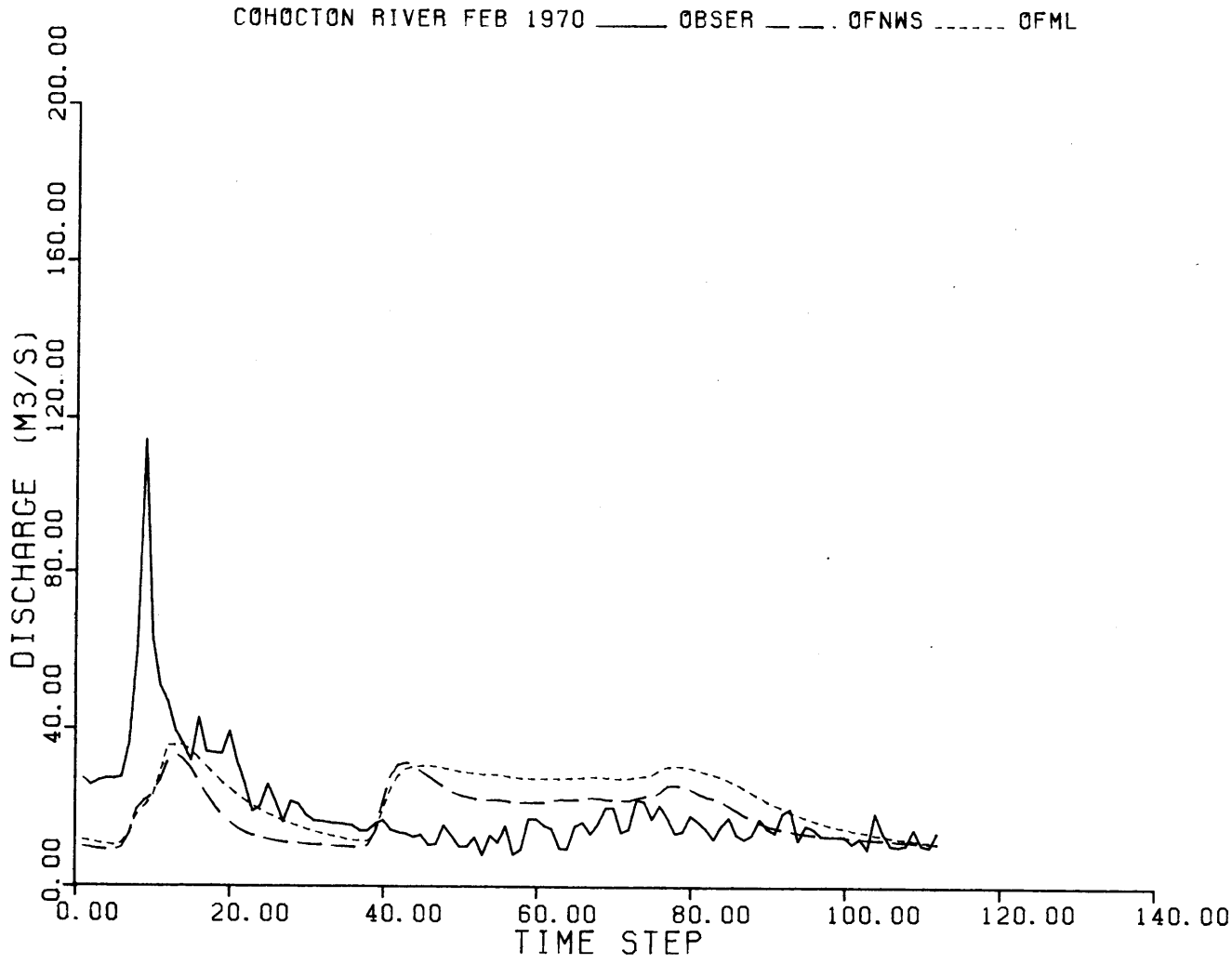


Figure B.21 Six-Hour Lead Off-Line Forecasts for the Cohocton River, February 1970.

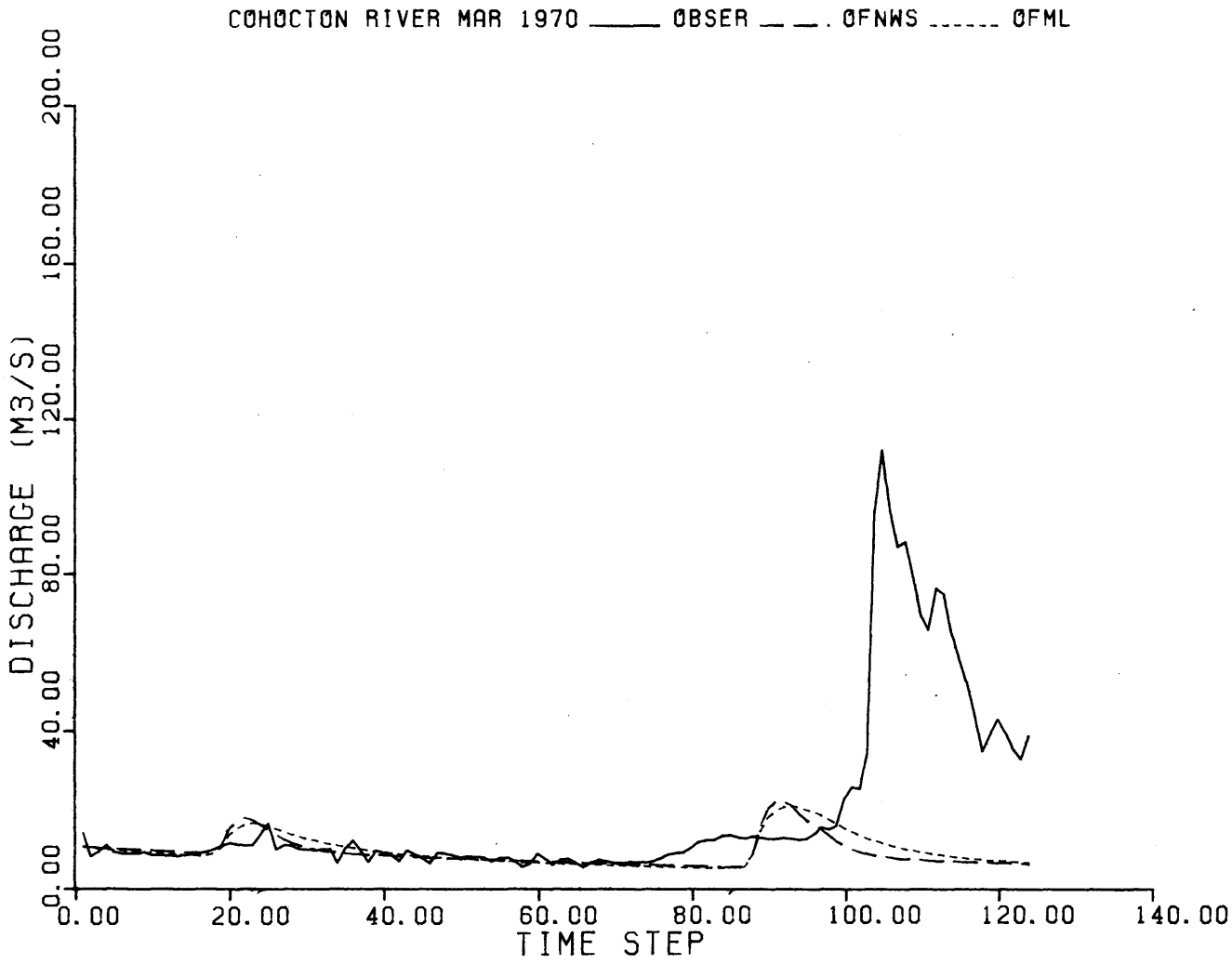


Figure B.22 Six-Hour Lead Off-Line Forecasts for the Cohocton River, March 1970.

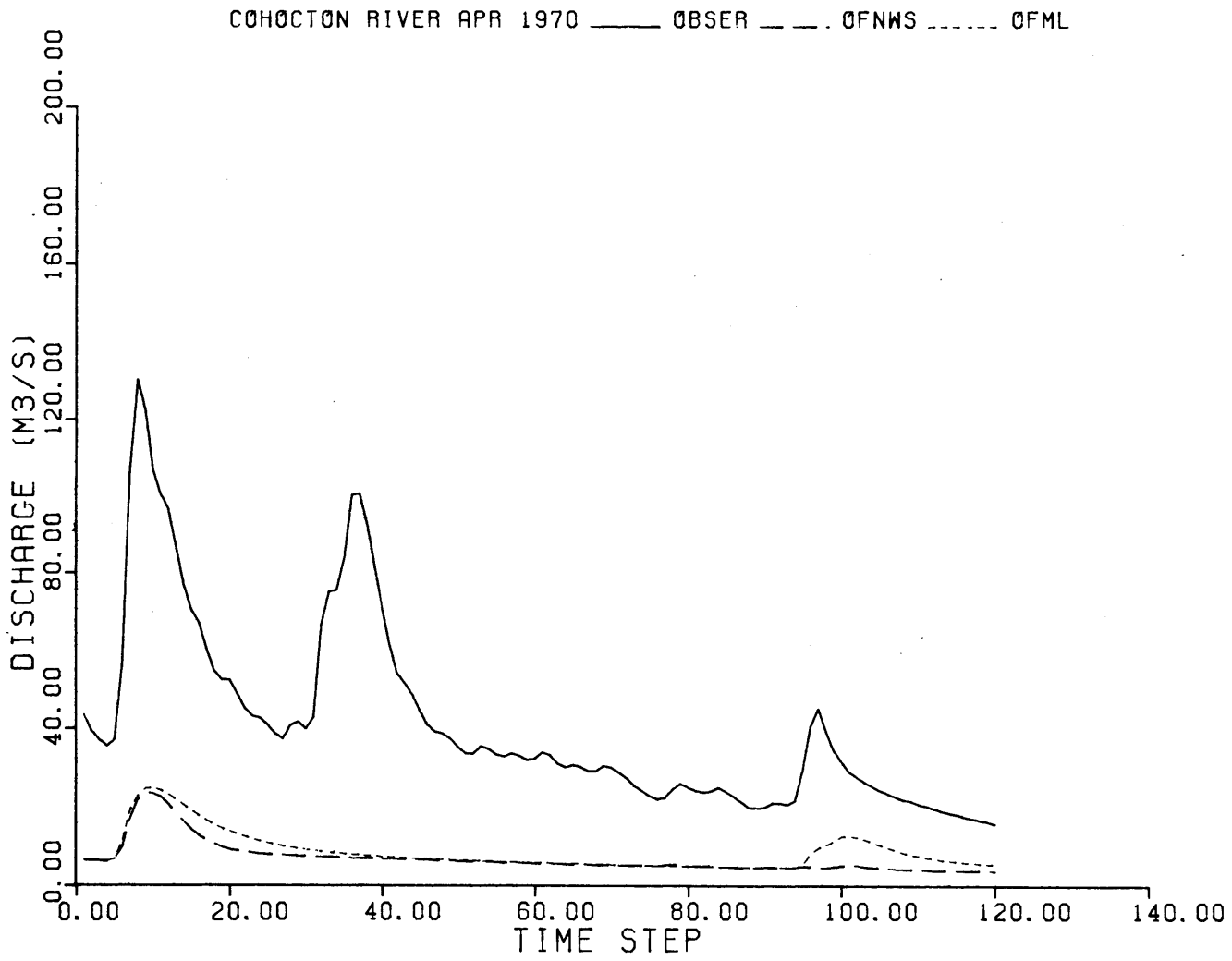


Figure B.23 Six-Hour Lead Off-Line Forecasts for the Cohocton River, April 1970.

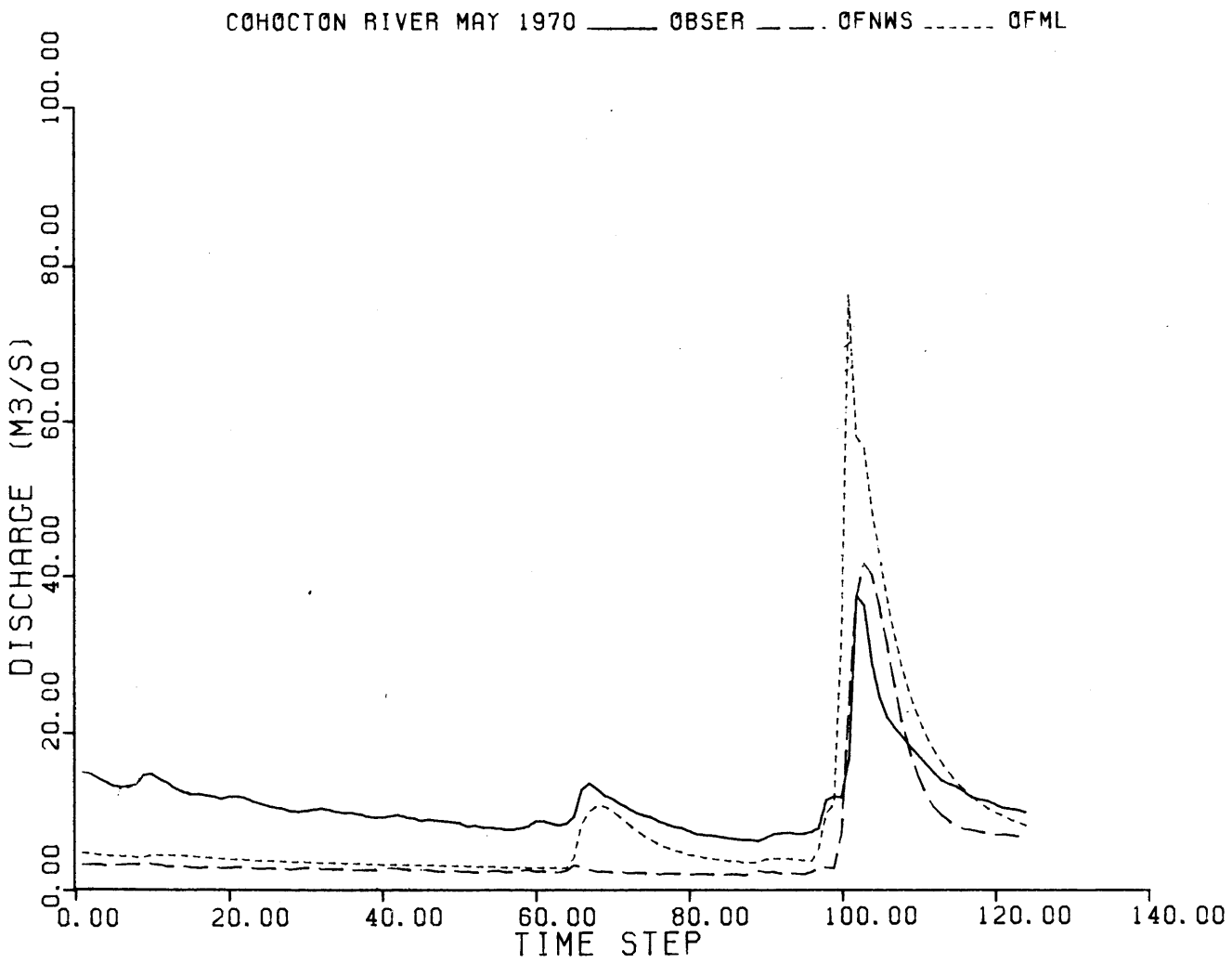


Figure B.24 Six-Hour Lead Off-Line Forecasts for the Cohocton River, May 1970.

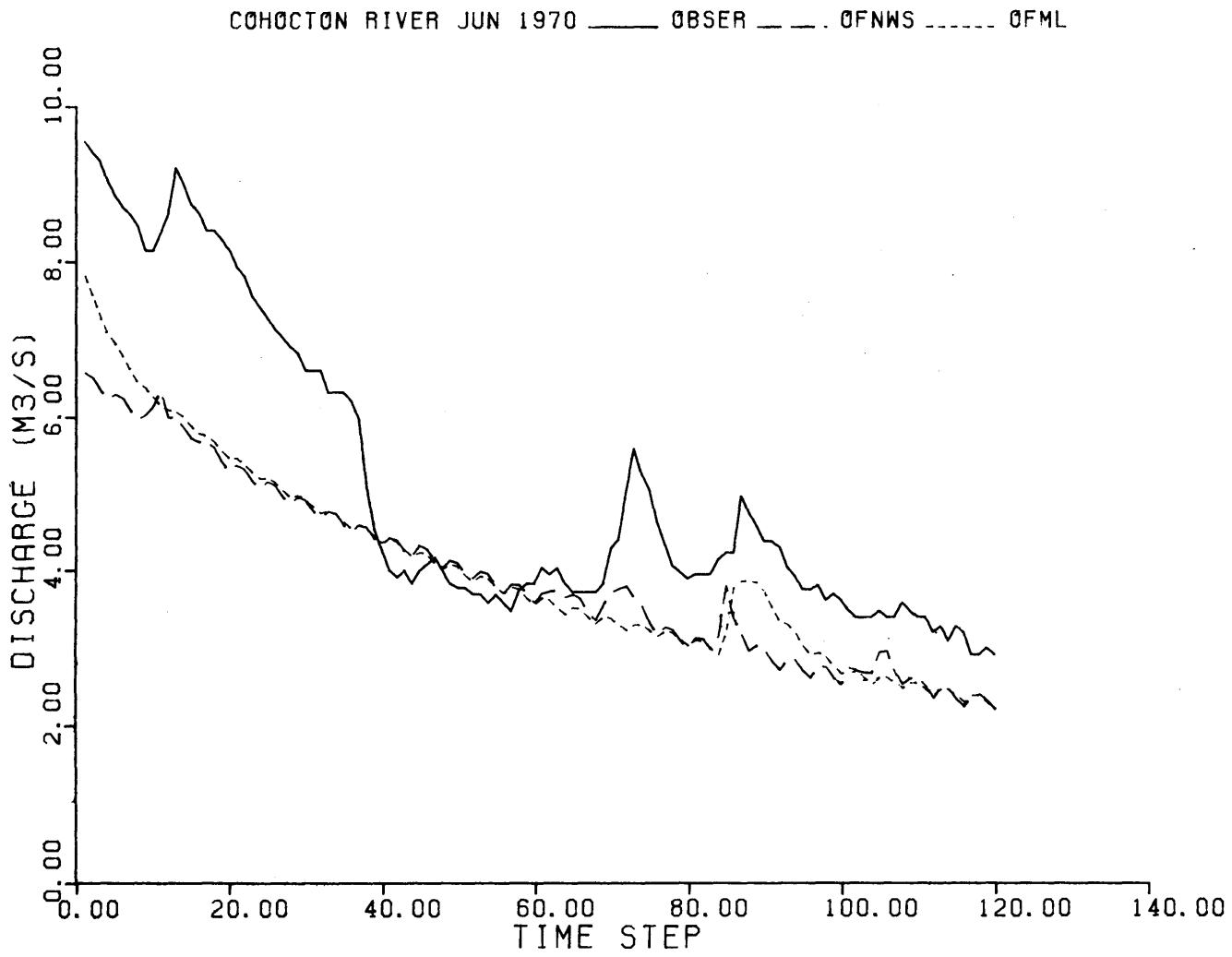


Figure B.25 Six-Hour Lead Off-Line Forecasts for the Cohocton River, June 1970.

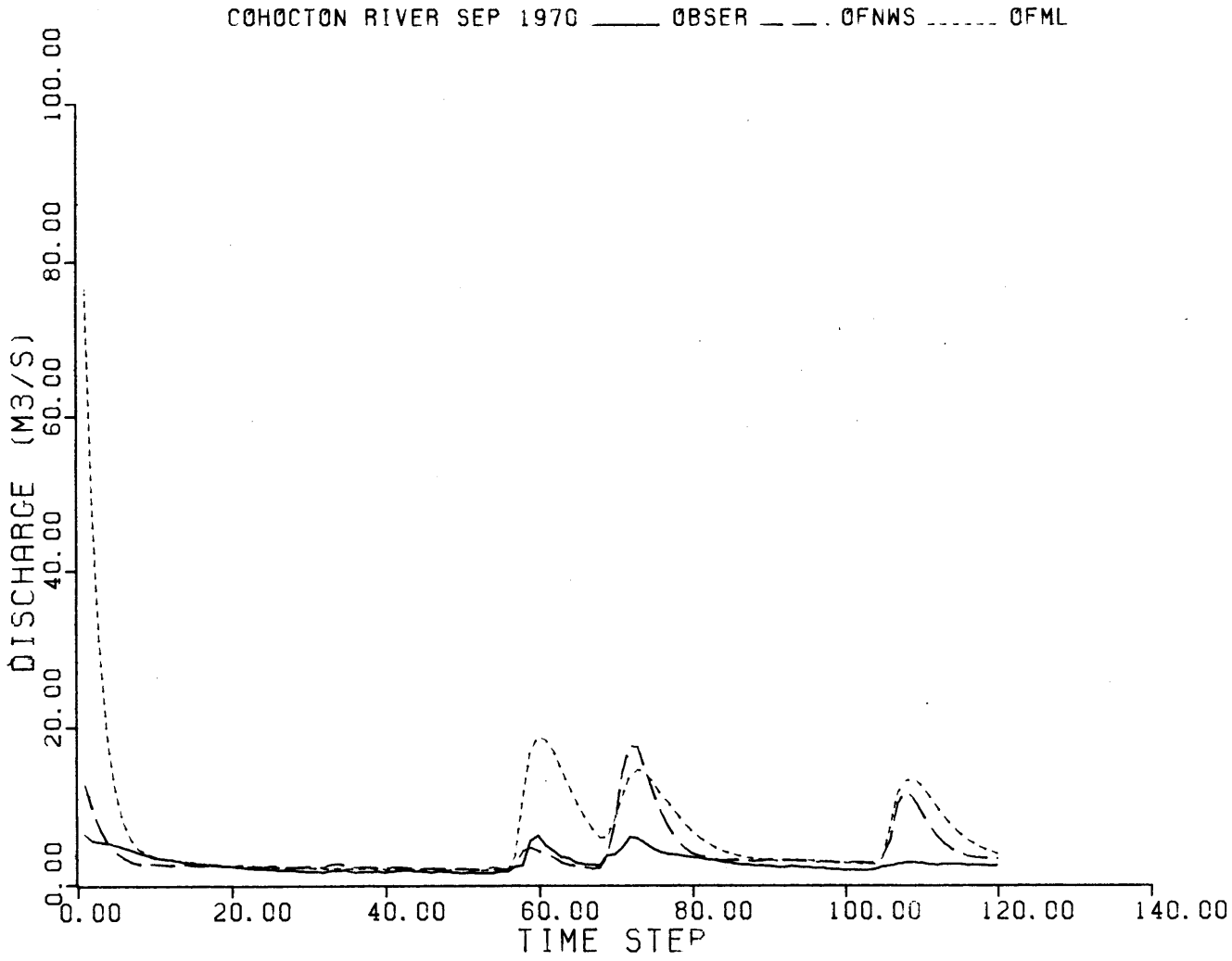


Figure B.26 Six-Hour Lead Off-Line Forecasts for the Cohocton River, September 1970.

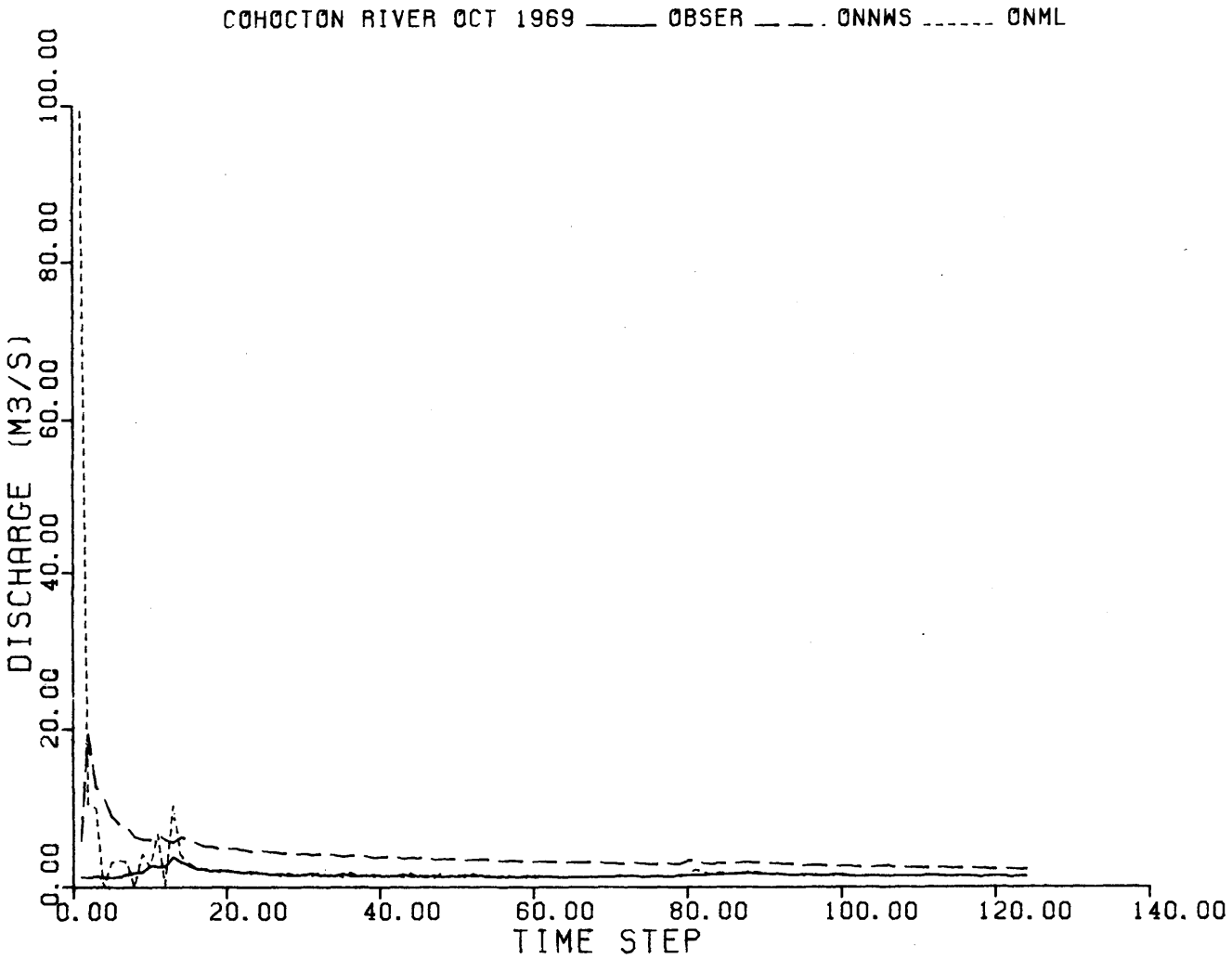


Figure B.27 Six-Hour Lead On-Line Forecasts for the Cohocton River, October 1969.

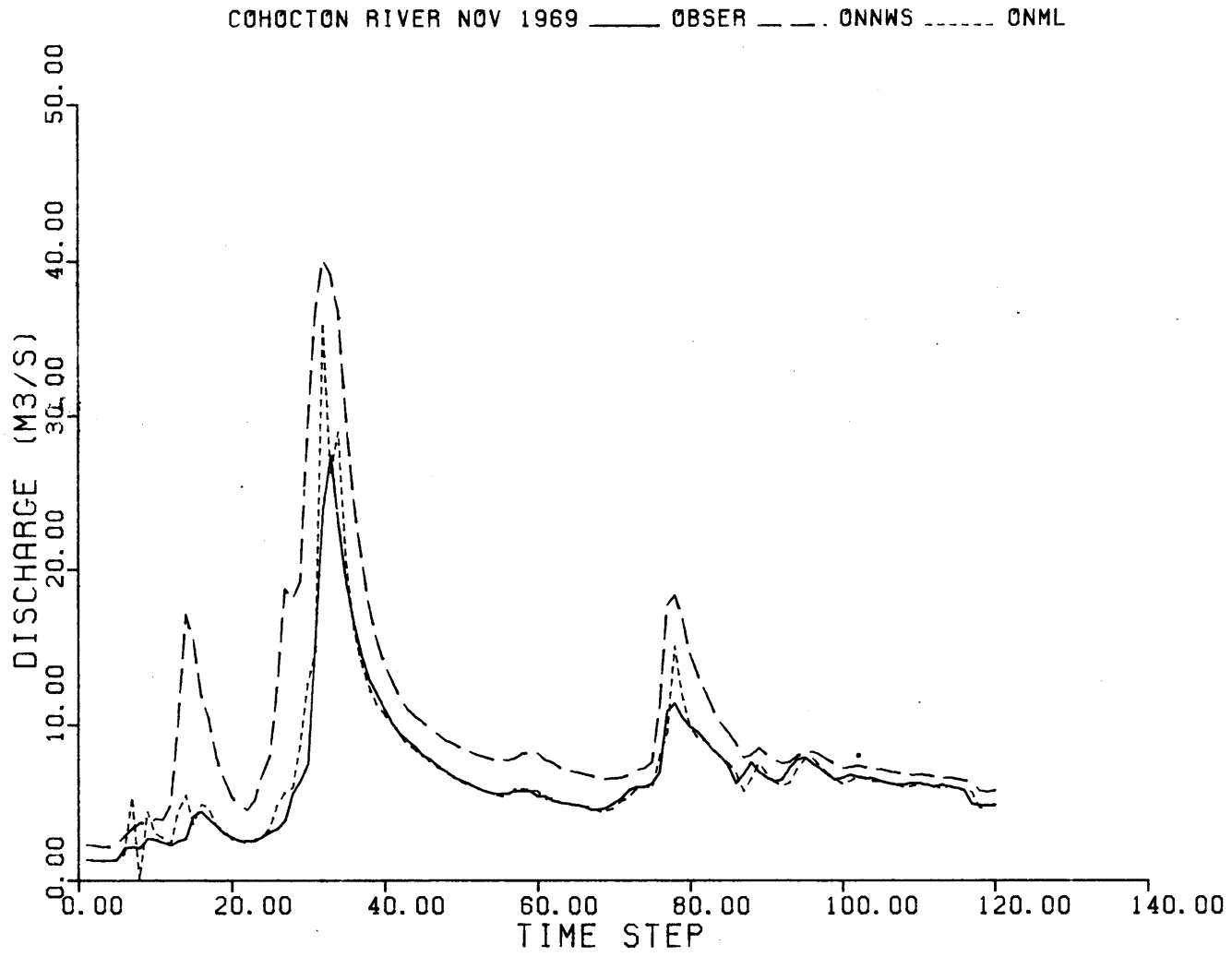


Figure B.28 Six-Hour Lead On-Line Forecasts for the Cohocton River, November 1969.

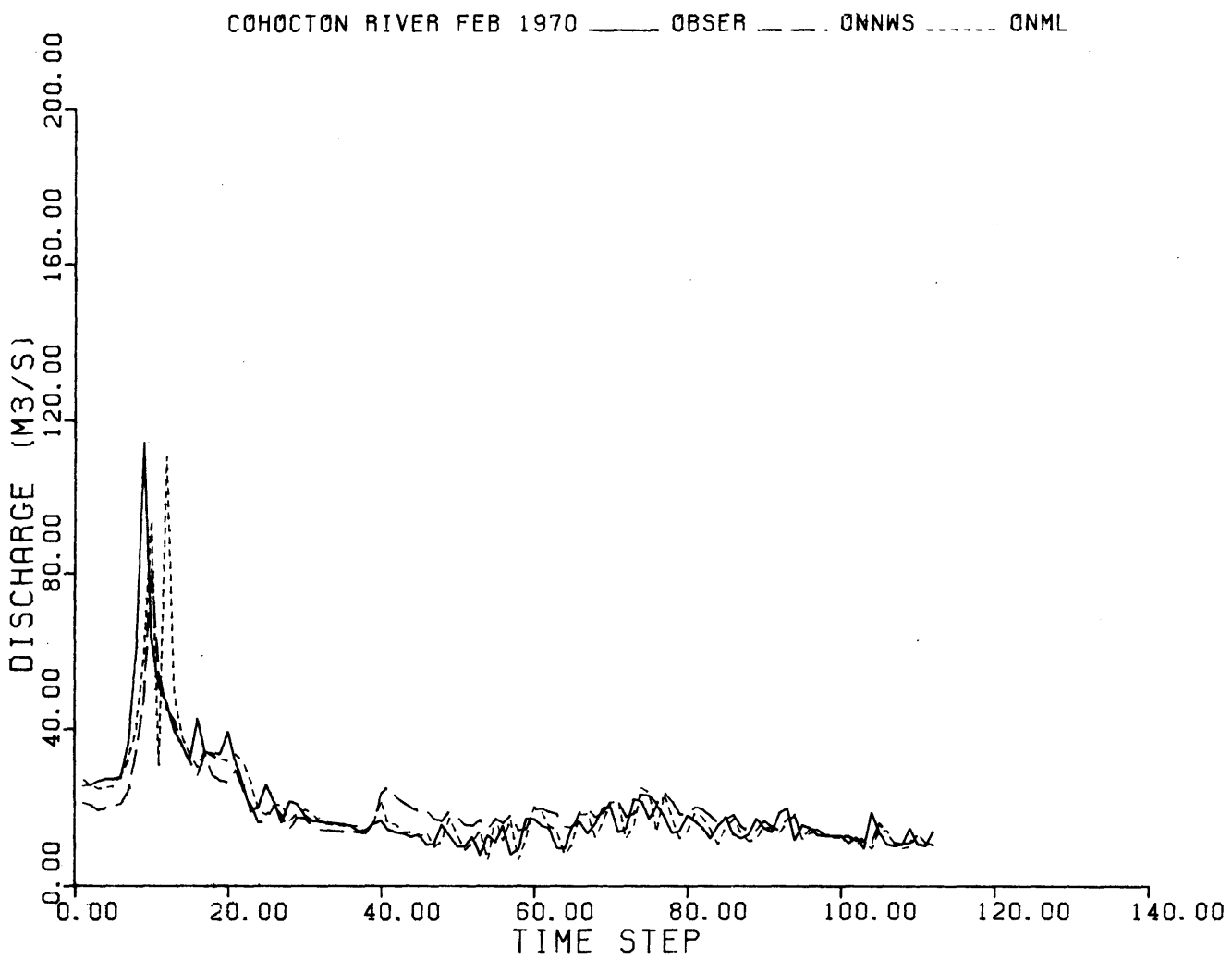


Figure B.29 Six-Hour Lead On-Line Forecasts for the Cohocton River, February 1970.

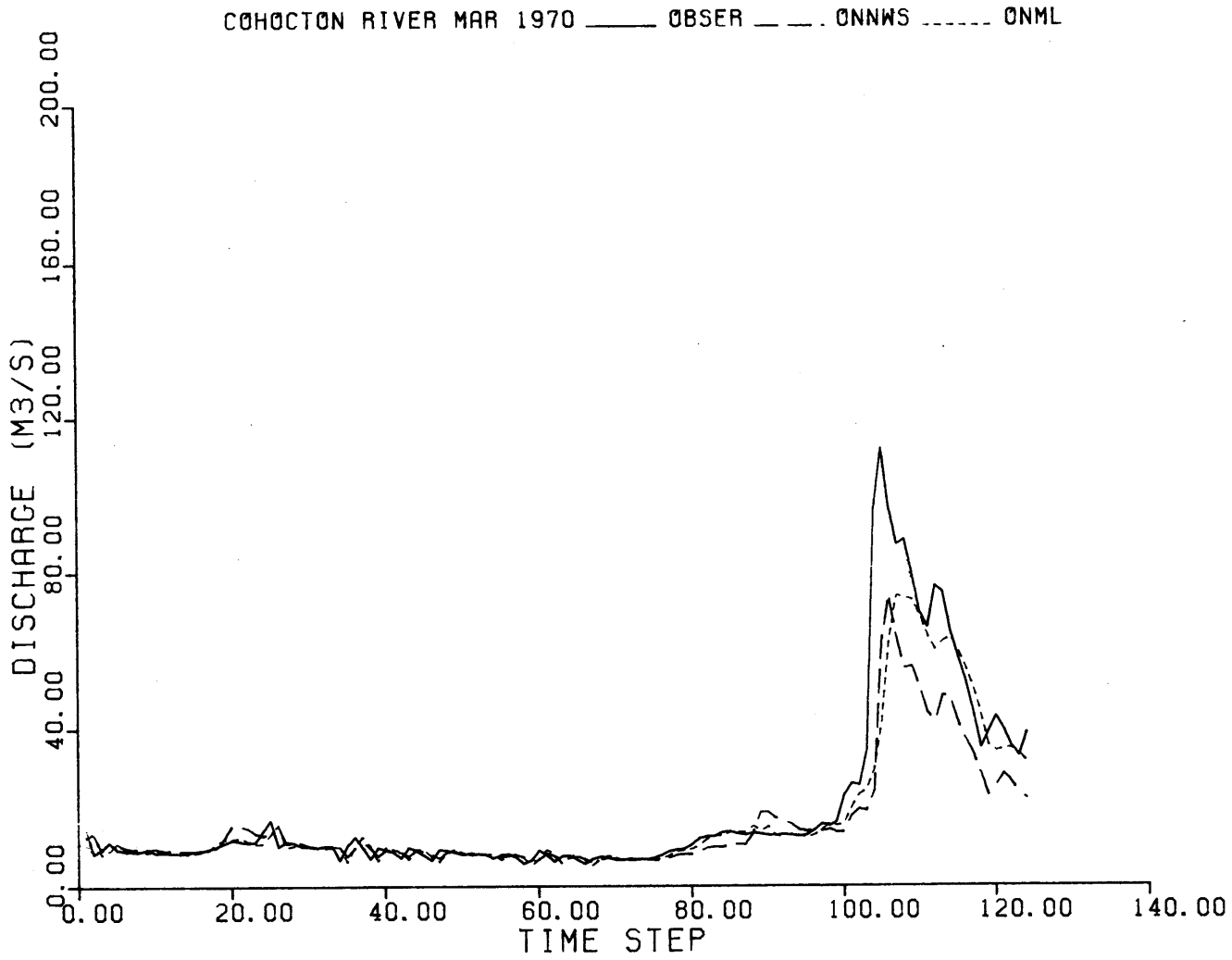


Figure B.30 Six-Hour Lead On-Line Forecasts for the Cohocton River, March 1970.

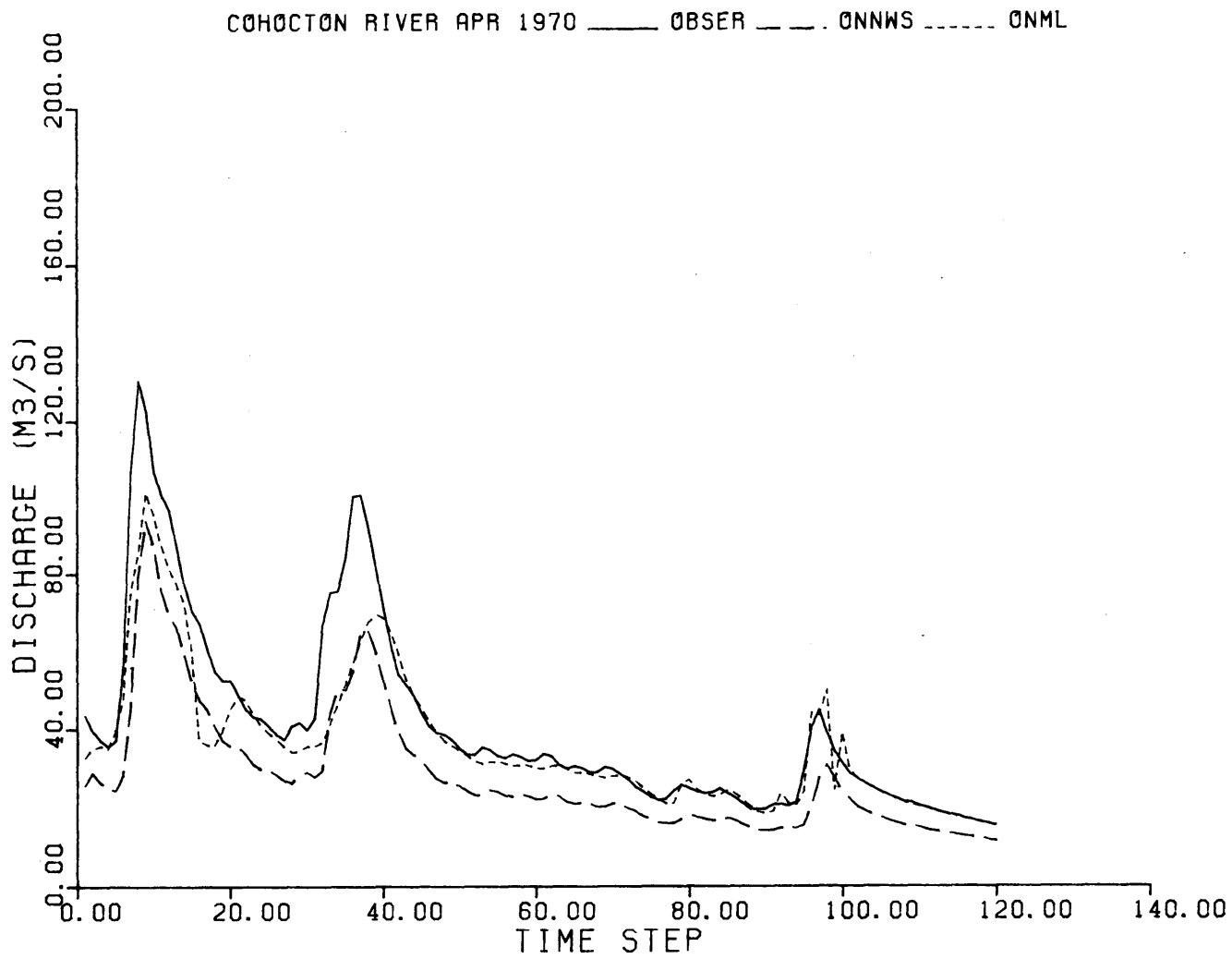


Figure B.31 Six-Hour Lead On-Line Forecasts for the Cohocton River, April 1970.

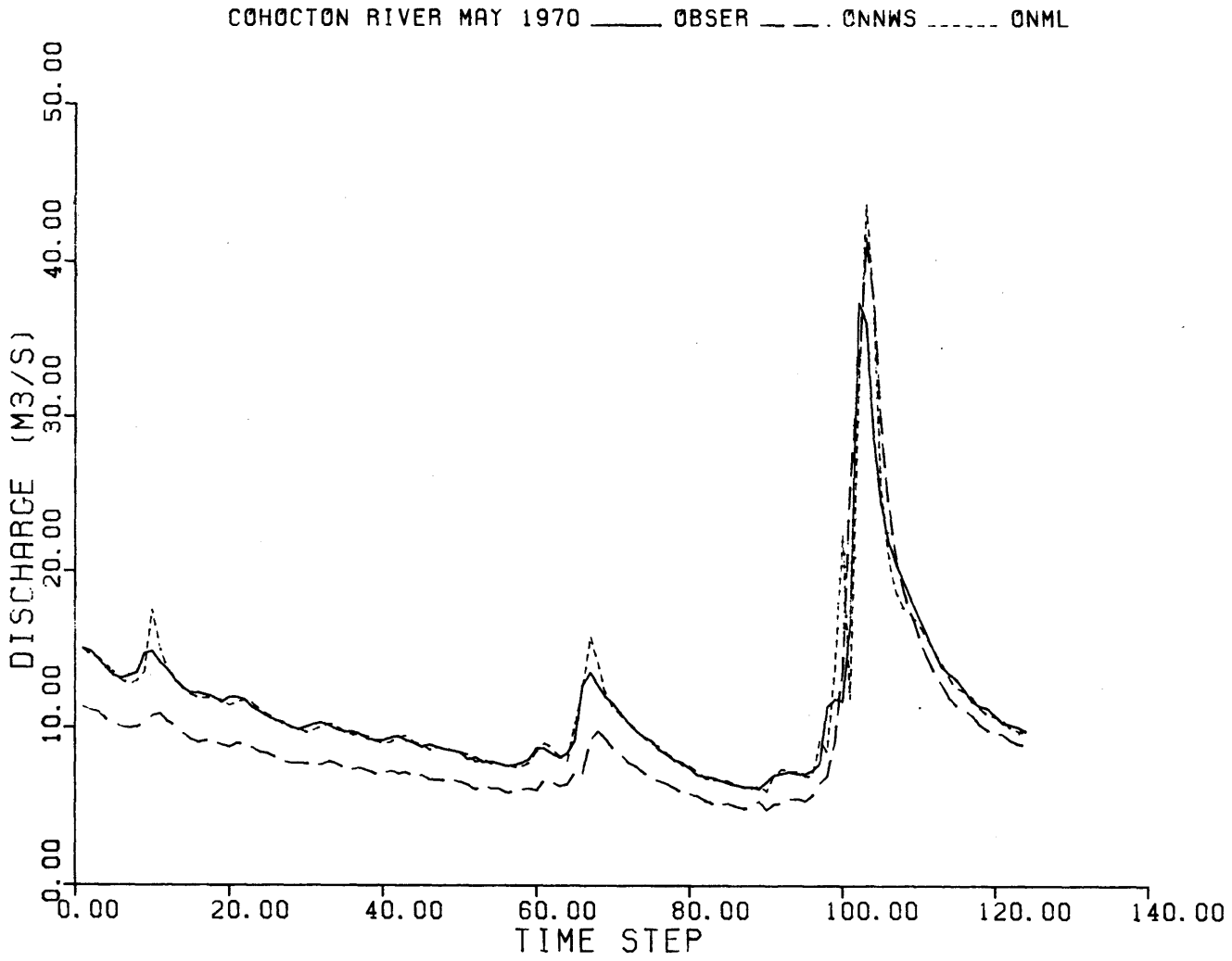


Figure B.32 Six-Hour Lead On-Line Forecasts for the Cohocton River, May 1970.

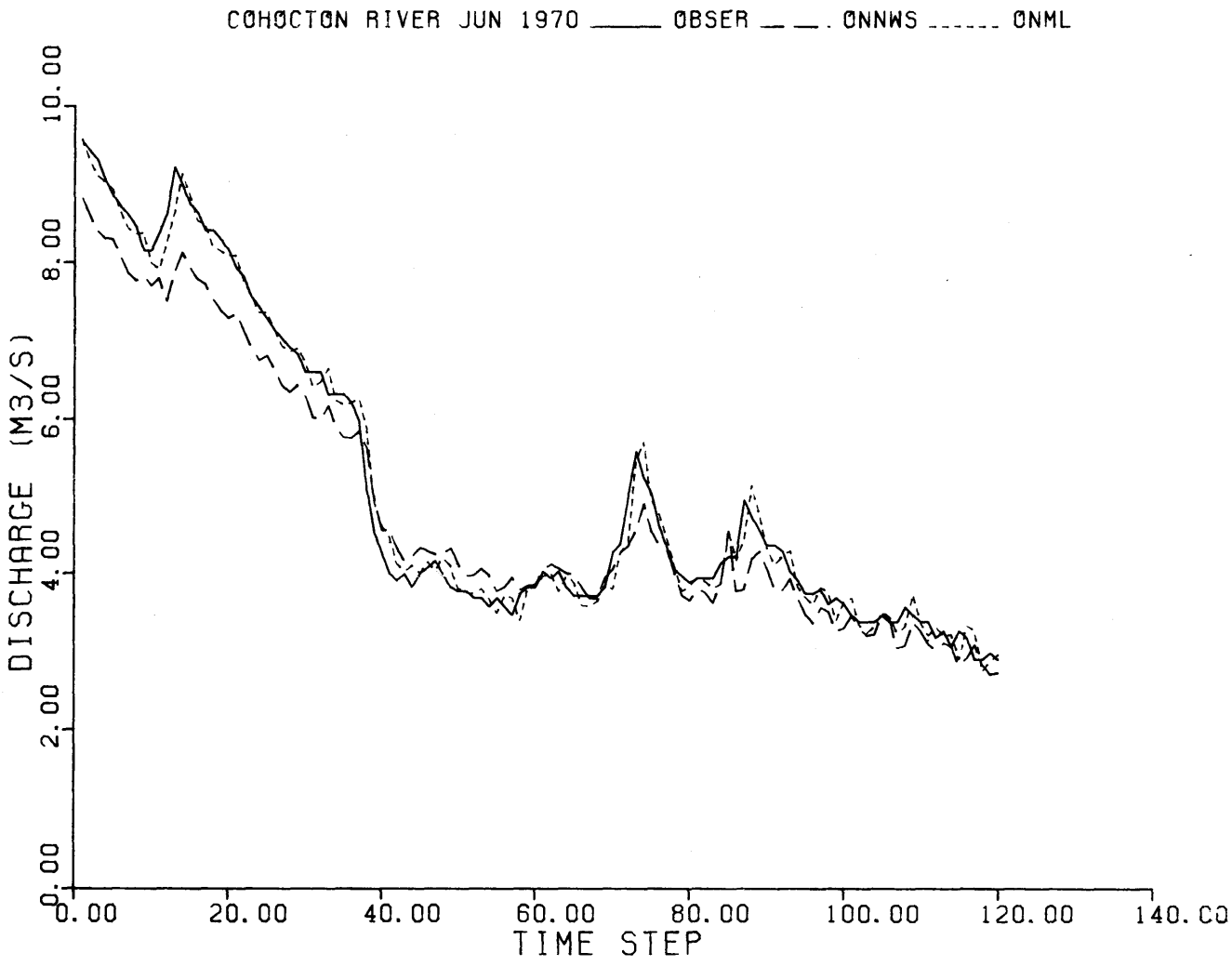


Figure B.33 Six-Hour Lead On-Line Forecasts for the Cohocton River, June 1970.

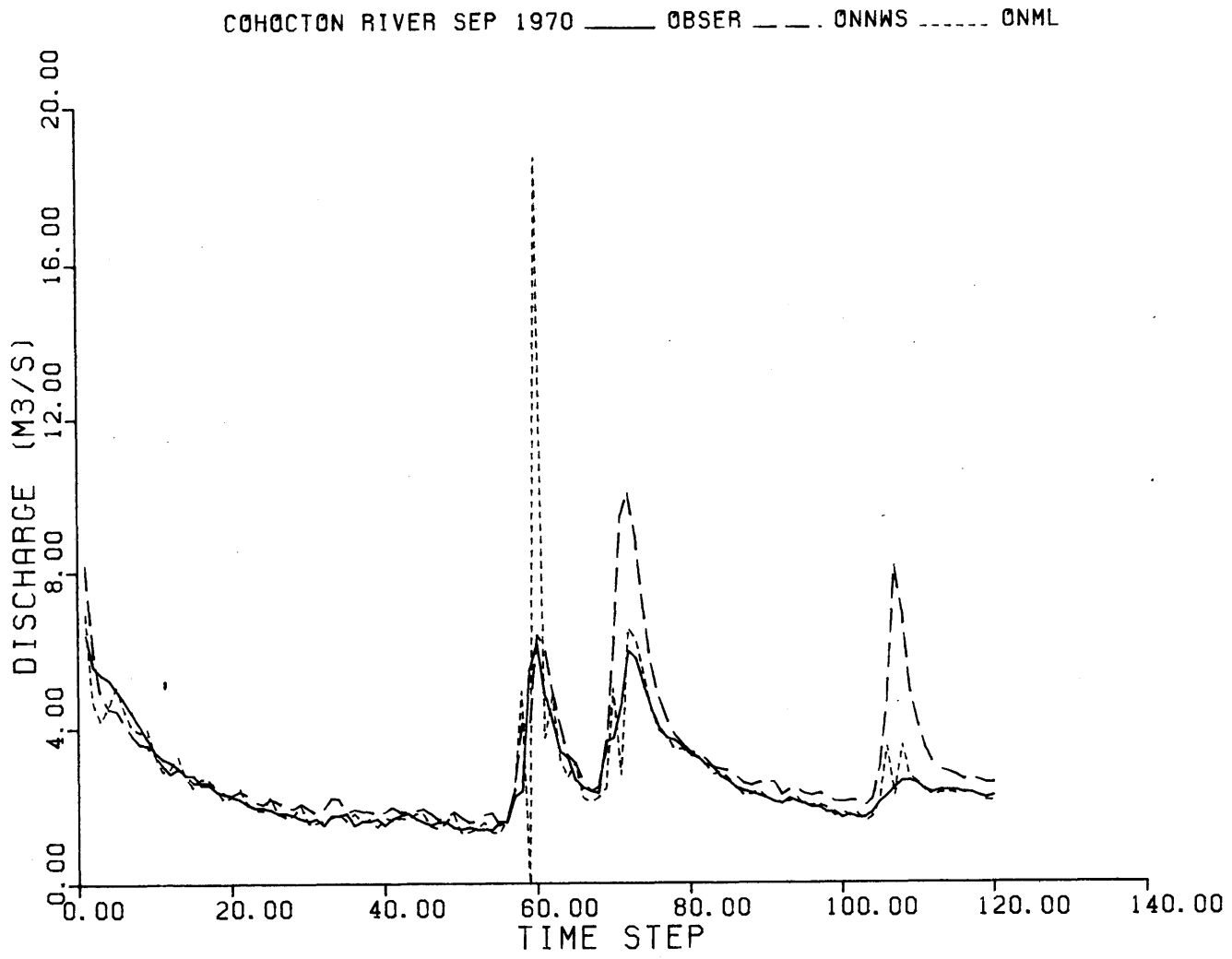


Figure B.34 Six-Hour Lead On-Line Forecasts for the Cohocton River, September 1970.

APPENDIX C

Stochastic Model of Base Flow With Constant Channel Losses.

Let the deterministic continuous time equations of the model of base flow be

$$dx_1(t)/dt = -d_1' \; x_1(t) \qquad (C.1)$$

$$dx_2(t)/dt = -d_1'' \; x_2(t) \qquad (C.2)$$

Integrating Equations (C.1) and (C.2) from time t to time $t+\Delta t$ and adding the discrete time white gaussian noise processes, the discrete time system equations become,

$$x_1(t + \Delta t) = A_1 \; x_1(t) + W_1(t) \qquad (C.3)$$

$$x_2(t + \Delta t) = A_2 \; x_2(t) + W_2(t) \qquad (C.4)$$

$$z(t) = x_1(t) + x_2(t) - S_s + v(t) \qquad (C.5)$$

In which all symbols, except S_s, the constant channel losses, were defined in Chapter 5. S_s is, of course, equivalent to the NWSRFS acronym for SSOUT. The only change in the Kalman filter equations used in the model of base flow of Chapter 5 is in the calculation of $z(t|t-1)$:

$$z(t|t-1) = x_1(t|t-1) + x_2(t|t-1) - S_s \qquad (C.6)$$

The parameter S_s should not be confused with a state variable. That parameter must be estimated by maximum likelihood jointly with the remaining parameters of the model.